



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

**Διαχείριση Πόρων σε Συστήματα Κινητών Επικοινωνιών
και Κατανεμημένα Υπολογιστικά Συστήματα**

Γεώργιος Σ. Αλυφαντής

ΑΘΗΝΑ
ΑΠΡΙΛΙΟΣ 2008



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

**Διαχείριση Πόρων σε Συστήματα Κινητών Επικοινωνιών
και Κατανεμημένα Υπολογιστικά Συστήματα**

Γεώργιος Σ. Αλυφαντής

ΑΘΗΝΑ
ΑΠΡΙΛΙΟΣ 2008

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Διαχείριση Πόρων σε Συστήματα Κινητών Επικοινωνιών και Κατανεμημένα

Υπολογιστικά Συστήματα

Γεώργιος Σ. Αλυφαντής

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: Λάζαρος Μεράκος, Καθηγητής ΕΚΠΑ

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ ΠΑΡΑΚΟΛΟΥΘΗΣΗΣ:

Λάζαρος Μεράκος, Καθηγητής ΕΚΠΑ

Αλέξης Δελής, Αναπληρωτής Καθηγητής ΕΚΠΑ

Ευστάθιος Χατζηευθυμιάδης, Επίκουρος Καθηγητής ΕΚΠΑ

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Λάζαρος Μεράκος,

Καθηγητής ΕΚΠΑ

Γεώργιος Πολύζος,

Καθηγητής ΟΠΑ

Χρήστος Δουλιγέρης,

Αναπληρωτής Καθηγητής ΠΑΠΕΙ

Ιωάννης Σταυρακάκης,

Καθηγητής ΕΚΠΑ

Αλέξης Δελής,

Αναπληρωτής Καθηγητής ΕΚΠΑ

Ευστάθιος Χατζηευθυμιάδης,

Επίκουρος Καθηγητής ΕΚΠΑ

Γεώργιος Σταφάνου,

Επίκουρος Καθηγητής ΕΚΠΑ

Ημερομηνία Εξέτασης 21/04/2008

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια, τα κινητά κι ασύρματα δίκτυα, καθώς κι οι τεχνολογίες καταναμημένων συστημάτων, όπως συστήματα caching και δίκτυα διανομής περιεχομένου Ιστού, βρίσκονται στο επίκεντρο της προσοχής επιστημόνων κι ερευνητικών ομάδων σε όλο τον κόσμο. Οι επιστημονικοί τομείς που εμπλέκονται είναι διάφοροι και ποικίλουν ανάλογα με τη σκοπιά από την οποία εξετάζονται τα ζητήματα που ανακύπτουν. Στην παρούσα Διδακτορική διατριβή, εξετάζονται εφαρμογές της Θεωρίας Παιγνίων σε διάφορα προβλήματα διαχείρισης πόρων σε ό,τι αφορά τον κινητό υπολογισμό και τα καταναμημένα υπολογιστικά συστήματα. Συγκεκριμένα, έμφαση δίνεται σε προβλήματα, στα οποία είναι έμφυτος ο ανταγωνισμός μεταξύ των χρηστών, όπως διαμοιρασμός εύρους ζώνης σε ένα ασύρματο δίκτυο ή αποθηκευτικής χωρητικότητας σε ένα σύστημα caching. Προβλήματα που μελετήθηκαν αποτελούν η εξισορρόπηση φόρτου κι η προδραστική διαχείριση πόρων σε ασύρματα δίκτυα, η ευέλικτη φασματική πρόσβαση, ο έλεγχος ισχύος, καθώς κι η καταναμημένη διαχείριση αποθηκευτικής χωρητικότητας σε συστήματα caching. Με τη βοήθεια της Θεωρίας Παιγνίων κατέστη δυνατό να μοντελοποιηθούν τα εξεταζόμενα προβλήματα κι έτσι να αναδειχθούν ξεκάθαρα τα ζητήματα ανταγωνισμού κι αλληλεπίδρασης που υπεισέρχονται. Επίσης, με τη βοήθεια των μοντέλων που προέκυψαν και κατάλληλων αριθμητικών αλγορίθμων που προτάθηκαν ήταν δυνατόν να υπολογιστούν οι ισορροπίες των παιγνίων που αντιστοιχούν τα προβλήματα αυτά, καθώς και πιθανά βέλτιστα σημεία λειτουργίας.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Διάχυτος Υπολογισμός

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Κινητός υπολογισμός, καταναμημένα συστήματα, διαχείριση πόρων, θεωρία παιγνίων, θεωρία βελτιστοποίησης

ABSTRACT

In the recent years, wireless and mobile networks, as well as distributed systems technologies, such as WWW caching systems and content distribution networks, have attracted the attention of scientists and research groups all over the world. There are different scientific fields studying the various problems that may arise, depending on the aspect from which they are observed. In this thesis, applications of Game Theory to a variety of resource management problems in mobile computing and distributed systems are studied. Specifically, the thesis focuses on problems in which the competition between users is inherent, such as bandwidth sharing in a wireless network, or disk space sharing in a caching system. Problems that have been studied include load balancing and proactive resource management in wireless networks, dynamic spectrum access, power control, as well as distributed disk space management in caching systems. Through Game Theory, the aforementioned problems have been modeled and competition/interaction issues have been clearly shown. Moreover, based on the derived models and appropriate numerical algorithms, the equilibria of the games corresponding to these problems, as well as possible optimum operating points, have been computed.

SUBJECT AREA: Pervasive Computing

KEYWORDS: Mobile computing, Distributed systems, resource management, game theory, optimization theory

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΕΧΟΜΕΝΑ	9
ΠΡΟΛΟΓΟΣ	15
ΚΕΦΑΛΑΙΟ 1	17
ΕΙΣΑΓΩΓΗ.....	17
ΚΕΦΑΛΑΙΟ 2	23
ΑΝΤΙΚΕΙΜΕΝΟ ΤΗΣ ΔΙΑΤΡΙΒΗΣ.....	23
2.1 ΔΙΑΧΕΙΡΙΣΗ ΠΟΡΩΝ ΣΤΟΝ ΚΙΝΗΤΟ ΥΠΟΛΟΓΙΣΜΟ	25
2.1.1 Εξισορρόπηση Φόρτου σε Ασύρματα Κυψελωτά Δίκτυα	25
2.1.2 Αποφυγή Διακοπών Υπηρεσίας Λόγω Μεταπομπών	27
2.1.3 Ευέλικτη Πρόσβαση σε Μη-αδειοδοτημένες Φασματικές Περιοχές ...	29
2.1.4 Αποτελεσματικότητα και Δικαιοσύνη στην Πρόσβαση στο Φυσικό Μέσο	31
2.1.5 Συνολική Άποψη της Μελέτης για τη Διαχείριση Πόρων στον Κινητό Υπολογισμό	33
2.2 ΔΙΑΧΕΙΡΙΣΗ ΚΑΤΑΝΕΜΗΜΕΝΩΝ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ.....	35
2.2.1 Διαχείριση Συστημάτων WWW Caching.....	35
2.2.2 Οικονομικό Πλαίσιο για τη Διαχείριση Ιεραρχιών Caching.....	37
2.3 ΔΟΜΗ ΤΗΣ ΔΙΑΤΡΙΒΗΣ.....	37
ΚΕΦΑΛΑΙΟ 3	39
ΕΙΣΑΓΩΓΗ ΣΕ ΕΝΝΟΙΕΣ ΘΕΩΡΙΑΣ ΠΑΙΓΝΙΩΝ ΚΙ ΕΦΑΡΜΟΓΕΣ.....	39
3.1 ΒΑΣΙΚΑ ΣΤΟΙΧΕΙΑ ΘΕΩΡΙΑΣ ΠΑΙΓΝΙΩΝ	39
3.2 ΣΥΝΑΡΤΗΣΕΙΣ ΩΦΕΛΕΙΑΣ	41
3.3 ΜΗ ΣΥΝΕΡΓΑΤΙΚΗ ΕΠΙΛΥΣΗ ΠΑΙΓΝΙΩΝ – ΙΣΟΡΡΟΠΙΑ ΚΑΤΑ NASH	43
3.4 ΚΑΤΑΝΕΜΗΜΕΝΗ ΣΥΓΚΛΙΣΗ ΣΕ ΙΣΟΡΡΟΠΙΕΣ ΠΑΙΓΝΙΩΝ.....	46
3.4.1 Santa Fe Bar Problem.....	47
3.4.2 Παίγνια Διασποράς.....	49
3.5 ΣΥΝΕΡΓΑΤΙΚΗ ΕΠΙΛΥΣΗ ΠΑΙΓΝΙΩΝ – ΛΥΣΗ ΔΙΑΠΡΑΓΜΑΤΕΥΣΗΣ ΤΟΥ NASH	50
3.6 ΔΗΜΟΠΡΑΣΙΕΣ	53
3.7 ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΘΕΩΡΙΑΣ ΠΑΙΓΝΙΩΝ ΣΕ ΔΙΚΤΥΑΚΑ ΠΡΟΒΛΗΜΑΤΑ.....	55
3.7.1 Εκχώρηση εύρους ζώνης σε Εικονικά Μονοπάτια (Virtual Paths) για δίκτυα πολλαπλών χρηστών	56
3.7.2 Παιγιοθεωρητική ανάλυση των αρχών ελέγχου συμφόρησης	58
3.7.3 Αρχιτεκτονική μη-συνεργατικών δικτύων	61
3.7.4 Έλεγχος Ισχύος σε CDMA δίκτυα	61
3.7.5 Άλλες εργασίες	65

ΜΕΡΟΣ Α	67
ΔΙΑΧΕΙΡΙΣΗ ΠΟΡΩΝ ΣΤΟΝ ΚΙΝΗΤΟ ΥΠΟΛΟΓΙΣΜΟ	67
ΚΕΦΑΛΑΙΟ 4	69
ΔΙΑΧΕΙΡΙΣΗ ΘΕΣΗΣ ΚΙΝΗΤΩΝ ΧΡΗΣΤΩΝ ΓΙΑ ΕΞΙΣΟΡΡΟΠΗΣΗ ΦΟΡΤΟΥ ΣΕ ΑΣΥΡΜΑΤΑ ΚΥΨΕΛΩΤΑ ΔΙΚΤΥΑ	69
4.1 ΕΙΣΑΓΩΓΗ	69
4.2 ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΑΙ ΣΥΝΕΙΣΦΟΡΑ	71
4.3 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΚΑΙ ΛΕΙΤΟΥΡΓΙΑ ΤΟΥ ΠΡΟΤΕΙΝΟΜΕΝΟΥ ΣΥΣΤΗΜΑΤΟΣ	72
4.3.1 <i>Λειτουργία του Συστήματος</i>	73
4.3.2 <i>Μοντελοποίηση Θέσης</i>	74
4.3.3 <i>Διεπαφή Τερματικού Πελάτη του Χρήστη</i>	75
4.3.4 <i>Δημιουργία Πρότασης Μετακίνησης</i>	75
4.3.5 <i>Ζητήματα Σχετικά με τις Προτάσεις Μετακίνησης</i>	78
4.4 ΜΙΚΡΟΟΙΚΟΝΟΜΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ ΚΙ ΕΦΑΡΜΟΓΗ ΣΤΟ ΜΙΤΟΣ	79
4.5 ΈΝΑ ΣΧΗΜΑ ΠΡΟΓΝΩΣΕΩΝ ΓΙΑ ΤΟ ΜΙΤΟΣ	81
4.6 ΜΟΝΤΕΛΟ ΠΡΟΣΟΜΟΙΩΣΗΣ	83
4.6.1 <i>Αποτελέσματα Προσομοίωσης για το μE-MITOS</i>	89
4.6.2 <i>Αποτελέσματα Προσομοίωσης για το SFBP-MITOS</i>	91
4.6.3 <i>Αποτελέσματα Προσομοίωσης για το Plain-Vanilla MITOS</i>	93
4.6.4 <i>Συζήτηση</i>	94
4.6.5 <i>Σύνοψη Σύγκρισης των Προτεινόμενων Σχημάτων του MITOS</i>	99
4.7 ΣΥΓΚΡΙΣΗ ΤΟΥ ΜΙΤΟΣ ΜΕ ΆΛΛΑ ΣΥΣΤΗΜΑΤΑ	99
4.8 ΣΥΜΠΕΡΑΣΜΑΤΑ	101
ΚΕΦΑΛΑΙΟ 5	103
ΠΡΟΔΡΑΣΤΙΚΗ ΔΙΑΧΕΙΡΙΣΗ ΕΥΡΟΥΣ ΖΩΝΗΣ ΣΕ ΚΥΨΕΛΩΤΑ ΔΙΚΤΥΑ ΜΕΣΩ ΔΙΑΠΡΑΓΜΑΤΕΥΣΕΩΝ	103
5.1 ΕΙΣΑΓΩΓΗ	103
5.2 ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ	104
5.3 ΑΞΙΟΛΟΓΗΣΗ ΤΩΝ ΠΟΡΩΝ ΑΠΟ ΤΟ ΚΤ ΚΑΙ ΤΟ ΣΒ	105
5.4 ΠΡΟΤΕΙΝΟΜΕΝΗ ΑΡΧΙΤΕΚΤΟΝΙΚΗ	106
5.4.1 <i>Μηχανισμός Διαπραγμάτευσης</i>	107
5.4.2 <i>Σχήμα Διαχείρισης Πόρων</i>	109
5.5 ΠΡΟΣΟΜΟΙΩΣΕΙΣ	118
5.5.1 <i>Σχεδιασμός Προσομοίωσης</i>	118
5.5.2 <i>Αποτελέσματα Προσομοιώσεων</i>	119
5.5.3 <i>Κέρδος του Δικτύου</i>	122
5.6 ΣΥΜΠΕΡΑΣΜΑΤΑ	123
ΚΕΦΑΛΑΙΟ 6	125
ΠΡΟΔΡΑΣΤΙΚΗ ΔΙΑΧΕΙΡΙΣΗ ΕΥΡΟΥΣ ΖΩΝΗΣ ΣΕ ΚΥΨΕΛΩΤΑ ΔΙΚΤΥΑ ΜΕ ΔΗΜΟΠΡΑΣΙΕΣ	125
6.1 ΕΙΣΑΓΩΓΗ	125
6.2 ΜΟΝΤΕΛΟ ΣΥΣΤΗΜΑΤΟΣ	126
6.3 ΟΛΟΚΛΗΡΩΜΕΝΟ ΣΥΣΤΗΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΑΣΥΡΜΑΤΩΝ ΠΟΡΩΝ	127
6.4 ΔΙΑΜΟΙΡΑΣΜΟΣ ΠΟΡΩΝ ΜΕΣΩ ΔΗΜΟΠΡΑΣΙΩΝ	130

6.5	ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΚΑΙ ΛΕΙΤΟΥΡΓΙΑ ΤΟΥ ΠΛΑΙΣΙΟΥ ΔΙΑΧΕΙΡΙΣΗΣ ΠΟΡΩΝ.....	133
6.5.1	<i>Δημοπρασία κατά την Έναρξη Χρονοθυρίδας</i>	<i>134</i>
6.5.2	<i>Δημοπρασία για την Εκκίνηση Νέας Εφαρμογής</i>	<i>136</i>
6.5.3	<i>Δημοπρασία μετά από Μεταπομπή.....</i>	<i>137</i>
6.5.4	<i>Χρέωση</i>	<i>138</i>
6.6	ΠΛΑΙΣΙΟ ΠΡΟΒΛΕΨΗΣ ΓΙΑ ΠΡΟΔΡΑΣΤΙΚΗ ΔΕΣΜΕΥΣΗ ΠΟΡΩΝ.....	139
6.6.1	<i>Πρόβλεψη Κινητικότητας Χρήστη.....</i>	<i>139</i>
6.6.2	<i>Πρόβλεψη Συμπεριφοράς Εφαρμογών.....</i>	<i>146</i>
6.7	ΠΡΟΣΟΜΟΙΩΣΕΙΣ	149
6.7.1	<i>Αποτελέσματα Προσομοιώσεων</i>	<i>150</i>
6.8	ΣΥΜΠΕΡΑΣΜΑΤΑ.....	155
ΚΕΦΑΛΑΙΟ 7		157
ΚΑΤΑΝΕΜΗΜΕΝΗ ΠΡΟΣΒΑΣΗ ΣΕ ΔΙΚΤΥΑ ΜΕ ΕΥΕΛΙΚΤΗ ΧΡΗΣΗ ΦΑΣΜΑΤΟΣ		
.....		157
7.1	ΕΙΣΑΓΩΓΗ	157
7.2	ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΑΙ ΣΥΝΕΙΣΦΟΡΑ.....	158
7.3	ΜΟΝΤΕΛΟ ΣΥΣΤΗΜΑΤΟΣ.....	159
7.4	ΚΑΤΑΝΕΜΗΜΕΝΟΣ, ΜΗ-ΣΥΝΕΡΓΑΤΙΚΟΣ ΈΛΕΓΧΟΣ ΙΣΧΥΟΣ.....	161
7.5	ΜΗ-ΣΥΝΕΡΓΑΤΙΚΗ ΕΠΙΛΟΓΗ ΖΩΝΗΣ ΣΥΧΝΟΤΗΤΩΝ	165
7.5.1	<i>Nash Ισορροπία του BSG.....</i>	<i>168</i>
7.5.2	<i>Αξιολόγηση Αποδοτικότητας του NE του BSG.....</i>	<i>169</i>
7.5.3	<i>Ζητήματα Σύγκλισης</i>	<i>173</i>
7.6	ΚΑΤΑΝΕΜΗΜΕΝΕΣ ΣΤΡΑΤΗΓΙΚΕΣ ΓΙΑ ΤΟ BSG	174
7.6.1	<i>Επιλογή Ζώνης Συχνότητας Βάσει Προγνώσεων.....</i>	<i>174</i>
7.6.2	<i>Στοχαστική Επιλογή Ζώνης Συχνότητας</i>	<i>178</i>
7.6.3	<i>Μία Εναλλακτική Στρατηγική Επιλογής Ζώνης Συχνότητας.....</i>	<i>180</i>
7.6.4	<i>Ένα Απλό Πρωτόκολλο για Δυναμική Πρόσβαση στο Φάσμα</i>	<i>181</i>
7.7	ΠΡΟΣΟΜΟΙΩΣΕΙΣ	182
7.7.1	<i>Αποτελέσματα Προσομοιώσεων</i>	<i>184</i>
7.8	ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ.....	190
ΚΕΦΑΛΑΙΟ 8		193
ΕΛΕΓΧΟΣ ΙΣΧΥΟΣ ΣΕ CDMA ΔΙΚΤΥΑ		193
8.1	ΕΙΣΑΓΩΓΗ	193
8.2	ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΑΙ ΣΥΝΕΙΣΦΟΡΑ.....	194
8.3	ΜΟΝΤΕΛΟ ΣΥΣΤΗΜΑΤΟΣ.....	197
8.4	ΜΗ-ΣΥΝΕΡΓΑΤΙΚΟΣ ΈΛΕΓΧΟΣ ΙΣΧΥΟΣ ΜΕ ΠΕΡΙΟΡΙΣΜΟΥΣ SIR	198
8.5	ΑΞΙΟΛΟΓΗΣΗ ΑΠΟΔΟΣΗΣ ΤΟΥ ΜΗ-ΣΥΝΕΡΓΑΤΙΚΟΥ ΣΧΗΜΑΤΟΣ ΕΛΕΓΧΟΥ ΙΣΧΥΟΣ	201
8.6	ΈΛΕΓΧΟΣ ΙΣΧΥΟΣ ΜΕ ΔΙΑΙΤΗΣΙΑ ΚΑΙ ΠΕΡΙΟΡΙΣΜΟΥΣ SIR	205
8.6.1	<i>Η Λύση Διαπραγμάτευσης του Nash για το Παίγνιο Ελέγχου Ισχύος</i>	<i>206</i>
8.6.2	<i>Προσδιορισμός της Λύσης Διαπραγμάτευσης του Nash</i>	<i>210</i>
8.6.3	<i>Ζητήματα Σταθερότητας του NBS</i>	<i>215</i>
8.7	ΑΡΙΘΜΗΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ.....	216
8.8	ΣΥΜΠΕΡΑΣΜΑΤΑ.....	222

ΜΕΡΟΣ Β	223
ΔΙΑΧΕΙΡΙΣΗ ΠΟΡΩΝ ΣΕ ΚΑΤΑΝΕΜΗΜΕΝΑ ΣΥΣΤΗΜΑΤΑ.....	223
ΚΕΦΑΛΑΙΟ 9	225
ΚΑΤΑΝΕΜΗΜΕΝΟ ΣΧΗΜΑ ΓΙΑ ΤΟ ΔΙΑΜΟΙΡΑΣΜΟ ΧΩΡΗΤΙΚΟΤΗΤΑΣ ΣΥΣΤΗΜΑΤΩΝ WWW CACHING.....	225
9.1 ΕΙΣΑΓΩΓΗ	225
9.2 ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ	226
9.3 ΜΟΝΤΕΛΟ ΣΥΣΤΗΜΑΤΟΣ	227
9.4 ΎΠΑΡΞΗ ΚΑΙ ΜΟΝΑΔΙΚΟΤΗΤΑ ΤΗΣ NASH ΙΣΟΡΡΟΠΙΑΣ	231
9.5 ΣΥΓΚΛΙΣΗ ΣΤΗ NASH ΙΣΟΡΡΟΠΙΑ.....	233
9.6 ΥΛΟΠΟΙΗΣΗ ΚΑΤΑΝΕΜΗΜΕΝΟΥ ΑΛΓΟΡΙΘΜΟΥ.....	236
9.6.1 Παραλλαγή Αλγορίθμου.....	236
9.6.2 Ζητήματα Υλοποίησης	237
9.7 ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΡΟΣΟΜΟΙΩΣΗΣ.....	241
9.8 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ.....	244
ΚΕΦΑΛΑΙΟ 10	247
ΟΙΚΟΝΟΜΙΚΟ ΠΛΑΙΣΙΟ ΓΙΑ ΤΗ ΔΙΑΧΕΙΡΙΣΗ ΣΥΣΤΗΜΑΤΩΝ CACHING	247
10.1 ΕΙΣΑΓΩΓΗ	247
10.2 ΕΙΣΑΓΩΓΗ ΣΕ ΟΙΚΟΝΟΜΙΚΕΣ ΈΝΝΟΙΕΣ.....	248
10.3 ΜΟΝΤΕΛΟ ΣΥΣΤΗΜΑΤΟΣ	251
10.4 ΜΟΝΟΠΩΛΙΟ ΜΕ ΧΡΕΩΣΗ ΜΕ ΔΙΑΚΡΙΣΗ.....	254
10.4.1 Καθορισμός Βέλτιστης Εκχώρησης Χωρητικότητας.....	256
10.4.2 Καθορισμός Τιμών και Ζητήματα Κινήτρων	258
10.5 ΜΟΝΟΠΩΛΙΟ ΜΕ ΧΡΕΩΣΗ ΧΩΡΙΣ ΔΙΑΚΡΙΣΗ.....	266
10.5.1 Καθορισμός Βέλτιστης Εκχώρησης Χωρητικότητας.....	267
10.5.2 Καθορισμός Τιμών και Ζητήματα Κινήτρων	269
10.6 ΜΟΝΟΠΩΛΙΟ ΜΕ ΡΥΘΜΙΣΤΗ ΓΙΑ ΜΕΓΙΣΤΟΠΟΙΗΣΗ ΚΟΙΝΩΝΙΚΗΣ ΕΥΗΜΕΡΙΑΣ....	271
10.6.1 Καθορισμός Βέλτιστης Εκχώρησης Χωρητικότητας.....	272
10.7 ΣΥΓΚΡΙΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΤΩΝ ΔΙΑΦΟΡΕΤΙΚΩΝ ΤΥΠΩΝ ΜΟΝΟΠΩΛΙΩΝ.....	274
10.8 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΗ ΕΡΓΑΣΙΑ.....	282
ΚΕΦΑΛΑΙΟ 11	285
ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΚΑΤΕΥΘΥΝΣΕΙΣ.....	285
11.1 ΣΥΜΠΕΡΑΣΜΑΤΑ.....	285
11.2 ΜΕΛΛΟΝΤΙΚΕΣ ΕΡΕΥΝΗΤΙΚΕΣ ΚΑΤΕΥΘΥΝΣΕΙΣ	287
ΠΑΡΑΡΤΗΜΑ Ι	293
ΣΧΗΜΑ ΠΟΛΛΑΠΛΩΝ ΔΗΜΟΠΡΑΣΙΩΝ.....	293
ΠΑΡΑΡΤΗΜΑ ΙΙ	297
ΕΠΙΔΡΑΣΗ ΠΛΗΘΟΥΣ ΠΡΟΓΝΩΣΤΩΝ ΑΝΑ ΧΡΗΣΤΗ ΣΤΗΝ ΑΠΟΤΕΛΕΣΜΑΤΙΚΟΤΗΤΑ ΤΟΥ SFBP-MITOS	297
ΠΑΡΑΡΤΗΜΑ ΙΙΙ	301
ΠΡΟΒΛΕΨΗ ΕΞΩ-ΚΥΨΕΛΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ.....	301

ΠΑΡΑΡΤΗΜΑ IV	305
ΒΕΛΤΙΣΤΗ ΧΡΗΣΗ ΦΑΣΜΑΤΟΣ	305
ΠΑΡΑΡΤΗΜΑ V	309
ΑΡΙΘΜΗΤΙΚΟΣ ΑΛΓΟΡΙΘΜΟΣ ΥΠΟΛΟΓΙΣΜΟΥ ΤΟΥ NBS.....	309
V.1 ΚΑΘΟΡΙΣΜΟΣ ΕΦΙΚΤΟΥ ΣΗΜΕΙΟΥ ΕΚΚΙΝΗΣΗΣ	309
V.2 ΚΑΘΟΡΙΣΜΟΣ ΕΦΙΚΤΩΝ ΚΑΤΕΥΘΥΝΣΕΩΝ ΑΎΞΗΣΗΣ.....	310
V.3 ΕΠΙΛΟΓΗ ΒΉΜΑΤΟΣ	313
ΌΡΟΙ.....	315
ΑΚΡΩΝΥΜΙΑ.....	319
ΑΝΑΦΟΡΕΣ.....	321

ΠΡΟΛΟΓΟΣ

Η παρούσα διδακτορική διατριβή εκπονήθηκε στο Τμήμα Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών και μελετά εφαρμογές της Θεωρίας Παιγνίων σε διάφορα προβλήματα διαχείρισης πόρων σε ό,τι αφορά τον κινητό υπολογισμό και τα κατανομημένα υπολογιστικά συστήματα.

Στα πλαίσια της παρούσας διατριβής, μελετούνται διάφορα προβλήματα στα οποία είναι έμφυτος ο ανταγωνισμός μεταξύ των χρηστών για τους εκάστοτε πόρους, όπως εύρος ζώνης σε ένα ασύρματο δίκτυο ή αποθηκευτική χωρητικότητα σε ένα σύστημα caching. Προβλήματα που μελετήθηκαν αποτελούν η εξισορρόπηση φόρτου κι η προδραστική διαχείριση πόρων σε ασύρματα δίκτυα, η ευέλικτη φασματική πρόσβαση, ο έλεγχος ισχύος, καθώς κι η κατανομημένη διαχείριση αποθηκευτικής χωρητικότητας σε συστήματα caching. Με τη βοήθεια της Θεωρίας Παιγνίων κατέστη δυνατό να μοντελοποιηθούν τα εξεταζόμενα προβλήματα κι έτσι να αναδειχθούν ξεκάθαρα τα ζητήματα ανταγωνισμού κι αλληλεπίδρασης που υπεισέρχονται. Επίσης, με τη βοήθεια των μοντέλων που προέκυψαν και κατάλληλων αριθμητικών αλγορίθμων που προτάθηκαν ήταν δυνατόν να υπολογιστούν οι ισορροπίες των παιγνίων που αντιστοιχούν στα προβλήματα αυτά, καθώς και πιθανά βέλτιστα σημεία λειτουργίας.

Ολοκληρώνοντας την εκπόνηση της διδακτορικής μου διατριβής, νιώθω την ανάγκη να ευχαριστήσω κάποιους ανθρώπους, οι οποίοι βοήθησαν στην ολοκλήρωση της. Προτίστως, ευχαριστώ τους καθηγητές της τριμελούς επιτροπής μου, Καθηγητή Λάζαρο Μεράκο, Αναπληρωτή Καθηγητή Αλέξη Δελή και τον Επίκουρο Καθηγητή Ευστάθιο Χατζηευθυμιάδη για την υποστήριξη και τις συμβουλές που μου παρείχαν.

Ιδιαίτερα, θα ήθελα να ευχαριστήσω τον Καθηγητή Λάζαρο Μεράκο, ο οποίος με την κριτική σκέψη και το ήθος του υπήρξε για μένα πηγή παραδειγματισμού. Η βοήθεια και συνδρομή του σε συγκεκριμένα στάδια της διατριβής υπήρξε καθοριστική. Θα ήθελα, επίσης, να ευχαριστήσω θερμά τον Επίκουρο Καθηγητή Ευστάθιο Χατζηευθυμιάδη, ο οποίος μου παρείχε συνεχή στήριξη, σε πολλά

επίπεδα, καθ' όλη τη διάρκεια της εργασίας. Η επιμονή κι ο ζήλος που τον διακρίνουν συνέβαλλαν τα μέγιστα στο να διατηρηθεί το ενδιαφέρον, για τη διαρκή αναζήτηση νέων ιδεών, ενεργό μέχρι και την ολοκλήρωση της. Οι συχνές δε συζητήσεις μαζί του υπήρξαν πηγή έμπνευσης κι αφετηρία για νέες ιδέες.

Κατά τη διάρκεια της εκπόνησης της διδακτορικής μου διατριβής, ήμουν μέλος της ομάδας Διάχυτου Υπολογισμού. Ως μέλος της ομάδας αυτής, είχα την ευκαιρία να συζητήσω επιστημονικά ή μη θέματα με ανθρώπους οι οποίοι ήταν πάντα πρόθυμοι να με συμβουλέψουν και να με βοηθήσουν. Η συνεργασία μαζί τους υπήρξε ιδιαίτερα αρμονική κι εποικοδομητική και θα ήθελα να ευχαριστήσω όλους γι' αυτό. Ιδιαίτερα ένα μεγάλο ευχαριστώ στους Βασίλη Τσέτσο, Οδυσσέα Σέκκα, Γιάννη Πρίγγουρη, Χρίστο Αναγνωστόπουλο, Γιάννη Μαριά, Μίλτο Κυριακάκο, Βασίλη Μπαούση, Γιάννη Αγγελόπουλο και Φάνη Κοντό.

Θα ήθελα, επίσης, να ευχαριστήσω θερμά το Κοινωνικό Ίδρυμα Αλέξανδρος Ωνάσης, η οικονομική συνδρομή του οποίου υπήρξε σημαντική βοήθεια για εμένα, δίνοντας μου την ευκαιρία να αφοσιωθώ απερίσπαστος στην εκπόνηση της διατριβής μου.

Τέλος, θα ήθελα να ευχαριστήσω τους ανθρώπους από το φιλικό κι οικογενειακό μου περιβάλλον, και κυρίως τους γονείς μου, των οποίων η συνδρομή σε όλα τα επίπεδα, αυτά τα χρόνια, ήταν αμέριστη και καθοριστική για εμένα.

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

Τα τελευταία χρόνια, έχει παρατηρηθεί εντυπωσιακή ανάπτυξη στις τεχνολογίες επικοινωνιών, κι ιδιαίτερα σε ό,τι αφορά την περιοχή των ασύρματων και κινητών επικοινωνιών. Πλέον, σε κάθε σχεδόν περιοχή στον κόσμο είναι δυνατή η πρόσβαση μέσω συστημάτων 2^{ης} (Global System Mobile – GSM) και 3^{ης} (Universal Telecommunications System – UMTS) γενιάς για υπηρεσίες φωνής, αλλά και πρόσβαση στο Διαδίκτυο. Μεγάλη άνθηση επίσης έχει παρατηρηθεί και στα τοπικά ασύρματα δίκτυα (wireless local area networks – WLAN), τα οποία πλέον συναντώνται σε πολλούς δημόσιους χώρους, όπως αεροδρόμια και εμπορικά κέντρα, αλλά και σε άλλους χώρους, όπως εταιρίες και Πανεπιστημιακές εγκαταστάσεις.

Η εξάπλωση των WLAN οφείλεται σε μεγάλο μέρος στο ότι λειτουργούν σε περιοχές συχνοτήτων οι οποίες μπορούν να χρησιμοποιηθούν χωρίς την ανάγκη έκδοσης άδειας (Industrial, Scientific, and Medical – ISM). Είναι εύκολο κανείς να εγκαταστήσει ένα ή περισσότερα σημεία πρόσβασης (access point – AP), αφού δεν έχει κανένα κόστος και περιορισμό στη χρήση, μπορώντας έτσι να εξυπηρετηθεί αποδοτικά πλήθος κινητών χρηστών και μάλιστα με αρκετά υψηλές ταχύτητες. Σε πολλές περιπτώσεις η πρόσβαση είναι ελεύθερη στο κοινό, ενώ σε άλλες περιπτώσεις, συναντάται η εμπορική εκμετάλλευση των εγκαταστάσεων WLAN από τους τοπικούς παρόχους, επιφέροντας τους σημαντικά κέρδη.

Σημαντικό ρόλο στην επιτυχή εξάπλωση της Διαδικτυακής πρόσβασης μέσω WLAN, έχει παίξει κι η αυξημένη διείσδυση του ευρυζωνικού Διαδικτύου (broadband Internet). Σε ανεπτυγμένες χώρες, ένα αρκετά μεγάλο ποσοστό του πληθυσμού έχει πρόσβαση στο Διαδίκτυο με ταχύτητες από 700 Kb/s, μέχρι και πάνω από 24 Mb/s, μέσω ευρυζωνικών συνδέσεων *asynchronous digital subscriber loop* (ADSL). Οι χρήστες αυτοί συνδέονται με το δίκτυο κατά κανόνα με κάποιον δρομολογητή ο οποίος συνήθως είναι εξοπλισμένος με ασύρματη διεπαφή WLAN, λειτουργώντας ως AP. Έτσι, σε αστικά, πυκνοκατοικημένα περιβάλλοντα η ραδιο-κάλυψη από WLAN είναι εκτενής. Το γεγονός αυτό καθιστά την τεχνολογία

των WLAN ικανή για την πρόσβαση στο Διαδίκτυο από το ευρύ κοινό. Για να πραγματοποιηθεί όμως κάτι τέτοιο, είναι απαραίτητο οι κάτοχοι σύνδεσης ADSL να επιτρέπουν σε άγνωστους (περαστικούς) να χρησιμοποιούν τη σύνδεση τους, παρέχοντας τους πρόσβαση μέσω του WLAN.

Αξίζει να σημειωθεί ότι κατά κανόνα οι ιδιώτες, κάτοχοι μίας σύνδεσης, αποφεύγουν να επιτρέπουν την ελεύθερη πρόσβαση σε άγνωστους χρήστες. Αυτό συμβαίνει διότι δεν επιθυμούν να μοιράζουν το εύρος ζώνης τους, καθώς το θέλουν για αποκλειστική χρήση, αλλά και για λόγους ασφαλείας, φοβούμενοι κακόβουλες επιθέσεις. Παρόλ' αυτά, τελευταία, έχει δοθεί έμφαση στο πως μπορεί αυτή η κατάσταση να ανατραπεί κι οι χρήστες να παρέχουν πρόσβαση στο Διαδίκτυο σε άλλους χρήστες μέσω του WLAN τους.

Οι προσεγγίσεις που εξετάζονται για την ελεύθερη παροχή πρόσβασης από ιδιώτες βασίζονται κυρίως στην παροχή κινήτρων. Τα κίνητρα μπορεί να είναι οικονομικά, π.χ. η σύνδεση να παρέχεται σε κάποιο κόστος κι ο κάτοχος της σύνδεσης να λαμβάνει ένα ποσοστό από τα κέρδη. Το βασικότερο κίνητρο όμως έχει σχέση με την ίδια ανάγκη του χρήστη, όταν είναι εκτός σπιτιού, να έχει σύνδεση στο Διαδίκτυο. Ένας χρήστης μπορεί να παρέχει πρόσβαση σε κάποιους άλλους χρήστες, ώστε όταν αυτός με τη σειρά του βρεθεί σε άλλη περιοχή να μπορέσει να συνδεθεί μέσω κάποιου άλλου χρήστη. Τέτοια σχήματα έχουν μελετηθεί στη βιβλιογραφία, με χαρακτηριστική την εργασία [1]. Επίσης, υπάρχουν κι εμπορικά συστήματα που βασίζονται στο κίνητρο της αμοιβαίας συνεισφοράς, όπως αυτό που παρέχει η FON [2] κι η Whisher [3]. Αξίζει να σημειωθεί ότι η διείσδυση των συστημάτων αυτών αυξάνεται ραγδαία, με τους κόμβους που είναι διαθέσιμοι να είναι χιλιάδες παγκοσμίως και να αυξάνονται καθημερινά. Προβλέπεται ότι σε λίγα χρόνια, θα είναι δυνατόν σχεδόν παντού στον κόσμο να μπορεί κανείς να συνδεθεί ελεύθερα μέσω WLAN. Η τάση δε αυτή είναι τόσο έντονη, που οι πάροχοι κινητών επικοινωνιών ανησυχούν για επικείμενες σημαντικές απώλειες κερδών.

Από τα παραπάνω, λοιπόν, είναι φανερό ότι η εξάπλωση των ασύρματων τεχνολογιών για πρόσβαση στο Διαδίκτυο είναι εκτεταμένη, κι αναμένεται στο προσεχές μέλλον να αυξηθεί περαιτέρω. Η μελέτη, λοιπόν, ζητημάτων που άπτονται των ασυρμάτων δικτύων και του κινητού υπολογισμού παρουσιάζει

ιδιαίτερο ενδιαφέρον. Κάτι τέτοιο αντανακλάται κι από την πολύ πλούσια βιβλιογραφία που έχει παραχθεί ως αποτέλεσμα της έρευνας στο πεδίο αυτό. Το ενδιαφέρον αυξάνεται περισσότερο αν αναλογιστεί κανείς τις όλο ένα νεοεμφανιζόμενες Διαδικτυακές υπηρεσίες, αλλά και το αναδυόμενο πρότυπο του *διάχυτου υπολογισμού* (pervasive computing), που θεωρείται η εξέλιξη του κινητού υπολογισμού. Σύμφωνα με αυτό το πρότυπο, το περιβάλλον του χρήστη θα είναι εφοδιασμένο με διάχυτη υπολογιστική ευφυΐα, η οποία θα πηγάζει από σύνολο αυτόνομων συσκευών που θα αλληλεπιδρούν ασύρματα, έχοντας ως σκοπό την εκτέλεση διάφορων εργασιών.

Η μελέτη, λοιπόν, των ασυρμάτων και κινητών συστημάτων αποτελεί ένα πολύ ενδιαφέρον θέμα. Αν αναλογιστεί κανείς την πολυπλοκότητα και τους περιορισμούς που υφίστανται, λόγω τεχνολογικών, ρυθμιστικών, αλλά κι οικονομικών ζητημάτων, το πρόβλημα αποτελεί μία σημαντική πρόκληση. Ιδιαίτερο ενδιαφέρον, επίσης, παρουσιάζει κι η αποδοτική παροχή των υπηρεσιών αυτών καθ' αυτών προς τους χρήστες. Εκτός δηλαδή από την επίτευξη καλής λειτουργίας της δικτυακής σύνδεσης, σημαντικό είναι κι η υπηρεσίες να παρέχονται με ικανή ποιότητα.

Μία από τις σημαντικότερες υπηρεσίες του Διαδικτύου σήμερα είναι ο Παγκόσμιος Ιστός (World Wide Web – WWW). Είναι γενικά παραδεκτό ότι η υπηρεσία WWW συντέλεσε αποφασιστικά προς την υλοποίηση και καθιέρωση της παγκόσμιας Πληροφοριακής Υπερλεωφόρου (Information SuperHighway) περισσότερο από οποιαδήποτε άλλη τεχνολογική εξέλιξη, από τις πρώτες στιγμές εμφάνισης του Διαδικτύου. Αποτελεί μία υπηρεσία που χρησιμοποιείται κατά κόρον και θεωρείται δίκαια ως μία από τις σημαντικότερες στο Διαδίκτυο.

Η εντυπωσιακή ανάπτυξη και διαρκώς αυξανόμενη χρήση της υπηρεσίας WWW κατέστησαν επιβεβλημένη τη βελτιστοποίηση στην πρόσβαση του πληροφοριακού περιεχομένου. Σημαντική είναι η συνεισφορά προς αυτήν την κατεύθυνση της τεχνολογίας των WWW cache και των *δικτύων διανομής περιεχομένου* (content distribution network – CDN). Η συμβολή τους έγκειται στην μείωση του απαιτούμενου χρόνου παράδοσης των περιεχομένων που ζητούνται από τους χρήστες, καθώς τα περιεχόμενα αυτά τοποθετούνται κοντά σε αυτούς.

Τα ζητήματα που ανακύπτουν στη διαχείριση των WWW cache και των CDN είναι πολυάριθμα. Έχουν σχέση με την επιλογή διαφόρων παραμέτρων, αλλά και γενικότερων στρατηγικών, όπως σε ποιους διακομιστές (server) να τοποθετηθούν κάποια αντίγραφα, ή από ποιον διακομιστή να ζητηθεί ένα αντίγραφο. Αυτά τα προβλήματα έχουν αποτελέσει αντικείμενο μελέτης από μεγάλο πλήθος ερευνητών καθώς κι αντικείμενο έρευνας κι ανάπτυξης προϊόντων από εμπορικούς οργανισμούς παγκοσμίως. Ανοικτά ζητήματα όμως παραμένουν αρκετά, καθώς υπάρχει σημαντικό περιθώριο για πιθανές βελτιώσεις υπάρχοντων αρχιτεκτονικών και στρατηγικών, αλλά και για πρόταση νέων.

Η παρούσα διατριβή εστιάζει σε δύο βασικά προβλήματα: 1) τη μελέτη της διαχείρισης πόρων στον κινητό υπολογισμό, εξετάζοντας το πρόβλημα σε διαφορετικά επίπεδα κι από διαφορετικές οπτικές γωνίες, και 2) τη μελέτη ζητημάτων που άπτονται στα κατανεμημένα συστήματα με έμφαση στην αποτελεσματική παροχή της υπηρεσίας WWW. Τα παραπάνω προβλήματα είναι φαινομενικά ανεξάρτητα, καθώς το πρώτο αφορά κατά βάση την ασύρματη δικτυακή πρόσβαση των χρηστών, ενώ το δεύτερο την υπηρεσία WWW. Παρόλ' αυτά και τα δύο σε συνδυασμό έχουν άμεση σχέση με την εξυπηρέτηση των αναγκών ενός σύγχρονου χρήστη. Η ασύρματη σύνδεση με το Διαδίκτυο είναι ένα από τα βασικά χαρακτηριστικά του σύγχρονου κόσμου, απαντώντας στην απαίτηση "always on" [4]. Επομένως, η ανάγκη για αποδοτική διαχείριση των ασύρματων πόρων είναι επιτακτική. Από την άλλη, καθώς η υπηρεσία WWW είναι από τις πλέον βασικές στη ζωή ενός σύγχρονου χρήστη του Διαδικτύου, η αποτελεσματική παροχή της είναι χωρίς αμφιβολία σημαντική.

Σε ό,τι αφορά τη διαχείριση πόρων στον κινητό υπολογισμό, τα θέματα που πραγματεύεται η παρούσα διατριβή εντάσσονται σε δύο βασικές κατηγορίες, με βάση τη χρονική κλίμακα που αφορούν:

α) Διαχείριση των ραδιο-πόρων (radio resource management – RRM), που αφορά σε αποφάσεις σχετικά με την εκχώρηση πόρων σε βραχυπρόθεσμη βάση, ώστε η πρόσβαση των χρηστών να είναι όσο το δυνατόν αποτελεσματική, αλλά και δίκαιη. Τέτοιο ζήτημα αποτελεί ο έλεγχος ισχύος εκπομπής (power control), αλλά και δυναμική πρόσβαση στο φάσμα (dynamic spectrum access – DSA),

β) διαχείριση πόρων σε μεσοπρόθεσμη βάση. Τέτοια ζητήματα έχουν σχέση με την πρόληψη συμβάντων διακοπής εξυπηρέτησης λόγω μεταπομπής (handover), ή την προσπάθεια εξισορρόπησης φόρτου σε περιόδους αιχμής.

Σε σχέση με τη διαχείριση κατανεμημένων συστημάτων, η παρούσα διατριβή πραγματεύεται ζητήματα που είναι συναφή με την τεχνολογία WWW caching. Συγκεκριμένα, εστιάζει σε θέματα που αφορούν στο δίκαιο διαμοιρασμό πόρων αποθήκευσης σε τέτοια συστήματα, ώστε να αποφεύγεται η μονοπώληση της χρήσης τους από μία περιορισμένη, επιθετική μόνο μερίδα χρηστών.

ΚΕΦΑΛΑΙΟ 2

ΑΝΤΙΚΕΙΜΕΝΟ ΤΗΣ ΔΙΑΤΡΙΒΗΣ

Στο κεφάλαιο αυτό, περιγράφεται το αντικείμενο της παρούσας διατριβής. Βασικό χαρακτηριστικό της διατριβής αποτελεί ότι οι χρήστες αντιμετωπίζονται ως οντότητες, οι οποίες αυτόνομα λαμβάνουν αποφάσεις σχετικά με τον τρόπο που θα χρησιμοποιούν τους πόρους του συστήματος (είτε αυτό είναι το κινητό δίκτυο, είτε ένας WWW caching proxy, ή ένα CDN). Συγκεκριμένα, γίνεται η υπόθεση ότι οι χρήστες είναι οντότητες εγωιστικές (μη-συνεργατικές), οι οποίες λαμβάνουν αποφάσεις με γνώμονα το δικό τους και μόνον όφελος, χωρίς να ενδιαφέρονται αν οι αποφάσεις που λαμβάνουν κι οι αντίστοιχες ενέργειες έχουν αρνητικές επιπτώσεις στους υπόλοιπους χρήστες του συστήματος. Εναλλακτικά, οι χρήστες θα μπορούσαν να θεωρηθούν συνεργατικές οντότητες που ως σκοπό δεν έχουν την προσωπική τους μόνον ευημερία, αλλά τη βελτιστοποίηση της λειτουργίας του εκάστοτε συστήματος. Κάτι τέτοιο, όμως, είναι αναμενόμενο να συμβαίνει σε συστήματα με «ομοιόμορφους» χρήστες (π.χ., εταιρικά δίκτυα). Οι χρήστες δημόσιων μη αδειοδοτημένων δικτύων αναμένεται να είναι «ανομοιογενείς». Έτσι, είναι μάλλον δύσκολο να υποτεθεί ότι θα μπορούν να συνεργαστούν και να μην φερθούν εγωιστικά. Επιπλέον, η υπόθεση της μη-συνεργατικότητας των χρηστών μπορεί να θεωρηθεί και σαν το πιο απαισιόδοξο σενάριο.

Αξίζει να σημειωθεί ότι με τον όρο χρήστης μπορεί να εννοηθεί είτε ο ίδιος ο χρήστης, είτε κάποιος πράκτορας (agent) ο οποίος λαμβάνει αποφάσεις εκ μέρους του. Ο πράκτορας είναι ένα λογισμικό το οποίο εκτελείται, είτε στο τερματικό του χρήστη, είτε στην υπάρχουσα δικτυακή και πληροφοριακή υποδομή. Οι αποφάσεις που μπορεί να λαμβάνει ένας πράκτορας αφορούν κυρίως σε πολύ βραχυπρόθεσμο χρονικό ορίζοντα, απαλλάσσοντας έτσι το χρήστη από τον απαιτούμενο φόρτο αποφάσεων. Σημαντικό είναι ότι ο πράκτορας δε διαθέτει κάποια ευφυΐα, ώστε να παίρνει δικές του πρωτοβουλίες. Αντιθέτως, ακολουθεί ρητά κανόνες πολιτικής που έχει θέσει ο ίδιος ο χρήστης, ώστε να λαμβάνονται αυτοματοποιημένα οι βέλτιστες (στα πρότυπα του χρήστη) αποφάσεις.

Η χρήση πρακτόρων που αντιπροσωπεύουν πραγματικούς χρήστες μπορεί να θεωρηθεί ως μία τεχνολογία που θα οδηγήσει σε ένα νέο επίπεδο τη διαχείριση δικτυακών κι υπολογιστικών συστημάτων. Αυτό γιατί με τη δυνατότητα λήψης αποφάσεων σε μεγάλους ρυθμούς, πολλές παράμετροι οι οποίες σήμερα είναι στατικές και προκαθορισμένες (π.χ., η τιμολόγηση και τα κόστη σύνδεσης στο δίκτυο) θα μπορούν να μεταβάλλονται δυναμικά. Επίσης, θα καταστεί εφικτό η παροχή πόρων να γίνεται εξατομικευμένα στον κάθε χρήστη, με βάση τους κανόνες που έχει θέσει ο ίδιος. Αυτό θα έχει ως αποτέλεσμα και την μεγαλύτερη ικανοποίηση του χρήστη από τη λαμβανόμενη ποιότητα υπηρεσίας (quality of service – QoS) σε σχέση με το αντίστοιχο κόστος που πληρώνει. Επίσης, θα επιφέρει την αύξηση των κερδών των παρόχων, λόγω της *χρέωσης με διάκριση* (price discrimination), που ως γνωστόν αποτελεί τρόπο για την μεγιστοποίηση του καθαρού κέρδους [49].

Αξίζει να σημειωθεί ότι, κατά κανόνα, οι αποφάσεις που λαμβάνει ένας χρήστης σε ένα επικοινωνιακό ή υπολογιστικό σύστημα επηρεάζουν άμεσα ή έμμεσα τους άλλους χρήστες. Η εν λόγω αμοιβαία επιρροή έχει ως αποτέλεσμα τη δημιουργία πολύπλοκων δομών κι αλληλεπιδράσεων, που θυμίζουν έντονα πραγματικές κοινωνίες και *προβλήματα κοινωνικών αλληλεπιδράσεων* (social interaction problems). Τέτοιου είδους προβλήματα έχουν ήδη προσελκύσει το ενδιαφέρον μαθηματικών κι οικονομολόγων, κυρίως τον προηγούμενο αιώνα, με αποτέλεσμα την ανάπτυξη της Θεωρίας Παιγνίων [72]. Η Θεωρία Παιγνίων είναι μία μαθηματική θεωρία για την μοντελοποίηση, αλλά και τη μελέτη προβλημάτων αμοιβαίας επιρροής κι αντικρουόμενων συμφερόντων μεταξύ εγωιστικών οντοτήτων, που ονομάζονται και παίκτες κατά την ορολογία της εν λόγω θεωρίας. Η Θεωρία Παιγνίων έχει βρει εφαρμογές σε διάφορες Επιστήμες, με βασικότερη τα Οικονομικά, αλλά και σε άλλες όπως η Ψυχολογία [37]. Η εφαρμογή της σε ζητήματα Πληροφορικής κι Επικοινωνιών έχει, επίσης, προσελκύσει το ενδιαφέρον των ερευνητών, κυρίως τις τελευταίες δεκαετίες, με αξιόλογα αποτελέσματα σε διάφορα προβλήματα. Στην παρούσα διατριβή, η Θεωρία Παιγνίων αποτέλεσε βασικό εργαλείο για τη μελέτη των προβλημάτων που εξετάστηκαν.

Όπως αναφέρθηκε και πιο πάνω, η παρούσα διατριβή καλύπτει δύο βασικές περιοχές: 1) τη διαχείριση πόρων στον κινητό υπολογισμό, και 2) διαχείριση

κατανεμημένων υπολογιστικών συστημάτων. Πιο κάτω, περιγράφεται η συνεισφορά της διατριβής στα προαναφερθέντα πεδία.

2.1 Διαχείριση Πόρων στον Κινητό Υπολογισμό

Σε ό,τι αφορά τον κινητό υπολογισμό, στην παρούσα διατριβή, μελετήθηκαν διάφορα προβλήματα με κοινό άξονα τη διαχείριση των πόρων των κινητών δικτύων. Καθώς βασική υπόθεση υπήρξε η μη-συνεργατικότητα των χρηστών (ή πρακτόρων), τα εξεταζόμενα προβλήματα μελετήθηκαν υπό το πρίσμα της Θεωρίας Παιγνίων. Πιο συγκεκριμένα, μελετήθηκαν τα παρακάτω προβλήματα:

- 1) Εξισορρόπηση φόρτου σε ασύρματα κυψελωτά δίκτυα (cellular networks)
- 2) Αποφυγή διακοπών υπηρεσίας λόγω μεταπομπών
- 3) Ευέλικτη πρόσβαση σε μη-αδειοδοτημένες φασματικές περιοχές, και
- 4) Αποτελεσματικότητα και δικαιοσύνη στην πρόσβαση στο φυσικό μέσο

2.1.1 Εξισορρόπηση Φόρτου σε Ασύρματα Κυψελωτά Δίκτυα

Συχνό φαινόμενο αποτελεί κάποιες περιοχές κάλυψης ενός WLAN να είναι πιο δημοφιλείς από κάποιες άλλες. Αυτό έχει ως αποτέλεσμα οι χρήστες που συνδέονται με κάποιο AP το οποίο εξυπηρετεί μία δημοφιλή περιοχή να έχουν προβληματική σύνδεση με το Διαδίκτυο επηρεάζοντας αρνητικά ο ένας τον άλλον. Συγκεκριμένα, μπορεί να είναι συστηματικά δημοφιλείς (π.χ., κοντά σε κάποιο café σε ένα εμπορικό κέντρο), είτε προσωρινά (π.χ., σε μία πύλη αεροδρομίου πριν από την αναχώρηση ενός αεροπλάνου). Την ίδια στιγμή, υπάρχουν περιοχές οι οποίες δεν είναι δημοφιλείς (συστηματικά ή προσωρινά) με αποτέλεσμα σημαντικό μέρος των ασύρματων πόρων να μένει αδιάθετο. Αξίζει να σημειωθεί ότι, αν ένα μέρος από τους χρήστες που βρίσκονται σε δημοφιλείς περιοχές μεταβεί σε μη δημοφιλείς περιοχές, είναι βέβαιο ότι όλοι θα αποκομίσουν κάποιο όφελος σε σχέση με την ποιότητα υπηρεσίας που θα λαμβάνουν.

Στα πλαίσια της παρούσας διατριβής, υλοποιήθηκε ένα πρωτότυπο συστήματος, το MITOS, το οποίο στοχεύει στην εκμετάλλευση των αδιάθετων πόρων ενός WLAN με σκοπό τη βελτίωση της ποιότητας υπηρεσίας των κινητών χρηστών. Ο χρονικός ορίζοντας στον οποίο στοχεύει το MITOS δεν είναι ιδιαίτερα

μακροπρόθεσμος. Το MITOS, συγκεκριμένα, δεν έχει σκοπό να αντιμετωπίσει καταστάσεις συμφόρησης οι οποίες είναι κατά μεγάλο ποσοστό αναμενόμενες ή συστηματικές. Τέτοιες καταστάσεις είναι προτιμότερο να αντιμετωπίζονται με προσεκτικό σχεδιασμό του δικτύου και τοποθέτηση των AP στα κατάλληλα σημεία. Έτσι, αν μία περιοχή είναι συστηματικά κορεσμένη, είναι σημαντικό να τοποθετηθούν εκεί περισσότερα AP.

Το MITOS έχει ως σκοπό την αντιμετώπιση μάλλον απρόβλεπτων και, ουσιαστικά, προσωρινών καταστάσεων συμφόρησης που είναι αρκετά δύσκολο να αντιμετωπιστούν κατά το σχεδιασμό του WLAN. Ο μόνος τρόπος να αποφευχθούν είναι μέσω πλεονάζουσας υποδομής (overprovisioning), ώστε προληπτικά να αντιμετωπίσει καταστάσεις συμφόρησης. Μία τέτοια στρατηγική όμως δεν είναι ιδιαίτερα αποδοτική σε ό,τι αφορά το κόστος. Το MITOS, έχοντας μία διαρκή εικόνα του φόρτου σε κάθε AP, καθώς και της θέσης του κάθε κινητού χρήστη και του AP που είναι συνδεδεμένος, έχει τη δυνατότητα να διακρίνει αν κάποια AP υπολειτουργούν και ποιοι χρήστες θα μπορούσαν να επωφεληθούν από αυτά.

Η βασική ιδέα του MITOS είναι η παροχή προτάσεων προς τους χρήστες οι οποίοι υποφέρουν λόγω δικτυακής συμφόρησης. Τέτοιες προτάσεις περιέχουν οδηγίες για μία νέα θέση, από την οποία ο χρήστης θα μπορεί να συνδεθεί με ένα λιγότερο κορεσμένο AP, αλλά κι οδηγίες για να μεταβεί εκεί. Η παραγωγή προτάσεων από το MITOS λαμβάνει υπόψη το όφελος που θα αποκομίσει ο χρήστης λόγω της μετακίνησης σε μία λιγότερο κορεσμένη θέση, αλλά και το κόστος που συνεπάγεται μία τέτοια μετακίνηση. Έτσι, οι οδηγίες που προτείνει το MITOS σε ένα χρήστη αφορούν κατά κανόνα κοντινές περιοχές σε σχέση με την τρέχουσα θέση του, κάνοντας τέτοιες προτάσεις αρκετά ελκυστικές.

Πέραν από την υλοποίηση του πρωτότυπου MITOS, αξιολογήθηκε η βελτίωση που επιφέρει το MITOS σε ένα υποθετικό WLAN και μελετήθηκαν διάφορα ζητήματα που αφορούν στις πιθανές συμπεριφορές των χρηστών και το πώς αυτές επηρεάζουν το σύστημα. Ένα βασικό ερώτημα ήταν πως θα επηρέαζε η «προθυμία» των χρηστών να ακολουθούν προτάσεις, καθώς υπήρχε ο φόβος μη αποτελεσματικής εξισορρόπησης του φόρτου κι εμφάνισης ταλαντώσεων, όταν ένα μεγάλο ποσοστό χρηστών ακολουθεί τις οδηγίες του MITOS. Από πειράματα που έγιναν, αποδείχθηκε ότι αυτό είναι υπαρκτό πρόβλημα. Γι' αυτό το λόγο,

προτάθηκαν εναλλακτικές στρατηγικές παραγωγής προτάσεων, οι οποίες προστατεύουν το σύστημα από τέτοια φαινόμενα. Όλες οι στρατηγικές αξιολογήθηκαν κι εμφάνισαν «ανοσία» σε πιθανές παραμέτρους του συστήματος, αλλά και τη συμπεριφορά των χρηστών.

2.1.2 Αποφυγή Διακοπών Υπηρεσίας Λόγω Μεταπομπών

Ένα από τα κυριότερα προβλήματα για τους αλγόριθμους διαχείρισης πόρων σε κινητά κυψελωτά δίκτυα αποτελεί η εμφάνιση των μεταπομπών. Όταν συμβαίνει μία μεταπομπή, η σύνδεση που έχει ένας κινητός χρήστης με τον σταθμό βάσης (ΣΒ) συνεχίζεται σε ένα νέο ΣΒ. Αν η μεταπομπή είναι επιτυχής, η σύνδεση συνεχίζεται απρόσκοπτα, διαφανώς για το χρήστη. Είναι όμως πιθανόν η μεταπομπή να είναι αποτυχημένη, γεγονός που μπορεί να συμβεί όταν ο νέος ΣΒ δε διαθέτει αρκετούς πόρους ώστε να υποδεχθεί τη σύνδεση του χρήστη. Σε αυτήν την περίπτωση, έχουμε τη λεγόμενη διακοπή σύνδεσης λόγω μεταπομπής, η οποία θεωρείται πιο δυσάρεστη κι από την αδυναμία έναρξης μίας σύνδεσης (π.χ. λόγω συμφόρησης του δικτύου) [60]. Γι' αυτόν το λόγο, οι μηχανικοί δικτύων κι υπηρεσιών αναζητούν τρόπους να ελαχιστοποιήσουν ή ακόμα και για εξαλείψουν τα συμβάντα διακοπών συνδέσεων λόγω μεταπομπής.

Για την αποφυγή ή τουλάχιστον ελάττωση της συχνότητας των διακοπών λόγω μεταπομπής, είναι αναγκαία η προδραστική (proactive) δέσμευση πόρων στους ΣΒ με σκοπό την υποδοχή χρηστών από γειτονικές κυψέλες. Προς αυτήν την κατεύθυνση έχουν προταθεί διάφοροι αλγόριθμοι *ελέγχου αποδοχής κλήσεων* (call admission control – CAC), οι οποίοι βασίζονται σε *κανάλια ασφαλείας* (guard channels). Σύμφωνα με αυτούς τους αλγόριθμους, μία νέα σύνδεση δε γίνεται πάντα αποδεκτή από το ΣΒ, ακόμα κι αν υπάρχει αδιάθετο εύρος ζώνης (ή κανάλια), καθώς ένα μέρος από το εύρος ζώνης προορίζεται για εκχώρηση μόνο σε συνδέσεις από μεταπομπή. Το ποσό του εύρους ζώνης ασφαλείας μπορεί να είναι είτε σταθερό είτε να μεταβάλλεται δυναμικά, ανάλογα με τη συχνότητα των μεταπομπών. Το βασικό μειονέκτημα αυτών των αλγόριθμων είναι η «απρόσωπη» λειτουργία τους, δηλαδή το ότι εξετάζουν τις συνέπειες των μεταπομπών αθροιστικά, κι όχι σε επίπεδο ατομικού χρήστη.

Οι χρήστες στον κινητό υπολογισμό είναι κατά κανόνα «ανομοιογενείς». Η διαφορετικότητα αυτή εμφανίζεται σε διάφορα επίπεδα: 1) ο κάθε χρήστης μπορεί να έχει διαφορετικό πρότυπο κίνησης σε σχέση με τους υπόλοιπους, 2) ο χρήστης μπορεί να έχει διαφορετικές ανάγκες από τους άλλους χρήστες σε ό,τι αφορά το εύρος ζώνης (π.χ. ένας χρήστης μπορεί να χρησιμοποιεί κυρίως WWW, ενώ ένας άλλος video), και 3) κάθε χρήστης μπορεί να αντιλαμβάνεται διαφορετικά τη διακοπή μίας σύνδεσης του, π.χ. ένας να ενοχλείται πάρα πολύ, ενώ ένας άλλος λιγότερο. Είναι, λοιπόν, επιθυμητό η προδραστική δέσμευση πόρων να γίνεται ανά χρήστη κι όχι απρόσωπα, καθώς έτσι είναι εφικτό το δίκτυο να μοιράσει τους πόρους του πιο αποδοτικά, με βάση τις ανάγκες των χρηστών κάθε στιγμή.

Για την ατομική δέσμευση πόρων με δεδομένο το μοναδικό τρόπο κίνησης του κάθε χρήστη, είναι επιβεβλημένη η χρήση τεχνικών για πρόβλεψη της θέσης του χρήστη στο προσεχές μέλλον, ώστε να επιτυγχάνεται δέσμευση πόρων στους σωστούς ΣΒ. Στα πλαίσια της διατριβής αυτής αναπτύχθηκε ένας αλγόριθμος πρόβλεψης κίνησης χρήστη, με δεδομένη την ιστορία της κίνησης του κι αναπτύχθηκαν μέθοδοι διαπραγμάτευσης μεταξύ δικτύου και κινητού χρήστη, για την προδραστική εκχώρηση των πόρων.

Προτείνονται δύο μέθοδοι με τις οποίες μπορούν οι κινητοί χρήστες να προδεσμεύσουν πόρους σε ΣΒ στους οποίους προβλέπεται να συνδεθούν. Η πρώτη είναι βασισμένη στην ιδέα των *διαπραγματεύσεων* (bargaining). Σε αυτό το πλαίσιο, ο ΣΒ παίζει το ρόλο του «πωλητή» του εύρους ζώνης, ενώ ο κινητός χρήστης το ρόλο του «αγοραστή». Οι δύο οντότητες λειτουργούν ανταγωνιστικά, με το ΣΒ να προσπαθεί να «πουλήσει» όσο το δυνατόν πιο ακριβά το εύρος ζώνης, και με τον κινητό χρήστη να προσπαθεί να το πάρει όσο πιο «φθηνά» γίνεται. Η διαπραγμάτευση είναι μία διαδικασία που πραγματοποιείται σε βήματα, επαναλαμβανόμενα, ανά τακτά χρονικά διαστήματα, συγκλίνοντας πιθανώς σε μία αμοιβαία αποδεκτή τιμή εύρους ζώνης. Ο τρόπος με τον οποίο διαπραγματεύονται ο ΣΒ κι ο κινητός χρήστης επηρεάζεται από την τρέχουσα κατάσταση του δικτύου, αλλά κι από την χρονική κρισιμότητα του εύρους ζώνης για τον κινητό χρήστη.

Η δεύτερη μέθοδος που προτείνεται είναι βασισμένη σε *δημοπρασίες* (auctions). Στην περίπτωση αυτή ανταγωνιζόμενες οντότητες είναι οι κινητοί χρήστες μεταξύ τους, οι οποίοι προσπαθούν να πάρουν μερίδιο από τους πόρους των ΣΒ που

βρίσκονται γειτονικά τους. Ένας ΣΒ παίζει το ρόλο του *πλειστηριαστή* (auctioneer), ο οποίος συλλέγει προσφορές (bids), επιλέγει τους νικητές, εκχωρώντας τους πόρους που ζήτησαν, κι εισπράττοντας τα αντίστοιχα κέρδη. Σημαντικό στην προαναφερθείσα διαδικασία είναι ο τρόπος με τον οποίον οι κινητοί χρήστες πραγματοποιούν τις προσφορές τους. Λαμβάνονται υπόψη τρεις βασικοί παράγοντες: 1) η πιθανότητα να γίνει η μεταπομπή στο εξεταζόμενο προσεχές μέλλον, 2) το πλήθος των ενεργών συνδέσεων που θα έχει ο χρήστης κι η αντίστοιχη πιθανότητα, και 3) οι προσωπικές προτιμήσεις του χρήστη σχετικά με την κρισιμότητα κάθε τύπου εφαρμογής και των συνεπειών διακοπής της. Έτσι, οι προσφορές διαφοροποιούνται από χρήστη σε χρήστη, αλλά κι από μία χρονική στιγμή σε μία άλλη, με αποτέλεσμα το προτεινόμενο σχήμα να είναι πλήρως δυναμικό, αλλά κι ικανό να παρέχει διαφοροποίηση μεταξύ χρηστών (καλό και για τους ίδιους, αλλά και για το δίκτυο που έτσι μπορεί να αυξήσει τα κέρδη του).

Η αξιολόγηση των μεθόδων βασισμένων στις διαπραγματεύσεις αλλά και στις δημοπρασίες έδειξε ότι επιτυγχάνεται αξιόλογη μείωση της πιθανότητας συμβάντων διακοπής λόγω μεταπομπής. Ειδικότερα με τη χρήση των δημοπρασιών, διαφάνηκε η διαφοροποίηση που επιτυγχάνεται μεταξύ των χρηστών.

2.1.3 Ευέλικτη Πρόσβαση σε Μη-αδειοδοτημένες Φασματικές Περιοχές

Εκτός από τη διαχείριση πόρων για την εξισορρόπηση φόρτου σε ένα κινητό δίκτυο και της αποφυγής ή ελάττωσης των συμβάντων διακοπής λόγω μεταπομπών, στην παρούσα διατριβή, μελετήθηκε το πρόβλημα της *ευέλικτης ή δυναμικής φασματικής πρόσβασης*. Η δυναμική φασματική πρόσβαση αποτελεί ένα νέο ερευνητικό πεδίο, το οποίο έχει συγκεντρώσει την προσοχή πολλών ερευνητών παγκοσμίως.

Το έναυσμα για την ευέλικτη φασματική πρόσβαση δόθηκε από πρόσφατες μελέτες που διεξήγαγε η US Federal Communications Commission (FCC), οι οποίες απεκάλυψαν ότι, ακόμα και σε μεγάλες αστικές περιοχές, κατά μέσον όρο μόνο το 30% του αδειοδοτημένου φάσματος χρησιμοποιείται. Γι' αυτό το λόγο η FCC εξέδωσε μία ανακοίνωση θέτοντας νέους κανόνες για την πρόσβαση στο αδειοδοτημένο φάσμα [79]. Το Defense Advanced Research Project Agency

(DARPA), επίσης, ξεκίνησε το πρόγραμμα neXt Generation (XG) για την ανάπτυξη νέων ευέλικτων μηχανισμών και τεχνολογιών για το διαμοιρασμό του φάσματος [80]. Σύμφωνα και με τις δύο πρωτοβουλίες, αναγνωρίζεται η χρήση συσκευών ευέλικτης φασματικής πρόσβασης, οι οποίες είναι ικανές να αναγνωρίζουν και να χρησιμοποιούν δυναμικά ανενεργές ή υποχρησιμοποιούμενες ζώνες συχνοτήτων, οδηγώντας έτσι σε αποδοτική χρήση του υπάρχοντος φάσματος. Η εξέλιξη αυτή ανοίγει νέους δρόμους για τις ασύρματες υπηρεσίες που λειτουργούν σε μη-αδειοδοτημένες περιοχές συχνοτήτων και πάσχουν από την έλλειψη φασματικών πόρων.

Το πρόβλημα της δυναμικής φασματικής πρόσβασης είναι πολυδιάστατο, αναδεικνύοντας ένα σύνολο από ζητήματα διαχείρισης πόρων. Τα βασικότερα ζητήματα, όπως έχουν προσδιοριστεί και στην εργασία [5] είναι τα παρακάτω:

- 1) ανίχνευση φάσματος (spectrum sensing): εντοπισμός ανενεργών περιοχών φάσματος και διαμοιρασμός με τρόπο που να μην προξενεί επιζήμιες παρεμβολές στους πρωτεύοντες χρήστες,
- 2) διαχείριση φάσματος (spectrum management): επιλογή της καταλληλότερης φασματικής περιοχής, ούτως ώστε να μπορούν να ικανοποιηθούν οι επικοινωνιακές ανάγκες του χρήστη,
- 3) κινητικότητα φάσματος (spectrum mobility): διάφανη διατήρηση της επικοινωνίας κατά τη μετάβαση σε μία νέα περιοχή φάσματος, και
- 4) διαμοιρασμός φάσματος (spectrum sharing): διαμοιρασμός των φασματικών πόρων μεταξύ των συνυπαρχόντων δευτερευόντων χρηστών με δικαιοσύνη κι αποτελεσματικότητα.

Η μελέτη στην παρούσα διατριβή εστιάστηκε κυρίως στα ζητήματα (2) και (4), δηλαδή στην επιλογή κατάλληλων φασματικών περιοχών, αλλά και το διαμοιρασμό του φάσματος ανάμεσα σε συνυπαρχόντες χρήστες. Προς αυτήν την κατεύθυνση, μοντελοποιήθηκε η δυναμική φασματική πρόσβαση για δίκτυα *πολλαπλής πρόσβασης με διαίρεση κωδίκων* (code division multiple access – CDMA). Συγκεκριμένα, θεωρήθηκε ότι οι ασύρματοι χρήστες είναι μη-συνεργατικές οντότητες, που προσπαθούν να χρησιμοποιήσουν το υπάρχον φάσμα με τον καλύτερο τρόπο για το δικό τους και μόνον όφελος, αδιαφορώντας για τις

συνέπειες που επιφέρουν οι αποφάσεις τους στους υπόλοιπους χρήστες. Το μοντέλο αναλύθηκε με χρήση της Θεωρίας Παιγνίων και προτάθηκαν δύο εναλλακτικές κατανεμημένες στρατηγικές, ώστε να οδηγείται το σύστημα σε ισορροπία. Με τις εν λόγω στρατηγικές είναι δυνατόν το σύστημα να λειτουργεί αποδοτικά, χωρίς να χρειάζεται παρέμβαση από κάποια κεντρική οντότητα. Με βάση την αξιολόγηση που πραγματοποιήθηκε, φάνηκε ότι οι δύο προτεινόμενες στρατηγικές επιδεικνύουν αρκετά υψηλή ταχύτητα σύγκλισης, αλλά και ταχύτητα προσαρμογής μετά από συγκεκριμένα κρίσιμα συμβάντα (π.χ. κατάληψη μίας φασματικής περιοχής από έναν πρωτεύοντα χρήστη).

2.1.4 Αποτελεσματικότητα και Δικαιοσύνη στην Πρόσβαση στο Φυσικό Μέσο

Το τελευταίο ζήτημα που πραγματεύεται η παρούσα διατριβή, σε σχέση με τη διαχείριση πόρων στον κινητό υπολογισμό, αφορά στην πρόσβαση στο φυσικό ασύρματο μέσο. Η μελέτη εστιάζεται στο πρόβλημα του πως μπορεί μία ζώνη συχνοτήτων, η οποία διαμοιράζεται από ένα σύνολο χρηστών, να χρησιμοποιηθεί αποτελεσματικά και δίκαια.

Ο κάθε χρήστης επιθυμεί το σήμα του να λαμβάνεται στο δέκτη (π.χ. ΣΒ) που τον ενδιαφέρει με συγκεκριμένη ποιότητα, η οποία έχει βέβαια σχέση με την ισχύ εκπομπής του χρήστη. Όσο μεγαλύτερη είναι η ισχύς εκπομπής, τόσο μεγαλύτερη θα είναι κι η ισχύς λήψης στο δέκτη, με αποτέλεσμα καλύτερη αποκωδικοποίηση, μικρότερο ρυθμό σφαλμάτων bit (bit-error-rate – BER), μικρότερο ρυθμό σφαλμάτων πακέτων (packet-error-rate – PER), και τελικά υψηλότερη ρυθμοαπόδοση (throughput). Πρέπει να σημειωθεί, όμως, ότι η αύξηση της εκπομπής ισχύος ενός χρήστη έχει ως αποτέλεσμα την αύξηση των παρεμβολών προς τους υπόλοιπους χρήστες. Κάθε σήμα, που συμβάλλει με ένα σήμα το οποίο πρέπει να αποκωδικοποιηθεί σε έναν δέκτη, αποτελεί θόρυβο κι αιτία σφαλμάτων στη διαδικασία της αποκωδικοποίησης. Τα σφάλματα δε αυξάνονται καθώς αυξάνεται κι η ισχύς των λαμβανόμενων σημάτων-θορύβων.

Τα σήματα των χρηστών που εκπέμπουν ταυτόχρονα επηρεάζουν, λοιπόν, αρνητικά το ένα το άλλο. Ένας τρόπος να αποφευχθεί το συγκεκριμένο πρόβλημα είναι να χρησιμοποιηθεί μία τεχνική πολλαπλής πρόσβασης η οποία να αποτρέπει

τέτοιες παρεμβολές. Παραδείγματα τέτοιων τεχνικών είναι η *πολλαπλή πρόσβαση με διαίρεση συχνότητας* (frequency division multiple access – FDMA), καθώς κι η *πολλαπλή πρόσβαση με διαίρεση χρόνου* (time division multiple access – TDMA). Και στις δύο τεχνικές, η ζώνη συχνοτήτων χωρίζεται σε κανάλια (κανάλια συχνοτήτων στο FDMA – χρονοθυρίδες στο TDMA), κάθε ένα από τα οποία εκχωρείται σε έναν χρήστη αποκλειστικά. Έτσι, ποτέ δύο χρήστες δεν εκπέμπουν στην ίδια ακριβώς συχνότητα την ίδια χρονική στιγμή.

Μία πιο σύγχρονη τεχνική πολλαπλής πρόσβασης, όμως, η οποία τείνει να επικρατήσει της FDMA και TDMA, είναι η CDMA. Η CDMA έχει υιοθετηθεί στα κινητά συστήματα 3^{ης} γενιάς [81], καθώς θεωρείται πιο ευέλικτη από τις άλλες δύο τεχνικές. Σύμφωνα με τη CDMA τεχνική, οι χρήστες του συστήματος εκπέμπουν στην ίδια συχνότητα, την ίδια χρονική στιγμή. Δεν υφίστανται σαφώς διαχωρισμένα κανάλια, ούτε στη διάσταση του χρόνου, ούτε των συχνοτήτων. Τα σήματα των χρηστών διαχωρίζονται με τη βοήθεια μοναδικών ψηφιακών κωδικών που τους εκχωρούνται. Τα ψηφιακά δεδομένα του χρήστη, πριν εκπεμφθούν, πολλαπλασιάζονται με τον εν λόγω κωδικό παράγοντας ένα ψηφιακό σήμα με πολύ μεγαλύτερο εύρος ζώνης από το αρχικό. Ο δέκτης λαμβάνει όλα τα σήματα των χρηστών ταυτόχρονα και τα αποκωδικοποιεί, χρησιμοποιώντας τους κωδικούς που τους έχει εκχωρήσει. Συγκεκριμένα, πολλαπλασιάζοντας το συνολικό λαμβανόμενο σήμα με τον κωδικό ενός χρήστη, αναδεικνύεται το αρχικό σήμα του χρήστη, ενώ συμπιέζεται σημαντικά η ισχύς των σημάτων των υπολοίπων χρηστών.

Το πόσο καλά θα μπορέσει να αναδειχθεί το αρχικό ψηφιακό σήμα του χρήστη μετά τον πολλαπλασιασμό του συνολικού σήματος με τον κωδικό, εξαρτάται κυρίως από τρεις παράγοντες: 1) την απολαβή του CDMA κώδικα (processing gain) [81], 2) την ισχύ με την οποία φθάνει το σήμα του χρήστη στο δέκτη, και 3) την ισχύ με την οποία φθάνουν στο δέκτη τα σήματα-παρεμβολές των υπολοίπων χρηστών. Στο CDMA, λοιπόν, σε αντίθεση με το FDMA και το TDMA, οι χρήστες αλληλεπιδρούν μεταξύ τους μέσω των εκπομπών τους, έχοντας τη δυνατότητα να προκαλέσουν πρόβλημα ο ένας στον άλλο, αν αυξήσουν πολύ την ισχύ εκπομπής τους. Με άλλα λόγια η ποιότητα (BER, PER ή ρυθμοαπόδοση) που αντιλαμβάνονται οι χρήστες καθορίζεται άμεσα από το την ισχύ εκπομπής τους, τη

λεγόμενη κατανομή ισχύος (power allocation). Γι' αυτό το λόγο, στο CDMA, βασικός είναι ο ρόλος του έλεγχου ισχύος που σκοπό έχει να οδηγήσει σε κατανομές ισχύος που είναι δίκαιες κι αποτελεσματικές.

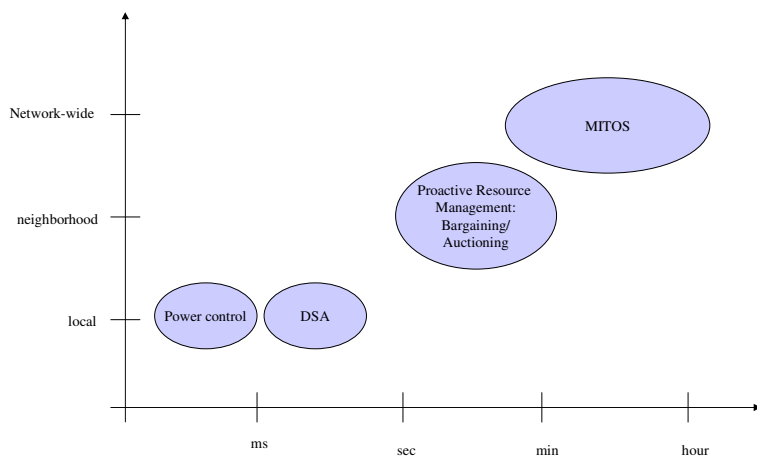
Το πρόβλημα ελέγχου ισχύος, στην παρούσα διατριβή, μοντελοποιήθηκε θεωρώντας ότι οι χρήστες είναι μη-συνεργατικοί, δηλαδή ότι ενδιαφέρονται αποκλειστικά για το προσωπικό τους όφελος, επιλέγοντας την ισχύ εκπομπής τους αυτόνομα. Το όφελος τους διαστασιοποιήθηκε με γνώμονα τη ρυθμοαπόδοση που επιτυγχάνουν, αλλά και την αντίστοιχη ισχύ εκπομπής (οι χρήστες είναι κατά βάση κινητοί κι επιθυμούν όσο το δυνατόν αυξημένο χρόνο ζωής της μπαταρίας του τερματικού τους). Το μοντέλο μελετήθηκε με τη βοήθεια της Θεωρίας Παιγνίων κι υπολογίστηκαν οι ισορροπίες που μπορεί να επιτευχθούν. Επειδή όμως οι συγκεκριμένες ισορροπίες δεν ήταν αποτελεσματικές, εφαρμόστηκε μία εναλλακτική λύση βασισμένη στη συνεργατική Θεωρία Παιγνίων, οδηγώντας τους χρήστες σε κατανομές ισχύος που είναι βέλτιστες και δίκαιες, επιτυγχάνοντας σημαντική βελτίωση σε σχέση με τη μη-συνεργατική περίπτωση.

2.1.5 Συνολική Άποψη της Μελέτης για τη Διαχείριση Πόρων στον Κινητό Υπολογισμό

Σε αυτήν την παράγραφο, συνοψίζεται η μελέτη που έγινε στα πλαίσια της παρούσας διατριβής σε ό,τι αφορά τη διαχείριση πόρων στον κινητό υπολογισμό. Τα θέματα που εξετάστηκαν, όπως αναλύθηκαν και στις προηγούμενες παραγράφους, είναι 1) η εξισορρόπηση του φόρτου σε ασύρματα κυψελωτά δίκτυα, 2) η αποφυγή διακοπών υπηρεσίας λόγω μεταπομπών, 3) η ευέλικτη πρόσβαση σε μη-αδειοδοτημένες περιοχές συχνοτήτων, και 4) η αποτελεσματικότητα και δικαιοσύνη στην πρόσβαση στο φυσικό μέσο. Τα παραπάνω ζητήματα μπορούν να θεωρηθούν ως συμπληρωματικά σε σχέση με τη διαχείριση πόρων ασύρματης πρόσβασης σε κινητά περιβάλλοντα.

Η συμπληρωματικότητα των μελετημένων ζητημάτων έγκειται σε τρεις παράγοντες: α) μπορούν να αντιστοιχηθούν λογικά σε διαφορετικά δικτυακά επίπεδα, β) αφορούν στη διαχείριση πόρων σε διαφορετικές χρονικές κλίμακες, και 3) αφορούν στη διαχείριση πόρων σε διαφορετικές χωρικά κλίμακες. Εξετάστηκαν θέματα που άπτονται του *φυσικού επιπέδου*, του *επιπέδου ελέγχου πρόσβασης στο μέσο*

(medium access control – MAC), αλλά κι υψηλότερων επιπέδων. Επίσης, άλλα ζητήματα αφορούσαν στη βραχυπρόθεσμη διαχείριση πόρων, ενώ άλλα σε μάλλον μεσοπρόθεσμο ορίζοντα. Τέλος, κάποια ζητήματα αφορούσαν στη τοπική διαχείριση πόρων και κάποια ενέπλεκαν και τη γειτονική υποδομή, ενώ κάποια άλλα αφορούσαν το κινητό δίκτυο στο σύνολο του.



Εικόνα 2.1. Λογική ταξινόμηση των προτεινόμενων στη διατριβή λύσεων διαχείρισης πόρων στον κινητό υπολογισμό, σε σχέση με τη χρονική και χωρική κλίμακα που εφαρμόζονται.



Εικόνα 2.2. Λογική ταξινόμηση των προτεινόμενων στη διατριβή λύσεων διαχείρισης πόρων στον κινητό υπολογισμό, σε σχέση με τη χωρική κλίμακα που εφαρμόζονται και το δικτυακό επίπεδο που αντιστοιχούν.

Στην Εικόνα 2.1 και στην Εικόνα 2.2, απεικονίζεται η λογική ταξινόμηση των λύσεων που προτείνονται στην παρούσα διατριβή, σχετικά με τη διαχείριση πόρων στο κινητό υπολογισμό. Στη μεν, οι λύσεις τοποθετούνται βάσει της χωρικής και

χρονικής κλίμακας εφαρμογής τους, ενώ, στη δε, τοποθετούνται με βάση τη χωρική κλίμακα εφαρμογής τους αναφορικά με το δικτυακό επίπεδο στο οποίο αντιστοιχούν. Η εν λόγω ταξινόμηση καθιστά φανερή τη συμπληρωματικότητα των ζητημάτων και των λύσεων που μελετήθηκαν στα πλαίσια της παρούσας διατριβής.

2.2 Διαχείριση Κατανεμημένων Υπολογιστικών Συστημάτων

Σε αυτό το υποκεφάλαιο, παρουσιάζονται τα ζητήματα που μελετήθηκαν σχετικά με τη διαχείριση κατανεμημένων υπολογιστικών συστημάτων. Ο κύριος άξονας της μελέτης αφορούσε συστήματα που σχετίζονται με την υπηρεσία WWW, όπως αρχιτεκτονικές WWW caching.

2.2.1 Διαχείριση Συστημάτων WWW Caching

Σε μία ιεραρχία WWW caching, οι caching proxies, αυτόνομα, αποθηκεύουν (cache) αντίγραφα των αντικειμένων που έχουν ζητηθεί από τους χρήστες του (τοπικού) δικτύου, ούτως ώστε, όταν κάποιο αντικείμενο ζητηθεί εκ νέου, αυτό να παραδίδεται κατευθείαν από τον caching proxy, χωρίς να γίνεται επικοινωνία με τον *πηγαίο διακομιστή WWW περιεχομένου* (origin server). Αυτή η τεχνική μπορεί να επιφέρει σημαντική επιτάχυνση στη πρόσβαση στο Διαδίκτυο, καθώς ένα μέρος των αντικειμένων που ζητούν οι χρήστες βρίσκονται ήδη πολύ κοντά σε αυτούς. Σημαντικό είναι επίσης ότι το WWW caching μπορεί να εφαρμοστεί σε διάφορα επίπεδα σε ιεραρχία (π.χ. Level 1 – L1, Level 2 – L2, ...).

Ένας βασικός περιορισμός στο WWW caching είναι ότι οι caching proxies καλούνται να αποθηκεύσουν κάθε αντίγραφο περιεχομένου που έχει ζητηθεί από τους χρήστες στο δίκτυο τους. Είναι προφανές, όμως, ότι το σύνολο όλων των περιεχομένων που μπορεί ή πρόκειται να ζητηθούν μπορεί εύκολα να ξεπεράσει τη αποθηκευτική χωρητικότητα του caching proxy. Επομένως, υπεισέρχεται ένα ζήτημα επιλογής μεταξύ αντιγράφων που θα αποθηκεύονται κι άλλων που θα απορρίπτονται.

Με δεδομένο ότι ο αντικειμενικός σκοπός ενός caching proxy είναι να ελαχιστοποιεί την καθυστέρηση που αντιλαμβάνονται οι χρήστες (ή ισοδύναμα να μεγιστοποιεί την πιθανότητα ένα ζητούμενο περιεχόμενο να είναι διαθέσιμο), τα αποθηκευμένα

αντικείμενα θα πρέπει να είναι τα πλέον δημοφιλή. Προς αυτήν την κατεύθυνση, έχουν προταθεί διάφοροι αλγόριθμοι αντικατάστασης (replacement algorithms), με σκοπό να απορρίπτονται αντίγραφα τα οποία δεν είναι δημοφιλή (η δημοτικότητα ορίζεται διαφορετικά για κάθε αλγόριθμο), ώστε να ελευθερώνεται χώρος για πιο δημοφιλή αντίγραφα. Πρέπει να σημειωθεί, όμως, ότι οι αλγόριθμοι αυτοί λειτουργούν απρόσωπα, δηλαδή, αντιμετωπίζουν τη ροή αιτήσεων συνολικά, προσπαθώντας να μεγιστοποιήσουν την *πιθανότητα ευστοχίας* (hit rate), χωρίς να εξετάζουν τη συνεισφορά, αλλά και την καθυστέρηση που αντιλαμβάνεται ο κάθε χρήστης ξεχωριστά.

Αξίζει να σημειωθεί, ότι από παρατηρήσεις σε δεδομένα από πραγματικούς caching proxies, χρήστες οι οποίοι επιδεικνύουν «επιθετική» συμπεριφορά, τείνουν να μονοπωλήσουν τον αποθηκευτικό χώρο του proxy, απολαμβάνοντας με αυτόν τον τρόπο υψηλές πιθανότητες ευστοχίας. Την ίδια στιγμή όμως, οι υπόλοιποι (μη-επιθετικοί χρήστες), υποχρεωμένοι να χρησιμοποιούν ένα μικρό μέρος της χωρητικότητας του proxy, υποφέρουν από χαμηλές πιθανότητες ευστοχίας κι αυξημένες καθυστερήσεις. Είναι, λοιπόν, αναγκαίο για να είναι πιο δίκαιη η χρήση της αποθηκευτικής χωρητικότητας του proxy να μην εφαρμόζονται καθολικές πολιτικές αντικατάστασης, αλλά η διαδικασία αυτή να λαμβάνει υπόψη τη συμπεριφορά του κάθε χρήστη ατομικά.

Στην παρούσα διατριβή, προτάθηκε ένας κατανεμημένος αλγόριθμος, ο οποίος είχε ως σκοπό να αποτρέψει την μονοπώληση της αποθηκευτικής χωρητικότητας ενός L1 caching proxy από μειοψηφίες «επιθετικών» χρηστών, ώστε να δώσει τη δυνατότητα σε «μικρότερους» χρήστες να αντιληφθούν υψηλότερες πιθανότητες ευστοχίας. Το προτεινόμενο σχήμα επιβράβευε, επίσης, τους χρήστες οι οποίοι έχουν συνεισφέρει στη συνολική πιθανότητα ευστοχίας (έχοντας ζητήσει την αποθήκευση αντιγράφων που αποδείχθηκαν δημοφιλή). Η αξιολόγηση του προτεινόμενου αλγορίθμου με προσομοίωση βασισμένη σε πραγματικές ροές αιτήσεων, έδειξε ότι μειώθηκε η διακύμανση της πιθανότητας ευστοχίας μεταξύ των χρηστών, γεγονός που υποδηλώνει ότι το caching γινόταν με πιο δίκαιο τρόπο.

2.2.2 Οικονομικό Πλαίσιο για τη Διαχείριση Ιεραρχιών Caching

Η ιδέα της δίκαιης παροχής της υπηρεσίας WWW caching, επεκτάθηκε και σε υψηλότερο επίπεδο στην ιεραρχία, συγκεκριμένα σε L2 caching. Το κίνητρο για τη μελέτη αυτή προέρχεται από την εργασία [6] που μελετά το φαινόμενο της «κακομεταχείρισης» σε κατανομημένες ομάδες από cache. Σύμφωνα με την προαναφερθείσα εργασία, υπό συγκεκριμένες συνθήκες, είναι δυνατόν μία cache με επιθετική συμπεριφορά να επιτύχει βελτιωμένη απόδοση για την ίδια, επηρεάζοντας, όμως, αρνητικά την απόδοση των άλλων. Το ίδιο πρόβλημα μπορεί να παρατηρηθεί κι όταν μία L2 cache χρησιμοποιείται από ένα σύνολο από L1 cache, στην ιεραρχία caching, και μία από αυτές, έχοντας επιθετική συμπεριφορά, υπονομεύει τις πιθανότητες ευστοχίας των άλλων cache προς δικό της όφελος.

Για την αποφυγή εγωιστικών συμπεριφορών από τις L1 cache, προτάθηκε ένα πλαίσιο, στο οποίο η L2 cache παρείχε τον αποθηκευτικό της χώρο, με χρέωση. Συγκεκριμένα, η L2 cache χώριζε το σύνολο του αποθηκευτικού της χώρου σε διακριτά τμήματα, καθένα από τα οποία «ενοικίαζε» στις ενδιαφερόμενες L1 cache. Το περιγραφόμενο πλαίσιο μοντελοποιήθηκε ως μία οικονομία (market), με σαφείς ρόλους για την κάθε οντότητα (η L2 cache ήταν ο πάροχος, ενώ οι L1 cache οι αγοραστές). Εξετάσθηκαν διάφορα μοντέλα οικονομιών, όπως η *χρέωση με διάκριση*, όπου η L2 cache επιτρεπόταν να χρεώνει με διαφορετική τιμή κάθε L1 cache, ούτως ώστε να μεγιστοποιήσει τα κέρδη της. Μελετήθηκε, επίσης, κι η περίπτωση της χρέωσης χωρίς διακρίσεις, όπου όλες οι L1 cache χρεώνονταν το ίδιο. Βάσει του μαθηματικού μοντέλου που αναπτύχθηκε, μελετήθηκαν, επίσης, διάφορα ζητήματα κινήτρων (incentive) που ανακύπτουν σε μία τέτοια οικονομία.

2.3 Δομή της Διατριβής

Στο παρόν υποκεφάλαιο, περιγράφεται η δομή του υπολοίπου της διατριβής. Στο Κεφάλαιο 3, γίνεται μία εισαγωγή σε έννοιες σχετικά με τη Θεωρία Παιγνίων, αλλά και σε σχετικούς συμβολισμούς. Καθώς η Θεωρία Παιγνίων αποτέλεσε βασικό εργαλείο στη μελέτη των προβλημάτων που παρουσιάστηκαν πιο πάνω, η ύπαρξη του αντίστοιχου κεφαλαίου κρίνεται αναγκαία, στην περίπτωση που ο αναγνώστης δεν είναι οικείος με τις έννοιες αυτές.

Από το Κεφάλαιο 4 μέχρι και το Κεφάλαιο 10, παρουσιάζεται το κύριο μέρος της μελέτης που διενεργήθηκε στα πλαίσια της παρούσας διατριβής. Για λόγους πιο αποτελεσματικής παρουσίασης, τα προαναφερθέντα κεφάλαια χωρίστηκαν σε δύο ομάδες. Συγκεκριμένα, το Μέρος Α περιλαμβάνει την ομάδα κεφαλαίων, από Κεφάλαιο 4, ως και Κεφάλαιο 8, και παρουσιάζει ζητήματα σχετικά με τη διαχείριση πόρων στον κινητό υπολογισμό. Στο Κεφάλαιο 4, παρουσιάζεται η μελέτη για το πρόβλημα της αντιμετώπισης τοπικού φόρτου, σε ασύρματα κυψελωτά δίκτυα (δες παράγραφο 2.1.1). Στο Κεφάλαιο 5, γίνεται η παρουσίαση του προβλήματος των διακοπών υπηρεσίας, λόγω μεταπομπών σε ασύρματα κυψελωτά δίκτυα (δες παράγραφο 2.1.2), καθώς και μίας λύσης βασισμένης σε διαπραγματεύσεις. Στο Κεφάλαιο 6, παρουσιάζεται μία εναλλακτική λύση στο ίδιο πρόβλημα με χρήση δημοπρασιών. Στο Κεφάλαιο 7, παρουσιάζεται η μελέτη που πραγματοποιήθηκε για την ευέλικτη πρόσβαση σε μη-αδειοδοτημένες φασματικές περιοχές (δες παράγραφο 2.1.3). Τέλος, στο Κεφάλαιο 8, περιγράφεται η μελέτη σχετικά με την απόδοση και δικαιοσύνη στη πρόσβαση στο φυσικό μέσο (δες παράγραφο 2.1.4).

Το Μέρος Β της διατριβής πραγματεύεται τα θέματα που έχουν σχέση με τη διαχείριση κατανεμημένων υπολογιστικών συστημάτων WWW caching και περιλαμβάνει το Κεφάλαιο 9 και το Κεφάλαιο 10. Στο Κεφάλαιο 9, παρουσιάζονται ένα μοντέλο για τη δίκαιη διαχείριση πόρων σε συστήματα WWW caching (δες παράγραφο 2.2.1). Στο Κεφάλαιο 10, παρουσιάζεται ένα Οικονομικό μοντέλο για το διαμοιρασμό της αποθηκευτικής χωρητικότητας σε μία ιεραρχία caching (δες παράγραφο 2.2.2).

ΚΕΦΑΛΑΙΟ 3

ΕΙΣΑΓΩΓΗ ΣΕ ΕΝΝΟΙΕΣ ΘΕΩΡΙΑΣ ΠΑΙΓΝΙΩΝ ΚΙ ΕΦΑΡΜΟΓΕΣ

Στο Κεφάλαιο αυτό, γίνεται μία αναφορά σε εισαγωγικές έννοιες, σε σχέση με τη Θεωρία Παιγνίων. Η θεωρία αυτή αποτέλεσε βασικό εργαλείο στη μελέτη των προβλημάτων που εξετάστηκαν, στα πλαίσια της παρούσας διατριβής. Πιο κάτω αναφέρονται βασικές πληροφορίες, όπως ιστορικά στοιχεία, διαφορετικές αρχές επίλυσης παιγνίων, καθώς και παραδείγματα εφαρμογών της θεωρίας σε δικτυακά προβλήματα.

3.1 Βασικά Στοιχεία Θεωρίας Παιγνίων

Η Θεωρία Παιγνίων μπορεί να οριστεί ως η μελέτη των μαθηματικών μοντέλων της αντιπαράθεσης (conflict) και συνεργασίας μεταξύ ευφυών και λογικών ληπτών αποφάσεων. Η Θεωρία Παιγνίων παρέχει γενικές μαθηματικές τεχνικές για την ανάλυση καταστάσεων, όπου δύο ή περισσότερα άτομα λαμβάνουν αποφάσεις που μπορούν να επηρεάσουν την ευημερία (welfare) των υπολοίπων. Οι καταστάσεις που εξετάζονται, κατά κανόνα, δεν αφορούν σε δραστηριότητες ψυχαγωγίας (όπως ο όρος παίγνιο αφήνει να εννοηθεί). Οι όροι «ανάλυση αντιπαράθεσης» (conflict analysis) ή «διαδραστική θεωρία αποφάσεων» (interactive decision theory) περιγράφουν πιο κατάλληλα το αντικείμενο μελέτης της θεωρίας.

Στη Θεωρία Παιγνίων, η έννοια του «παιγνίου» υποδηλώνει μία κοινωνική κατάσταση, όπου εμπλέκονται δύο ή περισσότερα άτομα. Τα άτομα αυτά καλούνται «παίκτες» (players). Για τους παίκτες γίνεται, κατά κανόνα, η υπόθεση ότι είναι λογικοί (rational). Ένας λήπτης αποφάσεων χαρακτηρίζεται λογικός εάν λαμβάνει αποφάσεις που βρίσκονται σε συνέπεια με τις επιδιώξεις του. Στη Θεωρία Παιγνίων, κάθε παίκτης έχει ως στόχο του την μεγιστοποίηση της ωφέλειας του (payoff), η οποία μετράται σε μία συγκεκριμένη κλίμακα.

Η ιδέα ότι ένας λογικός λήπτης αποφάσεων θα πρέπει να λαμβάνει αποφάσεις που μεγιστοποιούν την αναμενόμενη ωφέλεια ανήκει αρχικά στον Bernoulli (1738), ενώ η σύγχρονη θεώρηση του προβλήματος ανήκει στους John von Neumann και Oscar Morgenstern (1947). Έδειξαν, με ιδιαίτερα ασθενείς υποθέσεις, ότι για κάθε λογικό λήπτη αποφάσεων πρέπει να υπάρχει ένας τρόπος ανάθεσης αριθμητικών τιμών ωφέλειας (utility numbers) στα διάφορα πιθανά ενδεχόμενα που μπορεί να προκύψουν, ούτως ώστε να διαλέγει πάντοτε τη λύση που μεγιστοποιεί την αναμενόμενη του ωφέλεια. Αυτό το αποτέλεσμα καλείται θεώρημα μεγιστοποίησης της αναμενόμενης ωφέλειας (expected utility maximization theorem).

Τα λογικά αξιώματα τα οποία υποστηρίζουν το θεώρημα μεγιστοποίησης της αναμενόμενης ωφέλειας είναι υποθέσεις ασθενούς συνέπειας (weak consistency assumptions). Η βασική υπόθεση είναι το αξίωμα αντικατάστασης (substitution axiom), το οποίο μπορεί να αποδοθεί ως εξής: εάν ένας λήπτης αποφάσεων προτιμάει τη λύση 1 από την 2, όταν συμβεί το συμβάν A και προτιμάει την λύση 1 ακόμη κι αν δεν συμβεί η A, τότε θα προτιμήσει τη λύση 1 ακόμη κι αν δεν γνωρίσει εάν στο συμβάν A συμβεί ή όχι. Αυτή η υπόθεση είναι σε θέση να εξασφαλίσει ότι υπάρχει κάποια κλίμακα ωφέλειας στην οποία ο λήπτης αποφάσεων πάντα προτιμά τις επιλογές που μεγιστοποιούν την αναμενόμενη συνάρτηση ωφέλειας. Όταν υπάρχει αβεβαιότητα, η αναμενόμενη ωφέλεια μπορεί να καθοριστεί και να υπολογιστεί μόνο αν έχουν ανατεθεί πιθανότητες στα αβέβαια ενδεχόμενα συμβάντα.

Εν γένει, η μεγιστοποίηση της αναμενόμενης ωφέλειας δεν συμβαδίζει με την μεγιστοποίηση της χρηματικής απολαβής (monetary payoff). Ένα άτομο το οποίο αποφεύγει το ρίσκο (risk averse) μπορεί να ωφελείται περισσότερο από μία επιπλέον νομισματική μονάδα, όταν είναι φτωχός, παρά όταν είναι πλούσιος. Αυτή η παρατήρηση καθορίζει ότι, για ορισμένους λήπτες αποφάσεων, η ωφέλεια μπορεί να είναι μη γραμμική συνάρτηση της χρηματικής απολαβής. Γενικότερα, η ωφέλεια ενός ατόμου μπορεί να εξαρτάται από πολλαπλές μεταβλητές κι όχι μόνο από τη χρηματική αξία.

Η πρώτη εργασία στην περιοχή της Θεωρίας Παιγνίων αποδίδεται στον Γάλλο μαθηματικό Emil Borel ο οποίος πραγματεύτηκε την Θεωρία των Παιγνίων Στρατηγικής (Theory of Games of Strategy). Η εργασία των von Neumann και

Morgestern βασίστηκε στα αποτελέσματα του Borel. Οι λογικές ρίζες της Θεωρίας Παιγνίων εντοπίζονται στην Θεωρία Αποφάσεων του Bayes (Bayesian Decision Theory). Η Θεωρία Παιγνίων μπορεί να θεωρηθεί ως επέκταση της Θεωρίας Αποφάσεων (για δύο ή περισσότερους λήπτες αποφάσεων), ή ως η αναγκαία λογική εκπλήρωση της. Τα βασικά αποτελέσματα της Θεωρίας Αποφάσεων αναφέρονται στο ερώτημα αν ένα απλό ποσοτικό μοντέλο μπορεί να δώσει μία λογική περιγραφή της ανθρώπινης συμπεριφοράς. Ο οποιοσδήποτε λήπτης αποφάσεων, που ικανοποιεί ορισμένα αξιώματα, θα συμπεριφέρεται έτσι ώστε να μεγιστοποιεί την μαθηματική αναμενόμενη τιμή μίας συνάρτησης ωφέλειας, αναφορικά προς μία υποκειμενική πιθανοτική κατανομή. Δηλαδή, η συμπεριφορά ενός λογικού λήπτη αποφάσεων θα πρέπει να μπορεί να περιγραφεί από μία συνάρτηση ωφέλειας που δίνει ένα ποσοτικό χαρακτηρισμό των προτιμήσεων του αναφορικά με αποτελέσματα, και μία υποκειμενική πιθανοτική κατανομή που υποδηλώνει τις πεποιθήσεις του για όλους τους σχετικούς αστάθμητους παράγοντες. Όταν νέες πληροφορίες καθίστανται διαθέσιμες στο λήπτη αποφάσεων, οι υποκειμενικές πιθανότητες αναθεωρούνται βάσει του νόμου Bayes.

Οι αποφάσεις κάτω από αβεβαιότητα περιγράφονται από δύο μοντέλα: ένα πιθανοτικό μοντέλο (probability model) κι ένα μοντέλο καταστάσεων – μεταβλητών (state variable model). Σε κάθε περίπτωση, θεωρείται ότι ο λήπτης αποφάσεων επιλέγει μεταξύ διαφόρων λοταριών, αλλά τα δύο μοντέλα διαφέρουν στον ορισμό των λοταριών. Στο πιθανοκρατικό μοντέλο οι λοταρίες είναι κατανομές πιθανότητας πάνω σε ένα σύνολο τιμών. Στο μοντέλο καταστάσεων μεταβλητών οι λοταρίες είναι συναρτήσεις από ένα σύνολο από δυνατές καταστάσεις σε ένα σύνολο τιμών. Το κάθε μοντέλο είναι κατάλληλο για συγκεκριμένες εφαρμογές.

3.2 Συναρτήσεις Ωφελείας

Η συνάρτηση ωφελείας του παίκτη αντιπροσωπεύει τις προτιμήσεις του, πραγματοποιώντας μια αντιστοίχιση από μία κατάσταση του κόσμου (ή έκβαση του παιγνίου) σε έναν πραγματικό αριθμό. Όσο μεγαλύτερος είναι ο αριθμός αυτός, τόσο περισσότερο ικανοποιημένος είναι ο παίκτης από τη συγκεκριμένη κατάσταση. Στη Θεωρία Παιγνίων η συνάρτηση αυτή είναι γνωστή ως η συνάρτηση

ωφελείας Neumann-Morgenstern. Συγκεκριμένα, δεδομένου ότι το S αντιπροσωπεύει το σύνολο των καταστάσεων που μπορεί να αντιληφθεί ο παίκτης, η συνάρτηση ωφελείας του παίκτη i έχει την ακόλουθη μορφή:

$$u_i: S \rightarrow \mathfrak{R} \quad (3.1)$$

Πρέπει να σημειωθεί ότι οι καταστάσεις ορίζονται ως αυτές οι καταστάσεις που μπορεί να αντιληφθεί ο παίκτης. Για παράδειγμα, αν ένα ρομπότ έχει μόνον έναν αισθητήρα που το τροφοδοτεί με μία δυαδική είσοδο, π.χ. 1, αν έχει φως, και 0, αν είναι σκοτάδι, το ρομπότ θα έχει μία συνάρτηση ωφελείας ορισμένη σε δύο καταστάσεις μόνο, ανεξάρτητα από το πόσο περίπλοκος μπορεί να είναι ο πραγματικός κόσμος.

Στην πραγματικότητα, οι παίκτες έχουν σύνθετες εισόδους και δεν είναι ιδιαίτερα πρακτικό να οριστεί μία έξοδος για κάθε είσοδο. Έτσι, οι παίκτες πρέπει να αντιστοιχίσουν τις ακατέργαστες εισόδους τους σε ένα μικρότερο σύνολο καταστάσεων. Το να δημιουργηθεί αυτή η συνάρτηση μπορεί να αποδειχθεί δύσκολο, καθώς απαιτεί εις βάθος κατανόηση του μελετούμενου προβλήματος. Δεδομένης μίας συνάρτησης ωφελείας, είναι δυνατόν να οριστεί μία κατάταξη των πιθανών καταστάσεων, σε σχέση με το πόσο επιθυμητές είναι. Συγκρίνοντας τις τιμές ωφελείας δύο καταστάσεων, μπορεί να καθοριστεί ποια προτιμάει ο χρήστης. Η εν λόγω κατάταξη έχει τις ακόλουθες ιδιότητες:

- Αυτοπαθής (Reflexive): $u_i(s) \geq u_i(s)$
- Μεταβατική (transitive): Αν $u_i(a) \geq u_i(b)$ και $u_i(b) \geq u_i(c)$, τότε $u_i(a) \geq u_i(c)$.
- Συγκρίσιμη (Comparable): $\forall_{a,b}$ είτε $u_i(a) \geq u_i(b)$, είτε $u_i(a) \leq u_i(b)$.

Οι συναρτήσεις ωφελείας μπορούν να χρησιμοποιηθούν για την περιγραφή της συμπεριφοράς κάθε παίκτη, ποσοτικοποιώντας τις διάφορες ανταλλαγές (tradeoff) που αντιμετωπίζει, μαζί με την τιμή (ή αναμενόμενη τιμή) των ενεργειών του. Για παράδειγμα, μπορεί ένα ρομπότ να λαμβάνει μία συγκεκριμένη αμοιβή για να παραδώσει ένα πακέτο, αλλά η συγκεκριμένη ενέργεια παράδοσης να επιφέρει ένα αντίστοιχο κόστος. Αν μεταφραστούν όλες αυτές οι αμοιβές και τα κόστη σε αριθμούς μπορούν εύκολα να μελετηθούν οι ανταλλαγές μεταξύ τους.

Εφόσον έχει οριστεί μία συνάρτηση ωφελείας για όλους τους παίκτες, αυτό που έχουν να κάνουν είναι να επιλέξουν τις ενέργειες που μεγιστοποιούν την ωφέλεια τους. Όπως και στα Οικονομικά, η λέξη *εγωιστής* χρησιμοποιείται για να χαρακτηρίσει ένα λογικό παίκτη που στοχεύει στη μεγιστοποίηση της ωφελείας του. Πρέπει να σημειωθεί ότι αυτή η χρήση της λέξης διαφέρει ελαφρώς σε σχέση με την καθημερινή χρήση της που συχνά υπονοεί μία διάθεση πρόκλησης βλάβης στους άλλους. Ένας πραγματικά εγωιστής παίκτης ενδιαφέρεται αποκλειστικά για τη δική του ωφέλεια.

Πρέπει να σημειωθεί ότι η υπόθεση εγωιστών παικτών δεν αποκλείει τη μελέτη συνεργατικών δομών. Μία τέτοια δομή μπορεί να εξεταστεί ως μία κατάσταση όπου οι συναρτήσεις ωφελείας των παικτών έχουν οριστεί με τρόπο τέτοιο που οι παίκτες φαίνονται σαν να συνεργάζονται. Για παράδειγμα, αν ένας παίκτης λαμβάνει υψηλότερη ωφέλεια όταν βοηθάει τους άλλους παίκτες, η συμπεριφορά που θα προκύψει θα μπορεί να χαρακτηριστεί ως συνεργατική, από κάποιον εξωτερικό παρατηρητή, παρόλο που ο παίκτης λειτουργεί εγωιστικά.

3.3 Μη Συνεργατική Επίλυση Παιγνίων – Ισορροπία κατά Nash

Στόχος της Θεωρίας Παιγνίων, πέραν της μοντελοποίησης καταστάσεων ανταγωνισμού, ορίζοντας, για παράδειγμα το χώρο των στρατηγικών και τις συναρτήσεις ωφελείας των χρηστών, αποτελεί κι η επίλυση των παιγνίων που μοντελοποιούνται. Λέγοντας επίλυση, εννοείται ο υπολογισμός των στρατηγικών που μπορούν να επιλέξουν οι παίκτες σε μία ισορροπία. Για το σκοπό αυτό, έχουν προταθεί διάφορες αρχές επίλυσης που οδηγούν στο αποτέλεσμα (ή αποτελέσματα), στο οποίο είναι δυνατόν να ισορροπήσει ένα παίγνιο. Μία από τις πιο γνωστές αρχές αποτελεί η ισορροπία κατά Nash (Nash equilibrium – NE). Αποτελεί ένα από τα πιο ισχυρά, δημοφιλή, αλλά κι αμφιλεγόμενα εργαλεία της Θεωρίας Παιγνίων [13].

Η ισορροπία κατά Nash ορίζεται ως η κατάσταση στην οποία κανείς παίκτης δεν έχει συμφέρον να αλλάξει στρατηγική, δεδομένου ότι οι υπόλοιποι χρήστες θα παραμείνουν στη δική τους στρατηγική. Αν οριστεί η έννοια της *καλύτερης*

απόκρισης s_i^{BR} , για το παίκτη i , δεδομένων των στρατηγικών των άλλων παικτών $(s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_N) \equiv \mathbf{s}_{-i}$ ως

$$s_i^{BR} = \operatorname{argmax}_{s \in S_i} u_i(s_1, \dots, s_{i-1}, s, s_{i+1}, s_N) \equiv \operatorname{argmax}_{s \in S_i} u_i(s, \mathbf{s}_{-i}), \quad (3.2)$$

τότε η ισορροπία κατά Nash ορίζεται ως εξής:

Ορισμός 3.1: Μία έκβαση του παιγνίου $(s_1^{NE}, \dots, s_N^{NE})$ αποτελεί μία ισορροπία κατά Nash, αν και μόνον αν το s_i^{NE} είναι μία καλύτερη απόκριση στις στρατηγικές $(s_1^{NE}, \dots, s_{i-1}^{NE}, s_{i+1}^{NE}, \dots, s_N^{NE})$, για κάθε i . Δηλαδή, για όλα τα i , ισχύει

$$u_i(s_i^{NE}, \mathbf{s}_{-i}^{NE}) \geq u_i(s, \mathbf{s}_{-i}^{NE}), \quad \forall s_i \in S_i$$

Με άλλα λόγια, στο NE, κανείς παίκτης δεν έχει κάποιο συμφέρον να αποκλίνει, αν γνωρίζει ποιες στρατηγικές έχουν επιλέξει οι άλλοι παίκτες.

Έστω, για παράδειγμα το παίγνιο συντονισμού (coordination game) [14] που φαίνεται στην Εικόνα 3.1, όπου δύο επιχειρήσεις (A και B) πρέπει να αποφασίσουν ταυτόχρονα αν θα παράγουν υπολογιστές με μικρές δισκέτες (στρατηγική μ) ή μεγάλες δισκέτες (στρατηγική M). Αν κι οι δυο επιχειρήσεις επιλέξουν την ίδια δισκέτα, πετυχαίνουν κέρδη, ενώ αν διαλέξουν διαφορετικές θα έχουν απώλειες. Είναι προφανές, λοιπόν, ότι είναι καλύτερο να συντονιστούν, και μάλιστα σε μικρές δισκέτες, παρά σε μεγάλες (αφού τα κέρδη είναι μεγαλύτερα). Το παίγνιο αυτό είναι παίγνιο συντονισμού, διότι κι οι δύο εταιρείες έχουν κίνητρο να ακολουθήσουν την ίδια στρατηγική.

(B)

	M	μ
(A) M	1,1	-1,-1
μ	-1,-1	2,2

Εικόνα 3.1. Στρατηγική μορφή παιγνίου συντονισμού

Το συγκεκριμένο παίγνιο έχει δύο ισορροπίες κατά Nash. Η πρώτη είναι κι οι δύο επιχειρήσεις να παράγουν μεγάλες δισκέτες, δηλαδή $(s_A, s_B) = (M, M)$, ενώ η δεύτερη να παράγουν κι οι δύο μικρές, δηλαδή $(s_A, s_B) = (\mu, \mu)$. Το ότι τα δύο εν λόγω σημεία είναι ισορροπίες κατά Nash είναι εύκολο να δειχθεί, καθώς καμία

επιχείρηση δεν έχει συμφέρον να αποκλίνει από αυτά, αν η άλλη παραμείνει σταθερή στη στρατηγική της. Το ζήτημα, όμως, που δημιουργείται είναι ποια από τις υπάρχουσες ισορροπίες θα αποτελέσει τη λύση του παιγνίου. Η απάντηση είναι ότι κι οι δύο ισορροπίες μπορούν να προκύψουν. Δεν ισχύει ότι η μία ισορροπία είναι πιο ισορροπία από την άλλη. Υπάρχει μία λογική, η οποία είναι εκτός παιγνίου και λέγεται *focal point* (σημείο εστίασης), σύμφωνα με την οποία είναι δυνατόν να συντονίζονται οι παίκτες, ώστε να πετυχαίνουν συγκεκριμένες ισορροπίες, οι οποίες είναι περισσότερο επωφελείς από άλλες.

Στο παίγνιο, στην Εικόνα 3.1, φαίνεται ξεκάθαρα ότι κι οι δύο παίκτες θέλουν να πάνε στο (μ, μ) , διότι έτσι θα έχουν περισσότερα κέρδη. Κάθε εταιρεία είναι, επίσης, σε θέση να πιστεύει ότι κι ο αντίπαλος της έχει την ίδια επιθυμία, οπότε, κι οι δύο μπορούν να επιλέξουν ταυτόχρονα το συνδυασμό στρατηγικών (μ, μ) και, έτσι, να πετύχουν κι οι δύο υψηλότερα κέρδη (χωρίς να χάνει κανείς). Με άλλα λόγια, η ισορροπία (μ, μ) είναι κατά Pareto κυρίαρχη (Pareto dominant) από την ισορροπία (M, M) . Η επιλογή της (μ, μ) είναι, λοιπόν, ξεκάθαρη.

Γενικά, όταν είναι δυνατόν να καταταχθούν κατά Pareto (Pareto ranking) οι υπάρχουσες ισορροπίες, οι παίκτες μπορούν εύκολα να συντονιστούν και να επιτύχουν την καλύτερη δυνατή. Κατά κανόνα, όμως, οι πιθανές ισορροπίες δεν είναι υποχρεωτικό να μπορούν να καταταχθούν κατά Pareto. Αυτό σημαίνει ότι μία ισορροπία μπορεί να είναι καλύτερη για ένα παίκτη από μία άλλη, αλλά χειρότερη για κάποιον άλλο παίκτη. Σε τέτοιες περιπτώσεις, δεν είναι δυνατόν να συντονιστούν οι παίκτες, καθώς τα συμφέροντα τους αντικρούονται. Επιθυμητό χαρακτηριστικό, λοιπόν, σε παίγνια είναι η ύπαρξη μοναδικής ισορροπίας Nash, ώστε να μπορούν να μελετηθούν κι επιλυθούν ξεκάθαρα.

Ένα ακόμα σημαντικό στοιχείο των ισορροπιών Nash αποτελεί η αποτελεσματικότητά τους. Ένα σύνηθες κριτήριο αποτελεσματικότητας είναι η κατά Pareto αποτελεσματικότητα (Pareto efficiency), η οποία ορίζεται παρακάτω.

Ορισμός 3.2 (Κατά Pareto αποτελεσματικότητα): Το σημείο $\mathbf{u} \in U$, όπου U είναι το σύνολο των επιτεύξιμων ωφελειών, είναι κατά Pareto αποτελεσματικό αν, για κάθε $\mathbf{u}' \in U$, $\mathbf{u}' \geq \mathbf{u}$, τότε $\mathbf{u}' = \mathbf{u}$.

Με άλλα λόγια, ένα σημείο είναι κατά Pareto αποτελεσματικό αν δεν είναι δυνατόν να βρεθεί ένα άλλο σημείο, το οποίο να δίνει ανώτερα επίπεδα ωφέλειας για όλους τους παίκτες. Δυστυχώς, οι ισορροπίες κατά Nash, συνήθως, δεν είναι κατά Pareto αποτελεσματικές, καθώς μπορούν να βρεθούν άλλα σημεία περισσότερο επωφελή για όλους τους παίκτες. Έτσι, ενώ οι ισορροπίες κατά Nash έχουν το πλεονέκτημα ότι αποτελούν σταθερά σημεία λειτουργίας σε κατανεμημένα περιβάλλοντα, πάσχουν σε ό,τι αφορά την αποτελεσματικότητα. Οπότε μερικές φορές είναι χρήσιμο να καταφεύγει κανείς σε κεντρικό έλεγχο του συστήματος, ώστε να επιβάλλει πιο επωφελή σημεία λειτουργίας για τους παίκτες.

3.4 Κατανεμημένη Σύγκλιση σε Ισορροπίες Παιγνίων

Εκτός από αποτελεσματικότητα των Nash ισορροπιών ενός παιγνίου, σημαντικό ζήτημα, επίσης, αποτελεί και το πώς μπορούν στην πράξη να επιτευχθούν αυτές οι ισορροπίες, κατανεμημένα. Κατά κανόνα, μία κατά Nash ισορροπία είναι μία κατάσταση που προκύπτει φυσιολογικά, όταν οι παίκτες λειτουργούν βάσει της ιδέας της καλύτερης απόκρισης (δες (3.2)). Σε μία τέτοια κατάσταση, κάθε παίκτης, συνήθως ακολουθιακά, επιλέγει τη στρατηγική που εξυπηρετεί βέλτιστα το συμφέρον του, δεδομένων των στρατηγικών των άλλων παικτών. Το ίδιο κάνουν κι οι υπόλοιποι παίκτες, με αποτέλεσμα, σταδιακά, να οδηγείται το σύστημα σε μία κατάσταση από την οποία κανείς παίκτης δεν έχει συμφέρον να αποκλίνει, δηλαδή μία ισορροπία Nash. Υπάρχουν, όμως, περιπτώσεις στις οποίες δεν είναι δυνατόν, μέσω της μεθόδου της καλύτερης απόκρισης, να επιτευχθεί μία ισορροπία Nash, αλλά παρατηρούνται ταλαντώσεις.

Μία περίπτωση στην οποία παρατηρείται το πρόβλημα της αδυναμίας σύγκλισης σε μία ισορροπία Nash (παρόλο που μπορεί να υπάρχει), είναι όταν οι παίκτες αποφασίζουν την ίδια στιγμή ποια στρατηγική θα ακολουθήσουν, δεδομένης κοινής πληροφορίας, και χωρίς να έχουν τρόπο να συνεννοηθούν, εκ των προτέρων, μεταξύ τους. Ένα γνωστό πρόβλημα στο οποίο παρατηρείται το συγκεκριμένο φαινόμενο είναι το *Santa Fe bar problem* (SFBP) [27], το οποίο περιγράφεται στην παράγραφο 3.4.1, και το παίγνιο διασποράς (dispersion game – DG) [32], που περιγράφεται στην παράγραφο 3.4.2.

3.4.1 Santa Fe Bar Problem

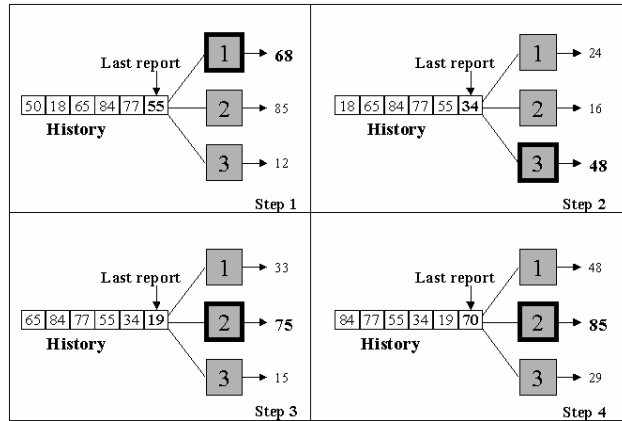
Το SFBP [27], [28] είναι γνωστό στους παιγνιοθεωρητικούς ως μία αφαίρεση του γενικότερου προβλήματος διαμοιρασμού πεπερασμένων πόρων. Η μόνη αλληλεπίδραση μεταξύ των παικτών πραγματοποιείται μέσω της από κοινού χρήσης των διαμοιραζόμενων πόρων. Το SFBP το εισήγαγε ο οικονομολόγος B. Arthur, στη μελέτη του σχετικά με την περιορισμένη λογική (bounded rationality) και τον επαγωγικό συμπερασμό [27]. Το σενάριο που περιγράφει το παιγνιοθεωρητικό μοντέλο έχει ως εξής:

Ν άτομα αποφασίζουν ανεξάρτητα κάθε εβδομάδα για το αν θα πάνε σε ένα μπαρ που είναι ανοικτό μία συγκεκριμένη βραδιά... Ο χώρος είναι περιορισμένος κι η βραδιά είναι ευχάριστη αν το μπαρ δεν είναι υπερβολικά γεμάτο – ειδικότερα, αν λιγότεροι από το 60% των πιθανών Ν ατόμων έχουν προσέλθει ... ένα άτομο θεωρεί ότι αξίζει να πάει στο μπαρ αν περιμένει ότι θα πάνε λιγότεροι από 60%. Οι επιλογές δεν επηρεάζονται από προηγούμενες επισκέψεις· δεν υπάρχει καμία συνεργία ή προηγούμενη επικοινωνία μεταξύ των εμπλεκόμενων ατόμων· η μόνη διαθέσιμη πληροφορία είναι ο αριθμός των επισκεπτών τις προηγούμενες εβδομάδες.

Όπως ειπώθηκε και στην εργασία [27], για να οδηγηθούν τα επίπεδα προσέλευσης σε ισορροπία, είναι αναγκαίο τα εμπλεκόμενα άτομα να έχουν διαφορετικό «τρόπο σκέψης» (δηλαδή οι εκτιμήσεις τους θα πρέπει να διαφέρουν). Διαφορετικά, αν η προσέλευση στο μπαρ ήταν μικρότερη από 60% την προηγούμενη εβδομάδα, όλοι θα επέλεγαν να επισκεφθούν το μπαρ την τρέχουσα εβδομάδα. Αυτό όμως θα προκαλούσε χαμηλή προσέλευση την επόμενη εβδομάδα, καθώς η προσέλευση ξεπέρασε το 60% κοκ. Ως εκ τούτου, το σύστημα θα υποφέρει από εκτεταμένες ταλαντώσεις. Για την αντιμετώπιση του προβλήματος, στην εργασία [27], προτείνεται η χρήση προγνωστών προσέλευσης, οι οποίοι εκτιμούν τη μελλοντική προσέλευση, βάσει της παρατηρούμενης ιστορίας της. Διαφορετικοί προγνώστες μπορούν να χρησιμοποιηθούν από τα διάφορα άτομα για να επιτευχθεί επαρκής ετερογένεια στις προβλέψεις τους και, επομένως, η επιθυμητή διαφοροποίηση στον «τρόπο σκέψης τους».

Κάθε άτομο επιλέγει τυχαία έναν αριθμό από προγνώστες από το σύνολο των διαθέσιμων προγνωστών. Έτσι, κάθε άτομο κατέχει ένα συγκεκριμένο σύνολο από προγνώστες, πιθανώς διαφορετικό από το σύνολο που έχει επιλέξει ένας άλλος

χρήστης, παρέχοντας του ένα διαφορετικό τρόπο σκέψης. Κάθε δεδομένη στιγμή, μόνο ένας προγνώστης ανά άτομο είναι ενεργός. Ένας προγνώστης ενεργοποιείται αν ήταν αυτός με την ακριβέστερη πρόβλεψη σε σχέση με την προσέλευση που τελικά σημειώθηκε (δες παράδειγμα στην Εικόνα 3.2).



Εικόνα 3.2. Ενεργοποίηση κι απενεργοποίηση προγνωστών. Στο βήμα 1, λαμβάνεται η τελευταία μέτρηση φόρτου (55) κι ο Προγνώστης 1 ενεργοποιείται. Στο βήμα 2, λαμβάνεται μία νέα μέτρηση φόρτου (34). Από τις προβλέψεις του προηγούμενου βήματος, παρατηρείται ότι ο Προγνώστης 3 ήταν πιο κοντά σε αυτή (12 στο βήμα 1), και γι' αυτό σημειώνεται ως ενεργός. Στη συνέχεια, δεδομένης της ιστορίας και της νέας μέτρησης, οι προγνώστες επιστρέφουν νέες προβλέψεις. Στο βήμα 3, όταν λαμβάνεται η νέα μέτρηση (19), ο Προγνώστης 2 παρατηρείται ότι είναι ο πιο ακριβής (16 στο βήμα 2), και, έτσι, ενεργοποιείται. Ο Προγνώστης 2 επιλέγεται ως ενεργός και για το βήμα 4.

Πειράματα στην εργασία [27] έδειξαν ότι η εν λόγω προσέγγιση επιτυγχάνει να διατηρήσει την προσέλευση του μπαρ κοντά στην επιθυμητή ισορροπία του 60%, με μικρές σχετικά αποκλίσεις γύρω από αυτή. Η ίδια προσέγγιση αποδείχθηκε επίσης αποτελεσματική στην περίπτωση με περισσότερα από ένα μπαρ, που είναι γνωστή ως το “New York City Bar Problem”, [28].

Μπορεί εύκολα να παρατηρηθεί ότι το SFBP έχει άμεση σχέση με προβλήματα κατανομημένης εξισορρόπησης φόρτου σε δίκτυα κι υπολογιστικά συστήματα. Σε τέτοια προβλήματα, αν οι παίκτες που διαμοιράζονται ένα σύνολο κοινών πόρων (π.χ. σύνολο ζωνών συχνοτήτων, ή σύνολο διακομιστών σε ένα δίκτυο διανομής περιεχομένου), η χρήστη στρατηγικών καλύτερης απόκρισης, πιθανότατα, θα οδηγήσει σε ταλαντώσεις, υπονομεύοντας την απόδοση του εκάστοτε συστήματος. Η λύση που προτείνεται για το SFBP, στην εργασία [27], στα πλαίσια της παρούσας διατριβής, αποτέλεσε βασικό εργαλείο για τη λύση διαφορετικών

προβλημάτων διαχείρισης πόρων, όπως περιγράφεται στο Κεφάλαιο 4, Κεφάλαιο 7, και Κεφάλαιο 11.

3.4.2 Παίγνια Διασποράς

Μία αρκετά μελετημένη κατηγορία παιγνίων αποτελούν τα παίγνια συντονισμού, παίγνια στα οποία οι παίκτες που αποφασίζουν ταυτόχρονα έχουν κέρδος μόνον όταν επιλέγουν την ίδια ενέργεια (δες Εικόνα 3.1). Μία συμπληρωματική κατηγορία που έχει συγκεντρώσει σχετικά μικρή προσοχή αφορά σε παίγνια στα οποία οι παίκτες έχουν κέρδος μόνον όταν επιλέγουν διαφορετικές ενέργειες· αυτά τα παίγνια αποκαλούνται μερικές φορές παίγνια *αποσυντονισμού* (*anticoordination games*). Η περισσότερη μελέτη σε σχέση με αυτά τα παίγνια έχει επικεντρωθεί μόνο στην περίπτωση των δύο παικτών (δες Εικόνα 3.3), όπου το παίγνιο συντονισμού και το παίγνιο αποσυντονισμού διαφέρουν μόνο στα ονόματα των ενεργειών των παικτών. Όμως, με αυθαίρετο πλήθος παικτών κι ενεργειών, τα δύο παίγνια διαφέρουν ουσιαστικά. Συγκεκριμένα, ενώ η γενίκευση του παιγνίου συντονισμού μπορεί να γίνει αρκετά ευθέως, η γενίκευση του παιγνίου αποσυντονισμού είναι πιο πολύπλοκη.

Στην εργασία [32], η γενίκευση του παιγνίου αποσυντονισμού ονομάστηκε παίγνιο διασποράς, καθώς αποτελεί ένα παίγνιο στο οποίο οι παίκτες επιδιώκουν να διαφοροποιούνται όσο γίνεται περισσότερο σε ό,τι αφορά τις ενέργειες που επιλέγουν. Παρόλο που μπορεί κανείς να μετασχηματίσει ένα παίγνιο διασποράς σε ένα παίγνιο συντονισμού, στο οποίο οι παίκτες συντονίζονται για να επιτύχουν τον πλήρη διασκορπισμό τους σε σχέση με τις υπάρχουσες ενέργειες, το πλήθος των επιδιωκόμενων αποτελεσμάτων αυξάνεται εκθετικά με το πλήθος των παικτών.

	(B)	
	M	μ
(A) M	-1,-1	1,1
μ	2,2	-1,-1

Εικόνα 3.3. Στρατηγική μορφή παιγνίου αποσυντονισμού

Τα DG μοντελοποιούν διάφορα προβλήματα που μπορούν να συναντηθούν στην πραγματικότητα. Ίσως η πιο φυσική εφαρμογή αποτελεί το πρόβλημα της εξισορρόπησης φόρτου. Αυτό το πρόβλημα μπορεί να μοντελοποιηθεί ως ένα DG, στο οποίο οι παίκτες είναι οι χρήστες, οι πιθανές ενέργειες είναι οι διαθέσιμοι πόροι (π.χ. διακομιστές), κι οι ισορροπίες του παιγνίου είναι οι εκβάσεις στις οποίες οι παίκτες είναι πλήρως διασκορπισμένοι. Μία άλλη φυσική εφαρμογή των DG είναι το πρόβλημα *niche selection* που μελετάται στα Οικονομικά και την εξελικτική βιολογία. Σε ένα γενικό πρόβλημα *niche selection*, καθένας από n ολιγοπωλιακούς παραγωγούς επιθυμεί να καταλάβει μία από k διαφορετικές γωνίες, κι οι παραγωγοί επιθυμούν να καταλάβουν «γωνίες» με λιγότερους ανταγωνιστές. Άλλα προβλήματα *niche selection* είναι το SFBP (δες παράγραφο 3.4.1) και τα *παίγνια μειονότητας* (minority games).

Μπορεί εύκολα να διαπιστωθεί ότι, αν οι παίκτες αποφασίζουν, ταυτόχρονα, με μία λογική καλύτερης απόκρισης, για το ποια ενέργεια θα πραγματοποιήσουν (ποιον πόρο θα επιλέξουν), στο σύστημα θα παρατηρηθούν εκτεταμένες ταλαντώσεις, πρόβλημα που συζητήθηκε και στο υποκεφάλαιο 3.4.1. Για το λόγο αυτό έχουν προταθεί διάφορες στρατηγικές, με σκοπό την αποφυγή του ανεπιθύμητου αυτού φαινομένου. Τέτοιες στρατηγικές βασίζονται σε μαθησιακούς κανόνες, π.χ. *Fictitious play* [20] ή *Reinforcement learning* [21], καθώς και σε πιθανοτικές προσεγγίσεις. Τέτοιου τύπου στρατηγικές χρησιμοποιήθηκαν, στα πλαίσια της παρούσας διατριβής, για τη λύση διαφορετικών προβλημάτων διαχείρισης πόρων, όπως περιγράφεται στο Κεφάλαιο 7.

3.5 Συνεργατική Επίλυση Παιγνίων – Λύση Διαπραγμάτευσης του Nash

Όπως έχει αναφερθεί στο υποκεφάλαιο 3.3, μία ισορροπία Nash αποτελεί τη λύση ενός μη συνεργατικού παιγνίου, και, στις περισσότερες περιπτώσεις, δεν αποτελεί ένα κατά Pareto αποτελεσματικό σημείο. Συνεπώς, η συνεργασία μεταξύ των παικτών μπορεί να αποφέρει από κοινού οφέλη. Παρόλ' αυτά η απόλυτη συνεργασία είναι δύσκολο να επιτευχθεί ή είναι απλά μη επιθυμητή. Σε τέτοιες περιπτώσεις, οι παίκτες μπορεί να ενδιαφέρονται να επιτύχουν την πιο συμφέρουσα (σε σχέση με το προσωπικό τους συμφέρον) συμφωνία που είναι

δυνατόν, ελαχιστοποιώντας παράλληλα το ρίσκο να μην επιτευχθεί καθόλου συμφωνία. Σε κάποιο βαθμό, αυτοί είναι αντικρουόμενοι στόχοι και ένας διαιτητής (arbitrator) ή ένα σχήμα διαιτησίας (arbitration scheme) μπορεί να είναι απαραίτητο για να προσφέρει μία ικανοποιητική λύση σε μία μη συνεργατική διαπραγμάτευση. Ο Luce και ο Raiffa [37] ορίζουν ένα σχήμα διαιτησίας ως μία συνάρτηση ή κανόνα που παρέχει μία μοναδική αμοιβή στους εμπλεκόμενους σε μία αντιπαράθεση παίκτες. Η αμοιβή είναι η λύση του παιγνίου με διαιτησία.

Για ένα δεδομένο παίγνιο, μπορεί να υπάρχουν διάφορα σχήματα διαιτησίας. Συνεπώς, η διαδικασία επιλογής μίας λύσης θα μπορούσε δυνητικά να είναι αυθαίρετη. Όμως, ο Nash, στην εργασία [103], παρουσίασε ένα σχήμα διαιτησίας που με μοναδικό τρόπο ικανοποιεί προκαθορισμένα αξιώματα δικαιοσύνης, αναγκαίες συνθήκες που κρίνονται λογικές με την ευρεία. Ο αξιωματικός ορισμός της λύσης διαπραγμάτευσης του Nash (Nash bargaining solution – NBS) περιγράφεται παρακάτω.

Ορισμός 3.3: Μία απεικόνιση (mapping) $F: G \rightarrow \mathfrak{R}^N$, όπου G είναι το σύνολο των επιτεύξιμων ωφελειών, σε σχέση με το αρχικό σημείο διαφωνίας \mathbf{u}^0 (status quo), είναι ένα NBS, αν ισχύουν τα ακόλουθα:

- 1) $F(U, \mathbf{u}^0) \in U_0$, όπου U_0 είναι το σύνολο των επιτεύξιμων ωφελειών που είναι ανώτερες από το status quo.
- 2) Το $F(U, \mathbf{u}^0)$ είναι κατά Pareto αποτελεσματικό.
- 3) Το F ικανοποιεί το αξίωμα της γραμμικότητας: αν $\boldsymbol{\varphi}: \mathfrak{R}^N \rightarrow \mathfrak{R}^N$, $\boldsymbol{\varphi}(\mathbf{u}) = \mathbf{u}'$ με $u'_j = a_j u_j + b_j$, $a_j > 0$, $j = 1, \dots, N$, τότε $F(\boldsymbol{\varphi}(\mathbf{u}), \boldsymbol{\varphi}(\mathbf{u}^0)) = \boldsymbol{\varphi}(F(\mathbf{u}, \mathbf{u}^0))$.
- 4) Το F ικανοποιεί το αξίωμα των αδιάφορων εναλλακτικών (irrelevant alternatives): αν $V \subset U$, $(V, \mathbf{u}^0) \in G$ και $F(U, \mathbf{u}^0) \in V$, τότε $F(U, \mathbf{u}^0) = F(V, \mathbf{u}^0)$.
- 5) Το F ικανοποιεί το αξίωμα της συμμετρίας: αν $\mathbf{u} \in U$, $u_i^0 = u_j^0$, κι αν $(u_1, \dots, u_i, \dots, u_j, \dots, u_N) \in U$, τότε $(u_1, \dots, u_j, \dots, u_i, \dots, u_N) \in U$, τότε $F(U, \mathbf{u}^0)_i = F(U, \mathbf{u}^0)_j$ για $i, j \in \{1, \dots, N\}$.

Το status quo είναι το διάνυσμα $\mathbf{u}^0 = (u_1^0, \dots, u_N^0)$, κι αναπαριστά την ωφέλεια που επιτυγχάνει ο κάθε παίκτης ελλείψει συνεργασίας (όταν δεν μπορεί να επιτευχθεί συμφωνία), π.χ. σε μία Nash ισορροπία. Το status quo, που στην ουσία αποτελεί

το εναρκτήριο σημείο στο χώρο των ωφελειών που συμφωνείται από όλους τους παίκτες, είναι συνήθως μη αποτελεσματικό κατά Pareto. Αναλυτικότερα, τα πέντε αξιώματα του NBS περιγράφονται παρακάτω.

Το αξίωμα 1 εξασφαλίζει ότι η λύση θα είναι ανώτερη από το status quo του παιγνίου, ώστε οι παίκτες να έχουν όφελος από τη συνεργασία.

Το αξίωμα 2 εξασφαλίζει ότι η λύση σε κάθε πρόβλημα διαπραγμάτευσης θα πρέπει να είναι κατά Pareto αποτελεσματική. Δηλαδή δεν θα πρέπει να υπάρχει άλλος δυνατός συνδυασμός ωφελειών που να είναι καλύτερος από τη λύση για έναν παίκτη κι όχι χειρότερος από την λύση για κάποιον άλλον παίκτη.

Το αξίωμα 3 καθορίζει ότι εάν ένα πρόβλημα διαπραγμάτευσης μπορεί να παραχθεί από ένα άλλο πρόβλημα διαπραγμάτευσης με συσχετιζόμενο (affine) μετασχηματισμό των συναρτήσεων ωφέλειας (ο μετασχηματισμός αυτός δεν θα επηρεάσει τα παιγνιοθεωρητικά χαρακτηριστικά των συναρτήσεων), η λύση του παιγνίου μπορεί να παραχθεί από τη λύση του αρχικού παιγνίου με τον ίδιο μετασχηματισμό. Δηλαδή, εάν μεταβληθεί ο τρόπος μέτρησης της ωφέλειας κατά την διαμόρφωση ενός παιχνιδιού διαπραγμάτευσης, αλλά οι νέες κλίμακες ωφέλειας είναι παιγνιοθεωρητικά ισοδύναμες με τις παλιές, η διαπραγματευτική λύση στον χώρο των αναθέσεων ωφέλειας θα πρέπει να μεταβληθεί τελείως ανάλογα.

Το αξίωμα 4 καθορίζει ότι αν συρρικνωθεί το πεδίο των ωφελειών, αλλά διατηρηθεί η λύση της διαπραγμάτευσης, το σημείο της διαπραγμάτευσης παραμένει το ίδιο. Αυτό σημαίνει ότι, αν οι παίκτες, συμφωνήσουν σε ένα σημείο s στο $V \subset U$, όταν θα μπορούσαν να είχαν επιλέξει κάθε άλλο σημείο στο U , όλα τα αποτελέσματα στο U εκτός του s μπορούν να απορριφθούν ως «αδιάφορα». Με άλλα λόγια, η λύση δεν πρέπει να εξαρτάται από «αδιάφορες» εναλλακτικές λύσεις, αλλά όταν οι παίκτες περιορίζονται σε ένα μικρότερο πεδίο που περιέχει το σημείο της διαπραγμάτευσης, πρέπει και πάλι να επιλέξουν την ίδια λύση. Το αξίωμα των ανεξάρτητων εναλλακτικών υπονοεί ότι το σημείο της λύσης διαπραγμάτευσης δεν επηρεάζεται επεκτείνοντας το πεδίο των ωφελειών, αν μπορεί να επιτευχθεί συμφωνία σε ένα περιορισμένο πεδίο.

Το αξίωμα 5 καθορίζει ότι το σημείο της λύσης διαπραγμάτευσης δεν έχει σχέση με τις ταυτότητες που μπορεί να έχουν ανατεθεί στους παίκτες, δηλαδή παίκτες με το ίδιο επίπεδο ωφελείας στο status quo και την ίδια συνάρτηση ωφελείας θα επιτύχουν το ίδιο αποτέλεσμα.

Η λύση στο πρόβλημα της διαπραγμάτευσης του Nash επιτυγχάνεται στο σημείο που αποτελεί το μέγιστο (σε ένα κυρτό και συμπαγές σύνολο σημείων) του γινόμενου των πλεοναζόντων ωφελειών των παικτών (surplus utility) πέραν του status quo.

$$\max_s \prod_{i=1}^N (u_i(\mathbf{s}) - u_i^0), \quad \mathbf{s} \in X, \quad X = \{\mathbf{r} \in S : u_i(\mathbf{r}) > u_i^0\} \quad (3.3)$$

Παρόλο που το σχήμα διαπραγμάτευσης του Nash αναπτύχθηκε στην προσπάθεια να επιτύχει ένα ικανοποιητικό αποτέλεσμα σε ένα μη συνεργατικό παίγνιο, ενίοτε δέχεται κριτική για το ότι δεν είναι δίκαιο υπό μερικές απόψεις. Σε όρους διαπραγμάτευσης, ισχυροί παίκτες με μεγαλύτερη διαπραγματευτική ισχύ μπορούν να επιτύχουν καλύτερα αποτελέσματα. Επί παραδείγματι, ένας εκατομμυριούχος έχει λιγότερα να ρισκάρει σε σχέση με έναν άπορο όταν διαπραγματεύεται \$1000. Συνεπώς, ο εκατομμυριούχος έχει την οικονομική ισχύ να θέσει πιο σοβαρές απειλές και απαιτήσεις, δεδομένου ότι μία ανεπιτυχής διαπραγμάτευση θα σημαίνει σχετικά μικρές απώλειες. Ο άπορος μπορεί να είναι πρόθυμος να δεχθεί κάποιες απώλειες για ένα σχετικά μικρό, εγγυημένο κέρδος. Ένας δίκαιος διαμοιρασμός των χρημάτων μπορεί να μη συνεπάγεται ότι ο κάθε παίκτης θα πάρει από \$500. Σε ένα σενάριο διαπραγμάτευσης, ένας ηθικά δίκαιος διαμοιρασμός του ποσού συνεπάγεται δίκαιο διαμοιρασμό των ωφελειών, όχι των χρημάτων ή των αγαθών [37]. Αυτό σημαίνει ότι σημαντική είναι η μορφή κι όχι η απόλυτη τιμή της συνάρτησης ωφελείας, και το σχήμα διαπραγμάτευσης του Nash, στη βάση του, ακολουθεί αυτή τη λογική.

3.6 Δημοπρασίες

Οι δημοπρασίες αποτελούν μία ειδική κατηγορία παιγνίων. Μία δημοπρασία είναι η θέσπιση μίας αγοράς με ένα σύνολο σαφών κανόνων που καθορίζουν την

εκχώρηση πόρων και τις τιμές, με βάση τις προσφορές που υποβάλλουν οι συμμετέχοντες στην αγορά [15]. Σε μία δημοπρασία, οι παίκτες υποβάλλουν προσφορές για ένα αγαθό κι ο μηχανισμός (που είναι γνωστός από όλους τους παίκτες) υπολογίζει το αποτέλεσμα, δηλαδή εκχωρήσεις των αγαθών κι αντίστοιχες χρεώσεις.

Ο ορισμός μίας δημοπρασίας περιλαμβάνει δύο βασικά βήματα. Το πρώτο είναι ο ορισμός του *κανόνα εκχώρησης πόρων* (allocation rule) και το δεύτερο ο ορισμός του *κανόνα χρέωσης* (pricing rule). Ο κανόνας εκχώρησης πόρων καθορίζει το νικητή (ή νικητές) της δημοπρασίας. Σε μία δημοπρασία για ένα «αδιαίρετο» (indivisible) αγαθό, ο νικητής είναι ο παίκτης που έχει υποβάλλει την υψηλότερη προσφορά. Υπάρχουν, όμως, και πιο πολύπλοκες δημοπρασίες, όπως στην περίπτωση πολλών διαφορετικών αγαθών (multi-commodity goods). Σε αυτού του τύπου τις δημοπρασίες που ονομάζονται συνδυαστικές (combinatorial auctions), μία προσφορά ορίζεται για ένα υποσύνολο των αγαθών κι η τελική εκχώρηση προκύπτει ως η λύση ενός συνδυαστικού προβλήματος. Επίσης υπάρχουν δημοπρασίες για απείρως διαιρέσιμα αγαθά, όπου ο παίκτης μπορεί να αποκτήσει ένα ποσοστό του αγαθού.

Αφού οριστεί ο κανόνας εκχώρησης, μένει να οριστεί ο κανόνας χρέωσης, δηλαδή το πόσο θα πληρώσει ένας παίκτης για τους πόρους που αποκτά. Υπάρχουν διάφοροι κανόνες χρέωσης στις δημοπρασίες: π.χ. κανόνας πρώτης τιμής (first price), όπου το ποσό που πληρώνει ο παίκτης είναι ίσο με την προσφορά που έκανε. Αυτός ο κανόνας συνηθίζεται στις παραδοσιακές «ανοδικές δημοπρασίες» (ascending auctions), όπου οι παίκτες αυξάνουν τις προσφορές τους μέχρις ότου να μείνει μόνον ένας παίκτης (ο νικητής), ή στις δημοπρασίες Ολλανδικού τύπου (Dutch auctions)¹, γνωστές κι ως «καθοδικές δημοπρασίες» (descending price auction), όπου η τιμή του αγαθού μειώνεται σταδιακά από τον πλειστηριαστή κι ο πρώτος παίκτης, που δηλώνει ενδιαφέρον για το αγαθό, το αποκτά. Ένα άλλο παράδειγμα κανόνα χρέωσης αποτελεί ο κανόνας της δεύτερης τιμής (second price), ή αλλιώς ο κανόνας Vickrey-Clarke-Groves (VCG) [16], [17], [18], όπου το ποσό που πληρώνει ο παίκτης είναι όσο οι προσφορές των παικτών που

¹ Το όνομα των δημοπρασιών Ολλανδικού τύπου προέκυψε από τις περίφημες δημοπρασίες που έλαβαν χώρα τον 17^ο αιώνα για τους ολλανδικούς βολβούς τουλίπας.

αποκλείστηκαν από το παίγνιο, εξ' αιτίας του. Ονομάζονται δημοπρασίες δεύτερης τιμής, καθώς στην περίπτωση ενός μοναδικού αδιαίρετου αγαθού, η τιμή που πληρώνει ο νικητής είναι ίση με τη δεύτερη υψηλότερη προσφορά. Σε ό,τι αφορά τη διαδικασία υποβολής των προσφορών, κάποια σχήματα χρειάζονται κάποιο χρόνο μέχρι να ολοκληρωθεί η δημοπρασία (όπως στις ανοδικές ή καθοδικές δημοπρασίες). Υπάρχουν, όμως, και δημοπρασίες ενός γύρου (one-shot), όπου όλοι οι παίκτες υποβάλλουν την προσφορά τους άπαξ και ταυτόχρονα.

Ένας καλά ορισμένος μηχανισμός δημοπρασιών πρέπει να χαρακτηρίζεται από διάφορες επιθυμητές ιδιότητες. Κατά πρώτον, είναι επιθυμητή η ύπαρξη μοναδικής κατά Nash ισορροπίας, ώστε να μπορεί να υπάρχει «έλεγχος» στο αποτέλεσμα της δημοπρασίας. Άλλες επιθυμητές ιδιότητες είναι οι εξής:

- 1) *συμβατότητα κινήτρων* (incentive compatibility) – το συμφέρον κάθε παίκτη να είναι να υποβάλλει ειλικρινείς προσφορές, δηλαδή να αποκαλύπτει το πόσο αξίζει γι' αυτόν (valuation) στην πραγματικότητα το αγαθό,
- 2) *ατομικός ορθολογισμός* (individual rationality) – το συμφέρον κάθε παίκτη να είναι να συμμετάσχει στο παίγνιο (η προκύπτουσα ωφέλεια να είναι πάντα θετική),
- 3) *αποτελεσματικότητα* (efficiency) – η αντικειμενική συνάρτηση στην τελική έκβαση του παιγνίου (στη Nash ισορροπία, αν υπάρχει) να λαμβάνει τη μέγιστη τιμή της.

Η αντικειμενική συνάρτηση μπορεί να είναι, συνήθως, το κέρδος του πλειστηριαστή, δηλαδή το άθροισμα των ωφελειών όλων των παικτών. Ένα ελκυστικό χαρακτηριστικό των δημοπρασιών δεύτερης τιμής, που δικαιολογεί την ευρεία χρήση τους, είναι ότι αποτελούν το μόνο είδος δημοπρασίας που έχει και την ιδιότητα 1) και την ιδιότητα 3) [19].

3.7 Εφαρμογές της Θεωρίας Παιγνίων σε Δικτυακά Προβλήματα

Σε αυτό το υποκεφάλαιο, παρατίθενται σύντομες περιγραφές εργασιών στις οποίες έχουν εφαρμοστεί λύσεις βασισμένες στη Θεωρία Παιγνίων για προβλήματα σχετικά με διαχείριση πόρων σε δίκτυα.

3.7.1 Εκχώρηση εύρους ζώνης σε Εικονικά Μονοπάτια (Virtual Paths) για δίκτυα πολλαπλών χρηστών

Μία σημαντική εργασία στον χώρο της μελέτης δικτυακών προβλημάτων με εφαρμογή παιγνιοθεωρητικών αρχών είναι η [121]. Η εργασία αυτή ασχολείται με το πρόβλημα του διαμοιρασμού του διαθέσιμου εύρους ζώνης ενός εικονικού μονοπατιού σε δίκτυο ATM (asynchronous transfer mode) μεταξύ διαφόρων χρηστών. Σύμφωνα με την παραδοσιακή τακτική, το παραπάνω πρόβλημα αποτελεί αντικείμενο μίας και μοναδικής διαχειριστικής οντότητας, πολιτική που έχει πλέον κριθεί αναποτελεσματική για μεγάλα ευρυζωνικά δίκτυα. Στην εν λόγω εργασία, οι χρήστες δεν είναι παθητικοί, αλλά, κατανεμημένα, είναι σε θέση να λαμβάνουν αποφάσεις σχετικές με τον διαμοιρασμό των πόρων.

Το σύστημα έχει μοντελοποιηθεί ως ένα σύνολο N χρηστών που μοιράζονται το δικτυακό πόρο συνολικού εύρους ζώνης B . Ο κάθε χρήστης δεσμεύει τμήμα του εύρους ζώνης για να εγκαταστήσει ένα εικονικό μονοπάτι για τις εισερχόμενες κλήσεις του, με σκοπό την ελαχιστοποίηση της πιθανότητας μπλοκαρίσματος (blocking probability). Αυτή η πιθανότητα είναι φθίνουσα συνάρτηση της χωρητικότητας που έχει δεσμευθεί στα εικονικά μονοπάτια του χρήστη. Η δέσμευση χωρητικότητας, όμως, γίνεται σταδιακά πιο δύσκολη (κι ακριβότερη) όσο περιορίζεται η διαθεσιμότητα των πόρων του συστήματος. Οι παραπάνω, αντικρουόμενες λογικές μπορούν να ποσοτικοποιηθούν, όπως περιγράφεται παρακάτω.

Έστω C_i η στρατηγική του παίκτη i δηλαδή το ποσό των πόρων (εύρους ζώνης) που δεσμεύει. Μπορεί να θεωρηθεί, χωρίς απώλεια της γενικότητας, ότι ένας χρήστης λαμβάνει πάντα την ποσότητα των πόρων την οποία αιτείται. Ο χώρος στρατηγικών (strategy space) του χρήστη i είναι το διάστημα $[0, B]$. Έστω, λοιπόν, $\mathbf{C} = (C_1, C_2, \dots, C_N)$ το διάνυσμα στρατηγικών (game strategy vector). Έστω, επίσης, C το συνολικό ποσό δεσμευμένης χωρητικότητας, δηλαδή $C = \sum_i C_i$. Η συνάρτηση κόστους για τον χρήστη i είναι, τότε, η J_i η μορφή της οποίας είναι η ακόλουθη:

$$J_i = J_i(C_i, C) = F_i(C_i, C) + G_i(C_i)$$

Η συνάρτηση F_i υποδηλώνει την διαθεσιμότητα των πόρων όπως την αντιλαμβάνεται ο χρήστης i . Η συνάρτηση G_i υποδηλώνει την επίδραση που έχει στις επιδόσεις του συγκεκριμένου χρήστη η δεσμευμένη χωρητικότητα C_i . Οι παραπάνω συναρτήσεις θεωρείται ότι έχουν τα εξής χαρακτηριστικά:

F1. Η $F_i(\cdot, \cdot)$ αυξάνεται μονότονα ως προς και τα δύο της ορίσματα.

F2. Η $F_i(C_i, C)$ είναι διαρκώς παραγωγίσιμη ως προς το C_i .

F3. Η $F_i(C_i, C)$ είναι αυστηρά κυρτή (strictly convex) ως προς το C_i .

F4. Η μερική παράγωγος $\partial F_i(C_i, C)/\partial C_i$ είναι μη-φθίνουσα ως προς το C κι επίσης αυστηρά αύξουσα ως προς το C_i .

F5. Το όριο $\lim_{C_i \rightarrow B} F_i(C_i, C) = \infty$.

G1. Η $G_i(C_i)$ είναι διαρκώς παραγωγίσιμη.

G2. Η $G_i(C_i)$ είναι αυστηρά φθίνουσα.

G3. Η $G_i(C_i)$ είναι κυρτή.

G4. Το όριο $\lim_{C_i \rightarrow 0} G_i(C_i) = \infty$.

Σύμφωνα με τις παραπάνω ιδιότητες, η συνάρτηση F_i αυξάνεται με τις αιτήσεις δέσμευσης εύρους ζώνης, από τον χρήστη i , καθώς και με το συνολικό όγκο δεσμεύσεων. Η ιδιότητα F5 επιβάλλει ότι ο κάθε χρήστης δεν μπορεί να αναλώσει μόνος του όλο το εύρος ζώνης. Η συνάρτηση G_i μειώνεται με το C_i δεικνύοντας έτσι τη βελτίωση στην επίδοση του χρήστη i . Η G4 δηλώνει ότι ο κάθε χρήστης έχει ανάγκη από κάποιο όγκο δεσμεύσεων.

Η ισορροπία κατά Nash του εν λόγω παιγνίου είναι ένα διάνυσμα στρατηγικών από το οποίο κανείς παίκτης δεν ενδιαφέρεται να αποκλίνει. Με δεδομένες τις παραπάνω ιδιότητες μπορεί να εξασφαλιστεί η ύπαρξη NE για το συγκεκριμένο πρόβλημα. Επίσης, αποδεικνύεται η ύπαρξη μοναδικού NE. Επιπλέον, αποδεικνύεται ότι η κατάσταση ισορροπίας είναι «δίκαια» για τους εμπλεκόμενους παίκτες/χρήστες, με την έννοια ότι όποιος χρήστης έχει περισσότερη ανάγκη από τον διαμοιραζόμενο πόρο, θα λάβει περισσότερο εύρος ζώνης στην κατάσταση ισορροπίας.

Στην ίδια εργασία μελετάται κι η δυναμική συμπεριφορά του παραπάνω μηχανισμού. Αποδεικνύεται ότι μπορεί να επιτευχθεί σύγκλιση στην κατάσταση ισορροπίας Nash με τη χρήση των επαναληπτικών σχημάτων Gauss-Seidel και

Jacobi. Μέσω των σχημάτων αυτών, οι χρήστες επανυπολογίζουν την χωρητικότητα την οποία έχουν δεσμεύσει, βελτιστοποιώντας τις συναρτήσεις κόστους τους, σε απάντηση της εκάστοτε κατάστασης του συστήματος. Μέσω των πειραμάτων τα οποία παρουσιάζονται στην εργασία [121] μελετάται η σχέση της σύγκλισης του προτεινόμενου σχήματος με τη συνάρτηση κόστους. Η μελετούμενη μορφή της συνάρτησης κόστους είναι η ακόλουθη:

$$F_i(C_i, C) = C_i \left(a_1 + \frac{a_2}{\left(1 - \frac{C}{B}\right)^n} \right)$$

Οι παράγοντες a_1 , a_2 και n είναι θετικοί πραγματικοί αριθμοί. Ο παράγοντας a_1 υποδηλώνει ένα σταθερό κόστος ανά μονάδα χωρητικότητας. Ο παράγοντας a_2 υποδηλώνει ένα μοναδιαίο κόστος οφειλόμενο σε παράγοντες όπως η δικτυακή συμφόρηση (η συνολική δεσμευμένη χωρητικότητα C τείνει προς την χωρητικότητα του καναλιού B). Η παράμετρος n υποδηλώνει πόσο νωρίς ανιχνεύεται η κατάσταση συμφόρησης.

Για την συνάρτηση G_i χρησιμοποιήθηκε η συνάρτηση απώλειας Erlang-B. Η διαδικασία άφιξης/γέννησης κλήσεων του χρήστη i θεωρείται Poisson με ρυθμό α_i . Η συνάρτηση Erlang-B που αντιστοιχεί στον χρήστη i είναι συνάρτηση του α_i και C_i και συμβολίζεται με $E(\alpha_i, C_i)$. Εάν το κ_i είναι το πάνω όριο στην πιθανότητα μπλοκαρίσματος κλήσης του χρήστη i , η G_i ορίζεται ως εξής:

$$G_i(C_i) = \begin{cases} \frac{1}{\kappa_i - E(\alpha_i, C_i)}, & \kappa_i > E(\alpha_i, C_i) \\ \infty, & otherwise \end{cases}$$

3.7.2 Παιγιοθεωρητική ανάλυση των αρχών ελέγχου συμφόρησης

Μία ιδιαίτερα σημαντική εργασία στο χώρο της παιγιοθεωρητικής μελέτης δικτυακών προβλημάτων είναι η [9]. Σύμφωνα με την εργασία αυτή, οι χρήστες του δικτύου είναι «εγωιστικές» οντότητες, οι οποίες προωθούν τα προσωπικά τους συμφέροντα, ενώ ο ρόλος του σχεδιαστή του δικτύου περιορίζεται στον καθορισμό της συμπεριφοράς των στοιχείων μεταγωγής του δικτύου (network switches). Ο στόχος της συγκεκριμένης μελέτης είναι ο αποτελεσματικός σχεδιασμός της

λογικής λειτουργίας, η οποία θα πρέπει να επιβληθεί στα στοιχεία του δικτύου, ώστε να επιτευχθούν αποδεκτές επιδόσεις, παρά την εγωιστική συμπεριφορά των τελικών χρηστών.

Η παιγνιοθεωρητική προσέγγιση εφαρμόζεται σε ένα απλό σύστημα ενός μεταγωγέα οι πόροι του οποίου μοιράζονται σε N χρήστες. Ο κάθε χρήστης στέλνει κίνηση Poisson προς τον μεταγωγέα με ρυθμό r_i . Ο αλγόριθμος μέσω του οποίου ρυθμίζεται η παράμετρος r_i καλείται έλεγχος ροής (flow control). Ο μεταγωγέας συμπεριφέρεται ως εκθετικός διακομιστής με “pre-emption”. Ένα μέτρο της συμφόρησης, την οποία υφίσταται κάποιος χρήστης, είναι ο μέσος αριθμός των πακέτων του συγκεκριμένου χρήστη τα οποία βρίσκονται στην ουρά του διακομιστή. Αυτό το μέτρο συμφόρησης (μέσο μέγεθος ουράς) συμβολίζεται με c_i κι εξαρτάται από τους ρυθμούς εισόδου στον μεταγωγέα (r_j), καθώς και την πολιτική που αυτός εφαρμόζει. Η ανάλυση βασίζεται στις εξής αρχές:

- 1) Η ικανοποίηση του χρήστη είναι συνάρτηση του ποσού και της ποιότητας υπηρεσίας την οποία παρέχει ο μεταγωγέας. Αυτό ποσοτικοποιείται φορμαλιστικά μέσω συναρτήσεων ωφέλειας $U_i(r_i, c_i)$ οι οποίες εκφράζουν την προτίμηση του χρήστη για συγκεκριμένα επίπεδα παροχής της υπηρεσίας.
- 2) Ο κάθε χρήστης εφαρμόζει, εγωιστικά, ένα αλγόριθμο ελέγχου ροής για να μεγιστοποιήσει την ωφέλεια του, μεταβάλλοντας το ρυθμό r_i . Όταν ο κάθε χρήστης συμπεριφέρεται με αυτόν τον τρόπο, το λειτουργικό σημείο του συστήματος αποτελεί ένα σημείο ισορροπίας κατά Nash.
- 3) Η επίδοση του συστήματος εκτιμάται μόνο βάσει του επιπέδου ευχαρίστησης που παρέχει στον χρήστη.
- 4) Ο αλγόριθμος βάσει του οποίου λειτουργεί ο μεταγωγέας τελεί υπό κεντρικό, διαχειριστικό έλεγχο.

Η σχεδιαστική προσπάθεια εστιάζεται στον καθορισμό πολιτικών λειτουργίας του μεταγωγέα, ώστε το σύστημα να εμφανίζει καλές επιδόσεις, παρά την εγωιστική συμπεριφορά των χρηστών. Η «καλή συμπεριφορά» συνίσταται στο να είναι η Nash ισορροπία δίκαια κι αποτελεσματική. Επίσης, η Nash ισορροπία θα πρέπει να είναι γρήγορα κι εύκολα προσβάσιμη, μέσω απλών τεχνικών κατανεμημένης

βελτιστοποίησης. Τέλος, το σύστημα θα πρέπει να προσφέρει ορισμένες εγγυήσεις επιδόσεων ακόμη κι αν λειτουργεί εκτός ισορροπίας.

Πέραν από την αναφορά στις βασικές αρχές που διέπουν τις Nash ισορροπίες, στην [9] μελετάται ο μηχανισμός λειτουργίας των μεταγωγέων που είναι γνωστός ως *proportional allocation* (αναλογική εκχώρηση) καθώς κι ο *fair share* (δίκαιου μεριδίου). Στον πρώτο μηχανισμό, το μέσο μήκος της ουράς ενός χρήστη είναι ανάλογο του ρυθμού αυτού του χρήστη. Η λογική αυτή απεικονίζεται στην πολιτική *first-in-first-out* (FIFO) των μεταγωγέων. Στον *fair share* μηχανισμό η λογική είναι αυτή ενός συστήματος χρονοδρομολόγησης με προτεραιότητες (*preemptive priority scheduling*). Εάν οι ρυθμοί r_i ταξινομηθούν κατά αύξουσα σειρά, ο χρήστης με ρυθμό r_1 έχει την μεγαλύτερη προτεραιότητα, ενώ ο ρυθμός του χρήστη με την αμέσως επόμενη προτεραιότητα είναι $r_2 - r_1$ κλπ.

Με το μηχανισμό *fair share* υπάρχει πάντα μία Nash ισορροπία, η οποία είναι και κατά Pareto αποτελεσματική (Ορισμός 3.2). Επίσης, η εν λόγω ισορροπία είναι “*envy-free*”, καθώς κανένας παίκτης δε «ζηλεύει» την ανάθεση ισορροπίας ενός άλλου, επομένως, και δίκαια. Σε αντιδιαστολή, η λογική *proportional allocation* οδηγεί σε ισορροπίες Nash που δεν είναι κατά Pareto αποτελεσματικές και μπορεί να μην είναι δίκαιες. Για την ίδια πολιτική, η σύγκλιση στην κατάσταση ισορροπίας δεν είναι εγγυημένη.

Όλοι οι συνήθεις αριθμητικοί αλγόριθμοι εντοπισμού μεγίστου, αν εφαρμοστούν στην περίπτωση *fair share* οδηγούν σε Nash ισορροπία. Επίσης, μελετάται μία παραλλαγή της κατάστασης ισορροπίας η οποία είναι γνωστή ως ισορροπία Stackelberg. Στο παίγνιο Stackelberg (αλλιώς γνωστό κι ως παίγνιο ηγέτη-ακολουθού), ο ρόλος κάποιου παίκτη αναβαθμίζεται, αφήνοντας τους υπολοίπους χρήστες να φτάσουν σε ισορροπία κι επιλέγοντας, εκ των προτέρων, την ενδεδειγμένη για αυτόν στρατηγική. Ο χρήστης αυτός χαρακτηρίζεται *ηγέτης* (*leader*), και γίνεται η υπόθεση ότι διαθέτει κάποια πληροφόρηση για τις συναρτήσεις ωφέλειας των υπολοίπων. Οι υπόλοιποι παίκτες ονομάζονται *ακόλουθοι* (*follower*) Στην κατάσταση ισορροπίας Stackelberg, η ωφέλεια του ηγέτη δεν είναι χαμηλότερη από αυτή στην απλή Nash ισορροπία. Η Nash ισορροπία, η οποία επιτυγχάνεται από τον *fair share* αλγόριθμο είναι, επίσης, ισορροπία Stackelberg.

3.7.3 Αρχιτεκτονική μη-συνεργατικών δικτύων

Μία άλλη σημαντική εργασία στον χώρο των παιγνιοθεωρητικών προσεγγίσεων δικτυακών προβλημάτων είναι η [10]. Η εργασία αυτή εστιάζει στο πρόβλημα της δρομολόγησης κι εισάγει δύο μεθοδολογίες για την συγκρότηση μη συνεργατικών δικτύων όπου οι χρήστες λαμβάνουν αποφάσεις με σκοπό τη μεγιστοποίηση της ωφέλειας τους. Οι μεθοδολογίες αναφέρονται στη φάση της διαστασιοποίησης (provisioning) του δικτύου, καθώς και στη φάση της λειτουργίας του.

Κατά τη φάση διαστασιοποίησης, υπολογίζονται οι βασικές λειτουργικές παράμετροι του δικτύου, με σκοπό τη διάθεση της χωρητικότητας των συνδέσεων, ώστε το σύστημα να οδηγηθεί σε μία αποτελεσματική Nash ισορροπία. Τα κριτήρια για την αποτελεσματικότητα του συστήματος είναι η τιμή (οριακό κόστος – marginal cost), την οποία αντιλαμβάνεται ο χρήστης, το συνολικό κόστος για τον κάθε χρήστη ή συνδυασμός των ανωτέρω. Η λύση του προβλήματος εκχώρησης χωρητικότητας δεν συμβαδίζει με τα αναμενόμενα, αφού η επαύξηση της χωρητικότητας των συνδέσεων μπορεί να οδηγήσει σε υποβάθμιση των επιδόσεων των χρηστών. Η επίδειξη αυτής της μη-αναμενόμενης συμπεριφοράς βασίζεται στο παράδοξο του Braess.

Κατά την διάρκεια της φάσης λειτουργίας ένας συντονιστής ελέγχει τη δρομολόγηση τμήματος της δικτυακής ροής. Ο συντονιστής γνωρίζει την μη-συνεργατική συμπεριφορά των χρηστών του δικτύου και λαμβάνει αποφάσεις δρομολόγησης, βάσει αυτής της συμπεριφοράς, προσπαθώντας πάντα να βελτιώσει την επίδοση του συστήματος. Το σενάριο αυτό είναι ακριβώς το σενάριο ενός παιγνίου Stackelberg, όπως αυτό συζητήθηκε στην παράγραφο 3.7.2. Ο συντονιστής (ηγέτης) μπορεί να επιβάλλει μία κατάσταση ισορροπίας, η οποία ταυτίζεται με την βέλτιστη κατάσταση του δικτύου.

3.7.4 Έλεγχος Ισχύος σε CDMA δίκτυα

Μία σημαντική εφαρμογή της Θεωρίας Παιγνίων στις επικοινωνίες αποτελεί η μελέτη του προβλήματος ελέγχου ισχύος σε CDMA δίκτυα. Σε αντίθεση με άλλες τεχνολογίες πρόσβασης στο ασύρματο μέσο, όπως, για παράδειγμα, το TDMA ή το FDMA, στο CDMA, οι χρήστες μπορούν να εκπέμπουν ταυτόχρονα στην ίδια συχνότητα, και τα σήματα τους διαχωρίζονται βάσει μοναδικών κωδίκων που τους

ανατίθενται. Με τους κωδικούς αυτούς, είναι δυνατόν στο δέκτη, τα σήματα των υπολοίπων χρηστών να συμπιεστούν αρκετά, ώστε να αναδειχθεί και να αποκωδικοποιηθεί το σήμα ενός συγκεκριμένου χρήστη.

Η επιτυχία της αποκωδικοποίησης, εξαρτάται από δύο παράγοντες: 1) την ισχύ με την οποία φθάνει το σήμα του χρήστη στο δέκτη και 2) την ισχύ των υπολοίπων χρηστών (παρεμβολή). Κι οι δύο αυτοί παράγοντες συνδυάζονται στο λόγο σήματος προς παρεμβολή (signal to interference ratio – SIR), ο οποίος καθορίζει και την απόδοση του συστήματος. Σε ένα CDMA δίκτυο, λοιπόν, κάθε χρήστης με την ισχύ του επηρεάζει (θετικά) την απόδοση του, αλλά επηρεάζει (αρνητικά) την απόδοση των άλλων χρηστών, καθώς το σήμα του εκλαμβάνεται ως παρεμβολή για τους υπόλοιπους, κι αντίστροφα. Επομένως, το όλο σύστημα μπορεί να παρομοιαστεί με ένα πρόβλημα κοινωνικής αλληλεπίδρασης και, έτσι, να αναλυθεί με τη χρήση της Θεωρίας Παιγνίων.

Στη βιβλιογραφία έχουν προταθεί διάφορα παιγνιοθερητικά μοντέλα για τον έλεγχο ισχύος σε CDMA δίκτυα, θεωρώντας, συνήθως μη συνεργατικούς (εγωιστές) χρήστες (παίκτες). Τα πιο γνωστά μοντέλα παιγνιοθερητικού ελέγχου ισχύος για κυψελωτά δίκτυα είναι αυτό της εργασίας [84] και της εργασίας [100].

3.7.4.1 Το Πλαίσιο Ελέγχου Ισχύος του Yates

Στην εργασία [84], μελετάται το πρόβλημα ελέγχου ισχύος στην *ανερχόμενη ζεύξη* (uplink). Κάθε κόμβος, j , προσπαθεί να επιτύχει ένα επιθυμητό SIR, γ_j , με την ελάχιστη δυνατή κατανάλωση ισχύος, p_j , στο δέκτη (ή δέκτες) που τον ενδιαφέρει, $\{v_j\}$ (ένας ή περισσότεροι ΣΒ). Κάθε κόμβος είναι σε θέση να παρατηρεί το SIR στους δέκτες που ανήκουν στο σύνολο $\{v_j\}$ (ή εναλλακτικά, παρατηρεί τη συνολική λαμβανόμενη ισχύ στα στοιχεία του $\{v_j\}$ και γνωρίζει τις δικές του απώλειες σήματος, $h_{j,\{v_j\}}$). Βάσει αυτών των παρατηρήσεων, οι κόμβοι υπολογίζουν μία συνάρτηση παρεμβολών $I(\mathbf{p})$ που ορίζεται ως ο λόγος του επιδιωκόμενου SIR προς το ουσιαστικό SIR που επιτυγχάνεται, όπου το \mathbf{p} είναι ένα διάνυσμα ισχύων εκπομπής.

Οι ιδιότητες της συνάρτησης $I(\mathbf{p})$, που είναι βασικές για τα αποτελέσματα του μοντέλου, είναι οι ακόλουθες:

- Θετικότητα: $I(\mathbf{p}) > 0$
- Μονοτονία: Αν $\mathbf{p} \geq \mathbf{p}^*$, τότε $I(\mathbf{p}) \geq I(\mathbf{p}^*)$
- Κλιμάκωση (scalability): Για όλα τα $\alpha > 1$, $\alpha I(\mathbf{p}) > I(\alpha \mathbf{p})$
όπου η έκφραση $\mathbf{p} > \mathbf{p}^*$ σημαίνει ότι $p_j > p_j^*, \forall j \in N$.

Σε γενικές γραμμές, η συνάρτηση $I_j(\mathbf{p})$, όπου $I(\mathbf{p}) = \sum_{j \in N} I_j(\mathbf{p})$, παίρνει την ακόλουθη μορφή (λόγος επιδιωκόμενου SIR προς SIR που επιτυγχάνεται).

$$I_j(\mathbf{p}) = \frac{\gamma_j}{\mu_j(\mathbf{p})}$$

Η ισχύς εκπομπής κάθε κινητού ενημερώνεται στο βήμα $k+1$ βάσει της ακόλουθης επαναληπτικής διαδικασίας:

$$p_j(k+1) = p_j(k) I(\mathbf{p}(k))$$

Κάνοντας την υπόθεση ότι δεν παραβιάζονται οι περιορισμοί χωρητικότητας, το μοντέλο αυτό συγκλίνει σε μία σταθερή κατάσταση, στις ακόλουθες περιπτώσεις:

- Σταθερή ανάθεση, όπου κάθε κινητό ανατίθεται σε έναν συγκεκριμένο ΣΒ ($|\{v_j\}| = 1$)
- Ανάθεση ελάχιστης ισχύος, όπου κάθε κινητό ανατίθεται στο ΣΒ για τον οποίο μεγιστοποιείται το SIR του ($|\{v_j\}| = 1$, αλλά το v_j μεταβάλλεται)
- Πολλαπλή συνδυαστική λήψη (Macro diversity), όπου όλοι οι ΣΒ συνδυάζουν τα σήματα των κινητών ($|\{v_j\}| > 1$)
- Περιορισμένη πολλαπλή συνδυαστική λήψη (Limited diversity), όπου ένα υποσύνολο των ΣΒ συνδυάζουν τα σήματα των κινητών ($|\{v_j\}| > 1$)
- Λήψη μέσω πολλαπλών συνδέσεων, όπου το επιδιωκόμενο SIR πρέπει να διατηρηθεί σε ένα σύνολο ΣΒ ($|\{v_j\}| > 1$).

Στην εργασία [84], αποδεικνύεται ότι η συνάρτηση παρεμβολής έχει τις ακόλουθες ιδιότητες:

- Αν ο αλγόριθμος έχει ένα σταθερό σημείο (fixed point), το σημείο αυτό θα είναι μοναδικό.
- Όταν η $l(\mathbf{p})$ είναι εφικτή, υπάρχει ένα σταθερό σημείο. Η $l(\mathbf{p})$ χαρακτηρίζεται ως εφικτή αν υπάρχει κάποιο $\mathbf{p} \in P$ (χώρος των κατανομών ισχύος), για το οποίο $l(\mathbf{p}) > 1$
- Αν η $l(\mathbf{p})$ είναι εφικτή, τότε ξεκινώντας από οποιοδήποτε σημείο $\mathbf{p}(0)$, εκτός του $\mathbf{p}(0) = \mathbf{0}$, ο αλγόριθμος συγκλίνει στο σταθερό σημείο, όταν οι αποφάσεις λαμβάνονται συγχρονισμένα.

Οι ιδιότητες αυτές ισχύουν και για τις πέντε περιπτώσεις ανάθεσης των κινητών στους ΣΒ που αναφέρθηκαν πιο πάνω (σταθερή ανάθεση, ανάθεση ελάχιστης ισχύος, κτλ).

3.7.4.2 Παίγνια Ελέγχου Ισχύος του Goodman

Ενώ ο Yates αντιμετώπισε τον κατανεμημένο έλεγχο ισχύος ως ένα γενικού τύπου πρόβλημα σταθερού σημείου, ο Goodman θεώρησε τον κατανεμημένο έλεγχο ισχύος, ως ένα κατανεμημένο αλληλεπιδραστικό πρόβλημα βελτιστοποίησης, δηλαδή ως ένα παίγνιο. Η βασική μορφή της αντικειμενικής συνάρτησης που θεωρείται είναι η εξής:

$$u_i(\mathbf{p}) = \frac{R}{p_i} f(\mu'_{i,b}),$$

όπου το R είναι ο ρυθμός των δεδομένων, f η πιθανότητα επιτυχούς εκπομπής ενός bit ως συνάρτηση ενός «τροποποιημένου» SIR, $\mu'_{i,b}$, το οποίο υπολογίζεται ως εξής:

$$\mu'_{j,b} = \frac{W}{R} \frac{h_{j,b} p_j}{\sum_{k \in N \setminus j} h_{k,b} p_k + \sigma_b^2}$$

όπου W είναι το εύρος ζώνης εκπομπής

Στην εργασία [100], η αντικειμενική συνάρτηση $u_i(\mathbf{p})$ ορίζεται έχοντας ως βάση μία συνάρτηση ρυθμοαπόδοσης για πακέτα μήκους L , όπου όλα τα bit στο πακέτο πρέπει να ληφθούν επιτυχώς για να θεωρηθεί επιτυχής η εκπομπή. Συγκεκριμένα, η συνάρτηση $u_i(\mathbf{p})$ ορίζεται ως εξής:

$$u_i(\mathbf{p}) = \frac{ER}{\rho_j} (1 - e^{-0.5\mu'_j})^L$$

όπου το E υποδηλώνει το ενεργειακό περιεχόμενο της μπαταρίας.

Στην εργασία [100] αποδεικνύεται ότι το εν λόγω παίγνιο έχει τουλάχιστον ένα σταθερό σημείο, εξασφαλίζεται ότι είναι μοναδικό κι ότι σε αυτό όλοι οι παίκτες φθάνουν στο ΣΒ με ίση ισχύ. Ωστόσο, αυτό το σταθερό σημείο δεν είναι βέλτιστο (κατά Pareto), και για το λόγο αυτό εισάγεται μία συνάρτηση γραμμικής χρέωσης c_j που φαίνεται παρακάτω, όπου το t είναι μία θετική σταθερά.

$$c_j(\mathbf{p}) = tR\rho_j, t > 0$$

Με την αφαίρεση της συνάρτησης c_j από την αντικειμενική συνάρτηση $u_i(\mathbf{p})$, προκύπτει η επόμενη τροποποιημένη αντικειμενική συνάρτηση.

$$u'_i(\mathbf{p}) = \frac{ER}{\rho_j} (1 - e^{-0.5\mu'_j})^L - tR\rho_j, t > 0$$

Με προσομοιώσεις φάνηκε ότι υπάρχει σταθερό σημείο για τον τροποποιημένο αλγόριθμο, το οποίο μάλιστα είναι πιο αποτελεσματικό κατά Pareto σε σχέση με το αντίστοιχο σημείο του αρχικού αλγορίθμου.

Στην εργασία [104], προτείνεται μία εναλλακτική επιλογή του σημείου λειτουργίας του συστήματος. Συγκεκριμένα, θέτοντας ως απαίτηση όλοι οι χρήστες να έχουν ίδιο SIR (δηλαδή να φθάνουν στο ΣΒ με ίση ισχύ), μεγιστοποιούνται οι αντικειμενικές συναρτήσεις $u_i(\mathbf{p})$, και προκύπτει μία νέα κατανομή ισχύων. Το εν λόγω σημείο λειτουργίας είναι και δίκαιο, αλλά, όπως φαίνεται κι από προσομοιώσεις, πιο αποτελεσματικό. Όπως αποδείχθηκε στα πλαίσια της παρούσας διατριβής, στην εργασία [108], το σημείο αυτό είναι κατά Pareto αποτελεσματικό.

3.7.5 Άλλες εργασίες

Στην εργασία [11] μελετάται η εφαρμογή των αρχών της Θεωρίας Παιγνίων σε θέματα τυχαίας πρόσβασης (random access) κι ελέγχου ισχύος. Μελετάται η

συμπεριφορά «εγωιστικά» συμπεριφερόμενων χρηστών σε ένα απλοποιημένο σύστημα Aloha. Στην περίπτωση του ελέγχου ισχύος, επιτυγχάνεται, μέσω παιγνιοθεωρητικών τεχνικών, ένα βέλτιστο σημείο λειτουργίας του συστήματος, χωρίς να απαιτείται η παρέμβαση ενός εξωτερικού ελεγκτή (external controller). Στην εργασία [12] μελετούνται αλγόριθμοι για τον επιμερισμό του κόστους που προκύπτει από μεταδόσεις πολυεκπομπής (multicast). Εξετάζονται δύο βασικοί μηχανισμοί, το οριακό κόστος κι η λύση Shapley.

Μέρος Α

Διαχείριση Πόρων στον Κινητό Υπολογισμό

ΚΕΦΑΛΑΙΟ 4

ΔΙΑΧΕΙΡΙΣΗ ΘΕΣΗΣ ΚΙΝΗΤΩΝ ΧΡΗΣΤΩΝ ΓΙΑ ΕΞΙΣΟΡΡΟΠΗΣΗ ΦΟΡΤΟΥ ΣΕ ΑΣΥΡΜΑΤΑ ΚΥΨΕΛΩΤΑ ΔΙΚΤΥΑ

4.1 Εισαγωγή

Στο παρόν Κεφάλαιο, παρουσιάζεται η μελέτη που αφορά στην εξισορρόπηση φόρτου σε WLAN για την αποφυγή συμφόρησης σε ώρες αιχμής. Σε περιβάλλοντα WLAN, είναι συχνό φαινόμενο να εμφανίζονται μικρής διάρκειας, τοπικά φαινόμενα συμφόρησης, τα οποία, όμως, είναι ικανά να ενοχλήσουν έντονα το χρήστη. Χρήστες που συνδέονται σε «δημοφιλή» AP (AP που βρίσκονται εγκατεστημένα σε πολυσύχναστες περιοχές) είναι πιθανόν να έχουν προβληματική σύνδεση, λόγω υψηλού φόρτου. Την ίδια στιγμή, όμως, μπορεί να υπάρχουν άλλα AP σε γειτονικές περιοχές που να είναι ελάχιστα (ή καθόλου) κορεσμένα. Σε αυτήν την περίπτωση, θα ήταν επιθυμητό να υπήρχε ένα σύστημα, το οποίο θα είχε τη δυνατότητα να εξισορροπήσει το φόρτο σε όλη την έκταση του WLAN, ούτως ώστε οι χρήστες να επωφελούνται από το συνολικό υπάρχον εύρος ζώνης. Με ένα τέτοιο σύστημα, ο πάροχος του WLAN θα μπορούσε να αξιοποιήσει πλήρως την υποδομή του και να μεγιστοποιήσει τα κέρδη του, ενώ οι χρήστες να απολαμβάνουν βελτιωμένη ποιότητα υπηρεσίας.

Στο πρόβλημα που περιγράφεται παραπάνω, η συνεργασία μεταξύ των χρηστών και του δικτύου μπορεί ωφελήσει αμφότερους. Ένας χρήστης που θα δεχόταν να μετακινηθεί σε μία κοντινή περιοχή, θα λάμβανε υπηρεσίες υψηλότερου επιπέδου· τη ίδια στιγμή, ο πάροχος δε θα χρειαζόταν να προβεί σε εγκατάσταση πλεονάζουσας υποδομής. Για παράδειγμα, έστω ένα τερματικό σε ένα αεροδρόμιο, απ' όπου χιλιάδες χρήστες περνούν καθημερινά, και στο οποίο έχει εγκατασταθεί πληθώρα από AP. Οι χρήστες μπορούν να χρησιμοποιούν το WLAN για να πραγματοποιούν χαμηλού κόστους κλήσεις voice over IP (VoIP), οι οποίες, όμως, χαρακτηρίζονται από ανελαστικότητα – δεν μπορούν να συνεχιστούν με ανεκτή ποιότητα, αν το εύρος ζώνης πέσει κάτω από κάποιο επίπεδο. Οι χρήστες, βέβαια, μπορούν να μετακινηθούν σε λιγότερο κορεσμένες θέσεις, ώστε να ολοκληρώσουν

με επιτυχία την κλήση τους. Αντί όμως να μετακινούνται τυχαία, εντός του τερματικού του αεροδρομίου, προς εύρεση αποθεμάτων εύρους ζώνης, μπορούν να καθοδηγούνται με ακρίβεια σε κατάλληλες θέσεις από το σύστημα πληροφοριών του αεροδρομίου.

Με κίνητρο το παραπάνω σενάριο, στα πλαίσια της παρούσας μελέτης, προτείνεται ένα σύστημα ευφυών χώρων (smart spaces system – SSS – S3), με το όνομα MITOS², το οποίο μπορεί να χρησιμοποιηθεί σε συνδυασμό με άλλα σχήματα διαχείρισης πόρων (π.χ. βέλτιστη επιλογή θέσης των AP κατά την φάση σχεδιασμού του WLAN). Το MITOS εστιάζει σε μία διαχειρίσιμη οντότητα η οποία έχει παραβλεφθεί από τα συστήματα κι αλγορίθμους διαχείρισης πόρων: τη θέση του χρήστη. Συγκεκριμένα, ο σκοπός του MITOS είναι να επηρεάζει τις θέσεις των χρηστών, ούτως ώστε να εξισορροπείται ο φόρτος στο WLAN, κι οι χρήστες να αντιλαμβάνονται βελτιωμένες υπηρεσίες. Το MITOS μπορώντας να αντιληφθεί την εμφάνιση συμφόρησης σε ένα τμήμα του WLAN και γνωρίζοντας τις θέσεις των χρηστών, είναι ικανό να παροτρύνει χρήστες που υφίστανται τις συνέπειες της συμφόρησης να μεταβούν σε άλλες θέσεις (μέσω προτάσεων μετακίνησης, relocation proposal – RP), με υψηλότερα αποθέματα εύρους ζώνης. Το MITOS παράγει επίσης οδηγίες πλοήγησης που αφορούν στις προτεινόμενες μεταβάσεις.

Η αποδοτικότητα του προτεινόμενου συστήματος, εξαρτάται από τη δυνατότητα και θέληση των χρηστών να «υπακούσουν» στα RP του MITOS. Από την οπτική γωνία του χρήστη, η μετακίνηση σε μία γειτονική θέση έχει αξία, αν το κέρδος από τη μετακίνηση ξεπερνά την αντίστοιχη ενόχληση. Ένας χρήστης αναμένεται να ακολουθήσει την προτροπή του MITOS, αν η ποιότητα υπηρεσίας που λαμβάνει δεν είναι αποδεκτή στην τρέχουσα θέση κι η απόσταση μέχρι την προτεινόμενη θέση είναι σχετικά μικρή. Ο πάροχος θα μπορούσε επίσης να χρησιμοποιήσει κι εναλλακτικά κίνητρα, ώστε να πείσει τους χρήστες να κινηθούν: να δίνει κάποιο “bonus” (π.χ. επιπλέον χρόνο σύνδεσης στην περίπτωση προπληρωμένης

² Η λέξη «μίτος» προέρχεται από τον αρχαίο μύθο του Θησέα, το γιο του Βασιλιά Αιγαία (από τον οποίο πήρε το όνομα του το Αιγαίο Πέλαγος) που πήγε στη φωλιά του Μινώταυρου, για να τον σκοτώσει. Προηγούμενες προσπάθειες είχαν αποτύχει, λόγω του λαβυρίνθου στον οποίο ζούσε ο Μινώταυρος. Η Αριάδνη, η φίλη του Θησέα, είχε την έξυπνη ιδέα να δώσει στο Θησέα ένα μεταξωτό νήμα (τον μίτο) που θα μπορούσε να ξετυλίξει, ώστε να βρει την έξοδο από το λαβύρινθο. Καθώς ο σκοπός του συστήματος MITOS είναι να ανακαλύπτει τρόπους, ώστε οι χρήστες με προβληματική συνδεσιμότητα να οδηγηθούν σε άλλες θέσεις με υψηλότερα αποθέματα εύρους ζώνης, ο μίτος της Αριάδνης συμβολίζει το σκοπό αυτό.

παροχής υπηρεσίας, ή έκπτωση μελλοντικά) στους χρήστες που ακολουθούν τις προτροπές.

Στην παρούσα μελέτη, μελετάται κι η επίδραση, στην αποτελεσματικότητα του MITOS, του ποσοστού των χρηστών που ακολουθούν τις προτάσεις του συστήματος. Τα αποτελέσματα δείχνουν ότι το MITOS μπορεί να είναι αποτελεσματικό ακόμα κι αν ένα σχετικά μικρό ποσοστό χρηστών ανταποκρίνεται θετικά στις εν λόγω προτροπές. Μελετήθηκε επίσης κι η περίπτωση στην οποία οι χρήστες είναι ιδιαίτερα «πρόθυμοι» στις προτροπές του συστήματος. Υπό αυτές τις συνθήκες είναι δυνατόν να υποβαθμιστεί η αποτελεσματικότητα του συστήματος, καθώς η συμφόρηση θα μεταφέρεται από θέση σε θέση μετά από κάθε (μαζική) μετακίνηση χρηστών. Για την αποφυγή αυτού του κινδύνου, το MITOS εμπλουτίστηκε με μηχανισμούς βασισμένους από τη Θεωρία Δημοπρασιών [48] και τη Θεωρία Παιγνίων [37].

4.2 Σχετική Βιβλιογραφία και Συνεισφορά

Το πρόβλημα της εξισορρόπησης φόρτου σε WLAN έχει απασχολήσει διάφορους κατασκευαστές δικτυακού ασύρματου εξοπλισμού, τα τελευταία χρόνια. Υπάρχουν AP, καθώς κι ασύρματες κάρτες, με χαρακτηριστικά εξισορρόπησης φόρτου [41]. Τα AP που υποστηρίζουν αυτό το χαρακτηριστικό διατηρούν μία μέτρηση του φόρτου τους κι ευρυσκέπτονται (broadcast) στους χρήστες της τρέχουσας κυψέλης σήματα «φάρους» που περιέχουν αυτήν την πληροφορία. Οι νέοι χρήστες μπορούν να χρησιμοποιήσουν αυτήν την πληροφορία για να συνδεθούν στο λιγότερο φορτωμένο AP. Τέτοιου είδους τεχνικές, όμως, έχουν τοπική εμβέλεια και μπορούν να καταναείμουν τους χρήστες μόνο μεταξύ αλληλοεπικαλυπτόμενων κυψελών.

Μία εναλλακτική λύση για την εξισορρόπηση φόρτου σε WLAN προτείνεται στην εργασία [58]. Στην προσέγγιση αυτή, γίνεται η υπόθεση ότι τα AP έχουν δυνατότητα μετακίνησης, όντας τοποθετημένα πάνω σε ειδικές ράγες κι όχι εγκατεστημένα σε σταθερά σημεία. Προτάθηκαν αλγόριθμοι που αναγνωρίζουν τις κατάλληλες θέσεις και ρυθμίσεις των AP, ανάλογα με τις εκάστοτε συνθήκες φόρτου, κι αποδείχθηκε ότι είναι δυνατόν να επιτευχθεί σημαντική μείωση στο

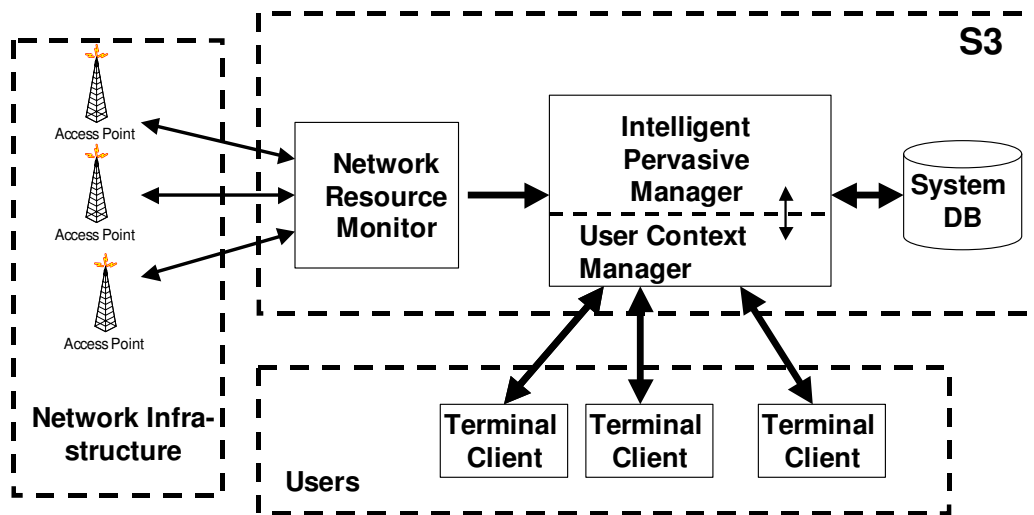
απαιτούμενο πλήθος AP, σε σύγκριση με ένα συμβατικό δίκτυο με στατικά τοποθετημένα AP. Παρόλο που η προαναφερθείσα πρόταση είναι ιδιαίτερα αποτελεσματική, έχει το βασικό μειονέκτημα της απαίτησης εξειδικευμένου υλικού (κινούμενα AP), αλλά κι εγκατάστασης συστημάτων ραγών, που δεν είναι ιδιαίτερα πρακτικό στις περισσότερες περιπτώσεις.

Μία ακόμα προσέγγιση για την αντιμετώπιση της τοπικής συμφόρησης σε WLAN hot-spot, η οποία θεωρείται κι η σχετικότερη με τη μελέτη που παρουσιάζεται στο παρόν κεφάλαιο, περιγράφεται στην εργασία [42]. Στην εν λόγω προσέγγιση, στην περίπτωση τοπικής συμφόρησης, το τερματικό βρίσκει ένα λιγότερο φορτωμένο AP στην γειτονιά του για να συνδεθεί (τεχνική channel switching). Αν δεν υπάρχει AP που να μπορεί να εγγραφεί βελτίωση στη σύνδεση, ένας διακομιστής (network-monitoring server) υποδεικνύει στο χρήστη ένα λιγότερο φορτωμένο, αλλά μακρινό, AP (τεχνική network directed roaming).

Σε αντίθεση με τις παραπάνω εργασίες, η προσέγγιση που προτείνεται στο παρόν κεφάλαιο, δεν απαιτεί εξειδικευμένα χαρακτηριστικά από το δικτυακό εξοπλισμό, αλλά μπορεί να λειτουργήσει πάνω από ένα οποιοδήποτε WLAN. Σε αντίθεση δε με την εργασία [42], γίνεται εκτενής αξιολόγηση της προτεινόμενης τεχνικής σε ό,τι αφορά την αποτελεσματικότητά της, υπό διάφορες συνθήκες και διάφορες συμπεριφορές χρηστών, όπως, για παράδειγμα, όταν η συμφόρηση πλήττει ταυτόχρονα πολλούς χρήστες, ή οι χρήστες δεν είναι ιδιαίτερα συνεργατικοί, απορρίπτοντας ποσοστό των προτάσεων του συστήματος.

4.3 Αρχιτεκτονική και Λειτουργία του Προτεινόμενου Συστήματος

Το MITOS αποτελεί το λειτουργικό συνδυασμό των ακόλουθων τεσσάρων διακριτών συστατικών: 1) *User Context Manager* (UCM), 2) *Intelligent Pervasive Manager* (IPM), 3) *Network Resource Monitor* (NRM) και 4) *Terminal Client* (TC). Τα παραπάνω συστατικά συνδυάζονται, όπως δείχνει η Εικόνα 4.1.



Εικόνα 4.1. Αρχιτεκτονική του Συστήματος MITOS

Ο UCM διαχειρίζεται πληροφορίες χρηστών (ταυτότητα χρήστη, τρέχουσα θέση, AP εξυπηρέτησης, RP και πληροφορίες πλοήγησης), οι οποίες διατηρούνται στη βάση δεδομένων του συστήματος (system database). Ο IPM συνδυάζει πληροφορίες δικτυακού φόρτου με πληροφορίες που προέρχονται από τον UCM (θέση χρήστη, AP εξυπηρέτησης, κτλ), ώστε να αναγνωρίζει τους χρήστες που πλήττονται από συμφόρηση. Εν συνεχεία, αποφασίζει που πρέπει να μεταβεί ο κάθε χρήστης (απ' όσους πλήττονται από τη συμφόρηση), και παρέχει την πληροφορία αυτή στον UCM, οποίος με τη σειρά του προωθεί το RP στον αντίστοιχο TC. Ο NRM συγκεντρώνει δεδομένα από τα AP, ούτως ώστε να υπολογιστούν οι μετρικές φόρτου για την ανίχνευση συμφόρησης. Σε περίπτωση που ανακαλυφθούν κάποια κορεσμένα AP, ο NRM στέλνει μία πλήρη αναφορά στον IPM, ώστε ο τελευταίος να προβεί στις απαραίτητες ενέργειες. Ο NRM βασίζεται στο Simple Network Management Protocol (SNMP). Ο TC εγκαθίσταται στο κινητό τερματικό και υλοποιεί τη διεπαφή μεταξύ του χρήστη και του συστήματος. Συγκεκριμένα, επικοινωνεί με τον UCM για την ανταλλαγή πληροφοριών σχετικών με τη θέση, αλλά και τη σύνδεση. Αντίστροφα, ο UCM, σε περίπτωση συμφόρησης, ειδοποιεί με κατάλληλο μήνυμα (RP) τον TC.

4.3.1 Λειτουργία του Συστήματος

Ο IPM και ο UCM πρέπει να είναι ενήμεροι σε σχέση με πληροφορίες χαμηλού επιπέδου, αλλά και πληροφορικές επιπέδου χρήστη. Αυτές οι πληροφορίες, σε

συνδυασμό με πληροφορίες σχετικά με την τοπολογική δομή του WLAN, διατηρούνται στη βάση δεδομένων του συστήματος. Συγκεκριμένα, οι πληροφορίες που αποθηκεύεται στη βάση δεδομένων του συστήματος μπορεί να διακριθούν σε 1) πληροφορίες χρήστη (π.χ. τρέχουσα θέση, AP εξυπηρέτησης και προτεινόμενη θέση), 2) πληροφορίες δικτυακού φόρτου και 3) πληροφορίες του δικτύου-κτηρίου (π.χ. τοποθεσίες του κτηρίου, τοπολογικές σχέσεις, κάλυψη των AP).

Όταν ο NRM διαπιστώνει την εμφάνιση συμφόρησης, στέλνει μία αναφορά στον IPM. Στη συνέχεια, ο τελευταίος ενημερώνει τη βάση δεδομένων του συστήματος βάσει της τελευταίας ενημέρωσης κι εκτελεί τις απαιτούμενες ενέργειες (όπως περιγράφονται στο υποκεφάλαιο 4.3.4). Ο TC περιοδικά εκπέμπει δεδομένα που αφορούν στην ποιότητα του λαμβανόμενου σήματος, ώστε ο UCM να μπορεί να εκτιμήσει την τρέχουσα θέση της κινητής συσκευής, καθώς και το AP στο οποίο είναι συνδεδεμένη.

Η υλοποίηση του UCM βασίζεται στο σύστημα εντοπισμού θέσης Nibble [24]. Το Nibble είναι ικανό να αναγνωρίζει θέσεις με ακρίβεια δωματίου, χρησιμοποιώντας μετρήσεις της ποιότητας του λαμβανόμενου σήματος στην κινητή συσκευή. Διατηρεί ένα κατάλληλα «εκπαιδευμένο» Bayesian δίκτυο κι επιστρέφει «συμβολική» πληροφορία (π.χ. ο χρήστης βρίσκεται στο «Δωμάτιο I-9»). Εναλλακτικά, το MITOS θα μπορούσε να χρησιμοποιήσει οποιοδήποτε άλλο σύστημα εντοπισμού θέσης για WLAN, όπως το Microsoft RADAR [44], το Ekahau [43], ή ένα απ' όσα αναφέρονται στις εργασίες [25] και [26].

4.3.2 Μοντελοποίηση Θέσης

Για την αναπαράσταση θέσης στο σύστημα, χρησιμοποιήθηκε ένας κατευθυντικός γράφος, ο οποίος από εδώ και στο εξής θα αναφέρεται ως Location Graph (LG) [53]. Οι κόμβοι στον LG αναπαριστούν θέσεις· μία ακμή που συνδέει δύο θέσεις υποδηλώνει ότι οι θέσεις συνορεύουν. Η ακρίβεια θέσης του μοντέλου του LG είναι συμβατή με το σύστημα εντοπισμού θέσης που χρησιμοποιήθηκε. Το Nibble επιστρέφει συμβολική πληροφορία (π.χ. ο χρήστης είναι στο δωμάτιο 'I-4'), αντί για απόλυτες συντεταγμένες (x,y), αποφεύγοντας έτσι την ανάγκη για ένα ενδιάμεσο επίπεδο που θα αντιστοιχούσε τις απόλυτες συντεταγμένες σε συμβολική πληροφορία.

Το «βάρος» που ανατίθεται σε μία ακμή του γράφου αναπαριστά την «απόσταση» μεταξύ δύο θέσεων, όπως μπορεί να είναι η φυσική τους απόσταση, ή γενικότερα, το κόστος (μέτρο δυσκολίας) για τη μετάβαση από τη μία θέση στην άλλη. Με τη χρήση του LG, η εύρεση των συντομότερων μονοπατιών (που χρειάζονται από το σύστημα για την παραγωγή των RP) είναι πολύ απλή. Στην υλοποίηση που έχει γίνει, ο αλγόριθμος Dijkstra [57] τρέχει μία φορά, κατά την εκκίνηση του συστήματος κι υπολογίζει τα συντομότερα μονοπάτια για κάθε συνδυασμό θέσης εκκίνησης και προορισμού.

4.3.3 Διεπαφή Τερματικού Πελάτη του Χρήστη

Όταν το MITOS στέλνει ένα RP σε έναν χρήστη που πλήττεται από συμφόρηση, ο TC είναι υπεύθυνος να ενημερώσει το χρήστη κατάλληλα. Η γραφική διεπαφή χρήστη του TC μπορεί να έχει τη μορφή ενός εικονιδίου τοποθετημένου σε μία γωνία της οθόνης. Το εικονίδιο μπορεί να είναι σε «ενεργή» ή σε «ανενεργή» κατάσταση· στη μεν περίπτωση το εικονίδιο αναβοσβήνει διακριτικά, υποδηλώνοντας ότι το AP στο οποίο είναι συνδεδεμένος ο χρήστης είναι κορεσμένα. Ο χρήστης τότε μπορεί να κάνει κλικ στο εικονίδιο, ώστε να λάβει πληροφορίες που αφορούν στην προτεινόμενη από το σύστημα μετακίνηση (δηλαδή προτεινόμενη θέση κι οδηγίες πλοήγησης). Αξίζει να σημειωθεί ότι τα ξεχωριστά RP που λαμβάνονται από το TC φιλτράρονται από τη διεπαφή χρήστη, καθώς χρησιμοποιούνται μόνο για να διατηρήσουν την ένδειξη συμφόρησης ενεργή, έχοντας, έτσι, ως αποτέλεσμα ελάχιστη αλληλεπίδραση μεταξύ του συστήματος και του χρήστη. Αυτό το χαρακτηριστικό καθιστά τη διεπαφή χρήστη φιλική και μη ενοχλητική, σύμφωνα και με τις αρχές του διάχυτου υπολογισμού [22], [23].

4.3.4 Δημιουργία Πρότασης Μετακίνησης

Σε αυτήν την παράγραφο, περιγράφεται πως το MITOS βοηθάει τους χρήστες στο να βελτιώσουν τη σύνδεση τους. Σε περίπτωση συμφόρησης, δηλαδή όταν ο φόρτος σε ένα ή περισσότερα AP ξεπεράσει ένα συγκεκριμένο κατώφλι, ο σκοπός του συστήματος είναι να αποσυμφορήσει τα υπερφορτωμένα AP κατανέμοντας το φόρτο στα γειτονικά AP (δηλαδή να παρακινήσει τους χρήστες που εξυπηρετούνται

από κορεσμένα AP να μετακινηθούν σε κοντινές θέσεις, όπου θα μπορέσουν να συνδεθούν με λιγότερο φορτωμένα AP). Για την αξιολόγηση μίας υποψήφιας θέσης για έναν χρήστη, λαμβάνονται υπόψη οι επόμενοι παράγοντες: 1) ο φόρτος L στην εξεταζόμενη θέση και 2) η απόσταση D που θα χρειαστεί ένας χρήστης να διανύσει για να μεταβεί στην εν λόγω θέση. Για μία μετακίνηση από τη θέση A στη θέση B, που συμβολίζεται με $(A \rightarrow B)$, υιοθετήθηκε η συνάρτηση κόστους $\text{Cost}(A \rightarrow B)$, η οποία είναι ένας γραμμικός συνδυασμός του φόρτου L_B στη θέση B και της απόστασης D_{AB} από τη θέση A στη θέση B (παρόμοια συνάρτηση κόστους έχει υιοθετηθεί και στην εργασία [38]):

$$\text{Cost}(A \rightarrow B) = \begin{cases} a \frac{D_{AB}}{\text{maxDistance}} + (1 - a) \cdot L_B, & D_{AB} \leq \text{maxDistance} \\ \infty, & \text{otherwise} \end{cases} \quad (4.1)$$

$$0 \leq a \leq 1$$

Το σύστημα προτείνει στο χρήστη τη θέση (B) με το μικρότερο κόστος, σε μία ακτίνα μικρότερη από maxDistance (δηλαδή θέσεις πέραν της απόστασης maxDistance αποκλείονται από το σύνολο των δυνατών επιλογών). Πρέπει να σημειωθεί ότι η τρέχουσα θέση του χρήστη δεν εξαιρείται από το σύνολο των επιλογών. Ο φόρτος, L_B , κι ο παράγοντας βάρους, a , παίρνουν τιμές από το διάστημα $[0, 1]$. Οι παράμετροι a και maxDistance θα μπορούσαν να ορίζονται κι από το χρήστη, ή να περιέχονται σε προφίλ χρηστών, υποδηλώνοντας, έτσι την προτίμηση μεταξύ αφθονίας σε εύρος ζώνης και γεωγραφικά περιορισμένων μετακινήσεων. Εναλλακτικά, το σύστημα θα μπορούσε να παρατηρεί τις αντιδράσεις του χρήστη σε σχέση με τα RP και δυναμικά να προσαρμόζει τον παράγοντα βάρους του χρήστη a (π.χ. να μειώνει το a , όταν ένα RP υιοθετείται). Τα RP που στέλνει το σύστημα συνοδεύονται κι από σχετικές πληροφορίες πλοήγησης, που προκύπτουν από τα προ-υπολογισμένα συντομότερα μονοπάτια μεταξύ των θέσεων στο WLAN.

Ο φόρτος L_B στην (4.1) υπολογίζεται αν μετρηθεί ο φόρτος στο AP στο οποίο θα συνδεόταν ο χρήστης αν βρισκόταν στη θέση B. Στην πρωτότυπη υλοποίηση που έχει γίνει, η μέτρηση αυτή πραγματοποιείται ως εξής. Ο NRM ζητά τις χρονικές

στιγμές nT , $n = 1, 2, \dots$, από τα υπάρχοντα AP το μέτρο του στιγμιαίου φόρτου τους για τη n -στη χρονοθυρίδα $[(n-1)T, nT)$, $n = 1, 2, \dots$. Το AP k απαντά στην ερώτηση του NRM στέλνοντας του το πλήθος των bytes που έχουν μεταφερθεί. Βάσει αυτής της πληροφορίας, ο NRM υπολογίζει το φόρτο $\gamma_k(n)$ του AP k για τη n -στη χρονοθυρίδα ως εξής:

$$\gamma_k(n) = \max(u_k(n), d_k(n)),$$

όπου το $u_k(n)$ ($d_k(n)$) αντιστοιχεί στο κλάσμα του χρόνου που το AP k λαμβάνει (στέλνει) δεδομένα μέσω της ανερχόμενης ζεύξης (κατερχόμενης ζεύξης – downlink) της ασύρματης διεπαφής για τη n -στη χρονοθυρίδα. Ειδικότερα, το $u_k(n)$ ($d_k(n)$) είναι ο αριθμός των bytes που έχουν ληφθεί (εκπεμφθεί) κατά τη n -στη χρονοθυρίδα προς τον αριθμό των bytes που είναι δυνατόν να λάβει (στείλει) η διεπαφή k στο μέγιστο ρυθμό, στη διάρκεια μίας χρονοθυρίδας.

Λόγω της «εκρηκτικότητας» (burstiness) της IP κίνησης, ο δείκτης φόρτου, $\gamma_k(n)$, είναι δυνατόν να παρουσιάζει σημαντική διακύμανση με το χρόνο, πράγμα που μπορεί να οδηγήσει σε άσκοπους «συναγερμούς» συμφόρησης. Για την αποφυγή του προβλήματος αυτού, στην υλοποίηση του συστήματος, έγινε προσπάθεια να εξομαλυνθεί η μέτρηση του φόρτου για το AP k , $\gamma_k(n)$, μέσω ενός χαμηλοπερατού φίλτρου, το οποίο βασίζεται στη μέση τιμή $L_k(n)$ από W διαδοχικά δείγματα από ένα αντίστοιχο χρονικό παράθυρο μήκους $W \times T$:

$$L_k(n) = \frac{1}{W} \sum_{i=0}^{W-1} \gamma_k(n-i) \quad (4.2)$$

Στην υλοποίηση, χρησιμοποιήθηκε $T = 500$ ms και $W = 16$. Πειράματα που έγιναν έδειξαν ότι οι προαναφερθείσες τιμές είναι ικανές να απορροφήσουν πρόσκαιρες αιχμές φόρτου, ενώ παράλληλα να διατηρήσουν την αποκρισιμότητα του συστήματος σε συμβάντα συμφόρησης³. Η τιμή του φόρτου ενός AP όπως υπολογίζεται από τη (4.2) είναι αυτή που χρησιμοποιείται στην (4.1) για τον υπολογισμό του κόστους μετακίνησης.

³ Η εξομάλυνση της τιμής του φόρτου μπορεί να επιτευχθεί και με εναλλακτικούς μηχανισμούς φιλτραρίσματος (π.χ. ένα εκθετικό φίλτρο κινούμενου μέσου – exponential moving average filter)

4.3.5 Ζητήματα Σχετικά με τις Προτάσεις Μετακίνησης

Ένα σημαντικό θέμα σε ό,τι αφορά τη δημιουργία των RP είναι πώς να αντιμετωπιστεί η περίπτωση των «επικαλυπτόμενων» κυψελών· δηλαδή η περίπτωση κατά την οποία κάποιες θέσεις καλύπτονται ταυτόχρονα από περισσότερα του ενός AP. Αυτό το ζήτημα, όμως, στην πράξη, δεν αποτελεί πρόβλημα, καθώς τέτοιες επικαλύψεις μπορούν να απεικονιστούν στο LG (και τις σχετικές δομές δεδομένων) και να ληφθούν υπόψη κατά τη δημιουργία των RP. Μπορούν επίσης να ληφθούν υπόψη κι άλλες ιδιαιτερότητες του προβλήματος (όπως ρυθμίσεις στην ισχύ εκπομπής που πραγματοποιούνται για την αποφυγή μεταπομπών, οι οποίες, όμως, μπορούν να βλάψουν τις συνδέσεις άλλων χρηστών σε γειτονικές κυψέλες), ώστε οι δημιουργούμενες προτάσεις να είναι όσο το δυνατόν πιο αποδοτικές.

Κατά την ανάπτυξη του πρωτοτύπου του MITOS [52], προέκυψε το ερώτημα αν θα πρέπει όλοι ανεξαιρέτως οι χρήστες που πλήττονται από συμφόρηση να λαμβάνουν ένα RP. Είναι σχετικά απίθανο όλοι οι χρήστες να λάβουν υπόψη ένα RP και να ακολουθήσουν τη σύσταση του συστήματος. Κάποιοι από αυτούς δε θα δείξουν ιδιαίτερη προθυμία να αποχωριστούν τη θέση τους για διάφορους λόγους, ενώ άλλοι μπορεί να είναι περαστικοί από το σημείο που σημειώθηκε η συμφόρηση έχοντας κατά νου να πάνε σε μία συγκεκριμένη θέση. Έτσι, στέλνοντας RP σε ένα τυχαία επιλεγμένο υποσύνολο χρηστών, είναι πολύ πιθανόν ότι κάποιοι που θα ήταν πρόθυμοι να μετακινηθούν δε θα λάβουν κάποια σχετική προτροπή, ενώ κάποιοι άλλοι θα λάβουν ένα RP που δε θα τους ενδιαφέρει. Από την άλλη μεριά, η συχνότητα των αποστελλόμενων RP, καθώς κι η απόσταση που απαιτείται να διανυθεί για την υιοθέσιά τους αποτελούν κι αυτά ένα ζήτημα. Στην παρούσα μελέτη, γίνεται η υπόθεση ότι επιθυμητές είναι οι λύσεις στις οποίες μπορεί να επιτευχθεί μειωμένη συχνότητα δημιουργίας RP με σχετικά μικρές αποστάσεις μετακίνησης για όσους τις ακολουθούν. Ένα άλλο ζήτημα είναι η επίδραση που μπορεί να έχει η μετάδοση των RP μέσω του ήδη κορεσμένου δικτύου. Αν το MITOS στέλνει RP συνεχώς, μπορεί να ειπωθεί ότι το κόστος για το δίκτυο θα είναι σημαντικό. Τέτοια θέματα μελετούνται στο υποκεφάλαιο 4.6, όπου συζητούνται αποτελέσματα από την προσομοίωση του συστήματος.

Στην υλοποίηση του MITOS, όλοι οι χρήστες μεταχειρίζονται κατά τον ίδιο τρόπο. Οι χρήστες που βρίσκονται στην ίδια θέση, σε περίπτωση συμφόρησης, θα λάβουν πανομοιότυπα RP. Έτσι, αν όλοι (ή σχεδόν όλοι) οι χρήστες ακολουθήσουν την πρόταση μετακίνησης που περιέχεται μέσα στο RP, όλη η κίνηση θα μεταφερθεί από το ένα AP στο άλλο, αντί ο φόρτος να μοιραστεί σε όλη την επικράτεια του δικτύου. Στο υποκεφάλαιο 4.4, συζητείται μία πιθανή λύση στο συγκεκριμένο πρόβλημα βασισμένη σε ένα μικροοικονομικό μοντέλο (με βάση τις δημοπρασίες). Στο υποκεφάλαιο 4.5, μελετάται ένα εναλλακτικό σχήμα, το οποίο βασίζεται σε ένα παιγνιοθεωρητικό μοντέλο, με την ονομασία “Santa Fe Bar Problem” (δες Κεφάλαιο 3, παράγραφο 3.4.1). Οι προαναφερθείσες λύσεις επιτυγχάνουν βελτιωμένα αποτελέσματα και πιο προβλέψιμη συμπεριφορά συστήματος, σε σχέση με ένα σύστημα στο οποίο απλά οι χρήστες διαφοροποιούνται μόνο μέσω της παραμέτρου a στην εξίσωση (4.1) [52].

4.4 Μικροοικονομικοί Αλγόριθμοι κι Εφαρμογή στο MITOS

Σε αυτό το υποκεφάλαιο, με βάση την εργασία [38], περιγράφεται ένας αλγόριθμος βασισμένος σε δημοπρασίες για το μ E-MITOS (μ E σημαίνει *microeconomic*). Η λεπτομερής περιγραφή του αλγορίθμου που σκιαγραφείται παρακάτω υπάρχει στο Παράρτημα I.

Ο μηχανισμός για τη δημιουργία RP στο μ E-MITOS δομείται γύρω από *πολλαπλές επαναλαμβανόμενες δημοπρασίες* (δηλαδή μία σειρά από δημοπρασίες πρώτης τιμής με κλειστές προσφορές – *first-price sealed-bid auctions*). Σε περίπτωση συμφόρησης, κάθε χρήστης που πλήττεται από αυτή, επιλέγει την προτιμότερη γι’ αυτόν θέση (με βάση την εξίσωση (4.1)). Οι χρήστες στη συνέχεια ομαδοποιούνται βάσει του AP που προτιμούν (το AP που εξυπηρετεί τη θέση που προτιμούν) και μία δημοπρασία διενεργείται μεταξύ των μελών κάθε ομάδας.

Κάθε χρήστης σε μία ομάδα υποβάλλει την προσφορά του για το προτιμώμενο AP. Η προσφορά του χρήστη πρέπει πάντα να είναι υψηλότερη από την τιμή του συγκεκριμένου AP, η οποία αναφέρεται κι ως *cost* (κόστος του AP). Το *cost* τίθεται αρχικά από το σύστημα σε κάθε συμβάν συμφόρησης (ίσο για όλα τα AP). Παρομοίως, η προσφορά ενός χρήστη φράσσεται κι ως προς τη μέγιστη τιμή της,

money, η οποία είναι η ίδια για όλους τους χρήστες και μεγαλύτερη από το *cost*. Η ακριβής τιμή της προσφοράς του χρήστη καθορίζεται από τον «παράγοντα προσφοράς» (*bidding factor*), β , ο οποίος παίρνει τιμές από το διάστημα [0,1]. Συγκεκριμένα, η τιμή της προσφοράς, *bid*, του χρήστη έχει ως εξής, [38]:

$$bid = cost + \beta \cdot (money - cost).$$

Όσο υψηλότερη είναι η τιμή του παράγοντα προσφοράς β , τόσο πιο επιθετικές είναι κι οι προσφορές που κάνει ο χρήστης. Η τιμή του παράγοντα προσφοράς ανανεώνεται ανάλογα με την έκβαση της δημοπρασίας: οι νικητές μειώνουν τον παράγοντα προσφοράς τους (δηλαδή γίνονται λιγότερο επιθετικοί), ενώ οι χαμένοι αυξάνουν τον παράγοντα προσφοράς τους, ώστε να αυξήσουν την πιθανότητα νίκης τους στις επόμενες δημοπρασίες.

Ο νικητής μίας δημοπρασίας μέσα σε μία ομάδα κερδίζει την προτίμηση του και δε θεωρείται πλέον ότι πλήττεται από συμφόρηση. Αν περισσότεροι από ένας χρήστες έχουν κάνει την ίδια μέγιστη προσφορά (δηλαδή υπάρχει ισοπαλία), ο νικητής επιλέγεται τυχαία. Μετά την ολοκλήρωση μίας δημοπρασίας, ο φόρτος κάθε AP ενημερώνεται για να απεικονίσει την πιθανή (υποθετική) μετακίνηση του χρήστη. Επιπλέον, το κόστος (*cost*) του AP τίθεται ίσο με την τιμή της νικητήριας προσφοράς.

Η ίδια διαδικασία (δηλαδή ομαδοποίηση, δημοπρασία κι εκχώρηση πόρων) επαναλαμβάνεται με τους υπόλοιπους χρήστες που πλήττονται από τη συμφόρηση, μέχρις ότου όλοι οι χρήστες κερδίσουν μία δημοπρασία⁴. Καθώς η τιμή κι ο υποθετικός φόρτος των AP αλλάζουν, οι προτιμήσεις των χρηστών σταδιακά αποκλίνουν. Οι συγκρούσεις επιλύονται και τελικά επιτυγχάνεται εξισορρόπηση φόρτου. Μετά τον τερματισμό του αλγορίθμου, κάθε χρήστης έχει «εικονικά» αγοράσει ένα AP, ή καλύτερα, είναι συνιδιοκτήτης του AP με κάποιους άλλους χρήστες. Οι χρήστες ενημερώνονται αναλόγως μέσω των αντίστοιχων RP που απεικονίζουν το αποτέλεσμα της δημοπρασίας.

Η λογική πίσω από τη χρήση του παράγοντα προσφοράς στο μοντέλο των δημοπρασιών που περιγράφεται πιο πάνω είναι να προσομοιώσει τη στρατηγική

⁴ Είναι επίσης δυνατόν για ιδιωτικούς πράκτορες που βρίσκονται στη σταθερή υποδομή να εκπροσωπούν τους χρήστες, υποβάλλοντας προσφορές εκ μέρους τους. Με αυτόν τον τρόπο, οι χρήστες δε χρειάζεται να υποβάλλουν προσφορές μέσω του, πιθανώς, κορεσμένου ασύρματου δικτύου. Έτσι, το μόνο μήνυμα του συστήματος που περνά μέσω του ασύρματου δικτύου είναι το RP.

ενός «λογικού» χρήστη, ο οποίος θα απέφευγε να κάνει προσφορές με το μέγιστο δυνατό ποσό (*money*). Συγκεκριμένα, αν ένας χρήστης πίστευε ότι θα έπρεπε να κάνει τη μέγιστη δυνατή προσφορά, για να κερδίσει τη δημοπρασία, τότε κι οι υπόλοιποι χρήστες θα σκέπτονταν με τον ίδιο τρόπο. Έτσι, όλοι οι χρήστες θα έκαναν τη μέγιστη δυνατή προσφορά κι ο νικητής θα επιλεγόταν τυχαία, κάτι που δεν είναι επιθυμητό.

4.5 Ένα Σχήμα Προγνώσεων για το MITOS

Σε αυτό το υποκεφάλαιο, περιγράφεται ένα παιγνιοθεωρητικό σχήμα, το οποίο βασίζεται στο SFBP (δες παράγραφο 3.4.1), για την επίτευξη της εξισορρόπησης φόρτου στη δικτυακή υποδομή. Το προτεινόμενο σχήμα βασίζεται σε προβλέψεις για τον δικτυακό φόρτο.

Υπάρχει στενή σχέση μεταξύ του SFBP και της παραλλαγής του MITOS που επιδιώκει εξισορρόπηση φόρτου στο δίκτυο. Οι ταλαντώσεις στο φόρτο είναι ανεπιθύμητες, καθώς υπονομεύουν την απόδοση του συστήματος. Με κίνητρο, λοιπόν, τις ιδιότητες του SFBP, το MITOS εμπλουτίστηκε με ευφυΐα πρόγνωσης, με σκοπό την επίτευξη της εξισορρόπησης φόρτου και την αποφυγή ταλαντώσεων.

Στο SFBP-MITOS, η μετρίσιμη ποσότητα, που αντιστοιχεί στην προσέλευση του μπαρ στο SFBP, είναι ο φόρτος του AP. Προς την επίτευξη της εξισορρόπησης του φόρτου στο δίκτυο, λαμβάνει χώρα, στο παρασκήνιο (διάφανα προς το χρήστη), μία διαδικασία πρόβλεψης. Αυτό επιτυγχάνεται βελτιώνοντας την αρχική αρχιτεκτονική του MITOS με ένα επιπλέον «συστατικό πρόγνωσης» για κάθε χρήστη.

Το συστατικό πρόγνωσης περιέχει K προγνώστες, οι οποίοι επιλέγονται τυχαία για τον κάθε χρήστη από ένα σύνολο M διαθέσιμων προγνωστών. Ο χρόνος διαιρείται σε χρονοθυρίδες διάρκειας T sec· το διάστημα $[nT, (n+1)T)$ αποτελεί τη χρονοθυρίδα $(n+1)$, $n = 0, 1, 2, \dots$. Οι μεταβλητές που ορίζονται παρακάτω αναφέρονται στη διαδικασία της πρόγνωσης του φόρτου ενός συγκεκριμένου AP από έναν συγκεκριμένο χρήστη· παρόλ' αυτά, για απλότητα στο συμβολισμό, δε χρησιμοποιούνται οι δείκτες που καθορίζουν το χρήστη και το AP. Έστω, λοιπόν, $P_i(n)$ η πρόβλεψη για το φόρτο για το χρόνο $(n+1)T$ που έγινε από τον προγνώστη

i τη χρονική στιγμή nT , $i = 1, 2, \dots, K$. Έστω $L(n)$ η αντίστοιχη μέτρηση φόρτου τη χρονική στιγμή nT , $n = 1, 2, \dots$, όπως υπολογίζεται μέσω της εξίσωσης (4.2). Το $P_i(n)$ και το $L(n)$ κανονικοποιούνται για να παίρνουν τιμές από το διάστημα $[0,1]$, υποδηλώνοντας το βαθμό κορεσμού του AP. Έστω θ_n ο δείκτης του προγνώστη που ελαχιστοποιεί την ποσότητα $|L(n) - P_i(n-1)|$, δηλαδή

$$\theta_n = \operatorname{argmin}_i \{|L(n) - P_i(n-1)|\}, n = 1, 2, \dots \quad (4.3)$$

Από εδώ και στο εξής, ο προγνώστης θ_n θα αναφέρεται ως ο «ενεργός» προγνώστης για τη χρονοθυρίδα n . Κάθε χρήστης, σε κάθε χρονοθυρίδα, χρησιμοποιεί την πρόβλεψη που του παρέχει ο ενεργός προγνώστης ως την «επίσημη» πρόβλεψη $P(n)$ του φόρτου για τη χρονική στιγμή $(n+1)T$:

$$P(n) = P_{\theta_n}(n), n = 1, 2, \dots \quad (4.4)$$

Ο Πίνακας 4.1 περιέχει τους τύπους των προγνωστών που έχουν υιοθετηθεί, [27]. Οι τύποι 1, 3, 4 αποτελούν, ουσιαστικά, παραμετροποιήσιμες οικογένειες προγνωστών. Με διαφορετικές τιμές στις εν λόγω παραμέτρους είναι δυνατόν να παραχθεί ένας μεγαλύτερος αριθμός διακριτών προγνωστών που θα σχηματίσουν το σύνολο των M διαθέσιμων προγνωστών, από το οποίο ο κάθε χρήστης επιλέγει τυχαία K . Οι μεταβλητές K και M αποτελούν σχεδιαστικές παραμέτρους. Όπως έχει περιγραφεί στην εργασία [54], μία μεγάλη τιμή του K (περισσότεροι προγνώστες που λειτουργούν ταυτόχρονα) έχει ως αποτέλεσμα χρήστες με αυξημένη «ευφυΐα»· παρόλ' αυτά, όταν το K αυξάνεται, περισσότεροι προγνώστες γίνονται κοινοί μεταξύ των χρηστών. Αυτό σημαίνει ότι η ετερογένεια των χρηστών (διαφορετικός τρόπος σκέψης) υπονομεύεται κι η συνολική απόδοση του συστήματος υποβαθμίζεται, λόγω ταλαντώσεων. Παρόμοια αποτελέσματα αναφέρονται και στο Παράρτημα II, όπου παρουσιάζονται αποτελέσματα προσομοιώσεων για διαφορετικούς αριθμούς προγνωστών ανά χρήστη. Το πρόβλημα της βέλτιστης επιλογής των τιμών των παραμέτρων K και M είναι αρκετά δύσκολο να αναλυθεί, καθώς εξαρτάται κι από την ίδια φύση των προγνωστών και της μεταβλητής που μελετάται.

Πίνακας 4.1. Τύποι προγνώστων

Τύπος Προγνώστη	Περιγραφή	Συνάρτηση Πρόγνωσης
1	Όμοια όπως πριν i χρονοθυρίδες (κυκλικός προγνώστης)	$P(n) = L(n - i), i = 0, 1, \dots$
2	Κατοπτρικά γύρω από το 0.5 (μισή χωρητικότητα) της τελευταίας χρονοθυρίδας	$P(n) = 0.5 + \text{sign}(0.5 - L(n)) \cdot L(n) - 0.5 = 1 - L(n)$
3	Ίσα με $const$ (στατικός προγνώστης)	$P(n) = const, const \in [0, 1]$
4	Μέση τιμή των τελευταίων y χρονοθυρίδων (προγνώστης μέσης τιμής)	$P(n) = \frac{1}{y} \sum_{i=0}^{y-1} L(n - i), y = 1, 2, \dots$
5	Μέση τιμή με βάρη για τις τελευταίες y χρονοθυρίδες (προγνώστης μέσης τιμής)	$P(n) = \sum_{i=0}^{y-2} \frac{1}{2^{i+1}} L(n - i) + \frac{1}{2^{y-1}} L(n - (y - 1))$
6	Παρεμβολή από τις τελευταίες δύο χροθυρίδες	$P(n) = L(n) + (L(n) - L(n - 1)) = 2 \cdot L(n) - L(n - 1)$

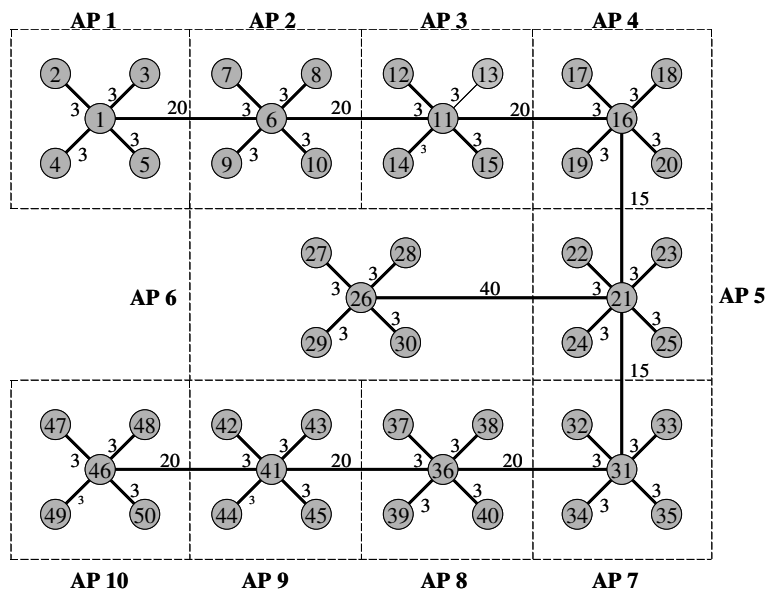
Σε περίπτωση συμφόρησης, η μετρική $P(n)$, όπως ορίζεται στην εξίσωση (4.4), χρησιμοποιείται στην εξίσωση (4.1) για να βρεθεί η καλύτερη θέση για να μεταβεί ο χρήστης. Συγκεκριμένα, τη χρονική στιγμή nT , αντί να χρησιμοποιηθεί η τρέχουσα τιμή του φόρτου $L(n)$ για τη μετρική L_B , στην εξίσωση (4.1), χρησιμοποιείται η μετρική $P(n)$, δηλαδή η πρόβλεψη της τιμής φόρτου για τη χρονική στιγμή $(n+1)T$. Η προκύπτουσα θέση με το ελάχιστο κόστος, μαζί με πληροφορίες πλοήγησης, διαβιβάζεται εν συνεχεία στο χρήστη, βάσει της καθιερωμένης διαδικασίας δημιουργίας RP από το MITOS.

4.6 Μοντέλο Προσομοίωσης

Ένα πλήρως λειτουργικό πρωτότυπο του MITOS υλοποιήθηκε κι εγκαταστάθηκε στο Τμήμα μας [52]. Για να εξάγουμε σαφή συμπεράσματα για την απόδοση του συστήματος, είναι αναγκαίο να διεξαχθούν μελέτες με πραγματικούς χρήστες για ένα αρκετά μεγάλο πληθυσμό χρηστών. Κάτι τέτοιο είναι, όμως, αρκετά δύσκολο να επιτευχθεί. Γι' αυτόν το λόγο, για να αξιολογήσουμε την απόδοση των

συστημάτων μE-MITOS και SFBP-MITOS, έγιναν εκτεταμένες προσομοιώσεις, για ένα συγκεκριμένο πλήθος χρηστών.

Στις προσομοιώσεις, υποτέθηκε μία διαρρύθμιση χώρων, όπως αυτή απεικονίζεται στην Εικόνα 4.2, με αρκετές διαφορετικές θέσεις, όπως διάδρομοι, γραφεία κι ένα αίθριο, που καλύπτονται από διαφορετικά AP. Η συγκεκριμένη διαρρύθμιση επελέγη ως μία απλοποιημένη εικόνα του κτηρίου του Τμήματος Πληροφορικής και Τηλεπικοινωνιών.



Εικόνα 4.2. Χωρική διαρρύθμιση στην προσομοίωση

Όπως φαίνεται στην Εικόνα 4.2, υπάρχουν 50 θέσεις και 10 AP που τις καλύπτουν. Οι θέσεις 1, 6, 11, 16, 21, 31, 36, 41 και 46 μπορούν να θεωρηθούν ως τμήματα διαδρόμων, η θέση 26 ως το αίθριο του κτηρίου κι οι υπόλοιπες θέσεις ως γραφεία. Κάθε AP καλύπτει μία περιοχή αποτελούμενη από 5 θέσεις (στην Εικόνα 4.2, η περιοχή κάλυψης ενός AP απεικονίζεται ως ένα τετράγωνο με διακεκομμένες γραμμές). Η ακμή που συνδέει δύο θέσεις έχει ένα βάρος, το οποίο υποδηλώνει την φυσική τους απόσταση. Η μετακίνηση των χρηστών από μία θέση σε μία άλλη γίνεται στοχαστικά και μοντελοποιείται ως μία αλυσίδα Markov με πίνακα πιθανοτήτων μεταβάσεων P . Αυτός ο πίνακας έχει κατασκευαστεί με τρόπο που, μακροπρόθεσμα, η πλειοψηφία των κινητών χρηστών συγκεντρώνεται στην περιοχή που σχηματίζουν οι θέσεις 5 και 6 (hot-spot). Για λόγους παρουσίασης

φαίνεται εδώ μόνο μία «σύνοψη» του πίνακα μεταβάσεων⁵ P , ο πίνακας F . Ο πίνακας μεταβάσεων F περιγράφει τη στοχαστική μετακίνηση των χρηστών μεταξύ περιοχών κάλυψης, δηλαδή το (i,j) στοιχείο, $F(i,j)$, είναι μία προσέγγιση της πιθανότητας μετάβασης από την περιοχή κάλυψης i στην περιοχή κάλυψης j .

$$F = \begin{bmatrix} 4/10 & 6/10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2/10 & 3/10 & 5/10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2/10 & 3/10 & 5/10 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2/10 & 4/10 & 4/10 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5/10 & 5/10 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4/10 & 0 & 4/10 & 2/10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5/10 & 3/10 & 2/10 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5/10 & 3/10 & 2/10 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6/10 & 4/10 \end{bmatrix}$$

Για να γίνει διάκριση των χρηστών με διαφορετικά χαρακτηριστικά κίνησης, η πιθανότητα μετάβασης $P(i,j)$ της μετακίνησης στη θέση j από τη θέση i μεταβλήθηκε ανάλογα με τον τύπο του χρήστη. Συγκεκριμένα, ο πίνακας P θεωρείται ως ένας βασικός πίνακας μεταβάσεων, από το οποίο παράγεται ένας πίνακας P' με βάση τον τύπο του χρήστη. Έτσι, αν ένας χρήστης επιδεικνύει υψηλή κινητικότητα, η πιθανότητα, $P(i,i)$, του να παραμείνει στην ίδια θέση μειώνεται, ενώ η πιθανότητα να φύγει από τη τρέχουσα θέση προσαρμόζεται κατάλληλα. Παρομοίως, ένας χρήστης με χαμηλή κινητικότητα έχει αυξημένη πιθανότητα να παραμείνει στην τρέχουσα θέση του. Ο Πίνακας 4.2 ορίζει τις τροποποιημένες πιθανότητες μετάβασης για τρεις τύπους κινητικότητας χρηστών: υψηλής, μέτριας και χαμηλής. Το Ω_i υποδηλώνει το σύνολο των περιοχών που γειτνιάζουν με την περιοχή i και το $|\Omega_i|$ το πλήθος των στοιχείων του συνόλου Ω_i . Στις προσομοιώσεις, για τους χρήστες υψηλής κινητικότητας, ίσχυε $b = 0.5$, για τους μέτριας κινητικότητας, $b = c = 1$, ενώ για τους χαμηλής κινητικότητας $c = 0.5$.

⁵ Ο συνοπτικός πίνακας F αποτελεί μία συγκεντρωτική αναπαράσταση του βασικού πίνακα P . Συγκεκριμένα, τα στοιχεία του πίνακα P ομαδοποιούνται βάσει της κάλυψης των AP κι οι αντίστοιχες πιθανότητες μετάβασης (δηλαδή από ένα AP σε ένα άλλο) προκύπτουν αναλόγως.

Πίνακας 4.2. Πιθανότητες μετάβασης διαφοροποιημένες βάσει του τύπου κινητικότητας του χρήστη

		Χωρική μετάβαση $i \rightarrow j$	Παραμονή στη θέση i
Επίπεδο Κινητικότητας	Υψηλό	$P'(i, j) = P(i, j) + (P(i, i) - P'(i, i))/ \Omega_i ,$ $j \in \Omega_i, i \in \{1, \dots, 50\}$	$P'(i, i) = b \cdot P(i, i),$ $i \in \{1, \dots, 50\}, 0 < b \leq 1$
	Μέτριο	$P'(i, j) = P(i, j)$	$P'(i, i) = P(i, i)$
	Χαμηλό	$P'(i, j) = c \cdot P(i, j),$ $j \in \Omega_i, i \in \{1, \dots, 50\}, 0 < c \leq 1$	$P'(i, i) = P(i, i) + \sum_{j \in \Omega_i} (P(i, j) - P'(i, j)),$ $i \in \{1, \dots, 50\}$

Τα χαρακτηριστικά των εφαρμογών που τρέχουν οι χρήστες αποτελούν ένα ακόμα σημαντικό ζήτημα. Θεωρούνται τρεις τύποι εφαρμογών με συγκεκριμένα χαρακτηριστικά σε ό,τι αφορά το απαιτούμενο εύρος ζώνης, τη διάρκεια και τη συχνότητα εμφάνισής τους. Για τους πρώτους δύο τύπους, δηλαδή το VoIP και το video, χρησιμοποιήθηκε ένα απλό μοντέλο Markov δύο καταστάσεων (ON/OFF)-δηλαδή, οι διάρκειες των εναλλασσόμενων περιόδων ON και OFF είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν την εκθετική κατανομή. Για την εφαρμογή WWW χρησιμοποιήθηκε επίσης ένα μοντέλο ON/OFF, με ανεξάρτητες περιόδους ON και OFF, οι οποίες ακολουθούν την κατανομή Pareto [56]. Ο Πίνακας 4.3 συνοψίζει τα χαρακτηριστικά των εφαρμογών που χρησιμοποιήθηκαν στις προσομοιώσεις. Υποτέθηκε ότι κατά τη διάρκεια των περιόδων ON, οι εφαρμογές εξέπεμπαν με σταθερό ρυθμό. Στις προσομοιώσεις, υποτέθηκε ότι οι χρήστες έχουν κοινά χαρακτηριστικά εφαρμογών, που απεικονίζει ο Πίνακας 4.3, δηλαδή οι ανάγκες τους σε εύρος ζώνης είναι παρόμοιες.

Πίνακας 4.3. Χαρακτηριστικά των εφαρμογών

Τύπος Εφαρμογής	Κατανομή Διάρκειας Περιόδου ON/OFF	Μέση διάρκεια περιόδου OFF (s)	Μέση διάρκεια περιόδου ON (s)	Ρυθμός Εκπομπής (Kb/s)
VoIP	Εκθετική	200	180	64
Video	Εκθετική	400	240	1000
WWW	Pareto (> 1 s)	8	10	180

Στις προσομοιώσεις, το σύστημα, σε τακτά χρονικά διαστήματα (κάθε *Congestion check period* – Πίνακας 4.4), ελέγχει εάν υπάρχουν κορεσμένα AP και σε αυτήν την περίπτωση στέλνει RP στους χρήστες που πλήττονται από τη συμφόρηση (οι οποίοι τελικά ενημερώνονται εάν το AP που τους εξυπηρετεί είναι κορεσμένο – δες παράγραφο 4.3.3). Τα RP γίνονται σεβαστά από τους χρήστες στοχαστικά. Η πιθανότητα αποδοχής ενός RP, P_a , τίθεται σε μία αρχική τιμή, π.χ. $P_a = 0.5$, αλλά διαφοροποιείται στη συνέχεια βάσει των χαρακτηριστικών της κινητικότητας του χρήστη. Συγκεκριμένα, για χρήστες υψηλής κινητικότητας υιοθετείται η εξίσωση (4.5), ενώ για τους χαμηλής κινητικότητας η (4.6):

$$P'_a = 1 - b \cdot (1 - P_a) \quad (4.5)$$

$$P'_a = c \cdot P_a \quad (4.6)$$

Οι σταθερές b και c παίρνουν τις ίδιες τιμές που περιέχει ο Πίνακας 4.2· δηλαδή γίνεται η προφανής υπόθεση ότι η προθυμία στην αποδοχή των RP είναι συσχετισμένη με το βαθμό κινητικότητας του χρήστη, δηλαδή ένας χρήστης με υψηλή κινητικότητα είναι περισσότερο πιθανό να αποδεχθεί ένα RP σε σύγκριση με έναν χρήστη με χαμηλή κινητικότητα. Γίνεται επίσης η υπόθεση ότι, αν ένας χρήστης αποδεχθεί ένα RP, ακολουθεί το συντομότερο μονοπάτι που του προτείνεται από το σύστημα. Επιπλέον, όταν φτάσει στην προτεινόμενη θέση, δεν επιχειρεί κάποια νέα μετακίνηση για ένα συγκεκριμένο τουλάχιστον χρόνο, ο οποίος αναφέρεται εδώ ως *Minimum Cell Residence Time after RP adoption*, (Πίνακας 4.4). Μετά από αυτόν τον ελάχιστο χρόνο παραμονής στην προτεινόμενη θέση, ο χρήστης ακολουθεί το στοχαστικό μοντέλο κινητικότητας Markov. Οι υποθέσεις που έγιναν σε σχέση με το μοντέλο κινητικότητας στις προσομοιώσεις είχαν ως αποτέλεσμα μία σχετικά αυξημένη κινητικότητα χρηστών, σε σχέση με τις μελέτες επισκόπησης για τους χρήστες WLAN που μπορούν να βρεθούν στις εργασίες [45], [46], [47]. Η αυξημένη κινητικότητα των χρηστών, όμως, συνετέλεσε στην αξιολόγηση της προτεινόμενης αρχιτεκτονικής υπό έντονο φόρτο εργασίας, δηλαδή το συνολικό υπολογιστικό φόρτο που θα πρέπει το σύστημα να διεκπεραιώσει ως αποτέλεσμα της αυξημένης κινητικότητας.

Έστω μία εγκατάσταση με N AP. Το $D_i(n)$ υποδηλώνει τον στιγμιαίο προσφερόμενο φόρτο (ζήτηση σε εύρος ζώνης) για το AP i τη χρονική στιγμή nT , $n = 1, 2, \dots, i = 1,$

2, ..., N. Ορίζεται ο συντελεστής διασποράς (coefficient of variation), $CoV(n)$, του $D_i(n)$ ως

$$CoV(n) = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^N [D_i(n) - \bar{D}(n)]^2}}{\bar{D}(n)}, i \in \{1, 2, \dots, N\}, \quad (4.7)$$

όπου $\bar{D}(n)$ είναι η μέση τιμή του προσφερόμενου φόρτου σε όλη την επικράτεια του δικτύου, τη χρονική στιγμή nT (η μέση τιμή για όλα τα AP):

$$\bar{D}(n) = \frac{1}{N} \sum_{i=1}^N D_i(n) \quad (4.8)$$

Στις προσομοιώσεις, χρησιμοποιήθηκε ως η βασική μετρική αξιολόγησης της εξισορρόπησης φόρτου των προτεινόμενων λύσεων η χρονική μέση τιμή του τετραγώνου του $CoV(n)$ για όλη τη διάρκεια της προσομοίωσης H :

$$Dev = \frac{1}{H} \sum_{n=1}^H (CoV(n))^2 \quad (4.9)$$

Ο Πίνακας 4.4 περιέχει τις παραμέτρους των προσομοιώσεων. Για το SFBP-MITOS, δημιουργήθηκε ένα σύνολο από $M = 15$ διαθέσιμους προγνώστες χρησιμοποιώντας του έξι βασικούς τύπους προγνώστων που περιέχει ο Πίνακας 4.1 με τις διαφορετικές τιμές παραμέτρων που δείχνει ο Πίνακας 4.4. Αυτό επέτρεψε να αυξηθεί η ετερογένεια των χρηστών, σε σχέση με την περίπτωση ενός συνδυασμού αποτελούμενου από έναν προγνώστη ανά τύπο, δηλαδή μόνο έξι προγνώστες. Στις προσομοιώσεις, κάθε χρήστης ανεξάρτητα επέλεγε τυχαία $K = 2$ προγνώστες από το σύνολο των διαθέσιμων 15.

Τα βασικά κριτήρια για την αξιολόγηση των προτεινόμενων λύσεων είναι 1) η μετρική απόκλισης Dev (4.9) και 2) ο συνολικός αριθμός των απεσταλμένων RP στη διάρκεια της προσομοίωσης. Συνεχόμενα RP που είναι όμοια (δηλαδή προτείνουν την ίδια θέση στο χρήστη) μετρώνται επίσης. Αυτό που αναζητείται είναι μία ρύθμιση που να επιτυγχάνει πολύ μικρή απόκλιση (Dev), δηλαδή ομαλή κατανομή του φόρτου, με το λιγότερο δυνατόν αριθμό απεσταλμένων RP (μειωμένο κόστος για το δίκτυο και το σύστημα). Η απόδοση αξιολογείται βάσει δύο παραμέτρων: την πιθανότητα αποδοχής ενός RP, P_a , και το κατώφλι φόρτου για ειδοποίηση συμφόρησης. Στις προσομοιώσεις, το P_a έπαιρνε τιμές από 0.0 μέχρι

1.0 (με διαστήματα 0.01 στην περιοχή τιμών [0.0, 0.05] και 0.05 στην περιοχή [0.05, 1.0]) και το κατώφλι φόρτου από 5% ως 100%. Στις επόμενες παραγράφους, συζητείται η απόδοση που επιτυγχάνεται από τα προτεινόμενα σχήματα.

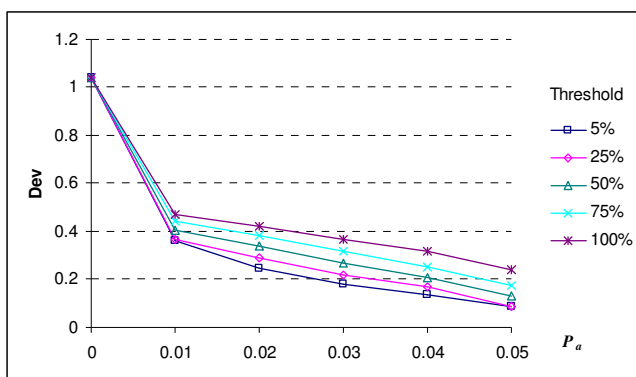
Πίνακας 4.4. Παράμετροι της προσομοίωσης

Παράμετρος Προσομοίωσης		Τιμή
Πλήθος χρηστών		100
Ποσοστό χρηστών υψηλής κινητικότητας		25
Ποσοστό χρηστών μέτριας κινητικότητας		50
Ποσοστό χρηστών χαμηλής κινητικότητας		25
Χωρητικότητα των AP (<i>BW</i>)		11 Mb/s
Παράγοντας βάρους <i>a</i> (δες εξίσωση 4.1)		0.6
Μέγιστη απόσταση (δες εξίσωση 4.1)		200 m
Πλήθος συνολικά διαθέσιμων προγνωστών (SFBP)		15
Μείγμα Διαθέσιμων Προγνωστών στην Προσομοίωση	Τύπος 1 (6 προγνώστες)	$i = 0, 1, 2, 3, 4$ και 5
	Τύπος 2 (1 προγνώστης)	-
	Τύπος 3 (4 προγνώστες)	$const = 0.1, 0.3, 0.7,$ και 0.9
	Τύπος 4 (2 προγνώστες)	$y = 4,$ and 10
	Τύπος 5 (1 προγνώστης)	$y = 10$
	Τύπος 6 (1 προγνώστης)	-
Πλήθος προγνωστών ανά χρήστη (SFBP)		2
Minimum Cell Residence Time after RP adoption		30 s
Περίοδος ελέγχου συμφόρησης		5 s
Διάρκεια προσομοίωσης (πραγματικός χρόνος)		~2.7 h (10,000 s)

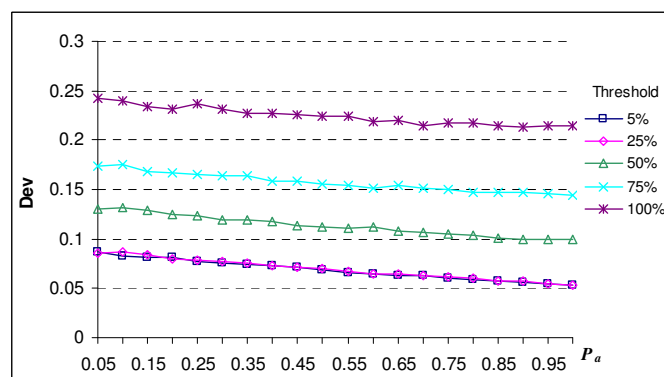
4.6.1 Αποτελέσματα Προσομοίωσης για το μE-MITOS

Στην Εικόνα 4.3 απεικονίζεται η απόκλιση και το πλήθος των απεσταλμένων RP που επιτυγχάνει το προτεινόμενο σύστημα μE-MITOS για διαφορετικές τιμές της πιθανότητας αποδοχής ενός RP. Στην Εικόνα 4.3(a) και Εικόνα 4.3(b),

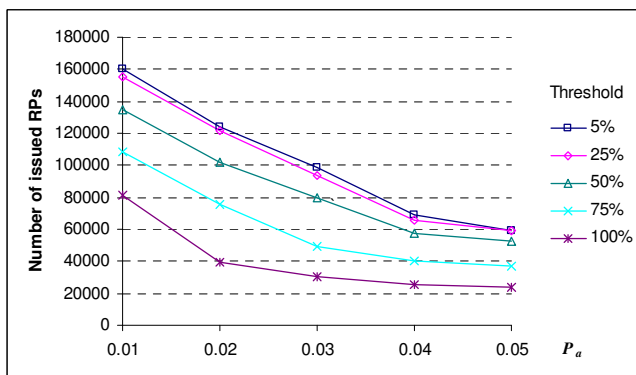
παρατηρείται ότι, όσο αυξάνει η πιθανότητα αποδοχής ενός RP, η απόκλιση σε γενικές γραμμές μειώνεται. Παρατηρείται επίσης ότι, όσο μειώνεται το κατώφλι συμφόρησης (δηλαδή αυξάνεται η απόκριση του συστήματος σε ό,τι αφορά τη συμφόρηση), μειώνεται κι η απόκλιση. Επίσης, αξίζει να σημειωθεί ότι η τιμή της απόκλισης, για το σενάριο χωρίς την ύπαρξη του MITOS (RP-free), είναι περίπου 1.05. Αυτό φαίνεται και στην Εικόνα 4.3(a) όπου για μηδενική πιθανότητα αποδοχής ενός RP (δηλαδή όταν κανένας χρήστης δεν λαμβάνει υπόψη τα RP που είναι ισοδύναμο με το RP-free σενάριο), $Dev \approx 1.05$. Έτσι, η βελτίωση που επιφέρει το μ E-MITOS είναι αρκετά σημαντική.



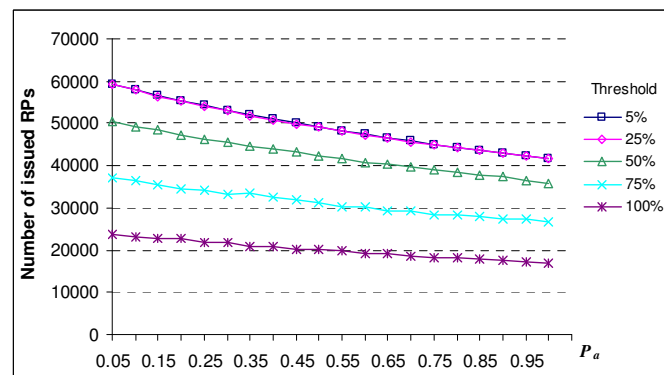
(a) Dev vs. $P_a \in [0, 0.05]$



(b) Dev vs. $P_a \in [0.05, 1.0]$



(c) πλήθος απεσταλμένων RP vs. $P_a \in [0.01, 0.05]$



(d) πλήθος απεσταλμένων RP vs. $P_a \in [0.05, 1]$

Εικόνα 4.3. μ E-MITOS: (a)-(b) Dev vs. P_a και (c)-(d) πλήθος απεσταλμένων RP vs. P_a .

Η σημαντική μείωση στην τιμή της μετρικής απόκλισης, ακόμα και για μικρές τιμές της πιθανότητας αποδοχής ενός RP, μπορεί να αποδοθεί σε δύο βασικούς λόγους: 1) οι χρήστες που παραμένουν στην κορεσμένη περιοχή λαμβάνουν περισσότερα από ένα RP, στη διάρκεια της περιόδου συμφόρησης. Αρχικά, ένας χρήστης μπορεί να αγνοήσει μία ένδειξη συμφόρησης στο TC του, αλλά στη συνέχεια να

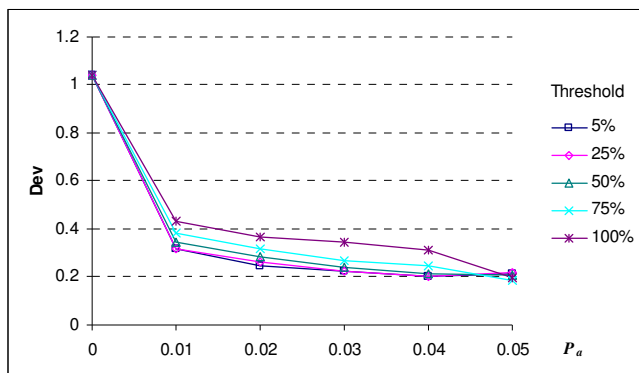
ακολουθήσει τις παρεχόμενες οδηγίες σε μία από τις επόμενες περιόδους. Έτσι, το ποσοστό των χρηστών που τελικά εγκαταλείπουν την κορεσμένη περιοχή (ροή εξόδου) είναι μακροπρόθεσμα μεγαλύτερο από την ονομαστική πιθανότητα αποδοχής των RP. Πρέπει να τονιστεί όμως ότι ο πληθυσμός των χρηστών στην κορεσμένη περιοχή δεν εξαντλείται σταδιακά, καθώς υπάρχει και μία «ροή εισόδου» λόγω του μοντέλου κινητικότητας Markov. 2) Η μετρική απόκλισης Dev επιδεικνύει μη γραμμική συμπεριφορά. Όσο λιγότερο εξισορροπημένο είναι το σύστημα, τόσο πιο μεγάλος είναι κι ο ρυθμός της μείωσης της μετρικής απόκλισης.

Στην Εικόνα 4.3(c) και Εικόνα 4.3(d) φαίνεται ο αριθμός των απεσταλμένων RP στη διάρκεια τη προσομοίωσης βάσει του P_a . Παρατηρείται ότι, καθώς αυξάνεται το P_a , ο αριθμός των RP σε γενικές γραμμές μειώνεται. Για παράδειγμα, για $P_a = 0.05$ και κατώφλι συμφόρησης 75%, αποστέλλονται 37,270 RP συνολικά. Αυτό το πλήθος αντιστοιχεί σε έναν μέσο ρυθμό 134.17 RP/h ανά χρήστη (ο Πίνακας 4.4 αναφέρει τον αριθμό των χρηστών και το χρόνο προσομοίωσης). Παρόλ' αυτά, για $P_a = 0.5$ και το ίδιο κατώφλι συμφόρησης, αποστέλλονται 31,182 RP (ή 112.26 RP/h ανά χρήστη). Τελικά, από την Εικόνα 4.3(a)-(d), παρατηρείται ότι υπάρχει μία ανταλλαγή μεταξύ της τιμής της απόκλισης που επιτυγχάνεται και του πλήθους των απεσταλμένων RP για διαφορετικές τιμές του κατωφλίου συμφόρησης. Για μικρές τιμές του κατωφλίου συμφόρησης, το σύστημα επιτυγχάνει πολύ μικρή απόκλιση, αλλά απαιτείται η αποστολή μεγάλου πλήθους RP, και το αντίστροφο.

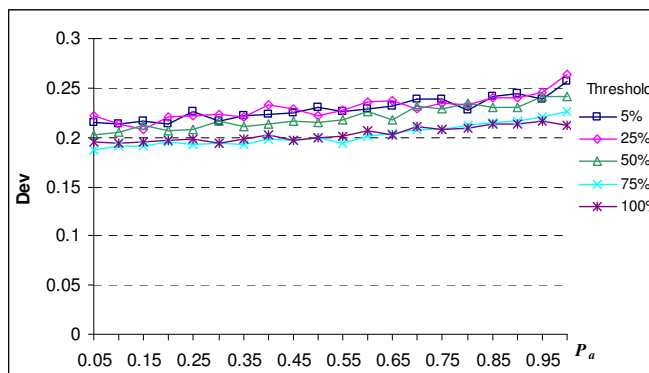
4.6.2 Αποτελέσματα Προσομοίωσης για το SFBP-MITOS

Στην Εικόνα 4.4, απεικονίζονται τα αποτελέσματα σχετικά με τη μετρική απόκλισης και το πλήθος των απεσταλμένων RP για το σύστημα SFBP-MITOS. Στην Εικόνα 4.4(b), παρατηρείται ότι σε αντίθεση με το μE -MITOS, η μετρική Dev είναι μάλλον αναισθητη στις μεταβολές της τιμής του κατωφλίου συμφόρησης και, καθώς αυξάνεται το P_a , η μετρική Dev επιδεικνύει αυξητική τάση. Αυτό οφείλεται στο ότι, όσο αυξάνεται ο αριθμός των χρηστών που είναι πρόθυμοι να μετακινηθούν (δηλαδή αυξάνεται το P_a), η συσχέτιση μεταξύ αυτών των χρηστών αυξάνεται (δηλαδή υπάρχει μία αυξανόμενη ομοιότητα στον «τρόπο σκέψης» τους [54]). Ως εκ τούτου, περισσότεροι χρήστες κάνουν πανομοιότυπες προβλέψεις και μετακινούνται στην ίδια θέση, δηλαδή οι χρήστες δεν κατανέμονται επαρκώς κι η

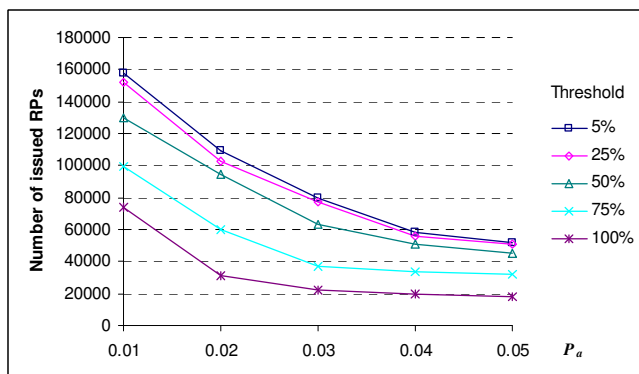
απόκλιση αυξάνεται. Πρέπει να σημειωθεί ότι η συσχέτιση των χρηστών δεν αποτελεί ζήτημα για το μE-MITOS, καθώς πιθανές συγκρούσεις χρηστών (κοινές προτιμήσεις) επιλύονται μέσω του σχήματος πολλαπλών δημοπρασιών. Τελικά, παρατηρείται από την Εικόνα 4.4(a)-(b) ότι οι τιμές της μετρικής Dev με το SFBP-MITOS είναι σημαντικά μειωμένες σε σχέση με το RP-free σενάριο (όπου $Dev \approx 1.05$).



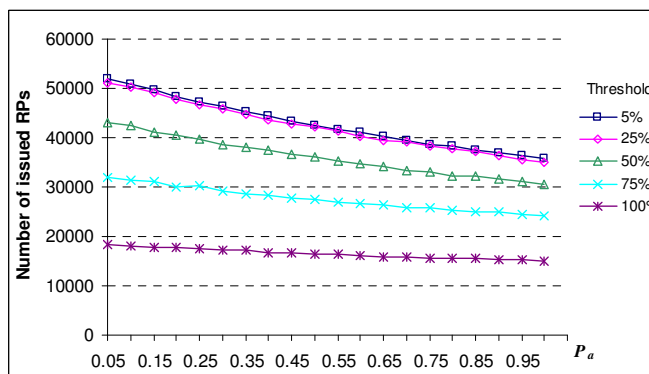
(a) Dev vs. P_a ($\in [0, 0.05]$)



(b) Dev vs. P_a ($\in [0.05, 1]$)



(c) πλήθος απεσταλμένων RP vs. P_a ($\in [0.01, 0.05]$)



(d) πλήθος απεσταλμένων RP vs. P_a ($\in [0.05, 1]$)

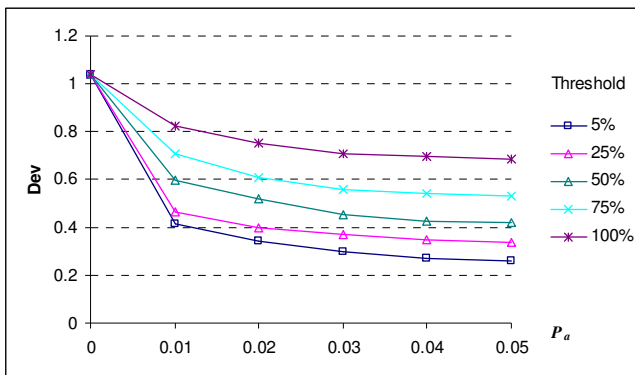
Εικόνα 4.4. SFBP-MITOS: (a)-(b) Dev vs. P_a και (c)-(d) πλήθος απεσταλμένων RP vs. P_a .

Η Εικόνα 4.4(c)-(d) δείχνει ότι το συνολικό πλήθος των απεσταλμένων RP μειώνεται καθώς αυξάνεται το P_a ή το κατώφλι συμφόρησης, όπως συμβαίνει και στην περίπτωση του μE-MITOS. Για παράδειγμα, για $P_a = 0.05$ και κατώφλι συμφόρησης 75%, αποστέλλονται συνολικά 32,006 RP (115.22 RP/h ανά χρήστη), ενώ για $P_a = 0.5$, αποστέλλονται 27,424 RP (ή 98.73 RP/h ανά χρήστη). Μπορεί επίσης να παρατηρηθεί ότι το SFBP-MITOS αποστέλλει λιγότερα RP σε σχέση με το μE-MITOS. Παρόλ' αυτά, το μE-MITOS επιτυγχάνει χαμηλότερες τιμές της μετρικής Dev , συγκριτικά με το SFBP-MITOS. Στην επόμενη παράγραφο,

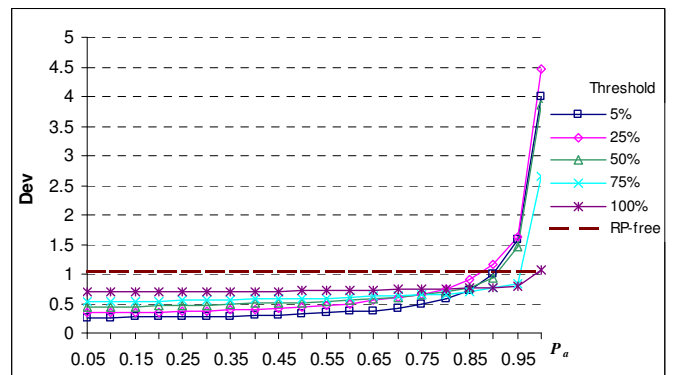
παρουσιάζονται αποτελέσματα από την προσομοίωση της απλής έκδοσης του MITOS (plain-vanilla MITOS – PV-MITOS), όπως αυτό παρουσιάζεται στο υποκεφάλαιο 4.5.

4.6.3 Αποτελέσματα Προσομοίωσης για το Plain-Vanilla MITOS

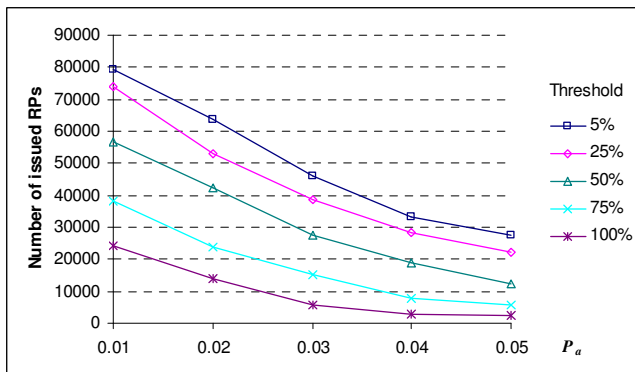
Πριν αναλυθούν περαιτέρω τα παραπάνω αποτελέσματα, είναι απαραίτητο να αξιολογηθεί και το σύστημα PV-MITOS. Η Εικόνα 4.5 δείχνει την απόκλιση που επιτυγχάνεται και το πλήθος των απεσταλμένων RP. Στην Εικόνα 4.5(b) παρατηρείται ότι το PV-MITOS παρουσιάζει ένα σημαντικό μειονέκτημα: καθώς αυξάνεται το P_a πάνω από 0.8, το Dev σκαρφαλώνει σε πολύ υψηλά επίπεδα. Συνεπώς, δεν αξίζει να υιοθετηθεί μία τέτοια προσέγγιση, καθώς επιδεικνύει υψηλότερη απόκλιση συγκριτικά με το RP-free σενάριο, λόγω των φαινομένων ταλαντώσεων που έχουν συζητηθεί στην παράγραφο 4.3.5.



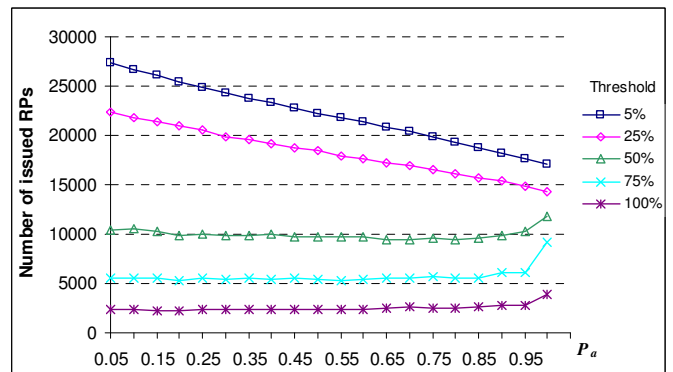
(a) Dev vs. P_a ($\in [0, 0.05]$)



(b) Dev vs. P_a ($\in [0.05, 1.0]$)



(c) πλήθος απεσταλμένων RP vs. P_a ($\in [0.01, 0.05]$)



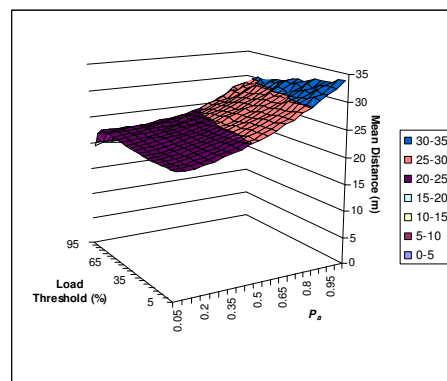
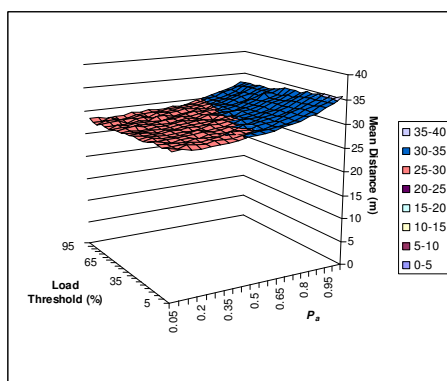
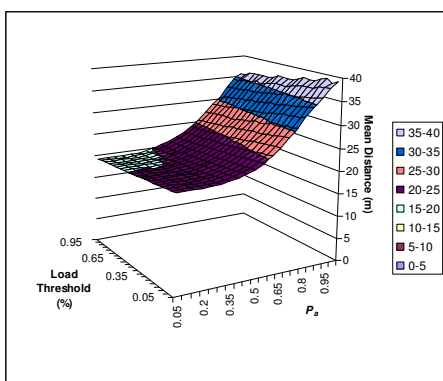
(d) πλήθος απεσταλμένων RP vs. P_a ($\in [0.05, 1.0]$)

Εικόνα 4.5. PV-MITOS: (a)-(b) Dev vs. P_a και (c)-(d) πλήθος απεσταλμένων RPs vs. P_a .

Στην Εικόνα 4.5(c)-(d), παρατηρείται ότι για μία δεδομένη τιμή της P_a , το πλήθος των απεσταλμένων RP μειώνεται καθώς αυξάνεται το κατώφλι συμφόρησης (κοινή παρατήρηση και για τα τρία μελετούμενα σχήματα). Αξίζει να σημειωθεί ότι ο μικρότερος αριθμός απεσταλμένων RP που επιτεύχθηκε με το PV-MITOS είναι 2,500 (ή 9 RP/h ανά χρήστη), ενώ οι άλλες δύο προσεγγίσεις αποστέλλουν περισσότερες από 15,000 RP (ή 54 RP/h ανά χρήστη).

4.6.4 Συζήτηση

Εκτός από την απόκλιση και το πλήθος των απεσταλμένων RP, μία αρκετά σημαντική ποσότητα αποτελεί κι η απόσταση μετακίνησης που χρειάζεται να διανύσει ένας χρήστης ακολουθώντας της οδηγίες ενός RP. Σύμφωνα με τα αποτελέσματα των προσομοιώσεων, κι οι τρεις προσεγγίσεις επιτυγχάνουν συγκρίσιμες επιδόσεις σε σχέση με τη μέση απόσταση μετακίνησης. Όπως φαίνεται στην Εικόνα 4.6, στην πλειοψηφία των περιπτώσεων, ο χρήστης χρειάζεται να μετακινηθεί σε ένα γειτονικό AP (απόσταση ~20m). Ένα μικρό μόνο ποσοστό μετακινήσεων ξεπερνά την απόσταση του ενός AP.



(a) Απόσταση μετακίνησης (μE-MITOS)

(b) Απόσταση μετακίνησης (SFBP-MITOS)

(c) Απόσταση μετακίνησης (PV-MITOS)

Εικόνα 4.6. Μέση απόσταση μετακίνησης για τα (a) μE-MITOS, (b) SFBP-MITOS και (c) PV-MITOS

Ένα σημαντικό ζήτημα, όπως συζητήθηκε στην παράγραφο 4.3.5, είναι το επιπλέον κόστος που μπορούν να επιφέρουν οι εκπομπές των RP στο ήδη κορεσμένο δίκτυο. Από τα αποτελέσματα της προσομοίωσης, μπορεί να ειπωθεί ότι το επιπλέον κόστος για το δίκτυο είναι μάλλον αμελητέο. Από την Εικόνα 4.4 (c)

μπορεί να παρατηρηθεί ότι το πλήθος των RP που αποστέλλει το MITOS στη διάρκεια της προσομοίωσης δεν μπορεί να ξεπεράσει τις 160,000, το οποίο αντιστοιχεί σε ένα ρυθμό εκπομπής 16 RP/s. Παρόλ' αυτά, το μέγεθος ενός RP είναι μάλλον περιορισμένο, καθώς η μόνη πληροφορία που είναι απαραίτητη είναι μία λίστα από θέσεις που απαρτίζουν το μονοπάτι από την τρέχουσα στην προτεινόμενη θέση. Τα αποτελέσματα στην Εικόνα 4.6 (η πλειοψηφία των μετακινήσεων γίνεται από ένα AP σε ένα από τα γειτονικά του) κι η διαρρύθμιση που χρησιμοποιήθηκε στις προσομοιώσεις (Εικόνα 4.3) υποδηλώνουν ότι, σε επίπεδο θέσεων, οι προτεινόμενες μετακινήσεις μπορούν να αναπαρασταθούν από μία λίστα από όχι περισσότερες από τέσσερις θέσεις. Εκτιμάται ότι ένα RP με μέγεθος 128 bytes είναι ικανό να μεταφέρει αυτήν την πληροφορία· αυτό αντιστοιχεί σε έναν εκτιμώμενο ρυθμό 16 Kb/s για την εκπομπή των RP για το δίκτυο στο σύνολο του. Το επιπλέον κόστος αυτό είναι πολύ μικρό σε σχέση με τη συνολική χωρητικότητα των AP (στις προσομοιώσεις: 11 Mb/s ανά AP) και τα χαρακτηριστικά των εφαρμογών των χρηστών (Πίνακας 4.3), επομένως η συμβολή του στη συμφόρηση του δικτύου είναι μάλλον αμελητέα.

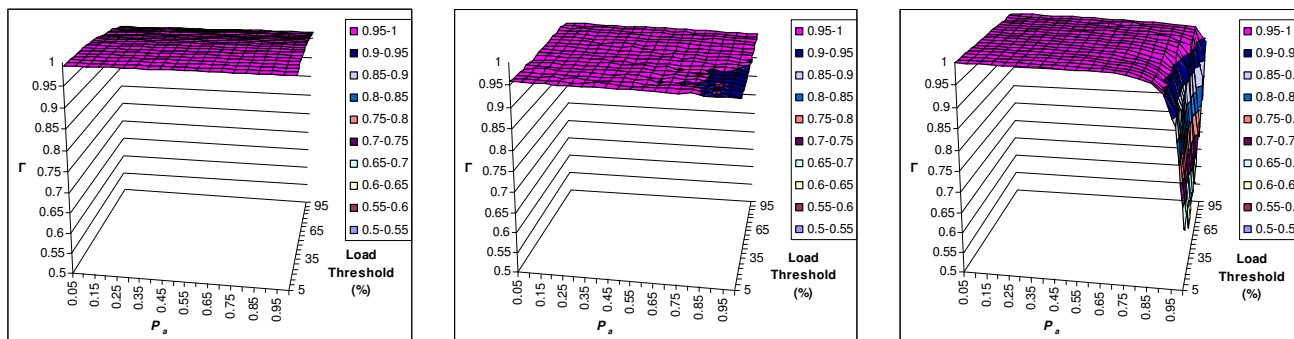
Ένα ακόμα σημαντικό μέτρο για την αξιολόγηση της απόδοσης των προτεινόμενων σχημάτων είναι η ρυθμοαπόδοση που επιτυγχάνεται, η οποία αντιπροσωπεύεται εδώ από τη μετρική Γ που ορίζεται παρακάτω. Αν BW η χωρητικότητα ενός AP, στη διάρκεια της n -στης χρονοθυρίδας, το AP μεταφέρει δεδομένα με ρυθμό $\min\{D_i(n), BW\}$, όπου $D_i(n)$ είναι ο προσφερόμενος φόρτος στο AP i , $i = 1, 2, \dots, N$, κατά τη n -στη χρονοθυρίδα. Η μετρική Γ που απεικονίζει τη ρυθμοαπόδοση του συστήματος ορίζεται ως το ποσοστό της συνολικά προσφερόμενης κίνησης προς όλα τα N AP, σε όλη τη διάρκεια της προσομοίωσης H , που το σύστημα κατάφερε να διεκπεραιώσει:

$$\Gamma = \frac{\sum_{n=1}^H \sum_{i=1}^N \min\{D_i(n), BW\}}{\sum_{n=1}^H \sum_{i=1}^N D_i(n)} \quad (4.10)$$

Εξ ορισμού, $0 \leq \Gamma \leq 1$. $\Gamma = 1$ υποδηλώνει μηδενική απώλεια δεδομένων.

Με το μοντέλο προσομοίωσης που υιοθετήθηκε, η τιμή της μετρικής Γ για το RP-free σενάριο ήταν προσεγγιστικά 0.86, δηλαδή το δίκτυο χωρίς καμία υποστήριξη

από το σύστημα MITOS επέτυχε να μεταφέρει επιτυχώς το 86% της προσφερόμενης κίνησης.



(a) Μετρική Γ (μ E-MITOS) (b) Μετρική Γ (SFBP-MITOS) (c) Μετρική Γ (PV-MITOS)

Εικόνα 4.7. Οι τιμές της μετρικής Γ για τα (a) μ E-MITOS, (b) SFBP-MITOS και (c) PV-MITOS

Η Εικόνα 4.7 απεικονίζει τις τιμές του Γ για τις διαφορετικές προσεγγίσεις του MITOS. Στην Εικόνα 4.7(a), παρατηρείται ότι το μ E-MITOS είναι ικανό να μεταφέρει σχεδόν όλη την προσφερόμενη κίνηση, καθώς η τιμή του Γ είναι σχεδόν ίση με 1, για κάθε τιμή του P_a και του κατωφλίου συμφόρησης. Στην Εικόνα 4.7(b), παρατηρείται ότι η τιμή του Γ για το SFBP-MITOS είναι σχεδόν σταθερή, ανεξάρτητα από την τιμή της P_a και του κατωφλίου συμφόρησης. Παρόλ' αυτά, οι τιμές του Γ που έχουν επιτευχθεί (~0.95) είναι λίγο υποδεέστερες αυτών που επιτεύχθηκαν με το μ E-MITOS. Η Εικόνα 4.7 (c) δείχνει τις τιμές του Γ για το PV-MITOS. Όπως ήταν αναμενόμενο, για υψηλές τιμές του P_a , η τιμή του Γ μειώνεται δραματικά, καθώς οι χρήστες δεν κατανέμονται επαρκώς (περισσότερο μάλλον ταλαντώνονται από το ένα AP σε κάποιο άλλο).

Τα παραπάνω αποτελέσματα δείχνουν ότι το μ E-MITOS και το SFBP-MITOS, λόγω της ικανότητάς τους να εξισορροπούν την προσφερόμενη κίνηση στα υπάρχοντα AP, μπορούν να προσφέρουν σημαντική βελτίωση στη ρυθμοαπόδοση σε σύγκριση με το RP-free σενάριο. Πιο κάτω, συνοψίζονται τα πλεονεκτήματα και τα μειονεκτήματα των τριών παραλλαγών του MITOS.

4.6.4.1 Πλεονεκτήματα του μ E-MITOS

Τα αποτελέσματα των προσομοιώσεων έδειξαν ότι το μ E-MITOS επιδεικνύει τα χαμηλότερα επίπεδα απόκλισης κι επιτυγχάνει μία σημαντική βελτίωση σε σχέση

με το RP-free σενάριο, καθώς το δίκτυο πετυχαίνει να μεταφέρει σχεδόν το σύνολο της προσφερόμενης κίνησης, ανεξάρτητα από την πιθανότητα αποδοχής του RP και το κατώφλι συμφόρησης. Επιπλέον, καθώς βασίζεται σε ένα σχήμα πολλαπλών δημοπρασιών, υπάρχει εγγύηση ότι η σύγκρουση μεταξύ των πρακτόρων που αναζητούν υπηρεσία υψηλού επιπέδου επιλύεται, όσο οι τιμές κι ο φόρτος ενός δημοφιλούς AP αυξάνουν σταδιακά, στη διάρκεια της διαδικασίας των επαναλαμβανόμενων δημοπρασιών. Έτσι αυτή η προσέγγιση είναι ικανή να παράγει εξισορροπημένες κατανομές φόρτου, λαμβάνοντας επίσης υπόψη τη συνεισφορά στο φόρτο του κάθε χρήστη ξεχωριστά.

4.6.4.2 Μειονεκτήματα του μE -MITOS

Παρόλο που η προσέγγιση αυτή επιτυγχάνει τα χαμηλότερα επίπεδα απόκλισης, αποστέλλει το μεγαλύτερο πλήθος από RP. Επιπλέον, οι μικροοικονομικοί αλγόριθμοι απαιτούν σημαντική αλληλεπίδραση μεταξύ των πρακτόρων για 1) την ομαδοποίηση τους, βάσει των προτιμήσεων τους, 2) τη δημιουργία των προσφορών και 3) τον καθορισμό των νικητών, για όλες τις απαιτούμενες δημοπρασίες μέχρις ότου όλοι οι χρήστες να έχουν κερδίσει. Αυτοί οι αλγόριθμοι καταναλώνουν αρκετούς υπολογιστικούς πόρους κι είναι ιδιαίτερα πολύπλοκοι. Σε ό,τι αφορά τις αποστάσεις μετακίνησης, για $P_a > 0.5$, οι μετακινήσεις των χρηστών για να φτάσουν στην προτεινόμενη θέση μπορεί να ξεπεράσει το ένα AP.

4.6.4.3 Πλεονεκτήματα του SFBP-MITOS

Το πλεονέκτημα του SFBP-MITOS είναι ότι οι πράκτορες μπορούν να αποφασίζουν αυτόνομα, δηλαδή δε χρειάζεται κανένας συντονισμός των πρακτόρων. Επιπλέον, οι προγνώστες είναι συστατικά απλά στην υλοποίησή τους, αλλά και στη χρήση τους, έχοντας ελάχιστες υπολογιστικές απαιτήσεις. Όπως παρατηρείται από τα αποτελέσματα των προσομοιώσεων, το σχήμα που βασίζεται στο SFBP μπορεί να επιτύχει σημαντικά μικρότερη απόκλιση κι υψηλότερη ρυθμοαπόδοση, σε σύγκριση με το RP-free σενάριο. Επιπλέον, το πλήθος των απεσταλμένων RP είναι πιο μικρό από αυτό του μE -MITOS. Με το SFBP-MITOS, οι αποστάσεις μετακινήσεων προς τις προτεινόμενες θέσεις φαίνονται να μην

επηρεάζονται από την πιθανότητα αποδοχής ενός RP και το κατώφλι συμφόρησης. Στις περισσότερες περιπτώσεις, οι χρήστες μετακινούνται σε γειτονικά AP.

4.6.4.4 Μειονεκτήματα του SFBP-MITOS

Παρόλο που το συγκεκριμένο σχήμα επιδεικνύει μία σταθεροποιητική συμπεριφορά, η στοχαστική του φύση δεν μπορεί να εξασφαλίσει την εξισορρόπηση φόρτου, σε αντίθεση με το μ E-MITOS. Η επιτυχία της προσέγγισης βασίζεται σε μεγάλο βαθμό στην ετερογένεια των πρακτόρων. Αν οι πράκτορες έχουν παρόμοιο τρόπο σκέψης, θα έχουν και παρόμοια συμπεριφορά, δηλαδή θα συμπεριφέρονται σαν ομάδες πρακτόρων, παρά ως ατομικές οντότητες. Με δεδομένο ένα σύνολο από διαθέσιμους προγνώστες, όσο αυξάνεται ο πληθυσμός των πρακτόρων, τόσο υπονομεύεται κι η ετερογένεια τους, με κίνδυνο την πρόκληση ταλαντώσεων. Αυτή η τάση αποκαλύπτεται και στην Εικόνα 4.4(b), όπου όσο αυξάνεται το P_a , τόσο αυξάνεται και το πλήθος των πρακτόρων που λαμβάνουν κοινές αποφάσεις, το οποίο έχει ως αποτέλεσμα να αυξάνεται η απόκλιση. Για να ελαχιστοποιηθεί αυτό το φαινόμενο, απαιτείται ένα πλούσιο σύνολο από διαθέσιμους προγνώστες.

4.6.4.5 Πλεονεκτήματα του Plain-Vanilla MITOS

Το PV-MITOS δεν επιτυγχάνει μικρότερη απόκλιση από το μ E-MITOS, ούτε την αποστολή λιγότερων RP, σε σχέση με το SFBP-MITOS. Παρόλ' αυτά, το βασικότερο του πλεονέκτημα είναι η απλότητα του: δεν χρειάζεται προγνώστες, ή πράκτορες που να συμμετέχουν σε δημοπρασίες. Το πρόβλημα που απεικονίζεται στην Εικόνα 4.5(a), δηλαδή η ταχεία αύξηση της απόκλισης (κι η αντίστοιχη μείωση της ρυθμοαπόδοσης που φαίνεται στην Εικόνα 4.7 (c)) θα μπορούσαν να θεραπευθούν μειώνοντας το ρυθμό εκπομπής των RP. Αποφεύγοντας την ειδοποίηση όλων των χρηστών (όταν $P_a \rightarrow 1$), η πιθανότητα αποδοχής των RP μπορεί να μειωθεί εικονικά· έτσι, η ανεπιθύμητη αυτή κατάσταση μπορεί να προληφθεί. Σε ό,τι αφορά την απόσταση μετακινήσεων των χρηστών προς τις προτεινόμενες θέσεις, το PV-MITOS επιδεικνύει μέτριες επιδόσεις.

4.6.5 Σύνοψη Σύγκρισης των Προτεινόμενων Σχημάτων του MITOS

Όπως φάνηκε παραπάνω, κάθε προσέγγιση έχει συγκεκριμένα πλεονεκτήματα και μειονεκτήματα. Αν το βασικό ζήτημα είναι να διατηρηθεί το σύστημα όσο πιο απλό γίνεται, θα πρέπει να επιλεγεί το PV-MITOS. Αν ο πάροχος του WLAN επιθυμεί να διατηρήσει το φόρτο όσο το δυνατόν πιο εξισορροπημένο (ώστε να μεγιστοποιηθεί η ρυθμοαπόδοση του συστήματος), θα πρέπει να χρησιμοποιηθεί το μ E-MITOS. Τελικά, αν ο πάροχος του WLAN στοχεύει στο να ελαχιστοποιήσει το ρυθμό των απεσταλμένων RP και να διατηρήσει το φόρτο του δικτύου σταθερό κι εξισορροπημένο, θα πρέπει να υιοθετηθεί το SFBP-MITOS. Ο Πίνακας 4.5 συνοψίζει τα συμπεράσματα που αφορούν στη σύγκριση των τριών εναλλακτικών σχημάτων που προτάθηκαν.

Πίνακας 4.5. Σύγκριση των προτεινόμενων σχημάτων

Σύστημα / Μετρική	PV-MITOS		SFBP	μ E-MITOS	RP-free
	Υψηλή P_a	Μικρή/μέτρια P_a			
Εξισορρόπηση φόρτου (<i>Dev</i>)	---	+	++	+++	--
Ρυθμοαπόδοση (Γ)					
Απεσταλμένα RP	+	+	++	-	NA
Απλότητα	++	++	+	---	+++
Απόσταση Μετακίνησης	+	++	+++	+	NA
<i>+++</i> : Εξαιρετική επίδοση, <i>---</i> : Πενιχρή επίδοση					

4.7 Σύγκριση του MITOS με Άλλα Συστήματα

Για την επιπλέον αξιολόγηση των ικανοτήτων εξισορρόπησης φόρτου του MITOS, οι προτεινόμενες λύσεις συγκρίθηκαν με την προσέγγιση που προτείνεται στην εργασία [42], η οποία θεωρείται ως η σχετικότερη με την τρέχουσα μελέτη. Στην εργασία [42], στην περίπτωση τοπικής συμφόρησης, το τερματικό βρίσκει ένα λιγότερο φορτωμένο AP στην γειτονιά του για να συνδεθεί, ανταλλάσσοντας μεταξύ ποιότητας σήματος και διαθέσιμου εύρους ζώνης (τεχνική channel switching). Αν δεν υπάρχει AP που να μπορεί να εγγυηθεί βελτίωση στη σύνδεση, ένας

διακομιστής (network-monitoring server) παρέχει πληροφορίες στο χρήστη υποδεικνύοντας ένα λιγότερο φορτωμένο, αλλά μακρινό, AP (τεχνική network directed roaming). Στην εργασία [42], υιοθετήθηκε μία μετρική που ονομάζεται *δείκτης εξισορρόπησης* (balance index), $B(n)$, ο οποίος σχετίζεται με το *συντελεστή διασποράς*, $CoV(n)$, (που έχει υιοθετηθεί στην παρούσα μελέτη) ως εξής:

$$(CoV(n))^2 = \frac{1}{B(n)} - 1 \quad (4.11)$$

Στην παρούσα μελέτη, δεν αξιολογούνται οι στιγμιαίες τιμές της μετρικής $CoV(n)$, αλλά η μέση χρονική τιμή Dev (όπως ορίζεται στην εξίσωση (4.9)). Στον Πίνακα 4.6, παρουσιάζονται οι τιμές της μετρικής Dev , για κάποιες ενδεικτικές τιμές του P_a και για κατώφλι συμφόρησης 75%.

Πίνακας 4.6. Τιμές της μετρικής απόκλισης (Dev) για διαφορετικά επίπεδα της πιθανότητας P_a για κατώφλι συμφόρησης = 75%

Σύστημα	P_a			
	0.05	0.5	0.95	1.0
μE-MITOS	0.174	0.156	0.146	0.144
SFBP-MITOS	0.188	0.199	0.221	0.226
PV-MITOS	0.535	0.589	0.848	2.640
RP-free	1.052			
Network Directed Roaming	~0.333			
Channel Switching	~0.538			
No Load Balancing	~0.818			

Στις τελευταίες τρεις γραμμές, ο Πίνακας 4.6 παρουσιάζει οι τιμές της μετρικής Dev για το σύστημα που προτείνεται στην εργασία [42], όπως έχουν αναφερθεί εκεί. Παρατηρείται ότι η τιμή του Dev για το σύστημα χωρίς εξισορρόπηση φόρτου (“no load-balancing”), που είναι ισοδύναμο με το αντίστοιχο RP-free σύστημα στην παρούσα μελέτη, ήταν περίπου 0.818, δηλαδή μικρότερο από την αντίστοιχη τιμή του Dev για το RP-free σενάριο (1.052). Αυτό σημαίνει ότι το περιβάλλον προσομοίωσης της παρούσας μελέτης οδηγούσε σε λιγότερο εξισορροπημένες κατανομές φόρτου σταθερής κατάστασης (steady-state distribution), σε σχέση με το αντίστοιχο της εργασίας [42]. Παρόλα αυτά, όπως δείχνει ο Πίνακας 4.6, το μE-

MITOS και το SFBP-MITOS υπερτερούν του σχήματος “network directed roaming” στην [42], παρόλο που τα χαρακτηριστικά κινητικότητας των χρηστών στην προσομοίωση της παρούσας μελέτης είναι μάλλον δυσμενή για τα αξιολογούμενα συστήματα MITOS. Επίσης, παρατηρείται ότι το σχήμα “channel switching” και το PV-MITOS επιτυγχάνουν συγκρίσιμες επιδόσεις για μέτριες πιθανότητες αποδοχής των RP.

4.8 Συμπεράσματα

Σε αυτό το κεφάλαιο, περιγράφηκε η υλοποίηση του MITOS, ενός συστήματος ευφυών χώρων, που μπορεί να λειτουργήσει πάνω από υπάρχουσες υποδομές WLAN. Ο κύριος σκοπός του MITOS είναι να επιτρέψει στο χρήστη να διατηρήσει το επίπεδο της ποιότητας στις υπηρεσίες του. Αυτός ο σκοπός επιτυγχάνεται με τη φυσική μετακίνηση των κινητών χρηστών, βάσει οδηγιών που αποστέλλονται από το σύστημα. Το MITOS ενσωματώνει την ευφυΐα που απαιτείται για την αποδοτική χρήση των ασύρματων πόρων με εξισορρόπηση της ζήτησης στο δίκτυο. Αξιολογήθηκε και συγκρίθηκε η απόδοση τριών παραλλαγών του MITOS, του απλού συστήματος (PV-MITOS), της μικροοικονομικής επέκτασης του (μE-MITOS), καθώς και της παιγνιοθεωρητικής του επέκτασης (SFBP-MITOS), μέσω προσομοιώσεων. Αξιολογήθηκε, επίσης, κι η θετική επίδραση της αποστολής προτάσεων μετακίνησης σε σχέση με ένα σύστημα χωρίς προτάσεις, δηλαδή μία κατάσταση όπου οι χρήστες μετακινούνται στην επικράτεια του δικτύου χωρίς επεμβάσεις από το σύστημα. Τα αποτελέσματα από τις προσομοιώσεις έδειξαν ότι και τα τρία συστήματα συνεισφέρουν σημαντικά στην αποσυμφόρηση των hot-spot, το οποίο έχει ως αποτέλεσμα τη βελτίωση της εμπειρίας του χρήστη, καθώς το δίκτυο επιτυγχάνει να αυξήσει σημαντικά τη ρυθμοαπόδοση του.

Το μE-MITOS επιτυγχάνει τη μικρότερη απόκλιση (καλύτερη εξισορρόπηση) και τη μέγιστη ρυθμοαπόδοση για το δίκτυο, σε σχέση με το PV-MITOS και το SFBP-MITOS. Παρόλ' αυτά, είναι το πιο πολύπλοκο κι αποστέλλει το μεγαλύτερο αριθμό από RP. Το SFBP-MITOS, επίσης, αποδεικνύεται ιδιαίτερα αποδοτικό, καθώς η προκύπτουσα με αυτό κατανομή φόρτου στο δίκτυο είναι αρκετά εξισορροπημένη (σχεδόν ίσα φορτωμένα AP), σε αντίθεση με το RP-free σενάριο. Φάνηκε επίσης

ότι το SFBP-MITOS αποστέλλει τον μικρότερο αριθμό από RP σε σύγκριση με τις άλλες δύο προσεγγίσεις.

Για την αξιολόγηση της επίδρασης του συστήματος στους κινητούς χρήστες, μετρήθηκε επίσης κι η απόσταση που πρέπει να διανύσει ο κινητός χρήστης, αν αποφασίσει να ακολουθήσει την προτροπή του συστήματος. Σε όλα σχεδόν τα μελετούμενα σενάρια, οι χρήστες χρειάστηκαν να μετακινηθούν σε ένα γειτονικό τους AP και σε λίγες περιπτώσεις σε ένα λίγο πιο μακρινό. Τέλος, το MITOS συγκρίθηκε και με υπάρχοντα συστήματα κι αποδείχθηκε πιο αποδοτικό σε ό,τι αφορά την εξισορρόπηση του φόρτου στο WLAN.

Μελλοντικά, θα ήταν επιθυμητό να ενσωματωθεί στην αρχιτεκτονική του MITOS ένα συστατικό ικανό να υπολογίζει την πιθανότητα αποδοχής των RP, παρατηρώντας τη συμπεριφορά των χρηστών. Μέσω αυτού του συστατικού, το σύστημα θα είναι ικανό να αυτό-ρυθμίζεται, στέλνοντας RP στοχαστικά, βάσει της υπολογισμένης πιθανότητας αποδοχής των RP. Έτσι, η συχνότητα αποστολής των RP θα μπορέσει να ελαχιστοποιηθεί. Επιπλέον, θα ήταν επιθυμητό να αξιολογηθεί το MITOS με βάση πραγματικούς χρήστες σε μία πραγματική υποδομή WLAN. Επίσης, θα ήταν επιθυμητό να μελετηθεί το MITOS σε ένα εμπορικό περιβάλλον (π.χ. όπου οι χρήστες αγοράζουν χρόνο σύνδεσης). Σε ένα τέτοιο περιβάλλον, το MITOS θα μπορούσε να συνδυαστεί με το σύστημα χρέωσης της υποδομής, ούτως ώστε οι χρήστες να παρακινούνται να αλλάξουν θέση με τη βοήθεια προσφορών (π.χ., επέκταση του υπολειπόμενου χρόνου σύνδεσης), όταν οι χρήστες δεν είναι ιδιαίτερα πρόθυμοι να αλλάξουν τη θέση τους. Ένα τέτοιο κίνητρο θα μπορούσε να αυξήσει την πιθανότητα αποδοχής των RP και τελικά να βελτιώσει την κατανομή του φόρτου στο δίκτυο. Τέλος, θα ήταν επιθυμητό να εμπλουτιστεί το MITOS με πιο λεπτομερείς μηχανισμούς παρακολούθησης του φόρτου (π.χ. παρακολούθηση του φόρτου στο ίδιο το τερματικό του χρήστη) που θα βελτιώνε την ποιότητα των πληροφοριών που συλλέγει το σύστημα και τελικά τις επιδόσεις του.

ΚΕΦΑΛΑΙΟ 5

ΠΡΟΔΡΑΣΤΙΚΗ ΔΙΑΧΕΙΡΙΣΗ ΕΥΡΟΥΣ ΖΩΝΗΣ ΣΕ ΚΥΨΕΛΩΤΑ ΔΙΚΤΥΑ ΜΕΣΩ ΔΙΑΠΡΑΓΜΑΤΕΥΣΕΩΝ

5.1 Εισαγωγή

Στο παρόν κεφάλαιο, παρουσιάζεται η μελέτη σχετικά με την προδραστική διαχείριση πόρων σε ασύρματα κυψελωτά δίκτυα. Συγκεκριμένα, μελετάται το πρόβλημα των μεταπομπών και μεθόδων που μπορούν να εφαρμοστούν για την ελαχιστοποίηση των συνεπειών τους. Λόγω μεταπομπής, μία εφαρμογή μπορεί να χρειαστεί να τερματιστεί, αν ο νέος ΣΒ δεν έχει επαρκείς πόρους για να υποστηρίξει τις απαιτήσεις ποιότητας υπηρεσίας της εφαρμογής. Αυτού του τύπου ο τερματισμός είναι ιδιαίτερα ενοχλητικός για το χρήστη κι αναφέρεται ως μπλοκάρισμα μεταπομπής [60].

Η πιθανότητα μπλοκαρίσματος από μεταπομπή μπορεί να περιοριστεί μέσω προδραστικών δεσμεύσεων πόρων στις γειτονικές κυψέλες του χρήστη [64]. Προσφάτως, προδραστικά σχήματα διαχείρισης πόρων, τα οποία εμπλέκουν πρόβλεψη κίνησης, έχουν υιοθετηθεί για περισσότερο ακριβείς δεσμεύσεις πόρων. Μετά από μία μεταπομπή, το *κινητό τερματικό* (ΚΤ) μπορεί να συνεχίσει κανονικά την εκτέλεση των εφαρμογών του, χωρίς να χρειάζεται να ανταγωνιστεί για τους αναγκαίους πόρους.

Το σχήμα που προτείνεται στο παρόν κεφάλαιο βασίζεται σε ένα μοντέλο χρέωσης για την προδραστική δέσμευση πόρων. Ο χρήστης διαπραγματεύεται με το δίκτυο, ούτως ώστε να προδεσμεύσει μία ποσότητα πόρων στην πιο πιθανή προς επίσκεψη επόμενη κυψέλη. Το δίκτυο πληρώνεται για αυτήν την υπηρεσία, ούτως ώστε να αντισταθμίσει το ότι αυτοί οι πόροι δε μπορούν να διατεθούν αλλού για κάποιο χρόνο. Το ποσό που πληρώνει ο χρήστης καθορίζεται από μία διαδικασία *διαπραγμάτευσης* που εξασφαλίζει την ιδιωτικότητα των *αξιολογήσεων* των ζητούμενων πόρων, και για το χρήστη και για το δίκτυο. Η ιδιωτικότητα των

αξιολογήσεων υιοθετείται, καθώς ο χρήστης μπορεί να λειτουργεί σε μη-έμπιστα δίκτυα.

5.2 Σχετική Βιβλιογραφία

Στο παρελθόν, δε γινόταν διάκριση μεταξύ των πιθανοτήτων μπλοκαρίσματος νέας κλήσης και μπλοκαρίσματος μεταπομπής. Η ανάλυση στην εργασία [60], όμως, έδειξε ότι αυτές οι δύο πιθανότητες είναι διαφορετικές κι ότι το μπλοκάρισμα από μεταπομπή αποτελεί έναν από τους βασικότερους λόγους για τη δυσαρέσκεια των χρηστών στα κυψελωτά δίκτυα. Η μείωση της πιθανότητας μπλοκαρίσματος από μεταπομπή μπορεί κυρίως να επιτευχθεί με τη χρήση πόρων ασφαλείας, που διατηρούνται από το σύστημα αποκλειστικά για κλήσεις από μεταπομπή. Στην εργασία [61], προτάθηκε ένας απλός, προσαρμοστικός αλγόριθμος *ελέγχου αποδοχής κλήσης*, που χρησιμοποιεί κανάλια ασφαλείας. Άλλα προσαρμοστικά σχήματα δέσμευσης κι αποδοχής έχουν μελετηθεί στις εργασίες [62] και [63].

Στην εργασία [64], ο χρόνος διαιρείται σε ίσα διαστήματα, κι υπολογίζεται η πιθανότητα κάθε ΚΤ να βρεθεί σε κάθε μία από τις κυψέλες στο “shadow cluster” (το σύνολο των ΣΒ στους οποίους ίσως συνδεθεί το ΚΤ στο προσεχές μέλλον) για τα μελλοντικά χρονικά διαστήματα. Οι ΣΒ ανταλλάσσουν πληροφορίες σχετικά με την προβλεπόμενη ζήτηση εύρους ζώνης για τα μελλοντικά χρονικά διαστήματα, ώστε να καθορίσουν το αν είναι εφικτό να αποδεχθούν νέες κλήσεις. Στην εργασία [68], μία προδραστική δέσμευση, για ένα συγκεκριμένο συνδυασμό κυψέλης-χρήστη, διατηρείται έγκυρη για ένα περιορισμένο διάστημα που ξεκινά τη στιγμή *Earliest Arrival Time* και τελειώνει τη στιγμή *Latest Arrival Time*. Οι δύο χρόνοι υπολογίζονται γεωμετρικά βάσει της κυψελωτής τοπολογίας του δικτύου. Υπάρχουν, επίσης, κι απλούστερα σχήματα που θεωρούν σταθερή ποσότητα εύρους ζώνης ασφαλείας [65], [66], τα οποία, όμως, είναι αποτελεσματικά μόνο υπό στατικές συνθήκες φόρτου και, συνήθως, οδηγούν σε ελλιπή χρήση του φάσματος.

Από τα παραπάνω, είναι εμφανές ότι η διαχείριση των πόρων ασφαλείας σε κυψελωτά δίκτυα έχει προσελκύσει το ενδιαφέρον των ερευνητών. Παρόλ’ αυτά,

απ' όσο είναι γνωστό, το εν λόγω πρόβλημα δεν έχει ακόμα μελετηθεί υπό το πρίσμα της Θεωρίας Παιγνίων κι αρχών χρέωσης.

5.3 Αξιολόγηση των Πόρων από το ΚΤ και το ΣΒ

Σε αυτό το υποκεφάλαιο, συζητάται πως τα ΚΤ κι οι ΣΒ αξιολογούν τους δικτυακούς πόρους για τις εξεταζόμενες προδραστικές δεσμεύσεις πόρων. Έστω p_i (\$/(Kb/s)) η τιμή ανά μονάδα του πόρου που συμφωνείται μεταξύ του ΚΤ και του ΣΒ για το στιγμιότυπο εφαρμογής i (για λεπτομέρειες σχετικά με τον καθορισμό του p_i δες το υποκεφάλαιο 5.4). Η συνάρτηση ωφελείας του ΚΤ δίδεται στην εξίσωση (5.1), ενώ του ΣΒ στην εξίσωση (5.2).

$$U_{MT} = Benefit_{MT,i} - p_i \quad (5.1)$$

$$U_{BS} = p_i - Cost_{BS} \quad (5.2)$$

Στην εξίσωση (5.1), το $Benefit_{MT,i}$ υποδηλώνει την αξία για το ΚΤ μίας μονάδας εύρους ζώνης. Η συγκεκριμένη εκτίμηση εξαρτάται από δύο παράγοντες: 1) το βάρος, $f \in [0,1]$, διαφορετικό για κάθε χρήστη, που υποδηλώνει την κρισιμότητα της εφαρμογής (μπορεί να προκύψει από ένα προφίλ χρήστη), και 2) την παρατηρημένη διάρκεια της εφαρμογής μέχρι την τρέχουσα χρονική στιγμή, d_i . Γίνεται η υπόθεση ότι μία εφαρμογή, που έχει ήδη διαρκέσει πολύ, θα συνεχιστεί και στο μέλλον, στην οποία περίπτωση το εύρος ζώνης που απαιτεί η συγκεκριμένη σύνοδος θεωρείται μεγάλης σημασίας. Αυτή η υπόθεση συνήθως υιοθετείται, με καλά αποτελέσματα, στη μελέτη των WWW cache [70].

$$Benefit_{MT,i} = f \cdot \frac{d_i}{d_{max}} \quad (5.3)$$

Στην εξίσωση (5.3), το d_{max} υποδηλώνει τη μεγαλύτερη παρατηρημένη από το ΚΤ διάρκεια εφαρμογής και χρησιμοποιείται για λόγους κανονικοποίησης. Για αξιολογήσεις στο διάστημα $[b_{min}, b_{max}]$ (δες υποκεφάλαιο 5.4), το $Benefit_{MT}$ μπορεί να μετασχηματιστεί γραμμικά για να πάρει τιμές στο επιθυμητό πεδίο, δηλαδή $Benefit'_{MT} = b_{min} + (b_{max} - b_{min}) \cdot Benefit_{MT}$.

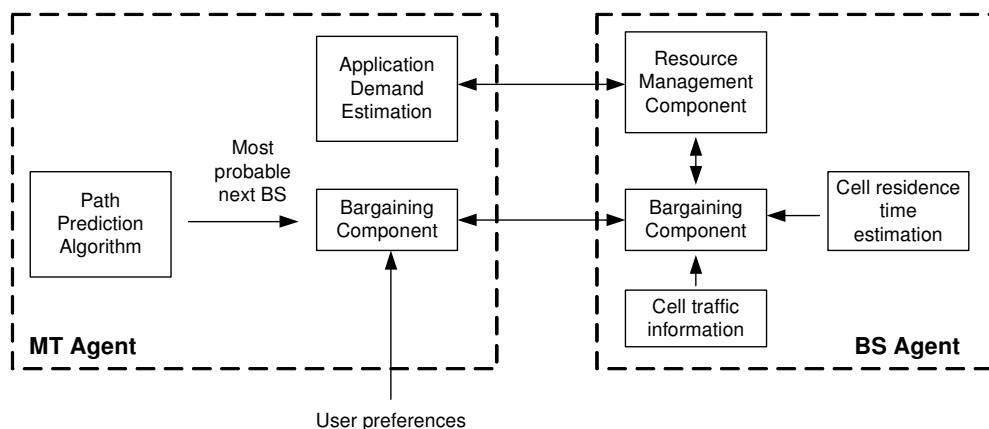
Στην εξίσωση (5.2), το $Cost_{BS}$ αντιπροσωπεύει την εκτίμηση του κόστους από το ΣΒ για να κάνει μία κράτηση. Αυτή η εκτίμηση είναι ανάλογη της χρησιμοποίησης των τοπικών πόρων, όπως φαίνεται στην εξίσωση (5.4).

$$Cost_{BS} = \frac{C_{Total} - C_{Free}}{C_{Total}} \quad (5.4)$$

Το C_{Free} υποδηλώνει την τρέχουσα ποσότητα ελεύθερων πόρων, ενώ το C_{Total} τη συνολική ποσότητα πόρων στο ΣΒ. Για εκτιμήσεις κόστους στο διάστημα $[s_{min}, s_{max}]$ (δες υποκεφάλαιο 5.4), το $Cost_{BS}$ μπορεί να μετασχηματιστεί κατάλληλα, δηλαδή $Cost'_{BS} = s_{min} + (s_{max} - s_{min}) \cdot Cost_{BS}$.

5.4 Προτεινόμενη Αρχιτεκτονική

Η προτεινόμενη αρχιτεκτονική βασίζεται σε τρεις βασικούς μηχανισμούς: 1) έναν αλγόριθμο πρόβλεψης κίνησης (ΑΠΚ), 2) ένα σχήμα διαπραγματεύσεων και 3) ένα πλαίσιο διαχείρισης πόρων. Κάθε ΣΒ, διατηρώντας κατάλληλα στατιστικά και εκτιμήσεις της συνολικής παρατηρούμενης κινητικότητας των χρηστών και του προσφερόμενου φόρτου, ενημερώνει την ελάχιστη αποδεκτή τιμή ανά μονάδα εύρους ζώνης, $Cost_{BS}$, που θα πρέπει να πληρώσει ένα ΚΤ για μία προδέσμευση, όπως έχει αναφερθεί και στο υποκεφάλαιο 5.3.



Εικόνα 5.1. Λειτουργία του ΚΤ και του ΣΒ

Όπως φαίνεται στην Εικόνα 5.1, ο πράκτορας του ΚΤ χρησιμοποιεί την έξοδο του ΑΠΚ⁶, ούτως ώστε να προσδιορίσει το ΣΒ που είναι ο πιο πιθανός για την επόμενη μεταπομπή. Αφού αξιολογήσει τον πόρο (δηλαδή υπολογίσει το $Benefit_{MT,i}$ για κάθε ενεργή σύνοδο i), το ΚΤ εκκινεί τη διαπραγμάτευση με το ΣΒ που υποδεικνύει ο ΑΠΚ.

5.4.1 Μηχανισμός Διαπραγμάτευσης

5.4.1.1 Επισκόπηση του Θεωρητικού Μοντέλου

Ο μηχανισμός διαπραγμάτευσης που υιοθετείται στην παρούσα μελέτη βασίζεται στο μοντέλο που προτείνεται στην εργασία [71] κι εντάσσεται στα μοντέλα διαπραγμάτευσης απείρου ορίζοντα με αμφίπλευρη αβεβαιότητα (infinite horizon bargaining model with two-sided uncertainty) [72]. Σε αυτή την κλάση των μοντέλων, ο πωλητής (ο παίκτης που κάνει τις προσφορές) κατέχει ένα αντικείμενο με αξία v_s για το ίδιο και v_b για τον αγοραστή. Ο πωλητής κι ο αγοραστής μπορούν να θεωρηθούν κι ως αντίπαλοι που προσπαθούν να μεγιστοποιήσουν τα οφέλη τους, συμμετέχοντας στη διαπραγμάτευση (μη-συνεργατικό παίγνιο δύο παικτών). Η ιδιότητα της αμφίπλευρης αβεβαιότητας αναφέρεται στο γεγονός ότι οι προαναφερθείσες αξιολογήσεις παραμένουν απόρρητες.

Ο πωλητής ξεκινά κάνοντας «ακριβές» προσφορές. Στη συνέχεια, σταδιακά κάνει τις προσφορές του όλο και πιο ελκυστικές, μέχρις ότου ο αγοραστής να δεχθεί να αγοράσει το αντικείμενο. Κι οι δύο παίκτες, όμως, επιβαρύνονται με κάποιο κόστος όταν καθυστερεί η ολοκλήρωση της διαπραγμάτευσης. Αυτό μοντελοποιήθηκε μειώνοντας τις απολαβές τους για κάθε γύρο της διαπραγμάτευσης, σύμφωνα με τον παράγοντα δ_b για τον αγοραστή και δ_s για τον πωλητή ($\delta_b, \delta_s \in (0, 1)$). Στο πρόβλημα που μελετάται εδώ, το κόστος της καθυστέρησης της ολοκλήρωσης της διαπραγμάτευσης εκφράζει την πίεση που ασκεί ο χρόνος στο ΚΤ (αγοραστή) και το ΣΒ (πωλητή), καθώς ο μεν μπορεί να φοβάται ότι η μεταπομπή μπορεί να λάβει χώρα πριν μία συμφωνία, ενώ ο δε μπορεί να ανησυχεί ότι το ΚΤ, ίσως να προτιμήσει έναν άλλο, φθηνότερο, ΣΒ, που ανήκει σε κάποιον άλλο πάροχο [75].

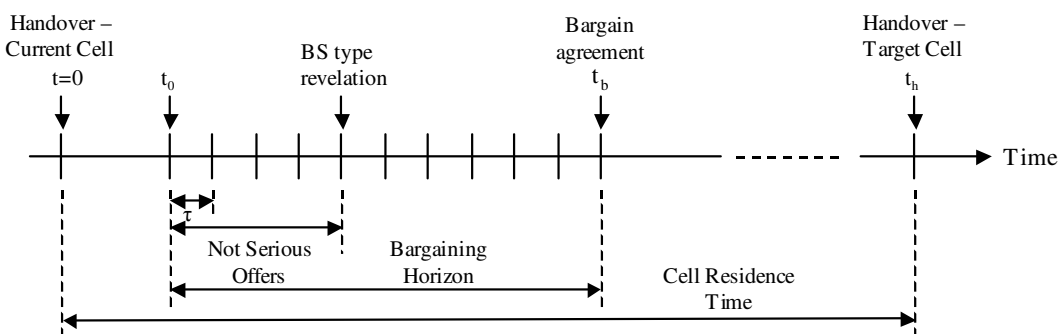
⁶ Η ανάλυση κι αξιολόγηση του ΑΠΚ δεν αποτελεί αντικείμενο του παρόντος κεφαλαίου. Γίνεται η υπόθεση ότι χρησιμοποιείται ο ΑΠΚ που προτείνεται στην εργασία [69]. Επιπλέον ζητήματα σχετικά με την πρόβλεψη κίνησης μπορούν να βρεθούν στις εργασίες [56], [67], [68].

Στην περίπτωση που δε μπορεί να επιτευχθεί μία συμφωνία, κι οι δύο παίκτες λαμβάνουν μηδενική απολαβή. Οι παράγοντες μείωσης (δ_b , δ_s) κι οι κατανομές των αξιολογήσεων των παικτών αποτελούν κοινή γνώση.

Στις προσομοιώσεις, υιοθετήθηκε η ακολουθιακή ισορροπία (sequential equilibrium) που περιγράφεται στην εργασία [71], δηλαδή οι συναρτήσεις που καθορίζουν τις βέλτιστες στρατηγικές των παικτών, δεδομένης της πληροφορίας σε σχέση με το πώς το παίγνιο έχει εξελιχθεί μέχρι στιγμής. Στην προαναφερθείσα ακολουθιακή ισορροπία, οι πληροφορίες αποκαλύπτονται σταδιακά με το χρόνο κι ο ρυθμός αποκάλυψης εξαρτάται από τα κόστη της καθυστέρησης. Οι διαπραγματευόμενοι που αναμένουν μεγαλύτερα οφέλη (που είναι επομένως περισσότερο ανυπόμονοι) φθάνουν σε συμφωνία πριν από αυτούς που περιμένουν μικρότερα οφέλη.

5.4.1.2 Εφαρμογή του Μοντέλου Διαπραγμάτευσης

Στην παρούσα μελέτη, οι παίκτες διαπραγματεύονται για την τιμή της μονάδας του εύρους ζώνης (π.χ. 1 Kb/s). Το ΚΤ ξεκινά μία διαπραγμάτευση με το ΣΒ που θεωρείται ο πιο πιθανός για μεταπομπή, αφού περάσει κάποιο χρόνο, t_0 , στην τρέχουσα κυψέλη. Πρέπει να σημειωθεί ότι, καθώς η σπουδαιότητα μεταξύ διαφορετικών τύπων εφαρμογών μπορεί να ποικίλει, για ένα συγκεκριμένο χρήστη, το ΚΤ κάνει μία ξεχωριστή διαπραγμάτευση για κάθε διαφορετικό τύπο εφαρμογής (π.χ. μεταφορά αρχείων, VoIP, streaming video).



Εικόνα 5.2. Μηχανισμός Διαπραγμάτευσης

Η Εικόνα 5.2 απεικονίζει τη διαδικασία διαπραγμάτευσης. Ο ΣΒ κάνει προσφορές κάθε τ sec, μέχρι το ΚΤ να αποδεχθεί μία προσφορά, τη χρονική στιγμή t_b . Τη

χρονική στιγμή t_h , το ΚΤ διενεργεί την αναμενόμενη μεταπομπή. Πρέπει να σημειωθεί ότι, στο παράδειγμα στην Εικόνα 5.2, $t_h > t_b$. Στην αντίθετη περίπτωση, η διαπραγμάτευση θα έπρεπε να διακοπεί και το ΚΤ κάνει τη μεταπομπή χωρίς να έχουν δεσμευθεί πόροι στο νέο ΣΒ. Αυτό μπορεί να συμβεί στην περίπτωση “no-gar” [72], όπου η συναλλαγή μπορεί να μην είναι δυνατόν να προσφέρει όφελος σε κάποιον από τους παίκτες⁷.

5.4.2 Σχήμα Διαχείρισης Πόρων

Η επιτυχής ολοκλήρωση μίας διαπραγμάτευσης υποδηλώνει ότι το ΚΤ κι ο ΣΒ έχουν συμφωνήσει σε ό,τι αφορά την τιμή της μονάδας του εύρους ζώνης. Η ποσότητα των πόρων, όμως, που θα πρέπει να δεσμεύσει ο ΣΒ δε αποτελεί αποτέλεσμα της διαπραγμάτευσης. Το ΚΤ ενημερώνει την ποσότητα των πόρων που θέλει να δεσμευθεί από τον υποψήφιο ΣΒ, ελέγχοντας συνεχώς, για κάθε τύπο εφαρμογής (για τον οποίο η διαπραγμάτευση έχει τερματιστεί επιτυχώς), αν η αντίστοιχη δέσμευση μπορεί να καλύψει τις ανάγκες των ενεργών εφαρμογών⁸. Αν όχι, το ΚΤ ενημερώνει τη λίστα των δεσμεύσεων, δηλώνοντας την επιπλέον ποσότητα πόρων που χρειάζεται. Το τι συμβαίνει τη στιγμή της μεταπομπής, καθώς και στον ενδιάμεσο χρόνο, περιγράφεται αναλυτικά στις επόμενες παραγράφους.

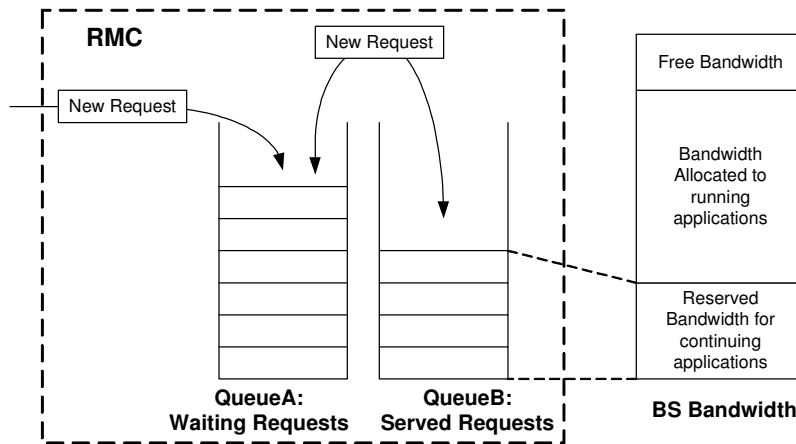
5.4.2.1 Συστατικό Διαχείρισης Πόρων

Το συστατικό διαχείρισης πόρων (Resource Management Component – RMC) βρίσκεται στο ΣΒ (δες Εικόνα 5.1) κι είναι υπεύθυνο για το χειρισμό των αιτήσεων δεσμεύσεων από τα ΚΤ. Ο σκοπός που εξυπηρετεί είναι να περιορίσει την πιθανότητα διακοπής, λόγω μεταπομπής, προσπαθώντας παράλληλα να μεγιστοποιήσει το χρηματικό όφελος του ΣΒ. Για το υπόλοιπο του κεφαλαίου, γίνεται η παρακάτω υπόθεση σε σχέση με τη πολιτική χρέωσης που ακολουθείται από το δίκτυο:

⁷ Η περίπτωση no-gar αναφέρεται στην κατάσταση όπου η αξιολόγηση του αγοραστή μπορεί να είναι χαμηλότερη από την αξιολόγηση του πωλητή, στην οποία περίπτωση δεν μπορεί να επιτευχθεί συμφωνία.

⁸ Γίνεται η υπόθεση ότι κάθε τύπος εφαρμογής έχει αυστηρές απαιτήσεις σε ό,τι αφορά τους πόρους που χρειάζεται. αν ο ΣΒ δε μπορεί να εκχωρήσει τους απαιτούμενους πόρους, η εφαρμογή διακόπτεται (ή μπλοκάρεται, σε περίπτωση μίας νέας εφαρμογής). Αν η συνθήκη αυτή δεν ικανοποιείται, ένας αριθμός από εφαρμογές πιθανώς θα τερματιστούν κατά τη μεταπομπή.

- Η χρέωση από το δίκτυο για τη βασική υπηρεσία σύνδεσης, για S μονάδες πόρων και t χρονικές μονάδες, είναι $c(t,S) = p \cdot S \cdot t$, όπου p ($\$/(\text{Kb/s}) \cdot \text{s}$) είναι η τυπική τιμή που χρεώνεται από το δίκτυο ανά μονάδα εύρους ζώνης.
- Η χρέωση για μία αίτηση προδέσμευσης, i , για s_i μονάδες πόρου, είναι $c_i = p_i \cdot s_i$, όπου p_i ($\$/(\text{Kb/s})$) είναι η τιμή μονάδας (που συμφωνείται από τα δύο μέρη μέσω της διαπραγμάτευσης). Πρέπει να σημειωθεί ότι το δίκτυο πληρώνεται μόνο αν οι δεσμευμένοι πόροι ήταν πράγματι δεσμευμένοι τη στιγμή της μεταπομπής.



Εικόνα 5.3. Στην ουρά *QueueA* αποθηκεύονται νέες αιτήσεις των ΚΤ. Στην ουρά *QueueB* διατηρούνται αιτήσεις που έχουν εξυπηρετηθεί.

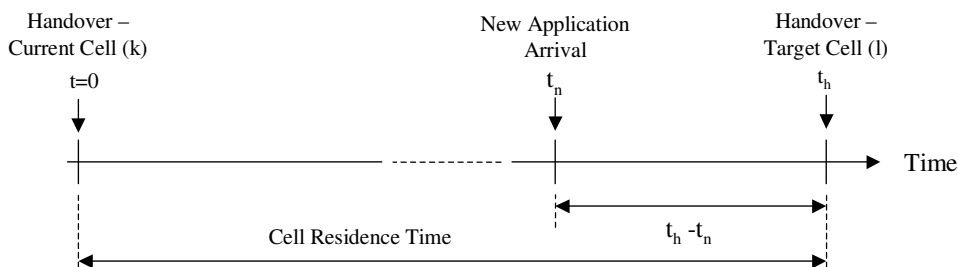
Όταν ένα ΚΤ στέλνει μία νέα αίτηση δέσμευσης πόρων, αυτή η αίτηση εισάγεται στην ουρά *QueueA* (δες Εικόνα 5.3). Όταν κάποια ποσότητα εύρους ζώνης απελευθερώνεται (π.χ. μία εφαρμογή τερματίζεται ή ένα ΚΤ κάνει μία μεταπομπή σε μία γειτονική κυψέλη), οι πιο «σημαντικές» αιτήσεις από την ουρά *QueueA* εισάγονται στην ουρά *QueueB*. Η «σημαντικότητα» μίας αίτησης καθορίζεται από δύο παράγοντες: 1) την αμοιβαία συμφωνημένη τιμή p_i , για τη συγκεκριμένη αίτηση i και 2) την «ηλικία» της αίτησης, δηλαδή το χρόνο από τη στιγμή που το ΚΤ μπήκε στην τρέχουσα κυψέλη (t_i), καθώς όσο πιο παλιά είναι μία αίτηση τόσο πιο σύντομα αναμένεται να λάβει χώρα η μεταπομπή. Συνεπώς, ορίζεται η μετρική Z_A , σύμφωνα με την οποία ταξινομούνται οι αιτήσεις της ουράς *QueueA*, ως εξής:

$$Z_A = p_i \cdot t_i$$

Η μεταφορά μίας αίτησης στην ουρά *QueueB* δεν υπονοεί ότι θα διατηρηθεί εκεί μέχρι την εκτέλεση της αντίστοιχης μεταπομπής. Αιτήσεις από την ουρά *QueueB* μπορεί να εκδιωχθούν (και να ξαναεισαχθούν στην ουρά *QueueA*) κατά την εκδήλωση συγκεκριμένων συμβάντων, π.χ. εκκίνηση μίας νέας εφαρμογής ή μεταπομπή ενός άλλου ΚΤ, όπως θα συζητηθεί και πιο κάτω.

5.4.2.2 Άφιξη Νέας Εφαρμογής

Έστω η κατάσταση που απεικονίζεται στην Εικόνα 5.4. Στο χρόνο t_n , το MT_y (που βρίσκεται στην κυψέλη l) θέλει να εκκινήσει μία καινούρια εφαρμογή, ζητώντας S μονάδες εύρους ζώνης. Σύμφωνα με την υιοθετημένη πολιτική χρέωσης, αν ο ΣΒ αποδεχθεί τη νέα εφαρμογή, θα χρεώσει το MT_y με $p \cdot S$ \$ ανά μονάδα χρόνου. Αντίθετα, όταν φτάσει το MT_x , ο ΣΒ θα λάβει άμεσα ένα κέρδος $p_x \cdot s_x$ \$, λόγω της επιτυχούς προδέσμησης. Αν δεν υπάρχει καθόλου ελεύθερη χωρητικότητα, ο ΣΒ πρέπει να αποφασίσει αν θα απομακρύνει την αίτηση του MT_x από την ουρά *QueueB*, ώστε να προσφέρει πόρους στη νέα εφαρμογή του MT_y ή αν θα μπλοκάρει την εφαρμογή που προσπαθεί να εκκινήσει το MT_y .



Εικόνα 5.4. Άφιξη νέας εφαρμογής

Δεδομένης μίας εκτίμησης του χρόνου παραμονής στην κυψέλη (*cell residence time* - CRT)⁹ και κάνοντας την υπόθεση ότι η νέα εφαρμογή δε θα είναι πολύ μικρής διάρκειας, το κέρδος για το δίκτυο, G , για το χρονικό διάστημα $[t_n, t_h]$, μπορεί να υπολογιστεί ως εξής:

⁹ Το CRT μπορεί να εκτιμηθεί με διάφορους τρόπους. Ένας απ' αυτούς είναι με τη χρήση ενός χαμηλοπερατού φίλτρου της μορφής: $CRT = a \cdot CRT_{new} + (1-a) \cdot CRT$, όπου $0 \leq a \leq 1$

$$G(t_n, t_h) = \begin{cases} p_x \cdot s_x, & \text{if new appl. is blocked} \\ p \cdot S \cdot (t_h - t_n), & \text{otherwise} \end{cases} \quad (5.5)$$

Ο ΣΒ θα επιλέξει την ενέργεια που οδηγεί στο μέγιστο κέρδος. Συγκεκριμένα, αν ο ΣΒ αποφασίσει να αποδεχθεί την πρόσβαση στη νέα εφαρμογή, θα πρέπει να αποβάλλει από την ουρά *QueueB* όσες αιτήσεις απαιτούνται, μέχρι μία επαρκής ποσότητα πόρων να γίνει διαθέσιμη. Είναι σημαντικό, όμως, να καθοριστεί ποιες αιτήσεις θα αποβληθούν. Η συσσωρευτική αξία των συγκεκριμένων αιτήσεων θα πρέπει να είναι όσο το δυνατόν πιο μικρή. Επίσης, η συνολική ποσότητα των πόρων που θα ελευθερωθούν πρέπει να ξεπερνά ή να είναι ίση με την ποσότητα των πόρων που απαιτεί η νέα εφαρμογή. Επιπλέον, το χρηματικό όφελος του συνόλου που εξετάζεται πρέπει να μην ξεπερνά την αξία της νέας εφαρμογής. Αυτό αποτελεί ένα τυπικό πρόβλημα βελτιστοποίησης με περιορισμούς, το οποίο διατυπώνεται ως εξής:

$$\min_y \sum_{j=1}^n p_j s_j y_j, \text{ s.t. } \sum_{j=1}^n s_j y_j \geq S, \sum_{j=1}^n \frac{p_j s_j}{t_j} y_j < p \cdot S$$

$$y_j \in \{0,1\}, j \in N = \{1,2,\dots,n\}$$

όπου N είναι το σύνολο των στοιχείων στην ουρά *QueueB*, t_j είναι ο αναμενόμενος CRT, ενώ y_j είναι δυαδικές μεταβλητές που υποδεικνύουν εάν η λύση περιέχει το αντίστοιχο στοιχείο j , ή όχι. Το S υποδηλώνει την ποσότητα των πόρων που πρέπει να αποδεσμευτούν. Αυτό αποτελεί ένα πρόβλημα γραμμικού ακέραιου προγραμματισμού, το οποίο έχει ομοιότητες με το γνωστό πρόβλημα σακιδίου 0-1 (0-1 knapsack problem) [73]. Για την επίλυση του εν λόγω προβλήματος, υιοθετήθηκε ένας προσεγγιστικός αλγόριθμος (linear programming relaxation) με πολυπλοκότητα $O(n)$. Τα στοιχεία στην ουρά *QueueB* ταξινομούνται βάσει του λόγου τιμής προς απαιτούμενη χωρητικότητα

$$Z_B = p_j / s_j.$$

Ο ψευδοκώδικας πιο κάτω συνοψίζει τον αλγόριθμο:

```

Event: A new application demands S units of bandwidth
/* It is assumed that there is not sufficient free bandwidth */
gatheredResources = 0
gatheredValue = 0
criticalElement = -1
for i = QueueB.length:(-1):1
    request = QueueB.getElem(i)
    gatheredResources = gatheredResources + request.wantedResources
    reqValue = (request.wantedResources*request.price)/timeLeftForHandover()
    gatheredValue = gatheredValue + reqValue
    if gatheredResources >= S then
        if gatheredValue < p*S then
            criticalElement = i
            break
        endif
    endif
endfor

if criticalElement <> -1 then
    for i = criticalElement :QueueB.length
        request = QueueB.getElem(i)
        QueueA.insertElem(request)
        QueueB.remove(i)
    endfor
    return OK
else
    return FAILED
endif

```

Εικόνα 5.5. Άφιξη νέας εφαρμογής

5.4.2.3 Μεταπομπή

Εδώ, μελετούνται οι ενέργειες που πραγματοποιούνται, στην περίπτωση μίας μεταπομπής, αν το RMC δεν έχει ακόμα εξυπηρετήσει όλες τις αιτήσεις προδέσμευσης του ΚΤ που υπέστη την μεταπομπή – δηλαδή, όταν όλες ή ένα μέρος των αιτήσεων παραμένουν στην ουρά *QueueA*. Σε μία τέτοια περίπτωση, είναι γνωστό στο RMC ότι οι αιτήσεις που είναι επί του παρόντος αποθηκευμένες στην ουρά *QueueB* (εξυπηρετημένες αιτήσεις) είναι ανενεργές, καθώς τα αντίστοιχα ΚΤ δεν έχουν ακόμα φτάσει. Είναι, επίσης, γνωστό ότι παρέχοντας

πρόσβαση στο ΚΤ στους πόρους που αιτήθηκε αποβάλλοντας κάποιες από τις αιτήσεις από την ουρά *QueueB* θα έχει ως αποτέλεσμα την προσυμφωνημένη πληρωμή. Οι αιτήσεις που αποβλήθηκαν, πιθανώς, θα έχουν την ευκαιρία να εξυπηρετηθούν εκ νέου πριν από την άφιξη των αντίστοιχων ΚΤ (π.χ. όταν πόροι απελευθερώνονται από άλλα ΚΤ). Έτσι, το RMC επιλέγει το σύνολο των αιτήσεων που θα αποβληθούν επιλύοντας το ακόλουθο πρόβλημα βελτιστοποίησης, παρόμοια με το συμβάν άφιξης νέας εφαρμογές (δες παράγραφο 5.4.2.2):

$$\min_y \sum_{j=1}^n p_j s_j y_j, \text{ s.t. } \sum_{j=1}^n s_j y_j \geq S$$

$$y_j \in \{0,1\}, j \in N = \{1,2,\dots,n\}$$

Ο αντίστοιχος αλγόριθμος παρουσιάζεται στην Εικόνα 5.6:

```

Event: A MT was handed-over and a request demanding S units of bandwidth is out of
QueueB
/* It is assumed that the RMC has not included the request in QueueB */
gatheredResources = 0
criticalElement = -1
for i = QueueB.length:(-1):1
    request = QueueB.getElem(i)
    gatheredResources = gatheredResources + request.wantedResources
    if gatheredResources >= S then
        criticalElement = i
        break
    endif
endfor

if criticalElement <> -1 then
    for i = criticalElement :QueueB.length
        request = QueueB.getElem(i)
        QueueA.insertElem(request)
        QueueB.remove(i)
    endfor
    return OK
else
    return FAILED
endif

```

Εικόνα 5.6. Μεταπομπή

5.4.2.4 Χειρισμός Σφαλμάτων του ΑΠΚ

Μία αστοχία του ΑΠΚ είναι σημαντικό ζήτημα καθώς μπορεί να οδηγήσει σε επιτυχείς δεσμεύσεις σε λάθος ΣΒ. Στο προτεινόμενο σχήμα, όταν διαπιστώνεται ένα σφάλμα του ΑΠΚ, το ΚΤ στέλνει μία αίτηση μεταφοράς δέσμευσης (*reservation transfer request*) στο νέο ΣΒ, υποδεικνύοντας ποιος είναι ο ΣΒ που προβλέφθηκε λανθασμένα. Ο τελευταίος παραδίδει στον πρώτο πληροφορίες σχετικά με τις αιτήσεις του ΚΤ, μαζί με τις συμφωνημένες τιμές. Εν συνεχεία, το RMC επιλέγει το σύνολο των αιτήσεων που θα αποβληθούν από την ουρά *QueueB* επιλύοντας το επόμενο πρόβλημα βελτιστοποίησης:

$$\min_y \sum_{j=1}^n p_j s_j y_j, \text{ s.t. } \sum_{j=1}^n s_j y_j \geq s_x, \sum_{j=1}^n p_j s_j y_j < p_x \cdot s_x$$
$$y_j \in \{0,1\}, j \in N = \{1,2,\dots,n\}$$

όπου το s_x υποδηλώνει το πλήθος των μονάδων εύρους ζώνης που απαιτούνται από την αίτηση και το p_x (\$/(Kb/s)) την τιμή που έχει συμφωνηθεί για τη δέσμευση στο λανθασμένο ΣΒ. Ο αντίστοιχος ψευδοκώδικας παρουσιάζεται στην Εικόνα 5.7 κι επαναλαμβάνεται για κάθε αίτηση που έχει γίνει στον λάθος ΣΒ.

Event: A MT was handed over but has previously made a path prediction error. The RMC tries to serve a request, demanding S units of bandwidth, which was served by the other BS. Agreed price: p_x

```
gatheredResources = 0
gatheredValue = 0
criticalElement = -1
for i = QueueB.length:(-1):1
    request = QueueB.getElem(i)
    gatheredResources = gatheredResources + request.wantedResources
    reqValue = request.wantedResources*request.price
    gatheredValue = gatheredValue + reqValue
    if gatheredResources >= S then
        if gatheredValue <  $p_x * S$  then
            criticalElement = i
            break
        endif
    endif
endfor

if criticalElement <> -1 then
    for i = criticalElement :QueueB.length
        request = QueueB.getElem(i)
        QueueA.insertElem(request)
        QueueB.remove(i)
    endfor
    return OK
else
    return FAILED
endif
```

Εικόνα 5.7. Συμβάν μεταπομπής

5.4.2.5 Διαχείριση Πόρων στο ΚΤ

Στην παρούσα παράγραφο, περιγράφεται πως το ΚΤ, μετά από μία μεταπομπή, χειρίζεται τους επιτυχώς δεσμευμένους πόρους. Οι εφαρμογές που έχουν κάνει επιτυχείς δεσμεύσεις έχουν προτεραιότητα στους δεσμευμένους πόρους. Είναι πιθανόν, όμως, ότι το δεσμευμένο εύρος ζώνης για μία εφαρμογή μπορεί να ξεπερνά την πραγματική ζήτηση τη στιγμή της μεταπομπής. Κάτι τέτοιο είναι

πιθανόν, καθώς η στρατηγική του ΚΤ (για κάθε τύπο εφαρμογής με μία επιτυχημένα ολοκληρωμένη διαπραγμάτευση), όπως έχει αναφερθεί και στην παράγραφο 5.4.2, είναι να διατηρήσει την ποσότητα των πόρων της αίτησης πάντα ίση ή μεγαλύτερη από τις ανάγκες των ενεργών εφαρμογών.

```
Event: Resource allocation to the applications by the MT after the handover

poolSize = 0
for i=1:numberOfApplicationTypes
    reservedCapacity = application[i].getReservedCapacity()
    currentCapacity = application[i].getCurrentCapacity()
    if currentCapacity <= reservedCapacity then
        application[i].continue()
        poolSize = poolSize + (reservedCapacity - currentCapacity)
    endif
endfor

application.sort() // sort the array of applications by the ratio Ui/si

for i = 1:numberOfApplicationTypes
    appl = application[i]
    if appl.getCurrentCapacity() <= poolSize then
        appl.continue()
        poolSize = poolSize - appl.getCurrentCapacity ()
    endif
endfor

releaseResources(poolSize)
```

Εικόνα 5.8. Κατανομή των δεσμευμένων πόρων στις ενεργές εφαρμογές από το ΚΤ

Στο προτεινόμενο σχήμα, το αδιάθετο εύρος ζώνης συγκεντρώνεται σε μία «δεξαμενή» κι οι εφαρμογές που δεν έχουν επιτύχει να εξασφαλίσουν πόρους επωφελοούνται από αυτήν. Αν η δεξαμενή περιέχει περισσότερους πόρους απ' όσους ζητούν οι εφαρμογές, όλες οι εφαρμογές προστατεύονται από διακοπή. Στην αντίθετη περίπτωση, όμως, το ΚΤ θα πρέπει να επιλέξει σε ποιες εφαρμογές θα εκχωρηθεί το προαναφερθέν εύρος ζώνης. Αυτό αποτελεί ένα κοινό πρόβλημα

σακιδίου, όπου η χωρητικότητα του σακιδίου είναι η συνολική ποσότητα πόρων στη δεξαμενή, η αξία των αντικειμένων είναι ίση με την αξιολόγηση του ΚΤ (δες (5.3)) και το βάρος τους είναι η αντίστοιχη ζήτηση σε εύρος ζώνης (δες Εικόνα 5.8). Οι εφαρμογές που δεν κατάφεραν να λάβουν μερίδιο από τους πόρους στη δεξαμενή αντιμετωπίζονται ως νέες εφαρμογές στην τρέχουσα κυψέλη, έτσι, χρειάζεται να ανταγωνιστούν για τους ζητούμενους πόρους.

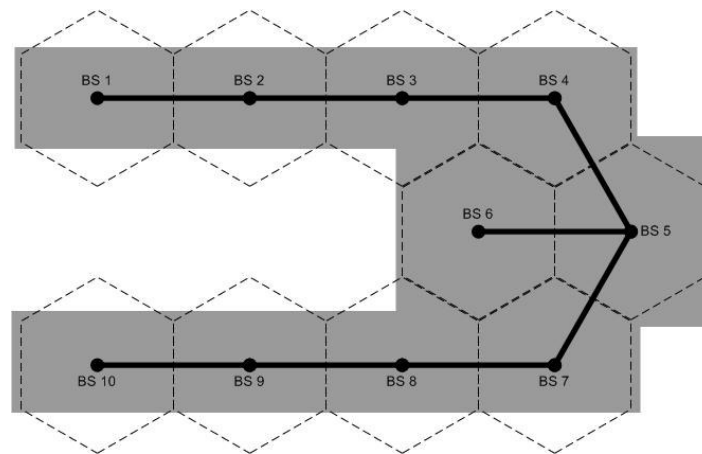
5.5 Προσομοιώσεις

5.5.1 Σχεδιασμός Προσομοίωσης

Για να ποσοτικοποιηθεί η αποτελεσματικότητα του προτεινόμενου σχήματος χρέωσης, σε ό,τι αφορά την πιθανότητα μπλοκαρίσματος νέων κι από μεταπομπή εφαρμογών, πραγματοποιήθηκε μία σειρά από προσομοιώσεις. Προσομοιώθηκαν η στοχαστική κίνηση των χρηστών σε μία περιοχή καλυπτόμενη από έναν αριθμό από ΣΒ και δικτυακές εφαρμογές με διαφορετικά χαρακτηριστικά.

5.5.1.1 Χωρική Διαρρύθμιση και Μοντέλο Κινητικότητας

Στις προσομοιώσεις, υιοθετήθηκε μία διαρρύθμιση χώρων, όπως αυτή στην Εικόνα 5.9 (παρόμοια διαρρύθμιση χρησιμοποιήθηκε και στην προσομοίωση του συστήματος MITOS, στο Κεφάλαιο 4). Έγινε η υπόθεση ότι υπήρχαν 10 ΣΒ στην υποδομή. Μία γραμμή που συνδέει δύο ΣΒ υποδηλώνει ότι οι αντίστοιχες κυψέλες είναι γειτονικές. Οι χρόνοι παραμονής σε μία κυψέλη ακολουθούν τη γενικευμένη κατανομή Γάμα, σύμφωνα και με τη μελέτη στην εργασία [74]. Γίνεται επίσης η υπόθεση ότι ο ΑΠΚ καταφέρνει να προβλέπει ορθά την επόμενη κυψέλη με πιθανότητα 0.8 [69].



Εικόνα 5.9. Χωρική διάταξη της περιοχής προσομοίωσης

5.5.1.2 Μοντέλο Εφαρμογών

Γίνεται η υπόθεση ενός απλού μοντέλου εφαρμογών, *file transfer protocol* (FTP), *hyper-text transport protocol* (HTTP), VoIP και video. Ο Πίνακας 5.1 συνοψίζει τα χαρακτηριστικά αυτών των εφαρμογών. Επιπλέον, γίνεται η υπόθεση ότι υπάρχει ένα μέγιστο πλήθος ενεργών στιγμιοτύπων ανά τύπο εφαρμογής που μπορούν να τρέχουν παράλληλα, ανά πάσα στιγμή, σε κάθε ΚΤ. Συγκεκριμένα, ένας χρήστης μπορεί να τρέξει παράλληλα μέχρι τρεις FTP, τρεις HTTP, μία VoIP και μία video εφαρμογή. Η διάρκεια μίας εφαρμογής μοντελοποιήθηκε ως μία τυχαία μεταβλητή που ακολουθεί την εκθετική κατανομή. Οι αφίξεις νέων εφαρμογών ακολουθούν την κατανομή Poisson.

Πίνακας 5.1. Χαρακτηριστικά Εφαρμογών

Τύπος Εφαρμογής	Απαιτούμενο Εύρος Ζώνης (Kb/s)	Μέση Διάρκεια Εφαρμογής (s)	Μέσος Χρόνος Ενδο-άφιξης (s)
FTP	100	300	1800
HTTP	50	3	60
VoIP	64	120	1200
Video	512	300	1800

5.5.2 Αποτελέσματα Προσομοιώσεων

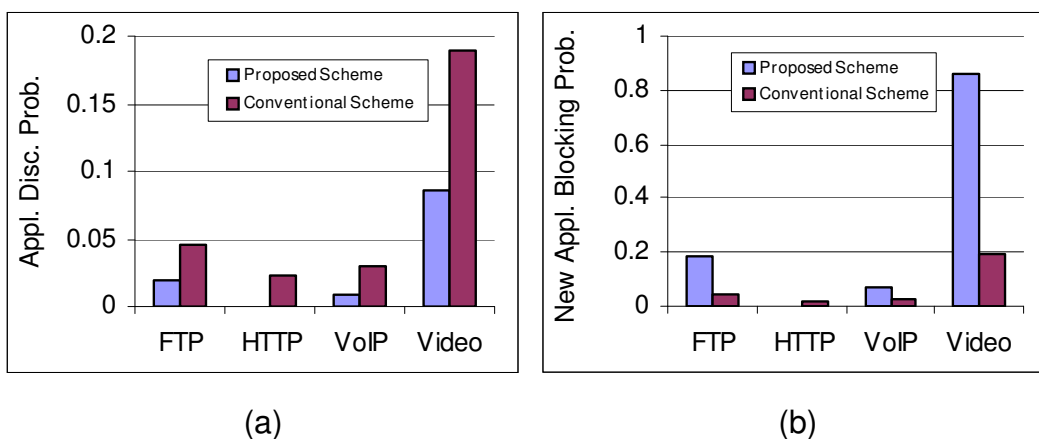
Για την αξιολόγηση του προτεινόμενου σχήματος, επιλέχθηκε η πιθανότητα μπλοκαρίσματος μίας νέας εφαρμογής P_n κι η πιθανότητα μπλοκαρίσματος μίας εφαρμογής προερχόμενης από μεταπομπή P_h . Η P_n ορίζεται ως το πλήθος των

αποτυχιών έναρξης εφαρμογής προς το πλήθος των προσπαθειών έναρξης εφαρμογής, ενώ η P_h ορίζεται ως το πλήθος των εφαρμογών που διακόπηκαν μετά από μία μεταπομπή προς το πλήθος των εφαρμογών που υπέστησαν μεταπομπή. Τελικά, εξετάζεται η επίδραση του προτεινόμενου σχήματος στα κέρδη του δικτύου.

Η χωρητικότητα κάθε ΣΒ (C_{TOTAL}) τέθηκε στις 20,000 μονάδες (Kb/s), ενώ υποτέθηκε η ύπαρξη 2,000 ΚΤ που περιφέρονταν στο δίκτυο. Αυτό το σενάριο είχε ως αποτέλεσμα ένα αρκετά κορεσμένο δίκτυο. Σε ό,τι αφορά τις παραμέτρους τις διαπραγμάτευσης, έγινε η υπόθεση ότι το v_b και το v_s κατανέμονται ομοιόμορφα στα διαστήματα (0.5,1.5) και (0,0.5), αντίστοιχα, ενώ, $\delta_s = \delta_b = 0.75$ (δες υποκεφάλαιο 5.4).

5.5.2.1 Επίδοση των Εφαρμογών

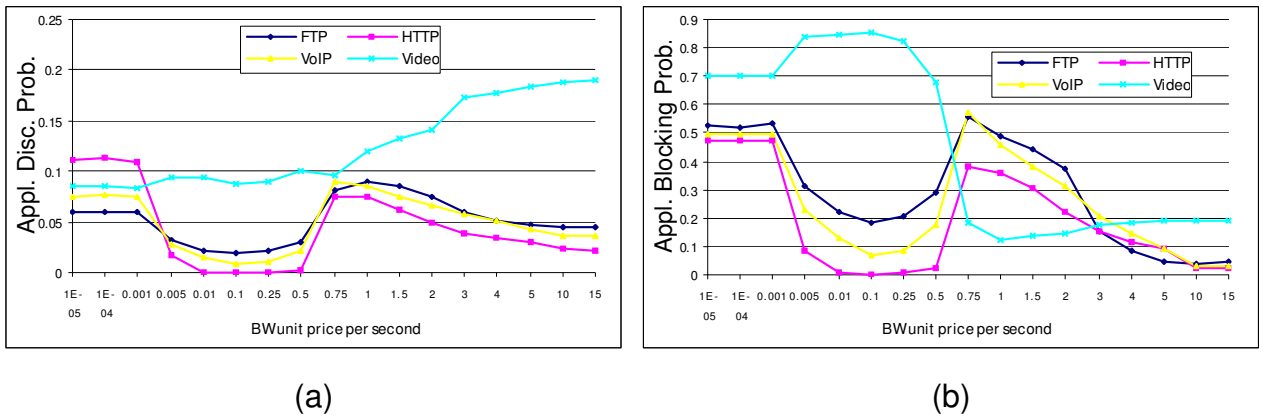
Για $\rho = 0.1$ \$/((Kb/s)·s), οι προσομοιώσεις έδειξαν τα αποτελέσματα που απεικονίζονται στην Εικόνα 5.10(a) κι Εικόνα 5.10(b). Το προτεινόμενο σχήμα μειώνει την πιθανότητα διακοπής μίας εφαρμογής, λόγω μεταπομπής, P_h , για όλους τους τύπους εφαρμογών, ενώ η πιθανότητα μπλοκαρίσματος μίας νέας εφαρμογής, P_n , είναι μεγαλύτερη, σε σχέση με την περίπτωση χωρίς προδέσμευση πόρων, όπως ήταν αναμενόμενο.



Εικόνα 5.10. (a) Πιθανότητα διακοπής εφαρμογής (b) Πιθανότητα μπλοκαρίσματος νέας εφαρμογής

Καθώς ο αλγόριθμος ελέγχου αποδοχής που περιγράφεται στην παράγραφο 5.4.2.2 λαμβάνει υπόψη την τυπική τιμή της μονάδας του εύρους ζώνης ανά δευτερόλεπτο (ρ), αναμένεται ότι η τιμή της ρ θα έχει άμεση επίδραση στο P_n και

στο P_h . Η τιμή του ρ λαμβάνεται, επίσης, υπόψη όταν το ΚΤ πρέπει να ανταγωνιστεί με τους ήδη παρόντες στη νέα κυψέλη, αν το RMC αρνηθεί την πρόσβαση στο εύρος ζώνης ασφαλείας (δες παράγραφο 5.4.2.3).



Εικόνα 5.11. Επίδραση της τυπικής τιμής στην (a) πιθανότητα διακοπής εφαρμογής (b) πιθανότητα μπλοκαρίσματος νέας εφαρμογής

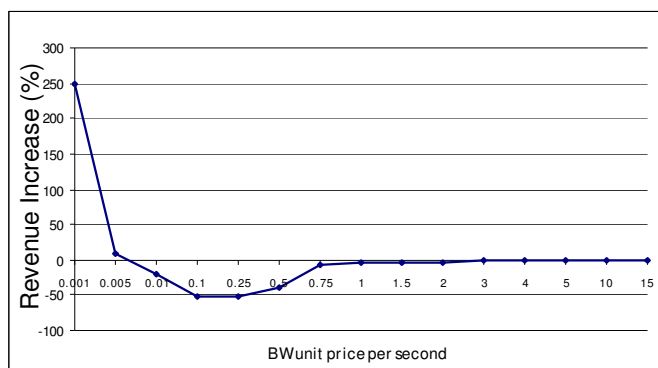
Οι κύριες παρατηρήσεις από την Εικόνα 5.11(a) και την Εικόνα 5.11(b) είναι οι ακόλουθες: 1) η επίδραση του ρ στις πιθανότητες P_h και P_n δεν είναι μονοτονική, 2) υπάρχει ένα πεδίο τιμών για το οποίο κι η P_h κι η P_n ελαχιστοποιούνται για το FTP, το VoIP και το HTTP και 3) καθώς αυξάνεται το ρ , η P_h κι η P_n συγκλίνουν στις αντίστοιχες πιθανότητες μπλοκαρίσματος τις περίπτωσης χωρίς προδέσμευση πόρων (δες Εικόνα 5.10(a) κι Εικόνα 5.10 (b)). Από τα παραπάνω, είναι εμφανές ότι ο προτεινόμενος μηχανισμός ευνοεί εφαρμογές χαμηλού ρυθμού. Σύμφωνα με τη μελέτη στην παράγραφο 5.4.2.1 και την παράγραφο 5.4.2.3, κάτι τέτοιο είναι λογικό, καθώς, στην περίπτωση συμφόρησης, μία αίτηση χαμηλού ρυθμού είναι πιο πιθανό να βρει τους απαραίτητους πόρους, σε σχέση με μία εφαρμογή υψηλού ρυθμού.

Η αύξηση στην τυπική τιμή ρ δεν ευνοεί μόνο τις νέες εφαρμογές, αλλά κι όσες εφαρμογές υφίστανται μεταπομπή και για κάποιο λόγο δεν έχουν επιτύχει να κάνουν μία δέσμευση στον μελλοντικό ΣΒ. Αυτό συμβαίνει γιατί τέτοιες εφαρμογές αντιμετωπίζονται ως νέες εφαρμογές στο ΣΒ. Έτσι, αυξάνοντας την τιμή της ρ προκαλείται μία μείωση στην πιθανότητα διακοπής μίας εφαρμογής. Αυξάνοντας περαιτέρω όμως την τιμή της ρ , το δίκτυο παύει να ευνοεί τις εφαρμογές που υφίστανται μεταπομπή κι έχουν κάνει επιτυχώς κάποια δέσμευση. Έτσι,

παρατηρείται μία ανταλλαγή μεταξύ μεγαλύτερων και μικρότερων τιμών της ρ , που είναι εμφανής στην Εικόνα 5.11(a) κι Εικόνα 5.11(a).

5.5.3 Κέρδος του Δικτύου

Μία ακόμα σημαντική πτυχή του μελετούμενου μηχανισμού αποτελεί το οικονομικό όφελος που αποκομίζει το δίκτυο παρέχοντας την υπηρεσία προδραστικής δέσμευσης πόρων στα ΚΤ. Το σύστημα, για να περιορίσει την πιθανότητα διακοπής των εφαρμογών, μειώνει το διαθέσιμο εύρος ζώνης για τις νέες εφαρμογές. Έτσι, το δίκτυο χάνει ένα ποσό χρημάτων καθώς το συνολικό πλήθος των εφαρμογών που γίνονται αποδεκτές μειώνεται. Το δίκτυο πληρώνεται για να προδεσμεύει πόρους· όπως θα συζητηθεί παρακάτω, όμως, αυτές οι πληρωμές δεν είναι πάντοτε επαρκείς για να καλύψουν τις προαναφερθείσες απώλειες. Αυτό ενισχύεται ακόμα πιο πολύ, αν λάβει κανείς υπόψη τις αστοχίες του ΑΠΚ, που ευθύνονται για άσκοπες δεσμεύσεις σε λάθος ΣΒ και τα ΚΤ που υπερεκτιμούν τις ανάγκες τους κάνοντας υπερβολικές δεσμεύσεις, περιορίζοντας, έτσι, ακόμα πιο πολύ το εύρος ζώνης για τις νέες εφαρμογές.



Εικόνα 5.12. Ποσοστιαία αύξηση των κερδών με τη χρήση του προτεινόμενου σχήματος

Όπως φαίνεται στην Εικόνα 5.12, για μικρές τιμές της τυπικής τιμής ρ το όφελος της χρήσης του προτεινόμενου μηχανισμού για το δίκτυο είναι σημαντικό. Για μεγαλύτερες τιμές της ρ , όμως, το όφελος μειώνεται σημαντικά και, τελικά, συγκλίνει στο μηδέν.

5.6 Συμπεράσματα

Στο κεφάλαιο αυτό, προτάθηκε ένα προδραστικό σχήμα δέσμευσης δικτυακών πόρων, για την ελάττωση των συμβάντων διακοπής σύνδεσης, λόγω μεταπομπής. Αυτή θεωρείται μία ιδιαίτερη υπηρεσία του δικτύου, για την οποία ο χρήστης πρέπει να πληρώσει. Μία ακολουθιακή διαδικασία διαπραγμάτευσης, η οποία μοντελοποιείται ως ένα μη-συνεργατικό παίγνιο δύο παικτών, του κινητού τερματικού και του στόχου ΣΒ, προηγείται μίας τέτοιας δέσμευσης. Με αυτή τη διαδικασία, και τα δύο μέρη συμφωνούν σε μία τιμή του πόρου (δηλαδή του εύρους ζώνης). Ακολούθως, το ΚΤ μπορεί να αιτηθεί μία προδέσμευση από το ΣΒ. Ο ΣΒ συγκεντρώνει τέτοιου είδους αιτήσεις από διαφορετικά τερματικά κι όταν λαμβάνουν χώρα συγκεκριμένα συμβάντα (π.χ. έναρξη μίας εφαρμογής, μεταπομπή, τερματισμός εφαρμογής), αποφασίζει ποιες αιτήσεις θα αποκτήσουν τους ζητούμενους πόρους.

Το προτεινόμενο σχήμα προσομοιώθηκε με σκοπό την αξιολόγηση του. Τα αποτελέσματα της προσομοίωσης έδειξαν ότι ο προτεινόμενος μηχανισμός είναι ικανός να ελαττώσει την πιθανότητα διακοπής εφαρμογής σε σύγκριση με το σενάριο χωρίς προδέσμευση πόρων. Αυτή η μείωση συνοδεύεται από μία αντίστοιχη αύξηση στην πιθανότητα μπλοκαρίσματος νέας εφαρμογής. Παρατηρήθηκε, ακόμα, ότι ο μηχανισμός ευνοεί εφαρμογές με σχετικά μικρές απαιτήσεις σε εύρος ζώνης. Μελετήθηκε, επίσης, η επίδραση της τιμής της βασικής υπηρεσίας σύνδεσης στις προαναφερθείσες πιθανότητες. Το κέρδος του δικτύου φάνηκε ότι εξαρτάται από την τυπική τιμή σύνδεσης. Για ένα συγκεκριμένο εύρος τιμών, το κέρδος του δικτύου είναι σημαντικά αυξημένο, ενώ οι δείκτες των επιδόσεων σχετικά με το μπλοκάρισμα των εφαρμογών δείχνουν εμφανή βελτίωση για τον κινητό χρήστη.

Μελλοντικά, η μελέτη θα εστιαστεί στο πως τέτοιοι μηχανισμοί μπορούν να εφαρμοστούν σε άλλους τύπους δικτυακών πόρων πέραν του εύρους ζώνης (π.χ. υπολογιστικούς πόρους, μνήμη κι αποθηκευτική χωρητικότητα) και στο να παραχθεί ένα ενιαίο πλαίσιο για την προδραστική διαχείριση των πόρων σε τέτοια κινητά κι ασταθή περιβάλλοντα.

ΚΕΦΑΛΑΙΟ 6

ΠΡΟΔΡΑΣΤΙΚΗ ΔΙΑΧΕΙΡΙΣΗ ΕΥΡΟΥΣ ΖΩΝΗΣ ΣΕ ΚΥΨΕΛΩΤΑ ΔΙΚΤΥΑ ΜΕ ΔΗΜΟΠΡΑΣΙΕΣ

6.1 Εισαγωγή

Στο παρόν κεφάλαιο, παρουσιάζεται ένα εναλλακτικό σχήμα προδραστικής διαχείρισης πόρων σε σχέση με αυτό που μελετήθηκε στο Κεφάλαιο 5. Σε αντίθεση με το σχήμα του προηγούμενου κεφαλαίου, όπου κάθε ΚΤ δρούσε ανταγωνιστικά με το ΣΒ, στην παρούσα μελέτη, οι ανταγωνιζόμενες οντότητες είναι τα ΚΤ μεταξύ τους, τα οποία προσπαθούν να πάρουν μερίδιο από τους πόρους των ΣΒ που βρίσκονται γειτονικά τους. Κάθε ΣΒ παίζει το ρόλο του *πλειστηριαστή*, συλλέγοντας προσφορές από τα ενδιαφερόμενα ΚΤ. Ο ΣΒ επιλέγει τους νικητές, εκχωρώντας τους πόρους που ζήτησαν, κι εισπράττοντας τα αντίστοιχα κέρδη.

Βασικός σκοπός της παρούσας μελέτης είναι να προσφέρει ένα πλαίσιο μέσω του οποίου η διαχείριση των πόρων να γίνεται εξατομικευμένα, λαμβάνοντας υπόψη την κρισιμότητα για κάθε χρήστη των πόρων που αιτείται. Συγκεκριμένα, για ένα ποσό πόρων σε έναν υποτιθέμενο ΣΒ λαμβάνεται υπόψη 1) η πιθανότητα να γίνει η μεταπομπή του ΚΤ στο προσεχές μέλλον, 2) το πλήθος των ενεργών συνδέσεων που αναμένεται να έχει ο χρήστης, και 3) οι προσωπικές προτιμήσεις του χρήστη σχετικά με την κρισιμότητα κάθε τύπου εφαρμογής και των συνεπειών διακοπής της. Έτσι οι προσφορές διαφοροποιούνται από χρήστη σε χρήστη, αλλά κι από μία χρονική στιγμή σε μία άλλη, με αποτέλεσμα το προτεινόμενο σχήμα να είναι πλήρως δυναμικό.

Για την ατομική δέσμευση πόρων με δεδομένο το μοναδικό τρόπο κίνησης του κάθε χρήστη, στα πλαίσια της παρούσας μελέτης, προτείνεται και μία τεχνική για την πρόβλεψη της μελλοντικής θέσης του χρήστη, ώστε να επιτυγχάνεται δέσμευση πόρων στους κατάλληλους ΣΒ.

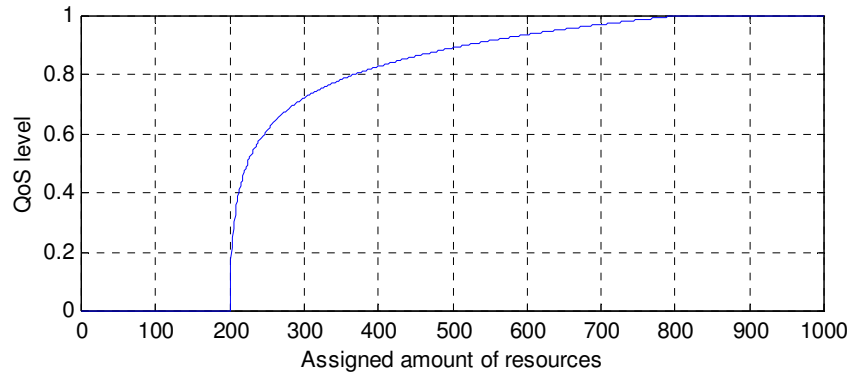
6.2 Μοντέλο Συστήματος

Στο παρόν υποκεφάλαιο περιγράφεται το μοντέλο του συστήματος που υιοθετήθηκε για το προτεινόμενο σχήμα προδραστικής διαχείρισης πόρων σε ασύρματα κυψελωτά δίκτυα. Θεωρείται ότι υπάρχει ένα σύνολο ΚΤ, $J = \{1, \dots, N\}$, τα οποία κινούνται στην έκταση ενός ασύρματου κυψελωτού δικτύου που απαρτίζεται από ένα σύνολο ΣΒ, $I = \{1, \dots, B\}$. Η χωρητικότητα του ΣΒ i υποδηλώνεται με R_i (b/s). Θεωρείται, επίσης, ότι, ανά πάσα στιγμή, κάθε χρήστης συνδέεται στο δίκτυο με τον κοντινότερο προς αυτόν ΣΒ (γεωγραφικά ή σε ό,τι αφορά τις απώλειες σήματος).

Γίνεται η υπόθεση ότι ένας χρήστης μπορεί να ξεκινήσει μία νέα εφαρμογή εφόσον ο ΣΒ, ο οποίος τον εξυπηρετεί, διαθέτει το απαιτούμενο ποσό πόρων. Η προϋπόθεση ύπαρξης πόρων για μία εφαρμογή δεν αφορά μόνο στην εκκίνηση της, αλλά και στη διατήρησή της, όταν ο χρήστης, λόγω της κινητικότητάς του, μεταπέμπεται σε κάποιον γειτονικό ΣΒ. Αν ο νέος ΣΒ δε διαθέτει το αναγκαίο ποσό πόρων, η εφαρμογή του χρήστη θα τερματιστεί [60].

Στο παρόν μοντέλο, γίνεται η υπόθεση ότι υπάρχει ένα σύνολο $C = \{1, \dots, K\}$ διαφορετικών τύπων εφαρμογών (π.χ. VoIP, video, FTP, WWW), με άνω και κάτω όρια στο ρυθμό τους, m_c και M_c , $\forall c \in C$, αντίστοιχα. Οι εφαρμογές που υπακούουν στο συγκεκριμένο μοντέλο χαρακτηρίζονται κι ως «ημι-ελαστικές», με τυπικά παραδείγματα να αποτελούν εφαρμογές με προσαρμοστική κωδικοποίηση (π.χ. VoIP, video) [76]. Για κάθε τύπο εφαρμογής c , υπάρχει μία συνάρτηση g_c που υποδηλώνει το επίπεδο της ποιότητας υπηρεσίας που επιτυγχάνεται (σε κλίμακα από 0 έως 1) δεδομένου του ρυθμού r που έχει εκχωρήσει το δίκτυο. Αυτή η συνάρτηση είναι κατά κανόνα αύξουσα και κοίλη [76], όπως η συνάρτηση που υποτίθεται εδώ:

$$g_c(r) = \begin{cases} 0, & r < m_c \\ \frac{\ln(r - m_c + 1)}{\ln(M_c - m_c + 1)}, & m_c \leq r \leq M_c, \forall c \in C \\ 1, & r > M_c \end{cases} \quad (6.1)$$



Εικόνα 6.1. Επίπεδο ποιότητας υπηρεσίας συναρτήσει του ποσού πόρων που έχουν εκχωρηθεί.
 $m_c = 200$, $M_c = 800$.

Η μορφή της συνάρτησης $g_c(r)$ απεικονίζεται στην Εικόνα 6.1. Όπως φαίνεται, όταν η εφαρμογή δε διαθέτει το ελάχιστο ποσό πόρων, το QoS που επιτυγχάνεται είναι μηδενικό. Όσο αυξάνει το ποσό πόρων μέχρι το μέγιστο που χρειάζεται η εφαρμογή, το QoS συγκλίνει στη μονάδα. Πέραν του μέγιστου ορίου, το QoS παύει να αυξάνεται. Συνεπώς, όταν ο ΣΒ δεν μπορεί να εξασφαλίσει τον ελάχιστο ρυθμό για μία εφαρμογή που ζητά να εκκινήσει, η εφαρμογή αυτή δεν ξεκινά. Αν επίσης, μετά την μεταπομπή ενός χρήστη, ο νέος ΣΒ δε διαθέτει πόρους που να μπορούν να εξασφαλίσουν τον ελάχιστο ρυθμό σε κάποια εφαρμογή του χρήστη, η εφαρμογή αυτή διακόπτεται.

6.3 Ολοκληρωμένο Σύστημα Διαχείρισης Ασύρματων Πόρων

Με δεδομένο το μοντέλο του συστήματος που περιγράφηκε στο προηγούμενο υποκεφάλαιο, προτείνεται εδώ ένα σχήμα διαχείρισης ασύρματων πόρων. Βασικός σκοπός του σχήματος αποτελεί ο περιορισμός της συχνότητας των διακοπών εφαρμογών, λόγω μεταπομπής, που, όπως έχει ήδη αναφερθεί, μπορεί να αποτελέσει σημαντικό παράγοντα ενόχλησης του χρήστη, σε πολύ μεγαλύτερο βαθμό μάλιστα από την αδυναμία εκκίνησης μίας νέας εφαρμογής. Σε αντίθεση με άλλα σχήματα που έχουν ήδη προταθεί, στόχος του παρόντος σχήματος είναι η διαχείριση να μη γίνεται απρόσωπα (π.χ. με σκοπό την ελαχιστοποίηση της συνολικής συχνότητας διακοπών, λόγω μεταπομπής), αλλά να λαμβάνονται υπόψη οι ανάγκες του κάθε χρήστη.

Προς την κατεύθυνση της εξατομικευμένης διαχείρισης πόρων, γίνεται η υπόθεση ότι κάθε χρήστης i έχει μία «ευαισθησία», για κάθε τύπο εφαρμογής c . Συγκεκριμένα, με $ncr_{i,c}$ υποδηλώνεται η σημαντικότητα (από 0 μέχρι 1) για το χρήστη i να μπορεί να ξεκινήσει μία νέα εφαρμογή τύπου c , ενώ με $hcr_{i,c}$ το να μη διακοπεί μία εφαρμογή τύπου c , μετά από μία μεταπομπή. Έτσι, για το χρήστη i , η σημαντικότητα του ρυθμού r που έχει εκχωρήσει το δίκτυο για κάποια εφαρμογή τύπου c είναι $nu_{i,c}(r) = ncr_{i,c} \cdot g_c(r)$, όταν πρόκειται για μία νέα εφαρμογή, ή $hu_{i,c}(r) = hcr_{i,c} \cdot g_c(r)$, για μία εφαρμογή που υφίσταται μεταπομπή.

Στο προτεινόμενο σχήμα, ο χρόνος διαιρείται σε χρονοθυρίδες διάρκειας T sec, με το διάστημα $[nT, (n+1)T)$ να αποτελεί τη χρονοθυρίδα $(n+1)$, $n = 0, 1, 2, \dots$. Στην αρχή κάθε χρονοθυρίδας, το δίκτυο συγκεντρώνει από κάθε χρήστη αιτήσεις για προ-δέσμευση πόρων σε διάφορους ΣΒ (εκτός του ΣΒ που εξυπηρετεί τη δεδομένη στιγμή το χρήστη). Αυτό έχει ως σκοπό, αν συμβεί μία μεταπομπή κατά τη διάρκεια της επικείμενης χρονοθυρίδας, ο χρήστης να βρει στο νέο ΣΒ τους πόρους που χρειάζονται οι εφαρμογές του. Οι συγκεκριμένες αιτήσεις είναι έγκυρες μόνο για τη διάρκεια μίας χρονοθυρίδας. Ο χρήστης στην αρχή της επόμενης χρονοθυρίδας θα πρέπει να υποβάλλει νέες αιτήσεις στο δίκτυο.

Σε αντίθεση με τις αιτήσεις προ-δέσμευσης πόρων που πραγματοποιούνται μόνο στην αρχή κάθε χρονοθυρίδας, στο προτεινόμενο σχήμα, ένας χρήστης μπορεί να αιτηθεί την εκκίνηση μίας νέας εφαρμογής ανά πάσα στιγμή. Το αντίθετο θα ήταν, φυσικά, μη ρεαλιστικό. Το δίκτυο αποφασίζει αν θα επιτρέψει τη σύνδεση στη νέα εφαρμογή και, αν ναι, το ρυθμό που θα της εκχωρήσει. Αν η εφαρμογή γίνει δεκτή, το δίκτυο εξασφαλίζει (τουλάχιστον για όσο ο χρήστης παραμένει στην ίδια κυψέλη) ότι η εφαρμογή δε θα διακοπεί. Ο ρυθμός της εφαρμογής διαπραγματεύεται στην αρχή κάθε χρονοθυρίδας, αλλά σε καμία περίπτωση δεν πέφτει κάτω από τον ελάχιστο ρυθμό που χρειάζεται για να λειτουργήσει.

Στο προτεινόμενο σχήμα, όλες οι αιτήσεις σε ένα ΣΒ, είτε αυτές αφορούν σε προδέσμευση πόρων, είτε σε εκκίνηση νέων εφαρμογών, αξιολογούνται από κοινού, καθώς όλες τους αποσκοπούν στο να αποσπάσουν ένα μέρος των πόρων του. Καθώς ο κάθε χρήστης δεν έχει κάποιον περιορισμό στο ποσό των πόρων που μπορεί να αιτηθεί και με δεδομένο ότι κάθε ΣΒ έχει πεπερασμένους πόρους, δεν είναι δυνατόν να εξυπηρετηθούν πάντα όλες οι αιτήσεις. Συγκεκριμένα,

κάποιες αιτήσεις, ίσως, θα πρέπει να απορριφθούν, ενώ κάποιες άλλες να εξυπηρετηθούν μερικώς. Το πρόβλημα είναι, λοιπόν, η ιεράρχηση / ταξινόμηση των αιτήσεων ως προς την σημαντικότητα τους, ώστε οι διαθέσιμοι πόροι να εκχωρηθούν αναλόγως. Για το σκοπό αυτό, θα μπορούσε να ληφθεί υπόψη η σημαντικότητα κάθε τύπου εφαρμογής για το χρήστη, όπως ορίστηκε πιο πάνω ($hu_{i,c}(r)$ για τις νέες ή $hu_{i,c}(r)$ για όσες προέρχονται από μεταπομπή).

Περαιτέρω διαμόρφωση της σημαντικότητας μίας εφαρμογής, υποψήφιας για μεταπομπή από ένα ΣΒ a σε ένα ΣΒ b , μπορεί να αποτελέσει κι η πιθανότητα $\rho_{HO,i}(a,b)$ ο χρήστης στη διάρκεια μίας χρονοθυρίδας να μεταπεμφθεί στο ΣΒ b , δεδομένης της τρέχουσας θέσης του. Επίσης, μπορεί να ληφθεί υπόψη, για κάθε πιθανό στιγμιότυπο n ενός τύπου εφαρμογής c , η πιθανότητα $\rho_{i,c}(n)$ να είναι ενεργό στη διάρκεια της χρονοθυρίδας (δηλαδή πόσο πιθανό είναι να υπάρχουν 1, 2, ... ενεργά στιγμιότυπα ενός τύπου εφαρμογής). Έτσι, η σημαντικότητα ενός στιγμιότυπου n του τύπου εφαρμογής c που είναι υποψήφιο για μεταπομπή από το ΣΒ a στο ΣΒ b διαμορφώνεται ως $\rho_{HO,i}(a,b) \cdot \rho_{i,c}(n) \cdot hcr_{i,c} \cdot g_c(r)$, για το χρήστη i . Ο υπολογισμός από το χρήστη των πιθανοτήτων $\rho_{HO,i}(a,b)$ και $\rho_{i,c}(n)$ αποτελεί θέμα του υποκεφαλαίου 6.6.

Αν κάθε χρήστης δήλωνε ειλικρινά στο ΣΒ τη σημαντικότητα της κάθε του αίτησης, ο δεύτερος θα μπορούσε να εκχωρήσει ένα δίκαιο μερίδιο των πόρων του στον πρώτο. Αν όμως οι χρήστες είναι εγωιστές, δεν έχουν κάποιο λόγο να δηλώσουν αληθή στοιχεία, αν κάτι τέτοιο δεν τους ωφελεί. Στη συγκεκριμένη περίπτωση, αν η αξιοπιστία των χρηστών δεν είναι δυνατόν να αξιολογηθεί, οι χρήστες έχουν συμφέρον να δηλώσουν μεγαλύτερη σημαντικότητα για τις εφαρμογές τους από την πραγματική, καθώς έτσι θα μπορέσουν να πάρουν μεγαλύτερο μερίδιο πόρων. Αν όμως όλοι οι χρήστες επιδεικνύουν παρόμοια συμπεριφορά, το δίκτυο πάλι θα αδυνατεί να ιεραρχήσει τις αιτήσεις ορθά, με αποτέλεσμα τη μη αποτελεσματική διαχείριση πόρων. Είναι, λοιπόν, φανερό ότι για να επιτευχθεί διαχείριση πόρων ευαίσθητη στις ανάγκες του κάθε χρήστη, θα πρέπει να εξασφαλιστεί η αξιοπιστία των δηλώσεων που κάνουν. Προς αυτήν την κατεύθυνση, το προτεινόμενο σχήμα υποστηρίζεται από ένα πλαίσιο δημοπρασιών, στο οποίο οι χρήστες πληρώνουν για να αποκτήσουν ένα ποσό πόρων κι είναι δομημένο με τρόπο που αποκλείει το ενδεχόμενο ψευδών δηλώσεων.

6.4 Διαμοιρασμός Πόρων Μέσω Δημοπρασιών

Γίνεται η υπόθεση ότι κάθε ΣΒ προσφέρει προς πώληση στους ενδιαφερόμενους χρήστες ένα πλήθος από «αντικείμενα». Σε όρους δικτυακών πόρων, κάθε αντικείμενο αντιστοιχεί σε ένα μέρος της χωρητικότητας του ΣΒ, π.χ. 1 Kb/s. Κάθε χρήστης μπορεί να αγοράσει πάνω από ένα αντικείμενο, ώστε να μπορέσει να καλύψει τις ανάγκες των εφαρμογών του. Συνολικά, βέβαια, μπορεί να πωληθεί ένα πλήθος αντικειμένων που δεν ξεπερνά τη συνολική χωρητικότητα του ΣΒ. Η όλη διαδικασία της αγοροπωλησίας δομείται γύρω από ένα πλαίσιο δημοπρασιών πολλαπλών αντικειμένων (multi-object auction) [48].

Κάθε χρήστης, για κάθε εφαρμογή (νέα ή από μεταπομπή), κάνει μία αίτηση-προσφορά προς το ΣΒ. Αυτή η αίτηση περιγράφει το πόσο πληρώνει ο χρήστης για κάθε μονάδα πόρου (αντικείμενο) που δυνητικά θα του εκχωρήσει ο ΣΒ. Για να υπάρχει συμφωνία με το μοντέλο της συνάρτησης ποιότητας υπηρεσίας g_c που έχει οριστεί στην προηγούμενο υποκεφάλαιο (δες εξίσωση (6.1)), ορίζεται παρακάτω η συνάρτηση $b_c(r)$ που εκφράζει το ποσό που θα πλήρωνε ο χρήστης (π.χ. σε \$) για να αποκτήσει από ένα ΣΒ μία μονάδα πόρου, δεδομένου ότι έχει αποκτήσει ήδη $r-1$ μονάδες (δηλαδή, $b_c(r) \sim dg_c(r)/dr$).

$$b_c(r) = \begin{cases} 0, & r < m_c \\ m_c, & r = m_c \\ \frac{1}{r - m_c + 1}, & m_c < r \leq M_c \\ 0, & r > M_c \end{cases}, \quad \forall c \in C, \text{ και } r = 0, 1, \dots \quad (6.2)$$

Συμφωνά με τη (6.2), ένας χρήστης δε θα πλήρωνε για πόρους που δε θα κάλυπταν τις ελάχιστες ανάγκες της εφαρμογής. Αν ήταν δυνατόν να αποκτήσει τους ελάχιστους δυνατούς πόρους (m_c), θα πλήρωνε μία χρηματική μονάδα ανά αντικείμενο, δηλαδή m_c μονάδες συνολικά. Πέραν της ελάχιστης χωρητικότητας, το ποσό που θα πλήρωνε για κάθε επιπλέον μονάδα φθίνει (καθώς φθίνει κι η αξία που έχει για το χρήστη σε ό,τι αφορά την ποιότητα υπηρεσίας). Για πόρους, πέραν του μέγιστου ποσού M_c που χρειάζεται η εφαρμογή, ο χρήστης δε θα πλήρωνε, καθώς δεν του προσφέρουν βελτίωση στην ποιότητα υπηρεσίας.

Η συνάρτηση b_c αποτελεί τη βασική συνάρτηση προσφοράς των χρηστών για μία εφαρμογή τύπου c και διαμορφώνεται από την ποιότητα υπηρεσίας του

συγκεκριμένου τύπου. Σε περίπτωση μίας νέας εφαρμογής, ένας χρήστης θα έκανε μία τέτοια προσφορά, αν η συγκεκριμένη εφαρμογή ήταν απόλυτα κρίσιμη για αυτόν, δηλαδή αν $ncr_{i,c} = 1$. Στη γενική, όμως, περίπτωση, η προσφορά του χρήστη i $b_{i,c}(r)$, για r μονάδες πόρου για μία νέα εφαρμογή τύπου c , διαμορφώνεται ως $b_{i,c}(r) = ncr_{i,c} \cdot b_c(r)$. Παρομοίως, στην περίπτωση εφαρμογής υποψήφιας για μεταπομπή από το ΣΒ a στο ΣΒ b , η προσφορά του χρήστη διαμορφώνεται ως $b_{i,c}(r) = p_{HO,i}(a,b) \cdot p_{i,c}(n) \cdot hcr_{i,c} \cdot b_c(r)$. Με άλλα λόγια, η προσφορά του χρήστη έχει ως βάση την προσφερόμενη ποιότητα υπηρεσίας, αλλά προσαρμόζεται με βάση την κρίσιμότητα των εκάστοτε πόρων σε σχέση με τις προτιμήσεις του συγκεκριμένου χρήστη, αλλά και του πόσο πιθανό είναι να χρειαστούν τελικά οι πόροι αυτοί.

Η κρίσιμότητα των πόρων ενός ΣΒ για κάθε χρήστη, όπως αναφέρεται και πιο πάνω, διαμορφώνεται βάσει των προτιμήσεων του χρήστη, αλλά και των πιθανοτήτων που καθορίζουν το πόσο χρήσιμοι είναι οι συγκεκριμένοι πόροι για το χρήστη. Καθώς, λοιπόν, κάθε «αντικείμενο» έχει διαφορετική αξία για κάθε χρήστη, είναι εύκολο για ένα ΣΒ να ταξινομήσει τις προσφορές (ανά αντικείμενο) και να εκχωρήσει τους διαθέσιμους πόρους στους χρήστες που έχουν κάνει τις μεγαλύτερες προσφορές (μέσω μίας δημοπρασίας). Έτσι, οι πόροι θα διατεθούν στους χρήστες που τους έχουν μεγαλύτερη ανάγκη. Η αξιοπιστία των δηλώσεων των χρηστών (μέσω των προσφορών τους) εξασφαλίζεται μέσω του μηχανισμού δημοπρασίας τύπου Vickrey (δες παράγραφο 3.6), ο οποίος έχει αυτό το βασικό χαρακτηριστικό, ότι δηλαδή οι (λογικοί) χρήστες αναγκάζονται να αποκαλύψουν την αξία που έχουν για αυτούς τα «αντικείμενα» που τους ενδιαφέρουν.

Έστω $\mathbf{b}_{i,c} = (b_{i,c}(1), \dots, b_{i,c}(m_c), \dots, b_{i,c}(M_c))$ το διάνυσμα που αντιστοιχεί σε μία προσφορά του χρήστη i για μία εφαρμογή τύπου c , όπου το $b_{i,c}(r)$ είναι όπως ορίζεται στην εξίσωση (6.2) ($b_{i,c}(r) \geq b_{i,c}(r+1), \forall r = 0, 1, \dots$). Αν υποθεθεί ότι έχουμε ένα σύνολο από τέτοιες προσφορές, ο ΣΒ παίρνει τα στοιχεία από τα διανύσματα που τις εκφράζουν και τα ταξινομεί ως προς την αξία τους, δημιουργώντας ένα συνολικό διάνυσμα προσφορών \mathbf{b} . Στη συνέχεια, δεδομένης της χωρητικότητας που διαθέτει ο ΣΒ (R_j), επιλέγει τα πιο σημαντικά στοιχεία του \mathbf{b} που μπορούν να εξυπηρετηθούν κι εκχωρεί την αντίστοιχη χωρητικότητα στους χρήστες. Η διαδικασία αυτή περιγράφεται στον ψευδοκώδικα πιο κάτω.

```

Input: freeCapacity

for (k=1:length(amountOfResources))
    amountOfResources[k] := 0
    minAmount[k] = bid[k].minAmount
    lost[k] = false
endFor

sumResources := 0
i = 1;
while (sumResources < freeCapacity && i <= length(b))

    if (lost[k] == false)
        relevantBid := b[i].relevantBid
        amountOfResources[relevantBid]++
        sumResources++
    endif

    if (sumResources == freeCapacity)
        for (k=1:length(amountOfResources))
            if (amountOfResources[k] < minAmount[k])
                lost[k] := true
                sumResources -= amountOfResources[k]
                amountOfResources[k] := 0
            endif
        endFor
    endif
    i++
endWhile

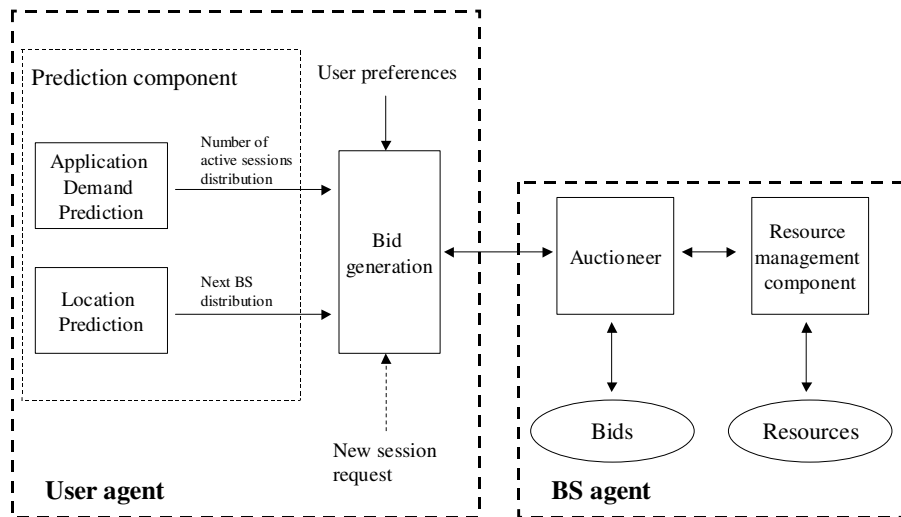
```

Εικόνα 6.2. Δημοπρασία για την εκχώρηση πόρων σε προσφορές χρηστών

Στον ψευδοκώδικα φαίνεται η δημοπρασία που γίνεται, όταν έχουν υποβληθεί k προσφορές. Οι προσφορές αυτές αποσυντίθενται κι οι επιμέρους προσφορές (ανά αντικείμενο) τοποθετούνται σε ένα ταξινομημένο διάνυσμα. Το διάνυσμα αυτό σαρώνεται κι αυξάνεται σταδιακά το πλήθος των πόρων που εκχωρούνται στις διάφορες εφαρμογές. Αν σε κάποιο σημείο δεν υπάρχουν άλλοι διαθέσιμοι πόροι, γίνεται έλεγχος, αν κάποιες προσφορές δεν έχουν επιτύχει να αποκτήσουν το ελάχιστο δυνατό ποσό πόρων που απαιτούν. Αν υπάρχουν τέτοιες προσφορές, οι αντίστοιχες εφαρμογές δε θα μπορέσουν να λειτουργήσουν, με συνέπεια οι εν λόγω πόροι να μείνουν αχρησιμοποίητοι. Γι' αυτό το λόγο, οι πόροι που έχουν εκχωρηθεί σε μία προσφορά, αν δε φτάνουν τους ελάχιστους απαραίτητους, αφήνονται ελεύθεροι, ώστε να εκχωρηθούν σε άλλες εφαρμογές, κι η συγκεκριμένη προσφορά δε λαμβάνεται πλέον υπόψη (οι επιμέρους προσφορές παρακάμπτονται).

6.5 Αρχιτεκτονική και Λειτουργία του Πλαισίου Διαχείρισης Πόρων

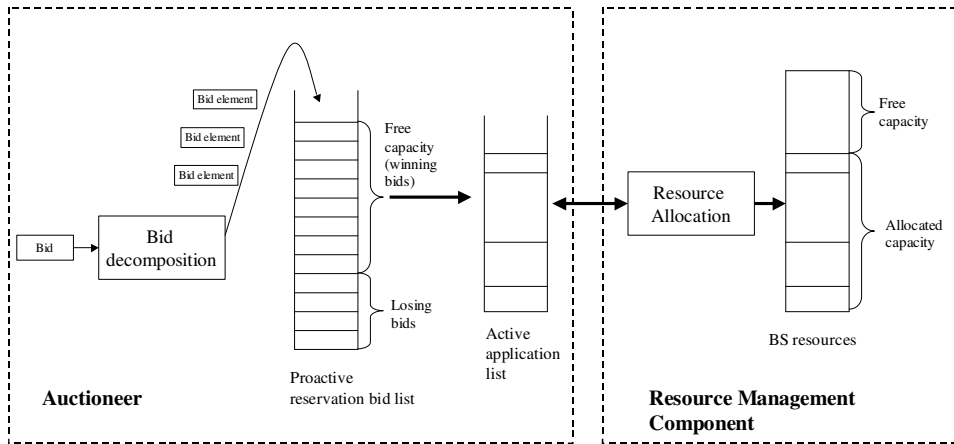
Η αρχιτεκτονική του προτεινόμενου σχήματος διαχείρισης πόρων φαίνεται στην Εικόνα 6.3. Το τερματικό του χρήστη είναι σε θέση να υπολογίζει πιθανότητες που αφορούν στην κίνηση του (για την πρόβλεψη του ΣΒ που πρόκειται να συμβεί μία μεταπομπή), καθώς και στη συμπεριφορά των εφαρμογών του (για την πρόβλεψη του πόσες ενεργές εφαρμογές είναι πιθανό να υφίστανται τη στιγμή μίας πιθανής μεταπομπής). Οι εκτιμήσεις αυτές, καθώς και πληροφορίες σχετικά με τις προτιμήσεις του χρήστη, τροφοδοτούν το συστατικό δημιουργίας προσφορών (bid generation), το οποίο όπως προδίδει και το όνομα του είναι υπεύθυνο να παράγει προσφορές εκ μέρους του χρήστη στους ΣΒ που του υποδεικνύει ο ΑΠΚ. Η αξία των προσφορών διαμορφώνεται λαμβάνοντας υπόψη τις προβλέψεις που παράγονται, αλλά και τις προτιμήσεις που έχει δηλώσει ο χρήστης. Κάθε ΣΒ είναι υπεύθυνος να υποστηρίξει τις διαδικασίες διαμοιρασμού των πόρων, μέσω δημοπρασιών, και να κάνει τις απαραίτητες δεσμεύσεις πόρων.



Εικόνα 6.3. Βασική αρχιτεκτονική του προτεινόμενου σχήματος διαχείρισης πόρων

Η αρχιτεκτονική του ΣΒ μπορεί να αναλυθεί περαιτέρω σε πιο λεπτομερή συστατικά, όπως αυτά φαίνονται στην Εικόνα 6.4. Όταν ο *Auctioneer* λάβει μία προσφορά, την αποσυνθέτει στις επιμέρους προσφορές που την απαρτίζουν (ανά μονάδα πόρου) και τις εισάγει στη λίστα με τις υπόλοιπες αντίστοιχες προσφορές, διατηρώντας την ταξινόμηση ως προς την αξία τους. Έχοντας δεδομένη την ελεύθερη χωρητικότητα (από το resource management component), μπορεί να εκτελέσει μία δημοπρασία, συγκεντρώνοντας τις πιο σημαντικές προσφορές,

σεβόμενος τον περιορισμό που τίθεται από τη διαθέσιμη χωρητικότητα. Συγκεντρώνοντας τις επιμέρους προσφορές (που συγκαταλέγονται σε αυτές που έχουν νικήσει – winning bids), υπολογίζει για κάθε βασική προσφορά το πλήθος των πόρων που έχει αποκτήσει. Την ουσιαστική εκχώρηση πόρων πραγματοποιεί το resource management component που απεικονίζει το αποτέλεσμα της δημοπρασίας στους πραγματικούς πόρους του ΣΒ.



Εικόνα 6.4. Αρχιτεκτονική του BS agent

Στις επόμενες παραγράφους περιγράφεται η λειτουργία του προτεινόμενου σχήματος διαχείρισης πόρων σε διάφορες περιπτώσεις, όπως κατά την έναρξη μίας χρονοθυρίδας, την εμφάνιση μίας νέας εφαρμογής, και τη μεταπομπή.

6.5.1 Δημοπρασία κατά την Έναρξη Χρονοθυρίδας

Όπως αναφέρεται και στην προηγούμενη παράγραφο, στην αρχή κάθε χρονοθυρίδας, κάθε ΣΒ συγκεντρώνει από τους ενδιαφερόμενους χρήστες προσφορές για δέσμευση πόρων. Με βάση αυτές τις προσφορές, ο ΣΒ μπορεί να πραγματοποιήσει μία δημοπρασία, ώστε να διαμοιράσει ανάλογα τους διαθέσιμους πόρους. Στο προτεινόμενο σχήμα, πραγματοποιείται μία δημοπρασία, στην αρχή κάθε χρονοθυρίδας, στην οποία συμμετέχουν όλες οι προσφορές που έχουν συγκεντρωθεί για εφαρμογές υποψήφιας προς μεταπομπή, αλλά κι οι προσφορές που αφορούν σε εφαρμογές που είχαν ενεργοποιηθεί σε κάποια προηγούμενη χρονοθυρίδα. Η δημοπρασία εκτελείται όπως περιγράφεται και στην Εικόνα 6.2, δίνοντας ως είσοδο τη συνολική χωρητικότητα του ΣΒ, R_j . Παρόλο που στην εν λόγω δημοπρασία λαμβάνονται υπόψη οι προσφορές για εφαρμογές που είναι υποψήφιας προς μεταπομπή, για τη συγκεκριμένη χρονοθυρίδα, ο σκοπός της

είναι μόνο να επαναδιαπραγματευθεί το ποσό των πόρων που θα εκχωρηθεί στις ήδη ενεργές εφαρμογές. Έτσι, είναι δυνατόν να προσαρμόζεται η κατανάλωση πόρων των εφαρμογών στις συνθήκες ζήτησης που προκύπτουν σε κάθε νέα χρονοθυρίδα, ώστε, αν η ζήτηση, λόγω πιθανών μεταπομπών, για πόρους αυξηθεί, οι ενεργές εφαρμογές να πάρουν λιγότερους πόρους.

```
sumActive := 0;
for (k=1:length(amountOfResources))
  if (lost[k] == true && bid[k].isActive == true)
    lost[k] := false
    amountOfResources[k] := bid[k].minAmount
    sumActive += amountOfResources[k]
  endif
endFor

if (sumActive > freeCapacity)
  for (k=1:length(amountOfResources))
    if (bid[k].isActive)
      oldAmount := amountOfResources[k]
      newAmount := bid[k].minAmount
      amountOfResources[k] := new_amount
      sumActive := sumActive - (oldAmount - newAmount)
      if (sumActive < freeCapacity)
        newAmount := (freeCapacity-sumActive) + newAmount
        amountOfResources[k] := new_amount
        break
      endif
    endif
  endFor
endIf
endForendIf
```

Εικόνα 6.5. Εκχώρηση πόρων στις ενεργές εφαρμογές που έχουν αποτύχει στη δημοπρασία

Πρέπει να σημειωθεί ότι, δεδομένου του ότι οι πόροι του ΣΒ είναι πεπερασμένοι, κάποιες από τις ήδη ενεργές εφαρμογές μπορεί να μην καταφέρουν να κερδίσουν πόρους στην δημοπρασία (δες Εικόνα 6.2). Στο προτεινόμενο σχήμα, όμως, αν μία νέα εφαρμογή γίνει δεκτή, θα παραμείνει ενεργή μέχρι να τερματίσει φυσικά (εκτός, βέβαια, αν γίνει κάποια αποτυχημένη μεταπομπή). Έτσι, στην αρχή κάθε χρονοθυρίδας, όπου διενεργείται η δημοπρασία, αν μία ενεργή εφαρμογή δεν κερδίσει, δε διακόπτεται. Αντίθετα, αφήνεται να συνεχίσει τη λειτουργία της, αλλά οι πόροι που λαμβάνει είναι οι ελάχιστοι που απαιτεί (δηλαδή m_c για μία εφαρμογή τύπου c). Επομένως, στο τέλος της δημοπρασίας, γίνεται ο έλεγχος, για κάθε προσφορά, αν πρόκειται για ενεργή εφαρμογή κι αν, επίσης, η προσφορά δεν είναι

επιτυχής. Αν ισχύει κάτι τέτοιο, η προσφορά αυτή δε θεωρείται πλέον αποτυχημένη και λαμβάνει το ελάχιστο πλήθος πόρων, όπως φαίνεται και στην Εικόνα 6.5.

Αφού ολοκληρωθεί η εκχώρηση πόρων στις ενεργές εφαρμογές, οι αντίστοιχες προσφορές απομακρύνονται και, για το υπόλοιπο της χρονοθυρίδας, διατηρούνται μόνο οι προσφορές που αφορούν σε εφαρμογές που είναι υποψήφιες για μεταπομπή. Επίσης, οι πόροι που εκχωρούνται στις ενεργές εφαρμογές διατηρούνται καθ' όλη τη διάρκεια της επικείμενης χρονοθυρίδας. Οι πόροι θα επαναδιαπραγματευθούν μόνο στην αρχή της επόμενης χρονοθυρίδας.

Επιπλέον της δημοπρασίας που διενεργείται στην αρχή κάθε χρονοθυρίδας για την εκχώρηση πόρων στις ενεργές εφαρμογές, δημοπρασίες μπορούν να πραγματοποιηθούν και κατά τη διάρκεια της χρονοθυρίδας. Συγκεκριμένα, μία δημοπρασία μπορεί να λάβει χώρα, όταν ένας χρήστης θέλει να εκκινήσει μία νέα εφαρμογή, όπως περιγράφεται στην επόμενη παράγραφο.

6.5.2 Δημοπρασία για την Εκκίνηση Νέας Εφαρμογής

Ένας χρήστης, ανά πάσα στιγμή, μπορεί να αιτηθεί από το ΣΒ που τον εξυπηρετεί πόρους για να εκκινήσει μία νέα εφαρμογή. Ο ΣΒ πρέπει να αποφασίσει για το αν θα επιτρέψει μία νέα σύνδεση στο χρήστη κι αν ναι τους πόρους που θα της εκχωρηθούν. Αυτή την απόφαση τη λαμβάνει με δεδομένες τις προσφορές που έχουν κάνει οι υπόλοιποι χρήστες για εφαρμογές που είναι υποψήφιες για μεταπομπή στην αρχή της τρέχουσας χρονοθυρίδας, μέσω μίας δημοπρασίας. Όσο πιο πολλές ή πιο μεγάλες είναι οι προσφορές για εφαρμογές υποψήφιες για μεταπομπή, τόσο πιο δύσκολα θα εκχωρούνται πόροι σε νέες εφαρμογές. Συγκεκριμένα, εισάγει την προσφορά της νέας εφαρμογής στο σύνολο με τις προσφορές για τις υποψήφιες για μεταπομπή εφαρμογές κι εκτελεί μία δημοπρασία, όπως αυτή περιγράφεται στην Εικόνα 6.2, με εισόδο το ποσό ελεύθερων πόρων (διαφορά της συνολικής χωρητικότητας R_j του ΣΒ από το σύνολο των πόρων που έχουν εκχωρηθεί σε ενεργές εφαρμογές).

Μετά την ολοκλήρωση της δημοπρασίας, η προσφορά της νέας εφαρμογής θα έχει κερδίσει ή χάσει. Στην μεν περίπτωση, η εφαρμογή θα πάρει κάποιο ποσό πόρων που εξαρτάται από τον ανταγωνισμό, κατά την τρέχουσα χρονοθυρίδα, θα εισαχθεί στη λίστα των ενεργών εφαρμογών και θα αποκτήσει τους αντίστοιχους πόρους

(δες Εικόνα 6.4). Στη δε, θα απορριφθεί και δε θα πάρει καθόλου πόρους. Και στις δύο περιπτώσεις, όταν η δημοπρασία ολοκληρωθεί, η προσφορά της νέας εφαρμογής αφαιρείται από το σύνολο των προσφορών, ώστε σε αυτό το σύνολο να υπάρχουν μόνο προσφορές για εφαρμογές που είναι υποψήφιες για μεταπομπή.

6.5.3 Δημοπρασία μετά από Μεταπομπή

Σε αυτήν την παράγραφο, περιγράφονται οι διαδικασίες που λαμβάνουν χώρα όταν ένα ΚΤ μεταπέμπεται σε ένα νέο ΣΒ. Όπως έχει προαναφερθεί, κάθε χρήστης, στην αρχή κάθε χρονοθυρίδας, βάσει προβλέψεων σε σχέση με το πόσο πιθανό είναι να κάνει μεταπομπή σε κάποιο ΣΒ, στη διάρκεια της χρονοθυρίδας, αλλά και να έχει ένα, δύο, κοκ ενεργά στιγμιότυπα για κάθε τύπο εφαρμογής, κάνει αντίστοιχες προσφορές στο δίκτυο. Έτσι, όταν λάβει χώρα μία μεταπομπή, αξιολογούνται οι προσφορές του αντίστοιχου χρήστη, σε σχέση με τις υπόλοιπες προσφορές που έχουν υποβληθεί, κι αποφασίζεται αν θα δοθούν πόροι σε αυτές κι αν ναι πόσοι.

Η διαδικασία με την οποία καθορίζεται το ποσό των πόρων που θα πάρει ένας χρήστης, όταν μεταπέμπεται σε ένα νέο ΣΒ, είναι μία δημοπρασία, όπως αυτή φαίνεται στην Εικόνα 6.2, με είσοδο την τρέχουσα ελεύθερη χωρητικότητα στο ΣΒ. Το αποτέλεσμα θα είναι πιθανώς κάποιες από τις προσφορές να είναι επιτυχημένες, ενώ κάποιες άλλες όχι. Όπως φαίνεται και στην Εικόνα 6.4, οι επιτυχημένες προσφορές δημιουργούν αντίστοιχες καταχωρήσεις στη λίστα των ενεργών εφαρμογών. Πρέπει να σημειωθεί ότι το πλήθος των προσφορών που νίκησαν μπορεί να μη συμπίπτει με το πλήθος των ενεργών εφαρμογών του χρήστη τη στιγμή της μεταπομπής. Συγκεκριμένα, μπορεί να ξεπερνάει το πλήθος των ενεργών εφαρμογών, ή το αντίστροφο. Στη μεν περίπτωση, όλες οι εφαρμογές του χρήστη συνεχίζουν απρόσκοπτα τη λειτουργία τους, ενώ οι πόροι που περισσεύουν αποδεδμεύονται από το χρήστη. Στη δε περίπτωση, κάποιες εφαρμογές θα πρέπει να διακοπούν. Σε κάθε περίπτωση, όσες εφαρμογές κατάφεραν να αποκτήσουν πόρους κατά τη μεταπομπή θεωρούνται κι αντιμετωπίζονται πλέον ως ενεργές εφαρμογές από το νέο ΣΒ. Επίσης, οι προσφορές που αντιστοιχούσαν στο χρήστη που έκανε τη μεταπομπή αφαιρούνται

από τη λίστα του *Auctioneer*, αφήνοντας έτσι μόνο προσφορές χρηστών που ακόμα αναμένονται να μεταπεμφούν.

6.5.4 Χρέωση

Το ποσό που θα πληρώσει κάποιος χρήστης για τη χρήση / δέσμευση πόρων του δικτύου (για μια ενεργή σύνδεση, ή ακόμα και για μία προδραστική δέσμευση πόρων) εξαρτάται από το χρόνο που έχει δεσμεύσει τους εν λόγω πόρους, αλλά και την τιμή που έχει προκύψει μέσω της δημοπρασίας. Η τιμή μπορεί να ανανεώνεται σε κάθε χρονοθυρίδα, αλλά καθ' όλη τη διάρκεια μίας χρονοθυρίδας παραμένει σταθερή. Συγκεκριμένα, στην αρχή της χρονοθυρίδας που επαναδιαπραγματεύονται οι πόροι των ενεργών εφαρμογών (παράγραφο 6.5.1), επανακαθορίζεται η τιμή του πόρου για τη νέα χρονοθυρίδα. Στην περίπτωση μίας νέας εφαρμογής, ο ΣΒ διενεργεί μία δημοπρασία για να αποφανθεί αν η νέα εφαρμογή θα γίνει δεκτή και πόσους πόρους θα πάρει (παράγραφο 6.5.2). Αν η νέα εφαρμογή γίνει δεκτή, θα πληρώσει ανάλογα με την προσφορά που έχει κάνει και τους πόρους που της εκχωρήθηκαν για όλο το υπόλοιπο της χρονοθυρίδας από τη στιγμή που έγινε δεκτή. Η ίδια λογική ισχύει και για τις εφαρμογές που γίνονται δεκτές από το νέο ΣΒ μετά από μία μεταπομπή (παράγραφο 6.5.3).

Το ποσό που πληρώνει ο χρήστης για κάθε εφαρμογή (είτε πρόκειται για μία ήδη ενεργή, είτε για μία νέα που γίνεται δεκτή, είτε για μία εφαρμογή από μεταπομπή) καθορίζεται από τον ανταγωνισμό. Συγκεκριμένα, εφαρμόζεται ο κανόνας των δημοπρασιών Vickrey για πολλά αντικείμενα [48]. Σύμφωνα με τον κανόνα αυτό, κάθε εφαρμογή που πέτυχε να αποκτήσει r μονάδες από τους πόρους, μετά από μια δημοπρασία, θα πληρώσει όσο αξίζει το άθροισμα των r πιο σημαντικών προσφορών που δεν έχουν κερδίσει (που θα είχαν κερδίσει αν δεν υπήρχε η συγκεκριμένη εφαρμογή) [48].

Όταν ολοκληρωθεί μία χρονοθυρίδα, πιθανώς, κάποιες προσφορές να έχουν παραμείνει στο ΣΒ. Αυτό σημαίνει ότι οι αντίστοιχοι χρήστες τελικά δεν εμφανίστηκαν κατά τη διάρκεια της χρονοθυρίδας, δηλαδή δεν συνέβη η αναμενόμενη μεταπομπή. Παρόλ' αυτά ο ΣΒ, όπως όφειλε, διατήρησε τις συγκεκριμένες προσφορές για όλη τη διάρκεια της χρονοθυρίδας, με αυτές τις προσφορές να έχουν ένα μερίδιο στις αποφάσεις που ελήφθησαν (ποιες νέες

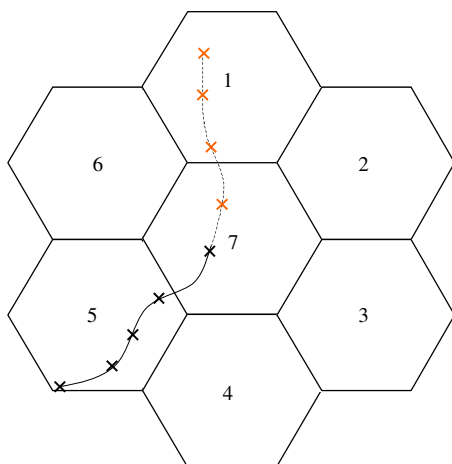
εφαρμογές να απορριφθούν, ποιες από μεταπομπή να διακοπούν κι όσες πέρασαν πόσους πόρους να πάρουν). Έτσι, στο τέλος της χρονοθυρίδας, διενεργείται μία τελευταία δημοπρασία, ώστε οι εναπομένουσες προσφορές να χρεωθούν για τους πόρους που κατέλαβαν στη διάρκεια της χρονοθυρίδας.

6.6 Πλαίσιο Πρόβλεψης για Προδραστική Δέσμευση Πόρων

Βασικό χαρακτηριστικό της προτεινόμενης αρχιτεκτονικής αποτελεί ότι ο χρήστης κάνει προδραστικές δεσμεύσεις πόρων, με σκοπό την αποφυγή διακοπών, λόγω μεταπομπής. Βασικό ρόλο στην όλη διαδικασία παίζουν οι προβλέψεις που κάνει ο χρήστης σε ό,τι αφορά την κίνηση του, αλλά και τη συμπεριφορά των εφαρμογών του. Το θέμα του παρόντος υποκεφαλαίου είναι οι μέθοδοι που χρησιμοποιούνται για τις προαναφερθείσες προβλέψεις.

6.6.1 Πρόβλεψη Κινητικότητας Χρήστη

Βασικό πρόβλημα στην προδραστική δέσμευση πόρων από το χρήστη αποτελεί η πρόβλεψη της θέσης του στο προσεχές μέλλον. Συγκεκριμένα, δεδομένου του ότι οι αιτήσεις-προσφορές των χρηστών λαμβάνουν χώρα περιοδικά, κάθε T sec, στην αρχή κάθε χρονοθυρίδας, κάθε χρήστης χρειάζεται να υπολογίσει την πιθανότητα να κάνει μία μεταπομπή σε κάποιο ΣΒ, στη διάρκεια της επικείμενης χρονοθυρίδας (δηλαδή μέσα στα επόμενα T sec), ώστε οι προσφορές να διαμορφώνονται αναλόγως. Γίνεται η υπόθεση ότι ο χρήστης γνωρίζει ανά πάσα στιγμή τη θέση του $\mathbf{p}_i(t) = (x_i(t), y_i(t))$ και μπορεί να διατηρεί την ιστορία της κίνησης του. Για ευκολία, γίνεται η θεώρηση ότι ο χρήστης ελέγχει τη θέση του περιοδικά, κάθε m sec, όπου το T είναι ακέραιο πολλαπλάσιο του m . Η Εικόνα 6.6 δείχνει μία ενδεικτική κατάσταση του προβλήματος. Η καμπύλη που ξεκινά από την κυψέλη 5 και περνάει στην κυψέλη 7, δείχνει την κίνηση του χρήστη μέχρι την τρέχουσα χρονική στιγμή. Η διακεκομμένη καμπύλη που συνεχίζει από την κυψέλη 7 και καταλήγει στην κυψέλη 1 δείχνει την κίνηση που προβλέπεται να κάνει ο χρήστης στο προσεχές μέλλον.



Εικόνα 6.6. Κίνηση χρήστη σε ένα κυψελωτό δίκτυο

Το πρόβλημα, λοιπόν, για το χρήστη i , είναι να υπολογίσει την πιθανότητα $\rho_{HO,i}$ να κάνει μία μεταπομπή μέσα στα επόμενα T sec, δεδομένης της τρέχουσας θέσης του, σε κάποιον ΣΒ b (για κάθε πιθανό ΣΒ b):

$$\rho_{HO,i}(b, n \cdot T) = \Pr[\exists k \in \{1, \dots, T/m\}, \mathbf{p}_i(k \cdot m + n \cdot T) \in \text{area}(b) \mid \mathbf{p}_i(0), \dots, \mathbf{p}_i(n \cdot T)] \quad (6.3)$$

Με άλλα λόγια, ο χρήστης πρέπει να υπολογίσει την πιθανότητα, δεδομένης της ιστορίας της κίνησης του, έστω μία από τις θέσεις που θα έχει μέσα στη χρονοθυρίδα διάρκειας T να βρίσκεται στην περιοχή που καλύπτει ο ΣΒ b ($\text{area}(b)$). Καθώς η πιθανότητα αυτή βασίζεται σε μελλοντικές θέσεις του χρήστη, δεν είναι δυνατός ο ακριβής υπολογισμός της, αλλά μόνο προσεγγίσεις είναι δυνατόν να επιτευχθούν. Συγκεκριμένα, είναι αναγκαίο να γίνουν διάφορες υποθέσεις σε ό,τι αφορά την κίνηση του χρήστη, ώστε να εκτιμηθούν οι μελλοντικές του θέσεις στη διάρκεια της επικείμενης χρονοθυρίδας και βάσει αυτών να υπολογιστεί τελικά η αναζητούμενη πιθανότητα.

Στην παρούσα μελέτη, προτείνεται μία μέθοδος που βασίζεται στον υπολογισμό μίας απλοποιημένης μορφής της πιθανότητας που ορίζεται στην εξίσωση (6.3). Συγκεκριμένα, δίνεται βάρος μόνο σε δύο από τις πιο σημαντικές παραμέτρους της κίνησης του χρήστη. Οι παράμετροι που λαμβάνονται υπόψη είναι η μέγιστη απόσταση $r_{max,i}$ στην οποία μπορεί να βρεθεί ο χρήστης στη διάρκεια μίας χρονοθυρίδας σε σχέση με το σημείο που βρίσκεται, όταν ξεκινά η χρονοθυρίδα,

καθώς κι οι κατευθύνσεις θ_i που ακολουθεί ο χρήστης στη διάρκεια μίας χρονοθυρίδας, δεδομένης της κατεύθυνσης που είχε στην αρχή της χρονοθυρίδας.

Οι παράμετροι $r_{max,i}$ και θ_i αποτελούν τυχαίες μεταβλητές με συγκεκριμένα χαρακτηριστικά και μπορούν να εξαχθούν από την ιστορία της κίνησης του χρήστη. Συγκεκριμένα, η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής $r_{max,i}$ για το χρήστη i , $p_{i,r_{max}}(r)$, ορίζεται ως εξής:

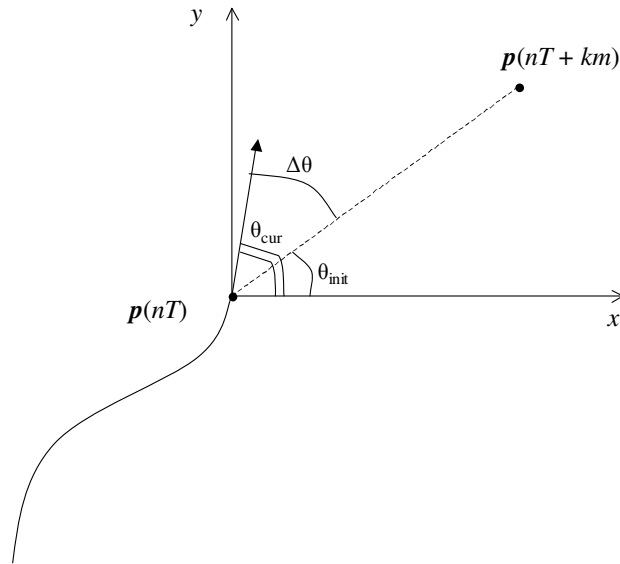
$$p_{i,r_{max}}(r) = \Pr[r_{max,i} = r] = \Pr\left[\max_{k=1,\dots,T/m} \{\|\mathbf{p}_i(nT + km) - \mathbf{p}_i(nT)\|\} = r, n = 0,1,\dots\right], \quad (6.4)$$

όπου το $\|\mathbf{p} - \mathbf{p}'\|$ εκφράζει την Ευκλείδεια απόσταση από το σημείο \mathbf{p} στο σημείο \mathbf{p}' .

Αντίστοιχα, η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής θ_i για το χρήστη i , $p_{i,\theta}(\theta)$, ορίζεται ως εξής:

$$p_{i,\theta}(\theta) = \Pr[\theta_i = \theta] = \Pr[\text{orient}(\mathbf{p}_i(nT + km), \mathbf{p}_i(nT), \theta_i(nT)) = \theta, n = 0,1,\dots], \quad (6.5)$$

όπου το $\text{orient}(\mathbf{p}, \mathbf{p}', \theta)$ εκφράζει τη γωνία μεταξύ της ευθείας που ενώνει τα σημεία \mathbf{p} και \mathbf{p}' με την ευθεία που έχει κατεύθυνση θ . Το μέγεθος $\theta_i(t)$ υποδηλώνει την κατεύθυνση του χρήστη τη χρονική στιγμή t κι ορίζεται ως η γωνία που αντιστοιχεί στην ευθεία που ορίζεται από το σημείο $\mathbf{p}_i(t-m)$ και το σημείο $\mathbf{p}_i(t)$.

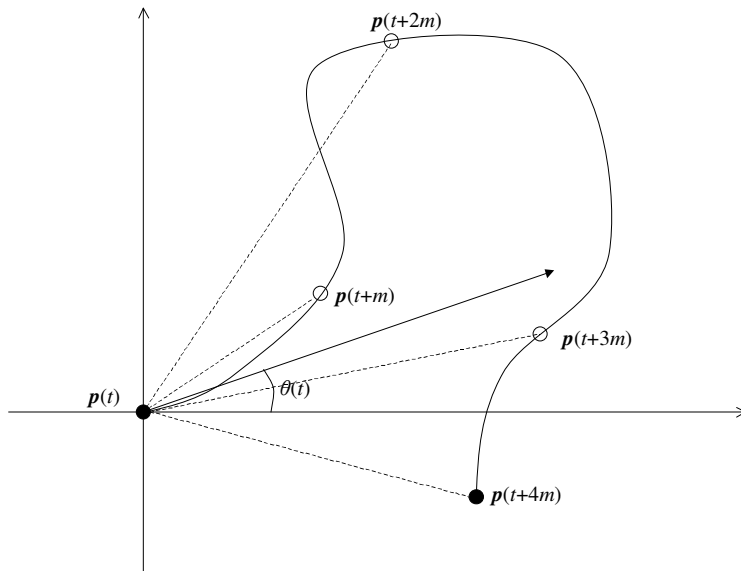


Εικόνα 6.7. Υπολογισμός παραμέτρων κίνησης χρήστη

Στην Εικόνα 6.7, φαίνεται η διαδικασία υπολογισμού της κατεύθυνσης $\text{orient}(\cdot; \cdot; \cdot)$ ενός σημείου $\mathbf{p}(nT + km)$, δεδομένης της θέσης του χρήστη στην αρχή της n -οστής

χρονοθυρίδας, $\mathbf{p}(nT)$, και της κατεύθυνσης που είχε τη δεδομένη στιγμή, θ_{init} . Η τρέχουσα κατεύθυνση του χρήστη τη στιγμή $nT + km$ είναι θ_{cur} , επομένως, η αναζητούμενη κατεύθυνση $orient(\cdot, \cdot, \cdot)$ θα είναι $\Delta\theta = \theta_{cur} - \theta_{init}$.

Οι πιθανότητες $p_{i,rmax}(r)$ και $p_{i,\theta}(\theta)$ μπορούν να υπολογιστούν μέσω παρατήρησης από τον ίδιο το χρήστη, ως εξής: 1) θέτοντας σε κάθε χρονοθυρίδα ως σημείο αναφοράς και κατεύθυνση αναφοράς του χρήστη τα αντίστοιχα μεγέθη τη στιγμή της έναρξης της χρονοθυρίδας και 2) υπολογίζοντας για κάθε θέση που βρίσκεται ο χρήστης στη διάρκεια της χρονοθυρίδας την απόσταση από το σημείο αναφοράς και την αντίστοιχη διαφορά κατεύθυνσης¹⁰. Η διαδικασία αυτή περιγράφεται και στην Εικόνα 6.8.



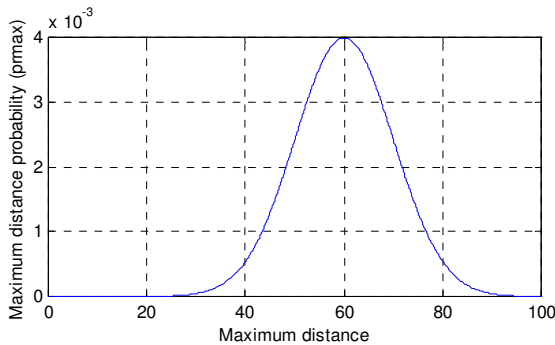
Εικόνα 6.8. Δειγματοληψία χαρακτηριστικών της κίνησης του χρήστη. Στο συγκεκριμένο παράδειγμα, στη διάρκεια μίας χρονοθυρίδας, λαμβάνονται τέσσερα δείγματα. Η μέγιστη απόσταση για την εξεταζόμενη χρονοθυρίδα προκύπτει στη θέση $\mathbf{p}(t+2m)$. Αφαιρώντας τη γωνία $\theta(t)$ (γωνία αναφοράς) από τη γωνία που σχηματίζει η ευθεία που ενώνει ένα σημείο με το σημείο αναφοράς ($\mathbf{p}(t)$) υπολογίζονται και δείγματα κατεύθυνσης

Γίνεται η υπόθεση ότι, δεδομένης μίας κατεύθυνσης θ , αν ένας χρήστης βρεθεί σε μία απόσταση r , τότε θα έχει περάσει κι από όλα τα ενδιάμεσα σημεία στην ίδια κατεύθυνση, από το σημείο αναφοράς μέχρι το εν λόγω σημείο. Με αυτήν την υπόθεση, είναι δυνατόν για μία δεδομένη κατεύθυνση θ να υπολογιστεί η πιθανότητα ο χρήστης να ξεπεράσει μία δεδομένη απόσταση r . Η προαναφερθείσα πιθανότητα $q_i(r, \theta)$ υπολογίζεται ως εξής:

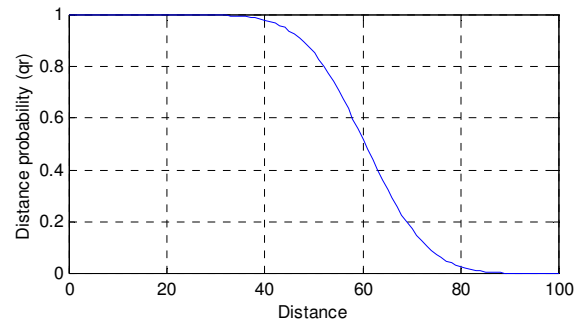
¹⁰ Είναι επίσης εφικτό μέσω παρατήρησης, να υπολογιστεί και η από κοινού συνάρτηση πυκνότητας πιθανότητας $p_{i,rmax,\theta}$. Στην παρούσα μελέτη γίνεται, όμως, η υπόθεση ότι οι τυχαίες μεταβλητές $r_{max,i}$ και θ_i είναι ασυσχέτιστες επομένως δεν μελετάται αυτή η δυνατότητα.

$$q_i(r, \theta) = p_{i,\theta}(\theta) \int_r^\infty p_{i,rmax}(a) da, \quad (6.6)$$

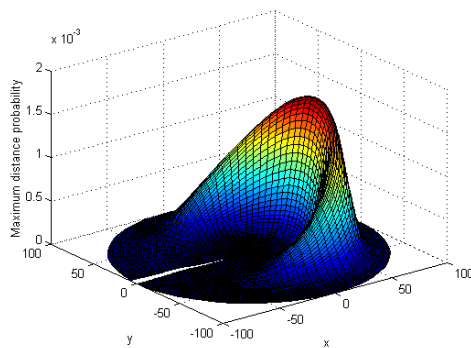
Στην Εικόνα 6.9(a), φαίνεται μία πιθανή μορφή της συνάρτησης πυκνότητας πιθανότητας $p_{i,rmax}$ και, στην Εικόνα 6.9(b), η αντίστοιχη πιθανότητα $q_i(r, \theta)$, για μία οποιαδήποτε κατεύθυνση θ . Στην Εικόνα 6.9(c), φαίνεται η πιθανότητα μέγιστης απόστασης για κάθε κατεύθυνση, ενώ, στην Εικόνα 6.9(d), η πιθανότητα $q_i(r, \theta)$.



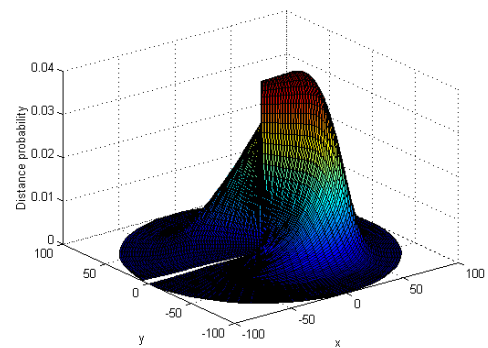
(a)



(b)



(c)



(d)

Εικόνα 6.9. Παράδειγμα κατανομών χαρακτηριστικών κίνησης του χρήστη. (a) $p_{i,rmax}(r)$ (b) $q_i(r, \theta)$ (c) $p_{i,rmax}(r, \theta)$ (d) $q_i(r, \theta)$

Η συνάρτηση $q_i(r, \theta)$ είναι χρήσιμη σε ό,τι αφορά την παρούσα μελέτη γιατί δίνει τη δυνατότητα υπολογισμού της πιθανότητας ο χρήστης, σε μία συγκεκριμένη κατεύθυνση να περάσει μία συγκεκριμένη νοητή γραμμή. Αυτή η νοητή γραμμή μπορεί να ορίζει μία νέα κυψέλη, επομένως μπορεί να χρησιμοποιηθεί στον υπολογισμό της πιθανότητας ο χρήστης να κάνει μία μεταπομπή. Για να μπορέσει να γίνει αυτό θα πρέπει πρώτα να ληφθεί υπόψη το σημείο κι η κατεύθυνση αναφοράς ($\mathbf{p}_{ref} = (x_{ref}, y_{ref})$ και θ_{ref}), του χρήστη, ώστε η παραπάνω πιθανότητα να μπορεί να δώσει τιμές για πραγματικές θέσεις στο χώρο που κινείται ο χρήστης. Συγκεκριμένα, μπορεί να χρησιμοποιηθεί μία συνάρτηση-μετασχηματισμός, $\mathbf{T} : [0,$

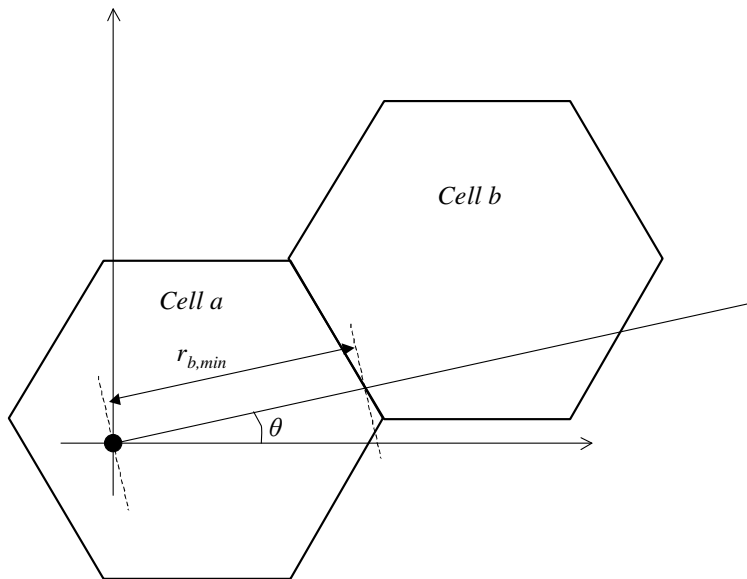
$2\pi] \times \mathbb{R}^+ \times \mathbb{R}^2 \times [0, 2\pi] \rightarrow \mathbb{R}^2$, για τη μετατροπή από πολικές σε καρτεσιανές συντεταγμένες, δεδομένου ενός σημείου και κατεύθυνσης αναφοράς, ως εξής:

$$\begin{aligned} T_x(\theta, r, x_{ref}, y_{ref}, \theta_{ref}) &= r \cdot \sin(\theta + \theta_{ref}) + x_{ref} \\ T_y(\theta, r, x_{ref}, y_{ref}, \theta_{ref}) &= r \cdot \cos(\theta + \theta_{ref}) + y_{ref} \end{aligned} \quad (6.7)$$

Έτσι, με τη βοήθεια της εξίσωσης (6.6) και (6.7), είναι εύκολο να υπολογιστεί η πιθανότητα $\rho_{i,b}(\theta, x_{ref}, y_{ref}, \theta_{ref})$ ένας χρήστης να βρεθεί σε μία κυψέλη b , θεωρώντας μία συγκεκριμένη κατεύθυνση θ , ως εξής:

$$\begin{aligned} \rho_{i,b}(\theta, x_{ref}, y_{ref}, \theta_{ref}) &= q_i(r_{b,min}, \theta), \\ r_{b,min} &= \min\{r \in [0, +\infty) : T(\theta, r, x_{ref}, y_{ref}, \theta_{ref}) \in area(b)\} \end{aligned} \quad (6.8)$$

Η απόσταση $r_{b,min}$ αντιστοιχεί στην ελάχιστη απόσταση στην κατεύθυνση θ μέχρι την περιοχή που καλύπτει ο ΣΒ b (το πρώτο σημείο τομής της ευθείας από το σημείο αναφοράς με την εν λόγω περιοχή), όπως φαίνεται και στην Εικόνα 6.10.



Εικόνα 6.10. Εύρεση ελάχιστης απόστασης από ένα ΣΒ για κατεύθυνση θ .

Έχοντας, λοιπόν, υπολογίσει την πιθανότητα $\rho_{i,b}(\theta, x_{ref}, y_{ref}, \theta_{ref})$ ο χρήστης i να κάνει μεταπομπή στο ΣΒ b στη διάρκεια μίας χρονοθυρίδας (με θέση στην αρχή της χρονοθυρίδας τη (x_{ref}, y_{ref}) και κατεύθυνση θ_{ref}), για συγκεκριμένη κατεύθυνση θ , είναι εφικτό να υπολογιστεί η πιθανότητα επίσκεψης στο ΣΒ b για κάθε πιθανή κατεύθυνση. Αν ονομάσουμε την πιθανότητα αυτή $\rho_{i,b}(x_{ref}, y_{ref}, \theta_{ref})$, μπορούμε να την υπολογίσουμε εύκολα ως

$$\rho_{i,b}(x_{ref}, y_{ref}, \theta_{ref}) = \int_0^{2\pi} \rho_{i,b}(\theta, x_{ref}, y_{ref}, \theta_{ref}) d\theta. \quad (6.9)$$

Στη συζήτηση πιο πάνω, έχει γίνει η υπόθεση ότι οι μεταβλητές $r_{i,max}$ και θ_i είναι συνεχείς τυχαίες μεταβλητές κι ως τέτοιες αντιμετωπίστηκαν στην ανάλυση. Επίσης, υποτέθηκε ότι η μεταβλητή $r_{i,max}$ μπορεί να πάρει οσοδήποτε μεγάλες τιμές. Πρέπει να σημειωθεί, όμως, ότι στην πράξη θα ήταν πιο απλή και πρακτική η χρήση διακριτών και φραγμένων τυχαίων μεταβλητών. Για το λόγο αυτό, θεωρούμε ότι η μεταβλητή $r_{i,max}$ λαμβάνει τιμές από το διάστημα $[0, R_{max}]$. Επίσης, το διάστημα $[0, R_{max}]$ χωρίζεται σε ίσα υποδιαστήματα μήκους Δr και το διάστημα $[0, 2\pi]$ σε υποδιαστήματα μήκους $\Delta\theta$. Έτσι, σε αντιστοιχία με τις συνεχείς τυχαίες μεταβλητές, ορίζονται οι παρακάτω μεταβλητές:

$$\hat{\rho}_{i,r_{max}}(l) = \int_{l \cdot \Delta r}^{(l+1) \cdot \Delta r} \rho_{i,r_{max}}(a) da, \text{ και}$$

$$\hat{\rho}_{i,\theta}(k) = \int_{k \cdot \Delta\theta}^{(k+1) \cdot \Delta\theta} \rho_{i,\theta}(\theta) d\theta$$

Επομένως, το αποτέλεσμα της πρόβλεψης $\rho_{i,b}(x_{ref}, y_{ref}, \theta_{ref}), \forall b \in I$, με τις πιο πάνω παραδοχές (διακριτές και φραγμένες τυχαίες μεταβλητές), βάσει των εξισώσεων (6.6), (6.8) και (6.9), μετατρέπεται στην ακόλουθη μορφή:

$$\hat{\rho}_{i,b}(x_{ref}, y_{ref}, \theta_{ref}) = \sum_{k=0}^{2\pi / \Delta\theta} \hat{\rho}_{i,b}(k, x_{ref}, y_{ref}, \theta_{ref}), \quad (6.10)$$

όπου

$$\hat{\rho}_{i,b}(k, x_{ref}, y_{ref}, \theta_{ref}) = \hat{q}_i(r_{b,min}, k),$$

$$r_{b,min} = \min\{l \cdot \Delta r, l = 0, 1, \dots, R_{max} / \Delta r : T(k, l, x_{ref}, y_{ref}, \theta_{ref}) \in area(b)\}$$

και

$$\hat{q}_i(l, k) = \hat{\rho}_{i,\theta}(k) \sum_{i=l}^{R_{max} / \Delta r} \hat{\rho}_{i,r_{max}}(i).$$

Συνοψίζοντας, λοιπόν, ο χρήστης παρατηρώντας την κίνηση του, υπολογίζει την κατανομή της μέγιστης απόστασης που μπορεί να μετακινηθεί, δεδομένου ενός σημείου αναφοράς. Υπολογίζει, επίσης, την κατανομή των κατευθύνσεων που μπορεί να βρεθεί, δεδομένου του σημείου και της κατεύθυνσης αναφοράς. Βάσει αυτών των κατανομών, υπολογίζεται η πιθανότητα, για μία συγκεκριμένη

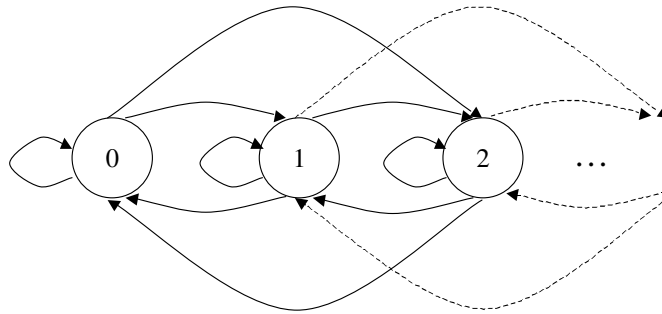
κατεύθυνση, ο χρήστης να ξεπεράσει μία δεδομένη απόσταση από το σημείο αναφοράς. Η πιθανότητα αυτή χρησιμοποιείται στη συνέχεια για να υπολογιστεί η πιθανότητα ο χρήστης να βρεθεί σε κάθε πιθανό ΣΒ, στη συγκεκριμένη κατεύθυνση. Τέλος, εξετάζοντας όλες τις πιθανές κατευθύνσεις, υπολογίζεται η πιθανότητα ο χρήστης να βρεθεί σε κάθε πιθανό ΣΒ. Οι πιθανότητες που προκύπτουν μπορούν να χρησιμοποιηθούν από το χρήστη για τη δημιουργία κατάλληλων προσφορών προς τους αντίστοιχους ΣΒ (δες υποκεφάλαιο 6.4).

6.6.2 Πρόβλεψη Συμπεριφοράς Εφαρμογών

Ένα ακόμα σημαντικό ζήτημα για την προδραστική δέσμευση πόρων από το χρήστη είναι η συμπεριφορά που επιδεικνύουν οι εφαρμογές του. Όπως έχει ήδη συζητηθεί, οι αιτήσεις-προσφορές των χρηστών λαμβάνουν χώρα περιοδικά, κάθε T sec, στην αρχή κάθε χρονοθυρίδας. Έτσι, ο χρήστης χρειάζεται να έχει μία εικόνα της ζήτησης σε πλήθος εφαρμογών που θα έχει τη στιγμή μίας πιθανής μεταπομπής σε κάποιο ΣΒ, στη διάρκεια της επικείμενης χρονοθυρίδας (δηλαδή μέσα στα επόμενα T sec). Μία λύση θα ήταν να βασιστεί στο πλήθος των εφαρμογών που έχει στην αρχή της χρονοθυρίδας. Η λύση αυτή, όμως, δεν είναι ιδιαίτερα ασφαλής, ούτε αποδοτική, καθώς το συγκεκριμένο πλήθος είναι πολύ πιθανό να αυξηθεί, ή να μειωθεί. Οι επιπτώσεις του φαινομένου αυτού είναι πιο έντονες, όσο αυξάνεται η διάρκεια της χρονοθυρίδας. Έτσι, για να μην γίνονται ελλιπείς ή υπερβολικές δεσμεύσεις, χρειάζεται ένα πλαίσιο για την πρόβλεψη της συμπεριφοράς των εφαρμογών του χρήστη.

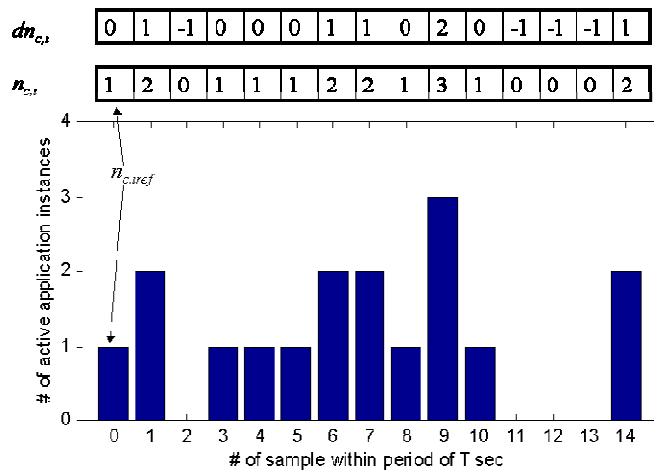
Στο προτεινόμενο σχήμα, ο χρήστης εξετάζει κάθε τύπο εφαρμογής ξεχωριστά, καθώς είναι λογικό εφαρμογές ενός συγκεκριμένου τύπου να έχουν παρόμοια συμπεριφορά. Στην Εικόνα 6.11, φαίνεται το μοντέλο καταστάσεων των εφαρμογών. Θεωρούμε, ότι ο χρήστης ελέγχει το πλήθος των ενεργών στιγμιότυπων κάθε τύπου εφαρμογής περιοδικά. Όπως και στο μοντέλο πρόβλεψης της κινητικότητας (παράγραφος 6.6.1), ο χρήστης ελέγχει κάθε m sec, όπου το T είναι ακέραιο πολλαπλάσιο του m . Η κατάσταση $n \in \{0, 1, \dots\}$ αντιστοιχεί σε n ενεργά στιγμιότυπα. Η μετάβαση $n \rightarrow n'$ υποδηλώνει το ενδεχόμενο, έχοντας n ενεργά στιγμιότυπα κατά τον τελευταίο έλεγχο, να προκύψουν n' στιγμιότυπα στον επόμενο έλεγχο (δηλαδή μετά από m sec). Καθώς η περίοδος δειγματοληψίας δεν έχει αμελητέα διάρκεια, δεν αποκλείονται πιθανές

μεταβάσεις σε μη γειτονικές καταστάσεις του μοντέλου, αντίθετα με μία τυπική birth-death διαδικασία όπου παρατηρούνται μόνο μεταβάσεις από μία κατάσταση στις γειτονικές τις [77].



Εικόνα 6.11. Μεταβάσεις μεταξύ καταστάσεων (πλήθος ενεργών στιγμιότυπων) για έναν τύπο εφαρμογής.

Για να εκτιμηθεί το πιθανό πλήθος των ενεργών στιγμιότυπων $n_{c,i}$, σε μία πιθανή μεταπομπή, δίνεται έμφαση στο πλήθος στιγμιότυπων που αναμένεται να είναι ενεργά, πλέον του πλήθους αναφοράς στην αρχή της χρονοθυρίδας, $dn_{c,i}$. Παρόμοια λογική είχε ακολουθηθεί και στο σχήμα πρόβλεψης θέσης του χρήστη με τη μελέτη της μεταβλητής $r_{max,i}$ και θ_i . Στην Εικόνα 6.12, φαίνεται ένα σενάριο που στην αρχή της χρονοθυρίδας παρατηρείται ένα στιγμιότυπο εφαρμογής και στη συνέχεια δειγματοληπτείται η διαφορά $dn_{c,i}$.



Εικόνα 6.12. Σενάριο παρατήρησης στιγμιότυπων ενός τύπου εφαρμογής στη διάρκεια μίας χρονοθυρίδας

Η παράμετρος $dn_{c,i}$ αποτελεί τυχαία μεταβλητή με συγκεκριμένα χαρακτηριστικά που μπορούν να εξαχθούν από την ιστορία της συμπεριφοράς της εφαρμογής τύπου c του χρήστη i . Συγκεκριμένα, η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής $dn_{c,i}$ για το χρήστη i , $p_{i,dnc}(n)$, ορίζεται ως εξής:

$$p_{i,dnc}(l) = \Pr[dn_{c,i} = l] = \Pr[n_{c,i}(nT + km) - n_{c,i}(nT) = l, k = 1, \dots, T/m, n = 0, 1, \dots] \quad (6.11)$$

Όπως είναι φανερό, η τυχαία μεταβλητή $dn_{c,i}$ παίρνει κι αρνητικές τιμές, καθώς στη διάρκεια της παρατήρησης το πλήθος των ενεργών στιγμιοτύπων μπορεί και να είναι μικρότερο από το πλήθος αναφοράς στην αρχή της χρονοθυρίδας. Για λόγους πρακτικότητας, γίνεται η υπόθεση ότι, για κάθε τύπο εφαρμογής c , υπάρχει ένα μέγιστο δυνατό πλήθος στιγμιοτύπων $N_{max,c}$ που μπορούν να είναι ταυτόχρονα ενεργά. Συνεπώς, η τυχαία μεταβλητή $dn_{c,i}$ μπορεί να πάρει τιμές από το διάστημα $\{-N_{max,c}, \dots, N_{max,c}\}$.

Γνωρίζοντας την κατανομή της τυχαίας μεταβλητής $dn_{c,i}$, είναι δυνατόν να υπολογιστεί η πιθανότητα στη διάρκεια μίας χρονοθυρίδας, δεδομένου του πλήθους ενεργών στιγμιοτύπων στην αρχή της χρονοθυρίδας $n_{c,i,ref}$, να έχουμε μία συγκεκριμένη αύξηση ή μείωση στο πλήθος των στιγμιοτύπων μέσα στη χρονοθυρίδα. Με αυτή τη γνώση, στη συνέχεια, είναι δυνατόν να υπολογιστεί η πιθανότητα $p_{i,nc}$ να έχουμε συγκεκριμένο πλήθος ενεργών στιγμιοτύπων μέσα στη χρονοθυρίδα, ως εξής:

$$p_{i,nc}(n_{c,i,ref}, l) = \begin{cases} p_{i,dnc}(l + n_{c,i,ref}), 1 \leq l + n_{c,i,ref} \leq N_{max,c} \\ \sum_{n=-N_{max,c}}^{n_{c,i,ref}} p_{i,dnc}(n), l = 0 \end{cases} \quad (6.12)$$

Έχοντας, λοιπόν, μία εκτίμηση για το πόσο πιθανό είναι να έχουμε 0, 1, 2, ..., στιγμιότυπα μίας εφαρμογής ενεργά, ανά πάσα στιγμή, σε μία χρονοθυρίδα, είναι δυνατόν να υπολογίσουμε και την αξία κάθε ατομικού στιγμιοτύπου. Συγκεκριμένα, αν η πιθανότητα να έχουμε ένα μόνο στιγμιότυπο είναι p_1 , αλλά να έχουμε δύο στιγμιότυπα p_2 , η πιθανότητα να υπάρχει ένα στιγμιότυπο θα είναι $p_1 + p_2$. Επομένως, όπως και στην παράγραφο 6.6.1, υπολογίζουμε μία συσσωρευτική πιθανότητα $q_{i,nc}$ η οποία αναπαριστά την πιθανότητα ύπαρξης συγκεκριμένου πλήθους στιγμιοτύπων, ως εξής:

$$q_{i,nc}(l) = \sum_{n=l}^{N_{max,c}} p_{i,dnc}(n) \quad (6.13)$$

Συνοψίζοντας, λοιπόν, ο χρήστης παρατηρώντας τη συμπεριφορά των εφαρμογών του, υπολογίζει την κατανομή της μεταβολής του πλήθους των ενεργών

στιγμιότυπων που μπορεί να σημειωθεί, δεδομένου ενός πλήθους αναφοράς (στην αρχή της χρονοθυρίδας). Βάσει αυτής της κατανομής, υπολογίζεται στη συνέχεια η πιθανότητα να έχει ταυτόχρονα ενεργό ένα πλήθος από στιγμιότυπα (ανεξάρτητα από το πλήθος αναφοράς). Τέλος, υπολογίζει και τη συσσωρευτική πιθανότητα για κάθε ξεχωριστό στιγμιότυπο εφαρμογής, η οποία μπορεί να χρησιμοποιηθεί από το χρήστη για τη δημιουργία κατάλληλων προσφορών προς το δίκτυο (δες υποκεφάλαιο 6.4).

6.7 Προσομοιώσεις

Στο παρόν υποκεφάλαιο, περιγράφονται αποτελέσματα προσομοιώσεων που έγιναν με σκοπό την αξιολόγηση του προτεινόμενου σχήματος προδραστικής διαχείρισης πόρων σε ασύρματα κυψελωτά δίκτυα. Ο Πίνακας 6.1 παρουσιάζει τις παραμέτρους της προσομοίωσης. Υπάρχουν 16 ΣΒ που είναι τοποθετημένοι με τρόπο που να σχηματίζουν ένα πλέγμα σε μία έκταση 4000m×4000m, ενώ, στην αρχή τα ΚΤ είναι διασκορπισμένα ομοιόμορφα, όπως φαίνεται στην Εικόνα 6.13.

Πίνακας 6.1. Λίστα των παραμέτρων της προσομοίωσης

N (πλήθος χρηστών)	50, 100, 200, 300, 400, 500
n_{cr_i}	$U(0,1)$
h_{cr_i}	$U(0, 1)$
B (πλήθος ΣΒ)	16
R_i (χωρητικότητα ΣΒ)	20,000 KB/s
Μοντέλο Κινητικότητας	Random Waypoint
u_{min} (Ελάχιστη ταχύτητα χρήστη)	0 km/h
u_{max} (Μέγιστη ταχύτητα χρήστη)	3 km/h
T (διάρκεια χρονοθυρίδας)	50 sec
Χρόνος προσομοίωσης	54,000 sec

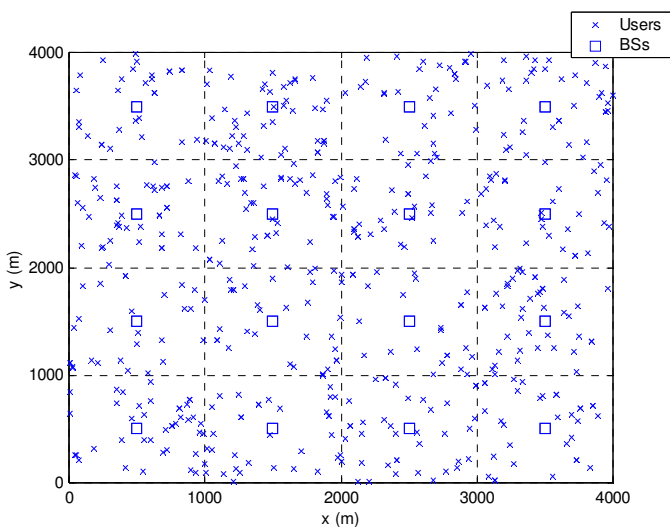
Οι χρήστες κινούνται τυχαία στην περιοχή προσομοίωσης, σύμφωνα με το μοντέλο κινητικότητας Random Waypoint (RWP) [78]. Σύμφωνα με το εν λόγω μοντέλο, ο χρήστης που βρίσκεται σε μία θέση, επιλέγει την επόμενη του θέση τυχαία (ομοιόμορφα) μέσα στην περιοχή που επιτρέπεται να κινηθεί. Στη συνέχεια, κινείται στην ευθεία που συνδέει την τρέχουσα θέση του με τη θέση που επέλεξε, με σταθερή ταχύτητα. Η ταχύτητα κίνησης επιλέγεται τυχαία από το διάστημα $[u_{min}, u_{max}]$, με ομοιόμορφο τρόπο. Όταν φθάσει στο σημείο που είχε επιλέξει, επιλέγει

μία νέα θέση και την ταχύτητα που θα κινηθεί. Σχετικές πληροφορίες με το μοντέλο παρέχει ο Πίνακας 6.1.

Γίνεται η υπόθεση ενός απλού μοντέλου εφαρμογών, VoIP, video, FTP και WWW. Ο Πίνακας 6.2 συνοψίζει τα χαρακτηριστικά αυτών των εφαρμογών. Επιπλέον, γίνεται η υπόθεση ότι υπάρχει ένα μέγιστο πλήθος ενεργών στιγμιότυπων ανά τύπο εφαρμογής που μπορούν να τρέχουν παράλληλα, ανά πάσα στιγμή, σε κάθε ΚΤ. Συγκεκριμένα, ένας χρήστης μπορεί να τρέξει παράλληλα μέχρι τρία στιγμιότυπα κάθε τύπου εφαρμογής. Η διάρκεια μίας εφαρμογής μοντελοποιήθηκε ως μία τυχαία μεταβλητή που ακολουθεί την εκθετική κατανομή με μέση τιμή. Οι αφίξεις νέων εφαρμογών ακολουθούν την κατανομή Poisson.

Πίνακας 6.2. Χαρακτηριστικά Εφαρμογών

Τύπος Εφαρμογής	Ελάχιστο Απαιτούμενο Εύρος Ζώνης (Kb/s)	Μέγιστο Απαιτούμενο Εύρος Ζώνης (Kb/s)	Μέση Διάρκεια Εφαρμογής (s)	Μέσος Χρόνος Ενδο-άφιξης (s)
VoIP	16	64	60	1200
Video	100	1000	300	900
FTP	1	10000	180	900
WWW	1	10000	5	30

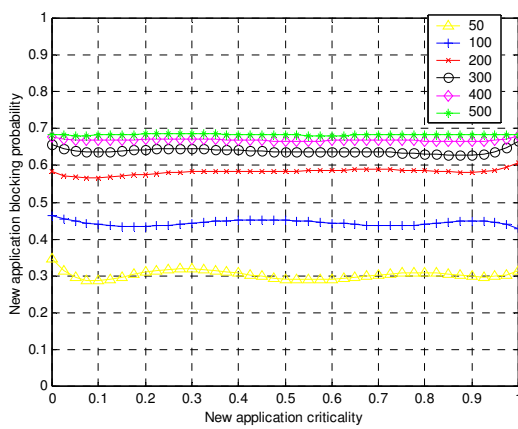


Εικόνα 6.13. Κατανομή χρηστών και ΣΒ στο δίκτυο

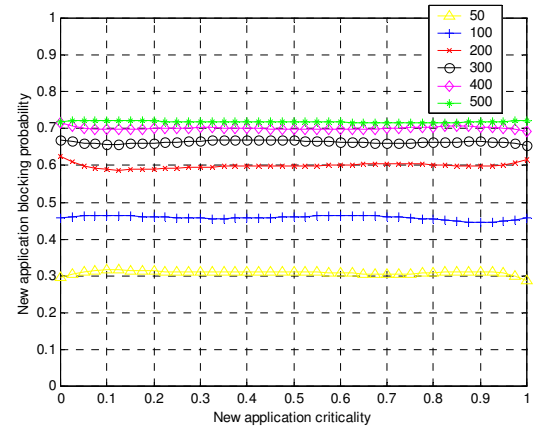
6.7.1 Αποτελέσματα Προσομοιώσεων

Καθώς ο σκοπός του προτεινόμενου σχήματος είναι να επιτύχει διαφοροποίηση μεταξύ των χρηστών, ανάλογα με τις ανάγκες τους, η μελέτη, σε ό,τι αφορά την

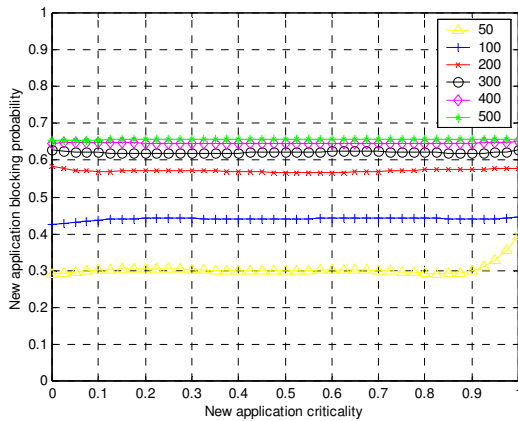
αποτελεσματικότητα του σχήματος, διενεργείται ακριβώς προς αυτήν την κατεύθυνση. Συγκεκριμένα, για την εν λόγω αξιολόγηση, επιλέγονται οι ακόλουθες μετρικές: 1) πιθανότητα μπλοκαρίσματος νέας εφαρμογής, 2) πιθανότητα διακοπής εφαρμογής, λόγω μεταπομπής, και 3) μέσος ρυθμός μετάδοσης που επιτυγχάνει μία εφαρμογή. Οι παραπάνω μετρικές αξιολογούνται ως προς τις τιμές κρισιμότητας νέας εφαρμογής, ncr_i , και εφαρμογής από μεταπομπή, hcr_i , για κάθε χρήστη i .



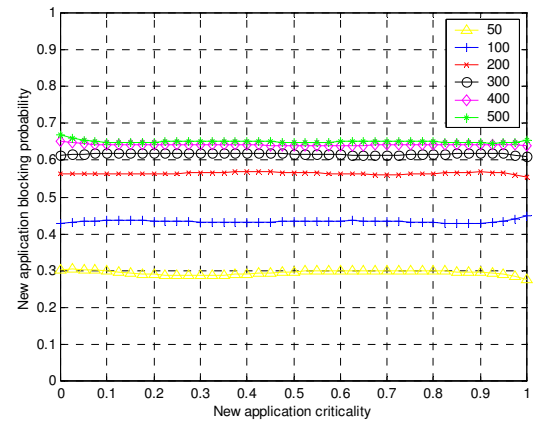
(a) VoIP



(b) Video



(c) FTP

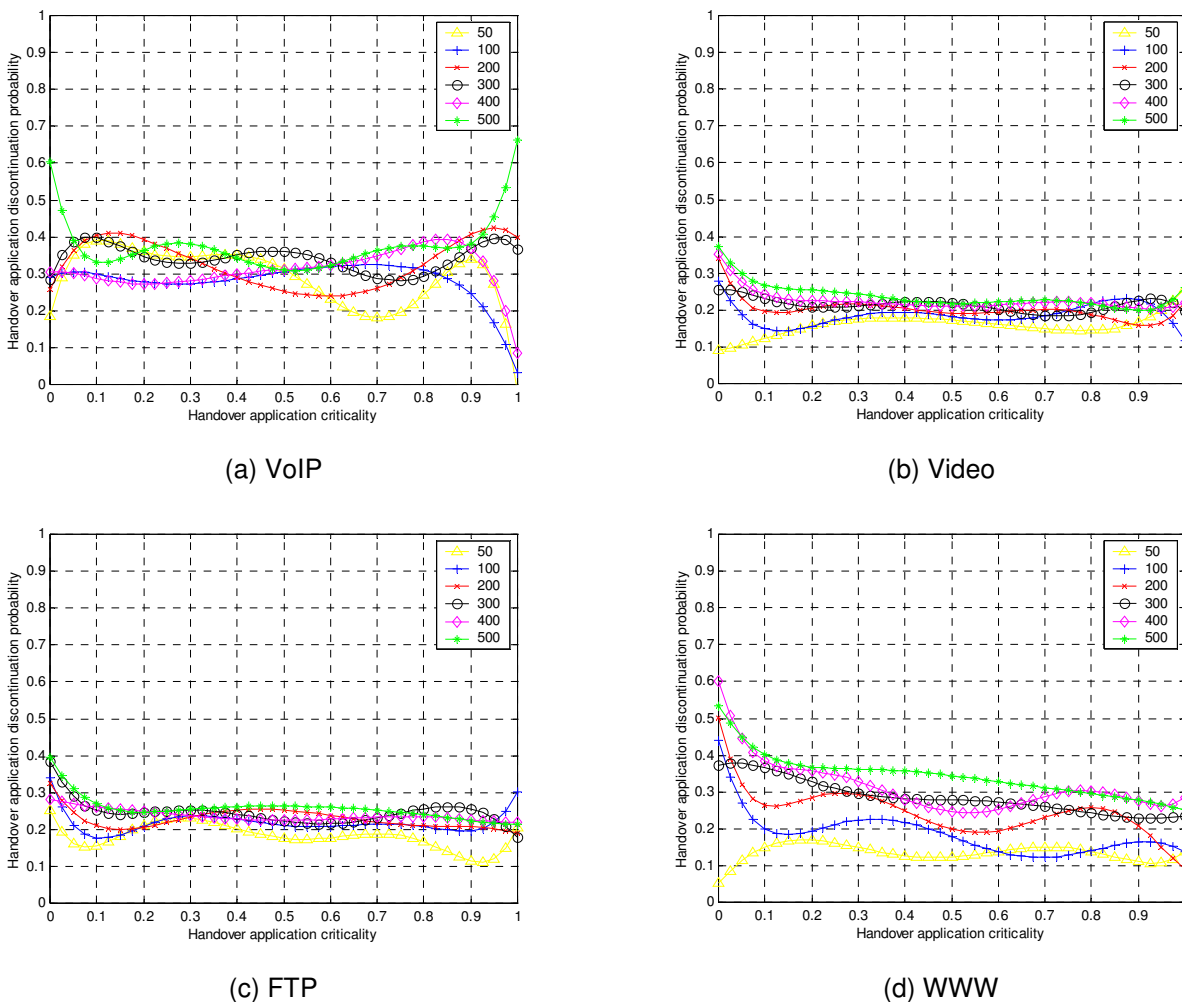


(d) WWW

Εικόνα 6.14. Πιθανότητα μπλοκαρίσματος νέας εφαρμογής, σε σχέση με την κρισιμότητα ncr_i

Στην Εικόνα 6.14, φαίνεται η πιθανότητα μπλοκαρίσματος νέας εφαρμογής, σε σχέση με την κρισιμότητα νέας εφαρμογής ncr_i , για τους τέσσερις τύπους εφαρμογών που έχουν προσομοιωθεί (VoIP, video, FTP, και WWW). Οι καμπύλες που απεικονίζονται έχουν προκύψει από *πολυωνυμική προσαρμογή* (polynomial

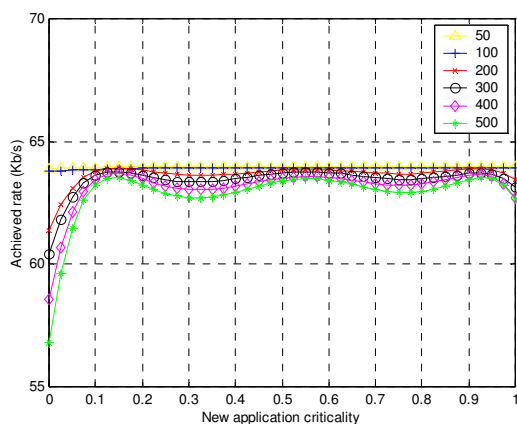
fitting) έκτου βαθμού στα αποτελέσματα, ώστε να μπορούν να απεικονιστούν πιο εύκολα οι τάσεις στο προτεινόμενο σύστημα. Παρατηρείται ότι η πιθανότητα μπλοκαρίσματος νέας εφαρμογής αυξάνεται, όσο αυξάνεται το πλήθος των χρηστών. Παρατηρείται ότι, για το video, η εν λόγω πιθανότητα είναι ελαφρά αυξημένη σε σχέση με τους άλλους τύπους εφαρμογών, το οποίο οφείλεται στο ότι το video απαιτεί τους περισσότερους πόρους κατ' ελάχιστο σε σχέση με τις άλλες εφαρμογές (δες Πίνακας 6.2), κάνοντας έτσι την απόρριψη του πιθανότερη. Παρατηρείται, επίσης, ότι ο παράγοντας κρισιμότητας νέας εφαρμογής ncr_i δεν επηρεάζει ιδιαίτερα τη μελετούμενη πιθανότητα.



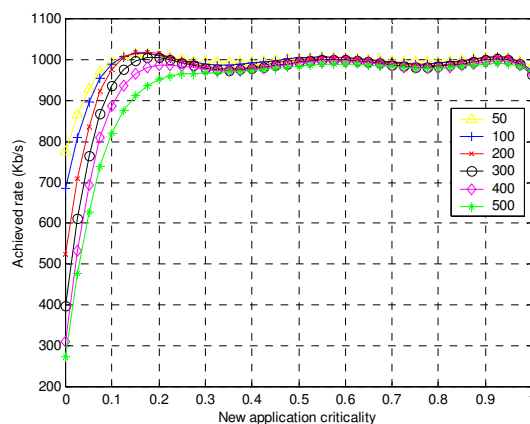
Εικόνα 6.15. Πιθανότητα διακοπής εφαρμογής, λόγω μεταπομπής, σε σχέση με την κρισιμότητα hcr_i

Στην Εικόνα 6.15, φαίνεται η πιθανότητα διακοπής μίας εφαρμογής, λόγω μεταπομπής, σε σχέση με την κρισιμότητα εφαρμογής από μεταπομπή hcr_i . Παρατηρείται, ότι ο εν λόγω παράγοντας (hcr_i) επηρεάζει τη μελετούμενη

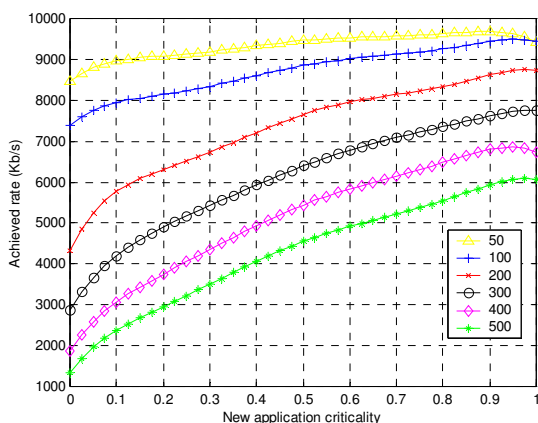
πιθανότητα (όσο μεγαλύτερο είναι το hcr_i , τόσο μικρότερη η πιθανότητα διακοπής εφαρμογής, λόγω μεταπομπής), κάτι που είναι περισσότερο φανερό για τις εφαρμογές τύπου WWW (Εικόνα 6.15(d)). Όπως ήταν αναμενόμενο, η συγκεκριμένη πιθανότητα αυξάνεται, καθώς αυξάνεται ο πληθυσμός των χρηστών στο δίκτυο.



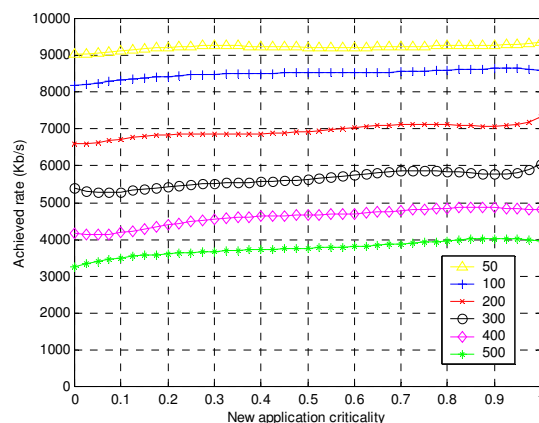
(a) VoIP



(b) Video



(c) FTP



(d) WWW

Εικόνα 6.16. Μέσος ρυθμός εφαρμογής, σε σχέση με την κρισιμότητα ncr_i

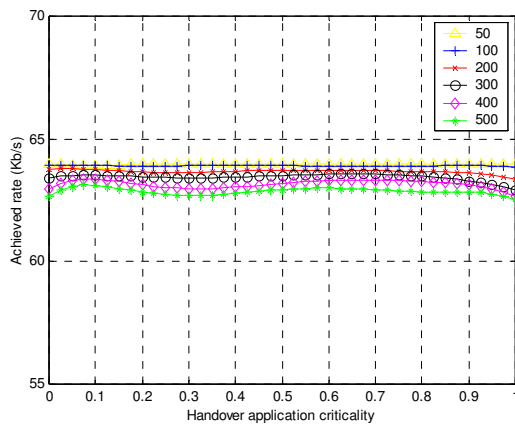
Στην Εικόνα 6.16, φαίνεται ο μέσος ρυθμός που επιτυγχάνεται, για κάθε τύπο εφαρμογής, σε σχέση με τον παράγοντα κρισιμότητας νέας εφαρμογής ncr_i . Η επίδραση του εν λόγω παράγοντα είναι ευδιάκριτη. Συγκεκριμένα, όσο αυξάνεται η κρισιμότητα μίας εφαρμογής για το χρήστη, τόσο υψηλότερο μέσο ρυθμό επιτυγχάνει. Για τις VoIP εφαρμογές, οι οποίες έχουν χαμηλές απαιτήσεις σε εύρος ζώνης (δες Πίνακας 6.2), ακόμα και για όχι πολύ υψηλές τιμές του παράγοντα

κρισιμότητας, οι χρήστες λαμβάνουν σχεδόν το μέγιστο που απαιτεί η εν λόγω εφαρμογή. Παρόμοια παρατήρηση μπορεί να γίνει και για το video, όπου, εκτός από τους χρήστες με πολύ χαμηλή κρισιμότητα, ο μέσος ρυθμός βρίσκεται πολύ κοντά στο μέγιστο.

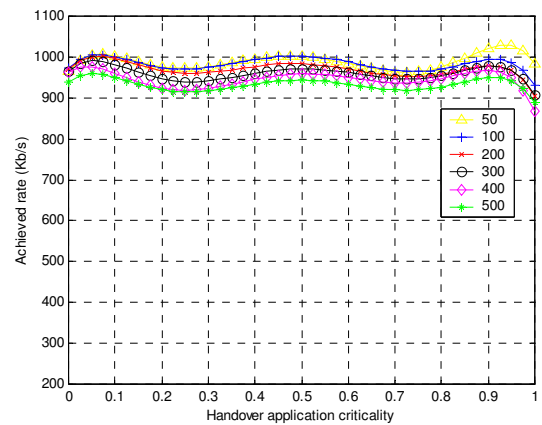
Αντίθετα με τις εφαρμογές VoIP και video, για τις εφαρμογές FTP και WWW, όπου ο μέγιστος ρυθμός είναι αρκετά πιο υψηλός (10 Mb/s), παρατηρείται ότι μόνον χρήστες με υψηλό παράγοντα κρισιμότητας καταφέρνουν να πλησιάσουν το μέγιστο. Είναι δε πιο εμφανής η επίδραση του πλήθους των χρηστών σε σχέση με τις άλλους δύο τύπους εφαρμογών. Μία ακόμα παρατήρηση που μπορεί να γίνει είναι ότι, παρόλο που οι εφαρμογές FTP και WWW έχουν τις ίδιες ελάχιστες και μέγιστες απαιτήσεις σε εύρος ζώνης (δες Πίνακας 5.1), η επίδραση του παράγοντα κρισιμότητας ncr_i είναι πιο έντονη στην εφαρμογή FTP. Αυτό μπορεί να αποδοθεί στις διαφορές των δύο εφαρμογών σε ό,τι αφορά τη συχνότητα εμφάνισης τους και τη διάρκεια τους. Οι WWW εφαρμογές έχουν μικρή διάρκεια κατά μέσον όρο (5 sec), σε αντίθεση με τις FTP (180 sec). Έτσι, αν μία WWW εφαρμογή καταφέρει να ξεκινήσει, κάποια στιγμή με χαμηλό ανταγωνισμό, παίρνοντας έτσι ένα ικανό ποσό πόρων, το πιθανότερο είναι να το διατηρήσει μέχρι και τον τερματισμό της. Αυτό γιατί, λόγω της μικρής διάρκειας ζωής της, είναι μάλλον ασύνηθες το να βρεθεί στο μεταίχμιο μεταξύ δύο χρονοθυρίδων (διάρκειας 50 sec, στις προσομοιώσεις – Πίνακας 6.1) , ώστε να χρειαστεί να επαναδιαπραγματευθεί το ποσό των πόρων της. Αντιθέτως, μία FTP εφαρμογή, κατά κανόνα, στη διάρκεια της ζωής της, θα υποστεί τη διαδικασία της επαναδιαπραγμάτευσης τουλάχιστον μία φορά, με αποτέλεσμα να αναγκαστεί να ελαττώσει τους πόρους που αρχικά απέκτησε, αν στο μεταξύ έχει αυξηθεί ο ανταγωνισμός.

Στην Εικόνα 6.17, φαίνεται ο μέσος ρυθμός που επιτυγχάνεται, για κάθε τύπο εφαρμογής, σε σχέση με τον παράγοντα κρισιμότητας εφαρμογής από μεταπομπή hcr_i . Από τα αποτελέσματα φαίνεται ότι ο συγκεκριμένος παράγοντας, μάλλον, δεν επηρεάζει το μέσο ρυθμό των εφαρμογών. Αυτό συμβαίνει γιατί ο παράγοντας hcr_i λαμβάνεται υπόψη μόνο σε συμβάντα μεταπομπής. Συγκεκριμένα, όπως περιγράφεται και στην παράγραφο 6.5.3, όταν ένα ΚΤ κάνει μία μεταπομπή, το αν θα διατηρήσει τις εφαρμογές του, καθώς και το ποσό των πόρων που θα αποκτήσουν, εξαρτάται από την τιμή του εν λόγω παράγοντα. Πρέπει όμως να

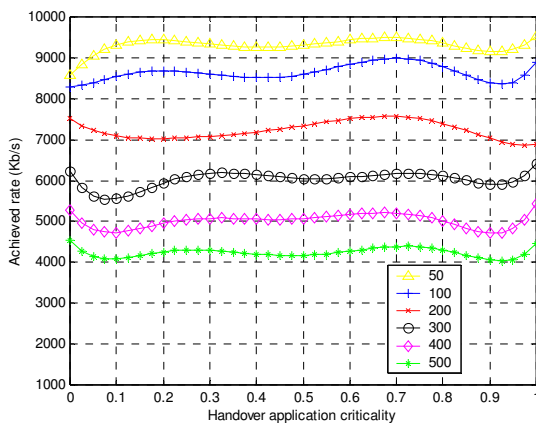
σημειωθεί ότι 1) η μεταπομπή αποτελεί ένα συμβάν με σχετικά χαμηλή συχνότητα (σε σχέση, για παράδειγμα, με ένα συμβάν έναρξης εφαρμογής), και 2) όταν παρέλθει η τρέχουσα χρονοθυρίδα (που διενεργήθηκε η μεταπομπή), το ποσό των πόρων θα επαναδιαπραγματευθεί, βάσει του παράγοντα ncr_i . Συνεπώς, είναι λογικό το ότι οι μέσοι ρυθμοί των εφαρμογών, είναι, μάλλον, αδιάφοροι σε ό,τι αφορά τον παράγοντα hcr_i , και σημαντικότερο ρόλο τελικά παίζει ο παράγοντας ncr_i .



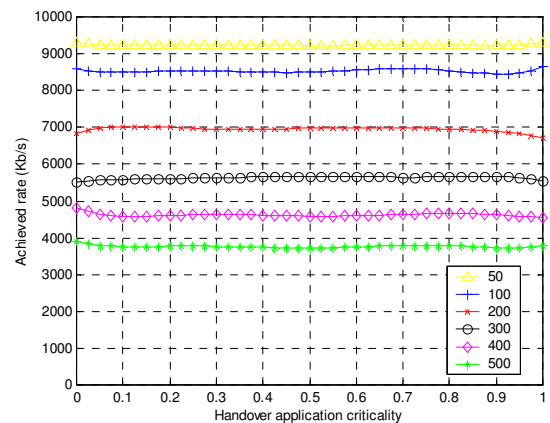
(a) VoIP



(b) Video



(c) FTP



(d) WWW

Εικόνα 6.17. Μέσος ρυθμός εφαρμογής, σε σχέση με την κρίσιμότητα hcr_i

6.8 Συμπεράσματα

Στο κεφάλαιο αυτό, προτάθηκε ένα προδραστικό σχήμα δέσμευσης δικτυακών πόρων, για την πρόληψη από διακοπές σύνδεσης, λόγω μεταπομπής. Συγκεκριμένα, προτάθηκε ένα σχήμα χρέωσης βασισμένο σε δημοπρασίες, στο

οποίο κάθε ΣΒ, ολοκληρωμένα, εξετάζει τις εφαρμογές που είναι ήδη ενεργές, αλλά και τις εφαρμογές που αναμένονται να αφιχθούν από άλλες κυψέλες. Για τον αποτελεσματικό διαμοιρασμό των πόρων στις εν λόγω εφαρμογές, σημαντικό ρόλο έπαιξαν οι δημοπρασίες, καθώς οι χρήστες έτσι είναι αναγκασμένοι να αποκαλύψουν τις πραγματικές προτεραιότητες τους. Στην παρούσα μελέτη προτάθηκε, επίσης, ένα πλαίσιο πρόβλεψης κινητικότητας των χρηστών, αλλά και της συμπεριφοράς των εφαρμογών. Βάσει του πλαισίου αυτού, οι χρήστες ήταν ικανοί να αξιολογήσουν, για τους ίδιους τους πόρους που θα χρειάζονται στο εγγύς μέλλον, ώστε να κάνουν αντίστοιχες προσφορές στο δίκτυο.

Κάθε ΣΒ συγκεντρώνει προσφορές από τα ΚΤ, είτε πρόκειται για νέες, είτε για εφαρμογές από μεταπομπή, και εξυπηρετεί τις πιο σημαντικές, δεδομένης της συνολικής χωρητικότητας τους. Από την αξιολόγηση του σχήματος, προέκυψε ότι όσο πιο σημαντικούς θεωρούν οι χρήστες τους πόρους, τόσο πιο πολλούς λαμβάνουν, το οποίο αποτελεί σημαντική βελτίωση σε σχέση με υπάρχοντα σχήματα προδραστικής διαχείρισης πόρων που κατά βάση λειτουργούν απρόσωπα και δεν είναι ικανά να λάβουν υπόψη τις προτεραιότητες των χρηστών ατομικά.

ΚΕΦΑΛΑΙΟ 7

ΚΑΤΑΝΕΜΗΜΕΝΗ ΠΡΟΣΒΑΣΗ ΣΕ ΔΙΚΤΥΑ ΜΕ ΕΥΕΛΙΚΤΗ ΧΡΗΣΗ ΦΑΣΜΑΤΟΣ

7.1 Εισαγωγή

Στο κεφάλαιο αυτό, παρουσιάζεται η μελέτη που πραγματοποιήθηκε, στα πλαίσια της παρούσας διατριβής, σχετικά με τη δυναμική διαχείριση φάσματος. Σύμφωνα με πρόσφατες μελέτες από την FCC, ακόμα και σε μεγάλες αστικές περιοχές, μόνο το 30% του αδειοδοτημένου φάσματος χρησιμοποιείται κατά μέσον όρο. Για το λόγο αυτό, η δυναμική πρόσβαση σε μη-αδειοδοτημένο φάσμα αποτελεί μία ενδιαφέρουσα πρόταση για την αποδοτικότερη χρήση του υπάρχοντος φάσματος. Προς αυτήν την κατεύθυνση, η FCC έχει εκδώσει το Notice of Public Rulemaking and Order που αφορά στις ασύρματες τεχνολογίες με επίγνωση του περιβάλλοντος (cognitive radio technologies) [79]. Επιπλέον, το DARPA έχει ξεκινήσει το πρόγραμμα neXt Generation (XG) για την ανάπτυξη νέων ευέλικτων μηχανισμών και τεχνολογιών για το διαμοιρασμό του εύρους ζώνης [80]. Και στις δύο πρωτοβουλίες, οι συσκευές ευέλικτης χρήσης του φάσματος αναγνωρίζουν και χρησιμοποιούν δυναμικά το αδρανές ή υποχρησιμοποιούμενο φάσμα, αυξάνοντας έτσι την αποδοτικότητα στη χρήση του φάσματος.

Στην παρούσα μελέτη, εξετάζεται το πρόβλημα της κατανεμημένης, ομορπορουνιστικής πρόσβασης σε CDMA δίκτυα. Η μελέτη εστιάζεται στο CDMA καθώς έχει υιοθετηθεί ευρέως σε δίκτυα τρίτης γενιάς (3G) [81] κι η υιοθεσία του στα δίκτυα 4G είναι, επίσης, αναμενόμενη [82], [83]. Γίνεται η υπόθεση ότι οι δευτερεύοντες χρήστες (μη έχοντες άδεια) ανταγωνίζονται εγωιστικά για την πρόσβαση στις υπάρχουσες ζώνες συχνοτήτων. Συγκεκριμένα, κάθε χρήστης επιλέγει μία ζώνη συχνοτήτων για να εκπέμψει και ρυθμίζει την ισχύ εκπομπής του, ώστε να μεγιστοποιήσει το όφελος του. Επίσης, γίνεται η υπόθεση ότι κάθε χρήστης έχει προσωπικές απαιτήσεις σε ό,τι αφορά την ποιότητα υπηρεσίας, εκφρασμένες σε όρους επιθυμητού λόγου σήματος προς παρεμβολή, SIR, στο δέκτη (ΣΒ). Επεκτείνονται γνωστά σχήματα ελέγχου ισχύος (π.χ. [84], [85], [86]),

ούτως ώστε η επιλογή μίας ζώνης συχνοτήτων (σε συνδυασμό με την ισχύ εκπομπής) να αποτελεί μέρος της στρατηγικής του χρήστη, ενώ ταυτόχρονα να λαμβάνονται υπόψη περιορισμοί στο SIR. Ο σκοπός της παρούσας μελέτης είναι η παροχή αποτελεσματικών, κατανεμημένων στρατηγικών για την επίτευξη των επιθυμητών επιπέδων SIR με την ελάχιστη δυνατή κατανάλωση ενέργειας. Προτείνεται μία στρατηγική βασισμένη σε προβλέψεις κι άλλη μία πιθανοτική στρατηγική.

7.2 Σχετική Βιβλιογραφία και Συνεισφορά

Σε αυτό το υποκεφάλαιο, παρατίθεται μία σύνοψη της σχετικής έρευνας που έχει πραγματοποιηθεί προσφάτως. Η δυναμική εκχώρηση καναλιών (*dynamic channel allocation* – DCA) έχει μελετηθεί σε μεγάλο βαθμό στη σχετική βιβλιογραφία. Ποικίλοι κεντριοποιημένοι ή κατανεμημένοι αλγόριθμοι DCA περιγράφονται και συγκρίνονται στην εργασία [87]. Οι μηχανισμοί αυτοί επικεντρώνονται στη γεωγραφική επαναχρησιμοποίηση, βάσει περιορισμών επαναχρησιμοποίησης καναλιών. Επιπλέον, η εκχώρηση καναλιών έχει μελετηθεί για κατανεμημένα ασύρματα δίκτυα peer-to-peer (π.χ. δίκτυα αισθητήρων), με στόχο την ελαχιστοποίηση της κατανάλωσης ενέργειας [88].

Στην εργασία [89], ορίστηκε η αρχιτεκτονική του cognitive radio και παρουσιάστηκε ένα πρωτότυπο cognitive radio, το CR1. Στην εργασία [90], μελετήθηκαν μοντέλα Markov συνεχούς χρόνου για τη δυναμική φασματική πρόσβαση σε δίκτυα με ευέλικτη χρήση φάσματος. Αποδείχθηκε ότι τα εν λόγω μοντέλα είναι ακριβή στο να προβλέπουν τη δυναμική της ανοικτής φασματικής πρόσβασης. Οι συγγραφείς της εργασίας [90] πρότειναν, επίσης, ένα κατανεμημένο πρωτόκολλο MAC με σχεδόν βέλτιστη απόδοση. Στην εργασία [91], μελετήθηκαν τα χαρακτηριστικά της ομορτουλιστικής διαθεσιμότητας του φάσματος και της εκμετάλλευσής της. Όμως, καμία από τις προαναφερθείσες εργασίες δεν μοντελοποίησε τη δυναμική φασματική πρόσβαση για CDMA δίκτυα, με περιορισμούς επιθυμητού QoS για κάθε χρήστη, το οποίο αποτελεί ένα ζήτημα που δεν ακόμη καλυφθεί από τη βιβλιογραφία.

Στην εργασία [92], η οποία θεωρείται η πιο σχετική με την παρούσα μελέτη, μελετήθηκε το πρόβλημα της παροχής εγγυήσεων QoS σε πολυμεσικές εφαρμογές, σε περιβάλλοντα δυναμικής φασματικής πρόσβασης. Προτάθηκε, για το λόγο αυτό, ένας καταναμημένος, μη-συνεργατικός, μηχανισμός επιλογής καναλιού. Παρόμοια με την προτεινόμενη σε αυτό το κεφάλαιο μελέτη, οι χρήστες, στην εργασία [92], δρουν εγωιστικά και για το λόγο αυτό υιοθετήθηκε μία παιγνιοθεωρητική ανάλυση. Παρόλ' αυτά, στην εργασία [92], έχει γίνει η υπόθεση ότι τα ΚΤ ενημερώνουν τη στρατηγική τους (επιλέγουν ζώνη συχνοτήτων) ακολουθιακά, μία υπόθεση που δεν είναι ιδιαίτερα ρεαλιστική σε ένα πλήρως καταναμημένο και μη-συνεργατικό περιβάλλον.

7.3 Μοντέλο Συστήματος

Σε αυτό το υποκεφάλαιο, μοντελοποιείται το πρόβλημα της δυναμικής φασματικής πρόσβασης σε δίκτυα CDMA. Γίνεται η υπόθεση ότι υπάρχουν δύο τύποι κόμβων: (1) τα μη-αδειοδοτημένα (δευτερεύοντα) ΚΤ και (2) οι ΣΒ. Έστω $J = \{1, \dots, M\}$ το σύνολο των ΚΤ και $I = \{1, \dots, B\}$ το σύνολο των ΣΒ. Κάθε ΚΤ συνδέεται στο ΣΒ με το ισχυρότερο σήμα.

Γίνεται η υπόθεση ότι υπάρχει ένα σύνολο από διαθέσιμες CDMA ζώνες συχνοτήτων, δηλαδή ζώνες συχνοτήτων στις οποίες κανένα αδειοδοτημένο (πρωτεύον) ΚΤ δεν εκπέμπει, $C = \{1, \dots, K\}$. Για λόγους απλότητας, θεωρείται ότι οι διαθέσιμες ζώνες συχνοτήτων έχουν ίσο εύρος συχνοτήτων. Οι πόροι κάθε ζώνης συχνοτήτων διαμοιράζονται μεταξύ των ΚΤ μέσω της τεχνικής CDMA. Αυτό σημαίνει ότι υπάρχουν αμοιβαίες παρεμβολές μεταξύ των ΚΤ που εκπέμπουν ταυτόχρονα, οι οποίες εξαρτώνται άμεσα από την ισχύ εκπομπής κάθε ΚΤ. Το ΚΤ j μπορεί να επιλέξει να εκπέμπει με ισχύ $p_j \in [0, p_{max}] \equiv P$, και στη ζώνη συχνοτήτων $c_j \in C$. Επομένως, το σύνολο στρατηγικών S_j του ΚΤ j ορίζεται ως το Καρτεσιανό γινόμενο $S_j = P \times C$. Το διάνυσμα $\mathbf{s} = (\mathbf{s}_1, \dots, \mathbf{s}_N)^T$, όπου $\mathbf{s}_j = (p_j, c_j)$, είναι ένα τυπικό προφίλ στρατηγικής στο χώρο των στρατηγικών $S = S_1 \times \dots \times S_N$.

Στο μοντέλο που υιοθετήθηκε, κάθε ΚΤ επιβάλλει πάνω και κάτω όρια στο SIR του, $\gamma_{m,j}$ and $\gamma_{M,j}$ ($\gamma_{m,j} < \gamma_{M,j}$). Το SIR αποτελεί τη βασική μετρική ποιότητας στις CDMA επικοινωνίες, καθώς σχετίζεται άμεσα με τη ρυθμοαπόδοση που μπορεί να

επιτευχθεί¹¹. Τέτοιες απαιτήσεις QoS είναι συμβατές με υπηρεσίες που αποδέχονται διακυμάνσεις στο ρυθμό τους, αλλά δεν ανέχονται καθυστερήσεις (π.χ., φωνή, video, ή μεταφορά αρχείων σε πραγματικό χρόνο).

Το SIR γ_j που επιτυγχάνει το ΚΤ j ορίζεται ως εξής:

$$\gamma_j = \frac{W}{R} \frac{h_{a_j,j} p_j}{\sum_{k \neq j, c_k = c_j} h_{a_j,k} p_k + \sigma_{c_j}^2}, \quad (7.1)$$

όπου το c_j υποδηλώνει την ζώνη συχνοτήτων στην οποία εκπέμπει το ΚΤ j και το a_j το ΣΒ που το εξυπηρετεί. Το W είναι ο ρυθμός chip (εύρος ζώνης εκπομπής) και R ο ρυθμός εκπομπής (b/s). Το $h_{a_j,j}$ υποδηλώνει τις απώλειες σήματος από το ΚΤ j μέχρι το ΣΒ a_j και $\sigma_{c_j}^2$ είναι η ισχύς του προσθετικού λευκού θορύβου Γκάους (*additive white Gaussian noise – AWGN*) στο ΣΒ στην ζώνη συχνοτήτων c_j . Το άθροισμα $\sum_{k \neq j, c_k = c_j} h_{a_j,k} p_k$ αντιστοιχεί σε σήματα που λαμβάνονται από το ΣΒ a_j , τα

οποία είναι παρεμβολές για το σήμα του ΚΤ j . Αυτή η παρεμβολή μπορεί να διαχωριστεί σε δύο μέρη: (1) την παρεμβολή από ΚΤ που εξυπηρετούνται από το συγκεκριμένο ΣΒ (ενδο-κυψελική παρεμβολή, *intra-cell interference*) και (2) παρεμβολή από ΚΤ που εξυπηρετούνται από άλλους ΣΒ (εξω-κυψελική παρεμβολή, *other-cell* ή *inter-cell interference*). Η συγκεκριμένη διάκριση γίνεται καθώς ο ΣΒ είναι ικανός να «ελέγχει» μόνον τα εξυπηρετούμενα από αυτόν ΚΤ· κάθε άλλο σήμα δε μπορεί παρά να γίνει αντιληπτό ως θόρυβος. Η εξω-κυψελική παρεμβολή υποδηλώνεται ως ρ_{c_j, a_j}^2 κι ορίζεται ως

$$\rho_{c_j, a_j}^2 = \sum_{k \neq j, c_k = c_j, a_k \neq a_j} h_{a_j,k} p_k + \sigma_{c_j}^2. \quad (7.2)$$

Το SIR του ΚΤ j μπορεί, επομένως, να γραφεί ισοδύναμα ως εξής:

$$\gamma_j = \frac{W}{R} \frac{h_{a_j,j} p_j}{\sum_{k \neq j, c_k = c_j, a_k = a_j} h_{a_j,k} p_k + \rho_{c_j, a_j}^2} \quad (7.3)$$

¹¹ Η ακριβής τους σχέση εξαρτάται από το σχήμα διαμόρφωσης. Για να διατηρηθεί, όμως, η γενικότητα στη μελέτη και να αποφευχθούν ζητήματα αντιστοίχισης, στην παρούσα μελέτη εξετάζεται μόνο το SIR ως μετρική που υποδεικνύει το επίπεδο του QoS.

Από την εξίσωση (7.1) ή (7.3), παρατηρείται ότι, για ένα δεδομένο επίπεδο παρεμβολών, ένα ΚΤ μπορεί να αυξήσει το SIR του αυξάνοντας την ισχύ εκπομπής του. Αυτό, όμως, έχει ως αποτέλεσμα τις αυξημένες παρεμβολές για τα άλλα ΚΤ. Είναι εύκολο να αποδειχθεί ότι, αν όλα τα ΚΤ αυξήσουν την ισχύ εκπομπής τους, το SIR τους θα αυξηθεί κι αυτό, αλλά ο ρυθμός με τον οποίο αυξάνεται το SIR μειώνεται, καθώς αυξάνεται η ισχύς εκπομπής τους. Έτσι, τα ΚΤ πρέπει να λαμβάνουν υπόψη, εκτός από το QoS (SIR) που θα επιτύχουν, και την κατανάλωση ισχύος, όταν επιλέγουν την ισχύ εκπομπής τους. Για να ποσοτικοποιηθεί αυτή η ανταλλαγή, γίνεται η υπόθεση ότι το ΚΤ j έχει μία συνάρτηση ωφελείας $u_j: S \rightarrow \mathfrak{R}$, η οποία εκφράζει τον αριθμό των bit που αποκωδικοποιούνται με επιτυχία ανά Joule καταναλισκόμενης ενέργειας [86], που ορίζεται ως εξής:

$$u_j(p, c) = \frac{LR}{Mp_j} (1 - e^{-0.5\gamma_j})^M \text{ (bits/Joule)} \quad (7.4)$$

Το L υποδηλώνει το μήκος ενός πλαισίου του ΚΤ (σε bit), το M το μήκος του πλαισίου με την επικεφαλίδα (σε bit) ($M > L$). Ο όρος $(1 - e^{-0.5\gamma_j})^M$ αντιστοιχεί στην πιθανότητα της επιτυχούς λήψης ενός πλαισίου η οποία αποτελεί μέτρο της ρυθμοαπόδοσης που επιτυγχάνεται, για ασύγχρονη FSK διαμόρφωση [95].

Από την εξίσωση (7.4), μπορεί να παρατηρηθεί ότι η ωφέλεια ενός ΚΤ είναι συνάρτηση των στρατηγικών που θα επιλέξει σε σχέση με την ζώνη συχνοτήτων κι ισχύ εκπομπής, αλλά και των ενεργειών των άλλων ΚΤ. Έτσι, παρατηρείται ότι υπάρχει αμοιβαία επίδραση μεταξύ τους, που σημαίνει ότι το πρόβλημα της επιλογής ζώνης συχνοτήτων κι ισχύος εκπομπής μπορούν να μελετηθούν με παιγνιοθεωρητικά. Συγκεκριμένα, η μελετούμενη κατάσταση μοντελοποιείται ως ένα μη-συνεργατικό παίγνιο, στο οποίο τα ΚΤ είναι οι παίκτες που εγωιστικά ανταγωνίζονται για το εύρος ζώνης της ανερχόμενης ζεύξης, επιλέγοντας ανεξάρτητα τη ζώνη συχνοτήτων κι ισχύ εκπομπής τους.

7.4 Κατανεμημένος, Μη-Συνεργατικός Έλεγχος Ισχύος

Σε αυτό το υποκεφάλαιο, μελετάται το πρόβλημα του κατανεμημένου ελέγχου ισχύος στο θεωρούμενο περιβάλλον CDMA. Συγκεκριμένα, το μελετούμενο

πρόβλημα μοντελοποιείται ως ένα μη-συνεργατικό παίγνιο, του οποίου οι παίκτες είναι τα ΚΤ που εκπέμπουν στην ίδια ζώνη συχνοτήτων.

Ένα βασικό ζήτημα στα μη-συνεργατικά παίγνια αποτελεί η ύπαρξη κι η μοναδικότητα μίας κατά Nash ισορροπίας [72]. Στο μελετούμενο παίγνιο, το NE, σε μία δεδομένη ζώνη συχνοτήτων c που διαμοιράζεται από ένα σύνολο από ΚΤ $J' = \{1, \dots, N\} \subseteq J$, το οποίο εξυπηρετείται από το ΣΒ a , ορίζεται πιο κάτω.

Ορισμός 7.1 (Ισορροπία Nash του Παιγνίου Ελέγχου Ισχύος): Σε μία δεδομένη ζώνη συχνοτήτων c που διαμοιράζεται από ΚΤ από το σύνολο $J' = \{1, \dots, N\} \subseteq J$ κι εξυπηρετείται από το ΣΒ a , ένα διάνυσμα ισχύος $\rho = (\rho_1, \dots, \rho_N)^T$ είναι μία κατά Nash ισορροπία του παιγνίου ελέγχου ισχύος αν, για κάθε ΚΤ $j \in J'$,

$$u_j(\rho_j, \mathbf{p}_{-j}, \mathbf{c}) \geq u_j(\rho'_j, \mathbf{p}_{-j}, \mathbf{c}), \forall \rho'_j \in S_j.$$

Με άλλα λόγια, σε ένα NE, με δεδομένα τα επίπεδα ισχύος των άλλων ΚΤ, στην ίδια ζώνη συχνοτήτων, κανένα ΚΤ δε μπορεί να βελτιώσει περαιτέρω την ωφέλεια του κάνοντας ατομικές αλλαγές στην εκπεμπόμενη ισχύ του. Στην εργασία [86], έχει αποδειχθεί ότι το παίγνιο ελέγχου ισχύος με τη συνάρτηση ωφελείας (7.4), χωρίς περιορισμούς SIR έχει ένα μοναδικό NE. Συγκεκριμένα, στο εν λόγω NE, όλα τα ΚΤ απολαμβάνουν το ίδιο SIR γ^* το οποίο είναι η λύση της επόμενης εξίσωσης:

$$0.5M\gamma + 1 = e^{0.5\gamma} \quad (7.5)$$

Η εξίσωση (7.5) προκύπτει από τη συνθήκη βελτιστότητας πρώτου βαθμού (first-order optimality condition) αναφορικά με τη συνάρτηση ωφελείας του ΚΤ (7.4) ως προς την ισχύ εκπομπής του, δηλαδή $\partial u_j(\cdot) / \partial \rho_j = 0$. Αν τίθενται περιορισμοί στο SIR από κάθε ΚΤ j , όπως υποτίθεται στην παρούσα μελέτη, δηλαδή $\gamma_{m,j} \leq \gamma_j \leq \gamma_{M,j}$, το SIR του ΚΤ j στο NE υπολογίζεται ως εξής [93]:

$$\gamma_j^* = \begin{cases} \gamma_{m,j}, & \gamma^* < \gamma_{m,j} < \gamma_{M,j} \\ \gamma^*, & \gamma_{m,j} < \gamma^* < \gamma_{M,j} \\ \gamma_{M,j}, & \gamma_{m,j} < \gamma_{M,j} < \gamma^* \end{cases} \quad (7.6)$$

Με βάση την εξίσωση (7.3) κι εξίσωση (7.6), είναι, επίσης, δυνατόν να υπολογιστεί η ισχύς εκπομπής των ΚΤ στο ΝΕ, επιλύοντας το γραμμικό σύστημα εξισώσεων που προκύπτει. Συγκεκριμένα, με δεδομένες τις απαιτήσεις σε SIR των ΚΤ που εξυπηρετούνται από το ΣΒ a κι εκπέμπουν στη ζώνη συχνοτήτων c , καθώς και της εξω-κυψελικής παρεμβολής $\rho_{c,a}^2$, η ισχύς στο ΝΕ q_j^* ($q_j^* = h_{a,j}p_j^*$), με την οποία το ΚΤ j φτάνει στο ΣΒ a , υπολογίζεται ως

$$q_j^* = \frac{\rho_{c,a}^2}{\left(\frac{W}{R\gamma_j^*} + 1\right) \left(1 - \sum_{k \in J'} \left(\frac{W}{R\gamma_k^*} + 1\right)^{-1}\right)}, \quad \forall j \in J'. \quad (7.7)$$

Από την εξίσωση (7.7) παρατηρείται ότι για να είναι η ισχύς στο ΝΕ πεπερασμένη και μη-αρνητική, ο ακόλουθος περιορισμός θα πρέπει να ικανοποιείται:

$$\sum_{k=1, c_k=c, a_k=a}^N \left(\frac{W}{R\gamma_k} + 1\right)^{-1} < 1 \quad (7.8)$$

Ο περιορισμός στην εξίσωση (7.8) αναπαριστά τον περιορισμό της χωρητικότητας της CDMA ζώνης συχνοτήτων, που ονομάζεται και «χωρητικότητα πόλου» (role capacity). Αξίζει να σημειωθεί ότι αυτό αποτελεί έναν έλεγχο που μπορεί να πραγματοποιηθεί από κάθε ΚΤ, αρκεί ο ΣΒ να ενημερώνει τα ΚΤ για την τιμή του αριστερού μέρους της ανισότητας (7.8), η οποία υποδεικνύει το επίπεδο κορεσμού της συγκεκριμένης ζώνης συχνοτήτων. Το ΚΤ j μπορεί να προσθέσει στην τιμή αυτήν τον όρο $(W/R\gamma_j + 1)^{-1}$, που αντιπροσωπεύει τη δική του συνεισφορά και να ελέγξει αν το αποτέλεσμα υπερβαίνει τη μονάδα. Αυτή η τεχνική μπορεί να αποτελέσει τη βάση για ένα σχήμα *ελέγχου αποδοχής κλήσεων* που είναι εύκολο στην υλοποίησή του, αλλά και πλήρως καταναμημένο. Αν ο συγκεκριμένος έλεγχος διενεργείται πριν από τη σύνδεση του ΚΤ, η κατανομή ισχύος στο ΝΕ που θα προκύψει θα είναι εφικτή.

Ένα σημαντικό ζήτημα, επίσης, αποτελεί πως μπορεί το κάθε ΚΤ να υπολογίσει την ισχύ που αντιστοιχεί στο ΝΕ, με την οποία πρέπει να εκπέμψει. Ενώ με την εξίσωση (7.5) κι εξίσωση (7.6) κάθε ΚΤ μπορεί να υπολογίσει ανεξάρτητα το SIR του στο ΝΕ, η εξίσωση (7.7) είναι μάλλον ακατάλληλη για τον καταναμημένο υπολογισμό της ισχύος στο ΝΕ. Αυτό συμβαίνει, καθώς, για να χρησιμοποιηθεί η (7.7), τα SIR των υπόλοιπων ΚΤ στο ΝΕ, όπως και το επίπεδο των εξω-κυψελικών

παρεμβολών χρειάζεται να είναι γνωστά σε κάθε ΚΤ. Για το λόγο αυτό, τα ΚΤ μπορούν να καταφύγουν σε μία κατανεμημένη, επαναληπτική διαδικασία, στην οποία κάθε ΚΤ προσαρμόζει την ισχύ του ανταποκρινόμενο στις παρεμβολές των άλλων ΚΤ, ούτως ώστε να επιτύχει το επιδιωκόμενο (NE) SIR. Αυτή η διαδικασία σκιαγραφείται πιο κάτω.

Γίνεται η υπόθεση ότι ο ΣΒ a (στον οποίο το ΚΤ j είναι συνδεδεμένο), κάθε T_p (περίοδος ελέγχου ισχύος), ευρεκπέμπει το επίπεδο της λαμβανόμενης ισχύος (παρεμβολής) στη ζώνη συχνοτήτων c (όπου εκπέμπει το ΚΤ j) $rp_{c,a}(t)$, όπου, για τη χρονική στιγμή t ,

$$rp_{c,a}(t) = \sum_{k=1, c_k=c_j}^N h_{a,k} p_k(t).$$

Το $h_{a,j}$ είναι οι εκτιμώμενες απώλειες σήματος¹² από το ΚΤ j στο ΣΒ a και $p_j(t)$ είναι η εκπεμπόμενη ισχύς του ΚΤ j , τη χρονική στιγμή t . Το ΚΤ j μπορεί, τότε, να προσαρμόσει την εκπεμπόμενη ισχύ του, βάσει του ακόλουθου τύπου:

$$p_j(t+1) = \min \left\{ \frac{R\gamma_j^*}{Wh_{a,j}} [rp_{c,a}(t) - h_{a,j}p_j(t)], p_{\max} \right\}, \quad (7.9)$$

όπου το p_{\max} υποδηλώνει τη μέγιστη ισχύ των πομπών. Μετά από ένα πεπερασμένο πλήθος βημάτων, η διαδικασία αυτή συγκλίνει στην κατανομή ισχύος του NE, αρκεί βέβαια η συνθήκη (7.8) να είναι αληθής.

Από τα αποτελέσματα, σε σχέση με το NE του παιγνίου ελέγχου ισχύος, που μελετήθηκε σε αυτό το υποκεφάλαιο, μπορεί να παρατηρηθεί ότι το SIR ενός ΚΤ στο NE δεν επηρεάζεται από την πλειοψηφία των παραμέτρων του συστήματος, όπως φαίνεται από τις εξισώσεις (7.5) και (7.6). Με άλλα λόγια, το ΚΤ επιτυγχάνει τα ίδια επίπεδα SIR (κι επομένως ρυθμοαπόδοση), ανεξάρτητα από τη ζώνη συχνοτήτων στην οποία εκπέμπει, το πλήθος των ΚΤ που εξυπηρετούνται από τον ίδιο ΣΒ, το (NE) SIR στο οποίο στοχεύουν, ή το επίπεδο της εξω-κυψελικής παρεμβολής. Η αντίστοιχη ισχύς στο NE, όμως, όπως φαίνεται καθαρά από την εξίσωση (7.7), είναι ευαίσθητη στους προαναφερθέντες παράγοντες. Αυτό σημαίνει

¹² Οι απώλειες $h_{a,j}$ συμπεραίνονται από τις απώλειες της κατερχόμενης ζεύξης, οι οποίες είναι γνωστές (ο ΣΒ εκπέμπει το «φάρο» του με γνωστή ισχύ). Γενικά, το κανάλι της ανερχόμενης και το κανάλι της κατερχόμενης ζεύξης είναι αρκετά κοντά, με αποτέλεσμα οι μεγάλης κλίμακας απώλειες να μπορούν να θεωρηθούν ισχυρά συσχετισμένες. Έτσι, μπορεί με ασφάλεια να υποτεθούν ίσες απώλειες για την ανερχόμενη και την κατερχόμενη ζεύξη.

ότι ένα ΚΤ θα μπορούσε να ωφεληθεί μεταπηδώντας σε μία άλλη ζώνη συχνοτήτων, η οποία είναι λιγότερο φορτωμένη, ή όπου τα επίπεδα εξω-κυψελικής παρεμβολής είναι χαμηλότερα, καθώς θα είναι δυνατόν να επιτύχει το ίδιο SIR στο ΝΕ με χαμηλότερη ισχύ εκπομπής. Έτσι, η ύπαρξη διαφορετικών διαθέσιμων ζωνών συχνοτήτων προσφέρει τη δυνατότητα στα ΚΤ να βελτιώσουν τη λειτουργία τους, ελαττώνοντας την κατανάλωση ενέργειας κι αυξάνοντας τη διάρκεια ζωής της μπαταρίας τους. Στο επόμενο υποκεφάλαιο, μελετάται το πρόβλημα της επιλογής ζώνης συχνοτήτων.

7.5 Μη-Συνεργατική Επιλογή Ζώνης Συχνοτήτων

Σε αυτό το υποκεφάλαιο, μελετάται το πρόβλημα της κατανομημένης επιλογής ζώνης συχνοτήτων στο εξεταζόμενο CDMA περιβάλλον με δυναμική φασματική πρόσβαση. Βάσει των παρατηρήσεων του προηγούμενου υποκεφαλαίου (δηλαδή ότι τα ΚΤ πάντα απολαμβάνουν το ίδιο SIR στο ΝΕ, αλλά το επιτυγχάνουν με διαφορετική ισχύ, αναλόγως με την κατάσταση της ζώνης συχνοτήτων εκπομπής), τα ΚΤ μπορούν αυτόνομα να μεταπηδούν από μία ζώνη συχνοτήτων σε μία άλλη, ώστε να ελαττώνουν τα ενεργειακά τους κόστη και να αυξήσουν την ωφέλεια τους. Το συγκεκριμένο πρόβλημα μοντελοποιείται ως ένα μη-συνεργατικό παίγνιο (band selection game – BSG).

Στο υποκεφάλαιο 7.4, αποδείχθηκε ότι το SIR ενός ΚΤ στο ΝΕ (και κατ' επέκταση η ρυθμοαπόδοση) είναι ανεξάρτητο από τη ζώνη συχνοτήτων εκπομπής, σε αντίθεση με την αντίστοιχη ισχύ εκπομπής. Έτσι, σε ό,τι αφορά την ωφέλεια του ΚΤ στο BSG, μόνον η ισχύς εκπομπής είναι σημαντική, που σημαίνει ότι η συνάρτηση ωφελείας (7.4) μπορεί να απλοποιηθεί. Από την εξίσωση (7.7), παρατηρείται ότι, όσο μεγαλύτερη είναι η τιμή του όρου

$$I_{c,a} = \sum_{k=1, c_k=c, a_k=a}^N \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1}, \quad (7.10)$$

ή της παρεμβολής $\rho_{c,a}^2$, τόσο μεγαλύτερη γίνεται κι η ισχύς του ΚΤ στο ΝΕ, καθώς η ισχύς του ΚΤ j είναι

$$q_j^* = \frac{\rho_{c,a}^2}{\left(\frac{W}{R\gamma_j^*} + 1\right) \left(1 - \sum_{k=1, c_k=c, a_k=a}^N \left(\frac{W}{R\gamma_k^*} + 1\right)^{-1}\right)} = \left(\frac{W}{R\gamma_j^*} + 1\right)^{-1} \frac{\rho_{c,a}^2}{1 - I_{c,a}}.$$

Μία δικαιολογημένη επιλογή για την απλοποιημένη συνάρτηση ωφελείας είναι, λοιπόν, η επόμενη μετρική φόρτου που αντανακλά το μέτρο της ισχύος εκπομπής:

$$Q_{c,a} = \frac{\rho_{c,a}^2}{1 - I_{c,a}}, \quad (7.11)$$

Πρέπει να σημειωθεί ότι η μετρική $Q_{c,a}$ υποδηλώνει την λαμβανόμενη ισχύ (στο NE) ενός ΚΤ στο ΣΒ, κανονικοποιημένη κατά το φόρτο του συγκεκριμένου ΚΤ, δηλαδή

$$Q_{c,a} = q_j^* / \left(\frac{W}{R\gamma_j^*} + 1\right)^{-1}.$$

Συγκεκριμένα, απεικονίζει την επίδραση των άλλων ΚΤ στην ισχύ (στο NE) ενός ΚΤ που εξυπηρετείται από το ΣΒ a , στη ζώνη συχνοτήτων c . Είναι φανερό ότι, αν $Q_{c,a} < Q_{c',a}$, η ζώνη συχνοτήτων c είναι προτιμότερη από τη ζώνη συχνοτήτων c' , καθώς η ισχύς στο NE θα είναι χαμηλότερη στην πρώτη σε σχέση με αυτή στη δεύτερη (ενώ το SIR που θα επιτευχθεί θα είναι το ίδιο). Έτσι, ένα ΚΤ που εκπέμπει σε μία ζώνη συχνοτήτων έχει συμφέρον να μεταπηδήσει στη ζώνη συχνοτήτων με την μικρότερη τιμή της παραπάνω μετρικής. Παρατηρείται, όμως, ότι η μετρική $Q_{c,a}$ είναι «έγκυρη» όταν η συνθήκη CAC (7.8) είναι αληθής (δηλαδή $I_{c,a} < 1$). Σε διαφορετική περίπτωση, το $Q_{c,a}$ παίρνει αρνητικές τιμές (ή μπορεί να γίνει κι άπειρο). Έτσι, η μετρική $Q_{c,a}$ είναι, μάλλον, ακατάλληλη για χρήση στο μελετούμενο περιβάλλον δυναμικής φασματικής πρόσβασης¹³. Για να ξεπεραστεί το συγκεκριμένο πρόβλημα, ορίζεται μία εναλλακτική μετρική, $L_{c,a}$, η οποία είναι αύξουσα κι ως προς το $I_{c,a}$ κι ως προς το $\rho_{c,a}^2$, ακόμα κι αν η συνθήκη CAC παραβιάζεται. Η μετρική $L_{c,a}$ ορίζεται ως εξής:

¹³ Σε ένα περιβάλλον δυναμικής φασματικής πρόσβασης, το σύνολο των διαθέσιμων ζωνών συχνοτήτων μπορεί να είναι ικανό να υποδεχθεί όλα τα ΚΤ, αν τα ΚΤ είναι ομοιόμορφα κατανομημένα στις ζώνες συχνοτήτων. Αν, όμως, δε συμβαίνει κάτι τέτοιο, και η πλειοψηφία των ΚΤ είναι συγκεντρωμένη σε συγκεκριμένες μόνο ζώνες συχνοτήτων, αυτές οι ζώνες συχνοτήτων μπορεί να μην καταφέρουν να προσφέρουν την απαραίτητη χωρητικότητα.

$$L_{c,a} = \rho_{c,a}^2 \cdot I_{c,a} \quad (7.12)$$

Η μετρική $L_{c,a}$ προσεγγίζει την επίδραση των άλλων ΚΤ, που εξυπηρετούνται από ένα ΣΒ και της εξω-κυψελικής παρεμβολής στη ζώνη συχνοτήτων c , στην ισχύ στο ΝΕ ενός ΚΤ που εξυπηρετείται από τον ίδιο ΣΒ, στην ίδια ζώνη συχνοτήτων.

Ένα ΚΤ θα μεταπηδήσει σε μία νέα ζώνη συχνοτήτων μόνο αν ο φόρτος σε αυτή θα είναι χαμηλότερος από το φόρτο στην τρέχουσα ζώνη συχνοτήτων του. Πριν αλλάξει ζώνη συχνοτήτων το ΚΤ πρέπει να υπολογίσει τη συνεισφορά του σε σχέση και τον παράγοντα φόρτου $I_{c,a}$ και την εξω-κυψελική παρεμβολή $\rho_{c,a}^2$ στην υποψήφια ζώνη συχνοτήτων c . Ο παράγοντας φόρτου, μετά από μία υποθετική μεταπήδηση σε μία ζώνη συχνοτήτων, $I'_{c,a}$, μπορεί να υπολογιστεί πολύ απλά ως εξής:

$$I'_{c,a} = I_{c,a} + \left(\frac{W}{R\gamma_j^*} + 1 \right)^{-1} \quad (7.13)$$

Ο ακριβής, όμως, υπολογισμός της επίδρασης της μεταπήδησης στη νέα ζώνη συχνοτήτων στην εξω-κυψελική παρεμβολή $\rho_{c,a}^2$ είναι, μάλλον, δύσκολος. Μπορεί να προσδιοριστεί μόνο αν υπάρχει καθολική γνώση για το σύστημα, δηλαδή αν το SIR στο ΝΕ κάθε ΚΤ κι οι απώλειες ισχύος από κάθε ΚΤ σε κάθε ΣΒ είναι γνωστά. Είναι φανερό ότι κάτι τέτοιο δεν είναι εφικτό στην πράξη, καθώς απαιτεί τη διάχυση σημαντικού όγκου πληροφοριών, μέσω του ασύρματου δικτύου. Επιπλέον, αυτές οι πληροφορίες και, ειδικότερα, οι τιμές των απωλειών σήματος, μεταβάλλονται πολύ γρήγορα σε κινητά περιβάλλοντα όπως το μελετούμενο.

Για την αντιμετώπιση του προβλήματος του υπολογισμού της εξω-κυψελικής παρεμβολής, μετά από μία υποθετική μεταπήδηση σε μία ζώνη συχνοτήτων, $\rho_{c,a}^{\prime 2}$, έγινε χρήση ενός απλού σχήματος πρόβλεψης που βασίζεται στην παρατήρηση ότι η εξω-κυψελική παρεμβολή είναι αύξουσα ως προς τον παράγοντα φόρτου $I_{c,a}$ και συγκεκριμένα ως προς όρο $1/(1 - I_{c,a})$ (για περισσότερες πληροφορίες, δες το Παράρτημα III). Βάσει της προηγούμενης παρατήρησης προβλέπεται η εξω-κυψελική παρεμβολή χρησιμοποιώντας μόνον τον παράγοντα φόρτου μετά από τη μεταπήδηση στη νέα ζώνη συχνοτήτων $I'_{c,a}$ και την τρέχουσα εξω-κυψελική παρεμβολή $\rho_{c,a}^2$, δηλαδή τοπική πληροφορία. Συγκεκριμένα, υιοθετείται ένα απλό

μοντέλο γραμμικής παρεκβολής (extrapolation) που προσεγγίζει την αυξητική τάση της εξω-κυψελικής παρεμβολής ως προς τον παράγοντα φόρτου.

$$\rho'_{c,a} = \begin{cases} \frac{1-l_{c,a}}{1-l'_{c,a}} \rho_{c,a}^2, & l'_{c,a} < 1 \\ l'_{c,a} \rho_{c,a}^2, & \text{otherwise} \end{cases} \quad (7.14)$$

Όταν $l'_{c,a} \geq 1$, αντί για τον παράγοντα $1/(1-l_{c,a})$, χρησιμοποιείται ο παράγοντας $l_{c,a}$, ώστε οι προβλέψεις να έχουν νόημα (μη-αρνητική ή άπειρη εξω-κυψελική παρεμβολή).

Ανακεφαλαιώνοντας, τα ΚΤ, πριν πραγματοποιήσουν μία αλλαγή ζώνης συχνοτήτων, συμβουλευόμαστε τη μετρική $L_{c,a}$, η οποία ποσοτικοποιεί την ποιότητα μίας ζώνης συχνοτήτων λαμβάνοντας υπόψη και τον παράγοντα φόρτου $l_{c,a}$ και την εξω-κυψελική παρεμβολή $\rho_{c,a}^2$. Πριν από την αλλαγή της ζώνης συχνοτήτων, το ΚΤ λαμβάνει επίσης υπόψη τη δική του συνεισφορά στον παράγοντα φόρτου $l_{c,a}$ και την εξω-κυψελική παρεμβολή $\rho_{c,a}^2$ στη νέα ζώνη συχνοτήτων, ούτως ώστε να είναι εξασφαλισμένο ότι η νέα ζώνη συχνοτήτων θα είναι καλύτερη από την τρέχουσα. Στο υπόλοιπο τους παρόντος υποκεφαλαίου, μελετούνται ζητήματα σχετικά με τις ισορροπίες που επιτυγχάνονται στο εξεταζόμενο παίγνιο και την αποτελεσματικότητά τους.

7.5.1 Nash Ισορροπία του BSG

Όπως έχει ήδη αναφερθεί στο υποκεφάλαιο 7.4, ένα σημαντικό ζήτημα στα μη-συνεργατικά παίγνια είναι τα NE που μπορούν να προκύψουν. Στο BSG, ένα NE είναι μία κατανομή ζωνων συχνοτήτων από την οποία τα ΚΤ δεν έχουν κίνητρο να αποκλίνουν, πραγματοποιώντας ατομικές αλλαγές στη στρατηγική τους. Παρακάτω, παρέχεται ο ορισμός του NE για το BSG.

Ορισμός 7.2 (Ισορροπία Nash του Παιγνίου Επιλογής Ζώνης Συχνοτήτων): Μία έκβαση του BSG είναι ένα NE αν, για κάθε ΚΤ j που εξυπηρετείται από το ΣΒ a και για όλες τις εκβάσεις που $c'_k = c_k$ και $c'_j \neq c_j$, ισχύει ότι $L_{c_j,a} \leq L_{c'_j,a}$.

Στο NE του BSG (Ορισμός 7.2), αν γίνει άθροιση για όλες τις διαθέσιμες ζώνες συχνοτήτων, παρατηρείται ότι, στο NE του BSG,

$$KL_{c_j,a} \leq \sum_{c'_j \in C} L_{c'_j,a}, \text{ ή } L_{c_j,a} \leq \frac{1}{K} \sum_{c'_j \in C} L_{c'_j,a}, \forall j \in J,$$

ή ισοδύναμα ότι

$$L_{c,a} = \frac{1}{K} \sum_{c' \in C} L_{c',a}, \forall c \in C. \quad (7.15)$$

Με άλλα λόγια, στο NE του BSG, ο φόρτος στις υπάρχουσες ζώνες συχνοτήτων είναι εξισορροπημένος, δηλαδή όλα τα ΚΤ παρατηρούν τον ίδιο φόρτο. Έτσι το NE αποτελεί ένα δίκαιο σημείο λειτουργίας του συστήματος. Στην επόμενη παράγραφο, μελετάται η αποτελεσματικότητα του NE του BSG, που αποτελεί μία επίσης σημαντική ιδιότητα.

7.5.2 Αξιολόγηση Αποδοτικότητας του NE του BSG

Για την αξιολόγηση της αποτελεσματικότητας του NE του BSG, γίνεται σύγκριση με «βέλτιστες» κατανομές ζωνών συχνοτήτων, οι οποίες ορίζονται ως αυτές που μεγιστοποιούν το λόγο του αριθμού των επιτυχημένα ληφθέντων bits προς την ενέργεια που έχει καταναλωθεί από το σύνολο των ΚΤ (δες Παράρτημα IV). Τέτοιες κατανομές αναμένεται να διαφέρουν από το NE του BSG, καθώς οι καθολικοί στόχοι του συστήματος δεν ταυτίζονται κατ' ανάγκη με τους ατομικούς, εγωιστικούς στόχους κάθε ΚΤ.

Στην παράγραφο 7.5.1, ορίστηκε η μετρική $Q_{c,a}$ (7.11) η οποία απεικονίζει την επίδραση του φόρτου $l_{c,a}$ που προσφέρεται στο ΣΒ a και της εξω-κυψελικής παρεμβολής $\rho_{c,a}^2$ στην ισχύ στο NE ενός ΚΤ που εξυπηρετείται από το συγκεκριμένο ΣΒ. Παρόλο που η συγκεκριμένη μετρική βρέθηκε ακατάλληλη για το μελετούμενο περιβάλλον (λόγω του περιορισμού στο φόρτο – $l_{c,a} < 1$), περιλαμβάνεται στην παρούσα αξιολόγηση. Συγκεκριμένα, μελετάται το NE του BSG που θα προέκυπτε αν, αντί της μετρικής $L_{c,a}$, χρησιμοποιούταν η μετρική $Q_{c,a}$.

Σε ένα τέτοιο ΝΕ, ο φόρτος θα είναι ομοιόμορφα κατανεμημένος στις υπάρχουσες ζώνες συχνοτήτων, δηλαδή θα ισχύει το παρακάτω¹⁴:

$$Q_{c_j,a} = \frac{1}{K} \sum_{c=1}^K Q_{c,a}, \forall j \in J, \text{ ή } Q_{c,a} = \frac{1}{K} \sum_{c'=1}^K Q_{c',a}, \forall c \in C.$$

Από τα παραπάνω, υπονοείται ότι $Q_{j,a} = Q_{k,a}, \forall j,k \in C$, ή

$$1 - I_{j,a} = \frac{\rho_{j,a}^2}{\rho_{k,a}^2} (1 - I_{k,a}). \quad (7.16)$$

Δεδομένου του συνολικού φόρτου $I_{tot,a}$ ($I_{tot,a} = \sum_{c=1}^K I_{c,a}$) που προσφέρεται από τα ΚΤ που εξυπηρετούνται από το ΣΒ a , προκύπτει ότι

$$I_{k,a} = 1 - \frac{\rho_{k,a}^2}{\sum_{c=1}^K \rho_{c,a}^2} (K - I_{tot,a}).$$

Όμως, πρέπει $I_{k,a} \geq 0, \forall k \in C$. Έτσι, ο φόρτος που προκύπτει στο ΝΕ, από εδώ και στο εξής αναφερόμενος ως $I_{k,a}^Q$, είναι

$$I_{k,a}^Q = \begin{cases} 1 - \frac{\rho_{k,a}^2}{\sum_{c=1}^{K''} \rho_{c,a}^2} (K'' - I_{tot,a}), & k \in \{1, \dots, K''\} \\ 0, & k \in \{K'' + 1, \dots, K\} \end{cases}, \quad (7.17)$$

όπου, χωρίς βλάβη της γενικότητας, οι ζώνες συχνοτήτων ταξινομούνται, βάσει του επιπέδου της εξω-κυψελικής παρεμβολής, δηλαδή $\rho_{1,a}^2 < \dots < \rho_{K,a}^2$. Το K'' είναι ο μεγαλύτερος ακέραιος από το σύνολο $\{1, \dots, K\}$ όπου $I_{k,a} \geq 0, \forall k \in \{1, \dots, K''\}$. Δηλαδή, οι τιμές $K, K-1, \dots$ δοκιμάζονται διαδοχικά, μέχρι η προκύπτουσα λύση να είναι μη-αρνητική.

Πρέπει να παρατηρηθεί ότι ο φόρτος $I_{k,a}^Q$ είναι παρόμοιας μορφής με τον βέλτιστο φόρτο $I_{k,a}^{opt}$ (δες και το Παράρτημα IV), όπου

¹⁴ Πρέπει να σημειωθεί ότι, όπως στο Παράρτημα IV, υιοθετείται η συνεχής χαλάρωση (continuous relaxation) του προβλήματος, αντιμετωπίζοντας τη μετρική φόρτου $I_{c,a}$ ως συνεχή μεταβλητή.

$$I_{k,a}^{opt} = \begin{cases} 1 - \frac{\sqrt{\rho_{k,a}^2}}{\sum_{c=1}^{K'} \sqrt{\rho_{c,a}^2}} (K' - I_{tot,a}), k \in \{1, \dots, K'\} \\ 0, k \in \{K' + 1, \dots, K\} \end{cases} \quad (7.18)$$

Στην εξίσωση (7.18), οι ζώνες συχνοτήτων ταξινομούνται επίσης βάσει του επιπέδου της εξω-κυψελικής παρεμβολής, δηλαδή $\rho_{1,a}^2 < \dots < \rho_{K',a}^2$ και K' είναι ο μεγαλύτερος ακέραιος από το σύνολο $\{1, \dots, K\}$ όπου $I_{k,a} \geq 0, \forall k \in \{1, \dots, K'\}$.

Τώρα, μπορεί να υπολογιστεί και το NE του BSG με τη μετρική $L_{c,a}$. Από την εξίσωση (7.15), προκύπτει ότι $L_{j,a} = L_{k,a}, \forall j, k \in C$, ή

$$I_{j,a} = \frac{\rho_{k,a}^2}{\rho_{j,a}^2} I_{k,a}. \quad (7.19)$$

Δεδομένου του συνολικού προσφερόμενου φόρτου από τα ΚΤ, δηλαδή $\sum_{c=1}^K I_{c,a} = I_{tot,a}$, προκύπτει ότι το φόρτος $I_{k,a}^L$ στο NE θα είναι

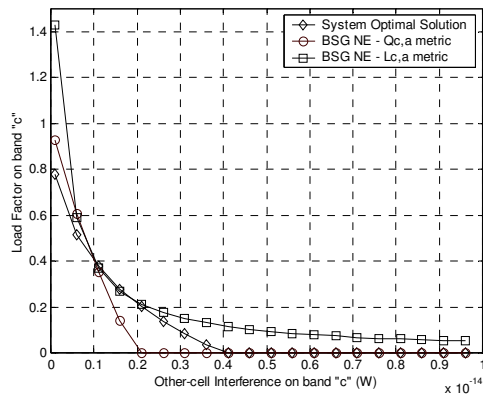
$$I_{k,a}^L = \frac{I_{tot,a}}{\rho_{k,a}^2 \sum_{c=1}^K \frac{1}{\rho_{c,a}^2}}. \quad (7.20)$$

Πρέπει να σημειωθεί ότι ο φόρτος στο NE με την μετρική $L_{c,a}$ ($I_{k,a}^L$) είναι διαφορετικής μορφής από τον αντίστοιχο με τη μετρική $Q_{c,a}$ ($I_{k,a}^Q$) και το βέλτιστο φόρτο ($I_{k,a}^{opt}$). Παρόλα αυτά, αν $\rho_{1,a}^2 = \dots = \rho_{K',a}^2 = \rho^2$ (δηλαδή οι ζώνες συχνοτήτων υπόκεινται σε ίσα επίπεδα εξω-κυψελικής παρεμβολής), η κατανομή του φόρτου μεταξύ των ζωνών συχνοτήτων είναι ομοιόμορφη για καθεμία από τις προαναφερθείσες μετρικές:

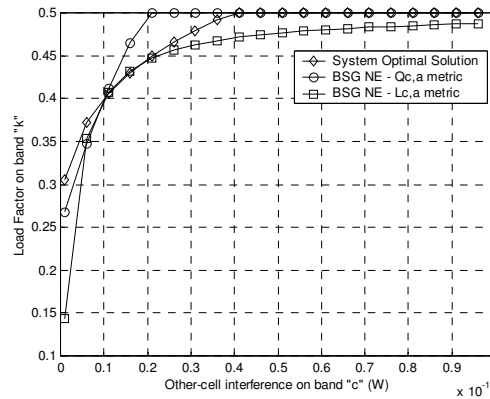
$$I_{j,a} = I_{k,a} = \frac{1}{K} I_{tot,a}, \forall j, k \in C.$$

Δηλαδή, η βέλτιστη λύση του συστήματος συμπίπτει με το NE του BSG. Με άλλα λόγια, στο σενάριο με συμμετρική παρεμβολή, η εγωιστική λειτουργία των ΚΤ οδηγεί σε βέλτιστη χρήση του φάσματος. Αν, όμως, δεν υφίσταται τέτοια

συμμετρία, οι επιμέρους κατανομές μπορούν να διαφέρουν, όπως φαίνεται στην Εικόνα 7.1 κι Εικόνα 7.2.



Εικόνα 7.1. Φόρτος σε μία ζώνη συχνοτήτων για διαφορετικές τιμές της εξω-κυψελικής παρεμβολής σε αυτήν τη ζώνη συχνοτήτων.



Εικόνα 7.2. Φόρτος σε μία ζώνη συχνοτήτων για διαφορετικές τιμές της εξω-κυψελικής παρεμβολής σε αυτήν τη ζώνη συχνοτήτων.

Στην Εικόνα 7.1 απεικονίζονται οι φόρτοι στο NE ($I_{k,a}^L$, $I_{k,a}^Q$, and $I_{k,a}^{opt}$) για μία δεδομένη ζώνη συχνοτήτων c . Η εξω-κυψελική παρεμβολή $\rho_{c,a}^2$ στη ζώνη συχνοτήτων c παίρνει τιμές από το σύνολο $[10^{-16}, 10^{-14}]$. Υπάρχουν ακόμα τέσσερις άλλες ζώνες συχνοτήτων με ίση εξω-κυψελική παρεμβολή της τάξης των 10^{-15} W. Ο συνολικός προσφερόμενος φόρτος είναι $I_{tot,a} = 2$. Η Εικόνα 7.2 απεικονίζει το φόρτο σε μία ζώνη συχνοτήτων $c' \neq c$ (δηλαδή μία από τις άλλες τέσσερις). Όπως ήταν αναμενόμενο, για εξω-κυψελική παρεμβολή ίση με 10^{-15} W (δηλαδή ίση εξω-κυψελική παρεμβολή και για τις πέντε ζώνες συχνοτήτων), το BSG NE (και για την μετρική $Q_{c,a}$ και για την μετρική $L_{c,a}$) οδηγεί στο βέλτιστο σημείο λειτουργίας του συστήματος, όπως φαίνεται από την Εικόνα 7.1 και την Εικόνα 7.2 (οι τρεις καμπύλες τέμνονται). Για μη-συμμετρική εξω-κυψελική παρεμβολή, διακρίνεται μία διαφορά στις κατανομές μεταξύ των διαφορετικών σχημάτων. Παρατηρείται, όμως, ότι αυτή η διαφορά δεν είναι σημαντική, γεγονός που επιτρέπει στο NE να θεωρηθεί σχεδόν βέλτιστο (near-optimal).

Η ευκολία της πραγματοποίησης του BSG αποτελεί ένα ακόμα χαρακτηριστικό υπέρ της κατανομημένης επιλογής ζώνης συχνοτήτων. Το κύριο πρόβλημα με τη βέλτιστη χρήση του φάσματος είναι πως μπορεί να επιτευχθεί στην πράξη.

Βέλτιστα σημεία λειτουργίας είναι αρκετά δύσκολο να επιτευχθούν σε μη-συνεργατικά περιβάλλοντα κι ακόμα πιο δύσκολο να διατηρηθούν, καθώς κάθε ΚΤ μπορεί να έχει συμφέρον να αποκλίνει από ένα τέτοιο σημείο λειτουργίας. Ο ΣΒ δεν είναι πάντα εφικτό να επιβάλλει κεντροποιημένο έλεγχο, καθώς δε μπορεί να επέμβει στη λειτουργία των ΚΤ, ούτως ώστε να οδηγήσει το σύστημα στη βέλτιστη κατανομή ζωνών συχνοτήτων. Από την άλλη, σε ένα κατανεμημένο περιβάλλον, το δίκτυο χρειάζεται μόνο να εκπέμπει καθολικές πληροφορίες του συστήματος, αφήνοντας τα ΚΤ να πάρουν ανεξάρτητα τις αποφάσεις τους.

7.5.3 Ζητήματα Σύγκλισης

Ένα σημαντικό ζήτημα στο μελετούμενο πρόβλημα αποτελεί η κατανεμημένη σύγκλιση στο NE του BSG. Σε αντίθεση με ένα σχήμα κεντροποιημένου ελέγχου, η κατανεμημένη φύση του μελετούμενου περιβάλλοντος είναι δυνατόν να καταστήσει τη σύγκλιση του παιγνίου επιλογής ζώνης συχνοτήτων προβληματική. Στο προτεινόμενο μοντέλο, μετά από μία ενημέρωση από το ΣΒ, τα ΚΤ λαμβάνουν τις αποφάσεις τους ταυτόχρονα, με κατανεμημένο τρόπο¹⁵. Έτσι, οι αλληλεπιδράσεις μεταξύ των ΚΤ πραγματοποιούνται μόνο από την συνδυασμένη χρήση των διαμοιραζόμενων πόρων. Για το λόγο αυτό, αν τα ΚΤ εφαρμόζουν στρατηγικές με βάση την αρχή της καλύτερης απόκρισης, αναμένεται το σύστημα να υποφέρει από ταλαντώσεις (δηλαδή τα ΚΤ να μεταπηδούν από μία ζώνη συχνοτήτων σε μία άλλη) [27], κάνοντας τη σύγκλιση στο NE του BSG προβληματική ή αδύνατη. Έτσι, τα ΚΤ χρειάζεται να υιοθετήσουν κατανεμημένες στρατηγικές που να λαμβάνουν υπόψη αυτό το χαρακτηριστικό του BSG, ούτως ώστε να αποφεύγονται οι ταλαντώσεις και να επιτυγχάνεται σύγκλιση σε ένα NE του BSG. Στο επόμενο υποκεφάλαιο προτείνονται δύο κατανεμημένες στρατηγικές. Η πρώτη είναι εμπνευσμένη από το SFBP [27] (δες Κεφάλαιο 3, παράγραφο 3.4.1), ενώ η δεύτερη βασίζεται σε στοχαστική λήψη αποφάσεων από το ΚΤ.

¹⁵ Δε γίνεται η υπόθεση ακολουθιακού σχήματος αποφάσεων, καθώς το εξεταζόμενο περιβάλλον είναι πλήρως κατανεμημένο και τα ΚΤ είναι εγωιστικά. Είναι, μάλλον, δύσκολο να επιβάλλει ο ΣΒ κάποια συγκεκριμένη ακολουθία αποφάσεων στα ΚΤ.

7.6 Κατανεμημένες Στρατηγικές για το BSG

Σε αυτό το υποκεφάλαιο, περιγράφονται δύο κατανεμημένες στρατηγικές για χρήση στο BSG. Η πρώτη στρατηγική είναι βασισμένη σε λήψη αποφάσεων με τη βοήθεια προβλέψεων, ενώ η δεύτερη σε στοχαστική λήψη αποφάσεων. Οι προτεινόμενες στρατηγικές προσβλέπουν στην αποφυγή ταλαντώσεων και τη σύγκλιση του BSG σε ένα NE.

7.6.1 Επιλογή Ζώνης Συχνοτήτων Βάσει Προγνώσεων

Στην παρούσα παράγραφο, περιγράφεται μία κατανεμημένη στρατηγική που μπορεί να υιοθετηθεί στο μελετούμενο περιβάλλον για σύγκλιση στο NE του BSG. Η αρχική ιδέα προέρχεται από το SFBP, που σκοπό έχει την αντιμετώπιση των ταλαντώσεων στα επίπεδα προσέλευσης σε ένα υποθετικό μπαρ. Το πρόβλημα των ταλαντώσεων παρατηρείται και στο BSG: τα ΚΤ που βρίσκουν μία μη φορτωμένη ζώνη συχνοτήτων θα την καταλάβουν, πράξη που στη συνέχεια θα οδηγήσει στον κορεσμό της συγκεκριμένης ζώνης συχνοτήτων. Αυτό με τη σειρά του θα κάνει όλα τα ΚΤ να εγκαταλείψουν τη ζώνης συχνοτήτων στον επόμενο γύρο κοκ. Έτσι, όπως και στο SFBP, για να αποφευχθούν οι ταλαντώσεις, είναι απαραίτητο τα ΚΤ να υιοθετούν διαφορετικό «τρόπο σκέψης».

Με βάση την προσέγγιση για τη λύση του SFBP, προτείνεται στην παρούσα μελέτη η επαύξηση των ΚΤ με ευφυΐα πρόγνωσης. Γίνεται η υπόθεση ότι ο χρόνος είναι διακριτός, με διάρκεια χρονοθυρίδας T_c (περίοδος ενημέρωσης πληροφορίας ζωνών συχνοτήτων – δες παράγραφο 7.6.4). Κάθε ΚΤ κατέχει ένα σύνολο από προγνώστες, οι οποίοι παίρνουν ως είσοδο την ιστορία του παράγοντα φόρτου $l_{c,a}$ ($l_{c,a}(\tau)$, $\tau = n, n-1, \dots$) για κάθε διαθέσιμη ζώνη συχνοτήτων $c \in C$, στο ΣΒ a που είναι συνδεδεμένο το ΚΤ. Κάθε προγνώστης, τότε, βάσει αυτής της ιστορίας, παράγει ως έξοδο το $P_{c,a}(n)$, που αποτελεί την πρόβλεψη του παράγοντα φόρτου για την ζώνη συχνοτήτων c στο γύρο $n+1$.

Σε συμφωνία με τη λύση του SFBP, ένας μόνο προγνώστης είναι ενεργός για κάθε χρήστη, κάθε στιγμή. Ένας προγνώστης ενεργοποιείται εάν είχε επιτύχει την πιο ακριβή πρόβλεψη στον προηγούμενο γύρο (Κεφάλαιο 3). Αυτή η διαδικασία περιγράφεται με ακρίβεια ως εξής. Έστω $O_{c,a,i}(n)$ η έξοδος του i -οστού προγνώστη

ενός ΚΤ που εξυπηρετείται από το ΣΒ a κατά τη χρονοθυρίδα n στη ζώνη συχνοτήτων c , δεδομένης της ιστορίας των μετρήσεων, για τις H τελευταίες περιόδους, με την πιο πρόσφατη να έχει ληφθεί τη χρονοθυρίδα n . Έστω $\theta(c,n)$ ο δείκτης του ενεργού προγνώστη, για τη ζώνη συχνοτήτων c τη χρονοθυρίδα n . Η προβλεπόμενη προσέλευση $P_{c,a}(n)$ (δηλαδή, η εκτίμηση για τη χρονοθυρίδα $n+1$), θα είναι

$$P_{c,a}(n) = O_{c,a,\theta(c,n)}(n), \forall c \in C. \quad (7.21)$$

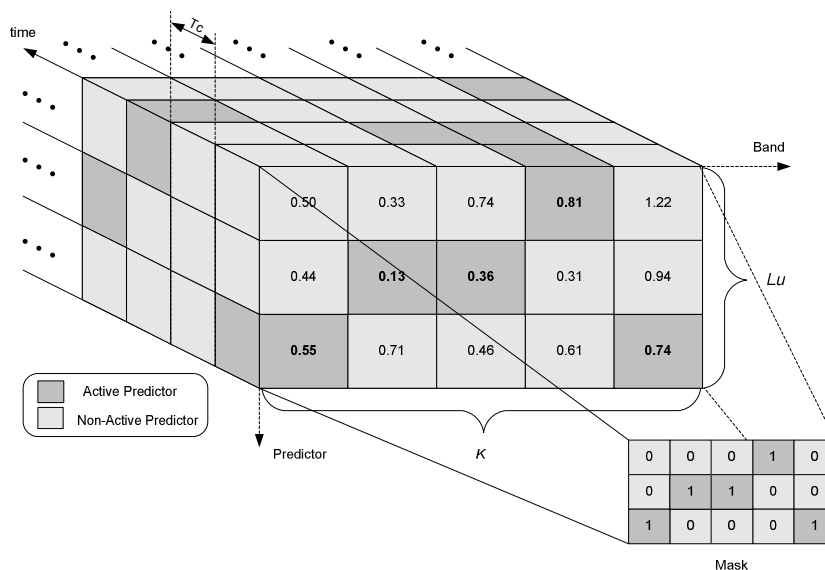
Υπό το φως της τρέχουσας μέτρησης του παράγοντα φόρτου στη ζώνη συχνοτήτων c , $R_{c,a}(n)$, η επιλογή του ενεργού προγνώστη, για τον επόμενο γύρο, πραγματοποιείται ως εξής:

$$\theta(c,n) = \arg \min_{i \in D_j} \{R_{c,a}(n) - O_{c,a,i}(n-1)\}, \forall c \in C. \quad (7.22)$$

Με άλλα λόγια, κάθε ΚΤ, μετά από μία ενημέρωση σχετικά με την κατάσταση των διαθέσιμων ζωνών συχνοτήτων, αποφασίζει ποιος θα είναι ο ενεργός προγνώστης μέχρι την επόμενη ενημέρωση, για κάθε ζώνη συχνοτήτων. Αυτή η διαδικασία απεικονίζεται, επίσης, στην Εικόνα 7.3. Στο μέτωπο της τρισδιάστατης δομής φαίνεται η τρέχουσα κατάσταση των L_U προγνωστών που έχουν ανατεθεί στο ΚΤ για τις K διαθέσιμες ζώνες συχνοτήτων, που μπορεί να αναπαρασταθεί από έναν δισδιάστατο δυαδικό πίνακα με διαστάσεις $L_U \times K$. Τα στοιχεία φόρτου στην τρισδιάστατη δομή υποδηλώνουν την ιστορία των προβλέψεων των προγνωστών.

Μέχρι στιγμής, δεν έχει γίνει αναφορά σε σχέση με τους τύπους των προγνωστών που μπορούν να χρησιμοποιηθούν από τα ΚΤ στο BSG. Ο Πίνακας 7.1 περιέχει τις κατηγορίες των προγνωστών που χρησιμοποιήθηκαν. Παρόμοιοι προγνώστες έχουν υιοθετηθεί και για το SFBP [27], αλλά και στο Κεφάλαιο 4 για το σύστημα MITOS. Για λόγους απλότητας, παραλείπονται οι δείκτες που αφορούν στη ζώνη συχνοτήτων και το ΣΒ.

Σύμφωνα με την εργασία [27], είναι σημαντικό να σημειωθεί ότι οι προγνώστες αυτοί καθ' αυτοί δεν είναι σημαντικοί, δηλαδή δεν υπάρχουν «καλοί» ή «κακοί» προγνώστες. Η ετερογένεια τους είναι πολύ πιο σημαντική. Η ποιότητα του συνόλου των διαθέσιμων προγνωστών δεν προκύπτει από τους ατομικούς προγνώστες που περιέχει, αλλά από την ποικιλία και τη διαφορετικότητα τους.



Εικόνα 7.3. Επισκόπηση της λειτουργίας βάσει προβλέψεων του SFBP. Κάθε στιγμή, δεδομένων των ενημερώσεων που παρέχει το σύστημα, παράγεται ένα νέο σύνολο από προβλέψεις (εδώ, 3 προβλέψεις για κάθε ζώνη συχνότητας). Σύμφωνα με την ακρίβεια των προβλέψεων στην προηγούμενη χρονοθυρίδα, επιλέγεται ο προγνώστης που θα ενεργοποιηθεί για κάθε ζώνη συχνότητας (δες (7.22)). Σε αυτό το παράδειγμα, φαίνεται ότι από τους ενεργούς προγνώστες, αυτός που προέβλεψε το χαμηλότερο φόρτο (δηλαδή, 0.13) είναι της δεύτερης ζώνης συχνότητας. Έτσι, αν τα ΚΤ αποφασίσει να φύγει από τη ζώνη συχνότητας που εκπέμπει, θα επιλέξει τη δεύτερη ζώνη συχνότητας.

Πίνακας 7.1. Τύποι Προγνώστων

Κατηγορία Προγνώστη	Περιγραφή	Συνάρτηση Πρόγνωσης
1	Όμοια όπως πριν i χρονοθυρίδες (κυκλικός προγνώστης)	$P(n) = I(n - i), i = 0,1,\dots$
2	Κατοπτρικά γύρω από το 0.5 (μισή της χωρητικότητας πόλου) της τελευταίας χρονοθυρίδας	$P(n) = 0.5 + \text{sign}(0.5 - \min\{I(n),1\}) \cdot \min\{I(n),1\} - 0.5 $ $= 1 - \min\{I(n),1\}$
3	Ίσα με $const$ (στατικός προγνώστης)	$P(t) = const, const \in [0,1]$
4	Μέση τιμή των τελευταίων y χρονοθυρίδων (προγνώστης μέσης τιμής)	$P(n) = \frac{1}{y} \sum_{i=0}^{y-1} I(n - i), y = 1,2,\dots$
5	Μέση τιμή με βάρη για τις τελευταίες y χρονοθυρίδες (προγνώστης μέσης τιμής)	$P(n) = \sum_{i=0}^{y-2} \frac{1}{2^{i+1}} I(n - i) + \frac{1}{2^{y-1}} I(n - (y - 1))$
6	Παρεμβολή από τις τελευταίες δύο χρονοθυρίδες	$P(n) = \max\{0, I(n) + (I(n) - I(n - 1))\}$ $= \max\{0, 2 \cdot I(n) - I(n - 1)\}$

Η προτεινόμενη κατανομημένη στρατηγική που βασίζεται στο SFBP έχει ως εξής: Ένα ΚΤ, αφού λάβει την ενημέρωση σχετικά με τον παράγοντα φόρτου και την εξω-κυβελική παρεμβολή, ελέγχει αν είναι ικανοποιημένο από την κατάσταση της τρέχουσας ζώνης συχνοτήτων του. Συγκεκριμένα, ελέγχει αν ο φόρτος $L_{c,a}$ στην τρέχουσα ζώνη συχνοτήτων είναι μικρότερος ή ίσος με τη μέση τιμή του φόρτου για όλες τις διαθέσιμες ζώνες συχνοτήτων (που αποτελεί το φόρτο στο NE – δεξ παράγραφο 7.5.1). Αν αυτό αληθεύει, το ΚΤ παραμένει στην ίδια ζώνη συχνοτήτων. Σε άλλη περίπτωση, μπορεί να επιλέξει μία άλλη λιγότερο φορτωμένη ζώνη συχνοτήτων. Εδώ, υπεισέρχεται η επιλογή ζώνης συχνοτήτων βάσει προβλέψεων. Το ΚΤ, βάσει των προγνωστών του, επιλέγει τη λιγότερο φορτωμένη ζώνη συχνοτήτων. Καθώς τα ΚΤ, κατά κανόνα, κατέχουν διαφορετικούς προγνώστες, είναι, μάλλον, απίθανο ότι τα «μη-ικανοποιημένα» ΚΤ θα μεταπηδήσουν μαζικά στην ίδια ζώνη συχνοτήτων. Κατά συνέπεια, οι ταλαντώσεις αποφεύγονται, και το BSG οδηγείται βαθμιαία στο NE.

Παρακάτω ορίζεται η βασισμένη σε προβλέψεις στρατηγική για το BSG, η οποία ονομάζεται *Santa Fe strategy for band selection* (SFBS).

Ορισμός 7.3 (Santa Fe strategy for Band Selection - SFBS): Δεδομένης μίας έκβασης του BSG, το ΚΤ j , που εξυπηρετείται από το ΣΒ a , χρησιμοποιώντας τη SFBS, θα κάνει τα ακόλουθα:

1. Αν $L_{c_j,a} \leq (F_a/K)$, παραμένει στη ζώνη συχνοτήτων c_j ,
2. Αλλιώς, επιλέγει την ζώνη συχνοτήτων c , όπου $c = \underset{c \in C}{\operatorname{argmin}} \{P_{c,a}(n)\rho_{c,a}'^2\}$

Όπου το $\rho_{c,a}'^2$ δίδεται στην εξίσωση (7.14) και $F_a = \sum_{c=1}^K L_{c,a}$.

Συνοψίζοντας, σύμφωνα με τη SFBS, σε περίπτωση που η μετρική φόρτου $L_{c,a}$ στη ζώνη συχνοτήτων c υπερβαίνει την εξισορροπημένη τιμή F_a/K , οι έξοδοι των προγνωστών λαμβάνονται υπόψη. Συγκεκριμένα, χρησιμοποιείται η εξίσωση (7.12), αλλά αντί του τρέχοντος φόρτου, $l_{c,a}(n)$, κάθε ΚΤ χρησιμοποιεί την τιμή του προβλεπόμενου φόρτου (δηλαδή $P_{c,a}(n)$). Το ΚΤ υπολογίζει επίσης την εξω-κυβελική παρεμβολή $\rho_{c,a}'^2$, όπως στην εξίσωση (7.14), ώστε να λάβει υπόψη και τη συνεισφορά του. Έτσι, κάθε ΚΤ, έχοντας μία διαφορετική εκτίμηση της μελλοντικής

κατανομής του φόρτου στις διαθέσιμες ζώνες συχνοτήτων, βρίσκει την ζώνη συχνοτήτων με την ελάχιστη (προβλεπόμενη) μετρική φόρτου. Στο επόμενο υποκεφάλαιο, περιγράφεται μία εναλλακτική στρατηγική, που βασίζεται στη στοχαστική λήψη αποφάσεων.

7.6.2 Στοχαστική Επιλογή Ζώνης Συχνοτήτων

Σε αυτό το υποκεφάλαιο, προτείνεται μία στοχαστική, κατανεμημένη στρατηγική για το BSG, από εδώ και στο εξής αναφερόμενη ως *probabilistic strategy for band selection* (PSBS). Σε αντίθεση με την SFBS, με την PSBS, τα ΚΤ δεν καταγράφουν την ιστορία των διαθέσιμων ζωνών συχνοτήτων ούτε χρησιμοποιούν προγνώστες φόρτου, αλλά απλά λαμβάνουν τις αποφάσεις τους, με βάση μόνο την τρέχουσα κατάσταση των διαθέσιμων ζωνών συχνοτήτων. Η διαφοροποίηση των αποφάσεων (η οποία είναι σημαντική για την αποφυγή των ταλαντώσεων) επιτυγχάνεται μέσω της στοχαστικότητας στην επιλογή ζώνης συχνοτήτων. Συγκεκριμένα, ένα ΚΤ, αντί να επιλέγει τη ζώνη συχνοτήτων που θα ήταν η καλύτερη για αυτό, *αν κανένα άλλο ΚΤ δε μεταπηδούσε σε άλλη ζώνη συχνοτήτων*, επιλέγει τυχαία από ένα σύνολο ζωνών συχνοτήτων. Το ΚΤ επίσης μπορεί να παραμείνει στην ίδια ζώνη συχνοτήτων παρόλο που μπορεί να υπάρχουν και καλύτερες ζώνες συχνοτήτων. Η στρατηγική PSBS, που έχει τις ρίζες της στα γνωστά DG [32] (δες παράγραφο 3.4.2), περιγράφεται αναλυτικά παρακάτω.

Το ΚΤ αφού λάβει μία ενημέρωση από το ΣΒ του (δες υποκεφάλαιο 7.6.4), ελέγχει αν είναι ικανοποιημένο από την τρέχουσα ζώνη συχνοτήτων. Συγκεκριμένα, όπως και στη στρατηγική SFBS, ελέγχει αν ο φόρτος $L_{c,a}$ στην τρέχουσα ζώνη συχνοτήτων είναι μικρότερος ή ίσος από τη μέση τιμή του φόρτου, για όλες τις διαθέσιμες ζώνες συχνοτήτων, δηλαδή το φόρτο στο ΝΕ. Αν αυτό αληθεύει, το ΚΤ παραμένει στην ίδια ζώνη συχνοτήτων (καθώς δεν μπορεί να υπάρξει περαιτέρω βελτίωση). Αλλιώς, το ΚΤ μπορεί να επιλέξει μία λιγότερο φορτωμένη ζώνη συχνοτήτων. Το ΚΤ, όμως, όπως έχει αναφερθεί και πιο πάνω, δεν επιλέγει απλά την καλύτερη ζώνη συχνοτήτων και μεταπηδά σε αυτή, γιατί αν όλα τα «μη-ικανοποιημένα» ΚΤ λειτουργούσαν κατ' αυτόν τον τρόπο, η τρέχουσα ζώνη συχνοτήτων θα εκκεωνόταν κι η νέα ζώνη συχνοτήτων θα έπρεπε να υποδεχθεί όλα τα ΚΤ. Για το λόγο αυτό, προτείνεται μία στρατηγική στην οποία 1) κάποια «μη-

ικανοποιημένα» ΚΤ παραμένουν στην ίδια ζώνη συχνοτήτων κι 2) όταν ένα ΚΤ αποφασίζει να εγκαταλείψει τη ζώνη συχνοτήτων του, επιλέγει τυχαία από ένα σύνολο ζωνών συχνοτήτων με χαμηλότερο φόρτο σε σχέση με την τρέχουσα ζώνη συχνοτήτων του. Έτσι, καθώς τα ΚΤ επιλέγουν τυχαία την επόμενη ζώνη συχνοτήτων τους, είναι μάλλον απίθανο να μεταπηδήσουν όλα στην ίδια ζώνη συχνοτήτων.

Παρακάτω, ορίζεται η στρατηγική PSBS.

Ορισμός 7.4 (Probabilistic Strategy for Band Selection - PSBS): Δεδομένης μίας έκβασης του BSG, το ΚΤ j , που εξυπηρετείται από το ΣΒ a , χρησιμοποιώντας τη στρατηγική PSBS, θα κάνει τα ακόλουθα:

1. Αν $L_{c_j,a} \leq (F_a/K)$, τότε επιλέγει τη ζώνη συχνοτήτων c_j με πιθανότητα 1,
2. Αλλιώς, επιλέγει τη ζώνη συχνοτήτων c_j με πιθανότητα $\nu = (F_a/K)/L_{c_j,a}$ και με πιθανότητα $1-\nu$ επιλέγει τυχαία μία από τις ζώνες συχνοτήτων συχνοτήτων c για τις οποίες
 - a) $L'_{c,b} < L_{c_j,a}$ και
 - b) $I'_{c,a} < 1$,

όπου $F_a = \sum_{c=1}^K L_{c,a}$ και $L'_{c,a} = \rho_{c,a}'^2 I'_{c,a}$. Επίσης το $I'_{c,a}$ είναι όπως ορίζεται στην εξίσωση (7.13) και το $\rho_{c,a}'^2$ όπως στην εξίσωση (7.14).

Με άλλα λόγια, ένα ΚΤ, αν δεν είναι ικανοποιημένο από το επίπεδο φόρτου στην τρέχουσα ζώνη συχνοτήτων του (δηλαδή η μετρική φόρτου $L_{c,a}$ ξεπερνά την εξισορροπημένη τιμή του φόρτου στο ΝΕ, F_a/K), εξετάζει την περίπτωση εγκατάλειψής της. Δεν αφήνει όμως πάντα την ζώνη συχνοτήτων του, αλλά μπορεί να παραμείνει σε αυτή με πιθανότητα ν , η οποία είναι αντιστρόφως ανάλογη του φόρτου της ζώνης συχνοτήτων. Αυτό εξασφαλίζει ότι δεν θα εγκαταλείψουν όλα τα ΚΤ την κορεσμένη ζώνη συχνοτήτων. Συγκεκριμένα, με αυτήν την πιθανότητα, ο αναμενόμενος φόρτος που θα απομείνει στη ζώνη συχνοτήτων, αφού όλα τα ΚΤ λάβουν τις αποφάσεις τους, θα είναι ο εξισορροπημένος φόρτος, F_a/K . Έτσι, σε κάθε βήμα, η κατανομή φόρτου πλησιάζει στο ΝΕ του BSG.

Αν ένα ΚΤ αποφασίσει να μεταπηδήσει σε μία νέα ζώνη συχνοτήτων, πρέπει να επιλέξει μία μεταξύ των υποψήφιων ζωνών συχνοτήτων. Σύμφωνα με την προτεινόμενη στρατηγική, οι υποψήφιες ζώνες συχνοτήτων είναι αυτές στις οποίες (a) ο φόρτος $L_{c,a}$ είναι χαμηλότερος ή ίσος με το επίπεδο φόρτου της τρέχουσας ζώνης συχνοτήτων και (b) η συνθήκη CAC αληθεύει. Το ΚΤ επιλέγει τυχαία μία ζώνη συχνοτήτων, με ίση πιθανότητα για κάθε ζώνη συχνοτήτων. Πρέπει να σημειωθεί ότι όταν αποτιμώνται οι (a) και (b), το ΚΤ λαμβάνει υπόψη τη δυνητική του συνεισφορά στις προαναφερθείσες συνθήκες (δηλαδή φόρτο, εξω-κυβελική παρεμβολή και χωρητικότητα (7.8)). Αυτός ο έλεγχος πραγματοποιείται, ώστε να αποφευχθούν δυνητικές ταλαντώσεις και τελικά το BSG να οδηγηθεί στο NE.

Στην επόμενη παράγραφο, περιγράφεται μία εναλλακτική στρατηγική που θα μπορούσε να χρησιμοποιηθεί στο BSG, αν ήταν δυνατόν να επιβληθεί μία συγκεκριμένη ακολουθία σύμφωνα με την οποία κάθε ΚΤ να λαμβάνει την απόφασή του.

7.6.3 Μία Εναλλακτική Στρατηγική Επιλογής Ζώνης Συχνοτήτων

Στις προηγούμενες παραγράφους, περιγράφηκαν δύο καταναμημένες στρατηγικές, (SFBS και PSBS) για το BSG. Και στις δύο στρατηγικές, υποτέθηκε ότι όλα τα ΚΤ, συγχρόνως, μετά από τη λήψη μίας ενημέρωσης από τον ΣΒ τους, λαμβάνουν τις αποφάσεις τους καταναμημένα. Μία εναλλακτική στρατηγική θα μπορούσε να βασιστεί σε ακολουθιακή λήψη αποφάσεων από τα ΚΤ, βάσει της αρχής της καλύτερης απόκρισης. Αυτή η προσέγγιση ονομάζεται και *elementary step system* (ESS). Η ESS έχει ελκυστικές θεωρητικές ιδιότητες (π.χ. σε κάθε βήμα το σύστημα «βελτιώνεται», ενώ σε παράλληλα καταναμημένα μοντέλα, όπως η SFBS, ή η PSBS, το σύστημα μπορεί επίσης να «χειροτερεύσει»). Η ESS έχει, όμως, κάποια σοβαρά μειονεκτήματα [94]: 1) καθώς μόνο ένα ΚΤ επιτρέπεται να επιλέξει μία νέα ζώνη συχνοτήτων σε κάθε βήμα, ο χρόνος σύγκλισης είναι τουλάχιστον $\Omega(N)$ και 2) η υλοποίηση μίας κεντρικής οντότητας για τον έλεγχο της σειράς, με την οποία τα ΚΤ θα λαμβάνουν τις αποφάσεις τους, δεν είναι ιδιαίτερα ελκυστική για ένα πραγματικά καταναμημένο σύστημα. Παρόλο που η ESS χαρακτηρίζεται από τα προηγούμενα μειονεκτήματα, συμπεριλαμβάνεται στη σύγκριση με τις προτεινόμενες στρατηγικές (PSBS και SFBS), καθώς έχει υιοθετηθεί στην εργασία

[92]. Πρέπει να σημειωθεί ότι στην εργασία [92] είχαν υποτεθεί διαφορετικές μετρικές φόρτου, λόγω της υπόθεσης της χρήσης TDMA. Η ιδέα του ESS μπορεί, όμως, να εφαρμοστεί και στο μελετούμενο περιβάλλον που βασίζεται σε CDMA.

7.6.4 Ένα Απλό Πρωτόκολλο για Δυναμική Πρόσβαση στο Φάσμα

Σε αυτήν την παράγραφο, περιγράφεται ένα απλό πρωτόκολλο που μπορεί υιοθετηθεί σε ένα πραγματικό δίκτυο, ούτως ώστε τα ΚΤ να ενημερώνονται σε σχέση με τις διαθέσιμες ελεύθερες ζώνες συχνοτήτων συχνοτήτων και του αντίστοιχου φόρτου τους. Γίνεται η υπόθεση ότι ο ΣΒ είναι ικανός να σαρώνει ένα σημαντικό εύρος συχνοτήτων και να ανιχνεύει ελεύθερες ζώνες συχνοτήτων, δηλαδή ζώνες συχνοτήτων που δεν καταλαμβάνονται από πρωτεύοντα ΚΤ. Ο ΣΒ είναι, επίσης, ικανός να υπολογίζει το βαθμό κορεσμού των ελεύθερων ζωνών συχνοτήτων. Σύμφωνα με όσα έχουν αναφερθεί στο υποκεφάλαιο 7.5, αυτή η πληροφορία απεικονίζεται μέσω των ποσοτήτων $\rho_{c,a}^2$ και $I_{c,a}$. Έτσι, ο ΣΒ a μπορεί να σχηματίσει για κάθε ελεύθερη ζώνη συχνοτήτων c ένα διάνυσμα της μορφής $\langle f_c, w_c, \rho_{c,a}^2, I_{c,a} \rangle$, όπου το f_c υποδηλώνει την κεντρική συχνότητα της ζώνης συχνοτήτων και το w_c το εύρος συχνοτήτων της ζώνης συχνοτήτων. Ο ΣΒ περιοδικά (κάθε T_c) ευρεκπέμπει όλα τα συγκεντρωμένα διανύσματα σε όλες τις ζώνες συχνοτήτων. Πρέπει να σημειωθεί ότι $T_c \gg T_p$, ούτως ώστε τα ΚΤ να μην μεταπηδήσουν σε μία νέα ζώνη συχνοτήτων, πριν συγκλίνει η διαδικασία ελέγχου ισχύος (δες εξίσωση (7.9)).

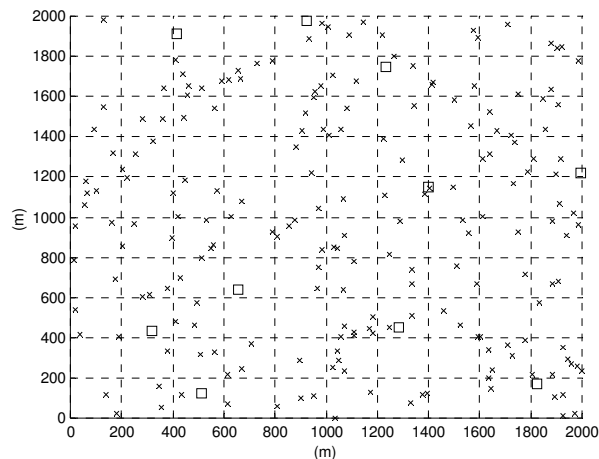
Εάν, ενώ κάποια δευτερεύοντα ΚΤ χρησιμοποιούν μία ζώνη συχνοτήτων, ο ΣΒ ανακαλύψει ότι ένα ή περισσότερα πρωτεύοντα ΚΤ ξεκίνησαν να εκπέμπουν, έχει υποχρέωση να ειδοποιήσει τα (δευτερεύοντα) ΚΤ να εγκαταλείψουν άμεσα τη ζώνη συχνοτήτων αυτή. Συγκεκριμένα, ένα μήνυμα ειδοποίησης ευρυ-εκπέμπεται στα ΚΤ της ζώνης συχνοτήτων, το οποίο περιέχει μία λίστα από *απαγορευμένες* ζώνες συχνοτήτων. Τα ΚΤ (που εκπέμπουν σε μία απαγορευμένη ζώνη συχνοτήτων) μεταπηδούν απευθείας σε μία *μη-απαγορευμένη* ζώνη συχνοτήτων¹⁶. Με αυτό το

¹⁶ Πρέπει να σημειωθεί ότι, παρόλο που τα ΚΤ θεωρούνται εγωιστικά και θα περίμενε κανείς ότι δεν έχουν λόγο να υπακούουν εντολές του ΣΒ σχετικά με απαγορευμένες ζώνες συχνοτήτων, η όλη ιδέα της ευέλικτης φασματικής πρόσβασης βασίζεται σε αυτήν την υπόθεση. Με άλλα λόγια, γίνεται ρητά η υπόθεση ότι αν ένας πρωτεύων χρήστης εκπέμπει σε μία ζώνη συχνοτήτων, κανένα δευτερεύον ΚΤ δε μπορεί να εκπέμπει σε

απλό πρωτόκολλο, το δίκτυο μπορεί να υποστηρίξει τη μη-συνεργατική, δυναμική πρόσβαση φάσματος. Κάθε ΚΤ απαντά στις ενημερώσεις που λαμβάνει, μεταπηδώντας στη ζώνη συχνοτήτων που μεγιστοποιεί το όφελος του, σύμφωνα με τη στρατηγική που έχει υιοθετήσει, π.χ. SFBS ή PSBS.

7.7 Προσομοιώσεις

Στο παρόν υποκεφάλαιο, περιγράφονται αποτελέσματα προσομοιώσεων που έγιναν με σκοπό την αξιολόγηση των στρατηγικών που προτάθηκαν στο παρόν κεφάλαιο, δηλαδή της SFBS και της PSBS. Ο Πίνακας 7.2 παρουσιάζει τις παραμέτρους της προσομοίωσης. Τα ΚΤ κι οι ΣΒ είναι διασκορπισμένοι ομοιόμορφα στην περιοχή προσομοίωσης. Η τυχαία γεωγραφική τους κατανομή απεικονίζεται στην Εικόνα 7.4.



Εικόνα 7.4. Γεωγραφική κατανομή των ΚΤ και των ΣΒ. Τα τετράγωνα σημάδια υποδηλώνουν ΣΒ, ενώ τα x υποδηλώνουν ΚΤ.

Στο συγκεκριμένο περιβάλλον προσομοίωσης, όλα τα ΚΤ λαμβάνουν αποφάσεις ταυτόχρονα, σε κάθε γύρο. Συγκεκριμένα, προσαρμόζουν την ισχύ εκπομπής τους σύμφωνα με την εξίσωση (7.9), κάθε T_p . επίσης, μετά από μία ενημέρωση, σε σχέση με την κατάσταση των διαθέσιμων ζωνών συχνοτήτων, κάθε T_c , όλα τα ΚΤ αποφασίζουν, αυτόνομα, εάν θα μεταπηδήσουν σε μία νέα ζώνη συχνοτήτων, σύμφωνα με την ακολουθούμενη στρατηγική, δηλαδή SFBS ή PSBS.

αυτή, αλλιώς η δυναμική φασματική πρόσβαση απλά δε θα μπορούσε να λειτουργήσει. Σε κάθε περίπτωση, τέτοια συμπεριφορά μπορεί να ελεγχθεί από το ΣΒ.

Πίνακας 7.2. Λίστα των παραμέτρων της προσομοίωσης

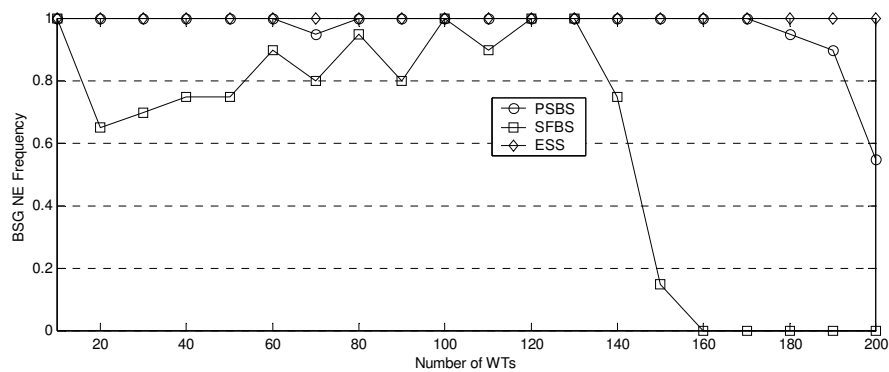
M , πλήθος bit στο πλαίσιο	80	
L , πλήθος ωφέλιμων bit στο πλαίσιο	64	
W , ρυθμός chip	10^6 Hz	
R , ρυθμός εκπομπής δεδομένων	10^4 bps	
σ^2 , ισχύς AWGN power στο δέκτη	5×10^{-15} Watt	
Τεχνική διαμόρφωσης	non-coherent FSK	
P_{max} , μέγιστη ισχύς εκπομπής	2 W	
N , πλήθος χρηστών	10:10:200	
$\gamma_{m,j}, \gamma_{M,j}$, περιορισμοί SIR των ΚΤ $j \in \{1, \dots, N\}$	Ομοιόμορφα κατανομημένοι στο διάστημα [1,20]	
C , πλήθος CDMA ζωνών συχνοτήτων	5	
B , πλήθος ΣΒ	10	
Εμβαδόν περιοχής προσομοίωσης	$2,000 \times 2,000$ m ²	
T_p , περίοδος ελέγχου ισχύος	1 ms	
T_c , περίοδος επιλογής ζώνης συχνοτήτων	200 ms	
Συνολικό πλήθος προγνωστών (SFBS)	15	
Μείγμα Διαθέσιμων Προγνωστών στην Προσομοίωση	Τύπος 1 (6 προγνώστες)	$i = \{0, 1, 2, 3, 4, 5\}$
	Τύπος 2 (1 προγνώστης)	-
	Τύπος 3 (4 προγνώστες)	$const = \{0.1, 0.3, 0.7, 0.9\}$
	Τύπος 4 (2 προγνώστες)	$y = \{4, 10\}$
	Τύπος 5 (1 προγνώστης)	$y = 10$
	Τύπος 6 (1 προγνώστης)	-
Πλήθος προγνωστών ανά χρήση (SFBS), k	5	
Βάθος ιστορίας (SFBS), H	10	
Χρόνος προσομοίωσης	36,000 ms	

Στις προσομοιώσεις, τα ΚΤ ξεκινούν εκπέμποντας στην ίδια ζώνη συχνοτήτων (π.χ. ζώνη συχνοτήτων 0). Άλλα ενδιαφέροντα σενάρια περιλαμβάνουν την άφιξη νέων ΚΤ, μετά τη σύγκλιση του συστήματος στο ΝΕ του BSG, ή την κατάληψη (ή εγκατάλειψη) μίας ζώνης συχνοτήτων από έναν πρωτεύοντα χρήστη. Όλες αυτές οι περιπτώσεις είναι, όμως, ισοδύναμες με την έννοια ότι το σύστημα μετατοπίζεται σε μία κατάσταση που δεν είναι ΝΕ και μετά προσπαθεί να οδηγηθεί σε ένα νέο ΝΕ. Ως εκ τούτου, οι προσομοιώσεις έδειξαν παρόμοιες τάσεις με το σενάριο που μελετάται εδώ και γι' αυτό δεν παρουσιάζονται τα αντίστοιχα αποτελέσματα. Επιπλέον, το σενάριο στο οποίο όλα τα ΚΤ ξεκινούν να εκπέμπουν στην ίδια ζώνη συχνοτήτων μπορεί να θεωρηθεί κι ως το πιο απαισιόδοξο σενάριο και το δυσκολότερο στο χειρισμό του.

7.7.1 Αποτελέσματα Προσομοιώσεων

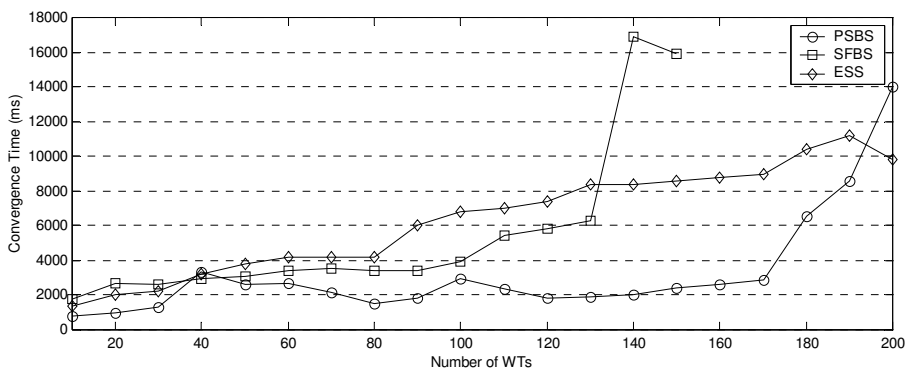
Ένα βασικό χαρακτηριστικό μίας κατανεμημένης στρατηγικής είναι η ικανότητα της να οδηγεί το σύστημα σε σταθερές καταστάσεις (ισορροπίες). Στο BSG, τέτοιες καταστάσεις είναι τα NE του BSG. Πραγματοποιήθηκε μία σειρά από πειράματα για να παρατηρηθεί η συμπεριφορά των προτεινόμενων στρατηγικών για διαφορετικά πλήθη χρηστών. Συγκεκριμένα, μετρήθηκε 1) η συχνότητα με την οποία οι προτεινόμενες στρατηγικές συνέκλιναν σε κάποιο NE του BSG, μέσα στο χρονικό παράθυρο της προσομοίωσης, 2) η ταχύτητα της σύγκλισης, δηλαδή ο χρόνος για να οδηγηθεί το σύστημα στο NE του BSG, 3) το πλήθος των μεταπηδήσεων από μία ζώνη συχνότητων σε μία άλλη, μέχρι να επιτευχθεί η σύγκλιση, κι 4) η ενεργειακή αποτελεσματικότητα του NE του BSG που επιτυγχάνεται (η αντίστοιχη μετρική ορίζεται παρακάτω).

Στην Εικόνα 7.5 απεικονίζεται η πιθανότητα (συχνότητα) με την οποία οι διαφορετικές στρατηγικές επιτυγχάνουν σύγκλιση σε ένα NE του BSG, εντός του χρόνου της προσομοίωσης. Με τη χρήση της στρατηγικής PSBS, ένα NE επιτυγχάνεται σχεδόν πάντα, σε αντίθεση με τη στρατηγική SFBS. Με την SFBS το ΚΤ μπορεί να «παγιδευτεί» σε ένα σύνολο από καταστάσεις, που εξαρτάται από την πορεία των ενεργειών των άλλων ΚΤ και των προγνωστών τους. Συγκεκριμένα, το σύστημα μπορεί να παγιδευτεί σε ένα σύνολο από (μη-NE) καταστάσεις, οι οποίες επισκέπτονται κυκλικά (οδηγώντας έτσι σε ταλαντώσεις). Αυτό παρατηρείται πιο συχνά όσο αυξάνεται το πλήθος των ΚΤ (δες Εικόνα 7.5), γιατί είναι πιο πιθανό τα ΚΤ να ακολουθούν έναν παρόμοιο «τρόπο σκέψης» (ίδιους προγνώστες). Στην Εικόνα 7.5, παρατηρείται επίσης ότι η στρατηγική ESS πάντοτε επιτυγχάνει να συγκλίνει σε ένα NE.



Εικόνα 7.5. Συχνότητα εμφάνισης NE

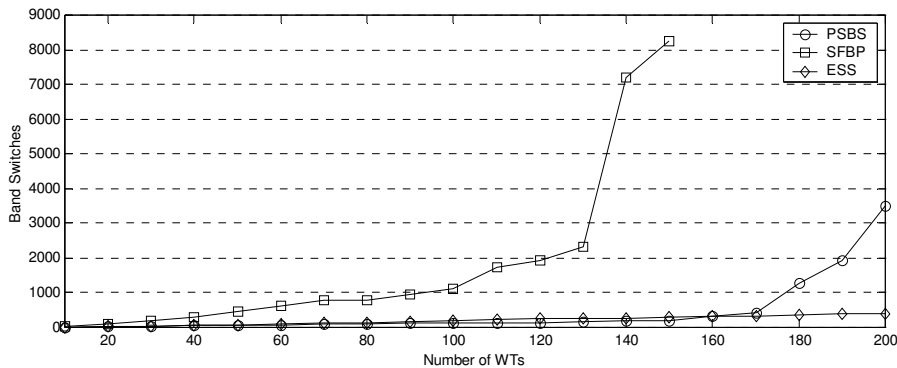
Ένα άλλο χαρακτηριστικό των στρατηγικών για το BSG αποτελεί η ταχύτητα σύγκλισης τους. Όσο πιο γρήγορα συγκλίνει το BSG στο NE (μία σταθερή κι αποδοτική κατάσταση – δεξ παράγραφο 7.5.2), τόσο το καλύτερο είναι για το σύστημα. Στην Εικόνα 7.6, απεικονίζεται ο χρόνος σύγκλισης στο NE για τις εξεταζόμενες στρατηγικές (πρέπει να σημειωθεί ότι λήφθηκαν υπόψη μόνο τα πειράματα όπου υπήρξε σύγκλιση σε ένα NE). Παρατηρείται ότι η στρατηγική PSBS χαρακτηρίζεται από ταχύτερη σύγκλιση σε σχέση με τη SFBS και την ESS. Αξίζει να σημειωθεί κι η ομαλή συμπεριφορά της στρατηγικής ESS σε συνάρτηση με το πλήθος των ΚΤ.



Εικόνα 7.6. Ταχύτητα σύγκλισης

Η δυναμική πρόσβαση στις διαθέσιμες ζώνες συχνοτήτων, μέχρι να επιτευχθεί ένα NE, απαιτεί ένα πλήθος από μεταπηδήσεις από μία ζώνη συχνοτήτων σε μία άλλη από τα εμπλεκόμενα ΚΤ. Είναι λογικό να γίνεται η υπόθεση ότι η μεταπήδηση από μία ζώνη συχνοτήτων σε μία άλλη είναι μία σχετικά δαπανηρή λειτουργία, καθώς εμπλέκει το υλικό (hardware), αλλά και γιατί, μετά από μία μεταπήδηση σε μία νέα

ζώνη συχνοτήτων, το ΚΤ θα πρέπει να προσαρμοστεί στις νέες συνθήκες του καναλιού (μέσω της επαναληπτικής διαδικασίας ελέγχου ισχύος), μία κατάσταση που συνήθως δεν είναι ιδιαίτερα αποδοτική σε σχέση με την κατανάλωση ενέργειας. Στην Εικόνα 7.7, απεικονίζονται τα αποτελέσματα σε ό,τι αφορά το πλήθος των μεταπηδήσεων από μία ζώνη συχνοτήτων σε μία άλλη που έγιναν από το σύνολο των ΚΤ καθ' όλη τη διάρκεια των προσομοιώσεων. Παρατηρείται ότι η ESS, και κυρίως η PSBS, συγκλίνουν σε ένα ΝΕ με πολύ λιγότερες μεταπηδήσεις από μία ζώνη συχνοτήτων σε μία άλλη, σε σχέση με την SFBS.



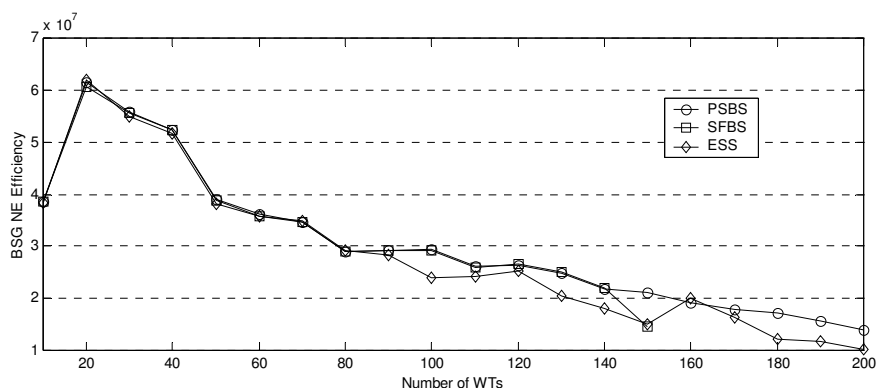
Εικόνα 7.7. Πλήθος μεταπηδήσεων ζώνης συχνοτήτων

Ένα ακόμα σημαντικό κριτήριο αξιολόγησης αποτελεί η αποδοτικότητα των ΝΕ που προκύπτουν. Ως μετρική αποδοτικότητας υιοθετείται ένα μέτρο της «κοινωνικής ευημερίας» (social welfare). Συγκεκριμένα, μετράται ο συνολικός όγκος των δεδομένων που μεταφέρθηκε επιτυχώς από το δίκτυο ως προς την ενέργεια που καταναλώθηκε από όλα τα ΚΤ (που είναι ισοδύναμο με το πλήθος των bits που ελήφθησαν επιτυχώς ανά μονάδα ενέργειας που καταναλώθηκε):

$$Eff = \frac{R \sum_{j=1}^N (1 - e^{-0.5\gamma_j})^M}{\sum_{j=1}^N p_j} \quad (\text{bits/Joule}) \quad (7.23)$$

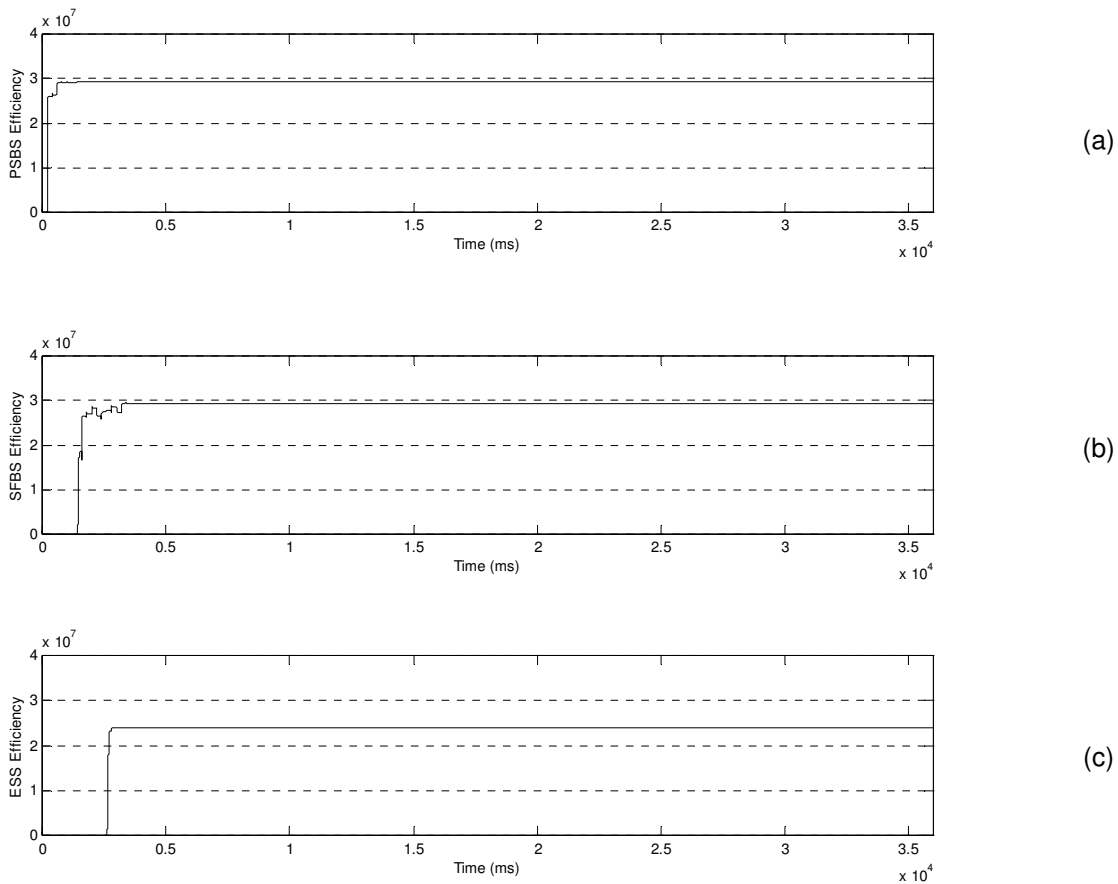
Τα σχετικά αποτελέσματα απεικονίζονται στην Εικόνα 7.8. Το ΝΕ που επιτυγχάνεται από τις τρεις εξεταζόμενες προσεγγίσεις επιδεικνύουν συγκρίσιμη αποδοτικότητα. Η αποδοτικότητα του ΝΕ που επιτυγχάνεται μέσω της στρατηγικής ESS είναι, όμως, είναι κάπως χαμηλότερη σε σύγκριση με τις στρατηγικές SFBS και PSBS. Αυτό αποδίδεται στο γεγονός ότι η ESS αποτελεί μία μάλλον

συντηρητική στρατηγική (μόνο ένα ΚΤ μπορεί να μεταπηδήσει σε μία νέα ζώνη συχνοτήτων σε κάθε βήμα). Έτσι, πιο αποδοτικά ΝΕ που θα μπορούσαν να προκύψουν αν αρκετά ΚΤ μεταπηδούσαν ταυτόχρονα σε διαφορετικές ζώνες συχνοτήτων δεν επιτυγχάνονται. Πρέπει να σημειωθεί ότι, όσο αυξάνει το πλήθος των ΚΤ, η μετρική της αποδοτικότητας μειώνεται. Αυτό είναι αναμενόμενο, καθώς το αυξανόμενο πλήθος των αμοιβαία παρεμβαλλόμενων ΚΤ, οδηγεί σε μία αύξηση στην ισχύ εκπομπής, ώστε να μπορέσουν να επιτευχθούν τα επιδιωκόμενοι SIR.



Εικόνα 7.8. Αποτελεσματικότητα του ΝΕ

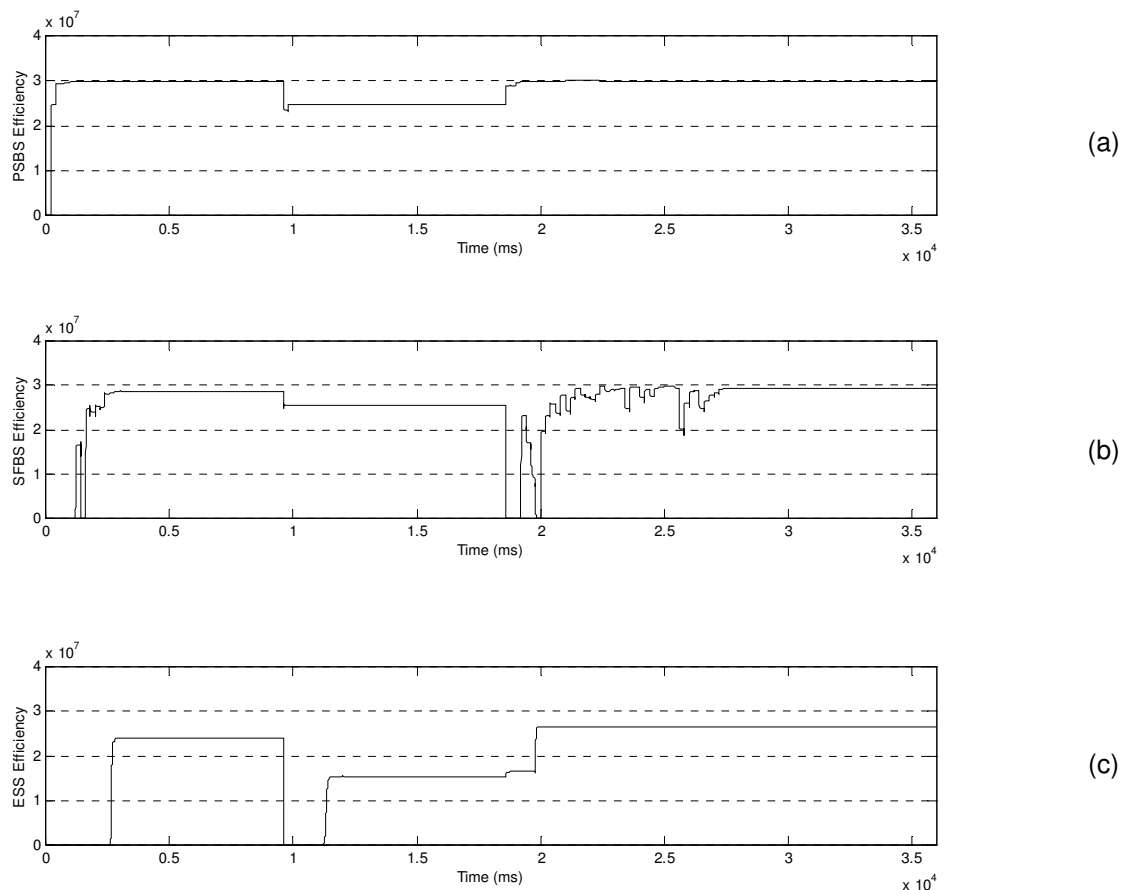
Παρουσιάζονται, επίσης, ενδεικτικά αποτελέσματα σε σχέση με την συμπεριφορά των αξιολογούμενων στρατηγικών στην πορεία του χρόνου. Συγκεκριμένα, μελετάται η εξέλιξη της μετρικής αποδοτικότητας (7.23) 1) κατά την εκκίνηση του συστήματος και 2) στο σενάριο όπου μία από τις διαθέσιμες ζώνες συχνοτήτων ξαφνικά καταλαμβάνεται από ένα πρωτεύον ΚΤ και μετά από λίγο αφήνεται και πάλι ελεύθερη. Στην Εικόνα 7.9 απεικονίζεται η εξέλιξη της μετρικής αποδοτικότητας, από την εκκίνηση του συστήματος (όλα τα ΚΤ ξεκινούν εκπέμποντας στη ζώνη συχνοτήτων 0), για πληθυσμό από 100 ΚΤ. Παρατηρείται η αργή σύγκλιση της στρατηγικής ESS κι ότι η αποδοτικότητα μετά τη σύγκλιση είναι κατώτερη, σε σχέση με τις στρατηγικές PSBS και SFBS (αποτελέσματα που είναι ήδη παρατηρημένα στην Εικόνα 7.6 κι Εικόνα 7.8, αντίστοιχα). Επιπλέον, παρατηρείται ότι η στρατηγική PSBS επιδεικνύει την γρηγορότερη σύγκλιση μεταξύ των τριών εξεταζόμενων στρατηγικών, καθώς κι ομαλότερη σύγκλιση σε σχέση με τη στρατηγική SFBS.



Εικόνα 7.9. Εξέλιξη της μετρικής αποτελεσματικότητας. Όλα τα ΚΤ είναι στην ίδια ζώνη συχνοτήτων κατά την εκκίνηση του συστήματος. (a) PSBS, (b) SFBS (c) ESS.

Η Εικόνα 7.10 απεικονίζει τη συμπεριφορά των εξεταζόμενων στρατηγικών για ένα σενάριο, στο οποίο ένα πρωτεύον ΚΤ καταλαμβάνει μία συγκεκριμένη ζώνη συχνοτήτων (τη ζώνη συχνοτήτων 0 στο παρόν παράδειγμα) για μία συγκεκριμένη χρονική περίοδο και μετά την ξαναφήνει ελεύθερη. Συγκεκριμένα, γίνεται η υπόθεση ότι υπάρχουν 100 δευτερεύοντα ΚΤ και στο χρόνο 9.5 sec ένα πρωτεύον ΚΤ καταλαμβάνει τη ζώνη συχνοτήτων 0 (δηλαδή τέσσερις ζώνες συχνοτήτων παραμένουν διαθέσιμες για χρήση από τα ΚΤ), μέχρι το χρόνο 18.5 sec, όπου όλες οι ζώνες συχνοτήτων διατίθενται εκ νέου για χρήση. Παρατηρείται ότι η στρατηγική PSBS επανέρχεται γρήγορα μετά κι από τα δύο συμβάντα (δηλαδή κατάληψη και παραχώρηση της ζώνης συχνοτήτων). Παρατηρείται, επίσης, ότι η στρατηγική SFBS δεν είναι το ίδιο γρήγορη με τη στρατηγική PSBS. Επιπλέον, η συμπεριφορά της μπορεί να χαρακτηριστεί ως μάλλον ασταθής, σε σύγκριση με την ομαλή συμπεριφορά των PSBS και ESS. Παρατηρείται, ακόμη, η μάλλον αργή απόκριση

της στρατηγικής ESS στις μεταβαλλόμενες συνθήκες κατοχής των ζωνών συχνότητας.



Εικόνα 7.10. Εξέλιξη της μετρικής αποτελεσματικότητας. Μία ζώνη συχνότητας καταλαμβάνεται από έναν πρωτεύοντα χρήστη στο χρόνο 9.5 sec κι απελευθερώνεται στο χρόνο 18.5 sec (0 sec = system startup). (a) PSBS, (b) SFBS (c) ESS.

7.2 Συζήτηση

Με βάση τα αποτελέσματα που παρουσιάστηκαν πιο πάνω, μπορεί κανείς να καταλήξει σε μερικά γενικά συμπεράσματα σε σχέση με την απόδοση των μελετούμενων κατανεμημένων στρατηγικών. Παρατηρείται ότι η ESS δεν συμπεριφέρεται πολύ αποδοτικά στη φάση που μεσολαβεί της εκκίνησης του συστήματος και της σύγκλισης, καθώς κι όταν μία ζώνη συχνότητας καταλαμβάνεται ξαφνικά από κάποιο πρωτεύον ΚΤ. Αυτό οφείλεται στο ότι μόνο ένα ΚΤ κάθε φορά επιτρέπεται να επιλέξει μία νέα ζώνη συχνότητας. Έτσι, ξεκινώντας από ένα ιδιαίτερα φορτωμένο σημείο, τα πληττόμενα ΚΤ δεν έχουν τη δυνατότητα να αντιδράσουν άμεσα. Αυτό οδηγεί σε ένα μεγάλο πλήθος βημάτων κι

ένα σημαντικό απαιτούμενο ποσό ενέργειας μέχρι να επιτευχθεί η σύγκλιση. Οι στρατηγικές SFBS και PSBS συγκλίνουν σημαντικά πιο γρήγορα. Από την άλλη, η ESS επιδεικνύει ομαλή συμπεριφορά σε σχέση με την εξέλιξη του συστήματος μέχρι τη σύγκλιση σε ένα NE (δες Εικόνα 7.9 κι Εικόνα 7.10) και πάντα επιτυγχάνει σύγκλιση σε ένα NE στο χρόνο της προσομοίωσης.

Η στρατηγική SFBS χαρακτηρίζεται από ταχύτερη σύγκλιση από την ESS (αλλά όχι όσο γρήγορη όσο η PSBS), αλλά δεν επιτυγχάνει πάντα να συγκλίνει σε ένα NE. Ένα ακόμα μειονέκτημα της στρατηγικής SFBS είναι το μεγάλο πλήθος των μεταπηδήσεων που πραγματοποιούνται από μία ζώνη συχνοτήτων σε μία άλλη μέχρι να επιτευχθεί σύγκλιση στο NE. Ο Πίνακας 7.3 συνοψίζει τα συμπεράσματα που εξάχθηκαν.

Πίνακας 7.3. Συγκριτική αξιολόγηση των στρατηγικών SFBS, PSBS και ESS

Στρατηγική/Χαρακτηριστικό	SFBS	PSBS	ESS
Συχνή Σύγκλιση σε NE του BSG	+	++	+++
Ταχύτητα Σύγκλισης	++	+++	+
Πλήθος Μεταπηδήσεων Ζωνών Συχνοτήτων	--	++	++
Αποτελεσματικότητα των NE του BSG που επιτυγχάνονται	+++	+++	++
Ταχεία Ανάκαμψη από Συμβάντα Κατάληψης/Απελευθέρωσης Ζώνης Συχνοτήτων	++	+++	+
Ομαλή Συμπεριφορά	-	++	+++
Απλότητα	+	+++	+++
<i>+++:</i> Εξαιρετική απόδοση, <i>--:</i> Πενιχρή απόδοση			

7.8 Συμπεράσματα και Μελλοντική Εργασία

Στην παρούσα εργασία, μελετήθηκε η μη-συνεργατική, δυναμική φασματική πρόσβαση σε δίκτυα CDMA. Συγκεκριμένα, θεωρήθηκε ένα περιβάλλον με διάφορες CDMA ζώνες συχνοτήτων, οι οποίες μπορούσαν να προσπελαστούν οπορτουμιστικά από τα ΚΤ. Η κύρια συνεισφορά είναι η μελέτη του προβλήματος της μη-συνεργατικής επιλογής ζώνης συχνοτήτων, για το οποίο προτάθηκαν δύο διαφορετικές στρατηγικές: (1) μία στρατηγική βασισμένη σε προβλέψεις, με βάση το SFBS και (2) μία στοχαστική στρατηγική (PSBS). Αξιολογήθηκε η απόδοση και των δύο στρατηγικών και τα αποτελέσματα έδειξαν ότι κατέχουν ελκυστικά χαρακτηριστικά, όπως ταχεία κι αποτελεσματική σύγκλιση σε ευσταθή σημεία

λειτουργίας. Αυτό σε συνδυασμό με την απλότητα τους τις καθιστά αξιόλογες υποψήφιους για υιοθεσία σε πραγματικά δίκτυα με ευέλικτη χρήση φάσματος.

Μελλοντικά, θα μελετηθεί η εισαγωγή άλλων οικονομικών και παιγνιοθεωρητικών μοντέλων, π.χ. μοντέλα βασισμένα σε δημοπρασίες ή μαθησιακοί αλγόριθμοι που εξετάζονται στην τεχνητή νοημοσύνη. Θα ήταν επιθυμητό, επίσης, να αξιολογηθούν περαιτέρω οι προτεινόμενες στρατηγικές, αποτιμώντας την επίδραση της κινητικότητας στις επιδόσεις τους και στη σταθερότητα του συστήματος.

ΚΕΦΑΛΑΙΟ 8

ΕΛΕΓΧΟΣ ΙΣΧΥΟΣ ΣΕ CDMA ΔΙΚΤΥΑ

8.1 Εισαγωγή

Στο παρόν κεφάλαιο, παρουσιάζεται η μελέτη που έγινε, στα πλαίσια της διατριβής, σε σχέση με την αποτελεσματική πρόσβαση σε ασύρματα CDMA δίκτυα. Το CDMA επιλέχθηκε, καθώς έχει υιοθετηθεί ως η ασύρματη διεπαφή για τα δίκτυα 3^{ης} γενιάς (3G) [81] κι η υιοθεσία του στα δίκτυα 4G είναι, επίσης, αναμενόμενη [82], [83]. Σε αντίθεση με το TDMA και το FDMA, οι χρήστες μπορούν να εκπέμπουν ταυτόχρονα στην ίδια συχνότητα, παρεμβάλλοντας ο ένας στον άλλο.

Στα δίκτυα CDMA, είναι γνωστό ότι μειώνοντας τις παρεμβολές ελέγχοντας την εκπεμπόμενη ισχύ του κάθε χρήστη αυξάνεται η χωρητικότητα του συστήματος κι επιμηκύνεται η χρόνος ζωής της μπαταρίας των τερματικών [84], [97], [98], [99]. Μία δημοφιλής προσέγγιση στο πρόβλημα του ελέγχου ισχύος των χρηστών σε δίκτυα CDMA βασίζεται σε παιγνιοθεωρητικά μοντέλα [100], [101], [102], [85]. Σε τέτοια μοντέλα, κάθε χρήστης επιζητά τη μεγιστοποίηση της ωφέλειας του με εγωιστικό τρόπο και το παίγνιο δυνητικά καταλήγει σε μία Nash ισορροπία [72]. Ένα τέτοιο σημείο λειτουργίας δεν είναι, όμως, ιδιαίτερα αποδοτικό από «κοινωνικής» άποψης, καθώς υπάρχουν, συνήθως, άλλα σημεία στα οποία όλοι οι χρήστες μπορούν να επιτύχουν μεγαλύτερη ωφέλεια [86].

Η παρούσα μελέτη έχει ως στόχο τον καθορισμό δίκαιων κι αποτελεσματικών σημείων λειτουργίας στο πρόβλημα ελέγχου ισχύος. Γίνεται η υπόθεση ότι οι υπηρεσίες δεδομένων των χρηστών είναι ελαστικές σε ό,τι αφορά το εύρος ζώνης που απαιτούν, αλλά μη-ανεκτικές σε καθυστερήσεις, δηλαδή υπηρεσίες με ελάχιστη και μέγιστη απαίτηση σε ρυθμοαπόδοση, όπου κάποια διακύμανση στο ρυθμό μετάδοσης δεν αποτελεί πρόβλημα, αλλά μόνον σύντομες καθυστερήσεις επιτρέπονται (π.χ. φωνή, video, μεταφορά αρχείων σε πραγματικό χρόνο). Γι' αυτό το λόγο, όπως και στο Κεφάλαιο 7, ο κάθε χρήστης θέτει περιορισμούς στη ρυθμοαπόδοση που επιθυμεί. Με βάση το μη-συνεργατικό παίγνιο ελέγχου ισχύος, μοντελοποιείται το αντίστοιχο συνεργατικό με χρήση της λύσης διαπραγματεύσεως

του Nash [103], η οποία εξ' ορισμού είναι κατά Pareto αποτελεσματική και δίκαιη. Η λύση διαπραγμάτευσης του Nash για το παίγνιο ελέγχου ισχύος οδηγεί σε ένα μη-γραμμικό πρόβλημα βελτιστοποίησης, το οποίο, όπως αποδεικνύεται, έχει μοναδική λύση. Για τον καθορισμό της λύσης προτείνεται ένας κατάλληλος αριθμητικός αλγόριθμος.

8.2 Σχετική Βιβλιογραφία και Συνεισφορά

Σε αυτό το υποκεφάλαιο, παρατίθεται μία σύντομη επισκόπηση της βιβλιογραφίας σε ό,τι αφορά το βέλτιστο και δίκαιο έλεγχο ισχύος σε CDMA δίκτυα. Λόγω των αμοιβαίων παρεμβολών των χρηστών, στις CDMA επικοινωνίες, το πρόβλημα του ελέγχου ισχύος μπορεί να θεωρηθεί ως ένα πρόβλημα κοινωνικής αλληλεπίδρασης, και, έτσι, να μοντελοποιηθεί με τη βοήθεια της Θεωρίας Παιγνίων. Ένα παιγνιοθεωρητικό μοντέλο ελέγχου ισχύος που είναι ευρέως γνωστό έχει προταθεί στην εργασία [86]. Σύμφωνα με το εν λόγω μοντέλο, οι χρήστες επιλέγουν την ισχύ εκπομπής τους, κατανομημένα, βασίζοντας την επιλογή τους σε μία συνάρτηση ωφελείας που μετράει το πλήθος των bit που αποκωδικοποιούνται ορθά στο ΣΒ, ανά μονάδα καταναλισκόμενης ενέργειας. Αποδείχθηκε ότι το συγκεκριμένο παίγνιο έχει ένα μοναδικό NE, το οποίο, όμως, δεν είναι ιδιαίτερα αποτελεσματικό, καθώς υπάρχουν άλλες κατανομές ισχύος που αποφέρουν υψηλότερα επίπεδα ωφέλειας για όλους τους χρήστες.

Στην εργασία [86], παρατηρήθηκε ότι, αν όλοι οι χρήστες μειώσουν την ισχύ τους σε σχέση με το NE, κατά ένα συγκεκριμένο παράγοντα, είναι δυνατόν να επιτευχθεί μία αύξηση στην ωφέλεια όλων τους. Στην εργασία [114], παρατηρήθηκε ότι, δεδομένου ενός επιδιωκόμενου SIR για κάθε χρήστη, αν οι χρήστες ανεχθούν μία ελαφρά μείωση στο SIR τους, μπορεί να επιτευχθεί μία σημαντική μείωση ισχύος. Εκμεταλλευόμενοι τη συγκεκριμένη παρατήρηση, οι συγγραφείς της εργασίας [114] πρότειναν ένα παίγνιο ελέγχου ισχύος, στο οποίο κάθε χρήστης επιλέγει την ισχύ του, ούτως ώστε να ελαχιστοποιήσει μία συνάρτηση κόστους. Η συνάρτηση αυτή είναι το άθροισμα με βάρη (weighted sum) του κόστους εκπομπής αυτού καθ' αυτού και της απόκλισης από το επιδιωκόμενο SIR (σφάλμα SIR). Τα αντίστοιχα

βάρη (που υποδηλώνουν τη σχετική σημαντικότητα των δύο συστατικών κόστους) καθορίζονται από το χρήστη.

Μια εναλλακτική προσέγγιση για να παρακινηθούν οι χρήστες να μειώσουν την ισχύ τους είναι η χρήση χρέωσης. Για το σκοπό αυτό, σε αρκετές εργασίες, π.χ. [86], [112], [113], και [106], έχει προταθεί η χρέωση των χρηστών αναλογικά με την ισχύ εκπομπής τους. Οι χρήστες, στα συγκεκριμένα μοντέλα, επιλέγουν την ισχύ τους με σκοπό να μεγιστοποιήσουν την «καθαρή ωφέλεια» (ωφέλεια μείον κόστος). Στην εργασία [86], το δίκτυο μπορεί να επηρεάζει το NE του παίγνιου, επιλέγοντας τον παράγοντα χρέωσης (την τιμή ανά μονάδα ισχύος). Τέτοια παίγνια είναι, επίσης, γνωστά ως παίγνια ηγέτη-ακολουθού (ή Stackelberg) [72], με *ηγέτη* το δίκτυο κι *ακολουθούς* τους ασύρματους χρήστες. Συγκεκριμένα, το δίκτυο επιλέγει τον παράγοντα χρέωσης, ούτως ώστε να επιτευχθεί η μέγιστη δυνατή κατά Pareto βελτίωση, σε σύγκριση με το αρχικό NE. Πρέπει να σημειωθεί, όμως, ότι με το συγκεκριμένο παίγνιο δεν επιτυγχάνονται κατά Pareto αποτελεσματικά σημεία λειτουργίας. Ο συνδυασμός ελέγχου ισχύος με χρέωση έχει, επίσης, προταθεί για πολύ-κυψελικά συστήματα CDMA, στην εργασία [115]. Σε αντίθεση με την εργασία [86], οι χρήστες χρεώνονται βάσει της ρυθμοαπόδοσης που επιτυγχάνουν, κι όχι ανάλογα με την εκπεμπόμενη ισχύ τους. Επιπλέον, το δίκτυο εστιάζει στη μεγιστοποίηση του κέρδους του (κι όχι την ευημερία των χρηστών) επιλέγοντας τον παράγοντα χρέωσης, βάσει μίας από τις δύο προτεινόμενες στρατηγικές (καθολική ή minimax). Ως εκ τούτου, τα προκύπτοντα NE μπορεί να μην είναι κατά Pareto αποτελεσματικά.

Ένα σημαντικό ζήτημα στον έλεγχο ισχύος με χρέωση είναι ότι η δικαιοσύνη συνήθως υπονομεύεται. Για παράδειγμα, στην εργασία [86], στο NE με χρέωση, οι χρήστες δεν απολαμβάνουν όλοι το ίδιο SIR (όπως στο NE χωρίς χρέωση). Οι χρήστες που αντιμετωπίζουν υψηλές απώλειες σήματος χρεώνονται περισσότερο και λαμβάνουν λιγότερους πόρους, σε σχέση με χρήστες που είναι κοντά στο ΣΒ. Αυτό ονομάζεται και “near-far unfairness”, στην εργασία [112]. Στην εργασία [115], οι συγγραφείς συγκρίνουν τις προτεινόμενες στρατηγικές χρέωσης (καθολική προς minimax) σε ό,τι αφορά τη δικαιοσύνη, δείχνοντας ότι με τη στρατηγική minimax, το κέρδος του δικτύου ελαττώνεται, αλλά οι πόροι διαμοιράζονται πιο δίκαια στους χρήστες. Η απαίτηση για βέλτιστες και δίκαιες κατανομές πόρων σε CDMA δίκτυα

υπήρξε το θέμα στην εργασία [116]. Σε ό,τι αφορά την αποτελεσματικότητα, οι συγγραφείς θεωρούν την κατά Pareto αποτελεσματικότητα, ενώ, σε ό,τι αφορά τη δικαιοσύνη, θεώρησαν και την *max-min*, αλλά και την *proportional* δικαιοσύνη. Σε αντίθεση με τις υπόλοιπες εργασίες που αναφέρονται πιο πάνω, όμως, το μοντέλο που υιοθετήθηκε δε βασίζεται στην ωφέλεια, αλλά εστιάζει στο πρόβλημα τις εκχώρησης ρυθμών εκπομπής, το οποίο διατυπώνεται ως ένα πρόβλημα βελτιστοποίησης που λύνεται κεντρικά.

Ο κεντροποιημένος έλεγχος ισχύος μπορεί μερικές φορές να αποδειχθεί ιδιαίτερα επωφελής για τους χρήστες, καθώς μπορούν να επιλεγούν πιο αποτελεσματικά σημεία λειτουργίας από το ΝΕ. Στην εργασία [107], ο ΣΒ υπολογίζει μία «βέλτιστη» κατανομή ισχύος, την οποία, στη συνέχεια, μεταδίδει στους χρήστες. Στην εργασία [104], προτείνεται η έννοια του «υποβοηθούμενου από το δίκτυο ελέγχου ισχύος» (*network assisted power control*), η οποία, ουσιαστικά, είναι η ίδια προσέγγιση με αυτήν που προτείνεται στην εργασία [107]. Και στις δύο εργασίες ([104] και [107]), υπάρχει η απαίτηση ότι, στη βέλτιστη κατανομή ισχύος, όλοι οι χρήστες φτάνουν στο ΣΒ με ίση ισχύ, δηλαδή έχουν ίσο SIR κι ωφέλεια, αν παραβλεφθούν οι απώλειες σήματος. Το βέλτιστο σημείο λειτουργίας είναι, τότε, αυτό το οποίο μεγιστοποιεί τη συνάρτηση ωφελείας των χρηστών (ίδια για όλους τους χρήστες). Όμως, οι εργασίες [104] και [107] δεν παρέχουν κάποια αυστηρή ανάλυση σχετικά με την αποδοτικότητα των αποτελεσμάτων που εξέθεσαν. Συγκεκριμένα, οι συγγραφείς τις εργασίας [107] εικάζουν ότι η προκύπτουσα κατανομή ισχύος είναι κατά Pareto αποτελεσματική, ενώ, στην εργασία [104], το ζήτημα αυτό δε θίγεται καθόλου. Η κατά Pareto αποτελεσματικότητα της εν λόγω λύσης αποδείχθηκε στην εργασία [108].

Στην παρούσα μελέτη, προτείνεται ένα σχήμα ελέγχου ισχύος με διαιτησία, στο οποίο κάθε χρήστης μπορεί να επιλέγει πάνω και κάτω όρια στη ρυθμοαπόδοση που επιθυμεί (με παρόμοιο τρόπο όπως στην εργασία [116], όπου οι χρήστες θέτουν πάνω και κάτω όρια στο ρυθμό εκπομπής τους), εκφρασμένα μέσω αντίστοιχων ορίων SIR. Ο ρόλος του διαιτητή αναλαμβάνεται από το ΣΒ, ο οποίος υπολογίζει τη βέλτιστη κατανομή ισχύος και τη στέλνει στους χρήστες. Το προτεινόμενο σχήμα βασίζεται στο NBS. Το NBS, εξ' ορισμού, παρέχει κατά Pareto αποτελεσματικές λύσεις, οι οποίες, επίσης, είναι συμβατές με μία έννοια

δικαιοσύνης παρόμοιας χροιάς με την proportional δικαιοσύνη. Σε αντίθεση με την εργασία [116], όμως, στην παρούσα μελέτη, η κατά Pareto αποτελεσματικότητα θεωρείται στο πεδίο των ωφελειών, κι όχι στους ρυθμούς μετάδοσης που εκχωρούνται. Επίσης, σε αντίθεση με την πλειοψηφία των εργασιών που συζητήθηκαν στο παρόν υποκεφάλαιο, (π.χ. [86], [115]), η λύση που προτείνεται δε χρησιμοποιεί χρέωση για να αναγκάσει τους χρήστες να μειώσουν την ισχύ εκπομπής τους κατανεμημένα, αλλά είναι ικανή να επιβάλλει τη βέλτιστη κατανομή ισχύος κεντρικά από το ΣΒ.

8.3 Μοντέλο Συστήματος

Σε αυτό το υποκεφάλαιο, περιγράφεται το μοντέλο που έχει υιοθετηθεί για τη μελέτη του προβλήματος του ελέγχου ισχύος σε δίκτυα CDMA, το οποίο αποτελεί μία απλοποίηση του αντίστοιχου μοντέλου στο Κεφάλαιο 7. Έστω $I = \{1, \dots, M\}$ το σύνολο των χρηστών που μοιράζονται το εύρος ζώνης της ανερχόμενης ζεύξης μίας CDMA κυψέλης. Ο χρήστης i ελέγχει την ισχύ εκπομπής του p_i , επιλέγοντας την από το σύνολο $S_i = [0, +\infty)$. Έστω $\mathbf{p} = (p_1, \dots, p_M)^T$ ένα διάνυσμα ισχύος στο χώρο των στρατηγικών που ορίζεται από το καρτεσιανό γινόμενο $S = S_1 \times \dots \times S_M$. Γίνεται, επίσης, η υπόθεση ότι ο χρήστης i έχει συγκεκριμένες απαιτήσεις σε ποιότητα υπηρεσίας (QoS) με κάτω κι άνω όρια στη ρυθμοαπόδοση που επιθυμεί να επιτύχει, εκφρασμένες με αντίστοιχα κάτω κι άνω όρια στο SIR που επιτυγχάνει, $\gamma_{m,i}$ και $\gamma_{M,i}$, αντίστοιχα ($\gamma_{m,i} < \gamma_{M,i}$) –δες Κεφάλαιο 7. Έστω $\Gamma = \{ \mathbf{p} \in S : \gamma_{m,i} \leq \gamma_i(\mathbf{p}) \leq \gamma_{M,i}, \forall i \in I \}$ το σύνολο των κατανομών ισχύος που οδηγούν σε εφικτά SIR, όπου το γ_i υποδηλώνει το SIR του χρήστη i κι ορίζεται ως εξής:

$$\gamma_i = \gamma_i(\mathbf{p}) = \frac{W}{R} \frac{h_i p_i}{\sum_{k \neq i} h_k p_k + \sigma^2} \quad (8.1)$$

Το W είναι ο ρυθμός chip και το R ο ρυθμός μετάδοσης δεδομένων (b/s), ενώ ο λόγος W/R υποδηλώνει την απολαβή του CDMA κώδικα. Το h_i εκφράζει τις απώλειες σήματος από το ΚΤ μέχρι το ΣΒ και το σ^2 την ισχύ του προσθετικού λευκού θορύβου Γκάους στο δέκτη (ΣΒ). Πρέπει να σημειωθεί ότι το σύνολο Γ που ορίζεται από τους περιορισμούς SIR είναι ένα κυρτό πολύεδρο στο χώρο των

κατανομών ισχύος, καθώς οι SIR περιορισμοί καταλήγουν σε γραμμικές ανισότητες (δες εξίσωση (8.1)).

Γίνεται η υπόθεση ότι ο χρήστης i έχει μία συνάρτηση ωφελείας $u_i: \Gamma \rightarrow \mathfrak{R}$ που εκφράζει το επίπεδο της ικανοποίησης του από τη χρήση των πόρων του συστήματος. Η εν λόγω συνάρτηση συνυπολογίζει το QoS (ή ρυθμοαπόδοση) που επιτυγχάνει ο χρήστης ως αποτέλεσμα της εκπεμπόμενης ισχύος του και των παρεμβολών από τους άλλους χρήστες, καθώς και το αντίστοιχο ενεργειακό κόστος. Υιοθετήθηκε η συνάρτηση ωφελείας που έχει προταθεί στην εργασία [86], η οποία εκφράζει το πλήθος των bit που αποκωδικοποιούνται με επιτυχία ανά μονάδα καταναλισκόμενης ενέργειας. Συγκεκριμένα, η συνάρτηση ωφελείας u_i του χρήστη i είναι

$$u_i(\mathbf{p}) = \frac{LR}{M\rho_i} (1 - e^{-0.5\gamma_i})^M \text{ (bit/Joule)}, \quad (8.2)$$

όπου το L είναι το μήκος σε bit του πλαισίου του χρήστη και το M το μήκος του πλαισίου συμπεριλαμβανομένης της επικεφαλίδας ($M > L$). Ο όρος $(1 - e^{-0.5\gamma_i})^M$ αποτελεί μία προσέγγιση της πιθανότητας ορθής λήψης ενός πλαισίου για ασύγχρονη FSK διαμόρφωση, υποθέτοντας προσθετικό Γκαουσιανό κανάλι (additive Gaussian channel) κι έλλειψη κωδικοποίησης καναλιού¹⁷. Περισσότερες πληροφορίες σχετικά με τη συγκεκριμένη συνάρτηση ωφελείας μπορούν να βρεθούν στην εργασία [86]. Στο επόμενο υποκεφάλαιο, δεδομένων των φορμαλισμών που εισήχθησαν στο παρόν υποκεφάλαιο, μελετάται το πρόβλημα του ελέγχου ισχύος ως ένα μη-συνεργατικό παίγνιο.

8.4 Μη-Συνεργατικός Έλεγχος Ισχύος με Περιορισμούς SIR

Σε αυτό το υποκεφάλαιο, μελετάται το μη-συνεργατικό παίγνιο ελέγχου ισχύος. Το βασικό θέμα σε αυτού του είδους τα παίγνια είναι η ύπαρξη κι η μοναδικότητα ενός ΝΕ. Ένας ορισμός του ΝΕ, σε σχέση με το μελετούμενο πρόβλημα, δίδεται πιο κάτω.

¹⁷ Αυτό το απλοϊκό μοντέλο καναλιού θεωρείται επαρκές για τη μοντελοποίηση του παιγνίου ελέγχου ισχύος [86].

Ορισμός 8.1: Ένα διάνυσμα ισχύων $\mathbf{p} = (p_1, \dots, p_N)^T$ είναι ένα NE του παιγνίου ελέγχου ισχύος αν για κάθε χρήστη $i \in I$,

$$u_i(p_i, \mathbf{p}_{-i}) \geq u_i(p'_i, \mathbf{p}_{-i}),$$

$$\text{για όλα τα } p'_i \in S_i, \gamma_{m,i} \leq \gamma_i(p'_i, \mathbf{p}_{-i}) \leq \gamma_{M,i},$$

όπου $\mathbf{p}_{-i} = (p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_N)^T$.

Με άλλα λόγια, στο NE, δεδομένων των επιπέδων ισχύος των άλλων χρηστών, κανείς χρήστης δε μπορεί να βελτιώσει το επίπεδο της ωφελείας του κάνοντας ατομικές αλλαγές στην ισχύ εκπομπής του, δηλαδή αποτελεί ένα ευσταθές σημείο λειτουργίας.

Το επόμενο θεώρημα υποστηρίζει ότι υπάρχει ένα και μόνο NE στο μελετούμενο παίγνιο ελέγχου ισχύος.

Θεώρημα 8.1: Υπάρχει ένα και μοναδικό NE στο μη-συνεργατικό παίγνιο ελέγχου ισχύος.

Απόδειξη: Η απόδειξη του θεωρήματος προκύπτει από το *Θεώρημα 3.2* του Debreu [109], καθώς η συνάρτηση ωφελείας $u_i(\cdot)$ είναι ημι-κοίλη (quasiconcave) ως προς το p_i [86] κι ορίζεται στο κυρτό σύνολο Γ . ■

Όπως έχει αναλυθεί και στο Κεφάλαιο 7 (υποκεφάλαιο 7.4), στο NE του μελετούμενου παιγνίου, ο χρήστης i θα έχει SIR γ_i που θα είναι η λύση της ακόλουθης εξίσωσης:

$$0.5M\gamma_i + 1 = e^{0.5\gamma_i} \quad (8.3)$$

Η αναζητούμενη τιμή μπορεί να υπολογιστεί εύκολα αριθμητικά. Αξίζει να σημειωθεί ότι το ίδιο SIR γ^* επιλύει την εξίσωση (8.3) για κάθε χρήστη $i \in I$. Λόγω, όμως, των περιορισμών στο SIR ($\gamma_{m,i}$ και $\gamma_{M,i}$), το SIR στο NE για το χρήστη i διαμορφώνεται ως εξής (δες Κεφάλαιο 7):

$$\gamma_i^* = \begin{cases} \gamma_{m,i}, & \gamma^* \leq \gamma_{m,i} < \gamma_{M,i} \\ \gamma^*, & \gamma_{m,i} < \gamma^* \leq \gamma_{M,i} \\ \gamma_{M,i}, & \gamma_{m,i} < \gamma_{M,i} < \gamma^* \end{cases} \quad (8.4)$$

Με άλλα λόγια, ο χρήστης i επιλύοντας πρώτα την εξίσωση (8.3) και χρησιμοποιώντας στη συνέχεια τον κανόνα στην εξίσωση (8.4), μπορεί να επιλέξει αυτόνομα το NE SIR του γ_i^* . Η αντίστοιχη ισχύς NE ρ_i^* μπορεί να υπολογιστεί, επίσης, ανεξάρτητα μέσω της ακόλουθης επαναληπτικής, κατανεμημένης διαδικασίας (δες Κεφάλαιο 7):

$$\rho_i(t+1) = \frac{R\gamma_i^*}{Wh_i} [r\rho(t) - h_i\rho_i(t)], \quad (8.5)$$

όπου το

$$r\rho(t) = \sum_{k=1}^N h_k \rho_k(t) + \sigma^2,$$

υποδηλώνει την συνολική λαμβανόμενη ισχύ (παρεμβολή) τη χρονική στιγμή t , η οποία εκπέμπεται, ανά τακτά χρονικά διαστήματα, από το ΣΒ. Για λόγους σαφήνειας, παραλείπεται ο έμφυτος περιορισμός $\rho_i \leq \rho_{max}$, όπου το ρ_{max} υποδηλώνει τη μέγιστη ισχύ των πομπών. Αν αυτός ο περιορισμός παραβιάζεται, μετά από μία επανάληψη, υπονοείται ότι η ισχύς εκπομπής περιορίζεται στο ρ_{max} .

Εκτός από το SIR στο NE που μπορεί να υπολογιστεί από τις εξισώσεις (8.3) και (8.4), μέσω της εξίσωσης (8.1), είναι δυνατόν να υπολογιστεί κι η αντίστοιχη ισχύς κάθε χρήστη. Συγκεκριμένα, η ισχύς με την οποία φτάνει ο χρήστης i στο ΣΒ $q_i = h_i\rho_i$ είναι η ακόλουθη (δες Κεφάλαιο 7):

$$q_i = \frac{\sigma^2}{\left(\frac{W}{R\gamma_i} + 1\right) \left(1 - \sum_{k=1}^N \left(\frac{W}{R\gamma_k} + 1\right)^{-1}\right)}, \quad \forall i \in I \quad (8.6)$$

Με άλλα λόγια, η επαναληπτική διαδικασία που περιγράφεται στην εξίσωση (8.5), μετά από ένα πεπερασμένο πλήθος βημάτων, συγκλίνει στην κατανομή ισχύων NE που δίδεται στην εξίσωση (8.6) (για $\gamma_i = \gamma_i^*$). Πρέπει να παρατηρηθεί, όμως, ότι για να προκύψει εφικτή κατανομή ισχύος, πρέπει να ικανοποιείται ο ακόλουθος περιορισμός:

$$\sum_{k=1}^N \left(\frac{W}{R\gamma_k} + 1 \right)^{-1} < 1 \quad (8.7)$$

Σε κάθε άλλη περίπτωση, η ισχύς που προκύπτει (όπως υπολογίζεται μέσω της εξίσωσης (8.6)) γίνεται άπειρη ή αρνητική κι η επαναληπτική διαδικασία που περιγράφεται από την εξίσωση (8.5) δε συγκλίνει.

Ο περιορισμός στην εξίσωση (8.7) εκφράζει τον περιορισμό στη χωρητικότητα ενός συστήματος CDMA (που αναφέρεται κι ως *χωρητικότητα πόλου*), ο οποίος μπορεί εύκολα να υιοθετηθεί ως βάση για έναν απλό μηχανισμό CAC (δες Κεφάλαιο 7).

8.5 Αξιολόγηση Απόδοσης του Μη-Συνεργατικού Σχήματος Ελέγχου Ισχύος

Στο προηγούμενο υποκεφάλαιο, μελετήθηκε το μη-συνεργατικό παίγνιο ελέγχου ισχύος με ύπαρξη περιορισμών SIR κι υπολογίστηκε το NE. Σε αυτό το υποκεφάλαιο, μελετάται η αποτελεσματικότητά του. Μία κοινά αποδεκτή έννοια αποτελεσματικότητας αποτελεί η κατά Pareto αποτελεσματικότητα (*Ορισμός 3.2*).

Στο μη-συνεργατικό παίγνιο ελέγχου ισχύος που μελετάται στην εργασία [86] (εκδοχή του μελετούμενου παιγνίου χωρίς περιορισμούς), αποδείχθηκε ότι, αν όλοι οι χρήστες, αντί να εκπέμπουν με την ισχύ που αντιστοιχεί στο NE, περιορίζαν ελαφρά την ισχύ εκπομπής τους ταυτόχρονα, το σημείο λειτουργίας που θα προέκυπτε θα είναι βελτιωμένο, για όλους τους χρήστες, σε ό,τι αφορά την ωφέλεια, σε σύγκριση με το NE, δηλαδή είναι δυνατόν να επιτευχθεί ένα κατά Pareto κυρίαρχο σημείο λειτουργίας.

Στην παρούσα μελέτη, βάσει της ίδιας ιδέας, δηλαδή της μείωσης της ισχύος των χρηστών σε σχέση με το NE, αποδεικνύεται ότι είναι δυνατόν να βρεθούν κατά Pareto κυρίαρχες κατανομές ισχύος σε σύγκριση με το NE. Λόγω, όμως, των περιορισμών SIR στο μοντέλο, μειώνοντας την ισχύ όλων των χρηστών κατά τον ίδιο παράγοντα (η τεχνική που προτείνεται στην εργασία [86]) μπορεί να οδηγήσει σε μη εφικτές κατανομές ισχύος. Για το λόγο αυτό, ακολουθείται μία πιο περίπλοκη προσέγγιση, σύμφωνα με την οποία οι χρήστες διαχωρίζονται σε δύο ξεχωριστές ομάδες: (1) όσους μειώνουν την ισχύ τους σε σχέση με το NE κατά έναν παράγοντα $\mu \in (0,1]$ κι (2) όσους διατηρούν το SIR του NE τους προσαρμόζοντας

την ισχύ τους, ανάλογα με τη μείωση ισχύος των άλλων χρηστών. Οι χρήστες από την πρώτη ομάδα, δηλαδή, δέχονται να «θυσιάσουν» το NE SIR τους, μειώνοντας την ισχύ τους, ενώ οι χρήστες της δεύτερης ομάδας πάντα διατηρούν το NE SIR τους.

Έστω ότι για το χρήστη i ισχύει ότι $\gamma_{m,i} > \gamma^*$ (όπου γ^* είναι η λύση της εξίσωσης (8.3)), δηλαδή $\gamma_i^* = \gamma_{m,i}$ από την εξίσωση (8.4). Αν ένας ή περισσότεροι χρήστες, εκτός από το χρήστη i , πραγματοποιήσουν μία μείωση στην ισχύ τους, ο χρήστης i θα μπορέσει να διατηρήσει το SIR που είχε στο NE γ_i^* με μικρότερη ισχύ εκπομπής (λόγω της μείωσης των παρεμβολών), γεγονός που σαφώς τον ωφελεί. Η μείωση ισχύος του χρήστη i , εν συνεχεία, ωφελεί τους υπόλοιπους χρήστες, καθώς η παρεμβολή που αποδίδεται σε αυτόν μειώνεται. Επίσης, καθώς ο χρήστης i δε μπορεί να δεχθεί καμία μείωση στο SIR του (το SIR του στο NE είναι το χαμηλότερο αποδεκτό, δηλαδή $\gamma_i^* = \gamma_{m,i}$), μπορεί να προκύψει ο κανόνας ότι ο χρήστης i για τον οποίο $\gamma_{m,i} > \gamma^*$ πρέπει πάντα να διατηρεί το SIR που είχε στο NE. Έστω K το σύνολο των δεικτών των χρηστών που διατηρούν το SIR που είχαν στο NE.

Όπως έχει ήδη αναφερθεί, το όφελος που επιτυγχάνεται από το χρήστη $i \in K$ είναι προφανές – το SIR που είχε στο NE διατηρείται, ενώ η ισχύς για να το επιτύχει μειώνεται. Μελετάται τώρα η επίδραση της μείωσης της ισχύος για τους χρήστες που δεν ανήκουν στο σύνολο K (δηλαδή τους χρήστες που μπορούν να ανεχθούν μία μείωση στο SIR τους σε σχέση με το NE). Αν θέσουμε $q_i(\mu) = \mu q_i^*$, για κάθε χρήστη $i \in I - K$, από την εξίσωση (8.1) και την εξίσωση (8.6), προκύπτει η ακόλουθη μορφή του SIR $\gamma_i(\mu)$ του χρήστη i :

$$\gamma_i(\mu) = \frac{W}{R} \frac{\mu q_i^*}{(A+1)\mu \left(\sum_{k \in I-K-\{i\}} q_k^* \right) + A\mu q_i^* + (A+1)\sigma^2}, \quad \forall i \in I-K, \quad (8.8)$$

όπου

$$A = \sum_{k \in K} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1} \left/ \left[1 - \sum_{k \in K} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1} \right] \right.$$

Για να επιτευχθεί μία αύξηση στην ωφέλεια, μειώνοντας την ισχύ q_i^* κατά την παράγοντα μ , η παράγωγος της συνάρτησης ωφελείας $u_i(q^*, \mu)$ του χρήστη i , $\forall i \in I$

– K , ως προς μ , θα πρέπει να είναι αρνητική για $\mu=1$. Πιο κάτω δίδεται η παράγωγος της συνάρτησης ωφελείας $u_i(\mathbf{q}^*, \mu)$ του χρήστη i , $\forall i \in I - K$, ως προς μ :

$$\frac{du_i(\mathbf{q}^*, \mu)}{d\mu} = h_i \frac{LR}{Mq_i^*} \frac{(1 - e^{-0.5\gamma_i})^{M-1} e^{-0.5\gamma_i}}{\mu^2} \left(0.5M\mu \frac{d\gamma_i(\mu)}{d\mu} - (e^{0.5\gamma_i} - 1) \right), \quad (8.9)$$

όπου

$$\frac{d\gamma_i(\mu)}{d\mu} = \frac{1}{\mu} \gamma_i(\mu) \frac{(A+1)\sigma^2}{(A+1)\mu \sum_{k \in I-K-\{i\}} q_k^* + A\mu q_i^* + (A+1)\sigma^2}.$$

Έτσι, από την εξίσωση (8.9), για $\mu = 1$, αρκεί να αποδειχθεί ότι η συνάρτηση $C(\gamma_i^*)$, που ορίζεται παρακάτω, είναι αρνητική, ούτως ώστε να επιτευχθεί μία αύξηση στη συνάρτηση ωφελείας του χρήστη i .

$$C(\gamma_i^*) = 0.5M\gamma_i^* \frac{(A+1)\sigma^2}{(A+1) \sum_{k \in I-K-\{i\}} q_k^* + Aq_i^* + (A+1)\sigma^2} - (e^{0.5\gamma_i^*} - 1) \quad (8.10)$$

Μπορεί εύκολα να αποδειχθεί ότι, αν $\gamma_i^* = \gamma^*$, τότε $C(\gamma_i^*) < 0$. Αυτό δεν ισχύει, όμως, πάντα, αν $\gamma_{M,i} < \gamma^*$ (δηλαδή $\gamma_i^* = \gamma_{M,i}$), γιατί αν το $\gamma_{M,i}$ είναι αρκετά μικρό, ο χρήστης i μπορεί να μην επωφεληθεί μειώνοντας την ισχύ εκπομπής του. Για να διευκρινιστεί αυτό, δίδεται κι η επόμενη ισοδύναμη μορφή της συνάρτησης $C(\gamma_i^*)$:

$$C(\gamma_i^*) = 0.5M \frac{1 - \sum_{k \in I-\{i\}} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1}}{1 - \sum_{k \in K} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1}} \gamma_i^* - 0.5M \frac{\sum_{k \in I-\{i\}} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1}}{1 - \sum_{k \in K} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1}} (\gamma_i^*)^2 - (e^{0.5\gamma_i^*} - 1) \quad (8.11)$$

Πρέπει να σημειωθεί ότι αν η συνθήκη

$$M \left(1 - \sum_{k \in I-\{i\}} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1} \right) > 1 - \sum_{k \in K} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1}$$

αληθεύει, μπορεί εύκολα να αποδειχθεί ότι υπάρχει ένα κατώφλι SIR $\gamma_0 \in (0, \gamma^*)$, για το οποίο $C(\gamma_i^*) < 0$, $\forall \gamma_i^* \in (\gamma_0, \gamma^*)$. Η φυσική σημασία του γ_0 είναι το ελάχιστο SIR που πρέπει να έχει ένας χρήστης στο NE, ούτως ώστε να επωφεληθεί από μία μείωση στην ισχύ του. Αυτό αποδεικνύει ότι ένας χρήστης με χαμηλό SIR στο NE μπορεί να μην ωφελείται πάντοτε μειώνοντας την ισχύ του. Μπορεί επίσης να

αποδειχθεί ότι το κατώφλι SIR γ_0 αυξάνεται καθώς περισσότεροι χρήστες συμπεριλαμβάνονται στο σύνολο K .

Βάσει των παραπάνω παρατηρήσεων, περιγράφεται μία μέθοδος για τον εντοπισμό κατανομών ισχύος που είναι κατά Pareto κυρίαρχες, σε σχέση με την κατανομή ισχύος στο NE. Συγκεκριμένα, μπορούν να διακριθούν οι παρακάτω περιπτώσεις:

- I. Αν υπάρχει έστω κι ένας χρήστης με SIR στο NE ίσο με γ^* (δηλαδή τη λύση της εξίσωσης (8.3)), μπορεί απλά όλοι οι υπόλοιποι χρήστες να συμπεριληφθούν στο σύνολο K . Μειώνοντας την ισχύ του εξεταζόμενου χρήστη κατά έναν κατάλληλο παράγοντα, όλοι οι χρήστες (συμπεριλαμβανομένου και του ίδιου) θα επωφεληθούν.
- II. Στην περίπτωση που δεν υπάρχει χρήστης με SIR στο NE ίσο με γ^* , είτε $\gamma_i^* = \gamma_{m,i}$, είτε $\gamma_i^* = \gamma_{M,i}, \forall i \in I$. Ορίζεται το σύνολο K ως $K = \{j \in I : \gamma_j^* = \gamma_{m,j}\}$. Αν $K = I$, τότε κανείς χρήστης δε μπορεί να επωφεληθεί πραγματοποιώντας μία μείωση στην ισχύ εκπομπής του, δηλαδή, δε μπορεί να βρεθεί καμία κατά Pareto κυρίαρχη κατανομή ισχύος.
- III. Έστω ότι το σύνολο K ορίζεται όπως και στην περίπτωση II, αλλά $K \subset I$. Αν $C(\gamma_i^*) < 0, \forall i \in I - K$, τότε μία μείωση ισχύος από τους χρήστες στο σύνολο $I - K$ θα ωφελούσε το σύνολο των χρηστών. Διαφορετικά, επιλέγεται ο χρήστης με το μικρότερο SIR στο NE¹⁸ από το σύνολο $I - K$, συμπεριλαμβάνεται στο σύνολο K κι ελέγχεται εκ νέου εάν $C(\gamma_i^*) < 0, \forall i \in I - K$. Αυτή η επαναληπτική διαδικασία ολοκληρώνεται, αν $C(\gamma_i^*) < 0, \forall i \in I - K$, ή $K = I$. Στην πρώτη περίπτωση, μία κατά Pareto βελτίωση είναι πιθανή, ενώ στη δεύτερη δεν είναι.

Έστω ότι έχει καθοριστεί το σύνολο των χρηστών που μπορούν να μειώσουν την ισχύ τους σε σχέση με το NE (δηλαδή το σύνολο $I-K$) επιτυγχάνοντας αύξηση της ωφέλειας για το σύνολο των χρηστών. Πρέπει να σημειωθεί ότι, όταν αυτοί οι χρήστες μειώσουν την ισχύ τους, το SIR τους θα μειωθεί (καθώς $d\gamma_i(\mu)/d\mu > 0$) κι έτσι οι περιορισμοί SIR που έχουν θέσει (σε ό,τι αφορά το ελάχιστο επίπεδο SIR)

¹⁸ Επιλέγεται ο χρήστης με το χαμηλότερο SIR στο NE, ώστε το γ_0 να αυξηθεί όσο το δυνατόν λιγότερο, ούτως ώστε να καταστεί ευκολότερο για τους υπόλοιπους χρήστες να ξεπεράσουν το νέο γ_0 .

μπορούν να παραβιαστούν. Από την εξίσωση (8.8), υπάρχει ένας ελάχιστος παράγοντας μείωσης $\mu_{th} \in (0, 1]$, τέτοιος που η κατανομή SIR που προκύπτει είναι εφικτή, για $\mu \in [\mu_{th}, 1]$. Ο ελάχιστος παράγοντας μ_{th} υπολογίζεται ως εξής:

$$\mu_{th} = \max_{i \in I-K} \left\{ (A+1)\sigma^2 / \left[\left(\frac{W}{R\gamma_{m,i}} - A \right) q_i^* - (A+1) \sum_{k \in I-K-\{i\}} q_k^* \right] \right\} \quad (8.12)$$

Έτσι, μπορεί να βρεθεί ένας παράγοντας μείωσης της ισχύος εκπομπής $\mu \in [\mu_{th}, 1]$ με τον οποίο μπορεί να επιτευχθεί αύξηση της ωφέλειας για όλους τους χρήστες, χωρίς να παραβιαστούν οι περιορισμοί SIR.

Από τα παραπάνω, παρατηρείται ότι η κατανομή ισχύος στο NE, στο παίγνιο ελέγχου ισχύος με περιορισμούς SIR, δεν είναι, κατά κανόνα, κατά Pareto αποτελεσματική. Έτσι, μπορεί να εξαχθεί το συμπέρασμα ότι απαιτείται ένας πιο αποδοτικός μηχανισμός από το μη-συνεργατικό παίγνιο ελέγχου ισχύος. Στο επόμενο υποκεφάλαιο, μελετάται ένα σχήμα διαιτησίας ως μέσον επίτευξης του προαναφερθέντος στόχου.

8.6 Έλεγχος Ισχύος με Διαιτησία και Περιορισμούς SIR

Σε αυτό το υποκεφάλαιο, μελετάται ένα σχήμα κεντρικού ελέγχου ισχύος με σκοπό την επίτευξη πιο αποδοτικών σημείων λειτουργίας, σε σύγκριση με το μη-συνεργατικό παίγνιο ελέγχου ισχύος. Αντί να αφήνονται οι χρήστες να λειτουργούν εγωιστικά, ο ΣΒ επιλύει το πρόβλημα ελέγχου ισχύος με «βέλτιστο» τρόπο και στη συνέχεια μεταδίδει τη λύση στους χρήστες.

Στην παρούσα μελέτη, ο στόχος είναι να παραχθούν κατά Pareto αποτελεσματικές λύσεις (Ορισμός 3.2). Για ένα παίγνιο, όμως, με N παίκτες, τα κατά Pareto αποτελεσματικά σημεία σχηματίζουν μία υπερεπιφάνεια (hypersurface) $N-1$ διαστάσεων, δηλαδή υπάρχει άπειρο πλήθος τέτοιων σημείων. Έτσι, το ερώτημα που προκύπτει είναι ποιο από όλα τα κατά Pareto αποτελεσματικά σημεία θα πρέπει να επιλεγεί ως λύση σε ένα τέτοιο παίγνιο. Μία συνήθης προσέγγιση είναι η εισαγωγή επιπλέον κριτηρίων, όπως απαίτηση για δικαιοσύνη, όπως και με τη λύση διαπραγμάτευσης του Nash [103] (δες υποκεφάλαιο 3.5). Στη συνέχεια, ορίζεται το NBS για το παίγνιο ελέγχου ισχύος και, μετά, μελετώνται οι ιδιότητες

του προβλήματος βελτιστοποίησης που αντιστοιχεί στο NBS. Τελικά, παρέχεται ένας αριθμητικός αλγόριθμος για τον υπολογισμό της επιδιωκόμενης λύσης.

8.6.1 Η Λύση Διαπραγμάτευσης του Nash για το Παίγνιο Ελέγχου Ισχύος

Σε αυτήν την παράγραφο, μελετούνται οι ιδιότητες του προβλήματος βελτιστοποίησης που οδηγεί στο NBS του παιγνίου ελέγχου ισχύος. Συγκεκριμένα, αποδεικνύεται ότι το μελετούμενο πρόβλημα έχει μοναδική λύση, γεγονός που είναι θεμελιώδες για την εφαρμογή μίας αριθμητικής μεθόδου βασισμένης στην κλίση (gradient-based numerical method) για τον υπολογισμό της λύσης. Η λύση διαπραγμάτευσης του Nash για το παίγνιο ελέγχου ισχύος ορίζεται ως το σημείο όπου μεγιστοποιείται το γινόμενο των πλεοναζόντων ωφελειών των χρηστών (surplus utility) πέραν του status quo (ωφέλεια που επιτυγχάνεται στο NE), δηλαδή

$$\max_{\mathbf{p}} \prod_{j=1}^N (u_j(\mathbf{p}) - u_j^0), \quad \mathbf{p} \in Q_0, \quad Q_0 = \{\mathbf{r} \in \Gamma : \mathbf{u}(\mathbf{r}) > \mathbf{u}_0\},$$

Με άλλα λόγια, στο NBS, μεγιστοποιείται το γινόμενο των πλεοναζόντων ωφελειών των χρηστών, λαμβάνοντας υπόψη τον περιορισμό ότι το SIR κάθε χρήστη θα πρέπει να βρίσκεται μέσα στα αντίστοιχα όρια κι ότι η ωφέλεια κάθε χρήστη θα πρέπει να είναι ανώτερη από την ωφέλεια του στο status quo.

Έστω ένας γραμμικός μετασχηματισμός της συνάρτησης ωφέλειας $\boldsymbol{\varphi}: \mathfrak{R}^N \rightarrow \mathfrak{R}^N$, όπου $\boldsymbol{\varphi}(\mathbf{u}) = \mathbf{v}$, $v_i = (1/h_i)u_i$, $\forall i \in I$. Είναι εύκολο να παρατηρηθεί ότι η μετασχηματισμένες συναρτήσεις ωφέλειας v_i έχουν την ακόλουθη μορφή:

$$v_i(\mathbf{q}) = \frac{LR}{Mq_i} (1 - e^{-0.5\gamma_i})^M, \quad (8.13)$$

όπου $q_i = h_i p_i$. Με αυτόν τον μετασχηματισμό, ο οποίος δεν επηρεάζει τη λύση, λαμβάνεται υπόψη μόνον η λαμβανόμενη ισχύς στο ΣΒ. Προς αυτήν την κατεύθυνση, ο ΣΒ δε χρειάζεται να είναι ενήμερος για τις απώλειες του σήματος κάθε χρήστη. Επιλύει ως προς τη λαμβανόμενη κατανομή ισχύος και μετά μεταδίδει τη λύση στους χρήστες, οι οποίοι στη συνέχεια, προσαρμόζουν την ισχύ εκπομπής τους, βάσει των υπολογιζόμενων απωλειών σήματος. Συγκεκριμένα, ο ΣΒ καλείται να επιλύσει το ακόλουθο πρόβλημα βελτιστοποίησης:

$$(P) \quad \max_{\mathbf{q}} f(\mathbf{q}), \quad f(\mathbf{q}) = \prod_{j=1}^N (v_j(\mathbf{q}) - v_j^0), \quad \mathbf{q} \in Q_0$$

Πρέπει να σημειωθεί ότι η συνάρτηση ωφελείας $v_i(\cdot)$ του χρήστη i είναι ημι-κοίλη στο σύνολο Γ , όπως ορίζει και το *Λήμμα 8.1*.

Λήμμα 8.1: Η συνάρτηση ωφελείας $v_i: \Gamma \rightarrow \mathfrak{R}$, για το χρήστη $i \in I$, είναι ημι-κοίλη.

Απόδειξη: Για να αποδειχθεί ότι η συνάρτηση $v_i(\mathbf{q})$ είναι ημι-κοίλη, αρκεί να αποδειχθεί ότι, για κάθε δύο σημεία \mathbf{x} και \mathbf{y} στο Γ , η ακόλουθη συνθήκη είναι αληθής [105]:

$$v_i(\theta\mathbf{x} + (1 - \theta)\mathbf{y}) \geq \min\{v_i(\mathbf{x}), v_i(\mathbf{y})\}, \quad (8.14)$$

όπου $0 \leq \theta \leq 1$.

Πρέπει να σημειωθεί ότι η συνάρτηση ωφελείας $v_i(\cdot)$ είναι αδιάφορη στις ατομικές συνεισφορές ισχύος των χρηστών εκτός του χρήστη i . Για το χρήστη i , μόνο το άθροισμα αυτών των συνεισφορών (δηλαδή συνολική παρεμβολή) έχει ενδιαφέρον. Έστω, στο χώρο Γ , το υπερεπίπεδο $N-1$ διαστάσεων

$$\sum_{k \neq i} q_k = c, \quad c > 0$$

στο οποίο η παρεμβολή που υφίστανται ο χρήστης i από τους άλλους χρήστες είναι ίση με c . Παρατηρείται ότι ο χρήστης i είναι αδιάφορος σε ό,τι αφορά συγκεκριμένα σημεία του εν λόγω υπερεπιπέδου (δηλαδή τη συνεισφορά του κάθε χρήστη ξεχωριστά), που μπορεί να βρίσκεται μία κατανομή ισχύος. Ενδιαφέρεται για το ίδιο το υπερεπίπεδο (δηλαδή τη συνολική παρεμβολή). Αυτή η παρατήρηση επιτρέπει την απλοποίηση της ανάλυσης και τη μελέτη του προβλήματος σε μόνο δύο διαστάσεις (αντί για N), δηλαδή (1) την ισχύ του χρήστη i , q_i , και (2) την παρεμβολή που προκαλούν οι υπόλοιποι χρήστες,

$$I_i = \sum_{k \neq i} q_k.$$

Έστω $\mathbf{x}, \mathbf{y} \in \Gamma$, ή, για το χρήστη i , $(x_i, I_{x,i})$ και $(y_i, I_{y,i})$, σύμφωνα με τα παραπάνω. Έστω (L_i) : $I_i = aq_i + b$ η ευθεία που ορίζεται από τα σημεία \mathbf{x} και \mathbf{y} , όπου a και b είναι σταθερές που προκύπτουν κατάλληλα¹⁹. Σε αυτήν την απόδειξη, γίνεται η υπόθεση ότι $a > 0$ και $b + \sigma^2 > 0$. Οι υπόλοιπες περιπτώσεις μπορούν να

¹⁹ $a = (I_{y,i} - I_{x,i})/(y_i - x_i)$ και $b = (y_i I_{x,i} - x_i I_{y,i})/(y_i - x_i)$

μελετηθούν παρόμοια και δεν αναλύονται εδώ, για λόγους συντομίας. Για να μελετηθεί η συμπεριφορά της συνάρτησης ωφελείας $v_i(\cdot)$, στην ευθεία L_i , αντικαθίσταται η παρεμβολή l_i με την ποσότητα $aq_i + b$. Ως εκ τούτου, η συνάρτηση ωφελείας $v_i(\cdot)$ γίνεται συνάρτηση μίας μεταβλητής, ήτοι της q_i . Πιο κάτω, υπολογίζεται η παράγωγος της $v_i(q_i)$ ως προς το q_i :

$$\frac{dv_i(q_i)}{dq_i} = \frac{LR}{Mq_i^2} (1 - e^{-0.5\gamma_i})^{M-1} e^{-0.5\gamma_i} \left(0.5M \frac{W}{R} (b + \sigma^2) \frac{q_i}{(aq_i + b + \sigma^2)^2} - (e^{0.5\gamma_i} - 1) \right)$$

Είναι εύκολο να δειχθεί ότι, για $q_i = 0$, $dv_i(q_i)/dq_i = 0$. Όμως, αυτό το σημείο δεν αποτελεί τοπικό μέγιστο της $v_i(q_i)$, καθώς $v_i(0) = 0$ και $v_i(\varepsilon) > 0$, όπου ε είναι ένας μικρός θετικός αριθμός. Έτσι, πιθανά κρίσιμα σημεία ενδιαφέροντος (δηλαδή πιθανά μέγιστα) της συνάρτησης ωφελείας $v_i(q_i)$ είναι οι μη μηδενικές ρίζες της συνάρτησης

$$0.5M \frac{W}{R} (b + \sigma^2) \frac{q_i}{(aq_i + b + \sigma^2)^2} - (e^{0.5\gamma_i} - 1) = 0. \quad (8.15)$$

Είναι εύκολο να αποδειχθεί ότι συναρτήσεις αυτής της μορφής έχουν μοναδική ρίζα, \tilde{q}_i . Επιπλέον, είναι εύκολο να δειχθεί ότι, για $q_i < \tilde{q}_i$, $dv_i(q_i)/dq_i > 0$, ενώ για $q_i > \tilde{q}_i$, $dv_i(q_i)/dq_i < 0$. Με άλλα λόγια, καθώς αυξάνει το q_i , η συνάρτηση ωφελείας $v_i(q_i)$, στην ευθεία L_i , είναι αύξουσα, μέχρι το \tilde{q}_i . Παρομοίως, όταν το q_i παίρνει τιμές μεγαλύτερες του \tilde{q}_i , η συνάρτηση $v_i(q_i)$ είναι φθίνουσα. Έτσι, η $v_i(q_i)$ έχει ένα μοναδικό μέγιστο πάνω στην ευθεία L_i και, ως εκ τούτου, η συνθήκη (8.14) ικανοποιείται, για κάθε δύο σημεία $\mathbf{x}, \mathbf{y} \in \Gamma$. ■

Σημαντικό ζήτημα στη μελέτη ενός προβλήματος βελτιστοποίησης αποτελεί κι η δομή του συνόλου περιορισμών, π.χ. αν είναι κυρτό. Στο μελετούμενο πρόβλημα ελέγχου ισχύος, το σύνολο περιορισμών Q_0 αποτελείται από τις κατανομές ισχύος στις οποίες (i) το SIR είναι αποδεκτό κι (ii) οι ωφέλειες είναι ανώτερες από το status quo. Το επόμενο λήμμα δηλώνει ότι το σύνολο περιορισμών Q_0 είναι κυρτό.

Λήμμα 8.2: Το σύνολο των κατανομών ισχύος Q_0 που είναι κατά Pareto κυρίαρχες από την status quo κατανομή ισχύος είναι κυρτό.

Απόδειξη: Πρέπει να σημειωθεί ότι το $Q_0 = \{ \mathbf{r} \in \Gamma : \mathbf{u}(\mathbf{r}) > \mathbf{u}^0 \}$. Ισοδύναμα, $Q_0 = \{ \mathbf{r} \in \Gamma : \mathbf{v}(\mathbf{r}) > \mathbf{v}^0 \} = \{ \mathbf{r} \in \Gamma : v_i(\mathbf{r}) > v_i^0, \forall i \in I \}$.

Ως εκ τούτου,

$$Q_0 = \bigcap_{i=1}^N \{ \mathbf{r} \in \Gamma : v_i(\mathbf{r}) > v_i^0 \}.$$

Από το *Λήμμα 8.1*, το σύνολο $\{ \mathbf{r} \in \Gamma : v_i(\mathbf{r}) > v_i^0 \}$ είναι κυρτό ως ένα “superlevel set” της ημι-κοίλης συνάρτησης ωφελείας $v_i(\cdot)$ [105]. Έτσι, παρατηρείται ότι το σύνολο Q_0 είναι κυρτό ως η τομή κυρτών (superlevel) συνόλων. ■

Ένα, επίσης, σημαντικό θέμα είναι η μορφή της αντικειμενικής συνάρτησης του μελετούμενου προβλήματος βελτιστοποίησης. Το επόμενο θεώρημα δηλώνει ότι η αντικειμενική συνάρτηση $f(\mathbf{q})$ του προβλήματος (P) είναι ημι-κοίλη συνάρτηση κι ότι το πρόβλημα (P) έχει μοναδική λύση.

Θεώρημα 8.2: Η αντικειμενική συνάρτηση $f(\mathbf{q})$ του προβλήματος (P) είναι ημι-κοίλη στο σύνολο Q_0 . Επιπλέον, το πρόβλημα (P) έχει μοναδική λύση.

Απόδειξη: Όπως στο *Λήμμα 8.1*, για να αποδειχθεί ότι η αντικειμενική συνάρτηση $f(\mathbf{q})$ είναι ημι-κοίλη, αρκεί να δειχθεί ότι για κάθε δύο σημεία \mathbf{x} και \mathbf{y} στο Q_0 , ισχύει η επόμενη συνθήκη:

$$f(\theta \mathbf{x} + (1 - \theta) \mathbf{y}) \geq \min\{f(\mathbf{x}), f(\mathbf{y})\} \quad (8.16)$$

όπου $0 \leq \theta \leq 1$.

Από την απόδειξη στο *Λήμμα 8.1*, η συνάρτηση ωφελείας $v_i(\mathbf{q})$ είναι ευαίσθητη μόνο στην ισχύ q_i του χρήστη i και της συνολικής παρεμβολής l_i από τους άλλους χρήστες. Έστω, τότε, $\mathbf{x}, \mathbf{y} \in Q_0$ και (L_i): $l_i = a_i q_i + b_i$, για το χρήστη i , να είναι η ισοδύναμη ευθεία που ορίζεται από τα σημεία \mathbf{x} και \mathbf{y} , και, όπως στην απόδειξη στο *Λήμμα 8.1*, γίνεται η υπόθεση ότι $a_i > 0$ και $b_i + \sigma^2 > 0$. Μελετάται η συμπεριφορά της αντικειμενικής συνάρτησης $f(\mathbf{q})$ στην ευθεία που ορίζεται από τα σημεία \mathbf{x} και \mathbf{y} (L_i για το χρήστη i). Συγκεκριμένα, αντικαθίσταται η παρεμβολή l_i με την ποσότητα $a_i q_i + b_i$, και, ως εκ τούτου, η αντικειμενική συνάρτηση $f(\mathbf{q})$ εκφράζεται ως εξής:

$$f(\mathbf{q}) = \prod_{j=1}^N (v_j(q_j) - v_j^*) \quad (8.17)$$

Υπολογίζεται, τώρα, η παράγωγος της $f(\mathbf{q})$ ως προς το q_i ως εξής:

$$\frac{\partial f(\mathbf{q})}{\partial q_i} = \frac{dv_i(q_i)}{dq_i} \prod_{j \neq i} (v_j(q_j) - v_j^*)$$

Καθώς τα σημεία $\mathbf{x}, \mathbf{y} \in Q_0$, η ωφέλεια του χρήστη i ($\forall i \in I$) στην ευθεία που συνδέει το \mathbf{x} και το \mathbf{y} είναι υψηλότερη, σε σχέση με την ωφέλεια στο status quo, δηλαδή $v_i(\mathbf{q}) > v_i^0$. Επομένως, το πρόσημο της $\partial f(\mathbf{q})/\partial q_i$ εξαρτάται από το πρόσημο της $dv_i(q_i)/dq_i$. Έτσι, παρόμοια με την ανάλυση της απόδειξης στο *Λήμμα 8.1*, υπάρχει μία μοναδική ισχύς \tilde{q}_i που $\partial f(\mathbf{q})/\partial q_i > 0$ (γιατί $dv_i(q_i)/dq_i > 0$), για $q_i < \tilde{q}_i$ και $\partial f(\mathbf{q})/\partial q_i < 0$ (γιατί $dv_i(q_i)/dq_i < 0$), για $q_i > \tilde{q}_i$. Παρατηρείται, έτσι, ότι $\nabla f(\mathbf{q}) > \mathbf{0}$, για $\mathbf{q} < \tilde{\mathbf{q}}$, και $\nabla f(\mathbf{q}) < \mathbf{0}$ για $\mathbf{q} > \tilde{\mathbf{q}}$, όπου $\tilde{\mathbf{q}} = (\tilde{q}_1, \dots, \tilde{q}_N)$. Πρέπει να σημειωθεί ότι, σιωπηρά, έχει γίνει η υπόθεση ότι $\mathbf{x} < \mathbf{y}$. Άλλες περιπτώσεις μπορούν να αναλυθούν με παρόμοιο τρόπο. Έτσι, η $f(\mathbf{q})$ έχει ένα μοναδικό μέγιστο στην ευθεία που ορίζεται από τα σημεία \mathbf{x} και \mathbf{y} και, ως εκ τούτου, ικανοποιεί την ιδιότητα (8.16), για κάθε δύο σημεία \mathbf{x}, \mathbf{y} στο Q_0 , δηλαδή, η συνάρτηση $f(\mathbf{q})$ είναι ημι-κοίλη στο Q_0 . Ως εκ τούτου, η $f(\mathbf{q})$ έχει ένα μοναδικό καθολικό μέγιστο. ■

Συνοψίζοντας, σε αυτό το υποκεφάλαιο, μελετήθηκε το πρόβλημα βελτιστοποίησης (P), η λύση του οποίου είναι το NBS του παιγνίου ελέγχου ισχύος. Συγκεκριμένα, αποδείχθηκε ότι η συνάρτηση ωφελείας $v_i(\cdot)$ του χρήστη i είναι ημι-κοίλη στο σύνολο Γ κι ότι το σύνολο Q_0 (το σύνολο των κατανομών ισχύος, στις οποίες η ωφέλεια είναι αυστηρά ανώτερη από την ωφέλεια στο status quo) είναι κυρτό. Στη συνέχεια, βάσει αυτών των αποτελεσμάτων, ακολούθησε το συμπέρασμα ότι η αντικειμενική συνάρτηση του προβλήματος (P) είναι ημι-κοίλη στο σύνολο Q_0 κι ότι το πρόβλημα (P) έχει μοναδική λύση. Η ίδια η λύση, όμως, δεν έχει καθοριστεί. Αυτό αποτελεί το ζήτημα που μελετάται στην επόμενη παράγραφο.

8.6.2 Προσδιορισμός της Λύσης Διαπραγμάτευσης του Nash

Στην προηγούμενη παράγραφο, αποδείχθηκε ότι υπάρχει μοναδική λύση στο πρόβλημα (P), δηλαδή υπάρχει μοναδική NBS κατανομή ισχύος. Στην εκδοχή, του βασισμένου στο NBS, παιγνίου ελέγχου ισχύος χωρίς περιορισμούς SIR, η λύση είναι μία κατανομή ισχύων όπου όλοι οι χρήστες λαμβάνονται από το ΣΒ με την

ίδια ισχύ, δηλαδή $q_1 = \dots = q_N = \tilde{q}$, και $v_1(\tilde{q}) = \dots = v_N(\tilde{q}) = v(\tilde{q})$, όπου το \tilde{q} επιλέγεται ούτως ώστε να μεγιστοποιεί τη συνάρτηση $v(\tilde{q})$ [93].

Πρέπει, όμως, να σημειωθεί ότι, λόγω της έλλειψης συμμετρίας στην εκδοχή του παιγνίου με περιορισμούς (κάθε χρήστης μπορεί να έχει διαφορετικές απαιτήσεις σε SIR), χρειάζεται να καταφύγει κανείς σε πιο πολύπλοκες τεχνικές βελτιστοποίησης, βασισμένες σε αριθμητικούς αλγορίθμους, με εφαρμογή σε N -διάστατους χώρους κατανομών ισχύος, αντί για τη λύση ενός προβλήματος μίας μόνο μεταβλητής.

Ως πρώτο βήμα για την αριθμητική λύση του μελετούμενου προβλήματος, ορίζεται το πρόβλημα (P'), το οποίο προκύπτει από το πρόβλημα (P), παίρνοντας το λογαριθμικό μετασχηματισμό της αντικειμενικής του συνάρτησης.

$$(P') \quad \max_{\mathbf{q}} g(\mathbf{q}), \quad g(\mathbf{q}) = \sum_{j \in J} \ln(v_j(\mathbf{q}) - v_j^0), \quad \mathbf{q} \in Q_0$$

Πρέπει να σημειωθεί ότι η αντικειμενική συνάρτηση του προβλήματος (P'), $g = \ln(f)$, είναι ημι-κοίλη ως σύνθεση της αύξουσας λογαριθμικής συνάρτησης $\ln : \mathfrak{R}_+ \rightarrow \mathfrak{R}$ και της ημι-κοίλης συνάρτησης $f : \mathfrak{R}^N \rightarrow \mathfrak{R}$, έτσι, έχει μοναδικό μέγιστο. Επιπλέον, το καθολικό μέγιστο $\tilde{\mathbf{q}}$ της αντικειμενικής συνάρτησης $f(\cdot)$ (δηλαδή η λύση του προβλήματος (P)) είναι, επίσης, το καθολικό μέγιστο της αντικειμενικής συνάρτησης $g(\cdot)$ (δηλαδή η λύση του προβλήματος (P')) [34]. Συνεπώς, το πρόβλημα (P) και το πρόβλημα (P') είναι ισοδύναμα²⁰.

Ο επαναληπτικός αλγόριθμος που προτείνεται για τη λύση του προβλήματος (P') βασίζεται στη μέθοδο *υπό συνθήκης κλίσης* (conditional gradient) [111]. Η μέθοδος αυτή μπορεί να εφαρμοστεί όταν η αντικειμενική συνάρτηση (εδώ η $g(\cdot)$) είναι συνεχώς διαφορίσιμη (continuously differentiable) και το σύνολο περιορισμών είναι μη-κενό, κλειστό και κυρτό. Στο μελετούμενο πρόβλημα, το σύνολο των περιορισμών Q_0 είναι μη-κενό. Συγκεκριμένα, (i) το σύνολο Γ δεν είναι ποτέ κενό, όπως εγγυάται από το μηχανισμό CAC και (ii) το σύνολο $Q_0 (\subseteq \Gamma)$ είναι μη-κενό, καθώς είναι συνήθως εφικτό να βρεθεί μία κατά Pareto κυρίαρχη κατανομή ισχύος

²⁰ Ο μετασχηματισμός του (P) στο (P') κάνει την αριθμητική εύρεση της επιδιωκόμενης λύσης πιο εύκολη, καθώς είναι ευκολότερο να χειριστεί κανείς ένα άθροισμα, παρά ένα γινόμενο, συναρτήσεων, π.χ. κατά τον υπολογισμό των παραγώγων τους.

(δες υποκεφάλαιο 8.5). Το σύνολο Q_0 δεν είναι, όμως, κλειστό λόγω των αυστηρών ανισοτήτων ωφελείας που το ορίζουν. Παρόλ' αυτά, αυτό δεν επηρεάζει τη σύγκλιση του προτεινόμενου αλγορίθμου, καθώς η συγκεκριμένη μέθοδος που χρησιμοποιείται εγγυάται ότι το σημείο που προκύπτει, μετά από κάθε επανάληψη, παραμένει πάντοτε μέσα στο σύνολο περιορισμών.

Τα βασικά συστατικά της προτεινόμενης μεθόδου είναι: 1) ο καθορισμός εφικτών κατευθύνσεων αύξησης (feasible ascent directions) κι 2) η επιλογή βήματος, και σκιαγραφούνται στην παρούσα παράγραφο. Περισσότερες τεχνικές λεπτομέρειες, όπως η επιλογή ενός αρχικού εφικτού σημείου, καλύπτονται στο Παράρτημα V.

8.6.2.1 Προσδιορισμός Εφικτών Κατευθύνσεων Αύξησης

Δεδομένου ενός αρχικού εφικτού διανύσματος \mathbf{q} , μία *εφικτή κατεύθυνση* (feasible direction) στο \mathbf{q} είναι ένα διάνυσμα \mathbf{d} τέτοιο που το $\mathbf{q} + a\mathbf{d}$ να είναι εφικτό για κάθε $a > 0$ που είναι επαρκώς μικρό. Μία μέθοδος εφικτής κατεύθυνσης ξεκινά με ένα εφικτό διάνυσμα $\mathbf{q}^{(0)}$ (δες και το Παράρτημα V.1) και παράγει μία ακολουθία από εφικτά διανύσματα $\{\mathbf{q}^{(n)}\}$, ως εξής

$$\mathbf{q}^{(n+1)} = \mathbf{q}^{(n)} + a^{(n)}\mathbf{d}^{(n)},$$

όπου το $\mathbf{d}^{(n)}$ είναι μία εφικτή κατεύθυνση στο $\mathbf{q}^{(n)}$ και το βήμα επιλέγεται, ώστε να είναι θετικό, αλλά και $\mathbf{q}^{(n+1)} \in Q_0$. Το $\mathbf{d}^{(n)}$ είναι, επίσης, μία κατεύθυνση αύξησης αν

$$\nabla g(\mathbf{q}^{(n)})^T \cdot \mathbf{d}^{(n)} > 0.$$

Ένας ευθύς τρόπος να παραχθεί μία εφικτή κατεύθυνση $\mathbf{d}^{(n)} = \bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}$ από το σημείο $\mathbf{q}^{(n)}$ που να ικανοποιεί την συνθήκη αύξησης είναι να επιλυθεί το πρόβλημα βελτιστοποίησης

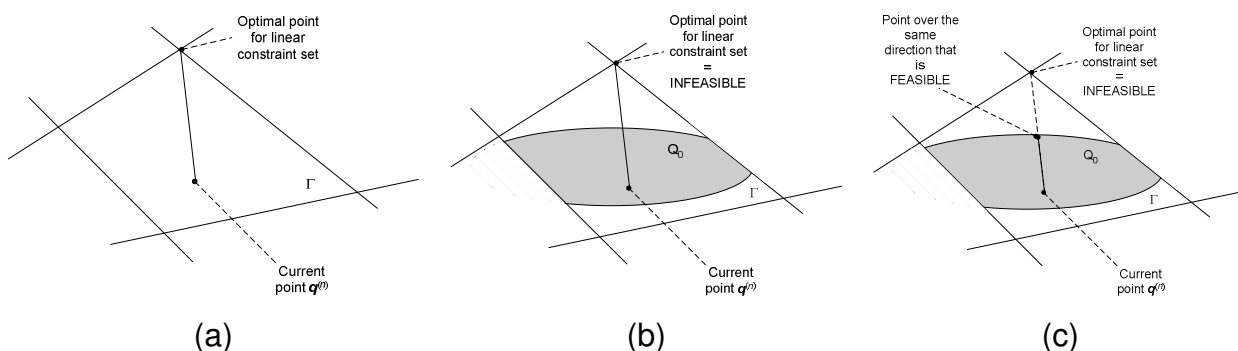
$$(D) \begin{cases} \max_{\mathbf{w}} \nabla g(\mathbf{q}^{(n)})^T \cdot (\mathbf{w} - \mathbf{q}^{(n)}) \\ \text{s.t. } \mathbf{w} \in Q_0 \end{cases}.$$

Η λύση του (D), $\bar{\mathbf{q}}^{(n)}$, είναι το πιο απομακρυσμένο σημείο του συνόλου Q_0 στην κατεύθυνση της παραγωγού (αύξησης) [111].

Αυτή η προσέγγιση συνήθως υιοθετείται όταν η συνάρτηση $g(\cdot)$ είναι μη-γραμμική, ενώ το σύνολο των περιορισμών Q_0 ορίζεται από γραμμικούς περιορισμούς. Σε αυτήν την περίπτωση, το πρόβλημα (D) είναι ένα γραμμικό πρόγραμμα που

μπορεί να επιλυθεί αρκετά εύκολα (π.χ. μέσω της μεθόδου Simplex). Στην περίπτωση, όμως, που εξετάζεται εδώ, το σύνολο των περιορισμών Q_0 περιλαμβάνει και μη-γραμμικούς περιορισμούς (κάθε χρήστης πρέπει να απολαμβάνει αυστηρά μεγαλύτερη ωφέλεια από την ωφέλεια στο status quo). Ως εκ τούτου, το πρόβλημα του εντοπισμού του πιο απομακρυσμένου σημείου του Q_0 , στην κατεύθυνση της παραγώγου, παρουσιάζει δυσκολίες.

Για να αντιμετωπιστεί το πρόβλημα της μη-γραμμικότητας του συνόλου περιορισμών, προτείνεται μία διαδικασία δύο φάσεων για την εύρεση εφικτών κατευθύνσεων. Πρώτα, λαμβάνονται υπόψη μόνο οι γραμμικοί περιορισμοί, δηλαδή οι περιορισμοί SIR που σχηματίζουν το κυρτό πολυέδρο Γ . Συγκεκριμένα, έχοντας φτάσει σε ένα εφικτό διάνυσμα $\mathbf{q}^{(n)}$, εντοπίζεται το σημείο $\bar{\mathbf{q}}^{(n)}$ που μεγιστοποιεί την αντικειμενική συνάρτηση του προβλήματος (D) στο πολυέδρο Γ (δηλαδή το πιο απομακρυσμένο σημείο του Γ στην κατεύθυνση της παραγώγου). Αυτή η κατεύθυνση θα μπορούσε να ήταν εφικτή, αν μόνον το σύνολο Γ αποτελούσε το σύνολο των περιορισμών. Κάθε εφικτό διάνυσμα $\bar{\mathbf{q}}^{(n)}$, όμως, πρέπει να είναι τέτοιο που όλοι οι χρήστες να απολαμβάνουν ωφέλεια που ξεπερνά αυστηρά την ωφέλεια του status quo (δηλαδή $v_i(\bar{\mathbf{q}}^{(n)}) > v_i^0, \forall i \in I$).



Εικόνα 8.1. Επιλογή κατεύθυνσης δύο φάσεων. (α) Βέλτιστη κατεύθυνση ως προς τους γραμμικούς περιορισμούς SIR μόνο, (β) Η βέλτιστη κατεύθυνση ως προς τους γραμμικούς περιορισμούς SIR δεν είναι εφικτή (δεν ανήκει στο σύνολο Q_0), (γ) Η αρχική διεύθυνση περιορίζεται ώστε να γίνει εφικτή

Σημαντικό στοιχείο αποτελεί ότι το σύνολο Q_0 είναι κυρτό (Λήμμα 8.2) κι ότι $Q_0 \subseteq \Gamma$. Παρατηρείται ότι, αν $\mathbf{q}^{(n)} \in Q_0$, αλλά $\bar{\mathbf{q}}^{(n)} \notin Q_0$, μπορεί να βρεθεί ένα σημείο $\bar{\mathbf{q}}^{(n)}$ στο ευθύγραμμο τμήμα που συνδέει τα σημεία $\mathbf{q}^{(n)}$ και $\bar{\mathbf{q}}^{(n)}$ ούτως ώστε $\bar{\mathbf{q}}^{(n)} \in Q_0$. Επιπλέον, κάθε σημείο στο ευθύγραμμο τμήμα που συνδέει τα σημεία $\mathbf{q}^{(n)}$

και $\bar{\mathbf{q}}^{(n)}$ θα είναι εφικτό. Έτσι, η δεύτερη φάση της διαδικασίας εύρεσης κατεύθυνσης είναι να βρεθεί το σημείο $\bar{\mathbf{q}}^{(n)}$, δηλαδή ένα εφικτό σημείο στο ευθύγραμμο τμήμα που συνδέει τα σημεία $\mathbf{q}^{(n)}$ και $\bar{\mathbf{q}}^{(n)}$. Το σημείο $\bar{\mathbf{q}}^{(n)}$ μπορεί να υπολογιστεί εύκολα μέσω μίας διαδικασίας διχοτόμησης, όπως παρουσιάζεται στο Παράρτημα V.2. Η κατεύθυνση $\mathbf{d}^{(n)}$ στο βήμα n είναι τότε

$$\mathbf{d}^{(n)} = \bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}.$$

Η Εικόνα 8.1 συνοψίζει τα δύο βήματα της διαδικασίας εύρεσης κατεύθυνσης, σε δισδιάστατο χώρο, για λόγους απεικόνισης. Με τον προτεινόμενο αλγόριθμο, εντοπίζεται το εφικτό σημείο που είναι το πιο απομακρυσμένο στο σύνολο Q_0 στην κατεύθυνση της παραγώγου, σε συμφωνία με τη λογική της αρχικής μεθόδου, όπως αυτή περιγράφεται στον ορισμό του προβλήματος (D).

8.6.2.2 Επιλογή Βήματος

Εδώ, συνοψίζεται ένας αλγόριθμος για την επιλογή βήματος, ο οποίος βασίζεται στον κανόνα του *Armijo* (Armijo rule) [111]. Γίνεται η υπόθεση ότι η κατεύθυνση $\mathbf{d}^{(n)}$ είναι τέτοια που το $\mathbf{q}^{(n)} + \mathbf{d}^{(n)}$ είναι εφικτό, που ισχύει για τη μέθοδο επιλογής κατεύθυνσης που περιγράφηκε πιο πάνω (δες παράγραφο 8.6.2.1). Έστω β και s σταθερές, με $\beta \in (0,1)$ και $s \in (0,1)$. Σύμφωνα με τον κανόνα του Armijo, το βήμα στην επανάληψη n είναι $a^{(n)} = \beta^{m(n)}$, όπου $m(n)$ είναι ο μικρότερος ακέραιος m όπου

$$g(\mathbf{q}^{(n)} + \beta^m \mathbf{d}^{(n)}) - g(\mathbf{q}^{(n)}) \geq s \beta^m [\nabla g(\mathbf{q}^{(n)})^T \cdot \mathbf{d}^{(n)}]. \quad (8.18)$$

Με άλλα λόγια, τα βήματα β, β^2, \dots , δοκιμάζονται το ένα μετά το άλλο, μέχρι η πιο πάνω ανισότητα να ικανοποιηθεί για $m = m(n)$. Έτσι, το βήμα $a^{(n)}$ επιλέγεται όχι μόνο για να παράγει θετική βελτίωση στην αντικειμενική συνάρτηση, αλλά μία αρκετά σημαντική βελτίωση. Στο Παράρτημα V.3, περιγράφεται κι ο σχετικός αριθμητικός αλγόριθμος.

8.6.2.3 Συζήτηση

Το πλεονέκτημα του αλγορίθμου που περιγράφεται πιο πάνω είναι ότι δεν βασίζεται σε κάποια δομή του συνόλου περιορισμών, εκτός της κυρτότητας του, κι η ακολουθία των εφικτών σημείων προκύπτει απλά με αναζήτηση προς τις

κατευθύνσεις αύξησης. Όπως έχει αποδειχθεί στο [111], η ακολουθία κατευθύνσεων $\{d^{(n)}\}$ που παράγεται από την μέθοδο conditional gradient σχετίζεται με την παράγωγο και συγκλίνει σε ακρότατο – εδώ το ολικό μέγιστο της συνάρτησης $g(\cdot)$. Συγκεκριμένα, έχει αποδειχθεί ότι κάθε οριακό σημείο της μεθόδου κλίσης υπό συνθήκη με επιλογή βήματος, βάσει του κανόνα του Armijo, είναι ακρότατο.

Σε ό,τι αφορά την ταχύτητα σύγκλισης της μεθόδου κλίσης υπό συνθήκη, όπως αναφέρεται στο [111], αν το σύνολο περιορισμών είναι ένα πολύεδρο, ο ασυμπτωτικός ρυθμός σύγκλισης της μεθόδου δεν είναι πολύ υψηλός. Αναφέρεται, όμως, ακόμα ότι, όσο αυξάνεται το πλήθος των περιορισμών (ακόμα και γραμμικών), η απόδοση της μεθόδου αυξάνεται. Στο μελετούμενο πρόβλημα, υπάρχουν $2N$ γραμμικοί (SIR) περιορισμοί, αλλά και N μη-γραμμικοί περιορισμοί (ωφελείας). Ως εκ τούτου, ο παρατηρούμενος ρυθμός σύγκλισης είναι αρκετά μεγάλος.

8.6.3 Ζητήματα Σταθερότητας του NBS

Όπως έχει ήδη συζητηθεί, το κύριο χαρακτηριστικό του NE είναι ότι κανείς χρήστης δεν μπορεί να επωφεληθεί, αλλάζοντας ατομικά την στρατηγική του, δηλαδή το NE αποτελεί ένα σταθερό σημείο λειτουργίας. Από την άλλη, το NBS, παρότι είναι κατά Pareto αποτελεσματικό, δε μπορεί να εγγυηθεί σταθερότητα, καθώς ένας χρήστης κάνοντας μία κατάλληλη αλλαγή στη στρατηγική του μπορεί να επωφεληθεί εις βάρος των άλλων χρηστών. Στο παίγνιο ελέγχου ισχύος, αν ένας χρήστης αποφασίσει να εκπέμψει με ισχύ που υπερβαίνει την ισχύ του NBS, θα αυξήσει το SIR του, ενώ, την ίδια στιγμή, το SIR των άλλων χρηστών θα μειωθεί. Ο ΣΒ χρειάζεται να αποθαρρύνει τέτοιου είδους συμπεριφορές και να επιβάλλει το σημείο λειτουργίας που αντιστοιχεί στο NBS. Στην εργασία [107], ο ΣΒ «τιμωρεί» τους αποκλίνοντες χρήστες αυξάνοντας σκοπίμως το BER τους. Συγκεκριμένα, ο ΣΒ μπορεί να προσφέρει στο χρήστη το ίδιο BER με το BER στο NBS, το οποίο σημαίνει ότι αν ένας χρήστης αυξήσει την ισχύ του, θα επιτύχει πάντα το BER του NBS. Έτσι, μία αύξηση στην ισχύ θα οδηγήσει μόνο σε επιπλέον κατανάλωση ενέργειας, χωρίς βελτίωση στο BER, γεγονός που σημαίνει ότι η ωφέλεια του

χρήστη θα μειωθεί. Επομένως κανείς λογικός χρήστης δεν θα είχε το συμφέρον να αποκλίνει την ισχύ του NBS.

8.7 Αριθμητικά Αποτελέσματα

Σε αυτό το υποκεφάλαιο, παρουσιάζονται τα αποτελέσματα από την ανάλυση της απόδοσης του προτεινόμενου σχήματος ελέγχου ισχύος. Ο Πίνακας 8.1 δίνει τις παραμέτρους του συστήματος που χρησιμοποιήθηκαν στην παρούσα ανάλυση²¹. Γίνεται η υπόθεση ότι ο χρήστης i έχει απαιτήσεις QoS, οι οποίες εκφράζονται μέσω αντίστοιχων πάνω και κάτω φραγμάτων στο SIR, $\gamma_{m,i}$ και $\gamma_{M,i}$, αντίστοιχα. Συγκεκριμένα, γίνεται η υπόθεση ότι

$$\gamma_{m,i} \sim U(\gamma_{min}, \gamma_{thres}) \text{ και } \gamma_{M,i} \sim U(\gamma_{thres}, \gamma_{max}),$$

όπου το $U(x, y)$ υποδηλώνει μία ομοιόμορφα κατανομημένη τυχαία μεταβλητή στο διάστημα $[x, y]$ και το $\gamma_{thres} \in [\gamma_{min}, \gamma_{max}]$. Με άλλα λόγια, το ελάχιστο επιθυμητό SIR $\gamma_{m,i}$ κάθε χρήστη επιλέγεται τυχαία μεταξύ ενός ελάχιστου SIR γ_{min} και του SIR γ_{thres} , ενώ το μέγιστο SIR $\gamma_{M,i}$ επιλέγεται τυχαία μεταξύ του SIR γ_{thres} κι ενός μέγιστου SIR γ_{max} . Στις τρέχουσες προσομοιώσεις, γίνεται η υπόθεση ότι $\gamma_{min} = 1$, $\gamma_{min} = 50$, ενώ $\gamma_{thres} = 5, 15, 25, 35$ και 45 . Μπορεί να αποδειχθεί εύκολα ότι, καθώς το γ_{thres} αυξάνεται, το SIR των χρηστών στο NE αυξάνεται (δες τον κανόνα (8.4)), κι ως εκ τούτου, λιγότεροι χρήστες μπορούν να γίνουν δεκτοί από το ΣΒ (δες (8.7)). Με άλλα λόγια, η τιμή του γ_{thres} υποδηλώνει την απαίτηση των χρηστών για πόρους. Ο Πίνακας 8.2, δείχνει τα όρια SIR που επιλέχθηκαν τυχαία για την αξιολόγηση.

Πίνακας 8.1. Λίστα παραμέτρων για το μονο-κυψελικό CDMA σύστημα

M , πλήθος bit στο πλαίσιο	80
L , πλήθος ωφέλιμων bit στο πλαίσιο	64
W , ρυθμός chip	3.84×10^6 Hz
R , ρυθμός εκπομπής δεδομένων	30×10^3 b/s
σ^2 , ισχύς AWGN power στο δέκτη	2×10^{-13} W
Τεχνική διαμόρφωσης	non coherent FSK
p_{max} , μέγιστη ισχύς εκπομπής	600 mW

²¹ Ρεαλιστικές τιμές για τις προσομοιώσεις προέκυψαν από τις εργασίες [81], [86] και [114]

Πίνακας 8.2. Απαιτήσεις SIR των χρηστών

Y_{thres} User	5		15		25		35		45	
	$Y_{m,i}$	$Y_{M,i}$	$Y_{m,i}$	$Y_{M,i}$	$Y_{m,i}$	$Y_{M,i}$	$Y_{m,i}$	$Y_{M,i}$	$Y_{m,i}$	$Y_{M,i}$
1	4.62	23.94	4.67	28.27	19.77	39.78	19.85	41.54	44.15	48.33
2	3.63	5.06	5.72	26.01	4.03	27.74	7.16	37.31	31.06	49.38
3	4.93	23.55	14.26	33.73	16.91	49.93	32.75	37.04	23.34	47.34
4	2.39	35.07	10.83	45.63	9.31	29.40	34.95	45.36	11.77	47.83
5	2.77	42.79	14.81	45.99	2.63	32.73	33.68	46.63	34.47	46.73
6	3.12	16.57	6.67	36.95	9.03	34.40	1.02	39.88	-	-
7	1.63	30.10	6.40	44.68	23.85	42.98	-	-	-	-
8	4.85	47.81	8.38	43.26	19.70	40.44	-	-	-	-
9	3.94	15.01	6.51	48.66	-	-	-	-	-	-
10	2.99	5.32	1.42	48.38	-	-	-	-	-	-
11	2.26	38.23	-	-	-	-	-	-	-	-
12	4.16	23.01	-	-	-	-	-	-	-	-

Οι μετρικές που χρησιμοποιήθηκαν για την αξιολόγηση του προτεινόμενου σχήματος, για κάθε χρήστη i , είναι (1) η ωφέλεια $v_i(\mathbf{q})$ που επιτυγχάνεται, (2) η ισχύς με την οποία ο χρήστης φτάνει το ΣΒ, q_i , (3) το SIR που επιτυγχάνεται, γ_i κι (4) η μέγιστη ακτίνα κυψέλης, $d_{i,max}$. Πρέπει να σημειωθεί ότι η συνάρτηση ωφελείας $v_i(\mathbf{q})$ υποδηλώνει την ωφέλεια του χρήστη i , ανεξάρτητα από τις συσχετιζόμενες απώλειες h_i (δες (8.13)). Επιπλέον, η μέγιστη ακτίνα κυψέλης, $d_{i,max}$, ορίζεται ως η μέγιστη απόσταση από την οποία ο χρήστης μπορεί να φτάσει το ΣΒ, εκπέμποντας με την ισχύ που αντιστοιχεί στο NE ή το NBS του παιγνίου. Συγκεκριμένα, γίνεται η υπόθεση ενός απλού μοντέλου απωλειών σήματος, $h_i = K_1/d_i^4$, όπου $K_1 = 0.097$ [106] και το d_i υποδηλώνει την απόσταση μεταξύ του χρήστη i και του ΣΒ. Δεδομένου του ότι ένας χρήστης μπορεί να εκπέμψει με ισχύ που δεν ξεπερνά μία μέγιστη τιμή p_{max} (Πίνακας 8.1), η απόσταση d_i από την οποία ο χρήστης i είναι ικανός να φτάσει το ΣΒ είναι φραγμένη. Έτσι, η μέγιστη ακτίνα κυψέλης $d_{i,max}$ είναι

$$d_{i,max} = \sqrt[4]{\frac{K_1 p_{max}}{q_i}}, \quad (8.19)$$

Όπου το q_i είναι ίσο με q_i^* ή \tilde{q}_i , για το NE ή το NBS, αντίστοιχα.

Ο Πίνακας 8.3, δείχνει τα επίπεδα ωφέλειας $v_i(q)$ που επιτυγχάνει κάθε χρήστης για διάφορες τιμές του γ_{thres} , στο NE και το NBS. Μπορεί να παρατηρηθεί ότι το προτεινόμενο συνεργατικό σχήμα προσφέρει σημαντική αύξηση στην ωφέλεια σε σύγκριση με το μη-συνεργατικό σχήμα. Αυτό είναι αναμενόμενο, λόγω της κατά Pareto κυριαρχίας του NBS σε σχέση με το NE του παιγνίου. Συγκεκριμένα, παρατηρείται ότι η σχετική αύξηση στην ωφέλεια είναι μεγαλύτερη για μικρές τιμές του γ_{thres} , όπως είναι αναμενόμενο. Αν το γ_{thres} αυξηθεί, οι απαιτήσεις SIR των χρηστών αυξάνονται, στην οποία περίπτωση, γίνεται πιθανότερο ότι το $\gamma_{m,i}$ μπορεί να ξεπεράσει το γ^* , που σημαίνει $\gamma_i^* = \gamma_{m,i}$ (δες (8.4)). Με άλλα λόγια, καθώς το γ_{thres} αυξάνεται, το σύνολο των χρηστών που μπορούν να ανεχθούν μία μείωση στο SIR τους, σε σχέση με το NE, σταδιακά περιορίζεται (δες υποκεφάλαιο 8.5), με αποτέλεσμα η μείωση στην ισχύ να γίνεται πιο μικρή και τα αντίστοιχα οφέλη λιγότερο σημαντικά. Αυτό μπορεί, καθαρά, να παρατηρηθεί για την περίπτωση, όπου $\gamma_{thres} = 45$, στην οποία μόνο ο χρήστης 4 έχει $\gamma_{m,i}$ μικρότερο από το γ^* (όπου $\gamma^* = 12.42$, στο παρόν σενάριο) κι έτσι μπορεί να θυσιάσει το SIR του, προς όφελος όλων των χρηστών. Παρόλο που το όφελος είναι θετικό, μπορεί να παρατηρηθεί ότι είναι σημαντικά μικρότερο, σε σχέση με τις περιπτώσεις μικρότερων γ_{thres} .

Πίνακας 8.3. Ωφέλεια στο NE και το NBS

γ_{thres} Χρήστης	5		15		25		35		45	
	NE	NBS	NE	NBS	NE	NBS	NE	NBS	NE	NBS
1	4.34E+17	2.02E+18	1.02E+18	1.92E+18	9.4E+17	1.46E+18	5.55E+17	9.31E+17	4.37E+17	4.57E+17
2	1.54E+15	1.02E+16	1.02E+18	1.94E+18	1.22E+18	1.34E+18	7.2E+17	9.28E+17	5.74E+17	6E+17
3	4.34E+17	1.99E+18	9.93E+17	2.6E+18	1.06E+18	1.66E+18	3.67E+17	6.16E+17	7.27E+17	7.6E+17
4	4.34E+17	1.99E+18	1.02E+18	2.62E+18	1.22E+18	1.82E+18	3.49E+17	5.85E+17	1.08E+18	1.11E+18
5	4.34E+17	1.99E+18	9.75E+17	2.55E+18	1.22E+18	1.32E+18	3.59E+17	6.02E+17	5.29E+17	5.52E+17
6	4.34E+17	1.99E+18	1.02E+18	1.92E+18	1.22E+18	1.8E+18	7.2E+17	7.5E+17	-	-
7	4.34E+17	2.04E+18	1.02E+18	1.92E+18	8.04E+17	1.25E+18	-	-	-	-
8	4.34E+17	1.99E+18	1.02E+18	2.31E+18	9.43E+17	1.47E+18	-	-	-	-
9	4.34E+17	2.03E+18	1.02E+18	1.92E+18	-	-	-	-	-	-
10	3.38E+15	2.23E+16	1.02E+18	1.92E+18	-	-	-	-	-	-
11	4.34E+17	1.99E+18	-	-	-	-	-	-	-	-
12	4.34E+17	2.01E+18	-	-	-	-	-	-	-	-

Ο Πίνακας 8.4 δείχνει τις τιμές της λαμβανόμενης ισχύος q_i του χρήστη i στο ΣΒ για το NE και το NBS. Παρατηρείται ότι η λαμβανόμενη ισχύς στο NE είναι σημαντικά υψηλότερη, σε σχέση με την αντίστοιχη ισχύ στο NBS. Με άλλα λόγια, η λειτουργία

στο σημείο NBS μπορεί να επιφέρει σημαντική εξοικονόμηση ενέργειας. Όπως και στην περίπτωση της ωφέλειας που επιτυγχάνεται, παρατηρείται ότι, καθώς η τιμή του γ_{thres} αυξάνεται, η μείωση που επιτυγχάνεται στην ισχύ γίνεται λιγότερο σημαντική. Αυτό αποδίδεται στο γεγονός ότι, όταν το γ_{thres} αυξάνεται, το ποσοστό των χρηστών που θα μπορούσαν να θυσιάσουν το SIR τους περιορίζεται, με αποτέλεσμα τη μειωμένη ελάττωση στην ισχύ (δες π.χ. για $\gamma_{thres} = 45$).

Πίνακας 8.4. Λαμβανόμενη ισχύς (W) στο NE και το NBS

γ_{thres} Χρήστης	5		15		25		35		45	
	NE	NBS	NE	NBS	NE	NBS	NE	NBS	NE	NBS
1	4.71E-13	5.29E-14	2E-13	5.87E-14	2.54E-13	1.63E-13	4.31E-13	2.57E-13	5.49E-13	5.25E-13
2	2.02E-13	3E-14	2E-13	5.9E-14	1.68E-13	8.26E-14	2.84E-13	1.33E-13	4.18E-13	4E-13
3	4.71E-13	5.27E-14	2.27E-13	8.66E-14	2.22E-13	1.42E-13	6.54E-13	3.9E-13	3.3E-13	3.16E-13
4	4.71E-13	5.27E-14	2E-13	7.16E-14	1.68E-13	9.81E-14	6.88E-13	4.1E-13	1.89E-13	1.72E-13
5	4.71E-13	5.27E-14	2.34E-13	8.96E-14	1.68E-13	8.21E-14	6.68E-13	3.98E-13	4.54E-13	4.34E-13
6	4.71E-13	5.27E-14	2E-13	5.87E-14	1.68E-13	9.72E-14	2.84E-13	1.24E-13	-	-
7	4.71E-13	5.31E-14	2E-13	5.88E-14	2.99E-13	1.92E-13	-	-	-	-
8	4.71E-13	5.27E-14	2E-13	6.38E-14	2.54E-13	1.63E-13	-	-	-	-
9	4.71E-13	5.3E-14	2E-13	5.88E-14	-	-	-	-	-	-
10	2.12E-13	3.14E-14	2E-13	5.87E-14	-	-	-	-	-	-
11	4.71E-13	5.27E-14	-	-	-	-	-	-	-	-
12	4.71E-13	5.28E-14	-	-	-	-	-	-	-	-

Πίνακας 8.5. SIR στο NE και το NBS

γ_{thres} Χρήστης	5		15		25		35		45	
	NE	NBS	NE	NBS	NE	NBS	NE	NBS	NE	NBS
1	12.42	9.20	12.42	9.33	19.77	19.77	19.85	19.85	44.15	44.15
2	5.06	5.05	12.42	9.38	12.42	9.30	12.42	9.60	31.06	31.06
3	12.42	9.15	14.26	14.26	16.91	16.91	32.75	32.75	23.34	23.34
4	12.42	9.15	12.42	11.55	12.42	11.20	34.95	34.95	12.42	11.77
5	12.42	9.16	14.81	14.81	12.42	9.23	33.68	33.68	34.47	34.47
6	12.42	9.15	12.42	9.33	12.42	11.09	12.42	8.89	-	-
7	12.42	9.23	12.42	9.34	23.85	23.85	-	-	-	-
8	12.42	9.16	12.42	10.20	19.70	19.70	-	-	-	-
9	12.42	9.21	12.42	9.34	-	-	-	-	-	-
10	5.32	5.31	12.42	9.33	-	-	-	-	-	-
11	12.42	9.15	-	-	-	-	-	-	-	-
12	12.42	9.18	-	-	-	-	-	-	-	-

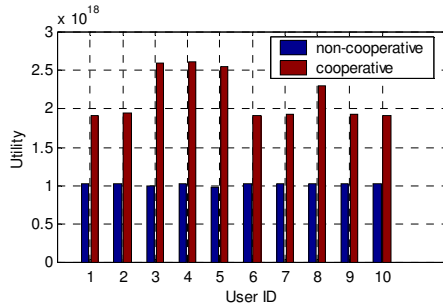
Ο Πίνακας 8.5 δείχνει το SIR που επιτυγχάνεται κι από τις δύο προσεγγίσεις. Πρέπει να σημειωθεί ότι το SIR στο NBS είναι υποδεέστερο του SIR στο NE. Σε μερικές περιπτώσεις, η μείωση στο SIR φτάνει και το 27%, για $\gamma_{thres} = 5$, αλλά η αντίστοιχη μείωση ισχύος ξεπερνά το 88%, στο ίδιο σενάριο. Παρατηρείται, επίσης, ότι, όσο αυξάνεται το γ_{thres} , η μείωση στο SIR γίνεται μικρότερη, καθώς περισσότεροι χρήστες διατηρούν το SIR που είχαν στο NE. Πρέπει να σημειωθεί, όμως, ότι σε όλα τα σενάρια κάθε επιδείνωση στο SIR είναι, μάλλον, περιορισμένης σημασίας, συγκρινόμενη με την αντίστοιχη εξοικονόμηση ενέργειας, λόγω της μειωμένης ισχύος εκπομπής (Πίνακας 8.4).

Ο Πίνακας 8.6 δείχνει τα αποτελέσματα σχετικά με τη μέγιστη ακτίνα κυψέλης. Παρατηρείται ότι στο NBS επιτυγχάνεται αξιοσημείωτη αύξηση στη μέγιστη ακτίνα κυψέλης, σε σχέση με το NE. Η εν λόγω αύξηση είναι πιο εμφανής για μικρότερες τιμές του της παραμέτρου γ_{thres} , για τους λόγους που συζητήθηκαν και πιο πάνω.

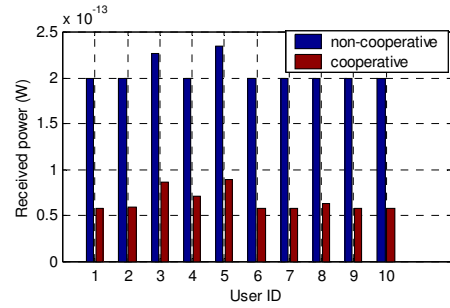
Πίνακας 8.6. Μέγιστη ακτίνα κυψέλης (m) στο NE και το NBS

γ_{thres} Χρήστης	5		15		25		35		45	
	NE	NBS	NE	NBS	NE	NBS	NE	NBS	NE	NBS
1	593.01	1024.15	734.49	997.67	691.69	772.75	606.29	689.97	570.66	576.99
2	732.39	1180.42	734.49	996.50	767.07	916.20	672.96	812.68	610.90	617.68
3	593.01	1025.22	711.88	905.32	715.73	799.62	546.27	621.67	648.03	655.22
4	593.01	1025.34	734.49	949.67	767.07	877.56	539.28	613.72	744.64	762.17
5	593.01	1025.18	705.87	897.69	767.07	917.72	543.25	618.23	598.35	604.99
6	593.01	1025.22	734.49	997.68	767.07	879.55	672.96	827.42	-	-
7	593.01	1023.22	734.49	997.53	664.48	742.35	-	-	-	-
8	593.01	1025.21	734.49	977.35	692.18	773.30	-	-	-	-
9	593.01	1023.74	734.49	997.61	-	-	-	-	-	-
10	723.70	1166.56	734.49	997.84	-	-	-	-	-	-
11	593.01	1025.34	-	-	-	-	-	-	-	-
12	593.01	1024.48	-	-	-	-	-	-	-	-

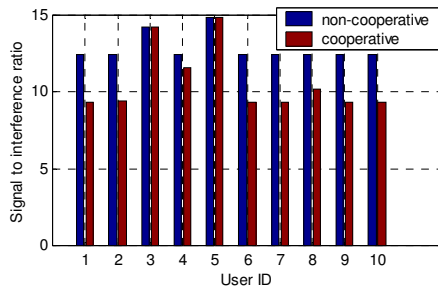
Στην Εικόνα 8.2-Εικόνα 8.5, παρουσιάζεται η αποτελεσματικότητα των δύο εξεταζόμενων προσεγγίσεων για $\gamma_{thres} = 15$, που θεωρείται ένα τυπικό σενάριο. Τα οφέλη από την υιοθεσία του συνεργατικού σχήματος είναι εμφανή.



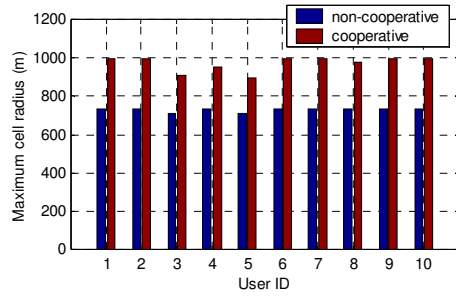
Εικόνα 8.2. Επίπεδο ωφέλειας στο NE και το NBS



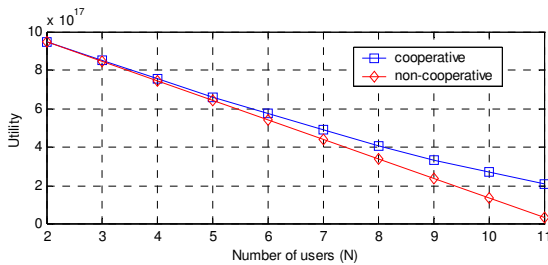
Εικόνα 8.3. Λαμβανόμενη ισχύς στο NE και το NBS



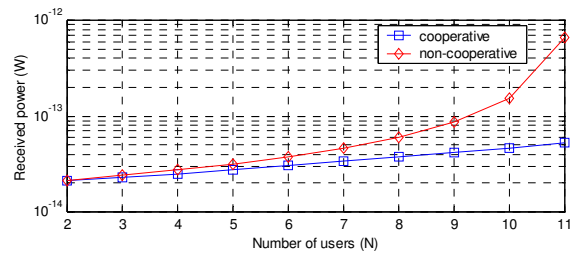
Εικόνα 8.4. SIR στο NE και το NBS



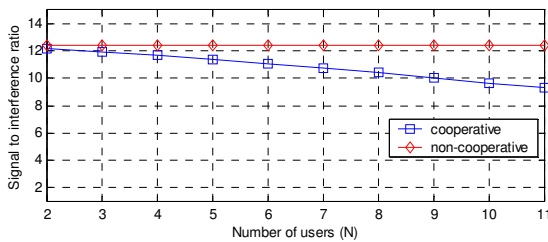
Εικόνα 8.5. Μέγιστη ακτίνα κυψέλης στο NE και το NBS



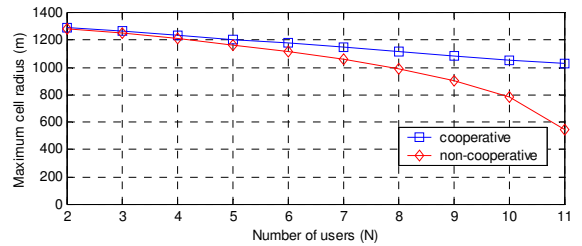
Εικόνα 8.6. SIR στο NE και το NBS



Εικόνα 8.7. Λαμβανόμενη ισχύς στο NE και το NBS



Εικόνα 8.8. SIR στο NE και το NBS



Εικόνα 8.9. Μέγιστη ακτίνα κυψέλης στο NE και το NBS

Ένα ακόμα σημαντικό ζήτημα αποτελεί η αποτελεσματικότητα της προτεινόμενης μεθόδου για διαφορετικά πλήθη χρηστών. Αυτό το ζήτημα εξετάζεται παρακάτω. Γίνεται η υπόθεση ότι το πλήθος των χρηστών κυμαίνεται μεταξύ δύο κι έντεκα κι ότι όλοι οι χρήστες έχουν ταυτόσημες απαιτήσεις σε SIR. Παρουσιάζονται οι ίδιες

μετρικές με τις παραπάνω, δηλαδή, ωφέλεια, λαμβανόμενη ισχύς στο ΣΒ, SIR και μέγιστη ακτίνα κυψέλης. Στην Εικόνα 8.6-Εικόνα 8.9, παρατηρείται ότι οι τιμές των εξεταζομένων μετρικών βελτιώνονται για κάθε πλήθος χρηστών. Όμως, φαίνεται ότι, όσο το πλήθος των χρηστών αυξάνεται, και το σύστημα φτάνει στα όρια της χωρητικότητας του, οι εν λόγω βελτιώσεις γίνονται πιο σημαντικές.

8.8 Συμπεράσματα

Στο κεφάλαιο αυτό, προτάθηκε μία λύση για το πρόβλημα ελέγχου ισχύος για CDMA ασύρματα δίκτυα δεδομένων. Έγινε η υπόθεση ότι οι υπηρεσίες των χρηστών είναι μη ελαστικές σε καθυστέρηση, αλλά ελαστικές σε εύρος ζώνης, όπου δηλαδή κάποια διακύμανση στο ρυθμό μετάδοσης δεν αποτελεί σημαντικό πρόβλημα, αλλά μόνο πολύ μικρές καθυστερήσεις είναι ανεκτές (π.χ. φωνή, video, μεταφορά αρχείων σε πραγματικό χρόνο κτλ). Τέτοιοι περιορισμοί εύρους ζώνης (ρυθμοαπόδοσης) αντιστοιχήθηκαν απευθείας σε αντίστοιχα όρια SIR. Το πρόβλημα μοντελοποιήθηκε αρχικά ως μη-συνεργατικό κι εν συνεχεία ως συνεργατικό παίγνιο, το οποίο επιλύθηκε βάσει της ιδέας του NBS. Η συνεργατική λύση, σε αντίθεση με τη μη-συνεργατική προσέγγιση, εγγυάται ότι το αποτέλεσμα του παιγνίου θα είναι κατά Pareto αποτελεσματικό και δίκαιο. Προτάθηκε, επίσης, ένας αριθμητικός αλγόριθμος για τον καθορισμό του NBS του παιγνίου.

Η συνεργατική λύση συγκρίθηκε με τη μη-συνεργατική και φάνηκε ότι η αναμενόμενη κατά Pareto βελτίωση μέσω του συνεργατικού σχήματος είναι σημαντική. Συγκεκριμένα, τα σημεία λειτουργίας του προτεινόμενου συνεργατικού σχήματος χαρακτηρίζονται από σημαντικά μειωμένη ισχύ εκπομπής και μία, μάλλον, περιορισμένη μείωση στο SIR. Αυτό είχε ως αποτέλεσμα μία δραστική μείωση στην κατανάλωση ισχύος (που οδηγεί σε αύξηση της ζωής του τερματικού), προκαλώντας μία ελαφρά ελάττωση στη ρυθμοαπόδοση. Επιπλέον, λόγω της μείωσης της ισχύος εκπομπής, η μέγιστη ακτίνα της κυψέλης αυξήθηκε σε σύγκριση με το μη-συνεργατικό σχήμα. Μελλοντικά, θα ήταν επιθυμητό να επεκταθεί η μελέτη σε περιβάλλοντα με πολλές κυψέλες και να εφαρμοστεί η συνεργατική θεωρία παιγνίων για να βελτιωθούν υπάρχοντα σχήματα ελέγχου ισχύος.

Μέρος Β

Διαχείριση Πόρων σε Κατανεμημένα Συστήματα

ΚΕΦΑΛΑΙΟ 9

ΚΑΤΑΝΕΜΗΜΕΝΟ ΣΧΗΜΑ ΓΙΑ ΤΟ ΔΙΑΜΟΙΡΑΣΜΟ ΧΩΡΗΤΙΚΟΤΗΤΑΣ ΣΥΣΤΗΜΑΤΩΝ WWW CACHING

9.1 Εισαγωγή

Στο κεφάλαιο αυτό, παρουσιάζεται η μελέτη που πραγματοποιήθηκε, στα πλαίσια της παρούσας διατριβής, σχετικά με το διαμοιρασμό της χωρητικότητας αποθήκευσης σε συστήματα WWW caching. Η χρήση των web cache για τη μείωση της δικτυακής κίνησης και της καθυστέρησης μεταφοράς της υπηρεσίας WWW τις έχει καταστήσει ένα σημαντικό συστατικό της υποδομής του Διαδικτύου. Οι αλγόριθμοι που χρησιμοποιούνται για τις web cache σχεδιάζονται με σκοπό να εκμεταλλεύονται συγκεκριμένα χαρακτηριστικά που υπάρχουν στις ακολουθίες των HTTP αιτήσεων. Για παράδειγμα, έχει παρατηρηθεί, από πραγματικά δεδομένα αιτήσεων, ότι κάποιοι χρήστες, οι οποίοι επιδεικνύουν «επιθετική» συμπεριφορά, τείνουν να μονοπωλήσουν το χώρο στο δίσκο της cache, απολαμβάνοντας υψηλές πιθανότητες ευστοχίας [123]. Την ίδια στιγμή, όμως, οι άλλοι χρήστες περιορίζονται στο να χρησιμοποιούν μία, μάλλον, μικρή ποσότητα από το χώρο στο δίσκο κι υποφέρουν από την αποβολή των «σημαντικών» για αυτούς αντικειμένων, με αποτέλεσμα πολυάριθμες αποτυχημένες αιτήσεις προς την cache (cache miss). Οι συμβατικοί αλγόριθμοι αντικατάστασης [70] εφαρμόζουν καθολικές πολιτικές, δηλαδή αντιμετωπίζουν όλους τους χρήστες ως σύνολο, χωρίς να λαμβάνεται υπόψη η ατομική συμπεριφορά των χρηστών. Ως εκ τούτου, δεν είναι ικανοί να παρέχουν την υπηρεσία caching με δίκαιο τρόπο.

Όπως αναφέρεται στην εργασία [117], οι επιτυχείς αιτήσεις προς την cache (cache hits) αποδίδονται κυρίως σε διαμοιραζόμενα αντικείμενα (“sharing hits”). Τα υπόλοιπα ονομάζονται “locality hits”. Παρατηρείται, δηλαδή, ότι οι χρήστες έχουν την τάση να ζητούν αντικείμενα που ανήκουν σε μία δημοφιλή «δεξαμενή» αντικειμένων, ενώ τα υπόλοιπα αντικείμενα σπάνια ζητούνται πάνω από μία φορά. Στην παρούσα μελέτη, προτείνεται ένας κατανεμημένος, παιγνιοθεωρητικός, αλγόριθμος που είναι ικανός να παρέχει την υπηρεσία caching στους χρήστες με

δίκαιο τρόπο, λαμβάνοντας, επίσης, υπόψη τη συνεισφορά κάθε χρήστη στη συνολική απόδοση του συστήματος caching. Συγκεκριμένα, σε ό,τι αφορά τη συμπεριφορά των χρηστών στη συνολική απόδοση του συστήματος, το προτεινόμενο σχήμα επιβραβεύει τους χρήστες που έχουν φέρει δημοφιλή αντικείμενα στην cache, επιτρέποντας τους να χρησιμοποιούν περισσότερο χώρο του δίσκου. Στην παρούσα μελέτη γίνεται η υπόθεση ότι η υπηρεσία caching παρέχεται εμπορικά (π.χ. από τον ISP) κι οι χρήστες χρεώνονται για να χρησιμοποιήσουν τους παρεχόμενους πόρους (χωρητικότητα δίσκου) αυτής της εμπορικής οντότητας [125].

Οι χρήστες θεωρούνται εγωιστές, δηλαδή δε συνεργάζονται μεταξύ τους. Δεν υπάρχει κάποια οντότητα υπεύθυνη να συντονίζει την πρόσβαση στην cache. Ο διακομιστής caching μεταδίδει στους χρήστες καθολικές πληροφορίες σε σχέση με την κατάσταση της cache, π.χ. τρέχων ελεύθερος χώρος, κι ειδικές για τους χρήστες πληροφορίες, π.χ. χώρος που καταλαμβάνει ο χρήστης στην cache. Αυτές οι πληροφορίες παραδίδονται στον χρήστη ενσωματωμένες στις HTTP απαντήσεις. Την επόμενη φορά που ο χρήστης στέλνει μία αίτηση, λαμβάνει υπόψη του τις πληροφορίες που συγκέντρωσε μέσω της τελευταίας HTTP απάντησης, ώστε να καθορίσει αν «συμφέρει» να ζητήσει την αποθήκευση, στην cache, του ζητούμενου αντικειμένου ή το κόστος που θα επιφέρει αυτή η αποθήκευση ξεπερνά το αντίστοιχο όφελος. Στη δεύτερη περίπτωση, ο χρήστης θέτει την επικεφαλίδα της HTTP αίτησης “cache-control: no-store” [70] υποδηλώνοντας ότι, αν το αντικείμενο δεν βρίσκεται ήδη στην cache, το αντικείμενο που θα αποκτηθεί από τον πηγαίο διακομιστή περιεχομένου να μην αποθηκευθεί στην cache.

9.2 Σχετική Βιβλιογραφία

Στη βιβλιογραφία, υπάρχουν μόνο λίγες πρόσφατες εργασίες που μελετούν με παιγνιοθεωρητικές μεθόδους το πρόβλημα του caching και replication περιεχομένου Ιστού. Η πιο σχετική εργασία με την παρούσα μελέτη είναι η [126], όπου ο ανταγωνισμός μεταξύ διαφορετικών χρηστών για την αποθηκευτική χωρητικότητα ενός διακομιστή caching μοντελοποιήθηκε και μελετήθηκε ως ένα μη

συνεργατικό παίγνιο. Όμως, η εργασία αυτή δεν παρείχε κάποιο αναλυτικό μοντέλο του προτεινόμενου σχήματος κι οι προσομοιώσεις βασίστηκαν σε συνθετικές ροές αιτήσεων. Άλλες, λιγότερο σχετικές, εργασίες επικεντρώνονται στο πρόβλημα της «κατανεμημένης εγωιστικής διατήρησης αντιγράφων» (distributed selfish replication – DSR). Το DSR ορίζεται ως η διατήρηση αντιγράφων αντικειμένων από διακομιστές που δρουν με εγωιστικό τρόπο. Το πρόβλημα αυτό έχει μελετηθεί ως ένα μη συνεργατικό παίγνιο στην εργασία [129]. Πιο πρόσφατα, στην εργασία [127], υπολογίστηκαν στρατηγικές ισορροπίας που αφορούν στην τοποθέτηση των αντικειμένων στο πρόβλημα DSR και προτάθηκαν δυο κατανεμημένοι αλγόριθμοι. Ακόμα μία εργασία σε σχέση με την εκχώρηση πόρων βάσει οικονομικών πλαισίων σε δίκτυα διανομής περιεχομένου είναι η [128]. Η μελέτη στην εργασία [128] εστιάζεται σε μία «αγορά» τριών μερών μεταξύ παρόχων περιεχομένου, διανομέων και πελατών.

9.3 Μοντέλο Συστήματος

Το μοντέλο συστήματος που υιοθετήθηκε έχει τις ρίζες του στο μοντέλο που προτείνεται στην εργασία [121]. Έστω $I = \{1, \dots, N\}$ το σύνολο των χρηστών που μοιράζονται τους πόρους ενός διακομιστή caching με δίσκο χωρητικότητας C . Ο χρήστης i ελέγχει τη χωρητικότητα y_i που καταλαμβάνει στην cache, ώστε να μπορεί να βελτιστοποιεί ένα συγκεκριμένο μέτρο απόδοσης. Το y_i επιλέγεται από το $S_i = [0, C]$. Έστω $\mathbf{y} = (y_1, \dots, y_N)$ ένα τυπικό διάνυσμα στρατηγικής (ή κατανομή χωρητικότητας στο δίσκο) στο χώρο των στρατηγικών $S = S_1 \times \dots \times S_N$ και

$$F = \sum_{i=1}^N y_i \text{ η συνολικά δεσμευμένη χωρητικότητα στην cache.}$$

Γίνεται η υπόθεση ότι κάθε χρήστης έχει μία συνάρτηση ωφελείας. Η συνάρτηση ωφελείας που αποτελείται από ένα μέρος οφέλους κι από ένα μέρος κόστους ποσοτικοποιεί το βαθμό ικανοποίησης του χρήστη σε ό,τι αφορά τη χρήση του διακομιστή caching. Η συνάρτηση κόστους επηρεάζεται από τα χαρακτηριστικά του συστήματος caching, π.χ. το σχήμα χρέωσης που έχει υιοθετηθεί. Κάθε χρήστης είναι ικανός να υπολογίζει το κόστος από πληροφορίες προερχόμενες από το σύστημα.

Γίνεται η υπόθεση ότι η συνάρτηση ωφελείας $u_i: S \rightarrow \mathfrak{R}$ του χρήστη i εξαρτάται μόνο από τη χωρητικότητα y_i που καταλαμβάνει ο χρήστης και τη συνολική χωρητικότητα F που έχει εκχωρηθεί, δηλαδή $u_i(\mathbf{y}) = u_i(y_i, F)$. Ο χρήστης i μπορεί να μεγιστοποιήσει τη συνάρτηση $u_i(y_i, F)$ ελέγχοντας το y_i , δεδομένης της συνολικής χωρητικότητας του συστήματος C και της «συναθροιστικής» χωρητικότητας F όλων των χρηστών. Γίνεται η υπόθεση ότι η $u_i(y_i, F)$ θα πρέπει να έχει τις ακόλουθες ιδιότητες:

- A1.** Η $u_i(y_i, F)$ είναι συνεχώς διαφορίσιμη ως προς το y_i .
- A2.** Η $\partial u_i(y_i, F)/\partial y_i$ είναι αυστηρώς φθίνουσα ως προς το y_i .
- A3.** Η $\partial u_i(y_i, F)/\partial F$ είναι μη-αύξουσα ως προς το F .

Οι πρώτες δύο υποθέσεις υποδηλώνουν ότι η u_i θα πρέπει να είναι αυστηρά κοίλη ως προς το y_i . Έστω $C_i = C - \sum_{j \neq i} y_j$ η διαθέσιμη χωρητικότητα της cache που βλέπει ο χρήστης i . Πρέπει να σημειωθεί ότι $F = y_i + C - C_i$. Τότε, η καλύτερη απόκριση του χρήστη i , για μία δεδομένη διαθέσιμη χωρητικότητα C_i , ορίζεται ως

$$r_i(C_i) = \operatorname{argmax}_{g \in S_i} u_i(g, g + C - C_i). \quad (9.1)$$

Με άλλα λόγια, το $r_i(C_i)$ αποτελεί τη βέλτιστη ποσότητα χωρητικότητας στην cache που θα πρέπει να καταλαμβάνει ο χρήστης i , αν η διαθέσιμη χωρητικότητα για τον i είναι C_i . Γίνεται επίσης κι η επόμενη υπόθεση σε ό,τι αφορά τη συνάρτηση απόκρισης του χρήστη.

- A4.** Δεδομένου κάθε C_i όπου $0 < C_i \leq C$, τότε $r_i(C_i) < C_i$

Η παραπάνω υπόθεση ισχύει, συνήθως, σε περιπτώσεις, όπου η συνάρτηση κόστους έχει τη μορφή τιμωρίας (π.χ. χρέωση) εμποδίζοντας τον κορεσμό των πόρων. Ένα παράδειγμα τέτοιας συνάρτησης ωφελείας που συμφωνεί με τις τέσσερις υποθέσεις είναι η “generalized power” [118]:

$$u_i(y_i, F) = (y_i)^{\beta_i} (C - F) \quad (9.2)$$

η οποία ερμηνεύεται ως όφελος $(y_i)^{\beta_i}$ διηρημένο με το κόστος $(C-F)^{-1}$. Αυτού του τύπου η συνάρτηση ωφελείας ικανοποιεί και τις τέσσερις υποθέσεις A1-A4, αν $0 < \beta_i \leq 1$.

Στην παρούσα μελέτη, προτείνεται μία παρόμοια συνάρτηση ωφελείας που επίσης λαμβάνει υπόψη την επίδραση που μπορεί να έχει ένας χρήστης στη συνολική απόδοση του συστήματος caching. Συγκεκριμένα, ο χρήστης i συσχετίζεται με ένα δείκτη δημοτικότητας (popularity index) a_i . Αυτός ο δείκτης υπολογίζεται δυναμικά από τον διακομιστή caching κι υποδηλώνει την «δημοτικότητα» των αντικειμένων που ο χρήστης i έχει φέρει στην cache. Ο δείκτης δημοτικότητας, για το χρήστη i , a_i , λαμβάνει τιμές από το διάστημα $[0,1)$.

Γίνεται η υπόθεση ότι κάθε αποθηκευμένο αντικείμενο συνοδεύεται από κάποια μετα-πληροφορία με σκοπό τη διευκόλυνση της λειτουργίας του συστήματος. Συγκεκριμένα, για κάθε αποθηκευμένο αντικείμενο $j \in \{1, \dots, K\}$, ο διακομιστή caching γνωρίζει το χρήστη i_j που έφερε το αντικείμενο στην cache, το μέγεθος s_j του αντικειμένου σε bytes και το πλήθος h_j των αιτήσεων που έχουν γίνει για το συγκεκριμένο αντικείμενο από τους χρήστες εκτός του χρήστη i_j . Ο δείκτης δημοτικότητας a_i του χρήστη i υπολογίζεται ως εξής:

$$a_i = \begin{cases} \frac{\sum_{j \in I_i} h_j \cdot s_j}{\sum_{j=1}^K h_j \cdot s_j}, & \sum_{j=1}^K h_j \cdot s_j \neq 0 \\ 0, & \text{otherwise} \end{cases} \quad (9.3)$$

όπου I_i είναι το σύνολο των αντικειμένων που ο χρήστης i έχει φέρει στην cache.

Η συνάρτηση ωφελείας του χρήστη i ορίζεται ως

$$u_i(y_i, F) = (y_i)^{\beta_i} (C - F)^{1-a_i}. \quad (9.4)$$

Η προτεινόμενη συνάρτηση ωφελείας είναι παρόμοιας μορφής με τη (9.2), εκτός του μέρους που αφορά στο κόστος. Συγκεκριμένα, το κόστος του χρήστη i μειώνεται κατά ένα παράγοντα $1-a_i$. Με άλλα λόγια, «δημοφιλείς» χρήστες χρεώνονται λιγότερο από τους μη δημοφιλείς χρήστες, αποκτώντας έτσι περισσότερη ελευθερία στο να χρησιμοποιούν το χώρο του δίσκου του διακομιστή caching.

Πρόταση 9.1: Η συνάρτηση ωφελείας $u_i(\cdot)$, για κάθε $i \in I$, ικανοποιεί τις υποθέσεις A1-A4.

Απόδειξη:

Υπόθεση A1

Μπορεί εύκολα να αποδειχθεί ότι η παράγωγος της $u_i(\cdot)$ ως προς το y_i , για κάθε $i \in I$, είναι συνεχής ως προς το y_i στο διάστημα S_i .

Υπόθεση A2

Η δεύτερη παράγωγος της $u_i(\cdot)$ ως προς το y_i , για κάθε $i \in I$, είναι αρνητική, που σημαίνει ότι η A2 ισχύει επίσης.

Υπόθεση A3

Η δεύτερη παράγωγος της $u_i(\cdot)$ ως προς το F είναι μη θετική, επομένως, η A3 ισχύει.

Υπόθεση A4

Δεδομένου ενός C_i , η συνθήκη βελτιστότητας $\nabla u_i = 0$ δίνει

$$\arg \max_{y_i \in S_i} u_i(y_i, F) = \frac{\beta_i}{1 + \beta_i - a_i} C_i < C_i$$

που σημαίνει ότι η A4, επίσης, ισχύει κι η πρόταση, τελικά, αποδεικνύεται. ■

Οι χρήστες είναι εγωιστές, γεγονός που υποδηλώνει ότι κάθε χρήστης ενδιαφέρεται μόνο για να μεγιστοποιήσει τη δική του συνάρτηση ωφελείας. Γίνεται η υπόθεση ότι η u_i αποτελεί ιδιωτική πληροφορία για το χρήστη i , ενώ άλλες πληροφορίες (π.χ. F , C) είναι κοινά γνωστές και παρέχονται από το διακομιστή caching. Τότε, δεδομένου του F , κάθε χρήστης υπολογίζει τη βέλτιστη στρατηγική επιλύοντας το ακόλουθο πρόβλημα βελτιστοποίησης (P).

$$(P) \begin{cases} \max_{y_i} u_i(y_i, F) \\ \text{s.t.} \\ y_i \in S_i, i \in I \end{cases}$$

Το γεγονός ότι η u_i εξαρτάται από το \mathbf{y} μόνο ως προς το y_i και το $\sum_{j \neq i} y_j$, οδηγεί το χρήστη i να θεωρήσει το σύνολο των υπόλοιπων χρηστών ως έναν μοναδικό αντίπαλο, κάνοντας και την ανάλυση που ακολουθεί πιο απλή. Το διάνυσμα $\mathbf{y}^* = (y_1^*, \dots, y_N^*) \in S$ αποτελεί μία ισορροπία Nash [72] αν και μόνον αν

$$y_i^* \in \operatorname{argmax}_{g \in \mathcal{S}_i} u_i \left(g, g + \sum_{j \neq i} y_j^* \right), \quad \forall i \in I. \quad (9.5)$$

Με άλλα λόγια, στο NE, δεδομένης της χωρητικότητας δίσκου που έχει εκχωρηθεί στους υπόλοιπους χρήστες, κανείς χρήστης δε μπορεί να βελτιώσει την ωφέλεια του πραγματοποιώντας ατομικές μεταβολές στην ποσότητα του χώρου που καταλαμβάνει. Στο επόμενο υποκεφάλαιο, μελετάται η ύπαρξη κι η μοναδικότητα ενός NE.

9.4 Ύπαρξη και Μοναδικότητα της Nash Ισορροπίας

Λόγω των ιδιοτήτων της συνάρτησης ωφελείας, που αναλύθηκαν στο προηγούμενο υποκεφάλαιο, η ύπαρξη του NE είναι εγγυημένη από το Θεώρημα 1 στην εργασία [119]. Για κάθε NE, το σύνολο των εξισώσεων σταθερού σημείου, που προκύπτουν από τη συνθήκη βελτιστότητας $\nabla u_i = 0$, μπορεί να γραφτεί ως ένα σύστημα εξισώσεων, ως εξής:

$$\begin{bmatrix} \beta_1 + 1 - a_1 & \beta_1 & \dots \\ \beta_2 & \beta_2 + 1 - a_2 & \dots \\ \dots & \dots & \dots \\ \beta_N & \dots & \dots \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix} = \begin{bmatrix} \beta_1 C \\ \beta_2 C \\ \dots \\ \beta_N C \end{bmatrix} \quad (9.6)$$

Η μοναδικότητα του NE μπορεί να αποδειχθεί αν διερευνηθούν οι ιδιότητες του συγκεκριμένου συστήματος.

Πρόταση 9.2: Υπάρχει ένα και μόνο NE $\mathbf{y}^* = (y_1^*, \dots, y_N^*)^T$ για το παίγνιο, το οποίο είναι εφικτό, δηλαδή $\sum_{i=1}^N y_i^* < C$, με $y_i^* = C \left(\frac{\beta_i}{1 - a_i} \right) \left[1 + \sum_{j=1}^N \left(\frac{\beta_j}{1 - a_j} \right) \right]^{-1}$.

Απόδειξη: Η μοναδικότητα του NE οφείλεται στο γεγονός ότι το σύστημα των εξισώσεων σταθερού σημείου είναι γραμμικό κι ο πίνακας των συντελεστών *πλήρους τάξης* (full rank). Η λύση του συστήματος $\mathbf{y}^* = (y_1^*, y_2^*, \dots, y_N^*)^T$ είναι η

$$y_i^* = C \left(\frac{\beta_i}{1 - a_i} \right) \left[1 + \sum_{j=1}^N \left(\frac{\beta_j}{1 - a_j} \right) \right]^{-1}, \quad \text{για κάθε } i \in I.$$

Προσθέτοντας τα στοιχεία του διανύσματος της λύσης, προκύπτει ότι

$$\sum_{i=1}^N y_i^* = C \left(\sum_{i=1}^N \frac{\beta_i}{1-a_i} \right) \left[1 + \sum_{j=1}^N \left(\frac{\beta_j}{1-a_j} \right) \right]^{-1} < C$$

Επομένως, η λύση του συστήματος είναι εφικτή κι η πρόταση αποδεικνύεται. ■

Στη συνέχεια μελετούνται ζητήματα σχετικά με τη δικαιοσύνη που επιτυγχάνεται στο εξεταζόμενο NE. Στην παρούσα μελέτη, γίνεται η υπόθεση ότι δικαιοσύνη σημαίνει ότι όταν ένας χρήστης έχει περισσότερη «ανάγκη» για χωρητικότητα στην cache θα πάρει και πιο πολλή χωρητικότητα στο NE. Η δικαιοσύνη στο NE σε σχέση με τη συνάρτηση απόκρισης του χρήστη ορίζεται παρακάτω (δες επίσης την εργασία [120]).

Ορισμός 9.1: Ο χρήστης i βρίσκεται σε «μεγαλύτερη ανάγκη» για χωρητικότητα απ' ότι ο χρήστης j , αν $r_i(C') \geq r_j(C')$ για κάθε ποσότητα διαθέσιμου χώρου $C' > 0$. Το NE $\mathbf{y}^* = (y_1^*, \dots, y_N^*)^T$ θεωρείται δίκαιο αν, για κάθε i και j , τέτοια που ο i βρίσκεται σε μεγαλύτερη ανάγκη από τον j , $y_i^* \geq y_j^*$.

Λήμμα 9.1: Για κάθε i , οι συναρτήσεις $r_i(C')$ και $C' - r_i(C')$ είναι μονότονα αύξουσες για $C' > 0$.

Απόδειξη: Όπως φαίνεται και στην απόδειξη στην Πρόταση 9.1, η συνάρτηση απόκρισης του χρήστη i είναι η

$$r_i(C') = \frac{\beta_i}{1 + \beta_i - a_i} C'.$$

Για $0 \leq a_i < 1$, $0 < \beta_i \leq 1$, φαίνεται εύκολα ότι $\beta_i / (1 + \beta_i - a_i) > 0$, επομένως, η συνάρτηση απόκρισης του χρήστη i είναι μονότονα αύξουσα. Φαίνεται, επίσης, ότι

$$C' - r_i(C') = \frac{1 - a_i}{1 + \beta_i - a_i} C',$$

έτσι, η συνάρτηση $C' - r_i(C')$ είναι επίσης μονότονα

αύξουσα. ■

Θεώρημα 9.1: Το NE της κατανομημένης εκχώρησης χωρητικότητας cache που δίνεται από το (P) είναι δίκαιο.

Απόδειξη: Έστω δύο χρήστες i και j ώστε $r_i(C') \geq r_j(C')$ για κάθε διαθέσιμη χωρητικότητα $C' > 0$. Εξ' ορισμού του NE και της συνάρτησης $r_i(\cdot)$, προκύπτει ότι

$$y_i^* = \arg \max_{g \in S_i} u_i(g, g + C - C_i^*) = r_i(C_i^*) \geq r_j(C_i^*).$$

Καθώς $C_i^* = C - \sum_{j \neq i} y_j^* = C - F^* + y_i^*$, δηλαδή,

$y_i^* = C_i^* - C + F^*$, η πιο πάνω ανισότητα μπορεί να πάρει τη μορφή $C_i^* - r_j(C_i^*) \geq C - F^*$. Πρέπει να σημειωθεί ότι $C - F^* > 0$ (Πρόταση 9.2). Από $y_j^* = r_j(C_j^*)$, προκύπτει ότι $C_j^* - r_j(C_j^*) = C - F^*$.

Από την ιδιότητα της μονότονης αύξησης της συνάρτησης $C - r_j(C)$ στο Λήμμα 9.1, $C_i^* \geq C_j^*$, και, επομένως, $y_i^* = F^* - C + C_i^* \geq F^* - C + C_j^* = y_j^*$. ■

9.5 Σύγκλιση στη Nash Ισορροπία

Στο παρόν υποκεφάλαιο, περιγράφεται ένα επαναληπτικό σχήμα Gauss-Seidel για σύγκλιση στο NE. Αυτό το σχήμα υποθέτει ότι οι χρήστες λειτουργούν επαναληπτικά, με ασύγχρονο τρόπο. Μετά από κάθε HTTP αλληλεπίδραση (αίτηση-απάντηση), οι χρήστες ενημερώνουν τη χωρητικότητα που καταλαμβάνουν στο δίσκο, ούτως ώστε να μεγιστοποιήσουν τις συναρτήσεις ωφελείας τους δεδομένης της τρέχουσας κατάστασης τους συστήματος. Αυτή η υπόθεση είναι συμβατή με το εξεταζόμενο σενάριο WWW caching, όπου οι χρήστες μπορούν να ενημερώνουν τις πληροφορίες που διαθέτουν για την κατάσταση του συστήματος, αφού λάβουν μία HTTP απάντηση.

Σε κάθε βήμα του σχήματος Gauss-Seidel (δηλαδή κάθε φορά που πραγματοποιείται μία HTTP αλληλεπίδραση), ο χρήστης ενημερώνεται σε ό,τι αφορά τη συνολική κατάσταση του συστήματος, καθώς και για τη δική του κατάσταση, κι επανυπολογίζει τη βέλτιστη χωρητικότητα που θα πρέπει να καταλαμβάνει στο δίσκο. Μόνο ο χρήστης που αλληλεπιδρά με το διακομιστή caching ενημερώνει τη στρατηγική του. Αυτού του είδους οι αλγόριθμοι ονομάζονται “nonlinear” ή “coordinate descent” [110]. Η βασική ιδέα πίσω από αυτήν την κατηγορία των αλγορίθμων είναι ότι σταθεροποιούνται όλα τα στοιχεία του \mathbf{y} , εκτός της i -οστής συνιστώσας, η οποία ενημερώνεται, ώστε να μεγιστοποιηθεί η $u_i(\mathbf{y})$. Αυτή η διαδικασία επαναλαμβάνεται, ορίζοντας, έτσι, έναν επαναληπτικό αλγόριθμο.

Στον *nonlinear* αλγόριθμο *Gauss-Seidel*, οι μεγιστοποιήσεις εκτελούνται διαδοχικά για κάθε συνιστώσα. Αυστηρά, ο αλγόριθμος μπορεί να περιγραφεί ως εξής:

$$y_i(t+1) = \arg \max_{y_i} u_i(y_1(t+1), \dots, y_{i-1}(t+1), y_i, y_{i+1}(t), \dots, y_N(t)) \quad (9.7)$$

Στην εξίσωση (9.7), υπονοείται ότι η σειρά με την οποία οι χρήστες ενημερώνουν τις στρατηγικές τους είναι προκαθορισμένη. Κάτι τέτοιο όμως δεν είναι υποχρεωτικό. Η σειρά μπορεί να επιλεγεί τυχαία ή και να μεταβάλλεται δυναμικά [110]. Πιο κάτω, αποδεικνύεται ότι, με το σχήμα Gauss-Seidel, το διάνυσμα των χωρητικότητων των χρηστών, στο βήμα t , $\mathbf{y}(t) = (y_1(t), \dots, y_N(t))^T$, συγκλίνει στο μοναδικό NE $\mathbf{y}^* = (y_1^*, \dots, y_N^*)^T$, καθώς $t \rightarrow \infty$. Από την εξίσωση (9.5), είναι εύκολο να δειχθεί ότι ο χρήστης i , στο βήμα $t+1$, δεδομένων των στρατηγικών των άλλων χρηστών απαντά ως εξής:

$$y_i(t+1) = \frac{\beta_i}{1 + \beta_i - a_i} \left(C - \sum_{j \neq i} y_j(t) \right) \quad (9.8)$$

Έστω $F(t) = \sum_{j=1}^N y_j(t)$ η συνολική χωρητικότητα που έχει δεσμευτεί από όλους τους χρήστες μετά και το βήμα t . Ορίζεται, τότε, η επόμενη μετρική:

$$S(\mathbf{y}(t)) = \sum_{j=1}^N |y_j(t) - y_j^*| + |F(t) - F^*|$$

Θα δειχθεί ότι η $S(\mathbf{y}(t))$ συγκλίνει στο 0 (δηλαδή οι ατομικές στρατηγικές συγκλίνουν στο NE).

Λήμμα 9.2: Αν υποτεθεί ότι ο χρήστης i εκτελεί το $(t+1)$ βήμα, τότε

$$F(t+1) \geq F^* \Rightarrow y_i(t+1) \leq y_i^*$$

κι αντίστροφα

$$F(t+1) \leq F^* \Rightarrow y_i(t+1) \geq y_i^*,$$

όπου

$$F^* = \sum_{j=1}^N y_j^* = C \left(\sum_{j=1}^N \frac{\beta_j}{1 - a_j} \right) \left[1 + \sum_{j=1}^N \frac{\beta_j}{1 - a_j} \right]^{-1}.$$

Επιπλέον, τα παραπάνω ισχύουν κι όταν όλες οι ανισότητες είναι αυστηρές.

Απόδειξη: Από την (9.8), η συνολική χωρητικότητα που δεσμεύεται από όλους τους χρήστες μετά το βήμα $(t+1)$ θα είναι

$$F(t+1) = \sum_{j \neq i} y_j(t) + \frac{\beta_i}{1 + \beta_i - a_i} \left(C - \sum_{j \neq i} y_j(t) \right)$$

ή

$$F(t+1) = \frac{\beta_i}{1 + \beta_i - a_i} C + \frac{1 - a_i}{1 + \beta_i - a_i} \sum_{j \neq i} y_j(t).$$

Αν υποθεθεί ότι $F(t+1) \geq F^*$, υπονοείται ότι

$$\sum_{j \neq i} y_j(t) \geq \frac{1 + \beta_i - a_i}{1 - a_i} C \left(\left[\sum_{j=1}^N \frac{\beta_j}{1 - a_j} \right] \left[1 + \sum_{j=1}^N \frac{\beta_j}{1 - a_j} \right]^{-1} - \frac{\beta_i}{1 + \beta_i - a_i} \right)$$

οπότε προκύπτει ότι

$$y_i(t+1) = \frac{\beta_i}{1 + \beta_i - a_i} \left(C - \sum_{j \neq i} y_j(t) \right) \leq \frac{\beta_i}{1 - a_i} C \left[1 + \sum_{j=1}^N \frac{\beta_j}{1 - a_j} \right]^{-1} = y_i^*$$

Παρομοίως, μπορεί να αποδειχθεί ότι $y_i(t+1) \geq y_i^*$, αν $F(t+1) \leq F^*$. Για τις περιπτώσεις αυστηρών ανισοτήτων, η απόδειξη παραμένει η ίδια. ■

Το επόμενο λήμμα δηλώνει ότι η ακολουθία $S(\mathbf{y}(t))$, $t \in \mathbb{N}$, είναι μη αύξουσα [121].

Λήμμα 9.3: Με το επαναληπτικό σχήμα Gauss-Seidel

$$S(\mathbf{y}(t+1)) \leq S(\mathbf{y}(t)), \text{ για όλα τα } t \in \mathbb{N}.$$

Το επόμενο λήμμα αναφέρει ότι η ακολουθία $S(\mathbf{y}(t))$ φθίνει αυστηρά μετά από ένα πεπερασμένο πλήθος βημάτων [121].

Λήμμα 9.4: Για όλα τα t υπάρχει ένα (πεπερασμένο) $k > t$, τέτοιο που, αν $(y_1(t), \dots, y_N(t)) \neq (y_1^*, \dots, y_N^*)$, τότε $S(\mathbf{y}(k)) < S(\mathbf{y}(t))$.

Η σύγκλιση του αλγορίθμου Gauss-Seidel στο μοναδικό NE υποστηρίζεται από το επόμενο θεώρημα [121].

Θεώρημα 9.2: Με το σχήμα Gauss-Seidel, το διάνυσμα των χωρητικότητας που έχουν δεσμευθεί από όλους τους χρήστες στο βήμα t , $\mathbf{y}(t) = (y_1(t), \dots, y_N(t))^T$, συγκλίνει στο μοναδικό NE (y_1^*, \dots, y_N^*) καθώς $t \rightarrow \infty$, δηλαδή $\lim_{t \rightarrow \infty} \mathbf{y}(t) = \mathbf{y}^*$.

9.6 Υλοποίηση Κατανεμημένου Αλγορίθμου

Στο προηγούμενο υποκεφάλαιο, αποδείχθηκε ότι το nonlinear, επαναληπτικό σχήμα Gauss-Seidel συγκλίνει στο NE. Αυτό το σχήμα, όμως, δεν είναι δυνατόν να εφαρμοστεί απaráλλακτο το παίγνιο caching, λόγω των επόμενων λόγων: 1) ένας χρήστης πραγματοποιώντας μία ατομική HTTP αλληλεπίδραση μπορεί μόνο να ζητά (ή να μη ζητά) την αποθήκευση του αντικειμένου που αιτήθηκε (δηλαδή ο χρήστης μπορεί να αυξήσει τη χωρητικότητα που καταλαμβάνει κατά ένα αντικείμενο τη φορά) και 2) οι χρήστες μπορούν μόνο να αυξήσουν τη χωρητικότητα τους, καθώς αν ξεπεράσουν τη χωρητικότητα ισορροπίας, δεν έχουν τη δυνατότητα να «συρρικνώσουν» την περιοχή που καταλαμβάνουν στο δίσκο, εκτός αν ο δίσκος γεμίσει κι εκτελεστεί ο αλγόριθμος αντικατάστασης (π.χ. Least Recently Used – LRU). Για αυτό το λόγο, προτείνεται μία παραλλαγή του σχήματος που εξετάστηκε στο προηγούμενο υποκεφάλαιο, η οποία λαμβάνει υπόψη τους προαναφερθέντες τεχνικούς περιορισμούς.

9.6.1 Παραλλαγή Αλγορίθμου

Ως ενδεικτικό παράδειγμα, έστω ο χρήστης i , στο βήμα t , χρησιμοποιεί τη συνάρτηση απόκρισης της εξίσωσης (9.8), με την πληροφορία που παρέχεται από την προηγούμενη HTTP απάντηση (ελεύθερος χώρος στο δίσκο $C - F$, δείκτης δημοτικότητας a_i και δεσμευμένος χώρος στο δίσκο y_i), ώστε να ανανεώσει τη χωρητικότητα που καταλαμβάνει. Αν η συνάρτηση απόκρισης υποδεικνύει ότι ο χρήστης θα πρέπει να απαντήσει με δεσμευμένη χωρητικότητα που ξεπερνά την τρέχουσα δεσμευμένη χωρητικότητα, ο χρήστης μπορεί μόνο να αυξήσει τη χωρητικότητά του στο δίσκο κατά το μέγεθος του ζητούμενου αντικειμένου, δηλαδή δε μπορεί να απαντήσει «πλήρως» στις στρατηγικές των άλλων χρηστών. Αντίστροφα, αν η συνάρτηση απόκρισης συμβουλεύει το χρήστη να ελαττώσει το χώρο που καταλαμβάνει, ο χρήστης μπορεί μόνο να δώσει εντολή στο διακομιστή caching να μην αποθηκεύσει το ζητούμενο αντικείμενο, δηλαδή να διατηρήσει τη χωρητικότητά του στο δίσκο σταθερή.

Βάσει των παραπάνω, προτείνεται μία παραλλαγή του επαναληπτικού nonlinear Gauss-Seidel σχήματος. Έστω ένα αρχικό διάνυσμα $\mathbf{y}(0)$ χωρητικοτήτων που είναι δεσμευμένες από τους χρήστες, π.χ. $\mathbf{y}(0) = (0, \dots, 0)^T$. Τότε, ο χρήστης i , στο βήμα

$t+1$, δεδομένων των πληροφοριών από την τελευταία HTTP απάντηση, υπολογίζει τη καλύτερη απόκριση $r_i(t+1) = \beta_i(1 + \beta_i - a_i)^{-1}(C - F(t) + y_i(t))$, σε σχέση με την κατάσταση του συστήματος. Στη συνέχεια, αν $r_i(t+1) > y_i(t)$, ο χρήστης δίνει την εντολή στο διακομιστή caching να αποθηκεύσει το ζητούμενο αντικείμενο (αν δεν είναι ήδη αποθηκευμένο). Διαφορετικά, ο χρήστης δίνει την εντολή στο διακομιστή caching απλά να φέρει το αντικείμενο, αλλά να μην το αποθηκεύσει. Ο διακομιστής caching επιστρέφει το ζητούμενο αντικείμενο κι ενημερώνει το χρήστη σε ό,τι αφορά την τρέχουσα κατάσταση του συστήματος.

9.6.2 Ζητήματα Υλοποίησης

Σε αυτό το υποκεφάλαιο, μελετάται πως μπορεί να υλοποιηθεί ο προτεινόμενος ασύγχρονος καταναμημένος αλγόριθμος. Για κάθε αποθηκευμένο αντικείμενο, ο διακομιστής caching διατηρεί τις ακόλουθες μετα-πληροφορίες: 1) το ID του χρήστη, λόγω του οποίου αποθηκεύτηκε το αντικείμενο, 2) το πλήθος των αιτήσεων για το συγκεκριμένο αντικείμενο προερχόμενες από άλλους χρήστες και 3) το μέγεθος του αντικειμένου. Ένα αποθηκευμένο αντικείμενο μπορεί να αναπαρασταθεί ως μία εγγραφή, όπως φαίνεται στην Εικόνα 9.1.

data	URL	timestamp	clientID	hits	size
------	-----	-----------	----------	------	------

Εικόνα 9.1. Δομή αποθηκευμένου αντικειμένου στην cache

Χρειάζεται επίσης να διατηρείται πληροφορία σχετικά με κάθε χρήστη του διακομιστή caching. Συγκεκριμένα, ο διακομιστής caching πρέπει να είναι ενήμερος σε ό,τι αφορά τη χωρητικότητα που καταλαμβάνει κάθε χρήστης και τον παράγοντα δημοτικότητας (δηλαδή τον αριθμητή του δείκτη δημοτικότητας στην εξίσωση (9.3)). Έτσι, ο διακομιστής caching κρατάει για κάθε χρήστη μία εγγραφή όπως αυτή που φαίνεται στην Εικόνα 9.2.

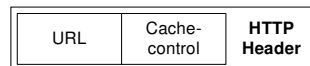
clientID	space	popularity
----------	-------	------------

Εικόνα 9.2. Δομή εγγραφής χρήστη

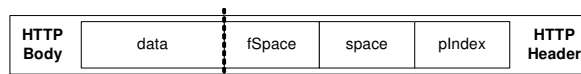
Γίνεται επίσης η υπόθεση ότι ο διακομιστής caching είναι ικανός να εκτελεί τις ακόλουθες τρεις λειτουργίες:

- 1) αναζήτηση αντικειμένου στην cache με βάση το URL του,
- 2) αναζήτηση για το λιγότερο πρόσφατα χρησιμοποιημένο αντικείμενο στην cache και
- 3) αναζήτηση εγγραφής χρήστη βάσει του ID του.

Πριν να παρουσιαστούν οι αλγόριθμοι που εκτελούνται από το διακομιστή caching και τους χρήστες, ορίζεται η δομή των μηνυμάτων αιτήσεων κι απαντήσεων. Η Εικόνα 9.3 δείχνει τη δομή του μηνύματος αίτησης κι η Εικόνα 9.4 του μηνύματος απάντησης.



Εικόνα 9.3. Δομή μηνύματος αίτησης



Εικόνα 9.4. Δομή μηνύματος απάντησης

Στην Εικόνα 9.4, το *fSpace* υποδηλώνει τον ελεύθερο χώρο στην cache, το *space* το χώρο που καταλαμβάνει ο χρήστης που έστειλε την αίτηση και το *pIndex* το δείκτη δημοτικότητας του συγκεκριμένου χρήστη. Όλα αυτά τα παιδιά των επικεφαλίδων μπορούν να ενσωματωθούν στο μήνυμα της απάντησης μέσω του HTTP extension framework [124].

Εν συνεχεία, παρουσιάζονται οι αλγόριθμοι που εκτελούνται από το διακομιστή caching. Οι ενέργειες του διακομιστή caching, όταν καλείται ο αλγόριθμος αντικατάστασης, παρουσιάζονται στην Εικόνα 9.5.

```

Function invokeLRU()
    // Specify the target free space
    targetFreeSpace = LRUFactor*C

    do
        // find the least recently used object
        LRU_obj = LRUObjectLookup()
        s = LRU_obj->size
        client_ID = LRU_obj->clientID
        hits_num = LRU_obj->hits
        client = clientLookup(client_ID)
        // update the client's popularity
        client->popularity -= hits_num*s
        // update the client's occupied disk space
        client->space -= s

        cacheFreeSpace += s
        remove(LRU_obj)
    while (cacheFreeSpace < targetFreeSpace)

endFunction

```

Εικόνα 9.5. Αλγόριθμος αντικατάστασης

Όταν ληφθεί μία αίτηση από έναν χρήστη, ο διακομιστής caching ακολουθεί τα βήματα που περιγράφονται στον ψευδοκώδικα στην Εικόνα 9.6.

```

Event receiveRequest (req)
    requestorID = getClientID(req)
    url = req->URL
    cache_instruction = req->cache_control
    requested_obj = cacheLookup(url)
    //object found in cache
    if (requested_obj != NULL) then
        requested_obj->hits++
        client_ID = requested_obj->clientID
        client = clientLookup(client_ID)
        s = requested_obj->size
        client->popularity += s
    else //object not found in cache
        requested_obj = fetchObject(URL)
        // client instructs for caching
        if (cache_instruction == TRUE) then
            requested_obj->clientID = requestorID
            requested_obj->hits = 0
            s = requested_obj->size
            // No adequate free disk space
            if (cacheFreeSpace - s < 0) then
                invokeLRU()
            endif
            cacheObject(requested_obj)
            // store the object into the cache
            client = clientLookup(requestorID)
            client->space += s
        endif
    endif
    reply->data= requested_obj->data
    reply->fSpace = cacheFreeSpace
    client = clientLookup(requestorID)
    reply->space = client->space
    reply->pIndex = client->popularity/requestOverlapping
    sendReply(requestorID,reply)
endEvent

```

Εικόνα 9.6. Χειρισμός αίτησης στο διακομιστή caching

9.7 Αποτελέσματα Προσομοίωσης

Η επίδραση του προτεινόμενου παιγνιοθεωρητικού μηχανισμού αξιολογήθηκε σε ένα περιβάλλον όπου ένα πλήθος χρηστών αλληλεπιδρά με ένα διακομιστή caching. Προσομοιώθηκε, επίσης, και το μη-παιγνιοθεωρητικό σενάριο. Οι προσομοιώσεις βασίστηκαν σε πραγματικά δεδομένα αιτήσεων που έχουν καταγραφεί από τους διακομιστές της DEC [122]. Ο Πίνακας 9.1 συνοψίζει τα χαρακτηριστικά των δεδομένων που χρησιμοποιήθηκαν, αλλά και του συστήματος caching που προσομοιώθηκε.

Πίνακας 9.1. Λίστα παραμέτρων προσομοίωσης

Πλήθος Χρηστών	10,366
Πλήθος Αιτήσεων	4,431,200
Αλγόριθμος Αντικατάστασης	LRU
Χωρητικότητα Cache	128MB, 256MB, 512MB, 1GB, 2.5GB, 5GB, 10GB

Οι μετρικές που καταγράφηκαν στη διάρκεια των προσομοιώσεων ήταν οι ακόλουθες: 1) η συνολική παρατηρούμενη πιθανότητα ευστοχίας (HR), 2) η πιθανότητα ευστοχίας που πέτυχε κάθε χρήστης i (HR_i), 3) ο συντελεστής διασποράς των HR_i , 4) το πλήθος των κλήσεων του μηχανισμού αντικατάστασης LRU και 5) η μέση χρήση της χωρητικότητας της cache. Η πιθανότητα ευστοχίας της cache (HR) υπολογίζεται ως

$$HR = \frac{\text{Number of Hits}}{\text{Number of Requests}},$$

ενώ η πιθανότητα ευστοχίας HR_i για το χρήστη i υπολογίζεται ως εξής:

$$HR_i = \frac{\text{Number of Hits of Client } i}{\text{Number of Requests of Client } i}$$

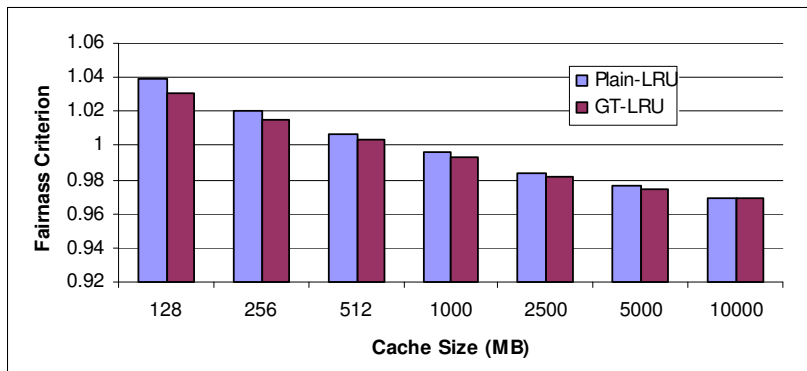
Ως δείκτης της δικαιοσύνης που επιτυγχάνει το εξεταζόμενο σχήμα, υιοθετήθηκε ο συντελεστής διασποράς των μετρήσεων HR_i . Υψηλές τιμές αυτού του κριτηρίου δικαιοσύνης (fairness criterion – FC) υποδηλώνει ότι οι χρήστες δεν απολαμβάνουν τα ίδια οφέλη από το caching, αλλά ότι κάποιοι μονοπωλούν το

χώρο στο δίσκο. Αντίστροφα, μία μικρή τιμή του FC υπονοεί ότι η συμπεριφορά που επιδεικνύουν οι περισσότεροι χρήστες είναι παρόμοια κι ότι όλοι απολαμβάνουν περίπου την ίδια πιθανότητα ευστοχίας. Συγκεκριμένα, η μετρική δικαιοσύνης ορίζεται ως

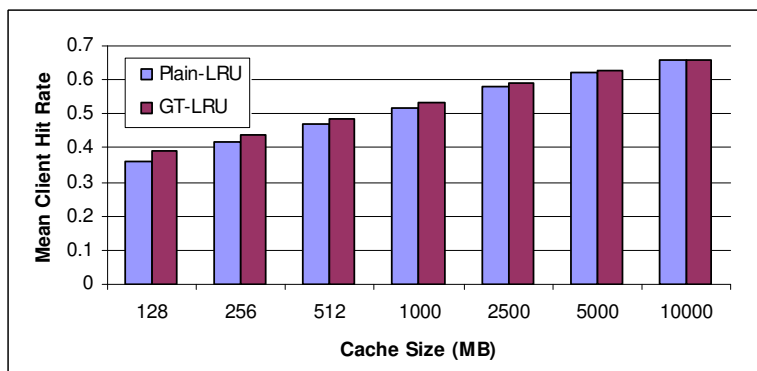
$$FC = \frac{1}{\bar{H}} \sqrt{\frac{1}{N-1} \sum_{i=1}^N (HR_i - \bar{H})^2}, \quad (9.9)$$

όπου το \bar{H} υποδηλώνει την μέση πιθανότητα ευστοχίας των χρηστών κι υπολογίζεται ως εξής:

$$\bar{H} = \frac{1}{N} \sum_{i=1}^N HR_i$$



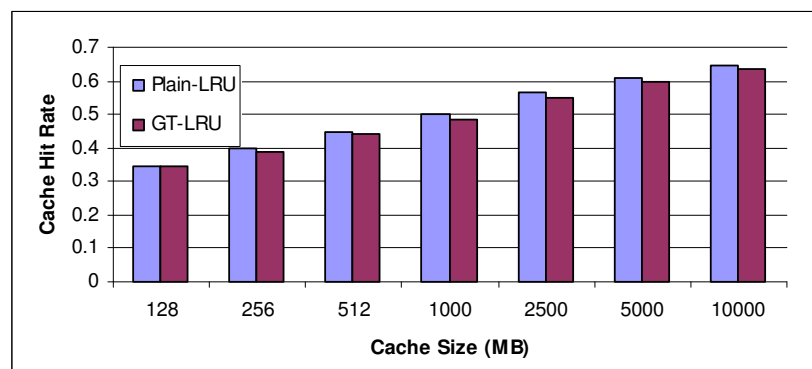
Εικόνα 9.7. Κριτήριο δικαιοσύνης



Εικόνα 9.8. Μέση πιθανότητα ευστοχίας των χρηστών

Όπως φαίνεται στην Εικόνα 9.7, ο παιγνιοθεωρητικός μηχανισμός επιτυγχάνει μικρότερες τιμές FC. Επιπλέον, όπως φαίνεται στην Εικόνα 9.8, ο προτεινόμενος

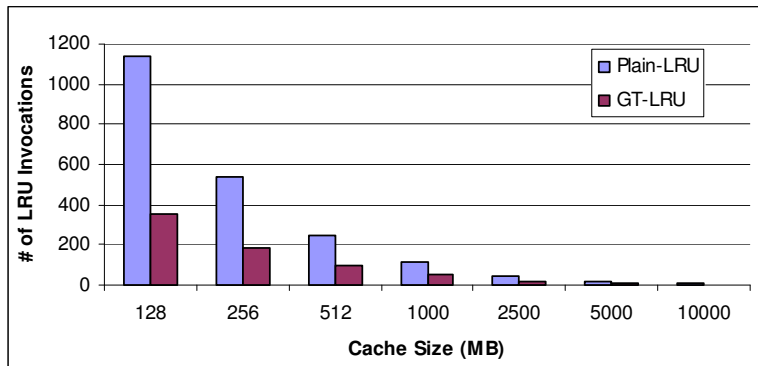
μηχανισμός επιτυγχάνει μια αύξηση στη μέση πιθανότητα ευστοχίας των χρηστών \bar{H} , σε σχέση με το απλό LRU σχήμα. Αυτή η αύξηση μπορεί να αποδοθεί στο γεγονός ότι, στην παιγνιοθεωρητική περίπτωση, οι «ιδιαίτερα επιθετικοί» χρήστες, που είναι σχετικά λίγοι, σταματούν να ζητούν caching, βάσει του προτεινόμενου μηχανισμού, παύοντας έτσι να μονοπωλούν το χώρο στο δίσκο. Κατά συνέπεια, οι «λιγότερο επιθετικοί» χρήστες που αποτελούν και την πλειοψηφία, επωφελούνται από τη διαθέσιμη χωρητικότητα στο δίσκο κι επιτυγχάνουν, έτσι, αυξημένη πιθανότητα ευστοχίας. Όπως αναμενόταν, όμως, η συνολική πιθανότητα ευστοχίας (HR) μειώνεται, όπως φαίνεται στην Εικόνα 9.9.



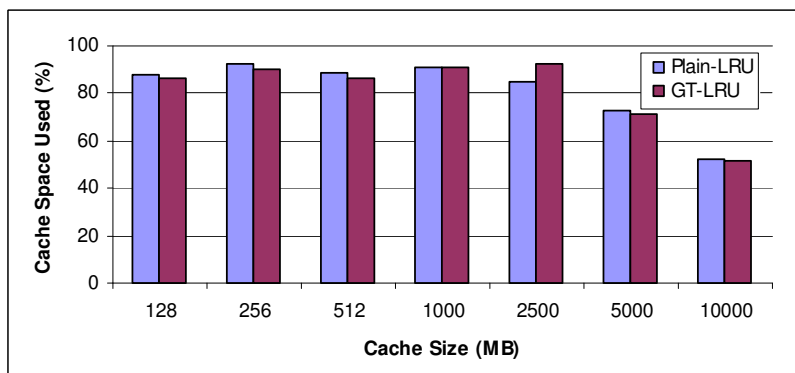
Εικόνα 9.9. Πιθανότητα ευστοχίας

Ο προτεινόμενος μηχανισμός πέτυχε, επίσης, να μειώσει δραματικά το πλήθος των κλήσεων του αλγορίθμου αντικατάστασης LRU, όπως φαίνεται στην Εικόνα 9.10. Συγκεκριμένα, το πλήθος των αντικαταστάσεων στο παιγνιοθεωρητικό σενάριο μειώθηκε σημαντικά, σε σχέση με την απλή περίπτωση, ειδικά για μικρό μέγεθος cache. Σε κάθε κλήση, το σχήμα αντικατάστασης απομάκρυνε τα λιγότερο πρόσφατα χρησιμοποιημένα αντικείμενα της cache με σκοπό να ελευθερώσει το 20% της εκχωρημένης χωρητικότητας.

Ποσοτικοποιήθηκε, επίσης, η επίδραση στη μέση χρήση του δίσκου (χρονική μέση τιμή) της εφαρμογής του παιγνιοθεωρητικού σχήματος. Η Εικόνα 9.11 δείχνει τα σχετικά ευρήματα. Γενικά, το επίπεδο της μέσης χρήσης της χωρητικότητας του δίσκου είναι πιο χαμηλό για το παιγνιοθεωρητικό σχήμα, σε σχέση με τον απλό μηχανισμό. Αυτό είναι αναμενόμενο, καθώς οι χρήστες δεν ζητούν την αποθήκευση των αντικειμένων αλόγιστα κι ο δίσκος της cache γεμίζει με πιο αργό ρυθμό.



Εικόνα 9.10. Πλήθος κλήσεων LRU



Εικόνα 9.11. Μέση χρήση του δίσκου της cache

9.8 Συμπεράσματα και Μελλοντική Εργασία

Στα πλαίσια του web caching, οι χρήστες, συνήθως, καταλαμβάνοντας περισσότερο χώρο στο δίσκο να βελτιώνουν την απόδοσή τους (πιθανότητα ευστοχίας). Καθώς, όμως, ο χώρος του δίσκου μίας cache αποτελεί έναν πεπερασμένο πόρο, αναδεικνύεται ένα πρόβλημα κοινωνικής αλληλεπίδρασης. Στην παρούσα μελέτη, το πρόβλημα μοντελοποιήθηκε ως ένα μη-συνεργατικό παίγνιο κι εξετάστηκε η ύπαρξη ενός ΝΕ. Ο προτεινόμενος παιγνιοθεωρητικός μηχανισμός αξιολογήθηκε βάσει προσομοιώσεων. Από τα αποτελέσματα που προέκυψαν φάνηκε μία αξιόλογη βελτίωση στη μετρική δικαιοσύνης. Η απόδοση που επιτυγχάνουν οι χρήστες είναι συγκρίσιμη και πιο προβλέψιμη, σε σύγκριση με την απλή περίπτωση. Με το παιγνιοθεωρητικό σχήμα, το πλήθος των αντικαταστάσεων μειώθηκε σημαντικά και η μέση πιθανότητα ευστοχίας των

χρηστών αυξήθηκε, ενώ παρατηρήθηκε μία σχετικά περιορισμένη μείωση στη συνολική πιθανότητα ευστοχίας.

Μελλοντικά, θα ήταν επιθυμητό να αξιολογηθεί η επίδραση παρόμοιων μηχανισμών σε συνεργατικό caching (π.χ. ιεραρχίες caching) καθώς και σε pre-fetching web περιεχομένου.

ΚΕΦΑΛΑΙΟ 10

ΟΙΚΟΝΟΜΙΚΟ ΠΛΑΙΣΙΟ ΓΙΑ ΤΗ ΔΙΑΧΕΙΡΙΣΗ ΣΥΣΤΗΜΑΤΩΝ CACHING

10.1 Εισαγωγή

Σε αυτό το κεφάλαιο, προτείνεται ένα Οικονομικό πλαίσιο για τη διαχείριση της υπηρεσίας caching σε ιεραρχικό / συνεργατικό περιβάλλον. Η χρήση ιεραρχικών / συνεργατικών δομών caching, εν γένει, αυξάνει το ποσό των πόρων για τους κόμβους που ανήκουν στην κοινότητα που απαρτίζει την εν λόγω δομή, βελτιώνοντας, έτσι, την ατομική και συλλογική επίδοσή τους. Παρόλ' αυτά, επιτρέπει την εμφάνιση προβλημάτων «κακομεταχείρισης», λόγω της έκθεσης των πόρων του ενός κόμβου στον άλλο [6]. Όταν η χωρητικότητα ενός κόμβου διαμοιράζεται μεταξύ διάφορων άλλων κόμβων (π.χ. μίας L2 cache από κάποιες L1 cache), είναι δυνατόν, αν ένας κόμβος είναι ιδιαίτερα «επιθετικός», σε σχέση με τους άλλους, να κυριαρχήσει επί της διαμοιραζόμενης χωρητικότητας, με αποτέλεσμα, να επωφεληθεί εις βάρος των άλλων κόμβων. Τέτοια προβλήματα μπορούν να παρατηρηθούν, εκτός από συστήματα caching, και σε άλλα περιβάλλοντα, όπως δίκτυα P2P, και CDN.

Όταν ο διαμοιρασμός ενός κοινού πόρου είναι προβληματικός, επειδή κάποιιοι μπορούν να τον καταχραστούν αδιαφορώντας για τους υπόλοιπους ενδιαφερόμενους, είναι ανάγκη να εισαχθεί ένα πλαίσιο διαχείρισης που να θέτει περιορισμούς και να οδηγεί σε αποτελεσματική διαχείριση των πόρων. Ένας τρόπος να επιτευχθεί κάτι τέτοιο είναι η εγκαθίδρυση μίας οικονομίας, σύμφωνα με την οποία κάποια (ή κάποιες) οντότητα κατέχει το διαμοιραζόμενο πόρο κι οι υπόλοιπες οντότητες, ως ενδιαφερόμενοι, προσπαθούν να αποκτήσουν ένα μέρος του. Σε μία τέτοια οικονομία υπάρχουν σαφείς κανόνες που διέπουν τις αλληλεπιδράσεις των οντοτήτων, καθώς και τιμές που οδηγούν σε διαφορετικά επίπεδα ζήτησης. Αναλόγως με το σενάριο, μπορούν να εφαρμοστούν διαφορετικά σχήματα αγορών, αλλά και τρόποι χρεώσεων.

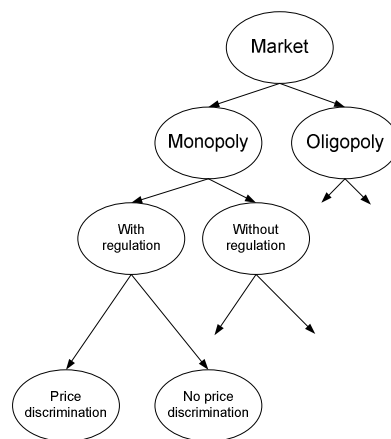
Στην παρούσα μελέτη, προτείνεται ένα οικονομικό σχήμα για τη διαχείριση της χωρητικότητας, για εφαρμογή σε ιεραρχίες caching. Στο μελετούμενο περιβάλλον, το οικονομικό αγαθό είναι ο αποθηκευτικός χώρος. Οι cache επιπέδου- v (πωλητές) παρέχουν χώρο από το δίσκο τους, ενώ οι cache επιπέδου- $(v-1)$ (αγοραστές) πληρώνουν για να αποκτήσουν μέρος του εν λόγω χώρου. Το μοντέλο που προτείνεται, στα πλαίσια της παρούσας μελέτης περιορίζεται στο δεύτερο επίπεδο caching ($v = 2$). Η γενίκευση του σε υψηλότερα επίπεδα αποτελεί θέμα μελλοντικής εργασίας. Ορίζεται μία μονοπωλιακή αγορά και βάσει αυτής προτείνονται διαφορετικά σχήματα εκχώρησης της χωρητικότητας: 1) βάσει χρέωσης με διάκριση, 2) βάσει χρέωσης χωρίς διάκριση και 3) με σκοπό τη μεγιστοποίηση της κοινωνικής ευημερίας. Για κάθε ένα από τα διαφορετικά σχήματα, υπολογίζονται αναλυτικά οι εκχωρήσεις χωρητικότητας που προκύπτουν από την εφαρμογή τους, ή, αν κάτι τέτοιο δεν είναι εφικτό, αναπτύσσονται κατάλληλοι αριθμητικοί αλγόριθμοι. Τα διαφορετικά σχήματα αξιολογούνται μέσω αριθμητικών παραδειγμάτων.

10.2 Εισαγωγή σε Οικονομικές Έννοιες

Με τον όρο αγορά ή οικονομία περιγράφεται ένα σύστημα που αποτελείται από πωλητές, αγοραστές και πιθανώς ένα ρυθμιστή (regulator). Οι πωλητές ανταλλάσσουν με τους αγοραστές τα οικονομικά αγαθά που κατέχουν με χρήματα. Σε μία ελεύθερη αγορά, ο τύπος της χρέωσης (κι οι τιμές αυτές καθαυτές) μπορεί να διαμορφωθεί από την αλληλεπίδραση μεταξύ πωλητών κι αγοραστών. Εναλλακτικά, ένας ρυθμιστής (δηλαδή ένας τρίτος, όπως το κράτος σε μία πραγματική οικονομία) μπορεί να επηρεάσει τις αποφάσεις των προαναφερθέντων οντοτήτων. Συγκεκριμένα, ένας ρυθμιστής μπορεί να επιβάλλει διάφορους περιορισμούς, συνήθως στους πωλητές, ούτως ώστε να αυξήσει την κοινωνική ευημερία ή να επιβάλλει κοινωνική δικαιοσύνη.

Ανάλογα με τους ρόλους των εμπλεκόμενων οντοτήτων και της δύναμης που μπορούν να διαθέτουν είναι δυνατόν να προκύψουν διαφορετικοί τύποι οικονομιών. Έστω, για λόγους απλότητας, ότι υπάρχει ένα μοναδικό αγαθό στην αγορά. Κατά πρώτον, είναι δυνατόν να γίνει διάκριση μεταξύ ενός μονοπωλίου (monopoly) κι ενός ολιγοπωλίου (oligopoly). Στη μεν περίπτωση, υπάρχει μόνον

έναν πάροχο που πωλεί το αγαθό στους αγοραστές και θέτει την τιμή. Στη δεύτερη περίπτωση, υπάρχουν περισσότεροι του ενός πωλητές. Οι πωλητές ρυθμίζουν την τιμή του αγαθού, με σκοπό να προσελκύσουν περισσότερους αγοραστές και να μεγιστοποιήσουν τα έσοδα τους (revenue). Ανεξάρτητα από τις δύο παραπάνω περιπτώσεις, ένας πωλητής μπορεί (ή δε μπορεί) να χρεώνει κάθε αγοραστή διαφορετικά, στην οποία περίπτωση ο πωλητής ασκεί *χρέωση με διάκριση*. Ο πωλητής θέτει την τιμή του αγαθού, λαμβάνοντας υπόψη τα χαρακτηριστικά των αγοραστών, με σκοπό τη μεγιστοποίηση των εσόδων του. Στην περίπτωση, όμως, που υπάρχει ένας ρυθμιστής, ο πωλητής μπορεί να μην επιτρέπεται να θέτει οποιαδήποτε τιμή επιθυμεί σε οποιοδήποτε αγοραστή. Για παράδειγμα, μπορεί να μην επιτρέπεται να ασκήσει χρέωση με διάκριση, αλλά να του επιβάλλεται να αντιμετωπίζει όλους τους χρήστες ομοιόμορφα. Στην Εικόνα 10.1, φαίνεται η διάκριση μεταξύ διαφόρων τύπων αγορών.



Εικόνα 10.1. Διαφορετικοί τύποι αγορών

Ανεξάρτητα από τον τύπο της αγοράς, οι αποφάσεις των αλληλεπιδρώντων πωλητών κι αγοραστών καθοδηγούνται, συνήθως, από διάφορα κίνητρα. Τα εν λόγω κίνητρα αναπαρίστανται με μαθηματικό τρόπο, μέσω συναρτήσεων ωφέλειας (ή κόστους) που ποσοτικοποιούν το όφελος (κόστος) για τον αγοραστή (πωλητή) για τη χρήση (παροχή) ενός ποσού του αγαθού. Η χρέωση των αγαθών εισάγεται για να επιτευχθεί η *αποτελεσματική* χρήση τους. Το όφελος του αγοραστή μειώνεται λόγω της χρέωσης που υφίσταται, ενώ ο πωλητής συλλέγοντας τις πληρωμές μειώνει το κόστος παροχής του αγαθού. Σύμφωνα με την Οικονομική ορολογία, η ικανοποίηση ενός αγοραστή, ο οποίος αγοράζει ένα ποσό ενός

αγαθού, ποσοτικοποιείται από το καθαρό όφελος (net benefit), δηλαδή την ωφέλεια που αποκομίζει έχοντας το αγαθό μείον τη χρέωση για να το αποκτήσει. Επομένως, η χρέωση είναι ένας μηχανισμός που έχει τη δυνατότητα να επηρεάσει τη συμπεριφορά του αγοραστή και την κατάσταση όλης της αγοράς. Χαμηλές τιμές αυξάνουν τη ζήτηση του αγοραστή, ενώ υψηλές την περιορίζουν. Τα έσοδα του πωλητή επηρεάζονται κι από τη ζήτηση που επιδεικνύουν οι αγοραστές, αλλά κι από την τιμή του αγαθού, παράγοντες που είναι αμφίδρομα συσχετισμένοι.

Έστω $u_i: S_i \rightarrow \mathfrak{R}$ η συνάρτηση ωφελείας του αγοραστή i (συνήθως μία αύξουσα και κυρτή συνάρτηση) που ορίζεται στο σύνολο S_i (π.χ. $[0, Y_{max}]$). Η συνάρτηση u_i ποσοτικοποιεί την ικανοποίηση του αγοραστή λόγω της κατοχής ενός συγκεκριμένου ποσού του αγαθού. Γίνεται η υπόθεση γραμμικής χρέωσης, δηλαδή ο αγοραστής χρεώνεται αναλογικά του ποσού y_i του αγαθού που παίρνει και δεν εφαρμόζονται πάγιες χρεώσεις. Έστω p_i η τιμή ανά μονάδα του αγαθού που πληρώνει ο αγοραστής i . Είναι εύκολο να φανεί ότι, δεδομένης της τιμής p_i , ο αγοραστής «απαντά» με ζήτηση $y_i(p_i)$ που είναι η λύση του ακόλουθου προβλήματος βελτιστοποίησης:

$$\text{(BUYER)} \quad \max_{y_i \in S_i} u_i(y_i) - p_i y_i$$

Από τη συνθήκη βελτιστότητας πρώτου βαθμού προκύπτει ότι

$$p_i = \frac{\partial u_i(y_i)}{\partial y_i}. \quad (10.1)$$

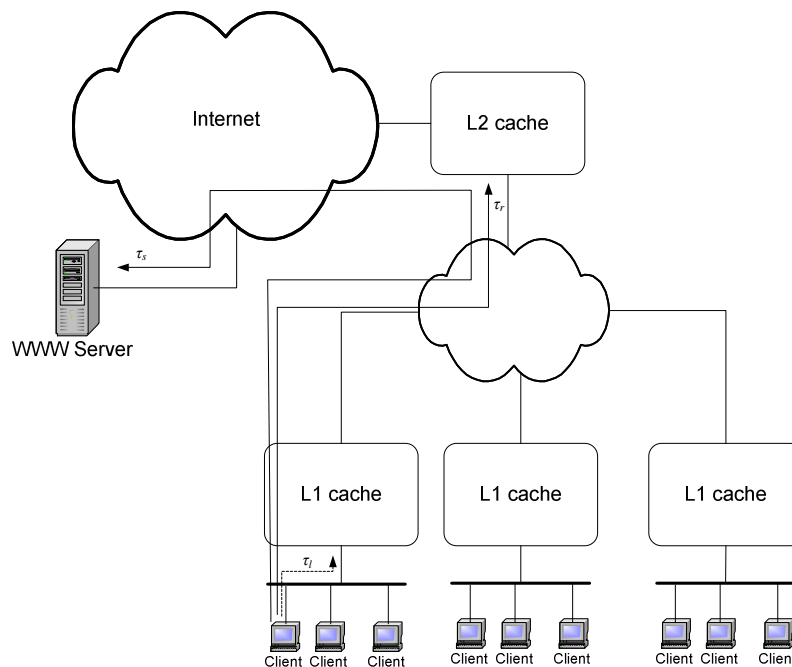
Αν η συνάρτηση ωφελείας u_i είναι κυρτή, η συνθήκη βελτιστότητας (10.1), εκτός από αναγκαία, είναι κι ικανή, που σημαίνει ότι υπάρχει ένα μόνο σημείο μεγιστοποίησης $y_i(p_i)$. Με άλλα λόγια, για κάθε τιμή p_i , ο αγοραστής απαντά με μία ζήτηση $y_i(p_i)$ που μεγιστοποιεί το όφελος του.

Με παρόμοιο τρόπο, ο πωλητής μπορεί να έχει κάποια ελευθερία στο να καθορίζει παραμέτρους της αγοράς. Για παράδειγμα, ο πωλητής μπορεί να είναι σε θέση να επιλέξει την τιμή του αγαθού, ή να υπολογίσει διαφορετικές τιμές για κάθε αγοραστή, με σκοπό τη μεγιστοποίηση των εσόδων του. Ανάλογα με τον τύπο της αγοράς (δες Εικόνα 10.1), ο πωλητής έχει διαφορετικές επιλογές, οι οποίες, στα πλαίσια της παρούσας μελέτης, θα εξετασθούν ξεχωριστά. Συγκεκριμένα, θεωρώντας ένα μονοπωλιακό περιβάλλον (δηλαδή έναν μόνο πωλητή – cache

επιπέδου-2), εξετάζεται η περίπτωση χρέωσης με διάκριση, αλλά κι ομοιόμορφης χρέωσης. Εξετάζεται, επίσης, η περίπτωση που ένας ρυθμιστής επιβάλλει στην cache επιπέδου-2 να διαμοιράσει τη χωρητικότητά της, ώστε να μεγιστοποιηθεί το κοινωνικό όφελος (δηλαδή το άθροισμα των ωφελειών των εμπλεκόμενων στην αγορά).

10.3 Μοντέλο Συστήματος

Στο παρόν υποκεφάλαιο, περιγράφεται το μοντέλο του συστήματος που υιοθετήθηκε, στα πλαίσια της παρούσας μελέτης. Έστω $I = \{1, \dots, N\}$ το σύνολο των L1 cache (αγοραστής) που πληρώνουν για επιπλέον αποθηκευτικό χώρο από την L2 cache. Έστω O το σύνολο των αντικειμένων που περιλαμβάνει ο Ιστός, και $O_i \subseteq O$ το σύνολο των αντικειμένων στα οποία αναφέρονται οι χρήστες της L1 cache i ²².



Εικόνα 10.2. Συνήθης ιεραρχική αρχιτεκτονική caching

Ένα αντικείμενο που αιτείται ένας χρήστης μπορεί να βρεθεί στην L1 cache i , στην οποία περίπτωση θεωρείται ότι επιτυγχάνεται μία *τοπική ευστοχία* (local hit), με αποτέλεσμα ο χρόνος εξυπηρέτησης να είναι $\tau_{l,i}$. Αν το αντικείμενο δεν υπάρχει

²² Γίνεται η υπόθεση ότι μία L1 cache εξυπηρετεί χρήστες που ενδιαφέρονται για ένα υποσύνολο του Ιστού, δηλαδή εξυπηρετεί μία ομάδα χρηστών που έχουν κάποιες ομοιότητες στα ενδιαφέροντα τους.

στην L1 cache, η αίτηση θα προωθηθεί στην L2 cache, κι αν βρεθεί εκεί, επιτυγχάνεται μία απομακρυσμένη ευστοχία (remote hit), με αντίστοιχο χρόνο εξυπηρέτησης $\tau_{r,i}$. Αν το αντικείμενο δε βρεθεί ούτε στην L2 cache, θα πρέπει να ζητηθεί από τον πηγαίο διακομιστή, με αποτέλεσμα ο χρόνος εξυπηρέτησης να γίνει $\tau_{s,i}$. Χωρίς βλάβη της γενικότητας, γίνεται η υπόθεση ότι οι χρόνοι $\tau_{l,i}$, $\tau_{r,i}$ και $\tau_{s,i}$ είναι σταθερές και $\tau_{l,i} < \tau_{r,i} < \tau_{s,i}$. Η Εικόνα 10.2 απεικονίζει ένα σύνηθες ιεραρχικό σχήμα caching.

Στο προτεινόμενο μοντέλο, ο αντικειμενικός σκοπός μίας L1 cache είναι να ελαχιστοποιήσει το χρόνο εξυπηρέτησης των χρηστών της, δηλαδή να ελαχιστοποιήσει το χρόνο που απαιτείται για ένα αντικείμενο να παραδοθεί στο χρήστη που το ζήτησε. Μπορεί εύκολα να δειχθεί ότι ο αναμενόμενος χρόνος, $E\{\tau_{ij}\}$, για να παραδοθεί ένα αντικείμενο σε ένα χρήστη της L1 cache i , υπολογίζεται ως εξής:

$$E\{\tau_{ij}\} = P_{h,l,i} \times \tau_{l,i} + (1 - P_{h,l,i}) P_{h,r,i} \times \tau_{r,i} + (1 - P_{h,l,i})(1 - P_{h,r,i}) \times \tau_{s,i} \quad (10.2)$$

όπου το $P_{h,l,i}$ υποδηλώνει την πιθανότητα ένα αντικείμενο να βρεθεί στην L1 cache (local hit probability) και το $P_{h,r,i}$ την πιθανότητα ένα αντικείμενο να βρεθεί στην L2 cache (remote hit probability). Αυτές οι πιθανότητες μπορούν να εκτιμηθούν με υπάρχοντα μοντέλα δημοτικότητας του Ιστού [70], όπως περιγράφεται πιο κάτω.

Έχει αποδειχθεί ότι η δημοτικότητα των αντικειμένων του Ιστού ακολουθούν την κατανομή Zipf [70]. Αυτή η συμπεριφορά έχει παρατηρηθεί σε διάφορα επίπεδα (ή περιβάλλοντα), π.χ. στο επίπεδο του προγράμματος πλοήγησης, στο επίπεδο του proxy, ακόμα και σε επίπεδο δημοτικότητας των αντικειμένων μίας ιστοσελίδας. Σύμφωνα με την κατανομή Zipf, αν καταταχθούν στοιχεία ενός συνόλου από αντικείμενα (π.χ., τα αντικείμενα μίας ιστοσελίδας) ως προς τη δημοτικότητά τους (με το αντικείμενο 1 να είναι το πιο δημοφιλές), η πιθανότητα να ζητηθεί το j -ιοστό αντικείμενο είναι η ακόλουθη:

$$\text{Pr}[\text{object } j \text{ is requested}] = K/j^a$$

όπου το $a \in (0,1)$ είναι ο παράγοντας της κατανομής Zipf, η τιμή του οποίου εξαρτάται από την εκάστοτε εφαρμογή και το K είναι ένας παράγοντας κανονικοποίησης ώστε να ισχύει

$$\sum_{j=1}^{|U|} \frac{K_j}{j^{a_i}} = 1,$$

όπου το U είναι το «σύμπαν» των αντικειμένων που εξετάζονται. Στο προτεινόμενο μοντέλο, το εν λόγω «σύμπαν» ορίζεται για κάθε L1 cache ως το σύνολο των αντικειμένων στα οποία αναφέρονται οι χρήστες της, δηλαδή $U \equiv O_i$ για την L1 cache i .

Έστω ότι μία L1 cache είναι ικανή να αποθηκεύσει τα C_i πιο δημοφιλή αντικείμενα από το σύνολο O_i . Βάσει της παραπάνω συζήτησης, είναι εμφανές ότι η πιθανότητα ευστοχίας από την L1 cache θα είναι η ακόλουθη:

$$\Pr[\text{hit}; C_i, O_i] = \sum_{j=1}^{C_i} \frac{K_j}{j^{a_i}}$$

Υπό συνήθεις συνθήκες, το μέγεθος s_j ενός αντικειμένου στο O_i είναι τάξεις μεγέθους μικρότερο από το μέγεθος της L1 cache (π.χ. KB συγκρινόμενο με GB). Αυτό επιτρέπει να προσεγγιστεί η πιθανότητα ευστοχίας της L1 cache θεωρώντας τα περιεχόμενα της σαν ένα συνεχές (continuum) δεδομένων, παρά ως διακριτά αντικείμενα. Έτσι, προκύπτει η επόμενη συνεχής έκφραση της πιθανότητας ευστοχίας της L1 cache:

$$\Pr[\text{hit}; C_i, O_i] = \int_0^{C_i} \frac{K_i}{x^{a_i}} dx$$

όπου

$$\int_0^{O_i} \frac{K_i}{x^{a_i}} dx = 1.$$

Ως εκ τούτου, η πιθανότητα $P_{h,i}$ λαμβάνει την ακόλουθη μορφή:

$$P_{h,i}(C_i) = \left(\frac{C_i}{O_i} \right)^{1-a_i} \quad (10.3)$$

Επιπλέον, γίνεται η υπόθεση ότι η L1 cache i έχει ένα ποσό χώρου y_i στο δίσκο της L2 cache, όπου είναι αποθηκευμένα τα αμέσως λιγότερο δημοφιλή αντικείμενα από αυτά που είναι αποθηκευμένα στην L1 cache. Η πιθανότητα ευστοχίας στην L2 cache $P_{h,r,i}$ είναι, επομένως, η ακόλουθη:

$$P_{h,r,i}(y_i, C_i) = \frac{\left((C_i + y_i)^{1-a_i} - C_i^{1-a_i} \right)}{O_i^{1-a_i}} \quad (10.4)$$

Από τις εξισώσεις (10.3) και (10.4), προκύπτει το συμπέρασμα ότι ο αναμενόμενος χρόνος εξυπηρέτησης $E\{\tau_i\}$ για τους χρήστες της L1 cache i είναι ο ακόλουθος:

$$E\{\tau_i; y_i, C_i\} = \left(\frac{C_i}{O_i}\right)^{1-a_i} \times \tau_{l,i} + \left(1 - \left(\frac{C_i}{O_i}\right)^{1-a_i}\right) \left(\frac{((C_i + y_i)^{1-a_i} - C_i^{1-a_i})}{O_i^{1-a_i}}\right) \times \tau_{r,i} \\ + \left(1 - \left(\frac{C_i}{O_i}\right)^{1-a_i}\right) \left(1 - \frac{((C_i + y_i)^{1-a_i} - C_i^{1-a_i})}{O_i^{1-a_i}}\right) \times \tau_{s,i}$$

Καθώς, λοιπόν, ο αντικειμενικός σκοπός της L1 cache είναι να ελαχιστοποιήσει το χρόνο εξυπηρέτησης των χρηστών της, μία φυσιολογική επιλογή συνάρτησης ωφελείας είναι η

$$u_i(y_i) = -E\{\tau_i; y_i, C_i\}. \quad (10.5)$$

Η συνάρτηση ωφελείας (10.5) αποτελεί τη βάση μέσω της οποίας οι οντότητες της αγοράς (οι L1 cache κι η L2 cache) θα υποστηρίζουν τις αποφάσεις τους. Καθώς στην παρούσα μελέτη, γίνεται η υπόθεση ύπαρξης χρέωσης, οι αποφάσεις των εμπλεκόμενων οντοτήτων θα επηρεαστεί κι από το εκάστοτε σχήμα χρέωσης που θα υιοθετηθεί. Στο επόμενο υποκεφάλαιο, μελετούνται συγκεκριμένοι τύποι αγορών που μπορούν να εφαρμοστούν στο εξεταζόμενο περιβάλλον, δηλαδή μονοπώλιο με και χωρίς διάκριση χρέωσης, καθώς κι εκχώρηση πόρων με σκοπό τη μεγιστοποίηση της κοινωνικής ευημερίας.

10.4 Μονοπώλιο με Χρέωση με Διάκριση

Όπως έχει ήδη συζητηθεί, με ένα μονοπώλιο με χρέωση με διάκριση, ο μοναδικός πωλητής μπορεί να θέσει διαφορετική χρέωση σε κάθε αγοραστή. Δεδομένων των χαρακτηριστικών των αγοραστών (δηλαδή των συναρτήσεων ωφέλειας τους)²³, ο πωλητής μπορεί να υπολογίσει ατομικές τιμές με αποτέλεσμα ένα διάνυσμα

²³ Πρέπει να σημειωθεί ότι αυτό δεν είναι πάντα εφικτό (ή και επιτρεπτό) να γνωρίζει ο πωλητής τη συνάρτηση ωφελείας του αγοραστή. Για τους λόγους της παρούσας ανάλυσης, όμως, γίνεται η υπόθεση ότι η εν λόγω πληροφορία είναι γνωστή στον πωλητή.

ζήτησης που μεγιστοποιεί τα έσοδα του²⁴. Αυτό μπορεί να εκφραστεί μαθηματικά, μέσω του ακόλουθου προβλήματος βελτιστοποίησης:

$$(\text{SELLER}) \begin{cases} \max_{\mathbf{y} \in \mathcal{S}_1 \times \dots \times \mathcal{S}_N} \mathbf{p}^T \mathbf{y} \\ \text{s.t.} \\ \sum_{j=1}^N y_j \leq C \end{cases}$$

Με άλλα λόγια, ο πωλητής (L2 cache) αντιμετωπίζει το πρόβλημα του να βρει το διάνυσμα εκχώρησης πόρων $\mathbf{y} = (y_1, \dots, y_N)^T$, το οποίο να μεγιστοποιεί το εσωτερικό γινόμενο $\mathbf{p}^T \mathbf{y}$, με τον περιορισμό ότι οι συνολικά εκχωρημένοι πόροι (χωρητικότητα δίσκου) δεν ξεπερνά το συνολικό διαθέσιμο ποσό πόρων C . Πρέπει να σημειωθεί ότι, σιωπηρά, γίνεται η υπόθεση ότι ο χώρος της L2 cache διαχωρίζεται σε διαφορετικά (απομονωμένα) διαμερίσματα που εκχωρούνται αποκλειστικά σε ξεχωριστές L1 cache, δηλαδή ο χώρος που παίρνει μία L1 cache δε μπορεί να χρησιμοποιηθεί από κάποια άλλη L1 cache. Πρέπει να σημειωθεί, επίσης, ότι η τιμή p_i κι η εκχώρηση (ή ζήτηση) σχετίζονται, όπως υπαγορεύεται από την εξίσωση (10.1).

Λόγω του περιορισμού της συνολικής διαθέσιμης χωρητικότητας, για την επίλυση του προβλήματος (SELLER), γίνεται εφαρμογή της θεωρίας Lagrange. Συγκεκριμένα, η επιδιωκόμενη λύση είναι αυτή που μεγιστοποιεί τη συνάρτηση Lagrange που ορίζεται ως εξής:

$$L(\mathbf{y}, \lambda) = \mathbf{p}^T \mathbf{y} - \lambda \left(\sum_{j=1}^N y_j - C \right) = \sum_{j=1}^N p_j y_j - \lambda \left(\sum_{j=1}^N y_j - C \right) \quad (10.6)$$

όπου το λ είναι ο πολλαπλασιαστής Lagrange που σχετίζεται με τον περιορισμό χωρητικότητας της L2 cache ($\lambda \geq 0$). Από τις συνθήκες βελτιστότητας πρώτης τάξης, προκύπτει ότι

$$\frac{\partial p_i(y_i)}{\partial y_i} y_i + p_i - \lambda = 0, \forall i \in I$$

Από την εξίσωση (10.1) προκύπτει ότι

²⁴ Γίνεται η υπόθεση ότι η L2 cache δεν έχει λειτουργικά έξοδα σε ό,τι αφορά το χώρο που εκχωρεί. Τα κόστη που έχει να αντιμετωπίσει σχετίζονται με την εγκατάσταση της υποδομής (sunk costs).

$$p_i = \frac{\partial u_i(y_i)}{\partial y_i} = A_i(C_i + y_i)^{-a_i} \quad (10.7)$$

και

$$\frac{\partial p_i(y_i)}{\partial y_i} = \frac{\partial^2 u_i(y_i)}{\partial y_i^2} = -a_i A_i(C_i + y_i)^{-a_i-1} \quad (10.8)$$

όπου

$$A_i = \left(1 - \left(\frac{C_i}{O_i}\right)^{1-a_i}\right) \left(\frac{1}{O_i}\right)^{1-a_i} (1-a_i)(\tau_{s,i} - \tau_{r,i}). \quad (10.9)$$

Ως εκ τούτου, η συνθήκη βελτιστότητας πρώτης τάξης γίνεται

$$A_i(C_i + y_i)^{-a_i-1} ((1-a_i)y_i + C_i) = \lambda. \quad (10.10)$$

Από τη θεωρία Lagrange, είτε $\lambda = 0$, είτε $\lambda > 0$. Στη δεύτερη περίπτωση, υπονοείται ότι ο περιορισμός της συνολικής χωρητικότητας είναι ενεργός (binding) στη λύση, δηλαδή,

$$\sum_{j=1}^N y_j = C.$$

Μπορεί εύκολα ναδειχθεί ότι $\lambda \neq 0$. Συγκεκριμένα, αν υποθεθεί το αντίθετο, δηλαδή ότι $\lambda = 0$, από την εξίσωση (10.10), ισχύει ότι $(1-a_i)y_i + C_i = 0$. Αυτό, προφανώς, δεν είναι δυνατόν, καθώς $y_i \geq 0$. Επομένως, προκύπτει το συμπέρασμα ότι $\lambda > 0$, δηλαδή όλη η χωρητικότητα της L2 cache θα πρέπει να εκχωρηθεί. Από την εξίσωση (10.10), μπορεί να παρατηρηθεί ότι για κάθε τιμή του πολλαπλασιαστή λ , προκύπτουν διαφορετικές εκχωρήσεις χωρητικότητας. Συγκεκριμένα, μπορεί εύκολα ναδειχθεί ότι, όσο αυξάνει το λ , το y_i αυξάνεται, και, επίσης, αυξάνεται η συνολική ζήτηση. Επομένως, το πρόβλημα της βέλτιστης επιλογής τιμών εντοπίζεται στην εύρεση της τιμής του λ για την οποία η συνολική ζήτηση για χωρητικότητα είναι ίση με τη συνολική χωρητικότητα της L2 cache. Αυτό το ζήτημα μελετάται στην επόμενη παράγραφο.

10.4.1 Καθορισμός Βέλτιστης Εκχώρησης Χωρητικότητας

Σε αυτήν την παράγραφο, εξετάζεται πως μπορεί η L2 cache να υπολογίσει την κατανομή χωρητικότητας που μεγιστοποιεί τα έσοδα της. Όπως έχει ήδη συζητηθεί, το πρόβλημα αυτό είναι ισοδύναμο με τον υπολογισμό του πολλαπλασιαστή Lagrange λ , ο οποίος, μέσω της εξίσωσης (10.10), έχει ως

αποτέλεσμα να εκχωρηθεί η χωρητικότητα της L2 cache στο σύνολο της. Αυτό θα ήταν εύκολο, αν η βέλτιστη κατανομή χωρητικότητας y_i για την L1 cache i μπορούσε να υπολογιστεί αναλυτικά από την εξίσωση (10.10). Όμως, αφού κάτι τέτοιο δεν είναι εφικτό, στην παρούσα μελέτη, προτείνεται μία κατάλληλη αριθμητική μέθοδος.

Πρέπει να σημειωθεί ότι το αριστερό μέλος της εξίσωσης (10.10), δηλαδή η συνάρτηση $f(y_i) = \partial L(\mathbf{y}, \lambda) / \partial y_i + \lambda$ είναι φθίνουσα ως προς το y_i . Επομένως, δεδομένου του συνόλου από το οποίο μπορεί να πάρει τιμές το y_i , δηλαδή $[0, O_i - C_i]$, ισχύει ότι $f(O_i - C_i) \leq f(y_i) \leq f(0)$. Έτσι, αν $f(O_i - C_i) \leq \lambda \leq f(0)$, υπάρχει μοναδικό $y_i(\lambda)$ στο οποίο $\partial L(\mathbf{y}, \lambda) / \partial y_i = 0$. Στην περίπτωση που $\lambda > f(O_i - C_i)$, τότε $y_i(\lambda) = O_i - C_i$. Παρομοίως, αν $\lambda < f(0)$, τότε $y_i(\lambda) = 0$.

Δεδομένου ενός πολλαπλασιαστή Lagrange λ , μπορεί να εφαρμοστεί η αριθμητική μέθοδος Newton-Raphson [130] στη συνάρτηση $\partial L(\mathbf{y}, \lambda) / \partial y_i = f(y_i) - \lambda$. Συγκεκριμένα, η αριθμητική μέθοδος έχει ως εξής:

$$y_i^{(n+1)} = y_i^{(n)} - \frac{f(y_i^{(n)}) - \lambda}{\partial f(y_i) / \partial y_i \big|_{y_i = y_i^{(n)}}}$$

και μπορεί να γραφεί, εναλλακτικά, ως εξής:

$$y_i^{(n+1)} = y_i^{(n)} - \frac{A_i (C_i + y_i^{(n)})^{-a_i - 1} ((1 - a_i) y_i^{(n)} + C_i) - \lambda}{-A_i a_i (C_i + y_i^{(n)})^{-a_i - 2} ((1 - a_i) y_i^{(n)} + 2C_i)}, \quad n = 0, 1, \dots \quad (10.11)$$

Με την επαναληπτική μέθοδο που περιγράφει η (10.11), μπορεί να επιλυθεί η (10.10), ως προς το y_i , για ένα δεδομένο λ , εφόσον $f(O_i - C_i) \leq \lambda \leq f(0)$. Η συγκεκριμένη επαναληπτική μέθοδος μπορεί να χρησιμοποιηθεί ως εργαλείο για την επίλυση του βασικού προβλήματος υπολογισμού του λ που έχει ως αποτέλεσμα μία εκχώρηση χωρητικότητας, στην οποία όλη η χωρητικότητα της L2 cache διατίθεται στις L1 cache.

Για τον υπολογισμό του βέλτιστου λ , ορίζεται η συνάρτηση $g(\lambda)$ ως εξής:

$$g(\lambda) = \sum_{j=1}^N y_j(\lambda) - C$$

Από την εξίσωση (10.10), φαίνεται εύκολα ότι, λόγω του γεγονότος ότι η συνάρτηση $f(y_i)$ είναι φθίνουσα ως προς το y_i , το $y_i(\lambda)$, δηλαδή η λύση της εξίσωσης (10.10), είναι επίσης φθίνον ως προς το λ . Επομένως, η συνάρτηση $g(\lambda)$ είναι

φθίνουσα ως προς το λ . Από τα παραπάνω, $y_i(0) = O_i - C_i$, και $y_i(+\infty) = 0$. Για το λόγο αυτό, υπάρχει ένα μοναδικό λ^* για το οποίο $g(\lambda^*) = 0$, εφόσον $(O_1 - C_1) + \dots + (O_N - C_N) \geq C$.

Για τον υπολογισμό της ρίζας, λ^* , της $g(\lambda)$, προτείνεται, στην παρούσα μελέτη, η χρήση της επαναληπτικής μεθόδου Newton-Raphson. Η μέθοδος αυτή έχει την παρακάτω μορφή:

$$\lambda^{(k+1)} = \lambda^{(k)} - \frac{\sum_{j=1}^N y_j(\lambda^{(k)}) - C}{\sum_{j=1}^N \left. \frac{\partial y_j(\lambda)}{\partial \lambda} \right|_{\lambda=\lambda^{(k)}}}, \quad k = 0, 1, \dots$$

Για την εφαρμογή της παραπάνω επαναληπτικής μεθόδου, κατά πρώτον, είναι αναγκαίο να υπολογιστεί η παράγωγος $\partial y_i(\lambda)/\partial \lambda$. Υπολογίζοντας την παράγωγο της (10.10) ως προς το λ , προκύπτει ότι

$$\frac{\partial y_i(\lambda)}{\partial \lambda} = - \frac{(C_i + y_i(\lambda))^{a_i+2}}{A_i a_i ((1 - a_i) y_i(\lambda) + 2C_i)} \quad (10.12)$$

Έτσι, η επαναληπτική μέθοδος για τον υπολογισμό του βέλτιστου λ λαμβάνει την ακόλουθη μορφή:

$$\lambda^{(k+1)} = \lambda^{(k)} - \frac{\sum_{j=1}^N y_j(\lambda^{(k)}) - C}{\sum_{j=1}^N \frac{(C_j + y_j(\lambda^{(k)}))^{a_j+2}}{A_j a_j ((1 - a_j) y_j(\lambda^{(k)}) + 2C_j)}}, \quad k = 0, 1, \dots \quad (10.13)$$

Το $y_i(\lambda)$ αναπαριστά τη ζήτηση του i -οστού αγοραστή, σε απάντηση στον πολλαπλασιαστή Lagrange λ , κι υπολογίζεται μέσω της επαναληπτικής διαδικασίας (10.11). Η εφαρμογή της (10.13) έχει ως αποτέλεσμα το βέλτιστο πολλαπλασιαστή Lagrange λ . Με άλλα λόγια, το προτεινόμενο σχήμα αποτελεί ένα ιεραρχικό σχήμα, με τη διαδικασία (10.13) να αποτελεί την πρωτεύουσα επαναληπτική διαδικασία και τη (10.11) την *εμφωλισμένη*.

10.4.2 Καθορισμός Τιμών και Ζητήματα Κινήτρων

Μετά τη σύγκλιση της διαδικασίας (10.13), δηλαδή αφού έχει βρεθεί ο βέλτιστος πολλαπλασιαστής Lagrange λ^* και καθοριστεί η αντίστοιχη κατανομή

χωρητικότητα, $\mathbf{y}(\lambda^*) = (y_1(\lambda^*), \dots, y_M(\lambda^*))$, μπορεί να υπολογιστεί και το βέλτιστο διάνυσμα χρέωσης. Συγκεκριμένα, από την εξίσωση (10.1), ισχύει ότι

$$p_i(\lambda^*) = \left. \frac{\partial u_i(y_i)}{\partial y_i} \right|_{y_i=y_i(\lambda^*)}$$

ή

$$p_i(\lambda^*) = A_i (C_i + y_i(\lambda^*))^{-a_i}$$

Από τα παραπάνω, τα έσοδα λόγω της L1 cache i είναι

$$profit_i(\lambda^*) = A_i (C_i + y_i(\lambda^*))^{-a_i} y_i(\lambda^*) \quad (10.14)$$

και τα συνολικά έσοδα είναι

$$profit(\lambda^*) = \sum_{j=1}^N A_j (C_j + y_j(\lambda^*))^{-a_j} y_j(\lambda^*). \quad (10.15)$$

Από τα αποτελέσματα που έχουν προκύψει μέχρι εδώ, είναι δυνατόν να διερευνηθεί η επίδραση κάποιων από τις παραμέτρους του μοντέλου στο αποτέλεσμα της αγοράς, και, επομένως, να προσδιοριστούν ζητήματα κινήτρων. Κατά πρώτον, εξετάζεται η επίδραση της συνολικής χωρητικότητας που διαθέτει η L2 cache. Όπως είναι ήδη γνωστό, στη λύση του προβλήματος (SELLER), όλη η χωρητικότητα της L2 cache πρέπει να έχει εκχωρηθεί. Όμως, δεν έχει εξετασθεί ακόμα ο τρόπος με τον οποίο σχετίζεται το όφελος της L2 cache με τη χωρητικότητα που διαθέτει, καθώς όταν αυξάνεται η χωρητικότητα, για να δοθεί πλέον όλη, θα πρέπει να ελαττωθεί η τιμή. Στην επόμενη πρόταση, δηλώνεται ότι τα έσοδα της L2 cache αυξάνονται με την αύξηση της συνολικής χωρητικότητας C .

Πρόταση 10.1: Στο μονοπώλιο με χρέωση με διάκριση, με L1 cache οι οποίες έχουν τη συνάρτηση ωφελείας $u_i(y_i) = -E\{T_i; y_i, C_i\}$, το όφελος της L2 cache αυξάνεται όταν αυξάνεται η χωρητικότητα C που διαθέτει.

Απόδειξη: Η παράγωγος της συνάρτησης των εσόδων που προέρχονται από την L1 cache i (δες εξίσωση (10.14)) ως προς το λ έχει ως εξής:

$$\frac{\partial profit_i(\lambda)}{\partial \lambda} = \frac{\partial y_i(\lambda)}{\partial \lambda} A_i (C_i + y_i(\lambda))^{-a_i-1} ((1 - a_i)y_i(\lambda) + C_i)$$

Από την (10.10) και (10.12) προκύπτει ότι

$$\frac{\partial profit_i(\lambda)}{\partial \lambda} = -\frac{(C_i + y_i(\lambda))^{a_i+2}}{A_i a_i ((1 - a_i) y_i(\lambda) + 2C_i)} \lambda$$

Παρατηρείται, έτσι, ότι το $\partial profit_i(\lambda)/\partial \lambda < 0$, δηλαδή τα έσοδα που προσφέρονται από την L1 cache i φθίνει ως προς το λ . Πρέπει να σημειωθεί ότι το βέλτιστο λ φθίνει ως προς τη συνολική χωρητικότητα C . Ως εκ τούτου, τα έσοδα $profit_i$ λόγω της L1 cache i αυξάνει ως προς τη συνολική χωρητικότητα C . Συνεπώς, το συνολικό κέρδος της L2 cache αυξάνεται όσο αυξάνεται η συνολική χωρητικότητα C . ■

Ως εκ τούτου, η L2 cache έχει κίνητρο να επενδύσει σε χωρητικότητα, ώστε να αυξήσει τα κέρδη της. Επιπλέον, για μία δεδομένη χωρητικότητα C , η L2 cache μπορεί να επωφεληθεί, λόγω της επικάλυψης των αντικειμένων που αποθηκεύονται για τις διαφορετικές L1 cache. Συγκεκριμένα, καθώς το σύνολο των αντικειμένων στα οποία αναφέρονται οι χρηστές διαφορετικών L1 cache μπορεί να μην είναι ανεξάρτητα, η L2 cache μπορεί να μην αποθηκεύει διπλά αντίγραφα αντικειμένων, εξοικονομώντας, έτσι, χωρητικότητα.

Στην παρούσα μελέτη, γίνεται η υπόθεση ότι η L1 cache γνωρίζει ακριβώς τα αντικείμενα (URL) που είναι αποθηκευμένα στην L2 cache. Τέτοια πληροφορία μπορεί να γίνει εύκολα διαθέσιμη, αν η L1 cache διατηρεί μία κατάλληλη δομή δεδομένων. Με τη δομή αυτή, η L1 cache, όταν θα λαμβάνει μία HTTP αίτηση από ένα χρήστη, θα γνωρίζει αν το ζητηθέν αντικείμενο βρίσκεται αποθηκευμένο τοπικά, απομακρυσμένα (στην L2 cache), ή θα πρέπει να μεταφερθεί από τον πηγαίο διακομιστή. Στην περίπτωση που υποτίθεται ότι υπάρχει στην L2 cache, η L1 cache θα αναμένει τη γρήγορη παράδοση του. Αν το αντικείμενο αργήσει να παραδοθεί από την L2 cache, η L1 cache μπορεί να συμπεράνει ότι το αντικείμενο δεν ήταν αποθηκευμένο εκεί, όπως θα έπρεπε, αλλά μεταφέρθηκε από τον πηγαίο διακομιστή. Σε αυτήν την περίπτωση, λέγεται ότι η L2 cache *εξαπατά* την L1 cache, δηλαδή, ενώ θα έπρεπε να έχει αποθηκεύσει κάποια αντικείμενα κατ' απαίτηση της L1 cache, δεν τα έχει αποθηκεύσει, π.χ. για να πωλήσει *εικονική* χωρητικότητα και να αυξήσει τα κέρδη της.

Έστω δ_i ένας παράγοντας που λαμβάνει τιμές από το διάστημα $[0,1]$ κι αναπαριστά την ικανοποίηση της L1 cache i για τις υπηρεσίες που της προσφέρει η L2 cache.

Συγκεκριμένα, αν $\delta_i = 1$, η L2 cache δεν έχει «απογοητεύσει» ποτέ την L1 cache i , ενώ μικρότερες τιμές του παράγοντα δ_i υποδηλώνουν ότι η L1 cache δεν είναι πλήρως ικανοποιημένη. Μία λογική επιλογή για το δ_i θα μπορούσε να είναι το ποσοστό των αιτήσεων που έχουν εξυπηρετηθεί από την L2 cache κι αφορούν σε αντικείμενα που θα έπρεπε να είναι εκεί αποθηκευμένα, δηλαδή ένα μέτρο της προσφερόμενης ποιότητας υπηρεσίας.

Για να ποσοτικοποιηθεί η επίδραση της προσφερόμενης ποιότητας υπηρεσίας στην αγορά, ορίζεται μία εναλλακτική μορφή της συνάρτησης ωφελείας της L1 cache, η οποία λαμβάνει υπόψη την προσφερόμενη ποιότητα. Συγκεκριμένα, η ωφέλεια μειώνεται κατά τον παράγοντα δ_i ως εξής:

$$u_i(y_i) = -\delta_i E\{\tau_i; y_i, C_i\} \quad (10.16)$$

Η επίδραση της υποβαθμισμένης ποιότητας υπηρεσίας ($\delta_i < 1$) για την L1 cache είναι εμφανής, όπως υποδεικνύεται από την εξίσωση (10.16). Αυτό που μένει να εξετασθεί, όμως, είναι η επίδραση στα έσοδα της L2 cache. Αν μία L1 cache δεν είναι πλήρως ικανοποιημένη, η ζήτηση που θα επιδείξει θα μειωθεί, που σημαίνει ότι ένα ποσό χωρητικότητας θα μείνει αδιάθετο. Καθώς, όμως, στο βέλτιστο σημείο λειτουργίας, πρέπει να διατεθεί όλη η χωρητικότητα, η L2 cache θα χρειαστεί να μειώσει τις τιμές, ούτως ώστε να την πωλήσει ολόκληρη. Στην επόμενη πρόταση, δηλώνεται ότι, αν η L2 cache δεν αποθηκεύει όλα τα αντικείμενα, κατ' απαίτηση των L1 cache, με αποτέλεσμα την υποβάθμιση της ποιότητας υπηρεσίας, θα παρατηρηθεί μείωση στα έσοδα της.

Πρόταση 10.2: Στο μονοπώλιο με χρέωση με διάκριση, όταν οι L1 cache έχουν τη συνάρτηση ωφελείας (10.16), τα έσοδα της L2 cache μειώνονται αν αυτή δεν αποθηκεύει όλα τα αντικείμενα που της έχουν ζητήσει οι L1 cache.

Απόδειξη: Αν $u_i(y_i) = -\delta_i E\{\tau\}$, τότε η (10.14) γίνεται

$$profit_i(\lambda^*(\boldsymbol{\delta})) = \delta_i A_i (C_i + y_i(\lambda^*(\boldsymbol{\delta})))^{-a_i} y_i(\lambda^*(\boldsymbol{\delta})),$$

ενώ η (10.15) γίνεται

$$profit(\lambda^*(\boldsymbol{\delta})) = \sum_{j=1}^N \delta_j A_j (C_j + y_j(\lambda^*(\boldsymbol{\delta})))^{-a_j} y_j(\lambda^*(\boldsymbol{\delta}))$$

Παρομοίως, η (10.10) παίρνει την ακόλουθη μορφή:

$$\delta_i A_i (C_i + y_i)^{-a_i-1} ((1 - a_i) y_i + C_i) = \lambda^*(\boldsymbol{\delta}) \quad (10.17)$$

Πρέπει να σημειωθεί ότι το λ^* αντικαταστάθηκε με το $\lambda^*(\boldsymbol{\delta})$, ούτως ώστε να τονιστεί ότι το λ που επιλύει το πρόβλημα (SELLER), τώρα, εξαρτάται από την προσφερόμενη ποιότητα υπηρεσίας, $\boldsymbol{\delta} = (\delta_1, \dots, \delta_N)$.

Η παράγωγος του $profit(\lambda^*(\boldsymbol{\delta}))$ ως προς το δ_i είναι

$$\begin{aligned} \frac{\partial profit(\lambda^*(\boldsymbol{\delta}))}{\partial \delta_i} &= \sum_{j=1, j \neq i}^N \delta_j A_j (C_j + y_j(\lambda^*(\boldsymbol{\delta})))^{-a_j-1} ((1 - a_j) y_j(\lambda^*(\boldsymbol{\delta})) + C_j) \frac{\partial y_j(\lambda^*(\boldsymbol{\delta}))}{\partial \delta_i} \\ &+ \delta_i A_i (C_i + y_i(\lambda^*(\boldsymbol{\delta})))^{-a_i-1} ((1 - a_i) y_i(\lambda^*(\boldsymbol{\delta})) + C_i) \frac{\partial y_i(\lambda^*(\boldsymbol{\delta}))}{\partial \delta_i} + A_i (C_i + y_i(\lambda^*(\boldsymbol{\delta})))^{-a_i} y_i(\lambda^*(\boldsymbol{\delta})). \end{aligned}$$

Επομένως, από την (10.17), προκύπτει ότι

$$\frac{\partial profit(\lambda^*(\boldsymbol{\delta}))}{\partial \delta_i} = \lambda^*(\boldsymbol{\delta}) \sum_{j=1}^N \frac{\partial y_j(\lambda^*(\boldsymbol{\delta}))}{\partial \delta_i} + A_i (C_i + y_i(\lambda^*(\boldsymbol{\delta})))^{-a_i} y_i(\lambda^*(\boldsymbol{\delta})).$$

Επιπλέον, καθώς ο περιορισμός της χωρητικότητας είναι πάντα ενεργός, προκύπτει ότι

$$\sum_{j=1}^N y_j(\lambda^*(\boldsymbol{\delta})) = C.$$

Επομένως,

$$\sum_{j=1}^N \frac{\partial y_j(\lambda^*(\boldsymbol{\delta}))}{\partial \delta_i} = 0.$$

Ως εκ τούτου,

$$\frac{\partial profit(\lambda^*(\boldsymbol{\delta}))}{\partial \delta_i} = A_i (C_i + y_i(\lambda^*(\boldsymbol{\delta})))^{-a_i} y_i(\lambda^*(\boldsymbol{\delta})).$$

Δηλαδή, $\partial profit(\lambda^*(\boldsymbol{\delta}))/\partial \delta_i > 0$. Με άλλα λόγια, το κίνητρο της L2 cache είναι να ικανοποιεί όσο το δυνατόν περισσότερο τις εξυπηρετούμενες L1 cache, καθώς έτσι μεγιστοποιούνται τα έσοδα της. ■

Ως εκ τούτου, η L2 cache έχει ως κίνητρο να διατηρεί τους πελάτες της (τις L1 cache) όσο πιο ικανοποιημένες είναι δυνατόν, που μεταφράζεται στο να αποθηκεύει τα αντικείμενα που της ζητούν, ώστε να είναι διαθέσιμα αν ζητηθούν εκ νέου. Το παραπάνω αποτέλεσμα αφορά στην υποβαθμισμένη παροχή ποιότητας υπηρεσίας από την L2 cache, χωρίς, όμως, να έχει κάποιο συγκεκριμένο λόγο να

το κάνει. Μπορεί κανείς να θεωρήσει την κατάσταση στην οποία η L2 cache προσποιείται ότι η διαθέσιμη χωρητικότητα είναι μεγαλύτερη από την πραγματική. Όπως εκφράζεται στην *Πρόταση 10.1*, τα έσοδα της L2 cache αυξάνονται όσο αυξάνεται κι ο διαθέσιμος χώρος της. Από την άλλη, αν η L2 cache πωλεί περισσότερη χωρητικότητα από αυτήν που στην πραγματικότητα διαθέτει, κάποια αντικείμενα δε θα αποθηκευθούν. Ως εκ τούτου, είναι πιθανόν να συμβούν κάποιες αστοχίες (miss), που είναι ικανές να υποβαθμίσουν την *αίσθηση* ποιότητας υπηρεσίας των L1 cache και, τελικά, να οδηγήσουν σε μία μείωση στα έσοδα, σύμφωνα με την *Πρόταση 10.2*. Είναι, επομένως, φανερό ότι η L2 cache που εξαπατά τις L1 cache αντιμετωπίζει μία ανταλλαγή. Το κίνητρο του να εξαπατά ή όχι (ή το πόσο να εξαπατά) δεν είναι ξεκάθαρο.

Με σκοπό την εξέταση του παραπάνω ζητήματος, έστω $Z = z_1 + \dots + z_N$ η επιπλέον (εικονική) χωρητικότητα που προσποιείται η L2 ότι έχει. Πρώτα, ποσοτικοποιείται ο παράγοντας ικανοποίησης δ_i της L1 cache i . Μία λογική επιλογή θα ήταν το ποσοστό των αιτήσεων που έχουν εξυπηρετηθεί από την L2 cache κι αφορούν σε αντικείμενα που θα έπρεπε να είναι εκεί αποθηκευμένα. Ως κατώτατο όριο, ας θεωρηθεί η περίπτωση που η L2 cache αποθηκεύει τα λιγότερο δημοφιλή αντικείμενα (δηλαδή, κάθε αντικείμενο στην L2 cache είναι λιγότερο δημοφιλές από κάθε αντικείμενο στην L1 cache). Τότε ο παράγοντας ικανοποίησης δ_i είναι η πιθανότητα να ζητηθεί από την L2 cache ένα αντικείμενο που θα έπρεπε να είναι αποθηκευμένο εκεί (εναλλακτικά διατυπωμένο, το συμπλήρωμα της πιθανότητας αστοχίας, λόγω του γεγονότος ότι η L2 cache κλέβει):

$$\delta_i(z_i) = 1 - \frac{(C_i + y_i(\mathbf{z}))^{1-a_i} - (C_i + y_i(\mathbf{z}) - z_i)^{1-a_i}}{O_i^{1-a_i}} \quad (10.18)$$

Πρέπει να σημειωθεί ότι η εκχώρηση y_i είναι συνάρτηση του \mathbf{z} (το \mathbf{z} επηρεάζει το δ , το οποίο με τη σειρά του επηρεάζει το y).

Λόγω της προαναφερθείσας ανταλλαγής, αναμένεται ότι υπάρχει ένα διάνυσμα $\mathbf{z}^* = (z_1^*, \dots, z_N^*)$ το οποίο μεγιστοποιεί τα έσοδα της L2 cache. Όμως, ένα τέτοιο σημείο είναι μάλλον δύσκολο να υπολογιστεί αναλυτικά ή κι αριθμητικά ακόμα. Παρόλ' αυτά, αυτό που είναι εφικτό να δειχθεί είναι ότι η L2 cache έχει πράγματι κίνητρο να εξαπατήσει, ή ισοδύναμα ότι υπάρχουν πιο αποδοτικά σημεία, για την L2 cache,

από την άποψη των εσόδων, σε σύγκριση με το σημείο όπου δεν εξαπατά, όπως διατυπώνεται στην επόμενη πρόταση.

Πρόταση 10.3: Στο μονοπώλιο με χρέωση με διάκριση, όταν οι L1 cache έχουν τη συνάρτηση ωφελείας (10.16), τα έσοδα της L2 cache μπορούν να αυξηθούν αν πωλεί περισσότερη χωρητικότητα στις L1 cache απ' όση έχει πραγματικά.

Απόδειξη: Η παράγωγος της $profit(\mathbf{z})$ ως προς το z_i υπολογίζεται ως εξής:

$$\begin{aligned} \frac{\partial profit(\mathbf{z})}{\partial z_i} &= \sum_{j=1}^N \delta_j(z_j) A_j (C_j + y_j(\mathbf{z}))^{-a_j-1} ((1 - a_j) y_j(\mathbf{z}) + C_j) \frac{\partial y_j(\mathbf{z})}{\partial z_i} \\ &+ \frac{\partial \delta_i(z_i)}{\partial z_i} A_i (C_i + y_i(\mathbf{z}))^{-a_i} y_i(\mathbf{z}) \end{aligned} \quad (10.19)$$

Παρακάτω φαίνεται η ισοδύναμη μορφή της (10.17), για το σενάριο της L2 cache που κλέβει:

$$\delta_i(z_i) A_i (C_i + y_i(\mathbf{z}))^{-a_i-1} ((1 - a_i) y_i(\mathbf{z}) + C_i) = \lambda^*(\mathbf{z}) \quad (10.20)$$

Από την (10.20), η (10.19) παίρνει την ακόλουθη μορφή:

$$\frac{\partial profit(\mathbf{z})}{\partial z_i} = \lambda^*(\mathbf{z}) \sum_{j=1}^N \frac{\partial y_j(\mathbf{z})}{\partial z_i} + \frac{\partial \delta_i(z_i)}{\partial z_i} A_i (C_i + y_i(\mathbf{z}))^{-a_i} y_i(\mathbf{z}) \quad (10.21)$$

Από τον περιορισμό της χωρητικότητας, δηλαδή

$$\sum_{j=1}^N y_j(\mathbf{z}) = C + \sum_{j=1}^N z_j,$$

προκύπτει ότι

$$\sum_{j=1}^N \frac{\partial y_j(\mathbf{z})}{\partial z_i} = 1.$$

Ως εκ τούτου, η (10.21) παίρνει την ακόλουθη μορφή:

$$\frac{\partial profit(\mathbf{z})}{\partial z_i} = \lambda^*(\mathbf{z}) + \frac{\partial \delta_i(z_i)}{\partial z_i} A_i (C_i + y_i(\mathbf{z}))^{-a_i} y_i(\mathbf{z}) \quad (10.22)$$

Η παράγωγος του δ_i ως προς το z_i υπολογίζεται ως εξής:

$$\begin{aligned} &\frac{\partial \delta_i(z_i)}{\partial z_i} \\ &= -\frac{1}{O_i^{1-a_i}} \left((1 - a_i) \frac{\partial y_i(\mathbf{z})}{\partial z_i} (C_i + y_i(\mathbf{z}))^{-a_i} - (1 - a_i) \left(\frac{\partial y_i(\mathbf{z})}{\partial z_i} - 1 \right) (C_i + y_i(\mathbf{z}) - z_i)^{-a_i} \right) \end{aligned} \quad (10.23)$$

Τότε, από την (10.20) και (10.23), η (10.22) γίνεται

$$\begin{aligned} & \frac{\partial profit(\mathbf{z})}{\partial z_i} \\ &= \delta_i(z_i) A_i (C_i + y_i(\mathbf{z}))^{-a_i-1} ((1-a_i)y_i(\mathbf{z}) + C_i) \\ & - \frac{1}{O_i^{1-a_i}} \left((1-a_i) \frac{\partial y_i(\mathbf{z})}{\partial z_i} (C_i + y_i(\mathbf{z}))^{-a_i} - (1-a_i) \left(\frac{\partial y_i(\mathbf{z})}{\partial z_i} - 1 \right) (C_i + y_i(\mathbf{z}) - z_i)^{-a_i} \right) A_i (C_i + y_i(\mathbf{z}))^{-a_i} y_i(\mathbf{z}). \end{aligned} \quad (10.24)$$

Από την (10.24), τώρα, υπολογίζεται η παράγωγος του $profit(\mathbf{z})$ στο σημείο $\mathbf{z} = (0, \dots, 0) = \mathbf{0}$. Πρέπει να σημειωθεί ότι

$$y_i(0, \dots, 0) = y_i,$$

$$\delta_i(0) = 1 \text{ και}$$

$$\lambda^*(0, \dots, 0) = \lambda^*.$$

$$\left. \frac{\partial profit(\mathbf{z})}{\partial z_i} \right|_{z=0} = A_i (C_i + y_i)^{-a_i-1} \left((1-a_i) \left(1 - \frac{(C_i + y_i)^{1-a_i}}{O_i^{1-a_i}} \right) y_i + C_i \right)$$

Να σημειωθεί ότι ο όρος

$$1 - \frac{(C_i + y_i)^{1-a_i}}{O_i^{1-a_i}}$$

υποδηλώνει την πιθανότητα αστοχίας, P_{miss} , του συνδυασμένου συστήματος caching L1 + L2, επομένως $P_{miss} \leq 1$. Έτσι,

$$\left. \frac{\partial profit(\mathbf{z})}{\partial z_i} \right|_{z=0} > 0,$$

δηλαδή, τα έσοδα της L2 cache αυξάνουν, στο σημείο $\mathbf{z} = \mathbf{0}$, το οποίο αποδεικνύει ότι η L2 cache μπορεί να αυξήσει τα έσοδα της πωλώντας μη-μηδενική «εικονική» χωρητικότητα στις L1 cache. ■

Συνεπώς, η L2 cache μπορεί να επωφεληθεί προσποριούμενη ότι έχει περισσότερη χωρητικότητα από την πραγματική. Όμως, η βέλτιστη εικονική χωρητικότητα, καθώς και κατανομή της ανά L1 cache είναι ένα, μάλλον, δύσκολο πρόβλημα, το οποίο δεν διερευνάται, στα πλαίσια της παρούσας μελέτης.

Συνοψίζοντας, στο παρόν υποκεφάλαιο, μελετήθηκε το πρόβλημα του caching επιπέδου-2, υπό μονοπωλιακό καθεστώς, όπου η L2 cache λειτουργεί ως ο μοναδικός πωλητής αποθηκευτικής χωρητικότητας προς τις ενδιαφερόμενες L1 cache. Η L2 cache ασκεί χρέωση με διάκριση, δηλαδή χρεώνει τις L1 cache

ατομικά, σύμφωνα με τις συναρτήσεις ωφελείας τους, με σκοπό τη μεγιστοποίηση των εσόδων της. Προτάθηκε ένας αριθμητικός αλγόριθμος που μπορεί να εφαρμοστεί από την L2 cache (κεντρικοποιημένη προσέγγιση), για τον καθορισμό των βέλτιστων τιμών και των αντίστοιχων κατανομών χωρητικότητας. Αναγνωρίστηκαν, επίσης, κίνητρα που ανακύπτουν στο εξεταζόμενο σύστημα, π.χ. ότι η L2 cache μπορεί να επωφεληθεί με αύξηση της χωρητικότητας της.

Στο επόμενο υποκεφάλαιο, γίνεται η υπόθεση ότι η L2 cache διατηρεί τον ρόλο της ορίζοντας η ίδια τις τιμές, χωρίς, όμως, να επιτρέπεται να ασκεί χρέωση με διάκριση, περιορισμός που μπορεί να τεθεί από ένα ρυθμιστή.

10.5 Μονοπώλιο με Χρέωση χωρίς Διάκριση

Στο παρόν υποκεφάλαιο, μελετάται το πρόβλημα του L2 caching με χρέωση χωρίς διάκριση. Είναι παρόμοιο με το πρόβλημα που μελετήθηκε στο υποκεφάλαιο 10.4, με τη διαφορά ότι η L2 cache δεν επιτρέπεται να θέσει διαφορετικές τιμές σε κάθε μία L1 cache ξεχωριστά. Αποφασίζει για μία ενιαία τιμή, η οποία μεγιστοποιεί τα έσοδα της κι εφαρμόζεται σε όλες τις L1 cache. Το αντίστοιχο πρόβλημα βελτιστοποίησης που αντιμετωπίζει η L2 cache ορίζεται ακολούθως:

$$(SELLER') \left\{ \begin{array}{l} \max_{y \in S_1 \times \dots \times S_N} \mathbf{p}^T \mathbf{y} \\ s.t. \\ \sum_{j=1}^N y_j \leq C, \\ \rho_i = \rho_j, \forall i, j \in I \end{array} \right.$$

Το πρόβλημα (SELLER') είναι παρόμοιο με το πρόβλημα (SELLER), με τη διαφορά ότι υπάρχει ο επιπλέον περιορισμός $\rho_i = \rho_j, \forall i, j \in I$. Επομένως, η ανάλυση που θα ακολουθήσει έχει αρκετές ομοιότητες με την ανάλυση που έγινε στο προηγούμενο υποκεφάλαιο. Για λόγους συντομίας, δε θα αναπαραχθούν αποτελέσματα που είναι κοινά και στα δύο προβλήματα. Συγκεκριμένα, με βάση τη μεθοδολογία του προηγούμενου υποκεφαλαίου, φαίνεται εύκολα ότι οι εξισώσεις (10.7)-(10.10) έχουν, επίσης, εφαρμογή στο τρέχον σενάριο. Επιπλέον, μπορεί εύκολα να δειχθεί ότι, στη λύση του προβλήματος (SELLER'), όλη η χωρητικότητα

πρέπει να εκχωρηθεί (δηλαδή ο περιορισμός της χωρητικότητας πρέπει να είναι ενεργός). Στην επόμενη παράγραφο, χαρακτηρίζεται η λύση του προβλήματος.

10.5.1 Καθορισμός Βέλτιστης Εκχώρησης Χωρητικότητας

Στην παράγραφο αυτή, υπολογίζεται η κατανομή χωρητικότητας που μεγιστοποιεί τα έσοδα της L2 cache. Σε αντίθεση με τη λύση του (SELLER), το πρόβλημα (SELLER') μπορεί να αντιμετωπιστεί πολύ ευκολότερα. Συγκεκριμένα, όπως θα δειχθεί παρακάτω, είναι δυνατόν να οριστεί η αναλυτική έκφραση της λύσης.

Από τον ορισμό του προβλήματος, $p_i = p_j, \forall i, j \in I$. Επομένως, από την (10.1) και (10.7) έχουμε ότι

$$A_i(C_i + y_i)^{-a_i} = A_j(C_j + y_j)^{-a_j}. \quad (10.25)$$

Επιπλέον, από την (10.10), προκύπτει ότι

$$A_i(C_i + y_i)^{-a_i-1}((1 - a_i)y_i + C_i) = A_j(C_j + y_j)^{-a_j-1}((1 - a_j)y_j + C_j). \quad (10.26)$$

Έτσι, από τη (10.25) και (10.26) προκύπτει το επόμενο αποτέλεσμα:

$$a_i y_i C_j = a_j y_j C_i \quad (10.27)$$

Επιπλέον, καθώς ο περιορισμός της συνολικής χωρητικότητας είναι ενεργός στη λύση, δηλαδή,

$$\sum_{j=1}^N y_j = C,$$

από την (10.27), προκύπτει ότι η βέλτιστη κατανομή χωρητικότητας y_i για την L1 cache i είναι

$$y_i = C \frac{C_i}{a_i \sum_{j=1}^N \frac{C_j}{a_j}} \quad (10.28)$$

Αν η L2 cache έχει γνώση των χαρακτηριστικών των L1 cache, η μεν μπορεί εύκολα να χρησιμοποιήσει την (10.28) για να υπολογίσει την κατανομή χωρητικότητας που μεγιστοποιεί τα έσοδα της. Στην αντίθετη περίπτωση, είναι δυνατόν να χρησιμοποιηθεί ένα επαναληπτικό καταμεμημένο σχήμα.

Αν υποθεθεί ότι η L2 cache δεν έχει καμία γνώση για τις συναρτήσεις ωφελείας των L1 cache²⁵, η L2 cache θα πρέπει να υπολογίσει τη βέλτιστη τιμή μέσω μίας επαναληπτικής διαδικασίας, που ονομάζεται *tatonnement* [49]. Αυξάνοντας την τιμή, η συνολική ζήτηση μειώνεται, ενώ μειώνοντας την τιμή αυξάνεται η ζήτηση. Ρυθμίζοντας την τιμή επαναληπτικά, μπορεί να επιτευχθεί μία τιμή ισορροπίας. Ως μέθοδος εντοπισμού της βέλτιστης τιμής, προτείνεται ένας απλός αλγόριθμος βασισμένος στη διχοτόμηση.

Φαίνεται εύκολα ότι, για μηδενική τιμή, δηλαδή, $p = 0$, η συνολική ζήτηση θα απειρισθεί. Παρομοίως, για μία αρκετά μεγάλη τιμή, έστω $p = P$, η συνολική ζήτηση των L1 cache θα πέσει κάτω από τη συνολική προσφορά, τείνοντας στο μηδέν. Μία τιμή, η οποία εξασφαλίζει μειωμένη ζήτηση, δε μπορεί να βρεθεί αναλυτικά, χωρίς γνώση των χαρακτηριστικών των L1 cache. Όμως, τέτοιες τιμές είναι εύκολο να βρεθούν πειραματικά (αυξάνοντας σταδιακά την τιμή μέχρι η ζήτηση να πέσει κάτω από κάποια χωρητικότητα). Έχοντας υπολογίσει, λοιπόν, το P , η εφαρμογή της διχοτομικής μεθόδου για τον υπολογισμό της τιμής, με την οποία η ζήτηση θα είναι ίση με τη συνολική χωρητικότητα, μπορεί να γίνει άμεσα.

Πιο κάτω συνοψίζεται ο διχοτομικός αλγόριθμος για τον καθορισμό της βέλτιστης τιμής p_{opt} , με καταναμημένο τρόπο.

1. Set $p_l := 0$, $p_r := P$, and $p = 0.5(p_l + p_r)$.
2. Post price p and calculate the aggregate demand for disk space of the L1 caches, i.e., $Y = \sum_{j=1}^N y_j(p)$
3. If $Y > C$, then set $p_l := p$, else set $p_r := p$. Then set $p_{old} := p$, and $p := 0.5(p_l + p_r)$.
4. If $|p_{old} - p| < \varepsilon$, then go to step 5, else go to step 2.
5. $p_{opt} := p$. Finish

Η σταθερά ε στο βήμα 4 είναι ένας αρκετά μικρός, θετικός, πραγματικός αριθμός, ο οποίος χρησιμοποιείται ως κριτήριο τερματισμού του επαναληπτικού αλγορίθμου.

²⁵ Κάνοντας μόνο την υπόθεση ότι οι συναρτήσεις ωφελείας των L1 cache είναι κοίλες συναρτήσεις, που αποτελεί κοινή υπόθεση σε ό,τι αφορά τη συμπεριφορά ενός αγοραστή σε μία οικονομία.

Με τον αλγόριθμο αυτόν, υπολογίζεται η βέλτιστη τιμή, p_{opt} , με την οποία πρέπει να χρεωθούν οι L1 cache ώστε να εκχωρηθεί ολόκληρη η χωρητικότητα της L2 cache.

Στην επόμενη παράγραφο, υπολογίζονται οι βέλτιστες τιμές που αντιστοιχούν στην κατανομή χωρητικότητας που υπολογίστηκε εδώ, καθώς και ζητήματα κινήτρων που ανακύπτουν.

10.5.2 Καθορισμός Τιμών και Ζητήματα Κινήτρων

Αφού καθοριστεί το βέλτιστο διάνυσμα εκχώρησης χωρητικότητας, $\mathbf{y}^* = (y_1^*, \dots, y_N^*)$, μπορεί να υπολογιστεί το αντίστοιχο βέλτιστο διάνυσμα τιμών. Συγκεκριμένα, από την (10.1) και την (10.28), προκύπτει ότι

$$p^* = p_i^* = A_i (C_i + y_i^*)^{-a_i} \quad (10.29)$$

Από τα παραπάνω, τα έσοδα της L2 cache, λόγω της L1 cache i , είναι

$$profit_i^* = p_i^* y_i^*.$$

Επομένως,

$$profit_i^* = A_i (C_i + y_i^*)^{-a_i} y_i^*$$

και

$$profit^* = \sum_{j=1}^N A_j (C_j + y_j^*)^{-a_j} y_j^*$$

Όπως στην παράγραφο 10.4.2, εξετάζεται εδώ η επίδραση των διαφόρων παραμέτρων του μοντέλου στην αγορά. Κατά πρώτον, μελετάται η επίδραση της συνολικής χωρητικότητας της L2 cache. Όπως στο μονοπώλιο με χρέωση με διάκριση, κι εδώ, τα έσοδα της L2 cache αυξάνονται όσο αυξάνεται η συνολική χωρητικότητα C , όπως αναφέρεται και στην παρακάτω πρόταση.

Πρόταση 10.4: Στο μονοπώλιο με χρέωση χωρίς διάκριση, με L1 cache που έχουν τη συνάρτηση ωφελείας $u_i(y_i) = -E\{\tau_i; y_i, C\}$, τα έσοδα της L2 cache αυξάνονται όσο αυξάνεται η συνολική χωρητικότητα C που διαθέτει.

Απόδειξη: Η απόδειξη είναι παρόμοια με την απόδειξη για την Πρόταση 10.1 ■

Συνεπώς, η L2 cache έχει το κίνητρο να επενδύσει σε χωρητικότητα, όπως και στο σενάριο χρέωσης με διάκριση, ώστε να αυξήσει τα έσοδα της. Επιπλέον, η L2 cache μπορεί να επωφεληθεί από την επικάλυψη των αντικειμένων που

αποθηκεύονται για τις διαφορετικές L1 cache, αποφεύγοντας την αποθήκευση διπλών αντιγράφων.

Όπως και στο προηγούμενο υποκεφάλαιο, εξετάζεται, εδώ, η επίδραση της υποβαθμισμένης ποιότητας υπηρεσίας που προσφέρεται από την L2 cache. Με τον όρο υποβαθμισμένη ποιότητα υπηρεσίας, εννοείται η παροχή της υπηρεσίας caching, όταν η L2 cache δεν αποθηκεύει πάντα αντικείμενα, όπως της ζητούν οι L1 cache. Ως εκ τούτου, αντικείμενα που υποτίθεται ότι θα έπρεπε να βρίσκονται στην L2 cache μεταφέρονται, τελικά, από τον πηγαίο διακομιστή. Ακολουθείται ο ίδιος συμβολισμός, όπως και στο προηγούμενο υποκεφάλαιο, δηλαδή με δ_i υποδηλώνεται ο παράγοντας ικανοποίησης της L1 cache i για τις υπηρεσίες που της παρέχει η L2 cache, ο οποίος μειώνει την τιμή της συνάρτησης ωφελείας κι οδηγεί στην εναλλακτική μορφή που περιγράφεται στην (10.16). Στη επόμενη πρόταση, δηλώνεται ότι, αν η L2 cache δεν αποθηκεύει όλα τα αντικείμενα, σύμφωνα με τις οδηγίες των L1 cache, με αποτέλεσμα κάποια υποβάθμιση στην προσφερόμενη ποιότητα υπηρεσίας, παρατηρείται μείωση στα έσοδα.

Πρόταση 10.5: Στο μονοπώλιο με χρέωση χωρίς διάκριση, με τις L1 cache να έχουν τις συναρτήσεις ωφελείας (10.16), τα έσοδα της L2 cache μειώνονται αν δεν αποθηκεύει όλα τα αντικείμενα που έχουν ζητήσει οι L1 cache.

Απόδειξη: Αν $u_i(y_i) = -\delta_i E\{T\}$, τότε η (10.29) λαμβάνει την ακόλουθη μορφή:

$$p_i(\boldsymbol{\delta}) = \delta_i A_i (C_i + y_i(\boldsymbol{\delta}))^{-a_i}$$

Αυτό σημαίνει ότι η βέλτιστη τιμή μειώνεται, ώστε να πωληθεί η ίδια συνολική χωρητικότητα. Ταυτόχρονα, φαίνεται, εύκολα, ότι η εκχώρηση χωρητικότητας δεν επηρεάζεται, δηλαδή, για κάθε L1 cache i , η βέλτιστη κατανομή χωρητικότητας θα είναι πάλι όπως στην (10.28), δηλαδή,

$$y_i(\boldsymbol{\delta}) = y_i = C \frac{C_i}{a_i \sum_{j=1}^N \frac{C_j}{a_j}}.$$

Επομένως, η L2 cache θα επιτύχει πάντα την ίδια εκχώρηση, αλλά με μία τιμή που αυξάνεται όσο αυξάνεται η ικανοποίηση των L1 cache. Ως εκ τούτου, οι L1 cache, για μία δεδομένη χωρητικότητα C πρέπει να είναι όσο το δυνατόν πιο ικανοποιημένες, για να μεγιστοποιηθούν τα έσοδα της L2 cache. ■

Ως εκ τούτου, η L2 cache έχει κίνητρο να διατηρεί τους πελάτες της (τις L1 cache) όσο πιο ικανοποιημένους γίνεται, το οποίο είναι ισοδύναμο με το να αποθηκεύονται όλα τα αντικείμενα όπως ζητούνται, ώστε να παραδίδονται όταν γίνει αίτηση για αυτά, όπως στο μονοπώλιο με χρέωση με διάκριση. Το αποτέλεσμα αυτό έρχεται σε αντίθεση, κατά μία έννοια, με την *Πρόταση* 10.4, η οποία υποστηρίζει ότι τα κέρδη της L2 cache αυξάνονται με την κατοχή περισσότερης χωρητικότητας. Αν, δηλαδή, η L2 cache αποθήκευε μέρος των αντικειμένων που ζητούν οι L1 cache και χρησιμοποιούσε την πλεονάζουσα χωρητικότητα, εκχωρώντας την σε άλλες L1 cache, ίσως να μπορούσε να αυξήσει τα έσοδα της. Η επόμενη πρόταση δηλώνει, ακριβώς, ότι η L2 cache εξαπατώντας τις L1 cache είναι δυνατόν να αυξήσει τα έσοδα της.

Πρόταση 10.6: Στο μονοπώλιο με χρέωση χωρίς διάκριση, με τις L1 cache να έχουν τη συνάρτηση ωφελείας (10.16), τα έσοδα της L2 cache μπορούν να αυξηθούν αν πωλεί περισσότερη χωρητικότητα στις L1 cache από αυτήν που πραγματικά διαθέτει.

Απόδειξη: Η απόδειξη είναι παρόμοια με την απόδειξη για την *Πρόταση* 10.3. ■

Στο παρόν υποκεφάλαιο, μελετήθηκε το πρόβλημα του caching επιπέδου-2, σε μονοπωλιακό περιβάλλον με χρέωση χωρίς διάκριση. Υπολογίστηκαν οι βέλτιστες κατανομές χωρητικότητας κι οι αντίστοιχες τιμές. Αναγνωρίστηκαν, επίσης, κίνητρα που ανακύπτουν στο συγκεκριμένο σύστημα, π.χ., ότι η L2 cache μπορεί να επωφεληθεί από αύξηση στη συνολική της χωρητικότητα.

Στο επόμενο υποκεφάλαιο, γίνεται η υπόθεση ότι η L2 cache ελέγχεται από ένα ρυθμιστή. Ο ρυθμιστής επιβάλλει στην L2 cache να διανέμει τη χωρητικότητα της στις L1 cache, με σκοπό τη μεγιστοποίηση της κοινωνικής ευημερίας.

10.6 Μονοπώλιο με Ρυθμιστή για Μεγιστοποίηση Κοινωνικής Ευημερίας

Στο παρόν υποκεφάλαιο, εξετάζεται το πρόβλημα του caching επιπέδου-2, στα πλαίσια μίας αγοράς με ύπαρξη ενός ρυθμιστή. Όπως έχει ήδη αναφερθεί, ο ρυθμιστής είναι μία τρίτη οντότητα, η οποία έχει τη δύναμη να επιβάλλει στους παίκτες της αγοράς συγκεκριμένες ενέργειες ή να θέσει συγκεκριμένους περιορισμούς, ώστε να οδηγηθεί η αγορά σε κοινωνικά αποδεκτές καταστάσεις.

Ένας συνήθης στόχος ενός ρυθμιστή αποτελεί η μεγιστοποίηση της κοινωνικής ευημερίας, δηλαδή, η μεγιστοποίηση του αθροίσματος των ωφελειών του πωλητή (πωλητών) και των αγοραστών. Ο ρυθμιστής μπορεί, λοιπόν, να επιβάλλει στον πωλητή (πωλητές) να θέσει τιμές που μεγιστοποιούν την κοινωνική ευημερία. Το αντίστοιχο πρόβλημα, εκφρασμένο με μαθηματικούς όρους, είναι το ακόλουθο:

$$(SW) \begin{cases} \max_{y \in S_1 \times \dots \times S_N} \sum_{j=1}^N u_j(y_j) \\ \text{s.t.} \\ \sum_{j=1}^N y_j \leq C \end{cases}$$

Όπως έχει ήδη αναφερθεί, στο προτεινόμενο μοντέλο, η L2 cache δεν έχει λειτουργικά κόστη που να είναι ανάλογα του ποσού των πόρων που παρέχει. Αν, βέβαια, κάτι τέτοιο δεν ίσχυε, το κόστος της L2 cache θα μπορούσε να ενσωματωθεί στην αντικειμενική συνάρτηση του προβλήματος (SW) (με αρνητικό πρόσημο). Επομένως, στο συγκεκριμένο σενάριο, δε γίνεται χρέωση των L1 cache και τα έσοδα της L2 cache είναι μηδενικά.

Λόγω του περιορισμού χωρητικότητας του προβλήματος (SW), για να βρεθεί η λύση του, χρειάζεται να μεγιστοποιηθεί η ακόλουθη συνάρτηση Lagrange:

$$L(\mathbf{y}, \lambda) = \sum_{j=1}^N u_j(y_j) - \lambda \left(\sum_{j=1}^N y_j - C \right) \quad (10.30)$$

όπου λ είναι ο πολλαπλασιαστής Lagrange που σχετίζεται με τον περιορισμό χωρητικότητας της L2 cache ($\lambda \geq 0$).

10.6.1 Καθορισμός Βέλτιστης Εκχώρησης Χωρητικότητας

Στην παρούσα παράγραφο, εξετάζεται πως η L2 cache μπορεί να υπολογίσει την κατανομή χωρητικότητας που μεγιστοποιεί την κοινωνική ευημερία, όπως επιβάλλεται από την πολιτική του ρυθμιστή. Από την συνθήκη βελτιστότητας πρώτου βαθμού, αναφορικά με τη συνάρτηση Lagrange που ορίζεται στην (10.30), δηλαδή,

$$\frac{\partial L(\mathbf{y}, \lambda)}{\partial y_i} = 0,$$

προκύπτει ότι

$$A_i(C_i + y_i)^{-a_i} = \lambda. \quad (10.31)$$

Ο πολλαπλασιαστής Lagrange λ αντιστοιχεί στον περιορισμό χωρητικότητας της L2 cache και στη λύση θα είναι είτε ενεργός, δηλαδή, θα έχει εκχωρηθεί όλη η χωρητικότητα, είτε όχι. Στη μεν περίπτωση, $\lambda > 0$, ενώ στη δε, $\lambda = 0$. Μπορεί, όμως, εύκολα να δειχθεί ότι $\lambda > 0$, καθώς, αν $\lambda = 0$, θα πρέπει, από την (10.31), να ισχύει ότι

$$(C_i + y_i)^{-a_i} = 0,$$

το οποίο δεν είναι δυνατόν. Επομένως ισχύει ότι, $\lambda > 0$. Συνεπώς, προκύπτει ότι, στη λύση, θα ισχύει

$$y_i(\lambda) = \max\left\{a_i \sqrt[a_i]{\frac{A_i}{\lambda}} - C_i, 0\right\}, \quad (10.32)$$

όπου, για το λ θα πρέπει να ισχύει ότι

$$\sum_{i=1}^N y_i(\lambda) = C.$$

Επομένως, το πρόβλημα του υπολογισμού της βέλτιστης κατανομής χωρητικότητας εντοπίζεται στον καθορισμό του πολλαπλασιαστή Lagrange λ , ο οποίος οδηγεί σε πλήρη εκχώρηση της διαθέσιμης χωρητικότητας. Ο αναλυτικός υπολογισμός του λ δεν είναι, όμως, ιδιαίτερα εύκολος, γι' αυτό κρίνεται αναγκαία η χρήση μίας αριθμητικής μεθόδου. Στα πλαίσια της παρούσας μελέτης, προτείνεται η χρήση της μεθόδου διχοτόμησης. Συγκεκριμένα, θεωρείται ένας ελάχιστος πολλαπλασιαστής λ_{\min} κι ένας μέγιστος λ_{\max} , με τον αναζητούμενο πολλαπλασιαστή να βρίσκεται στο διάστημα $(\lambda_{\min}, \lambda_{\max})$. Ως λ_{\min} μπορεί να επιλεγεί το μηδέν²⁶. Ως λ_{\max} μπορεί να επιλεγεί ο πολλαπλασιαστής, ο οποίος οδηγεί τη συνολική ζήτηση σε επίπεδα μικρότερα της διαθέσιμης χωρητικότητας C . Αν ετίθετο ως απαίτηση $y_i = 0$, τότε ο αντίστοιχος πολλαπλασιαστής λ_i θα ήταν

$$\lambda_i = \frac{A_i}{C^{a_i}}.$$

Επομένως, αν επιλεγεί ως λ_{\max} το

$$\lambda_{\max} = \max_{i \in I} \{\lambda_i\},$$

²⁶ Όπως έχει αναφερθεί παραπάνω, $\lambda \neq 0$. Η επιλογή του μηδενός ως ελαχίστου ορίου στη μέθοδο της διχοτόμησης διασφαλίζει τον περιορισμό αυτό (τα όρια στα οποία ορίζεται η μέθοδος δεν προκύπτουν ποτέ ως λύση).

εξασφαλίζεται ότι ο πολλαπλασιαστής Lagrange που αντιστοιχεί στη λύση του προβλήματος (SW) θα βρίσκεται στο διάστημα $(\lambda_{\min}, \lambda_{\max})$. Δεδομένων, λοιπόν, των ορίων λ_{\min} και λ_{\max} , ο διχοτομικός αλγόριθμος για τον υπολογισμό του επιδιωκόμενου λ είναι ο εξής:

1. Set $\lambda_l := \lambda_{\min}$, $\lambda_r := \lambda_{\max}$, and $\lambda = 0.5(\lambda_l + \lambda_r)$.
2. Compute the demand of each L1 cache (see (10.32)), and calculate the aggregate demand for disk space, i.e., $Y = \sum_{j=1}^N y_j(\lambda)$
3. If $Y > C$, then set $\lambda_l := \lambda$, else set $\lambda_r := \lambda$. Then set $\lambda_{old} := \lambda$, and $\lambda := 0.5(\lambda_l + \lambda_r)$.
4. If $|\lambda_{old} - \lambda| < \epsilon$, then go to step 5, else go to step 2.
5. $\lambda_{opt} := \lambda$. Finish

Πρέπει να σημειωθεί ότι, σε αντίθεση με τα προηγούμενα σενάρια που εξετάστηκαν (χρέωση με και χωρίς διάκριση), στο παρόν σενάριο, δεν υφίστανται θέματα εξαπάτησης από την L2 cache, με σκοπό την αύξηση των εσόδων της, καθώς εξ' ορισμού, εδώ, τα έσοδα της είναι μηδενικά. Επομένως, δεν αναλύονται ζητήματα κινήτρων, όπως έγινε στα προηγούμενα υποκεφάλαια.

10.7 Συγκριτική Αξιολόγηση των Διαφορετικών Τύπων Μονοπωλίων

Στο παρόν υποκεφάλαιο, γίνεται μία συνολική αξιολόγηση των τριών διαφορετικών τύπων αγορών που εξετάστηκαν, στην παρούσα μελέτη, με σκοπό την εφαρμογή τους στο πρόβλημα του caching επιπέδου-2, δηλαδή, το μονοπώλιο με χρέωση με διάκριση, μονοπώλιο με χρέωση χωρίς διάκριση και μονοπώλιο με ύπαρξη ρυθμιστή (με έμφαση στην κοινωνική ευημερία). Συγκεκριμένα, δόθηκε έμφαση σε ποσοτικά χαρακτηριστικά, όπως τα έσοδα της L2 cache και το όφελος που απολαμβάνουν οι L1 cache (δηλαδή το καθαρό όφελος), καθώς και ποιοτικά ζητήματα, όπως περιορισμοί των εξεταζόμενων μοντέλων κι άλλα χαρακτηριστικά.

Παρουσιάζεται ένα απλό παράδειγμα με τρεις L1 cache και μία L2 cache. Ο Πίνακας 10.1 παρουσιάζει τα χαρακτηριστικά της εξεταζόμενης υποδομής. Οι παράμετροι που παρουσιάζει ο Πίνακας 10.1 περιγράφουν μία βασική ρύθμιση, η οποία είναι συμμετρική (κι οι τρεις L1 cache έχουν τα ίδια χαρακτηριστικά). Με μία

τέτοια ρύθμιση, υποθέτοντας, επίσης, ότι οι όλες οι cache έχουν όμοιες συναρτήσεις ωφελείας, όλα τα εξεταζόμενα σχήματα θα δώσουν τα ίδια αποτελέσματα. Ως εκ τούτου, για να συγκριθούν τα εν λόγω σχήματα, είναι απαραίτητο να εισαχθεί κάποια διαφορετικότητα μεταξύ των L1 cache.

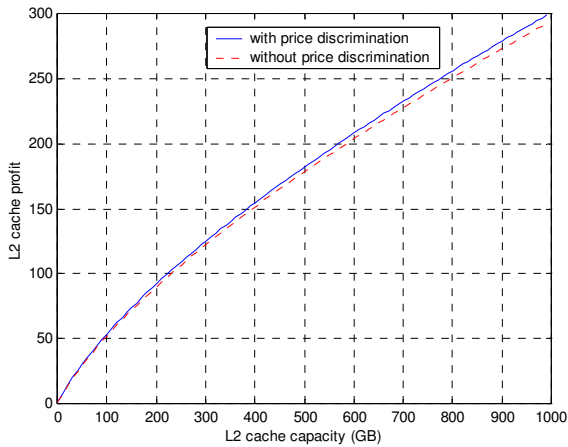
Πίνακας 10.1. Χαρακτηριστικά της υποδομής του παραδείγματος

	L1 Cache 1	L1 Cache 2	L1 Cache 3
Χωρητικότητα επιπέδου-1 (C_i)	1GB	1GB	1GB
Συνολική χωρητικότητα αντικειμένων αναφερόμενα από την cache επιπέδου-1 (O_i)	500GB	500GB	500GB
Παράγοντας δημοτικότητας Zipf (a_i)	0.3	0.3	0.3
Χρόνος εξυπηρέτησης επιπέδου-1 ($\tau_{l,i}$)	5ms	5ms	5ms
Χρόνος εξυπηρέτησης επιπέδου-2 ($\tau_{r,i}$)	10ms	10ms	10ms
Χρόνος εξυπηρέτησης από τον πηγαίο διακομιστή ($\tau_{s,i}$)	1,000ms	1,000ms	1,000ms
	L2 Cache		
Συνολική Χωρητικότητα (C)	1:10:1000GB		

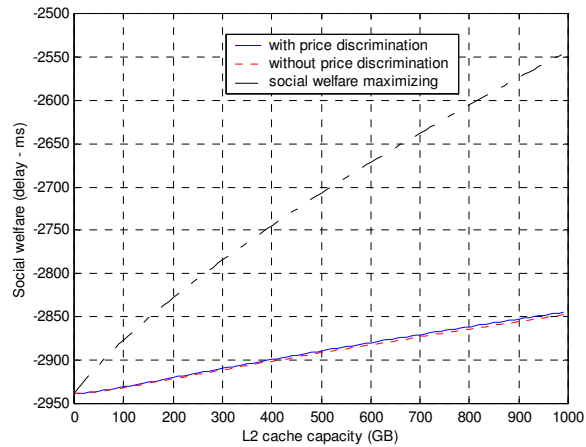
Με βάση τις παραμέτρους του πίνακα, ορίζεται ένα σύνολο σεναρίων με μεταβολή κάποιων από τις παραμέτρους. Συγκεκριμένα, για να διαφανεί η επίδραση της χωρητικότητας που διαθέτει κάθε L1 cache (C_i), στο χώρο που θα αποκτήσει από την L2 cache (y_i), καθώς και στην προκύπτουσα ωφέλεια, αντί του διανύσματος $(C_1, C_2, C_3) = (1GB, 1GB, 1GB)$, χρησιμοποιείται το $(C_1, C_2, C_3) = (1GB, 2GB, 3GB)$.

Στην Εικόνα 10.3 φαίνονται τα έσοδα για την L2 cache, για το μονοπώλιο με χρέωση με διάκριση και για το μονοπώλιο με χρέωση χωρίς διάκριση²⁷. Μπορούν να παρατηρηθούν τα εξής: 1) Όταν γίνεται χρέωση με διάκριση, η L2 cache λαμβάνει περισσότερα έσοδα και 2) τα έσοδα της L2 cache αυξάνονται όσο αυξάνεται κι η χωρητικότητα της, σύμφωνα και με την *Πρόταση 10.1* και την *Πρόταση 10.4*.

²⁷ Η περίπτωση της αγοράς με σκοπό τη μεγιστοποίηση της κοινωνικής ευημερίας δεν περιλαμβάνεται, καθώς σε αυτήν η L2 cache δε λαμβάνει έσοδα.



Εικόνα 10.3. Έσοδα της L2 cache



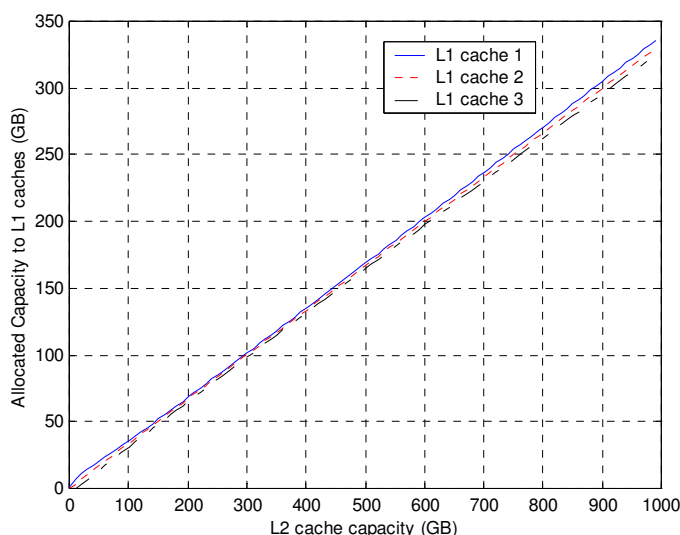
Εικόνα 10.4. Κοινωνική Ευημερία

Στην Εικόνα 10.4, φαίνεται η κοινωνική ευημερία που επιτυγχάνεται στο εξεταζόμενο σύστημα αναλόγως με το οικονομικό καθεστώς που υφίσταται σε κάθε περίπτωση. Ως κοινωνική ευημερία, ορίζεται το άθροισμα των καθαρών ωφελειών των παικτών της αγοράς, δηλαδή

$$SW(\mathbf{y}, \mathbf{p}) = \sum_{i=1}^N (u_i(y_i) - p_i y_i).$$

Οι αρνητικές τιμές που απεικονίζονται οφείλονται στο ότι η συνάρτηση ωφελείας ορίζεται ως το αρνητικό του χρόνου εξυπηρέτησης (δες (10.8)). Από την Εικόνα 10.4, μπορούν να παρατηρηθούν τα εξής: 1) Η αγορά με σκοπό τη μεγιστοποίηση της κοινωνικής ευημερίας, επιτυγχάνει κοινωνικά αποτελεσματικότερες κατανομές χωρητικότητας, σε σχέση με τις άλλες δύο αγορές (με χρέωση με και χωρίς διάκριση), 2) Όταν εφαρμόζεται χρέωση με διάκριση, επιτυγχάνονται ελαφρώς υψηλότερες τιμές κοινωνικής ευημερίας, σε σχέση με τη χρέωση χωρίς διάκριση. Η πρώτη παρατήρηση είναι εξ' ορισμού αναμενόμενη. Η δεύτερη, όμως, δεν είναι το ίδιο αναμενόμενη, καθώς η αγορά με χρέωση με διάκριση, είναι πλήρως χρησιμοθηρική (με σκοπό τη μεγιστοποίηση των εσόδων της, με τη μεγαλύτερη δυνατή εκμετάλλευση των αγοραστών). Παρόλ' αυτά, φαίνεται ότι η χρέωση με διάκριση είναι πιο ευέλικτη από την ομοιόμορφη χρέωση επιτυγχάνοντας και μεγαλύτερα κέρδη και μεγαλύτερη ευημερία για το κοινωνικό σύνολο. Σημαντικό, όμως, κριτήριο αποτελεί και το πως αντιλαμβάνεται η κάθε L1 cache ξεχωριστά το κάθε σχήμα χρέωσης, κάτι που δεν είναι δυνατόν να φανεί από την κοινωνική ευημερία (άθροισμα ωφελειών).

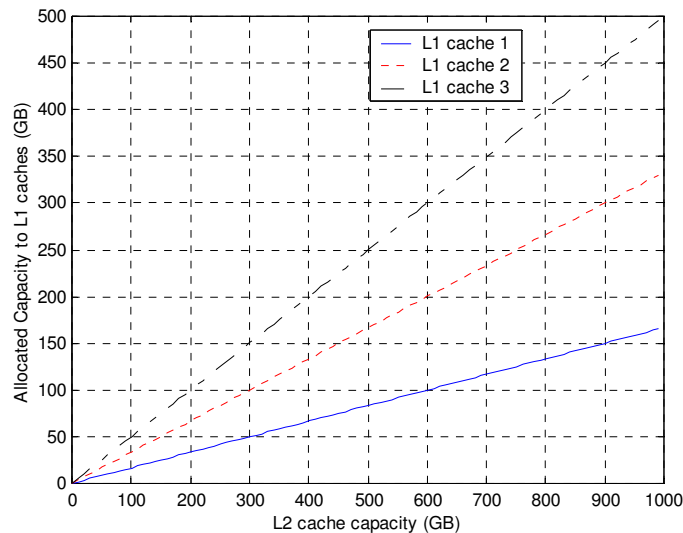
Στην Εικόνα 10.5, φαίνεται η χωρητικότητα που εκχωρείται σε κάθε μία από τις τρεις L1 cache, για την αγορά με χρέωση με διάκριση. Σύμφωνα με τα αποτελέσματα, η L1 cache 1 παίρνει περισσότερη χωρητικότητα από την L1 cache 2, η οποία παίρνει περισσότερη χωρητικότητα από την L1 cache 3. Το συγκεκριμένο αποτέλεσμα οφείλεται στο ότι, καθώς $C_1 < C_2 < C_3$, η L1 cache 1 έχει περισσότερη ανάγκη για χωρητικότητα από την L1 cache 2, η οποία έχει περισσότερη ανάγκη από την L1 cache 3, γεγονός που εκμεταλλεύεται η L2 cache, ώστε να μεγιστοποιήσει τα κέρδη της. Ένα, επίσης, αποτέλεσμα που αξίζει να σημειωθεί είναι το γεγονός ότι, για μικρές τιμές της χωρητικότητας της L2 cache, λόγω του χρησιμοθηρικού χαρακτήρα της L2 cache, η L1 cache 2 και L1 cache 3 δε λαμβάνουν καθόλου χωρητικότητα, γεγονός που μπορεί να χαρακτηρίσει το εν λόγω σχήμα άδικο.



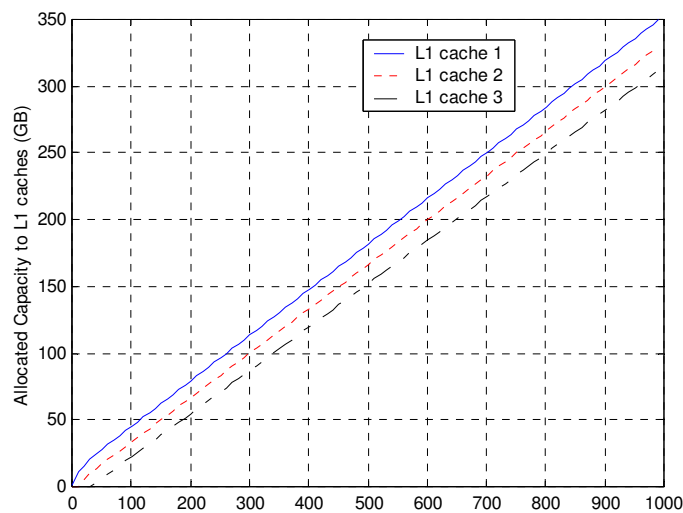
Εικόνα 10.5. Χρέωση με Διάκριση – Εκχωρημένη χωρητικότητα σε κάθε L1 cache

Στην Εικόνα 10.6, φαίνεται η χωρητικότητα που εκχωρείται σε κάθε μία από τις τρεις L1 cache, για την αγορά με χρέωση χωρίς διάκριση. Σε αντίθεση με την αγορά με χρέωση με διάκριση, μία L1 cache λαμβάνει περισσότερη χωρητικότητα όσο μεγαλύτερη είναι κι η χωρητικότητα επιπέδου-1 που διαθέτει (C_i) – δεξ (10.28). Έτσι η L1 cache 1 παίρνει λιγότερη χωρητικότητα από την L1 cache 2, η οποία παίρνει λιγότερη χωρητικότητα από την L1 cache 3. Έτσι, η χωρητικότητα, δε δίδεται στις cache που την έχουν περισσότερη ανάγκη, με αποτέλεσμα μικρότερα έσοδα, αλλά και κοινωνική ευημερία. Αυτό που αξίζει, όμως, να επισημανθεί είναι το γεγονός ότι η αγορά με ομοιόμορφη χρέωση είναι πιο «δίκαιη», καθώς και σε

μικρές χωρητικότητες της L2 cache, όλες οι L1 cache λαμβάνουν μη μηδενικό μερίδιο της χωρητικότητας.



Εικόνα 10.6. Χρέωση χωρίς Διάκριση – Εκχωρημένη χωρητικότητα σε κάθε L1 cache

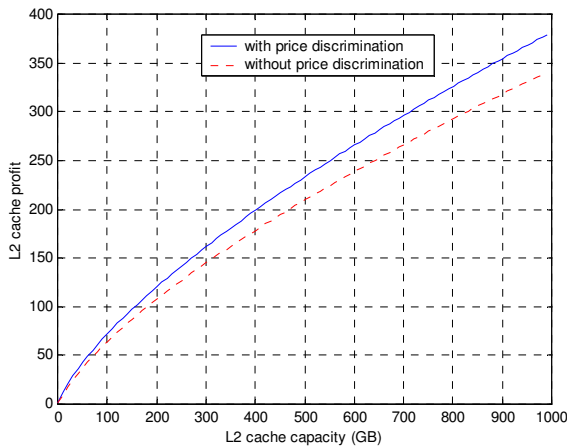


Εικόνα 10.7. Μεγιστοποίηση Κοινωνικής Ευημερίας – Εκχωρημένη χωρητικότητα σε κάθε L1 cache

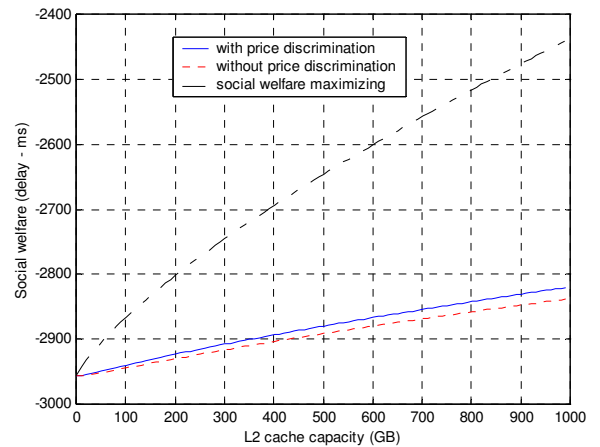
Στην Εικόνα 10.7, φαίνεται η χωρητικότητα που εκχωρείται σε κάθε μία από τις τρεις L1 cache, για την αγορά με σκοπό τη μεγιστοποίηση της κοινωνικής ευημερίας. Παρατηρείται σημαντική ομοιότητα μεταξύ του συγκεκριμένου σχήματος και της αγοράς με χρέωση με διάκριση. Συγκεκριμένα, περισσότερη χωρητικότητα λαμβάνουν αυτοί που την έχουν περισσότερη ανάγκη (με μικρότερη χωρητικότητα

επιπέδου-1) και, για χαμηλές τιμές χωρητικότητας της L2 cache, η L1 cache 2 κι η L1 cache 3 δε λαμβάνουν καθόλου χωρητικότητα.

Εκτός από το σενάριο στο οποίο οι L1 cache διαφέρουν στη χωρητικότητα επιπέδου-1 που διαθέτουν (C_i), μελετάται και το σενάριο, στο οποίο διαφέρουν στο συνολικό χώρο των αντικειμένων στα οποία αναφέρονται (O_i). Συγκεκριμένα, οι χωρητικότητες επιπέδου-1 τίθενται στις αρχικές τους τιμές, δηλαδή $(C_1, C_2, C_3) = (1\text{GB}, 1\text{GB}, 1\text{GB})$, ενώ για τις συνολικό χώρο των αντικειμένων χρησιμοποιείται το διάστημα $(O_1, O_2, O_3) = (250\text{GB}, 500\text{GB}, 750\text{GB})$.



Εικόνα 10.8. Έσοδα της L2 cache

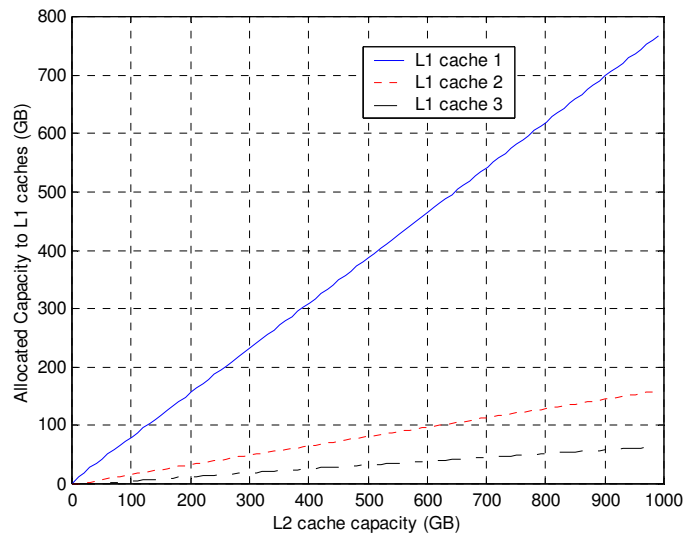


Εικόνα 10.9. Κοινωνική Ευημερία

Στην Εικόνα 10.8 φαίνονται τα έσοδα για την L2 cache, για το μονοπώλιο με χρέωση με διάκριση και για το μονοπώλιο με χρέωση χωρίς διάκριση. Όπως και για το προηγούμενο σενάριο (διαφορετικές χωρητικότητες επιπέδου-1), παρατηρείται ότι 1) όταν γίνεται χρέωση με διάκριση, η L2 cache λαμβάνει περισσότερα έσοδα και 2) τα έσοδα της L2 cache αυξάνονται όσο αυξάνεται κι η χωρητικότητα της. Στην Εικόνα 10.9, φαίνεται η κοινωνική ευημερία που επιτυγχάνεται στο εξεταζόμενο σύστημα, αναλόγως με το οικονομικό καθεστώς που υφίσταται σε κάθε περίπτωση. Μπορεί να παρατηρηθεί, όπως και στο προηγούμενο σενάριο, ότι 1) η αγορά με σκοπό τη μεγιστοποίηση της κοινωνικής ευημερίας, επιτυγχάνει κοινωνικά αποτελεσματικότερες κατανομές χωρητικότητας, σε σχέση με τις άλλες δύο αγορές (με χρέωση με και χωρίς διάκριση) και 2) όταν εφαρμόζεται χρέωση με διάκριση, επιτυγχάνονται ελαφρώς υψηλότερες τιμές

κοινωνικής ευημερίας, σε σχέση με τη χρέωση χωρίς διάκριση, λόγω της ευελιξίας της πρώτης.

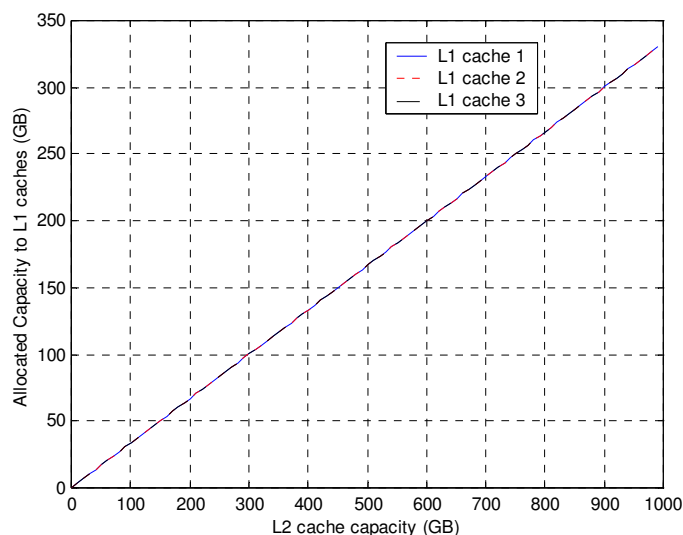
Στην Εικόνα 10.10, φαίνεται η χωρητικότητα που εκχωρείται σε κάθε μία από τις τρεις L1 cache, για την αγορά με χρέωση με διάκριση. Σύμφωνα με τα αποτελέσματα, η L1 cache 1 παίρνει περισσότερη χωρητικότητα από την L1 cache 2, η οποία παίρνει περισσότερη χωρητικότητα από την L1 cache 3. Το συγκεκριμένο αποτέλεσμα οφείλεται στο ότι $O_1 < O_2 < O_3$. Όσο πιο μεγάλος είναι ο χώρος των αντικειμένων στα οποία αναφέρεται μία L1 cache, τόσο μικρότερη ωφέλεια της προσφέρει ένα συγκεκριμένο ποσό χωρητικότητας. Έτσι, είναι πιο αποτελεσματικό να εκχωρείται περισσότερη χωρητικότητα στις L1 cache που έχουν μικρότερο χώρο αναφερόμενων αντικειμένων. Ένα, επίσης, αποτέλεσμα είναι ότι, για μικρές τιμές της χωρητικότητας της L2 cache, λόγω του χρησιμοθηρικού χαρακτήρα της L2 cache, η L1 cache 2 και L1 cache 3 δε λαμβάνουν καθόλου χωρητικότητα.



Εικόνα 10.10. Χρέωση με Διάκριση – Εκχωρημένη χωρητικότητα σε κάθε L1 cache

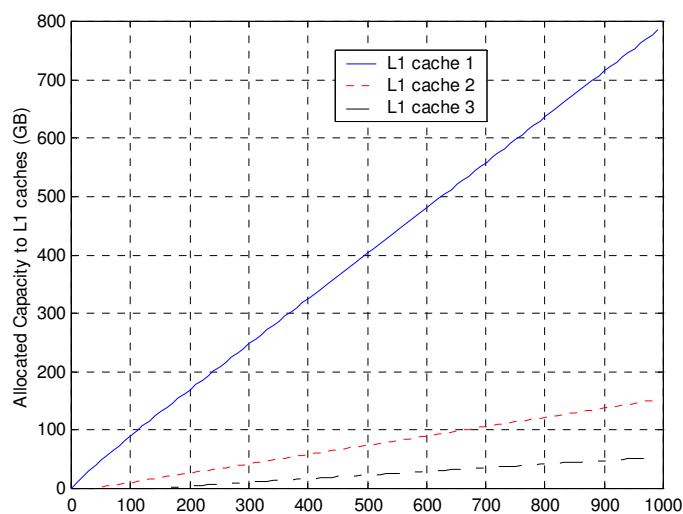
Στην Εικόνα 10.11, φαίνεται η χωρητικότητα που εκχωρείται σε κάθε μία από τις τρεις L1 cache, για την αγορά με χρέωση χωρίς διάκριση. Σε αντίθεση με την αγορά με χρέωση με διάκριση, ο χώρος των αντικειμένων στα οποία αναφέρεται μία L1 cache δεν επηρεάζει τη χωρητικότητα που λαμβάνει, όπως φαίνεται κι από την (10.28). Έτσι, κι οι τρεις L1 cache παίρνουν ακριβώς την ίδια χωρητικότητα, χωρίς να λαμβάνεται υπόψη το πόσο σημαντική είναι η χωρητικότητα για κάθε μία.

Ως εκ τούτου, τα έσοδα είναι μικρότερα για την L2 cache κι η κοινωνική ευημερία περιορισμένη. Αυτό που αξίζει, όμως, να επισημανθεί είναι το γεγονός ότι η αγορά με ομοιόμορφη χρέωση είναι πιο «δίκαιη», καθώς και σε μικρές χωρητικότητες της L2 cache, όλες οι L1 cache λαμβάνουν μη μηδενικό μερίδιο της χωρητικότητας.



Εικόνα 10.11. Χρέωση χωρίς Διάκριση – Εκχωρημένη χωρητικότητα σε κάθε L1 cache

Στην Εικόνα 10.12, φαίνεται η χωρητικότητα που εκχωρείται σε κάθε μία από τις τρεις L1 cache, για την αγορά με σκοπό τη μεγιστοποίηση της κοινωνικής ευημερίας. Παρατηρείται σημαντική ομοιότητα μεταξύ του συγκεκριμένου σχήματος και της αγοράς με χρέωση με διάκριση.



Εικόνα 10.12. Μεγιστοποίηση Κοινωνικής Ευημερίας – Εκχωρημένη χωρητικότητα σε κάθε L1 cache

10.8 Συμπεράσματα και Μελλοντική Εργασία

Στο κεφάλαιο αυτό, προτάθηκε ένα Οικονομικό πλαίσιο για τη διαχείριση της υπηρεσίας caching σε μία ιεραρχία. Στα πλαίσια της μελέτης, προτάθηκε ένα μοντέλο βασισμένο σε οικονομικές και παιγνιοθεωρητικές αρχές, βάσει του οποίου ορίστηκαν οι συμπεριφορές που διέπουν τις εμπλεκόμενες οντότητες. Σύμφωνα με το μοντέλο που προτάθηκε, οι cache επιπέδου-1, με σκοπό την επέκταση της χωρητικότητας τους, μπορούσαν να αγοράσουν χωρητικότητα από μία cache επιπέδου-2. Σε αντίθεση με το τυπικό ιεραρχικό μοντέλο caching, στο οποίο η χωρητικότητα μίας cache διατίθεται προς κοινή χρήση, στο προτεινόμενο μοντέλο, ο χώρος της L2 cache χωρίζεται σε τμήματα που χρησιμοποιούνται αποκλειστικά από κάθε L1 cache. Ως εκ τούτου, αποφεύγονται προβλήματα μονοπώλησης της χωρητικότητας από «επιθετικές» L1 cache, καθώς κάθε L1 cache μπορεί να επωφεληθεί μόνο από το δικό της χώρο.

Στα πλαίσια της μελέτης, με βάση το προτεινόμενο μοντέλο, εξετάσθηκαν διαφορετικά σχήματα αγορών που θα μπορούσαν να εφαρμοστούν. Συγκεκριμένα, μελετήθηκε το σχήμα χρέωσης με διάκριση, στο οποίο η L2 cache (πωλητής) μπορεί να θέτει διαφορετική τιμή της χωρητικότητας για κάθε L1 cache, με σκοπό τη μεγιστοποίηση των εσόδων της. Επίσης, μελετήθηκε η εφαρμογή χρέωσης χωρίς διάκριση (ομοιόμορφη χρέωση), όπου όλες οι L1 cache χρεώνονται το ίδιο ανά μονάδα χωρητικότητας. Επιπλέον, εξετάσθηκε κι ένα σχήμα χωρίς χρέωση, σκοπός του οποίου είναι να επιτευχθούν εκχωρήσεις χωρητικότητας που μεγιστοποιούν την κοινωνική ευημερία. Από την αξιολόγηση των εν λόγω σχημάτων, προέκυψε ότι η ομοιόμορφη χρέωση δεν είναι ιδιαίτερα ευέλικτη, με αποτέλεσμα τα έσοδα της L2 cache να μειώνονται, αλλά κι η κοινωνική ευημερία να είναι περιορισμένη, σε σχέση με τη χρέωση με διάκριση, καθώς η χωρητικότητα δε διανέμεται σύμφωνα με το πόσο ανάγκη την έχει κάθε L1 cache. Ένα πλεονέκτημα της ομοιόμορφης χρέωσης αποτελεί ότι είναι πιο «δίκαια» σε σχέση με τα άλλα δύο σχήματα, όταν η χωρητικότητα της L2 cache είναι σχετικά μικρή, καθώς όλες οι L1 cache λαμβάνουν μη μηδενική χωρητικότητα.

Μελλοντικά, θα ήταν επιθυμητό το μοντέλο που προτάθηκε, στα πλαίσια της παρούσας μελέτης, να επεκταθεί. Συγκεκριμένα, θα ήταν ενδιαφέρον να μελετηθεί η γενίκευση του πέραν του επιπέδου που εξετάσθηκε (L1 – L2). Επίσης,

ενδιαφέρον θα ήταν, εκτός από την μονοπωλιακή αγορά που μελετήθηκε εδώ, το μοντέλο να επεκταθεί και για αγορές ολιγοπωλιακού χαρακτήρα, όπου μία cache ενός επιπέδου θα μπορεί να εξυπηρετηθεί από περισσότερες από μία cache του υψηλότερου επιπέδου, στην οποία περίπτωση ανακύπτουν ζητήματα ανταγωνισμού μεταξύ των cache στο υψηλότερο επίπεδο. Επίσης, θα μπορούσαν να μελετηθούν σχήματα διαχείρισης των πόρων αποθήκευσης τέτοιων συστημάτων βασισμένα σε δημοπρασίες.

ΚΕΦΑΛΑΙΟ 11

ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΚΑΤΕΥΘΥΝΣΕΙΣ

11.1 Συμπεράσματα

Στη διατριβή αυτή, παρουσιάστηκε ένα σύνολο από μελέτες που αφορούν σε ζητήματα διαχείρισης πόρων σε κινητά-ασύρματα και κατανομημένα συστήματα. Εξετάσθηκαν διάφορα προβλήματα, σε διαφορετικά, συμπληρωματικά επίπεδα. Βασικός άξονας της μελέτης που διενεργήθηκε αποτελεί η μεθοδολογία που ακολουθήθηκε για τη μοντελοποίηση, αλλά κι επίλυση των εν λόγω προβλημάτων, στα οποία ενέχονται στοιχεία ανταγωνισμού κι αμοιβαίων αλληλεπιδράσεων. Συγκεκριμένα, χρησιμοποιήθηκε η Θεωρία Παιγνίων ως εργαλείο για τη μοντελοποίηση τους, καθώς και στοιχεία θεωρίας βελτιστοποίησης για την επίλυση των προβλημάτων στα οποία οδηγούσαν τα παιγνιοθεωρητικά μοντέλα που προτάθηκαν. Τα προβλήματα που εξετάσθηκαν, στα πλαίσια της παρούσας διατριβής, μπορούν να κατηγοριοποιηθούν σε δύο βασικές θεματικές ενότητες: 1) ζητήματα σχετικά με τον κινητό υπολογισμό και 2) ζητήματα σχετικά με κατανομημένα υπολογιστικά ζητήματα.

Στην πρώτη ενότητα, μελετήθηκαν προβλήματα που αφορούσαν σε διαφορετικά επίπεδα επικοινωνίας, αλλά και χρονική κλίμακα, καθιστώντας τα, έτσι, συμπληρωματικά μεταξύ τους. Συγκεκριμένα, μελετήθηκαν τα ακόλουθα προβλήματα:

α) Διαχείριση θέσης των χρηστών σε ασύρματα δίκτυα με σκοπό την εξισορρόπηση φόρτου (Κεφάλαιο 4): παρουσίαση συστήματος, το οποίο, βάσει πληροφοριών για τη θέση των χρηστών και την κατανομή του φόρτου στο δίκτυο, αποστέλλει προτάσεις σε χρήστες που πλήττονται από συμφόρηση, ώστε να μετακινηθούν σε περιοχές όπου η σύνδεση τους θα βελτιωθεί.

β) Προδραστική δέσμευση πόρων σε κινητά δίκτυα, με σκοπό την πρόληψη των αρνητικών συνεπειών των μεταπομπών (Κεφάλαιο 5 και Κεφάλαιο 6): Πρόταση δύο εναλλακτικών μεθόδων για τη προδέσμευση εύρους ζώνης σε γειτονικούς ΣΒ, ώστε το ΚΤ να μπορεί να συνεχίσει απρόσκοπτα τις κλήσεις / συνδέσεις του,

ακόμα και μετά από μία μεταπομπή. Η μία μέθοδος βασίζεται σε διαπραγματεύσεις μεταξύ του ΚΤ και το ΣΒ (Κεφάλαιο 5), ενώ η δεύτερη σε δημοπρασίες μεταξύ των ενδιαφερόμενων ΚΤ για τους πόρους ενός ΣΒ (Κεφάλαιο 6).

γ) Δυναμική πρόσβαση σε μη αδειοδοτημένο φάσμα (Κεφάλαιο 7): Μοντελοποίηση και μελέτη του προβλήματος του διαμοιρασμού ζωνών συχνοτήτων CDMA από μη αδειοδοτημένους ΚΤ και πρόταση κατανεμημένων αλγορίθμων για σύγκλιση σε ισορροπίες, και

δ) Πρόσβαση στο φυσικό μέσο, μέσω ελέγχου ισχύος (Κεφάλαιο 8): Μοντελοποίηση και μελέτη του προβλήματος ελέγχου ισχύος σε CDMA δίκτυα. Επίλυση του προβλήματος, με χρήση εννοιών από τη συνεργατική Θεωρία Παιγνίων, για την επίτευξη αποτελεσματικών και δίκαιων σημείων λειτουργίας.

Στη δεύτερη θεματική ενότητα, η οποία αφορά στη διαχείριση κατανεμημένων συστημάτων, η μελέτη επικεντρώθηκε σε κατανεμημένα συστήματα caching. Βασικό στοιχείο της μελέτης αποτέλεσε η προσπάθεια αποφυγής φαινομένων μονοπώλησης της χωρητικότητας συστημάτων caching από «επιθετικούς» χρήστες/κόμβους, ώστε η υπηρεσία να παρέχεται με δίκαιο τρόπο. Συγκεκριμένα, εξετάστηκαν τα ακόλουθα προβλήματα:

α) Κατανεμημένος έλεγχος στο caching επιπέδου-1 (Κεφάλαιο 9): Μοντελοποίηση και μελέτη του προβλήματος caching επιπέδου-1. Πρόταση κατανεμημένου αλγορίθμου για το δίκαιο διαμοιρασμό της αποθηκευτικής χωρητικότητας στους χρήστες.

β) Οικονομικό πλαίσιο για τη διαχείριση χωρητικότητας caching επιπέδου-2 (Κεφάλαιο 10): Μοντελοποίηση και μελέτη του προβλήματος ως μία μονοπωλιακή αγορά και πρόταση αλγορίθμων για τον υπολογισμό των βέλτιστων εκχωρήσεων χωρητικότητας.

Η χρήση της Θεωρίας Παιγνίων αποτέλεσε, όπως έχει αναφερθεί και πιο πάνω, το βασικό εργαλείο για τη μοντελοποίηση των προβλημάτων που μελετήθηκαν (διαχείριση πόρων στον κινητό υπολογισμό και σε κατανεμημένα συστήματα).

Αξίζει να σημειωθεί ότι η εφαρμογή της οδήγησε σε αξιόλογα αποτελέσματα, επιτρέποντας τη σαφή μοντελοποίηση προβλημάτων διαχείρισης πόρων, στα οποία οι ενδιαφερόμενοι χρήστες αλληλεπιδρώντας επηρεάζουν ο ένας τον άλλο. Τα μαθηματικά μοντέλα που ορίστηκαν, οδήγησαν με τη σειρά τους σε σαφείς στόχους βελτιστοποίησης, σε ό,τι αφορά τη λειτουργία του εκάστοτε συστήματος.

Σημαντικό στοιχείο της παιγνιοθεωρητικής ανάλυσης αποτελεί το γεγονός ότι γίνεται η βασική υπόθεση ότι οι χρήστες είναι ανταγωνιστές, με σκοπό την προσωπική τους και μόνον ωφέλεια, λαμβάνοντας αποφάσεις ατομικά. Έτσι, οι αλγόριθμοι που τελικά προτάθηκαν είναι στην πλειοψηφία τους καταναμημένοι, χωρίς την ανάγκη κεντρικού ελέγχου. Μπορεί, λοιπόν, να προκύψει το συμπέρασμα ότι η Θεωρία παιγνίων αποτελεί ένα σημαντικό εργαλείο για τη μελέτη διάφορων προβλημάτων στο χώρο του κινητού υπολογισμού και των καταναμημένων υπολογιστικών συστημάτων, καθιστώντας δυνατή τη δημιουργία καταναμημένων αλγορίθμων για την έξυπνη κι αποτελεσματική λειτουργία τους.

11.2 Μελλοντικές Ερευνητικές Κατευθύνσεις

Μελλοντικές κατευθύνσεις της έρευνας που διενεργήθηκε στα πλαίσια της παρούσας διατριβής μπορούν να προκύψουν με αφετηρία κάθε ένα από τα επιμέρους προβλήματα που μελετήθηκαν. Έτσι, σε ό,τι αφορά το πρόβλημα εξισορρόπησης του πληθυσμού των ασύρματων χρηστών που μελετήθηκε στο Κεφάλαιο 4, μία πιθανή κατεύθυνση έρευνας θα μπορούσε να είναι ο συνδυασμός του προτεινόμενου συστήματος (MITOS) με συστήματα χρέωσης για την τόνωση του κινήτρου των χρηστών για μετακίνηση. Συγκεκριμένα, θα μπορούσαν να εξετασθούν οι συνθήκες κάτω από τις οποίες ένας πάροχος ασύρματης πρόσβασης θα μπορούσε να αποφύγει το κόστος εγκατάστασης περαιτέρω εξοπλισμού, επιτυγχάνοντας να πείσει ένα μέρος των χρηστών να μετακινηθούν μέσω κάποιας προσφοράς. Θα μπορούσε, δηλαδή, να εκπονηθεί μία μελέτη οικονομικού χαρακτήρα που να λαμβάνει υπόψη παραμέτρους, όπως συμπεριφορά χρηστών, μέγεθος προσφορών κι ανταπόκριση από τους χρήστες, ικανοποίηση χρηστών κι αντίκτυπο στα τελικά κέρδη. Με την κατασκευή ενός κατάλληλου μαθηματικού μοντέλου θα ήταν δυνατόν να μελετηθούν οι παράμετροι

μίας τέτοιας αγοράς και να προταθούν στρατηγικές που πιθανώς να οδηγούν σε αύξηση των κερδών του παρόχου και, ταυτόχρονα, στην ικανοποίηση των κινητών χρηστών.

Σε ό,τι αφορά τη μελέτη της προδραστικής διαχείρισης πόρων σε ασύρματα δίκτυα που περιγράφηκε στο Κεφάλαιο 5 και το Κεφάλαιο 6, ενδιαφέρον θα είχε η εφαρμογή και προσαρμογή των προτεινόμενων αλγορίθμων με σκοπό τη διαχείριση πόρων, εκτός του εύρους ζώνης, όπως υπολογιστικοί πόροι, μνήμη και αποθηκευτική χωρητικότητα. Απώτερος σκοπός αυτής της μελέτης θα μπορούσε να ήταν η ανάπτυξη ενός ενιαίου πλαισίου για την προδραστική διαχείριση των πόρων (εύρος ζώνης, υπολογιστικοί πόροι, μνήμη, κτλ) σε κινητά δίκτυα. Ενδιαφέρον, επίσης, παρουσιάζει και η έρευνα σε ό,τι αφορά την απαραίτητη αλγοριθμική υποδομή και ευφυΐα για να υποστηριχθεί η προδραστική διαχείριση πόρων. Βασικό συστατικό αποτελεί, για παράδειγμα, η δυνατότητα πρόβλεψης θέσης. Στο Κεφάλαιο 6, προτάθηκε ένας ΑΠΚ ο οποίος χρησιμοποιήθηκε για τη στοχευμένη δέσμευση πόρων στην γειτονική δικτυακή υποδομή. Η περαιτέρω μελέτη του εν λόγω αλγορίθμου θα μπορούσε να προσφέρει στοιχεία για την αξιολόγηση του, αλλά και τη βελτίωση του.

Σε σχέση με τη μελέτη της δυναμικής φασματικής πρόσβασης που παρουσιάστηκε στο Κεφάλαιο 7, ενδιαφέρουσα μελλοντική κατεύθυνση θα αποτελούσε η εξέταση επιπλέον κατανεμημένων αλγορίθμων επιλογής ζώνης συχνοτήτων. Για παράδειγμα, θα μπορούσαν να μελετηθούν σχήματα βασισμένα στις δημοπρασίες, όπου οι ζώνες συχνοτήτων θα εκχωρούνται βάσει των προσφορών που υποβάλλουν τα ενδιαφερόμενα τερματικά. Επίσης, θα μπορούσαν να μελετηθούν σχήματα βασισμένα στην Εξελικτική Θεωρία Παιγνίων, όπου τα τερματικά θα μπορούν να επιλέγουν ζώνη συχνοτήτων με τη βοήθεια μαθησιακών αλγορίθμων, χωρίς την ανάγκη πλήρους πληροφόρησης από το ΣΒ για τις συνθήκες κάθε ζώνης συχνοτήτων. Μία, επίσης, παράμετρος η οποία δεν εξετάστηκε, στην παρούσα μελέτη, αποτελεί η κινητικότητα των χρηστών και το πώς μπορεί να επηρεάσει τη σταθερότητα των αναθέσεων ζωνών συχνοτήτων ισορροπίας.

Σε ό,τι αφορά την πρόσβαση σε CDMA δίκτυα μέσω ελέγχου ισχύος που μελετήθηκε κυρίως στο Κεφάλαιο 8, μελλοντικά, θα είχε ενδιαφέρον η μελέτη να επεκταθεί σε περιβάλλοντα με περισσότερες της μίας κυψέλης, με σκοπό την

εφαρμογή της λύσης διαπραγμάτευσης του Nash. Επίσης, θα ήταν ενδιαφέρον να εφαρμοστούν κι άλλα σχήματα επίλυσης συνεργατικών παιγνίων, πέραν της λύσης διαπραγμάτευσης του Nash, όπως, για παράδειγμα, το Shapley value και να συγκριθούν τα οφέλη από τη συνεργασία στον έλεγχο ισχύος, μεταξύ των διαφόρων λύσεων.

Σε σχέση με το πρόβλημα της διαχείρισης αποθηκευτικής χωρητικότητας στον Ιστό και, συγκεκριμένα, με το πρόβλημα caching που μελετήθηκε στο Κεφάλαιο 9 και Κεφάλαιο 10, μελλοντικά, θα ήταν ενδιαφέρουσα η μελέτη με σκοπό την πρόταση ενός ενιαίου πλαισίου διαχείρισης. Συγκεκριμένα, το εν λόγω σχήμα θα μπορούσε να είναι ιεραρχικό, λειτουργώντας, για παράδειγμα, ως μία επέκταση του Οικονομικού σχήματος που προτάθηκε στο Κεφάλαιο 10. Επίσης, ενδιαφέρον θα ήταν, εκτός από την μονοπωλιακή αγορά που μελετήθηκε, το μοντέλο να επεκταθεί και για αγορές ολιγοπωλιακού χαρακτήρα, όπου μία cache ενός επιπέδου θα μπορεί να εξυπηρετηθεί από περισσότερες από μία cache του υψηλότερου επιπέδου, στην οποία περίπτωση ανακύπτουν ζητήματα ανταγωνισμού μεταξύ των cache στο υψηλότερο επίπεδο. Επιπλέον, θα μπορούσαν να μελετηθούν σχήματα διαχείρισης χωρητικότητας αποθήκευσης με βάση δημοπρασίες, όπου οι κόμβοι ενός επιπέδου κάνουν προσφορές σε κόμβους του υψηλότερου επιπέδου παίρνουν αντίστοιχους πόρους.

Σε ό,τι αφορά τον Ιστό, άλλα προβλήματα, εκτός του caching, αποτελεί το pre-fetching περιεχομένου, αλλά και τα CDN. Σε αντίθεση με τις ιεραρχίες caching, ένα CDN αποτελείται, κατά κανόνα, από ένα σύνολο διακομιστών οι οποίοι, κατ' απαίτηση του πηγαίου διακομιστή περιεχομένου, αποθηκεύουν αντίγραφα του πρωτότυπου περιεχομένου. Τέτοιοι διακομιστές μπορεί είναι διασκορπισμένοι σε όλον τον κόσμο, ούτως ώστε η απόσταση του WWW περιεχομένου να είναι όσο το δυνατόν πιο μικρή για κάθε χρήστη σε οποιοδήποτε μέρος κι αν βρίσκεται. Ο αντικειμενικός σκοπός που εξυπηρετεί ένα CDN είναι, με άλλα λόγια, να ελαχιστοποιήσει το χρόνο παράδοσης των ζητούμενων περιεχομένων στους χρήστες, ώστε οι τελευταίοι να μένουν όσο το δυνατόν πιο ευχαριστημένοι.

Στα CDN, ανακύπτουν ένα σύνολο από ζητήματα, όταν κανείς εξετάζει δυνατές επιλογές με σκοπό την τελική ελαχιστοποίηση της καθυστέρησης που αντιλαμβάνεται ένας χρήστης. Τα βασικότερα ζητήματα είναι τα ακόλουθα:

- 1) δεδομένου ενός συνόλου διακομιστών, που θα πρέπει αυτοί να τοποθετηθούν,
- 2) ποιος διακομιστής θα πάρει ποια αντίγραφα από τον πηγαίο διακομιστή WWW περιεχομένου,
- 3) δεδομένου ότι ένα αντίγραφο μπορεί να βρίσκεται σε πάνω από έναν διακομιστή του CDN, από ποιον να ζητήσει ένας χρήστης το αντικείμενο που επιθυμεί,
- 4) αν τα αντικείμενα που είναι αποθηκευμένα σε έναν διακομιστή του CDN είναι έγκυρα ή θα πρέπει να αντικατασταθούν από κάποια άλλα νεότερης έκδοσης (data consistency).

Με αφετηρία την παρούσα διατριβή, μελλοντικά, η μελέτη μπορεί να εστιαστεί στο θέμα (3), δηλαδή στην επιλογή του διακομιστή του CDN από τον οποίο θα ζητηθεί ένα αντικείμενο. Το προαναφερθέν ζήτημα μπορεί να εξεταστεί ανεξάρτητα από τα υπόλοιπα τρία, θεωρώντας, συγκεκριμένα, ότι οι θέσεις των διακομιστών, όπως κι η κατανομή αντιγράφων σε αυτούς, είναι δεδομένες. Η από κοινού μελέτη και των τεσσάρων προβλημάτων δεν είναι αναγκαία, καθώς μπορούν να θεωρηθούν συμπληρωματικά μεταξύ τους. Η επιλογή της θέσης των διακομιστών είναι μία διαδικασία που δε λαμβάνει χώρα συχνά, αν όχι μόνο μία φορά. Η δε κατανομή αντιγράφων στους διακομιστές του CDN μπορεί να πραγματοποιηθεί περισσότερες από μία φορές, αλλά και πάλι, είναι μία διαδικασία με μάλλον περιορισμένη συχνότητα. Αντίθετα, η επιλογή διακομιστή για τη διεκπεραίωση μίας αίτησης ενός χρήστη είναι μία διαδικασία με σχετικά μεγάλη συχνότητα.

Η επιλογή του διακομιστή CDN από έναν χρήστη ανά αίτηση, αποτελεί μία απόφαση για την οποία χρειάζονται να ληφθούν υπόψη κυρίως δύο πληροφορίες: 1) ο φόρτος του κάθε διακομιστή που μπορεί να του παραδώσει το αιτούμενο αντίγραφο, και 2) η δικτυακή απόσταση του χρήστη από το διακομιστή στον οποίο εξετάζει να προωθήσει την αίτηση. Κι οι δύο προηγούμενοι παράγοντες επηρεάζουν το χρόνο που θα χρειαστεί από τη στιγμή που ο χρήστης θα κάνει την αίτηση μέχρι να λάβει το ζητηθέν αντικείμενο.

Με βάση τα αποτελέσματα της παρούσας διατριβής, θα ήταν δυνατόν για την επιλογή διακομιστή CDN από τελικούς χρήστες ή από πληρεξούσιους (proxies)

δικτύων να χρησιμοποιηθούν στρατηγικές, παρόμοιες με αυτές που χρησιμοποιήθηκαν για το πρόβλημα εξισορρόπησης πληθυσμού στο Κεφάλαιο 4 ή για την επιλογή ζώνης συχνοτήτων στο Κεφάλαιο 7. Μπορούν, επίσης, να μελετηθούν και πλήρως κατανεμημένες στρατηγικές, χωρίς την ύπαρξη ρητής/σαφούς ενημέρωσης από το CDN, σε ό,τι αφορά την κατανομή φόρτου στο σύνολο των διακομιστών του, βασισμένες στην Εξελικτική Θεωρία Παιγνίων.

ΠΑΡΑΡΤΗΜΑ Ι

ΣΧΗΜΑ ΠΟΛΛΑΠΛΩΝ ΔΗΜΟΠΡΑΣΙΩΝ

Σε αυτό το παράρτημα, παρέχεται μία λεπτομερής περιγραφή του αλγορίθμου που συνοψίζεται στο Κεφάλαιο 4 (υποκεφάλαιο 4.4), δηλαδή του σχήματος πολλαπλών δημοπρασιών που υποστηρίζει τη διαδικασία δημιουργίας των RP στο MITOS. Έστω $U = \{1, 2, \dots, V\}$ το σύνολο των χρηστών στο WLAN. Έστω, ακόμα, A το σύνολο των χρηστών που πλήττονται από συμφόρηση μία δεδομένη στιγμή, L_j η προσφερόμενη κίνηση στο AP_j και $ServAP_i$ το AP στο οποίο είναι συνδεδεμένος ο χρήστης i :

$$A = \{i \in U : L_j \geq Threshold, ServAP_i = AP_j, j = \{1, 2, \dots, N\}\}$$

Ως PR_j ορίζεται το σύνολο των χρηστών που προτιμούν το AP_j και $pref_i$ το AP που προτιμά ο χρήστης i :

$$PR_j = \{i \in U : pref_i = AP_j, j \in \{1, 2, \dots, N\}\}$$

Έστω bid_i η προσφορά του χρήστη i για το AP που προτιμά. Έστω, επίσης, M_j το σύνολο των χρηστών που ανταγωνίζονται για το AP_j και κάνουν την υψηλότερη προσφορά:

$$M_j = \left\{ i \in PR_j : bid_i = \max_{k \in PR_j} \{bid_k\} \right\}, j \in \{1, 2, \dots, N\}$$

Σε περίπτωση ισοπαλίας (δηλαδή αν το M_j δεν αποτελείται από ένα και μόνο στοιχείο), ο νικητής w_j επιλέγεται τυχαία (ανεξάρτητα και με ίση πιθανότητα) μεταξύ των χρηστών που ανήκουν στο σύνολο M_j .

- | |
|--|
| 1. $U := \{1, 2, \dots, V\}$ |
| 2. $\beta_i := \beta^0, i \in U, \beta^0 \in [0, 1]$ |

Εικόνα Ι.1. Αρχικοποίηση του συστήματος

- | |
|---|
| 1. $cost_j := c^0, j \in \{1, 2, \dots, N\}, c^0 > 0$ |
| 2. $money_i := m^0, i \in U, m^0 > c^0$ |

Εικόνα Ι.2. Αρχικοποίηση πριν από μία διαδικασία πολλαπλών δημοπρασιών

Η Εικόνα I.1 περιγράφει τη διαδικασία της επανεκκίνησης του συστήματος, η οποία πραγματοποιείται ανά τακτά χρονικά διαστήματα, κάθε T_{init} , ώστε να ενημερώνεται το σύνολο των χρηστών και να επαναφέρονται οι παράγοντες προσφοράς (bidding factors) στις αρχικές τιμές τους [38]. Στις προσομοιώσεις, $T_{init} = 60 \text{ sec}$. Η Εικόνα I.2 συνοψίζει τη διαδικασία της αρχικοποίησης που πραγματοποιείται, πριν από την πολλαπλή δημοπρασία, δηλαδή κάθε φορά που ανιχνεύεται συμφόρηση σε κάποιο σημείο του WLAN. Στο βήμα 1, σε κάθε AP τίθεται ένα αρχικό κόστος και, στο βήμα 2, κάθε χρήστης παίρνει ένα αρχικό ποσό χρημάτων για να συμμετάσχει στη διαδικασία των πολλαπλών δημοπρασιών για το συγκεκριμένο συμβάν συμφόρησης.

1. $A := \{i \in U : L_j \geq \text{Threshold}, \text{ServAP}_i = AP_j, j \in \{1, 2, \dots, N\}\}$
2. $PR_j := \{i \in A : \text{pref}_i = AP_j\}, j \in \{1, 2, \dots, N\}$
3. $\text{bid}_i := \text{cost}_i + \beta_i \cdot (\text{money}_i - \text{cost}_i), i \in U, AP_j = \text{pref}_i$
Pick the winner $w_j \in M_j$,
4. $M_j := \left\{ i \in PR_j : \text{bid}_i = \max_{k \in PR_j} \{\text{bid}_k\} \right\}, j \in \{1, 2, \dots, N\}$
5. $A := A - W, W = \bigcup_{j=1}^N \{w_j\}$
6. $L_j := L_j + \text{rate}_{w_j}, j \in \{1, 2, \dots, N\}$
7. $L_{\text{ServAP}_i} := L_{\text{ServAP}_i} - \text{rate}_i, i \in W$
8. $\text{cost}_j := \text{bid}_{w_j}, j \in \{1, 2, \dots, N\}$
9. $\beta_i := \begin{cases} \max(0, \beta_i - \delta), & i \in W \\ \min(1, C_i + \delta), & i \notin W \end{cases}, i \in U, \delta \in [0, 1]$
10. $RP_i := \text{pref}_i, i \in W$
11. If $A \neq \emptyset$ go to 2.
12. Finish

Εικόνα I.3. Πολλαπλή δημοπρασία για τη δημιουργία RP

Το αρχικό ποσό χρημάτων m^0 , καθώς και το αρχικό κόστος του AP c^0 , επιλέγονται αυθαίρετα. Εφόσον όλοι οι χρήστες διαθέτουν το ίδιο ποσό χρημάτων και τα AP έχουν το ίδιο κόστος, το αποτέλεσμα της δημοπρασίας είναι πάντοτε το ίδιο, ανεξάρτητα από τις απόλυτες τιμές που ανατίθενται στις αρχικές ποσότητες. Στις προσομοιώσεις, χρησιμοποιήθηκαν οι επόμενες τιμές παραμέτρων: $\beta^0 = 0.5$, $\delta = 0.05$, $c^0 = 20$ και $m^0 = 200$.

Στην Εικόνα 1.3, στο βήμα 1, δημιουργείται το σύνολο των χρηστών που πλήττονται από συμφόρηση. Στο βήμα 2, οι εν λόγω χρήστες ομαδοποιούνται με βάση το AP που προτιμούν (με βάση την εξίσωση (4.1)). Στο βήμα 3, κάθε χρήστης κάνει μία προσφορά για το AP που προτιμά. Στο βήμα 4, οι προσφορές που υποβλήθηκαν αξιολογούνται κι επιλέγονται οι νικητές. Στην περίπτωση ισοπαλίας, ο νικητής επιλέγεται τυχαία ανάμεσα στους χρήστες με την ίδια προσφορά. Στο βήμα 5, οι νικητές αφαιρούνται από το σύνολο των χρηστών που θεωρούνται πληττόμενοι από τη συμφόρηση. Στα βήματα 6 και 7, η τρέχουσα κίνηση που παράγουν οι χρήστες που έχουν νικήσει μεταφέρεται εικονικά από τα τρέχοντα AP που είναι συνδεδεμένοι, στα AP που έχουν «αγοράσει» (το *rate*, είναι ο φόρτος²⁸ που παράγει ο χρήστης i) – το μοντέλο αυτό είναι ευαίσθητο στο φόρτο του κάθε χρηστή ατομικά. Ως εκ τούτου, οι προτιμήσεις των υπολειπόμενων πληττομένων από τη συμφόρηση χρηστών (βάσει της εξίσωσης (4.1)) μπορούν να μεταβληθούν στη διάρκεια της πολλαπλής δημοπρασίας. Στο βήμα 8, το κόστος των AP τίθεται ίσο με τη μέγιστη προσφορά που έγινε για αυτά, δηλαδή το κόστος αλλάζει βάσει της ζήτησης. Στο βήμα 9, οι χρήστες προσαρμόζουν τη στρατηγική προσφορών τους. Συγκεκριμένα, αν έχασαν στη δημοπρασία, αυξάνουν τον παράγοντα προσφοράς τους β_i κατά δ , ώστε να αυξήσουν την πιθανότητα να νικήσουν σε μία μελλοντική δημοπρασία. Αν νίκησαν, μειώνουν τον παράγοντα προσφοράς για να περιορίσουν τα κόστη. Στο βήμα 10, κάθε νικητής παίρνει το τελικό RP, που σχετίζεται με το AP που πέτυχε να κερδίσει. Στο βήμα 11, ο αλγόριθμος ελέγχει αν το σύνολο των χρηστών που πλήττονται από τη συμφόρηση είναι κενό. Αν ναι, ο αλγόριθμος τερματίζει· διαφορετικά, ο αλγόριθμος επιστρέφει στο βήμα 2.

²⁸ Σε μία υλοποίηση του συστήματος MITOS, αυτή η πληροφορία μπορεί εύκολα να ληφθεί τοπικά από τη δικτυακή διεπαφή του τερματικού. Τέτοια πληροφορία φόρτου θεωρείται επαρκώς ακριβής και μπορεί φιλτραριστεί χρονικά για την εξομάλυνση των διακυμάνσεων λόγω ριπών κίνησης.

Ο αλγόριθμος που παρουσιάστηκε στο παρόν παράρτημα μπορεί να χρησιμοποιηθεί και σε ένα πλήρως κατανεμημένο περιβάλλον, όπου πράκτορες, εκ μέρους των χρηστών, θα μπορούσαν να συμμετέχουν στις δημοπρασίες. Σε ένα τέτοιο περιβάλλον, η συμπεριφορά του πράκτορα σε ό,τι αφορά τις προσφορές που κάνει (όπως καθορίζεται από τον παράγοντα προσφοράς) μπορεί να μεταβάλλεται αναλόγως με την κατάσταση του χρήστη. Ειδικότερα, ο χρήστης θα μπορούσε να δηλώσει (1) την σημαντικότητα της νίκης σε μία δημοπρασία (π.χ. θέτοντας απευθείας την τιμή του παράγοντα προσφορών) ή (2) τη σχετική σημασία που έχουν για αυτόν οι διαφορετικοί τύποι εφαρμογών (π.χ. ότι το video είναι λιγότερο σημαντικό από το VoIP, αλλά πιο σημαντικό από το WWW), ούτως ώστε ο πράκτορας να κάνει προσφορές με βάση τις απαιτήσεις των ενεργών εφαρμογών του. Επιπλέον, σε κάθε μία από τις διαφορετικές δημοπρασίες, θα μπορούσε να αλλάζει η τιμή του παράγοντα προσφοράς βάσει του προτιμώμενου AP (π.χ. να κάνει χαμηλές προσφορές για AP που έχει επισκεφθεί πιο νωρίς και φάνηκε πως ο ανταγωνισμός για αυτά είναι μειωμένος).

ΠΑΡΑΡΤΗΜΑ ΙΙ

ΕΠΙΔΡΑΣΗ ΠΛΗΘΟΥΣ ΠΡΟΓΝΩΣΤΩΝ ΑΝΑ ΧΡΗΣΤΗ ΣΤΗΝ ΑΠΟΤΕΛΕΣΜΑΤΙΚΟΤΗΤΑ ΤΟΥ SFBP-MITOS

Σε αυτό το παράρτημα, βάσει του μοντέλου προσομοίωσης που περιγράφηκε στο Κεφάλαιο 4 (υποκεφάλαιο 4.5), παρουσιάζονται αποτελέσματα που αφορούν στην απόδοση του συστήματος SFBP-MITOS, σε σχέση με τον αριθμό K των προγνωστών που εκχωρούνται σε κάθε χρήστη από ένα σύνολο από M διαθέσιμους προγνώστες. Στις προσομοιώσεις, $M = 15$, και δοκιμάστηκαν τιμές για το K από το σύνολο $\{2, 4, \dots, 30\}$. Πρέπει να σημειωθεί ότι το K μπορεί και να είναι μεγαλύτερο του M , καθώς κάθε προγνώστης μπορεί να επιλεγεί από ένα χρήστη περισσότερες από μια φορά. Αυτό σημαίνει ότι το πλήθος των διακριτών προγνωστών ανά χρήστη είναι κατά μέσον όρο μικρότερο από το πλήθος των προγνωστών που έχουν ανατεθεί, K .

Έστω \bar{K} το μέσο πλήθος των διακριτών προγνωστών ανά χρήστη. Τότε,

$$\bar{K} = \sum_{k=1}^{\max\{K, M\}} kP(k, K, M),$$

όπου $P(k, K, M)$ είναι η πιθανότητα ένας χρήστης να έχει πάρει k διακριτούς προγνώστες επιλέγοντας K προγνώστες από ένα σύνολο M διαθέσιμων προγνωστών. Μπορεί να αποδειχθεί ότι αυτή η πιθανότητα υπολογίζεται ως εξής:

$$P(k, K, M) = \binom{M}{k} \frac{\binom{K-1}{K-k}}{\binom{M+K-1}{K}},$$

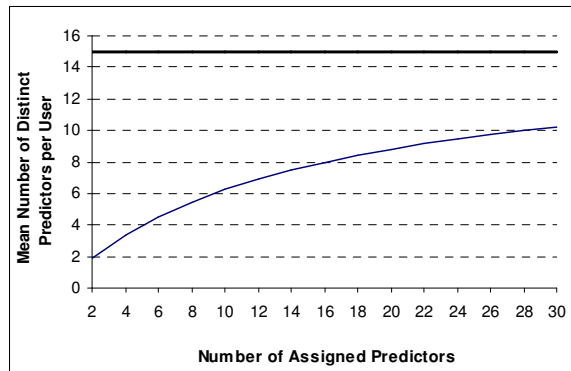
όπου

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Έτσι, το μέσο πλήθος διακριτών προγνωστών ανά χρήστη είναι

$$\bar{K} = \frac{M \cdot K}{M + K - 1} \quad (II.1)$$

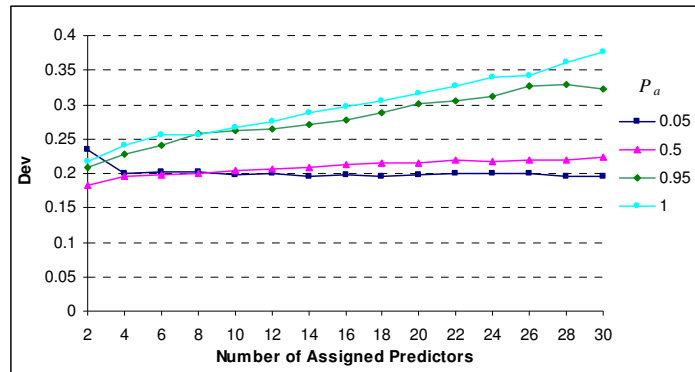
Όπως ήταν αναμενόμενο, από την εξίσωση (II.1) παρατηρείται ότι όσο αυξάνει το K , το μέσο πλήθος διακριτών προγνωστών ανά χρήστη \bar{K} ασυμπτωτικά τείνει στο M . Αυτό φαίνεται και στην Εικόνα II.1, όπου απεικονίζεται το \bar{K} , ως συνάρτηση του K , για $M = 15$.



Εικόνα II.1. Μέσο πλήθος διακριτών προγνωστών ανά χρήστη συναρτήσει του πλήθους των προγνωστών που ανατίθενται ($M = 15$)

Στην Εικόνα II.2, απεικονίζεται η μετρική Den , η οποία είναι κι η βασική μετρική για την αξιολόγηση των ικανοτήτων εξισορρόπησης φόρτου των μελετούμενων συστημάτων (εξίσωση (4.9)). Συγκεκριμένα, παρουσιάζονται οι τιμές του Den για μερικές ενδεικτικές τιμές της P_a , και για κατώφλι συμφόρησης 75%. Όπως ήταν αναμενόμενο, παρατηρείται αυξητική τάση της μετρικής Den , καθώς περισσότεροι προγνώστες ανατίθενται στους χρήστες. Συγκεκριμένα, φαίνεται ότι η κλίση είναι πιο απότομη για μεγαλύτερες τιμές της πιθανότητας αποδοχής των RP. Αυτό είναι λογικό, καθώς, όσο πιο μεγάλη τιμή έχει η P_a , τόσο μεγαλύτερο το πλήθος των χρηστών που αποδέχονται τα RP και τελικά ευάλωτο το σύστημα σε ταλαντώσεις. Επίσης, παρατηρείται ότι, για μέτριες τιμές της P_a (π.χ. 0.5), το σύστημα παραμένει μάλλον ανεπηρέαστο από το πλήθος των προγνωστών που ανατίθενται στους χρήστες. Τέλος, φαίνεται ότι για μικρές τιμές της πιθανότητας αποδοχής των RP (π.χ. 0.05) η απόδοση του συστήματος μπορεί να βελτιωθεί όταν ανατίθενται περισσότεροι προγνώστες στους χρήστες. Αυτό οφείλεται στο γεγονός ότι οι εμπλεκόμενοι χρήστες είναι λίγοι κι επομένως ένα μεγάλο πλήθος από

προγνώστες ανά χρήστη βελτιώνει τις ικανότητες πρόβλεψης των χρηστών, χωρίς, όμως, να υπονομεύει την ετερογένεια και τις επιδόσεις του συστήματος.



Εικόνα II.2. Μετρική Dev σε σχέση με το πλήθος των προγνωστών που ανατίθενται ($M = 15$)

ΠΑΡΑΡΤΗΜΑ ΙΙΙ

ΠΡΟΒΛΕΨΗ ΕΞΩ-ΚΥΨΕΛΙΚΗΣ ΠΑΡΕΜΒΟΛΗΣ

Σε αυτό το παράρτημα, περιγράφεται μία μέθοδος για την πρόβλεψη της εξω-κυψελικής παρεμβολής, πριν από μία υποθετική επιλογή μίας ζώνης συχνοτήτων από ένα ΚΤ. Ένα σημαντικό ζήτημα για ένα ΚΤ, όταν αξιολογεί μία πιθανή αλλαγή ζώνης συχνοτήτων, είναι αν η νέα ζώνη συχνοτήτων θα είναι καλύτερη από την τρέχουσα ζώνη συχνοτήτων. Συγκεκριμένα, το ΚΤ πρέπει να ελέγξει την κατάσταση της ζώνης συχνοτήτων που εξετάζει, αλλά, επίσης, να λάβει υπόψη τη δική του συνεισφορά, μετά από μία υποθετική επιλογή της ζώνης συχνοτήτων: το ΚΤ πρέπει να προβλέψει τη συνεισφορά του στο φόρτο του ΣΒ ($I'_{c,a}$), αλλά και την εξω-κυψελική παρεμβολή ($\rho^2_{c,a}$), η οποία θα αυξηθεί, όταν το ΚΤ εισέλθει στην εν λόγω ζώνη συχνοτήτων. Ο φόρτος $I_{c,a}$ μπορεί να υπολογιστεί ευθέως, όπως υποδεικνύεται στην εξίσωση (7.13). Η εξω-κυψελική παρεμβολή $\rho^2_{c,a}$, όμως, είναι, μάλλον, δύσκολο να υπολογιστεί από το ΚΤ.

Το ΚΤ θα μπορούσε, αναλυτικά, να καθορίσει την εξω-κυψελική παρεμβολή, αν το επιδιωκόμενο SIR κάθε ΚΤ, καθώς κι οι απώλειες σήματος από κάθε ΚΤ σε κάθε ΣΒ, ήταν γνωστά. Όμως, αυτή η πληροφορία είναι, μάλλον, δύσκολο να αποκτηθεί. Ειδικότερα, οι απώλειες σήματος για όλα τα N ΚΤ προς όλους τους B ΣΒ είναι ένας πίνακας $N \times B$ τα στοιχεία του οποίου αλλάζουν πολύ γρήγορα. Έτσι, η αναγκαία πληροφορία είναι αρκετά δύσκολο να διαδοθεί στο δίκτυο, καθώς απαιτεί την εκπομπή μεγάλου ποσού δεδομένων. Επιπλέον, και να αποκτηθεί η εν λόγω πληροφορία χάνει γρήγορα την αξία της. Έτσι, κρίνεται πιο κατάλληλο να υπολογιστεί προσεγγιστικά, ή να προβλεφθεί, βάσει μόνο τοπικής πληροφορίας.

Στο παρόν παράρτημα, προτείνεται μία προσεγγιστική πρόβλεψη της εξω-κυψελικής παρεμβολής από το ΚΤ, βάσει τοπικής πληροφορίας (παρεχόμενη από το ΣΒ). Έστω ένα απλό παράδειγμα με δύο ΣΒ, δηλαδή $I = \{1,2\}$ και N ΚΤ, δηλαδή $J = \{1, \dots, N\}$. Έστω ότι τα ΚΤ από 1 ως $N-1$ είναι συνδεδεμένα με το ΣΒ 1, ενώ το ΚΤ N είναι συνδεδεμένο με το ΣΒ 2. Τότε, μπορεί να υπολογιστεί η ισχύς εκπομπής των ΚΤ, ως εξής:

$$\rho_j = \begin{cases} \frac{1}{h_{j,1}} \frac{\sigma^2 + h_{N,1} \rho_N}{\left(\frac{W}{R\gamma_j^*} + 1\right) \left(1 - \sum_{k=1}^{N-1} \left(\frac{W}{R\gamma_k^*} + 1\right)^{-1}\right)}, \quad \forall j \in \{1, \dots, N-1\} \\ \frac{1}{h_{j,1}} \sigma^2 \frac{1 + \left[\sum_{k=1}^{N-1} \frac{h_{k,2}}{h_{k,1}} \left(\frac{W}{R\gamma_k^*} + 1\right)^{-1}\right] \times \frac{1}{1 - l_{c,1}}}{\frac{W}{R\gamma_j^*} \frac{h_{j,2}}{h_{j,1}} - \left[\sum_{k=1}^{N-1} \frac{h_{k,2}}{h_{k,1}} \left(\frac{W}{R\gamma_k^*} + 1\right)^{-1}\right] \times \frac{1}{1 - l_{c,1}}}, \quad j = N \end{cases} \quad (III.1)$$

Η ισχύς του ΚΤ N γίνεται αντιληπτή ως εξω-κυβελική παρεμβολή από τα ΚΤ 1 ως $N-1$. Συγκεκριμένα, το ΚΤ N συνεισφέρει με ένα συστατικό παρεμβολής στο ΣΒ 1 που είναι ίσο με $h_{N,1} \rho_N$. Όπως μπορεί να παρατηρηθεί, αυτή η παρεμβολή εξαρτάται από το φόρτο που παρέχεται στο ΣΒ 1 ($l_{c,1}$), και, ειδικότερα, από τον παράγοντα φόρτου $1/(1 - l_{c,1})$.

Χωρίς βλάβη της γενικότητας, γίνεται η υπόθεση ότι $\frac{h_{1,2}}{h_{1,1}} \leq \frac{h_{2,2}}{h_{2,1}} \leq \dots \leq \frac{h_{N-1,2}}{h_{N-1,1}}$. Έτσι,

είναι δυνατόν να φραχθεί η ισχύς εκπομπής ρ_N του ΚΤ N , ως εξής:

$$\rho_N \leq \frac{1}{h_{N,1}} \sigma^2 \frac{1 + \left[\frac{h_{N-1,2}}{h_{N-1,1}} \sum_{k=1}^{N-1} \left(\frac{W}{R\gamma_k^*} + 1\right)^{-1}\right] \times \frac{1}{1 - l_{c,1}}}{\frac{W}{R\gamma_N} \frac{h_{N,2}}{h_{N,1}} - \left[\frac{h_{N-1,2}}{h_{N-1,1}} \sum_{k=1}^{N-1} \left(\frac{W}{R\gamma_k^*} + 1\right)^{-1}\right] \times \frac{1}{1 - l_{c,1}}}, \quad \eta$$

$$\rho_N \leq \frac{1}{h_{N,1}} \sigma^2 \frac{1 + \left[\frac{h_{N-1,2}}{h_{N-1,1}} l_{c,1}\right] \times \frac{1}{1 - l_{c,1}}}{\frac{W}{R\gamma_N} \frac{h_{N,2}}{h_{N,1}} - \left[\frac{h_{N-1,2}}{h_{N-1,1}} l_{c,1}\right] \times \frac{1}{1 - l_{c,1}}}$$

Επίσης, καθώς $l_{c,1} \leq 1$, προκύπτει, τελικά, ότι

$$\rho_N \leq \frac{1}{h_{N,1}} \sigma^2 \frac{1 + \frac{h_{N-1,2}}{h_{N-1,1}} \frac{1}{1 - I_{c,1}}}{\frac{W}{R\gamma_N} \frac{h_{N,2}}{h_{N,1}} - \frac{h_{N-1,2}}{h_{N-1,1}} \frac{1}{1 - I_{c,1}}}. \quad (\text{III.2})$$

Έτσι, παρατηρείται ότι η παρεμβολή στο ΣΒ 1, λόγω ενός ΚΤ που είναι συνδεδεμένο στο ΣΒ 2 (δηλαδή, η ισχύς $h_{N,1}\rho_N$) εξαρτάται από τον παράγοντα φόρτου $I_{c,1}$ του ΣΒ 1 (συγκεκριμένα, είναι αύξουσα ως προς αυτόν). Παρατηρείται, επίσης, ότι εξαρτάται από το επιδιωκόμενο (NE) SIR του ΚΤ και τις απώλειες σήματος από το ΚΤ προς τους δύο ΣΒ. Πρέπει να σημειωθεί, από την (III.2), ότι όλοι οι παράγοντες, εκτός του φόρτου $I_{c,1}$ και του σ^2 , δεν είναι γνωστοί στο ΣΒ 1 και, κατά συνέπεια, σε όλα τα συνδεδεμένα σε αυτόν ΚΤ. Ένα ΚΤ που εξετάζει μία υποθετική μετακίνηση σε μία ζώνη συχνοτήτων c , είναι σε θέση να γνωρίζει μόνο τον τοπικό φόρτο (μετά τη μετακίνηση) $I'_{c,1} = I_{c,1} + (W/R\gamma_j^* + 1)^{-1}$ και το σ^2 . Με άλλα λόγια, δε μπορεί να υπολογίσει με ακρίβεια την εξω-κυψελική παρεμβολή $h_{N,1}\rho'_N$ που θα προκύψει, λόγω της μετακίνησης. Για το λόγο αυτό, χρησιμοποιείται μία γραμμική, προσεγγιστική πρόβλεψη της παρεμβολής από ένα ΚΤ διαφορετικής κυψέλης, βάσει μόνον του τοπικού φόρτου $I_{c,a}$. Συγκεκριμένα, η πρόβλεψη της παρεμβολής $h_{N,1}\rho'_N$ μπορεί να υπολογιστεί (μέσω παρεκβολής) από τον τοπικό φόρτο $I_{c,a}$ και την τρέχουσα εξω-κυψελική παρεμβολή $\rho^2_{c,a}$, ως εξής:

$$h_{N,1}\rho'_N = a \frac{1}{1 - I'_{c,1}}, \quad (\text{III.3})$$

$$a = h_{N,1}\rho_N \left(\frac{1}{1 - I_{c,1}} \right)^{-1}$$

Με άλλα λόγια, το ΚΤ υπολογίζει την κλίση a της ευθείας που θα χρησιμοποιηθεί ως η συνάρτηση πρόβλεψης, ούτως ώστε η προκύπτουσα ευθεία να περνά από τρέχουσα παρατηρούμενη εξω-κυψελική παρεμβολή, για τον τρέχοντα παράγοντα φόρτου $I_{c,a}$. Στη συνέχεια, υπολογίζει το φόρτο $I'_{c,a}$ συμπεριλαμβάνοντας τη δική του συνεισφορά και, τελικά, κάνει την πρόβλεψη της εξω-κυψελικής παρεμβολής, βάσει του φόρτου $I'_{c,a}$ μετά την αλλαγή ζώνης συχνοτήτων. Αν η ατομική

συνεισφορά φόρτου του ΚΤ είναι μικρή σε σχέση με το συνολικό φόρτο, η πρόβλεψη μπορεί να είναι αρκετά ακριβής.

Η μέθοδος πρόβλεψης που περιγράφηκε πιο πάνω μπορεί να γενικευθεί για περισσότερα του ενός ΚΤ (συνδεδεμένα σε άλλους ΣΒ). Με άλλα λόγια, είναι δυνατόν, με παρόμοιο τρόπο, να προβλεφθεί η εξω-κυψελική παρεμβολή $\rho^2_{c,a}$, βάσει της τρέχουσας τιμής της και του τρέχοντος φόρτου, και να επιτευχθεί μία αρκετά ακριβής προσέγγιση.

ΠΑΡΑΡΤΗΜΑ IV

ΒΕΛΤΙΣΤΗ ΧΡΗΣΗ ΦΑΣΜΑΤΟΣ

Σε αυτό το παράτημα, παρατίθεται μία μελέτη για το πώς μπορεί να επιτευχθεί η βέλτιστη χρήση του φάσματος σε ένα περιβάλλον CDMA. Πρέπει να σημειωθεί ότι η βελτιστότητα είναι μία μάλλον ασαφής κι υποκειμενική έννοια. Βασίζεται στους αντικειμενικούς σκοπούς που θέτει ο σχεδιαστής του συστήματος προς βελτιστοποίηση. Αφού καθοριστούν συγκεκριμένοι αντικειμενικοί σκοποί και περιορισμοί, τότε μπορεί να σχεδιαστεί ένα «βέλτιστο» σύστημα. Σε ένα σύστημα CDMA, ο αντικειμενικός σκοπός του σχεδιαστή θα μπορούσε να είναι η μεγιστοποίηση της ρυθμοαπόδοσης, βάσει περιορισμών χωρητικότητας του συστήματος. Ένας εναλλακτικός στόχος θα μπορούσε να είναι η ελαχιστοποίηση της καταναλισκόμενης ενέργειας, με σκοπό τη μεγιστοποίηση του χρόνου ζωής των ΚΤ.

Στην παρούσα μελέτη, θεωρείται η ανταλλαγή μεταξύ της συνολικής ρυθμοαπόδοσης που επιτυγχάνεται και της συνολικής ενέργειας που καταναλώνεται. Συγκεκριμένα, ορίζεται, με σκοπό να μεγιστοποιηθεί, η αντικειμενική συνάρτηση $G(\mathbf{c})$:

$$G(\mathbf{c}) = \frac{R \sum_{j=1}^N \left(1 - e^{-0.5\gamma_j^*}\right)^M}{\sum_{j=1}^N q_j}, \quad (\text{IV.1})$$

όπου \mathbf{c} είναι το διάνυσμα που αναπαριστά την κατανομή ζωνών συχνοτήτων των ΚΤ. Ο αριθμητής αναπαριστά τη ρυθμοαπόδοση του CDMA συστήματος (συνολικός ρυθμός δεδομένων πολλαπλασιασμένος με την πιθανότητα ορθής λήψης). Ο παρονομαστής αναπαριστά τη συνολική κατανάλωση ισχύος.

Στο μη-συνεργατικό παίγνιο ελέγχου ισχύος που περιγράφεται στο Κεφάλαιο 7 (υποκεφάλαιο 7.4), τα SIR των ΚΤ είναι σταθερά, ανεξάρτητα της ζώνης συχνοτήτων που έχει επιλεγεί, όπως φαίνεται και στην (7.5) και την (7.6). Έτσι, ο αριθμητής της συνάρτησης $G(\mathbf{c})$ είναι σταθερός, ως έκφραση σταθερών όρων, και το πρόβλημα της μεγιστοποίησης της $G(\mathbf{c})$ είναι, τελικά, ισοδύναμο με την

ελαχιστοποίηση του παρονομαστή της $G(\mathbf{c})$, δηλαδή τη συνολική ισχύ που απαιτείται για να επιτευχθούν τα προαναφερθέντα (σταθερά) SIR. Έτσι, ορίζεται το ακόλουθο πρόβλημα βελτιστοποίησης, η λύση του οποίου είναι η επιδιωκόμενη βέλτιστη κατανομή του ισχύων.

$$\min_{\mathbf{c} \in \mathcal{C}} \sum_{j=1}^N q_j \equiv \min_{\mathbf{c} \in \mathcal{C}} \sum_{c=1}^K \rho_{c,a}^2 \frac{\sum_{k=1, c_k=c}^N \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1}}{1 - \sum_{k=1, c_k=c}^N \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1}} \equiv \min_{\mathbf{c} \in \mathcal{C}} \sum_{c=1}^K \rho_{c,a}^2 \frac{I_{c,a}}{1 - I_{c,a}}$$

Με άλλα λόγια, παρατηρείται ότι το πρόβλημα της βέλτιστης χρήσης του φάσματος εκφυλίζεται στο πρόβλημα τις βέλτιστης κατανομής του φόρτου των ΚΤ στις υπάρχουσες ζώνες συχνοτήτων, δεδομένης της εξω-κυψελικής παρεμβολής που υπάρχει σε κάθε ζώνη συχνοτήτων. Για την απλοποίηση του προβλήματος, θεωρείται ότι η μεταβλητή $I_{c,a}$ είναι συνεχής. Έτσι, το απλοποιημένο πρόβλημα (ελαχιστοποίηση της συνολικής καταναλισκόμενης ενέργειας για ένα δεδομένο φόρτο στις υπάρχουσες ζώνες συχνοτήτων) παίρνει την ακόλουθη μορφή:

$$(P) \begin{cases} \min_{I_a \in \mathfrak{R}} \sum_{c=1}^K \rho_{c,a}^2 \frac{I_{c,a}}{1 - I_{c,a}}, \\ \text{s.t.} \\ \sum_{c=1}^K I_{c,a} = I_{tot,a}, \text{ and} \\ I_{c,a} \geq 0, \forall c \in \mathcal{C} \end{cases},$$

$$\text{όπου } I_{tot,a} = \sum_{c=1}^K I_{c,a} = \sum_{j=1}^N \left(\frac{W}{R\gamma_j^*} + 1 \right)^{-1}.$$

Λόγω της ύπαρξης περιορισμών, για να λυθεί το πρόβλημα (P), γίνεται η χρήση της θεωρίας των πολλαπλασιαστών Lagrange. Συγκεκριμένα, το σημείο που λύνει το (P), θα πρέπει να μεγιστοποιεί την ακόλουθη συνάρτηση Lagrange:

$$L(I_a, \lambda, \mu) = - \sum_{c=1}^K \rho_{c,a}^2 \frac{I_{c,a}}{1 - I_{c,a}} + \lambda \left(I_{tot,a} - \sum_{c=1}^K I_{c,a} \right) + \sum_{c=1}^K \mu_c I_{c,a} \quad (IV.2)$$

όπου λ και μ_c είναι οι πολλαπλασιαστές Lagrange που σχετίζονται με τον περιορισμό του συνολικού φόρτου και του περιορισμού ότι ο φόρτος σε κάθε ζώνη

συχνοτήτων δε θα πρέπει να είναι αρνητικός, αντίστοιχα. Από τη συνθήκη βελτιστότητας πρώτου βαθμού, δηλαδή, $\partial L(I_a, \lambda, \mu) / \partial I_{k,a} = 0$, προκύπτει ότι

$$\rho_{k,a}^2 \frac{1}{(1 - I_{k,a})^2} - \mu_k = -\lambda. \quad (IV.3)$$

Χωρίς βλάβη της γενικότητας, γίνεται η υπόθεση ότι οι διαθέσιμες ζώνες συχνοτήτων ταξινομούνται, βάσει της αντίστοιχης εξω-κυβελικής παρεμβολής, δηλαδή, $\rho_{1,a}^2 < \dots < \rho_{K',a}^2$. Μπορεί, εύκολα, να δειχθεί ότι, αν, στη λύση, ο περιορισμός της μη-αρνητικότητας είναι ενεργός για κάποιες ζώνες συχνοτήτων (δηλαδή, $I_{k,a} = 0$), αυτές οι ζώνες συχνοτήτων θα ανήκουν στο σύνολο $\{K'+1, \dots, K\}$. Αν υποθεθεί ότι το K' είναι γνωστό, από την (IV.3), ισχύει το παρακάτω, για κάθε $j, k \in \{1, \dots, K'\}$ ($\mu_j = \mu_k = 0$):

$$1 - I_{j,a} = \sqrt{\frac{\rho_{j,a}^2}{\rho_{k,a}^2}} (1 - I_{k,a}) \quad (IV.4)$$

και από τον περιορισμό του συνολικού φόρτου, δηλαδή, $\sum_{c=1}^{K'} I_{c,a} = I_{tot,a}$, ισχύει ότι

$$I_{k,a} = 1 - \frac{\sqrt{\rho_{k,a}^2}}{\sum_{c=1}^{K'} \sqrt{\rho_{c,a}^2}} (K' - I_{tot,a}), \quad k \in \{1, \dots, K'\}. \quad (IV.5)$$

Ως εκ τούτου, η λύση του (P) έχει την ακόλουθη μορφή:

$$I_{k,a}^{opt} = \begin{cases} 1 - \frac{\sqrt{\rho_{k,a}^2}}{\sum_{c=1}^{K'} \sqrt{\rho_{c,a}^2}} (K' - I_{tot,a}), & k \in \{1, \dots, K'\} \\ 0, & k \in \{K'+1, \dots, K\} \end{cases} \quad (IV.6)$$

όπου το K' είναι ο μεγαλύτερος ακέραιος από το σύνολο $\{1, \dots, K\}$, ώστε $I_{k,a} \geq 0$, $\forall k \in \{1, \dots, K\}$. Δηλαδή, οι τιμές $K, K-1, \dots$ δοκιμάζονται διαδοχικά, μέχρι η προκύπτουσα λύση να μην είναι αρνητική. Πρέπει να σημειωθεί, επίσης, ότι πάντα $I_{k,a}^{opt} < 1$. Αυτός είναι κι ο λόγος για τον οποίο δεν συμπεριλαμβάνεται ο αντίστοιχος περιορισμός στον ορισμό του προβλήματος βελτιστοποίησης (P).

ΠΑΡΑΡΤΗΜΑ V

ΑΡΙΘΜΗΤΙΚΟΣ ΑΛΓΟΡΙΘΜΟΣ ΥΠΟΛΟΓΙΣΜΟΥ ΤΟΥ NBS

Σε αυτό το παράρτημα, παρουσιάζονται οι λεπτομέρειες σχετικά με την υλοποίηση του αριθμητικού αλγορίθμου υπολογισμού που περιγράφηκε στο Κεφάλαιο 8 (παράγραφο 8.6.2). Συγκεκριμένα, περιγράφεται μία μέθοδος για τον καθορισμό ενός αρχικού, εφικτού σημείου, βάσει της συζήτησης στο υποκεφάλαιο 8.5, και παρέχονται περισσότερες λεπτομέρειες σε ό,τι αφορά την παραγωγή εφικτών κατευθύνσεων αύξησης (ascent directions), και την επιλογή κατάλληλων βημάτων, που συζητήθηκαν εν συντομία στις παραγράφους 8.6.2.1 και 8.6.2.2, αντίστοιχα.

V.1 Καθορισμός Εφικτού Σημείου Εκκίνησης

Για να εφαρμοστεί η μέθοδος *υπο συνθήκη κλίσης* που περιγράφεται επιγραμματικά στην παράγραφο 8.6.2, είναι αναγκαίο να υπάρχει ένα εφικτό σημείο εκκίνησης $q^{(0)}$. Αν το σύνολο των περιορισμών είναι ένα κυρτό πολύεδρο, είναι εύκολο να παραχθεί ένα εφικτό σημείο εκκίνησης. Όμως, αν το σύνολο των περιορισμών ορίζεται κι από μη γραμμικούς περιορισμούς, όπως το πρόβλημα που εξετάζεται στο Κεφάλαιο 8, τέτοια σημεία είναι, συνήθως, δύσκολο να καθοριστούν.

Προτείνεται μία μέθοδος, η οποία βασίζεται στη συζήτηση του υποκεφαλαίου 8.5. Συγκεκριμένα, αποδείχθηκε εκεί ότι αν υπάρχει έστω κι ένας χρήστης, i , με SIR στο NE που είναι ίσο με γ^* (λύση της (8.3)), τότε, αν όλοι οι άλλοι χρήστες διατηρήσουν του SIR που είχαν στο NE, μία μείωση της ισχύος του χρήστη i θα οδηγήσει σε μία νέα κατανομή ισχύος, όπου όλοι οι χρήστες θα απολαμβάνουν αυξημένη ωφέλεια, σε σχέση με το NE. Στην περίπτωση που κανείς χρήστης δεν έχει SIR στο NE ίσο με γ^* , η διαδικασία εύρεσης των χρηστών, οι οποίοι μειώνοντας την ισχύ τους όλοι οι χρήστες θα επιτύχουν ανώτερη ωφέλεια, είναι πιο πολύπλοκη και μπορεί να μη δώσει πάντα μία εφικτή κατανομή ισχύος. Ο ακόλουθος αλγόριθμος συνοψίζει τη διαδικασία της εύρεσης εφικτού σημείου εκκίνησης:

1. Define $Y := \{j \in I : \gamma_j^* = \gamma^*\}$. If $Y \neq \emptyset$, then set $K := I - Y$, and go to step 6. Otherwise, go to step 2.
2. Define set $K := \{j \in I : \gamma_j^* = \gamma_{m,j}\}$. If $K = I$, then go to step 11, else go to step 3.
3. If $C(\gamma_i^*) < 0, \forall i \in I - K$, then go to step 6, else go to step 4.
4. Set $k = \arg \min_{j \in I - K} \{\gamma_j^*\}$, and set $K := K \cup \{k\}$.
5. If $K \neq I$, then go to step 3, else go to step 11.
6. Set $A := \sum_{k \in K} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1} / \left[1 - \sum_{k \in K} \left(\frac{W}{R\gamma_k^*} + 1 \right)^{-1} \right]$, and

$$\mu_{th} := \max_{i \in I - K} \left\{ (A + 1) \sigma^2 / \left[\left(\frac{W}{R\gamma_{m,i}} - A \right) q_i^* - (A + 1) \sum_{k \in I - K - \{i\}} q_k^* \right] \right\}$$
7. Set $\mu := \mu_{th}$.
8. Set $q_i := \mu q_i^*, \forall i \in I - K$. Update $q_j, \forall j \in K$ accordingly (i.e., adjust q_j such that $\gamma_j = \gamma_j^*, \forall j \in K$).
9. If $v_i(\mathbf{q}) > v_i^0, \forall i \in I$, and $\mathbf{q} \in \Gamma$, then go to step 12
10. Set $\mu := 0.5 + 0.5\mu$. Go to step 8.
11. Finish (Feasible initial point not found). BS returns the NE power allocation \mathbf{q}^* .
12. $\mathbf{q}^{(0)} := \mathbf{q}$. Finish.

Στο βήμα 6, υπολογίζεται ο ελάχιστος παράγοντας μείωσης ισχύος μ_{th} , ο οποίος εξασφαλίζει εφικτές, σε επίπεδο SIR, κατανομές ισχύος και χρησιμοποιείται ως η αρχική τιμή του μ , στο βήμα 7. Στο βήμα 10, η αναδρομή $\mu := 0.5 + 0.5\mu$ παράγει μία ακολουθία παραγόντων μείωσης ισχύος που ασυμπτωτικά συγκλίνει στη μονάδα. Ο αλγόριθμος τερματίζει, είτε στο βήμα 11, αν δε βρέθηκε εφικτό σημείο εκκίνησης, ή, με επιτυχία, στο βήμα 12, όπου παράγεται ένα εφικτό σημείο εκκίνησης $\mathbf{q}^{(0)}$.

V.2 Καθορισμός Εφικτών Κατευθύνσεων Αύξησης

Εδώ, περιγράφεται η μέθοδος που υιοθετήθηκε για τον καθορισμό εφικτών κατευθύνσεων «βελτίωσης», στην επαναληπτική διαδικασία που οδηγεί στην εύρεση του NBS. Δεδομένου ενός εφικτού διανύσματος \mathbf{q} , μία *εφικτή κατεύθυνση*

στο \mathbf{q} είναι ένα διάνυσμα \mathbf{d} τέτοιο που το $\mathbf{q} + a\mathbf{d}$ είναι εφικτό για όλα τα $a > 0$ που είναι αρκούντως μικρά. Μία εφικτή κατεύθυνση ξεκινά με ένα εφικτό διάνυσμα $\mathbf{q}^{(0)}$ και παράγει μία ακολουθία από εφικτά διανύσματα $\{\mathbf{q}^{(n)}\}$ σύμφωνα με την αναδρομή

$$\mathbf{q}^{(n+1)} = \mathbf{q}^{(n)} + a^{(n)}\mathbf{d}^{(n)},$$

Αν το $\mathbf{q}^{(n)}$ δεν είναι στάσιμο, το $\mathbf{d}^{(n)}$ είναι μία εφικτή κατεύθυνση στο $\mathbf{q}^{(n)}$, που είναι και μία κατεύθυνση αύξησης, δηλαδή, $\nabla g(\mathbf{q}^{(n)})^T \cdot \mathbf{d}^{(n)} > 0$, και το βήμα επιλέγεται να είναι θετικό και τέτοιο που $\mathbf{q}^{(n)} \in Q_0$.

Αν το $\mathbf{q}^{(n)}$ είναι στάσιμο, η μέθοδος τερματίζει, δηλαδή $\mathbf{q}^{(n+1)} = \mathbf{q}^{(n)}$ (ισοδύναμα, $\mathbf{d}^{(n)} = 0$). Προτείνεται μία μέθοδος εφικτής κατεύθυνσης που να είναι αλγόριθμος αύξησης (ascent algorithm)· δηλαδή, το βήμα $a^{(n)}$ επιλέγεται, ούτως ώστε

$$g(\mathbf{q}^{(n)} + a^{(n)}\mathbf{d}^{(n)}) > g(\mathbf{q}^{(n)}), \quad \forall n \in \{0, 1, \dots\}.$$

Καθώς το Q_0 είναι κυρτό, υπάρχει ένας εναλλακτικός χαρακτηρισμός της μεθόδου εφικτής κατεύθυνσης· οι εφικτές κατευθύνσεις στο $\mathbf{q}^{(n)}$ μπορούν να είναι διανύσματα της μορφής

$$\mathbf{d}^{(n)} = \beta(\bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}), \quad \beta > 0,$$

όπου το $\bar{\mathbf{q}}^{(n)}$ είναι ένα εφικτό διάνυσμα. Έτσι, η μέθοδος εφικτής κατεύθυνσης μπορεί να γραφεί στη μορφή

$$\mathbf{q}^{(n+1)} = \mathbf{q}^{(n)} + a^{(n)}(\bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}),$$

όπου $a^{(n)} \in (0, 1]$, κι αν το $\mathbf{q}^{(n)}$ δεν είναι στατικό,

$$\bar{\mathbf{q}}^{(n)} \in Q_0, \quad \nabla g(\mathbf{q}^{(n)})^T \cdot (\bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}) > 0.$$

Πρέπει να σημειωθεί ότι, λόγω της κυρτότητας του συνόλου περιορισμών Q_0 , $\mathbf{q}^{(n)} + a^{(n)}(\bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}) \in Q_0$, για όλα τα $a^{(n)} \in (0, 1]$, αν $\mathbf{q}^{(n)} \in Q_0$, με αποτέλεσμα η παραγόμενη ακολουθία $\{\mathbf{q}^{(n)}\}$ να είναι εφικτή. Επιπλέον, αν το $\mathbf{q}^{(n)}$ δεν είναι στατικό, υπάρχει πάντα μία εφικτή κατεύθυνση $\bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}$ με την ιδιότητα της αύξησης.

Περιγράφεται, τώρα, πως επιλέγεται μία εφικτή κατεύθυνση για ένα δεδομένο εφικτό σημείο $\mathbf{q}^{(n)}$. Ο πιο ευθύς τρόπος να παραχθεί μία εφικτή κατεύθυνση $\bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}$ που ικανοποιεί τη συνθήκη αύξησης $\nabla g(\mathbf{q}^{(n)})^T \cdot (\bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}) > 0$ είναι να λυθεί το πρόβλημα βελτιστοποίησης

$$(D) \begin{cases} \max_{\mathbf{w}} \nabla g(\mathbf{q}^{(n)})^T \cdot (\mathbf{w} - \mathbf{q}^{(n)}) \\ \text{s.t. } \mathbf{w} \in Q_0 \end{cases},$$

με λύση το $\bar{\mathbf{q}}^{(n)}$,

$$\bar{\mathbf{q}}^{(n)} = \arg \max_{\mathbf{w} \in Q_0} \nabla g(\mathbf{q}^{(n)})^T \cdot (\mathbf{w} - \mathbf{q}^{(n)}),$$

που είναι το πιο μακρινό σημείο του Q_0 στην κατεύθυνση της παραγώγου.

Λόγω του γεγονότος ότι το Q_0 δεν ορίζεται μόνον από γραμμικούς περιορισμούς, προτείνεται μια διαδικασία εύρεσης κατεύθυνσης δύο φάσεων. Πρώτα, λαμβάνονται υπόψη μόνο οι γραμμικοί περιορισμοί, δηλαδή οι περιορισμοί στο SIR που σχηματίζουν το κυρτό πολύεδρο Γ . Συγκεκριμένα, έχοντας φτάσει σε ένα εφικτό διάνυσμα $\mathbf{q}^{(n)}$, εντοπίζεται το σημείο $\bar{\mathbf{q}}^{(n)}$ που μεγιστοποιεί την αντικειμενική συνάρτηση του προβλήματος (D) στο πολύεδρο Γ . Τότε, βάσει της κυρτότητας του Q_0 , και του γεγονότος ότι $Q_0 \subseteq \Gamma$, προσαρμόζεται η κατεύθυνση που προκύπτει από το $\bar{\mathbf{q}}^{(n)}$, ώστε να είναι εφικτή. Συγκεκριμένα, εντοπίζεται το σημείο $\bar{\mathbf{q}}^{(n)} \in Q_0$, το οποίο βρίσκεται στην ευθεία που ενώνει το $\mathbf{q}^{(n)}$ με το $\bar{\mathbf{q}}^{(n)}$. Η κατεύθυνση $\mathbf{d}^{(n)}$ στο βήμα n θα είναι τότε

$$\mathbf{d}^{(n)} = \bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}.$$

Πιο κάτω συνοψίζεται ο αλγόριθμος υπολογισμού εφικτής κατεύθυνσης αύξησης.

1. Calculate the gradient $\nabla g(\cdot)$ of the objective function $g(\cdot)$ at point $\mathbf{q}^{(n)}$
2. Solve the linear problem for determining the point $\bar{\mathbf{q}}^{(n)}$ that maximizes the inner product $\nabla g(\mathbf{q}^{(n)})^T \cdot \mathbf{q}$ over the polyhedron Γ
3. If $\bar{\mathbf{q}}^{(n)}$ is feasible (i.e., $\bar{\mathbf{q}}^{(n)} \in Q_0$), then go to step 8. Else go to step 4
4. Set $\mathbf{l}_q := \mathbf{q}^{(n)}$, $\mathbf{r}_q := \bar{\mathbf{q}}^{(n)}$, and $\mathbf{q} := (\mathbf{r}_q + \mathbf{l}_q)/2$
5. If $\mathbf{q} \in Q_0$ then set $\mathbf{l}_q := \mathbf{q}$. Else set $\mathbf{r}_q := \mathbf{q}$.
6. Set $\mathbf{q}_{old} := \mathbf{q}$ and $\mathbf{q} := (\mathbf{r}_q + \mathbf{l}_q)/2$
7. If $\|\mathbf{q}_{old} - \mathbf{q}\|_2 < \varepsilon$, then $\bar{\mathbf{q}}^{(n)} := \mathbf{q}$, and go to step 8. Else go to step 5.
8. $\mathbf{d}^{(n)} := \bar{\mathbf{q}}^{(n)} - \mathbf{q}^{(n)}$. Finish

Η σταθερά ε στο βήμα 7 είναι ένας αρκούντως μικρός, θετικός αριθμός, ο οποίος χρησιμοποιείται ως κριτήριο τερματισμού του αλγορίθμου διχοτόμησης. Η Φάση 1 αποτελείται από τα βήματα 1-3. Η Φάση 2 περιλαμβάνει τα βήματα 4-7, και λαμβάνει χώρα, μόνον αν η Φάση 1 δεν έδωσε εφικτό σημείο. Με το βήμα 8 τερματίζεται ο αλγόριθμος παρέχοντας μία εφικτή κατεύθυνση αύξησης $\mathbf{d}^{(n)}$. Με αυτόν τον αλγόριθμο, υπολογίζεται το εφικτό σημείο που είναι το πιο μακρινό του Q_0 στην κατεύθυνση της παραγώγου, σε συμφωνία με την ιδέα του αρχικού αλγορίθμου, όπως διατυπώνεται μέσω του ορισμού του προβλήματος (D).

V.3 Επιλογή Βήματος

Εδώ, περιγράφεται ο αλγόριθμος για την επιλογή βήματος στη μελετούμενη επαναληπτική διαδικασία. Οι περισσότεροι κανόνες για την επιλογή του βήματος $a^{(n)}$ σε μεθόδους κλίσης έχουν, επίσης, εφαρμογή σε μεθόδους εφικτής κατεύθυνσης. Μερικοί από τους πιο δημοφιλείς κανόνες είναι οι επόμενοι [111]: 1) *limited minimization*, 2) *Armijo*, και 3) *σταθερό βήμα*. Στην παρούσα μελέτη, προτείνεται μία μέθοδος επιλογής βήματος, με βάση τον κανόνα του Armijo. Πρώτα, γίνεται η υπόθεση ότι η κατεύθυνση $\mathbf{d}^{(n)}$ είναι τέτοια ώστε το $\mathbf{q}^{(n)} + \mathbf{d}^{(n)}$ είναι εφικτό.

Έστω β και s σταθερές, με $\beta \in (0,1)$, και $s \in (0,1)$. Τίθεται $a^{(n)} = \beta^{m(n)}$, όπου το $m(n)$ είναι ο μικρότερος ακέραιος m που

$$g(\mathbf{q}^{(n)} + \beta^m \mathbf{d}^{(n)}) - g(\mathbf{q}^{(n)}) \geq s\beta^m [\nabla g(\mathbf{q}^{(n)})^T \cdot \mathbf{d}^{(n)}].$$

Με άλλα λόγια, τα βήματα β, β^2, \dots , δοκιμάζονται διαδοχικά, μέχρις ότου να ικανοποιηθεί η παραπάνω συνθήκη για $m = m(n)$. Έτσι, το βήμα $a^{(n)}$, στην επανάληψη n , επιλέγεται όχι μόνο για να προκύψει θετική βελτίωση στην αντικειμενική συνάρτηση· σύμφωνα με τον παραπάνω έλεγχο, αυτή η βελτίωση θα πρέπει να είναι αρκετά μεγάλη.

Πιο κάτω, παρατίθεται ο αριθμητικός αλγόριθμος για την επιλογή βήματος, δεδομένης μίας εφικτής κατεύθυνσης $\mathbf{d}^{(n)}$.

1. Calculate the gradient $\nabla g(\cdot)$ and the value of the objective function $g(\cdot)$ at point $\mathbf{q}^{(n)}$. Set $m := 1$
2. Calculate the value of the objective function $g(\cdot)$ at point $\mathbf{q}^{(n)} + \beta^m \mathbf{d}^{(n)}$
3. If $g(\mathbf{q}^{(n)} + \beta^m \mathbf{d}^{(n)}) - g(\mathbf{q}^{(n)}) \geq s\beta^m [\nabla g(\mathbf{q}^{(n)})^T \cdot \mathbf{d}^{(n)}]$, then $a^{(n)} := \beta^m$. Finish.
4. Set $m := m + 1$. Go to step 2

Όροι

Άθροισμα με βάρη	Weighted sum
Ακολουθιακή ισορροπία	Sequential equilibrium
Ακόλουθος	Follower
Ανάλυση αντιπαράθεσης	Conflict analysis
Ανοδική δημοπρασία	Ascending auction
Ανταλλαγή	Tradeoff
Αντιπαράθεση	Conflict
Αντιστοιχία	Mapping
Αξιολόγηση	Valuation
Απολαβή CDMA κώδικα	Processing gain
Απομακρυσμένη ευστοχία	Remote hit
Αποτυχημένη αίτηση προς την cache	Cache miss
Αριθμητική μέθοδος βασισμένη στην κλίση	Gradient-based numerical method
Αστοχία	Miss
Ατομικός ορθολογισμός	Individual rationality
Αυτοπαθής	Reflexive
Δείκτης δημοτικότητας	Popularity index
Δεύτερης τιμής	Second price
Δημοπρασία	Auction
Δημοπρασία πολλαπλών αντικειμένων	Multi-object auction
Διαδραστική θεωρία αποφάσεων	Interactive decision theory
Διαδραστικός	Interactive
Διαιτησία	Arbitration
Διακομιστής	Server
Διακοπή λόγω μεταπομπής	Handover blocking
Διαπραγμάτευση	Bargain
Διαστασιοποίησης	Provisioning
Διάχυτος υπολογισμός	Pervasive computing
Δυναμική πρόσβαση στο φάσμα	Dynamic spectrum access
Εικονικό μονοπάτι	Virtual path
Εκθετικό φίλτρο κινούμενου μέσου	Exponential moving average filter
Έλεγχος αποδοχής κλήσης	Call admission control
Έλεγχος ισχύος	Power control
Έλεγχος πρόσβασης στο μέσο	Medium access control
Έλεγχος ροής	Flow control
Ενδο-κυψελική παρεμβολή	Intra-cell interference
Εξω-κυψελική παρεμβολή	Other-cell interference
Επιτυχής αίτηση προς την cache	Cache hit
Έσοδα	Revenue
Ευρυεκπομπή	Broadcast

Ηγέτης	Leader
Ημι-κοίλη	Quasiconcave
Θεωρία Αποφάσεων του Bayes	Bayesian decision theory
Ισορροπία Nash	Nash equilibrium
Καθαρό όφελος	Net benefit
Καθοδική δημοπρασία	Descending auction
Καλύτερη απόκριση	Best response
Κανάλι ασφαλείας	Guard channel
Κατά Pareto αποτελεσματικός	Pareto efficient
Κατά Pareto αποτελεσματικός	Pareto optimal
Κατά Pareto κατάταξη	Pareto ranking
Κατά Pareto κυρίαρχος	Pareto dominant
Κίνητρο	Incentive
Κλιμάκωση	Scalability
Κοινωνική ευημερία	Social welfare
Κυψελωτό	Cellular
Λογικός	Rational
Λύση διαπραγμάτευσης του Nash	Nash bargaining solution
Μέθοδος κλίσης υπό συνθήκη	Conditional gradient method
Μεταβατική	Transitive
Μεταγωγέας	Network switch
Μεταπομπή	Handover
Μονοπώλιο	Monopoly
Οικονομία	Market
Ολιγοπώλιο	Oligopoly
Οριακό κόστος	Marginal cost
Παίγνιο αποσυντονισμού	Anticoordination game
Παίγνιο διασποράς	Dispersion game
Παίγνιο μειονότητας	Minority game
Παίγνιο συντονισμού	Coordination game
Παρεκβολή	Extrapolation
Περιορισμένη λογική	Bounded rationality
Πιθανότητα μπλοκαρίσματος	Blocking probability
Πλειστηριαστής	Auctioneer
Πλεονάζουσα υποδομή	Overprovisioning
Πληρεξούσιος	Proxy
Πληρεξούσιος	Proxy
Πλήρους τάξης	Full rank
Ποιότητα υπηρεσίας	Quality of service
Πολυεκπομπή	Multicast
Πολυωνυμική προσαρμογή	Polynomial fitting
Πράκτορας	Agent
Πρόβλημα κοινωνικής αλληλεπίδρασης	Social interaction problem
Πρόβλημα σακιδίου	Knapsack problem
Πρόγραμμα πλοήγησης	Browser

Προδραστικός	Proactive
Προσφορά	Bid
Πρώτης τιμής	First price
Ρυθμοαπόδοση	Throughput
Ρυθμιστής	Regulator
Σημείο εστίασης	Focal point
Σταθερό σημείο	Fixed point
Στάσιμος	Stationary
Στοιχείο μεταγωγής δικτύου	Network switch
Συγκρίσιμος	Comparable
Συμβατότητα κινήτρων	Incentive compatibility
Συνεχές	Continuum
Συνεχώς διαφορίσιμος	Continuously differentiable
Συνθήκη βελτιστότητας	Optimality condition
Συντελεστής διασποράς	Coefficient of variation
Τοπική ευστοχία	Local hit
Τυχαία πρόσβαση	Random access
Υπερεπιφάνεια	Hypersurface
Χρέωση με διάκριση	Price discrimination
Χρονοδρομολόγηση	Scheduling
Χρονοθυρίδα	Timeslot
Χωρητικότητα πόλου	Pole capacity
Χώρος στρατηγικών	Strategy space
Ωφέλεια	Payoff, Utility

Ακρωνύμια

ADSL	Asynchronous digital subscriber loop
AP	Access point
ATM	Asynchronous transfer mode
AWGN	Additive white Gaussian noise
BER	Bit-error-rate
BSG	Band selection game
CAC	Call admission control
CDMA	Code division multiple access
CDN	Content distribution network
CRT	Cell residence time
DCA	Dynamic channel allocation
DG	Dispersion game
DSA	Dynamic spectrum access
DSR	Distributed selfish replication
ESS	Elementary step system
FDMA	Frequency division multiple access
FTP	File transfer protocol
GSM	Global system mobile
HTTP	Hyper-text transport protocol
IPM	Intelligent pervasive manager
ISM	Industrial, Scientific, and Medical
LG	Location graph
LRU	Least recently used
NBS	Nash bargaining solution
NE	Nash equilibrium
NRM	Network resource monitor
PER	Packet-error-rate
PSBS	Probabilistic strategy for band selection
QoS	Quality of service
RMC	Resource management component
RP	Relocation proposal
RRM	Radio resource management
RWP	Random waypoint
SFBP	Santa Fe bar problem
SFBS	Santa Fe strategy for band selection
SIR	Signal-to-interference ratio
SNMP	Simple network management protocol
TC	Terminal client
TDMA	Time division multiple access
UCM	User context manager
UMTS	Universal telecommunications system
VCG	Vickrey-Clarke-Groves
VoIP	Voice over IP

WLAN	Wireless local area network
WWW	World Wide Web
ΑΠΚ	Αλγόριθμος πρόβλεψης κίνησης
ΚΤ	Κινητό τερματικό
ΣΒ	Σταθμός βάσης

Αναφορές

- [1] E. C. Efstathiou, P. A. Frangoudis, G. C. Polyzos, "Stimulating Participation in Wireless Community Networks", Proc. INFOCOM 2006, April, 2006
- [2] "<http://www.fon.com>"
- [3] "<http://www.whisher.com>"
- [4] L. Kleinrock, "An Internet vision: the invisible global infrastructure", Ad Hoc Networks 1(1), 2003, pp. 3-11
- [5] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohantya, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey", Computer Networks, Volume 50, Issue 13, 2006, pp. 2127-2159
- [6] N. Laoutaris, G. Smaragdakis, A. Bestavros, I. Stavrakakis, "Mistreatment in Distributed Caching Groups: Causes and Implications", Proc. IEEE INFOCOM 2006, Barcelona, Spain
- [7] Davis, M. D. (1983), Game Theory, A Nontechnical Introduction, Dover Publications, New York.
- [8] Straffin, P. D. (1993), Game Theory and Strategy, The Mathematical Association of America, Washington, DC.
- [9] S.J.Shenker, "Making Greed Work in Networks: A Game Theoretic Analysis of Switch Service Disciplines", IEEE/ACM Transactions on Networking, Vol.3, No.6, December, 1995.
- [10] Y.A. Korilis, A.A. Lazar, and A. Orda, "Architecting Noncooperative Networks", IEEE JSAC, Vol. 13, No. 8, 1995.
- [11] A.B. MacKenzie, and S.B.Wicker, "Game Theory and the Design of Self-Configuring, Adaptive Wireless Networks", IEEE Communications Magazine, November 2001.
- [12] J. Feigenbaum, C.Papadimitriou, and S.Shenker, "Sharing the Cost of Multicast Transmissions", Proceedings of the 32nd Annual ACM Symposium on Theory of Computing (STOC00), May 2000.
- [13] Shaun P. Hargreaves Heap, Yanis Varoufakis, "Game Theory: A Critical Introduction", Routledge, 1995
- [14] M. J. Osborne and A. Rubinstein, "A course in game theory", MIT Press, 1994
- [15] R. P. McAfee and J. McMillan. "Auctions and bidding", Journal of Economic Literature, 25(2):699–738, Jun 1987.
- [16] W. Vickrey, "Counterspeculation, auctions, and competitive sealed tenders", Journal of Finance, 16(1):8–37, Mar 1961.
- [17] E. H. Clarke "Multipart pricing of public goods", Public Choice, 11:17–33, 1971.
- [18] T. Groves, "Incentives in teams", Econometrica, 41(3):617–631, Jul 1973.

- [19]J. Green and J.-J. Laffont, "Characterization of satisfactory mechanisms for the revelation of preferences for public goods", *Econometrica*, 45(2):427–438, Mar 1977.
- [20]J. Robinson, "An iterative method of solving a game", *Annals of Mathematics* 54, 1951, pp. 298–301
- [21]L. P. Kaelbling, M. L. Littman and A. P. Moore, "Reinforcement learning: A survey", *Journal of Artificial Intelligence Research* 4, 1996, pp. 237–285
- [22]M. Weiser, "*The computer for the twenty-first century*", *Scientific American*, September 1991, pp. 94-104. (Reprinted in *IEEE Pervasive Computing*, Jan-Mar 2002).
- [23]M. Satyanarayanan, "*Pervasive Computing: Vision and Challenges*", *IEEE Personal Communications*, August 2001, pp. 10-17
- [24]P. Castro, P. Chiu, T. Kremenek, and R. Muntz, "*A Probabilistic Room Location Service for Wireless Networked Environments*" in *Proceedings of Ubicomp 2001*, Atlanta, Georgia, USA, pp. 18-34.
- [25]D. Fox, J. Hightower, L. Liao, D. Schulz, and G. Borriello, "*Bayesian Filtering for Location Estimation*", *IEEE Pervasive Computing*, July 2003, pp. 24-33.
- [26]J. Hightower, and G. Borriello, "*Location Systems for Ubiquitous Computing*", *IEEE Computer*, August 2001, pp. 57-66.
- [27]W. B. Arthur, "*Inductive Reasoning and Bounded Rationality (the El Farol Problem)*", *Amer. Econ. Review (Papers and Proceedings)*, 1994, pp. 406-411.
- [28]A. Greenwald, B. Mishra, and R. Parikh, "*The Santa Fe Bar Problem Revisited: Theoretical and Practical Implications*", *Stonybrook Festival on Game Theory: Interactive Dynamics and Learning*, July 1998.
- [29]D. C. Parkes, and D. Steinig, "*The Santa Fe Bar Problem: A Study in Multiagent Learning*".
- [30]J. Farago, A. Greenwald, and K. Hall, "*Fair and Efficient Solutions to the Santa Fe Problem*", in *Proceedings of the Grace Hopper Celebration of Women in Computing 2002*. Vancouver, October 2002.
- [31]J. D. Thomas, and K. Sycara, "*Heterogeneity, Stability, and Efficiency in Distributed Systems*", in *Proceedings of the International Conference on Multi Agent Systems*, July 1998, pp. 293-300.
- [32]T. Grenager, R. Powers, and Y. Shoham, "*Dispersion Games: General Definitions and Some Specific Learning Results*", *AAAI 2002*.
- [33]A. Orda, R. Rom, and N. Shimkin, "*Competitive Routing in Multiuser Communication Networks*", *IEEE/ACM Transactions on Networking*, October 1993, pp. 510-521.
- [34]H. Yaiche, R. M. Mazumdar, and C. Rosenberg, "*A Game Theoretic Framework for Bandwidth Allocation and Pricing in Broadband Networks*", *IEEE/ACM Transactions on Networking*, October 2000, pp. 667-678.

- [35]K. V. Srikanth, “*Network Routing Problems Using Game Theory*”, Report for the Degree of Bachelor of Technology, 2001.
- [36]M. Littman, and J. Boyan, “*A Distributed Reinforcement Learning Scheme for network Routing*”, tech. report CMU-CS-93-165, Robotics Institute, Carnegie Mellon University, 1993.
- [37]R.D. Luce, and H. Raiffa, “*Games and Decisions: Introduction and Critical Survey*”, John Wiley and Sons, 1957.
- [38]D.F. Ferguson, Y. Yemini, and C. Nikolaou, “*Microeconomic Algorithms for Load Balancing in Distributed Systems*”, in Proceedings of the International Conference on Distributed Systems (ICDCS’ 88), pp. 491-499.
- [39]J.F. Kurose, and R. Simha, “*A Microeconomic Approach to Optimal Resource Allocation in Distributed Computer Systems*”, IEEE Transactions On Computing, May 1989, pp. 705-719.
- [40]J.F. Kurose, M. Schwartz, and Y. Yemini, “*A Microeconomic Approach to Optimization of channel Access Policies in Multiaccess Networks*”, in Proceedings of the 5th Int. Conf. Distrib. Comput. Syst., Denver, May 1985, pp 70-80.
- [41]Agere Systems, “*Firmware Update for ORINOCO PC Cards v7.28 – Spring 2001 release*”, April 2001.
- [42]A. Balachandran, P. Bahl, and G. Voelker, “*Hot-Spot Congestion Relief and User Service Guarantees in Public-area Wireless Networks*”, In Proceedings of ACM WMCSA ’02, NY, USA, June 2002, p 70.
- [43]P. Myllymaki, T. Roos, H. Tirri, P. Misikangas, J. Sievanen, “*A Probabilistic Approach to WLAN User Location Estimation*”, Proc. IEEE Workshop on Wireless LANs, 2001
- [44]P. Bahl, and V. N. Padmanabhan, “*RADAR: An In-Building RF-Based User Location and Tracking System*”, Microsoft Research, in Proceedings of INFOCOM 2000, pp. 775-784.
- [45]A. Balachandran, G. M. Voelker, P. Bahl, and P. V. Rangan, “*Characterizing User Behavior and Network Performance in a Public Wireless LAN*”, in Proceedings of SIGMETRICS 2002.
- [46]D. Tang, and M. Baker, “*Analysis of a Local-Area Wireless Network*”, in Proceedings of ACM MobiCom ’00, pp. 1-10.
- [47]D. Kotz, and K. Essien, “*Analysis of a Campus-wide Wireless Network*”, in Proceedings of ACM MobiCom ’02, 107-118.
- [48]V. Krishna, “*Auction Theory*”, Academic Press, 2002.
- [49]C. Courcoubetis, and R. Weber, “*Pricing Telecommunication Networks*”, Wiley, 2003
- [50]J. Liu, and V. Issarny, “*Service Allocation in Selfish Mobile Ad hoc Networks Using Vickrey Auction*”, in the Post-Proceedings of the EDBT 2004 Pervasive Information Management Workshop, Crete, Greece, pp. 385-394.

- [51]T. Sandholm “*Limitations of the Vickrey Auction in Computational Multiagent Systems*”, in Proceedings of the 2nd International Conference on Multi-agent Systems (ICMAS-96), Menlo Park, CA, 1996. AAAI Press.
- [52]G. Alyfantis, S. Hadjieftymiades, and L. Merakos, “*MITOS: A Smart Spaces System for Pervasive Computing*”, in the Post-Proceedings of the EDBT 2004 Pervasive Information Management Workshop, Crete, Greece, pp. 376-384.
- [53]J. Schiller, and A. Voisard, “*Location-Based Services*”, Morgan Kaufmann, 2004
- [54]N. F. Johnson, S. Jarvis, R. Jonson, P. Cheung, Y. R. Kwong, and P. M. Hui, “*Volatility and agent adaptability in a self-organizing market*”, preprint cond-mat/9802177, February 1998.
- [55]J. Hightower, and G. Borriello, “*Location Sensing Techniques*”, Technical Report, UW CSE 01-07-01, University of Washington, Department of Computer Science and Engineering, Seattle, WA, July 2001
- [56]S. Hadjieftymiades, and L. Merakos, “*Proxies + Path Prediction: Improving Web Service Provision in Wireless-Mobile Communications*”, ACM/Kluwer Mobile Networks and Applications (MONET), 8(4), 2003, pp. 389-399
- [57]T.H.Cormen, C.E.Leiserson, R.L.Rivest, and C.Stein, “*Introduction to Algorithms*”, The MIT Press, 2003.
- [58]Holger Claussen, “Autonomous Self-deployment of Wireless Access Networks in an Airport Environment”, Proc. WAC 2005, pp. 86-98
- [59]S. Hadjieftymiades, S. Papayiannis and L. Merakos, “Using Path Prediction to Improve TCP Performance in Wireless/Mobile Communications”, IEEE Communications Magazine, Vol.40 No.8, 2002
- [60]M. Sidi and D. Starobinski, “New Call Blocking versus Handoff Blocking in Cellular Networks”, Kluwer Wireless Networks, Vol. 3, Issue 1, 1997.
- [61]Y. Zhang and D. Liu, “An Adaptive Algorithm for Call Admission Control in Wireless Networks”, Proc. IEEE Global Communications Conference, San Antonio, TX, pp.3628-3632, 2001
- [62]O. Yu and V. Leung, “Adaptive Resource Allocation for Prioritized Call Admission over an ATM-based Wireless PCN”, IEEE JSAC, Vol. 15, no. 7, pp. 1208–25, 1997
- [63]C. Oliveira, J. Kim and T. Suda, “An Adaptive Bandwidth Reservation Scheme for High-speed Multimedia Wireless Networks”, IEEE JSAC, vol. 16, no. 6, pp. 858–74, 1998.
- [64]D. Levine, I. Akyildiz and M. Naghshineh, “A Resource Estimation and Call Admission Algorithm for Wireless Multimedia Networks Using the Shadow Cluster Concept”, IEEE/ACM Trans. Networking, vol. 5, no.1, pp. 1–12, 1997.
- [65]D. Hong and S. Rappaport, “Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Non-Prioritized Handoff Procedures,” IEEE Trans. Veh. Technol., vol. VT-35, no. 3, pp. 77–92, 1986.

- [66]R. Guerin, "Queuing-blocking system with two arrival streams and guard channels," IEEE Trans. Commun., vol. 36, pp. 153–63, 1988.
- [67]G. Liu and G. Maguire Jr, "A Class of Mobile Motion Prediction Algorithms for Wireless Mobile Computing and Communications", MONET, Vol.1, pp.113-121, 1996.
- [68]A. Aljadhari and T. Znati, "Predictive Mobility Support for QoS Provisioning in Mobile Wireless Environments", IEEE JSAC, Vol.19, No.10, 2001.
- [69]M. Kyriakakos, S. Hadjiefthymiades, N. Fragkiadakis and L. Merakos, "Enhanced Path Prediction for Network Resource Management in Wireless LANs", in IEEE Wireless Communications Magazine, Vol.10 No.6, 2003.
- [70]M. Rabinovich and O. Spatscheck, "Web Caching and Replication", Addison Wesley, 2001
- [71]P. Cramton, "Bargaining with Incomplete Information: An Infinite-Horizon Model with Two-Sided Uncertainty", Review of Economic Studies, 51, pp. 579–593, 1984.
- [72]D. Fudenberg and J. Tirole, "Game Theory", MIT Press, Cambridge (MA), 1991.
- [73]S. Martello and P. Toth, "Knapsack Problems", John Wiley & Sons, 1990.
- [74]M. Zonoozi and P. Dassanayake, "User Mobility Modeling and Characterization of Mobility Patterns", IEEE JSAC, Vol.15, No.7, 1997.
- [75]H. Lin, M. Chatterjee, Sajal K. Das and K. Basu, "ARC: an integrated admission and rate control framework for CDMA data networks based on non-cooperative games", Proc. MOBICOM 2003, pp. 326-338
- [76]Luiz A. DaSilva, "Pricing for QoS-Enabled Networks: A Survey", IEEE Communications Surveys and Tutorials, 2000, vol. 3, No. 2, pp. 2-8
- [77]L. Kleinrock and R. Gail, "Queueing Systems: Problems and Solutions", John Wiley & Sons, 1996
- [78]David B. Johnson and David A. Maltz, "Dynamic source routing in ad hoc wireless networks", In Tomasz Imielinski and Hank Korth, editors, Mobile Computing, volume 353, pages 153-181. Kluwer Academic Publishers, 1996
- [79]"Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies," the FCC Notice of Proposed Rulemaking and Order -ET Docket No. 03-108.
- [80]"http://www.darpa.mil/ato/programs/xg/rfc_vision.pdf,"
- [81]H. Holma and A. Toskala, "WCDMA for UMTS", Wiley, New York, 2002.
- [82]P. Nicopolitidis, G. I. Papadimitriou, M. S. Obaidat and A. S. Pomportsis, "Third generation and beyond wireless systems", Communications of the ACM 46(8), pp. 120-124, 2003
- [83]W. Lu, "4G mobile research in Asia", IEEE Communications Magazine, pp. 104–106, March 2003
- [84]R. D. Yates, "A framework for uplink power control in cellular radio systems", IEEE Journal on Selected Areas in Communications, 13(7), pp. 1341-1347, 1995.

- [85]H. Ji and C. Huang, "Non-cooperative uplink power control in cellular radio systems", *Wireless Networks*, (4), pp. 233-240, 1998.
- [86]C. U. Saraydar, N. B. Mandayam and D. J. Goodman, "Efficient power control via pricing in wireless data networks", *IEEE Transactions on Communications*, Vol. 50, pp. 291 - 303, Feb. 2002
- [87]I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems: A comprehensive survey", *IEEE Personal Communications*, Jun. 1996, vol. 3, pp. 10–31.
- [88]V. Rodoplu and T.H. Meng, "Minimum energy mobile wireless networks", *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 1333–1344, Aug. 1999.
- [89]J. Mitola III, "Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio", Doctoral Dissertation, KTH, Stockholm, Sweden, May 2000.
- [90]Y. Xing, R. Chandramouli, S. Mangold and S. Shankar, "Dynamic Spectrum Access in Open Spectrum Wireless Networks", *IEEE Journal on Selected Areas in Communications*, 24(3), pp. 626-637, 2006.
- [91]X. Liu and W. Wang, "On the Characteristics of Spectrum-Agile Communication Networks", *Proc. IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, 2005.
- [92]A. Larcher, H. Sun, M. Van Der Schaar and Z. Ding, "Decentralized Transmission Strategies for Delay-sensitive Applications over Spectrum Agile Networks", *Proc. Packet Video 2004*, Dec. 2004.
- [93]G. Alyfantis, S. Hadjiefthymiades, and L. Merakos, "A Cooperative Uplink Power Control Scheme for Elastic Data Services in Wireless CDMA Systems", *ACM SIGCOMM Computer Communications Review*, Vol.36, No.3, 2006
- [94]E. Even-Dar and Y. Mansour, "Fast Convergence of Selfish Rerouting", *Proc. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2005
- [95]J. Proakis, "Digital Communications", McGraw Hill, 2001
- [96]W. Feller, "An Introduction to Probability Theory and its Applications", Vol. 1, 3rd ed. New York, Wiley, 1968.
- [97]J. Zander, "Performance of optimum transmitter power control in cellular radio systems", *IEEE Transactions on Vehicular Technology*, 41(1): 57-62, February 1992.
- [98]S. Grandhi, R. Yates and D. J. Goodman, "Resource allocation for cellular radio systems", *IEEE Transactions on Vehicular Technology*, 46(3): 581-587, August 1997.
- [99]D. M. Novakovic and M. L. Dukic, "Evolution of the Power Control Techniques for DS-CDMA Toward 3G Wireless Communication Systems", *IEEE Communications Surveys and Tutorials* 3(4): (2000)

- [100] V. Shah, N. B. Mandayam and D. J. Goodman, "Power control for wireless data based on utility and pricing", Proc. PIMRC, pages 1427-1432, 1998.
- [101] D. Famolari, N. B. Mandayam and D. J. Goodman, "A new framework for power control in wireless data networks: Games, utility and pricing", Proc. Allerton Conference, September 1998.
- [102] N. Feng, N. B. Mandayam and D. J. Goodman, "Joint power and rate optimization for wireless data services based on utility functions", Proc. CISS, March 1999.
- [103] J. F. Nash, "The Bargaining Problem", *Econometrica*, 18, 1950, pp. 155-162
- [104] D. J. Goodman and N. B. Mandayam, "Network Assisted Power Control for Wireless Data", *Kluwer Mobile Networks and Applications* 6, pp. 409–415, 2001
- [105] A.W. Roberts and D.E. Varberg, "Convex Functions", Academic Press, Inc., 1973.
- [106] S. Gunturi and F. Paganini, "Game theoretic approach to power control in cellular CDMA", Proc. IEEE Vehicular Technology Conference, Orlando, Oct 2003.
- [107] A. B. MacKenzie and S. Wiker. "Game Theory in Communications: Motivation, explanation and application to power control", Proc. IEEE Globecom, November 2001.
- [108] G. Alyfantis, S. Hadjiefthymiades, and L. Merakos, "On Fair and Efficient Power Control in CDMA Wireless Data Networks", Proc. ICCCN 2006, Arlington, Virginia, USA, October, 2006.
- [109] G. Debreu, "A Social Equilibrium Existence Theorem", in Proceedings of the National Academy of Sciences, Vol. 38, 1952, pp 886-893.
- [110] D. P. Bertsekas and J. N. Tsitsiklis, "Parallel and Distributed Computation: Numerical Methods", Belmont, MA: Athena Scientific, 1997.
- [111] D. P. Bertsekas, "Nonlinear Programming", Belmont, MA: Athena Scientific, 1995.
- [112] M. Xiao, N. B. Shroff and E. K. P. Chong, "A Utility-based Power-Control Scheme in Wireless Cellular Systems", *IEEE/ACM Transaction on Networking*, vol. 11, No 2, April 2003, pp. 210-221.
- [113] T. Alpcan, T. Başar, R. Srikant and E. Altman, "CDMA Uplink Power Control as a Noncooperative Game", *Wireless Networks* 8, pp. 659–670, 2002, Kluwer Academic Publishers.
- [114] S. Koskie and Z. Gajic, "A Nash Game Algorithm for SIR-Based Power Control in 3G Wireless CDMA Networks", *IEEE/ACM Transactions on Networking*, October, 2005, vol. 13, No. 5, pp. 1017–1026

- [115] N. Feng, S.-C. Mau and N. B. Mandayam, "Pricing and Power Control for Joint User-centric and Network-centric Resource Allocation", IEEE Transactions on Communications, September 2004, vol. 52, No. 9, pp. 1547-1557
- [116] E. Altman, J. Galtier and C. Touati, "Fair power and transmission rate control in wireless networks", Proc. Third Annual Conference on Wireless On demand Network Systems and Services (WONS), Les Ménuires, France, January, 2006
- [117] B. M. Duska, D. Marwood and M. J. Feeley, "The Measured Access Characteristics of WWW Client Proxy Caches", Proc. USENIX Symposium on Internet Technologies and Systems, December 1997.
- [118] A. Bovopoulos and A. Lazar, "Decentralized algorithms for optimal flow control", Proc. Allerton Conference on Communications, Control, and Computing, 1987, pp. 979-988.
- [119] J. B. Rosen, "Existence and Uniqueness of Equilibrium Points for Concave N-Person Games", Econometrica, Vol.33, No.3, 1965.
- [120] S. H. Rhee and T. Konstantopoulos, "A Decentralized Model for Virtual Path Capacity Allocation", Proc. INFOCOM 1999.
- [121] A. Lazar, A. Orda and D. Pendarakis, "Virtual Path Bandwidth Allocation in Multiuser Networks", IEEE/ACM Transactions on Networking, Volume 5, Issue 6, December 1997, pp. 861 – 871
- [122] Digital Equipment Corporation, Digital's Web Proxy Traces. <ftp://ftp.digital.com/pub/DEC/traces>, August–September 1996.
- [123] C. Roadknight and I. Marshall, "Variations in cache behavior", Proc. 7th International WWW Conference (WWW7), Brisbane, Australia, April 1998.
- [124] H. Nielsen, P. Leach and S. Lawrence, "An HTTP Extension Framework", IETF Network Working Group, RFC 2774, February 2000.
- [125] P. Konstanty and M. Kozipski, "Web Cache charging policies", position paper in NLANR Web Caching Workshop, Boulder, USA, 1997.
- [126] S. Hadjiefthymiades, Y. Georgiadis and L. Merakos, "A Game Theoretic Approach to Web Caching", Proc. 3rd International IFIP-TC6 Networking Conference, Athens, Greece, 2004.
- [127] N. Laoutaris, O. Telelis, V. Zissimopoulos and I. Stavrakakis, "Distributed Selfish Replication", accepted in IEEE Transactions on Parallel and Distributed Systems, 2005.
- [128] Ö. Erçetin and L. Tassiulas, "Market-Based Resource Allocation for Content Delivery in the Internet", IEEE Transactions on Computers 52(12), pp. 1573-1585, 2003.

- [129] B. Chun, K. Chaudhuri, H. Wee, M. Barreno, C. Papadimitriou and J. Kubiawicz, "Selfish caching in distributed systems: a game-theoretic analysis", Proc. 23rd Annual ACM Symposium on Principles of Distributed Computing, Canada, 2004.
- [130] S.C. Chapra and R.P. Canale, "Numerical Methods for Engineers", McGraw-Hill, 1989

