



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Μελέτη Αλγορίθμων Διαχείρισης Πόρων σε Ασύρματα Δίκτυα Υψηλών Ταχυτήτων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Τερψιχόρη - Ελένη Ν. Βελιβασάκη

Επιβλέπων : Συμεών Παπαβασιλείου
Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2007



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ
ΠΛΗΡΟΦΟΡΙΚΗΣ

Μελέτη Αλγορίθμων Διαχείρισης Πόρων σε Ασύρματα Δίκτυα Υψηλών Ταχυτήτων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Τερψιχόρη - Ελένη Ν. Βελιβασάκη

Επιβλέπων : Συμεών Παπαβασιλείου
Επίκουρος Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 9^η Νοεμβρίου 2007.

.....
Συμεών Παπαβασιλείου
Επίκουρος Καθηγητής Ε.Μ.Π.

.....
Χρήστος Καψάλης
Καθηγητής Ε.Μ.Π.

.....
Γεώργιος Φικιώρης
Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2007

.....
Τερψιχόρη – Ελένη Βελιβασάκη

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Τερψιχόρη-Ελένη Βελιβασάκη, 2007

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Στην Ελένη Γαβρά

Περίληψη

Η παρούσα διπλωματική εργασία ασχολείται με το πρόβλημα της κατανομής πόρων σε ασύρματα δίκτυα CDMA/HDR, τα οποία υποστηρίζουν ποικίλες υπηρεσίες πολυμέσων διαφορετικών και συχνά αντικρουόμενων απαιτήσεων, δίνοντας ιδιαίτερη έμφαση σε υπηρεσίες πραγματικού χρόνου. Παρουσιάζονται βασικοί αλγόριθμοι που στοχεύουν στη λύση του παραπάνω προβλήματος.

Επίσης, οι πιθανοτικές απαιτήσεις χαμηλής καθυστέρησης εξυπηρέτησης που δελέασαν τους μέχρι τώρα σχεδιαστές αλγορίθμων για τις υπηρεσίες πραγματικού χρόνου αποδεικνύονται ανεπαρκείς για την ικανοποίηση των χρηστών τέτοιων υπηρεσιών. Αντιθέτως, οι πιθανοτικές απαιτήσεις βραχυπρόθεσμης απόδοσης επιδεικνύουν βελτιωμένες επιδόσεις στο συγκεκριμένο θέμα και ανταποκρίνονται στις ανάγκες των απαιτητικών υπηρεσιών πραγματικού χρόνου.

Απόρροια της συγκεκριμένης παρατήρησης είναι μια νεοεισαχθείσα στη βιβλιογραφία πολιτική δρομολόγησης που τηρεί ικανοποιητικά τις απαιτήσεις βραχυπρόθεσμης ρυθμαπόδοσης των υπηρεσιών πραγματικού χρόνου.

Τέλος, παρατίθενται αναλυτικά τα αποτελέσματα των προσομοιώσεων ορισμένων από τους αλγορίθμους που επιχειρούν τη δρομολόγηση χρηστών υπηρεσιών πραγματικού χρόνου και μέσω συγκριτικών διαγραμμάτων επαληθεύονται οι προσδοκίες περί ικανοποίησης των απαιτήσεων ποιότητας υπηρεσιών διαφορετικών κλάσεων από το νέο αλγόριθμο.

Λέξεις κλειδιά

CDMA, ασύρματο δίκτυο, κατανομή πόρων, αλγόριθμος δρομολόγησης, ομορτυνιστική πολιτική δρομολόγησης, υπηρεσία πραγματικού χρόνου, ποιότητα υπηρεσίας, καθυστέρηση εξυπηρέτησης, βραχυπρόθεσμη απόδοση, ρυθμαπόδοση.

Abstract

In this paper, the problem of efficient resource allocation in CDMA wireless networks supporting multimedia services with various and often diverse QoS prerequisites is addressed, emphasizing on real-time services' essential requirements satisfaction. Several algorithms trying to solve this problem are described.

An analytical framework for studying real-time users' short-term delay and throughput properties under fundamental opportunistic scheduling policies is provided. The corresponding results demonstrate that probabilistic delay constraints are insufficient indicators of real-time services' QoS prerequisites, while probabilistic short-term throughput requirements are more appropriate in asserting their performance expectations.

Based on these observations and results, a new scheduling policy is proposed for efficiently supporting heterogeneous services that include delay-tolerant non-real-time and delay-sensitive real-time services, over a wireless CDMA system under a common utility based framework.

Finally, via modeling and simulation it is demonstrated through comparative diagrams that significant performance improvements concerning both types of services' QoS requirements satisfaction are achieved through the proposed scheduling approach.

Key Words

CDMA, wireless networks, resource allocation, scheduling policy, opportunistic scheduling, real-time service, quality of service, delay tolerance, short-term throughput.

Ευχαριστήριο Σημείωμα

Ευχαριστώ τον επιβλέποντα καθηγητή μου Συμεών Παπαβασιλείου, ο οποίος εξ αρχής μου ενέπνευσε ενδιαφέρον για τη θεματική περιοχή των ασύρματων κινητών επικοινωνιών και μου έδωσε την ευκαιρία να τη γνωρίσω καλύτερα μέσω ενός ιδιαίτερα ενδιαφέροντος και επίκαιρου θέματος. Επίσης, το ήθος και η ευγένεια που τον διακρίνει, πέραν της βαθιάς γνώσης και της συνεχούς και επίμονης εργασίας, μου έδωσαν ένα παράδειγμα ζωής για την εικόνα του σύγχρονου επιστήμονα.

Ευχαριστώ τον Τιμόθεο Καστρινογιάννη, υποψήφιο διδάκτορα Ε.Μ.Π, ο οποίος στάθηκε πολύτιμος και υπομονετικός αρωγός στην προσπάθειά μου να γνωρίσω το θέμα της εργασίας, ώστε να τη φέρω επιτυχώς εις πέρας. Οι πολύτιμες συμβουλές του, η έμφαση στη λεπτομέρεια και η βαθιά γνώση του για το αντικείμενο με οδήγησαν στην τελειοποίηση της εργασίας.

Η άρτια συνεργασία μαζί τους μου προσέφερε μόνο θετικά στοιχεία και με όπλισε με σημαντικά εφόδια για το μέλλον, αλλά επίσης μου κέντρισε το ενδιαφέρον για περαιτέρω διερεύνηση του θέματος.

Πίνακας Περιεχομένων

1	ΕΙΣΑΓΩΓΗ	14
1.1	Η ΤΕΧΝΟΛΟΓΙΑ CDMA/HDR ([6], [7], [8])	15
1.2	ΤΟ ΑΣΥΡΜΑΤΟ ΠΕΡΙΒΑΛΛΟΝ	17
1.3	ΑΛΓΟΡΙΘΜΟΙ ΧΡΟΝΟΔΡΟΜΟΛΟΓΗΣΗΣ	18
1.4	ΜΟΝΤΕΛΟ ΣΥΣΤΗΜΑΤΟΣ	19
1.5	ΚΡΙΤΗΡΙΑ ΠΟΙΟΤΗΤΑΣ ΥΠΗΡΕΣΙΩΝ	22
2	ΚΥΡΙΟΙ ΑΛΓΟΡΙΘΜΟΙ ΧΡΟΝΟΔΡΟΜΟΛΟΓΗΣΗΣ ΧΡΗΣΤΩΝ ΥΠΗΡΕΣΙΩΝ ΜΗ ΠΡΑΓΜΑΤΙΚΟΥ ΧΡΟΝΟΥ	26
2.1	ΑΛΓΟΡΙΘΜΟΣ ΜΕΓΙΣΤΟΠΟΙΗΣΗΣ ΤΗΣ ΑΠΟΔΟΣΗΣ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ (MAX [5]).....	27
2.2	ΑΛΓΟΡΙΘΜΟΣ ΑΝΑΛΟΓΙΚΗΣ ΔΙΚΑΙΟΣΥΝΗΣ (PROPORTIONAL FAIRNESS – PF [2],[5]) ...	28
2.3	ΑΛΓΟΡΙΘΜΟΣ ΣΧΕΤΙΚΑ ΚΑΛΥΤΕΡΟΥ (RELATIVELY BEST – RB [5]).....	30
2.4	ΣΧΗΜΑ ΔΡΟΜΟΛΟΓΗΣΗΣ ΧΡΟΝΙΚΗΣ ΔΙΚΑΙΟΣΥΝΗΣ (TEMPORAL FAIRNESS SCHEDULING SCHEME [4]).....	32
2.5	ΣΧΗΜΑ ΔΡΟΜΟΛΟΓΗΣΗΣ ΧΡΗΣΤΙΚΗΣ ΔΙΚΑΙΟΣΥΝΗΣ (UTILITARIAN FAIRNESS SCHEDULING SCHEME [4]).....	34
2.6	ΣΧΗΜΑ ΔΡΟΜΟΛΟΓΗΣΗΣ ΕΓΓΥΗΣΗΣ ΕΛΑΧΙΣΤΗΣ ΑΠΟΔΟΣΗΣ (MINIMUM PERFORMANCE GUARANTEE SCHEDULING SCHEME [4]).....	37
3	ΚΥΡΙΟΙ ΑΛΓΟΡΙΘΜΟΙ ΧΡΟΝΟΔΡΟΜΟΛΟΓΗΣΗΣ ΧΡΗΣΤΩΝ ΥΠΗΡΕΣΙΩΝ ΠΡΑΓΜΑΤΙΚΟΥ ΧΡΟΝΟΥ	40
3.1	Ο ΑΛΓΟΡΙΘΜΟΣ «ΠΡΩΤΑ Η ΤΡΟΠΟΠΟΙΗΜΕΝΗ ΜΕΓΙΣΤΗ ΣΤΑΘΜΙΣΜΕΝΗ ΚΑΘΥΣΤΕΡΗΣΗ (MODIFIED LARGEST WEIGHTED DELAY FIRST – M-LWDF).....	41
3.2	Ο ΑΛΓΟΡΙΘΜΟΣ ΕΚΘΕΤΙΚΗΣ ΠΟΛΙΤΙΚΗΣ ΔΡΟΜΟΛΟΓΗΣΗΣ (THE EXPONENTIAL RULE – EXP).....	44
3.3	Ο ΑΛΓΟΡΙΘΜΟΣ ΟΠΟΡΤΟΥΝΙΣΤΙΚΗΣ ΔΡΟΜΟΛΟΓΗΣΗΣ ΕΚΧΩΡΗΣΗΣ ΒΑΣΙΣΜΕΝΗΣ ΣΤΗΝ ΠΙΘΑΝΟΤΗΤΑ ΠΡΟΣΒΑΣΗΣ (ACCESS PROBABILITY BASED ASSIGNMENT OPPORTUNISTIC SCHEDULING – APAOS).....	47
3.4	Ο ΑΛΓΟΡΙΘΜΟΣ ΟΠΟΡΤΟΥΝΙΣΤΙΚΗΣ ΔΡΟΜΟΛΟΓΗΣΗΣ ΒΑΣΙΣΜΕΝΗΣ ΣΕ ΕΛΑΣΤΙΚΗ ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΔΙΚΑΙΟΣΥΝΗ (ELASTIC SHORT-TERM FAIRNESS BASED OPPORTUNISTIC SCHEDULING – ESFS).....	51
4	ΤΟ ΚΡΙΤΗΡΙΟ ΤΗΣ ΒΡΑΧΥΠΡΟΘΕΣΜΗΣ ΡΥΘΜΑΠΟΔΟΣΗΣ	57
4.1	Η ΈΝΝΟΙΑ ΤΗΣ ΒΡΑΧΥΠΡΟΘΕΣΜΗΣ ΑΠΟΔΟΣΗΣ.....	58
4.2	ΜΑΘΗΜΑΤΙΚΗ ΑΠΟΔΕΙΞΗ ΑΝΕΠΑΡΚΕΙΑΣ ΤΟΥ ΚΡΙΤΗΡΙΟΥ ΚΑΘΥΣΤΕΡΗΣΗΣ [17]	59
4.3	ΟΙ ΠΙΘΑΝΟΤΗΤΕΣ ΑΣΤΟΧΙΑΣ ΓΙΑ ΤΙΣ ΠΟΛΙΤΙΚΕΣ MAX ΚΑΙ PF [17]	61
5	Ο ΑΛΓΟΡΙΘΜΟΣ USTF	73
5.1	Το ΜΑΘΗΜΑΤΙΚΟ ΠΡΟΒΛΗΜΑ.....	73
5.2	Η ΠΟΛΙΤΙΚΗ ΔΡΟΜΟΛΟΓΗΣΗΣ	74
6	ΠΡΟΣΟΜΟΙΩΣΕΙΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ	79
6.1	Το ΜΟΝΤΕΛΟ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ ΠΡΟΣΟΜΟΙΩΣΗΣ	79
6.2	ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ ΓΙΑ ΤΟ Σ1	81
6.3	ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ ΓΙΑ ΤΟ Σ2.....	89
6.4	Η ΣΗΜΑΣΙΑ ΤΗΣ «ΤΙΜΩΡΙΑΣ» ΣΤΟΝ ΑΛΓΟΡΙΘΜΟ USTF.....	104
7	ΣΥΜΠΕΡΑΣΜΑΤΑ	108
8	ΒΙΒΛΙΟΓΡΑΦΙΑ	110

Πίνακας Σχημάτων

Σχήμα 4. 1: Πιθανότητα αστοχίας καθυστέρησης για τους «καλούς» χρήστες συναρτήσει του κατωφλίου καθυστέρησης.	66
Σχήμα 4. 2: Πιθανότητα αστοχίας καθυστέρησης για τους «κακούς» χρήστες συναρτήσει του κατωφλίου καθυστέρησης.	67
Σχήμα 4. 3: Πιθανότητα αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης για τους «καλούς» χρήστες συναρτήσει του κατωφλίου ληφθέντων δεδομένων.	68
Σχήμα 4. 4: Πιθανότητα αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης για τους «κακούς» χρήστες συναρτήσει του κατωφλίου ληφθέντων δεδομένων.	69
Σχήμα 6. 1: Πιθανότητα αστοχίας καθυστέρησης εξυπηρέτησης συναρτήσει του κατωφλίου καθυστέρησης T_i ανά αλγόριθμο σε σύστημα όπου υπάρχουν μόνο RT χρήστες.	82
Σχήμα 6. 2: Πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης συναρτήσει του μεγέθους του παραθύρου παρατήρησης W_i ανά αλγόριθμο, όταν υπάρχουν μόνο RT χρήστες στο σύστημα.	83
Σχήμα 6. 3: Μέση ρυθμαπόδοση σε (Kbps) κάθε χρήστη συναρτήσει του κατωφλίου καθυστέρησης T_i ανά αλγόριθμο, όταν υπάρχουν μόνο RT χρήστες στο σύστημα.	85
Σχήμα 6. 4: Πιθανότητα αστοχίας καθυστέρησης ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης $T_i = 20$ χρονοθυρίδων.	86
Σχήμα 6. 5: Πιθανότητα αστοχίας καθυστέρησης ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης $T_i = 50$ χρονοθυρίδων.	87
Σχήμα 6. 6: Πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης ανά χρήστη και ανά αλγόριθμο για παράθυρο παρατήρησης $W_i = 20$ χρονοθυρίδων, όταν όλοι οι χρήστες είναι RT.	88
Σχήμα 6. 7: Πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης ανά χρήστη και ανά αλγόριθμο για παράθυρο παρατήρησης $W_i = 50$ χρονοθυρίδων, όταν όλοι οι χρήστες είναι RT.	89
Σχήμα 6. 8: Πιθανότητα αστοχίας καθυστέρησης συναρτήσει του αριθμού των χρηστών RT ανά αλγόριθμο για κατώφλι 20 χρονοθυρίδων.	90
Σχήμα 6. 9: Πιθανότητα αστοχίας καθυστέρησης συναρτήσει του αριθμού των χρηστών RT ανά αλγόριθμο για κατώφλι 50 χρονοθυρίδων.	91
Σχήμα 6. 10: Μέση πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης ανά αλγόριθμο συναρτήσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης $W_i = 20$ χρονοθυρίδων.	93
Σχήμα 6. 11: Μέση πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης ανά αλγόριθμο συναρτήσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης $W_i = 50$ χρονοθυρίδων.	93
Σχήμα 6. 12: Μέση ρυθμαπόδοση σε (Kbps) των RT χρηστών για κατώφλι καθυστέρησης $T_i = 20$ χρονοθυρίδες συναρτήσει του πλήθους τους.	95
Σχήμα 6. 13: Μέση ρυθμαπόδοση σε (Kbps) των RT χρηστών για κατώφλι $T_i = 50$ χρονοθυρίδες συναρτήσει του πλήθους τους.	96
Σχήμα 6. 14: Η μέση ρυθμαπόδοση σε (Kbps) των NRT χρηστών του συστήματος συναρτήσει του αριθμού των RT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες.	96
Σχήμα 6. 15: Η μέση ρυθμαπόδοση σε (Kbps) των NRT χρηστών του συστήματος συναρτήσει του πλήθους των RT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες.	97
Σχήμα 6. 16: Η μέση ρυθμαπόδοση σε (Kbps) του συστήματος συναρτήσει του αριθμού των RT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες.	97

Σχήμα 6. 17: Η μέση ρυθμαπόδοση σε (Kbps) του συστήματος συναρτήσει του αριθμού των RT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες.	98
Σχήμα 6. 18: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες και 1 RT χρήστη.....	98
Σχήμα 6. 19: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες και 3 RT χρήστες.	99
Σχήμα 6. 20: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες και 5 RT χρήστες.	99
Σχήμα 6. 21: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες και 1 RT χρήστη.....	100
Σχήμα 6. 22: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες και 3 RT χρήστες.	100
Σχήμα 6. 23: Η μέση ρυθμαπόδοση των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες και 5 RT χρήστες.	101
Σχήμα 6. 24: Η απόδοση (σε Kbps) ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης και παράθυρο παρατήρησης 20 χρονοθυρίδων και για σύστημα με 1 RT χρήστη.	102
Σχήμα 6. 25: Η απόδοση (σε Kbps) ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης και παράθυρο παρατήρησης 20 χρονοθυρίδων και για σύστημα με 3 RT χρήστες.	103
Σχήμα 6. 26: Η απόδοση (σε Kbps) ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης και παράθυρο παρατήρησης 20 χρονοθυρίδων και για σύστημα με 5 RT χρήστες.	103
Σχήμα 6. 27: Η πιθανότητα αστοχίας καθυστέρησης εξυπηρέτησης συναρτήσει του πλήθους των RT χρηστών για κατώφλι καθυστέρησης T_i 20 χρονοθυρίδων.....	104
Σχήμα 6. 28: Η πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης συναρτήσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης W_i 20 χρονοθυρίδων.....	105
Σχήμα 6. 29: Η μέση ρυθμαπόδοση (σε Kbps) των RT χρηστών συναρτήσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης W_i 20 χρονοθυρίδων.	105
Σχήμα 6. 30: Η μέση ρυθμαπόδοση (σε Kbps) των NRT χρηστών συναρτήσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης W_i 20 χρονοθυρίδων.....	105
Σχήμα 6. 31: Η μέση ρυθμαπόδοση (σε Kbps) του συστήματος συναρτήσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης W_i 20 χρονοθυρίδων.	106

1

ΕΙΣΑΓΩΓΗ

Οι σύγχρονες ανάγκες των επιχειρήσεων και της κοινωνίας γενικότερα κατέστησαν ζωτικής σημασίας την ασύρματη επικοινωνία, όχι μόνο σε επίπεδο φωνής, αλλά και σε επίπεδο εφαρμογών, που παραδοσιακά προσφέρονταν σε σταθερούς εργασιακούς χώρους. Ο επιχειρηματίας απαιτεί πλέον πλήρη συνδεσιμότητα με το γραφείο του σε όποιο σημείο και αν βρίσκεται, ενώ ο απλός καταναλωτής επιζητά σε όλο και μεγαλύτερο βαθμό υπηρεσίες πραγματικού χρόνου. Έτσι έφτασε η εποχή της λεγόμενης σύγκλισης των υπηρεσιών και της ευρυζωνικότητας που θα εξυπηρετήσει αυτή τη σύγκλιση.

Σαν αποτέλεσμα αυτών, γεννήθηκε η 3G ασύρματη τεχνολογία, η οποία υποστηρίζει ευρεία κλίμακα υπηρεσιών, μεταξύ των οποίων η αποθήκευση εφαρμογών και μεγάλων ποσοτήτων δεδομένων μέσω του ασύρματου διαύλου (Wireless application downloading), υπηρεσίες πολυμέσων (MMS, ήχο και βίντεο), ψηφιακές φωτογραφίες και βίντεο τραβηγμένα και μεταφερόμενα μέσω κινητών τερματικών, βελτιωμένες υπηρεσίες έκτακτης ανάγκης, κινητή συνδεσιμότητα στο Διαδίκτυο, παιχνίδια πραγματικού χρόνου, κινητό ηλεκτρονικό ταχυδρομείο, υπηρεσίες χαμηλής καθυστέρησης για συνομιλία και ανταλλαγή μηνυμάτων βίντεο (low latency push-to-chat and push-to-video message services) κ.ά.

Τα ασύρματα συστήματα που θα φιλοξενήσουν τέτοιες εφαρμογές θα πρέπει να υποστηρίζουν υψηλούς ρυθμούς μετάδοσης και να εξυπηρετούν κινητούς χρήστες, τα βασικά χαρακτηριστικά των οποίων είναι η ασυμμετρία κίνησης δεδομένων που παρουσιάζουν οι ζεύξεις ανόδου και καθόδου -με περισσότερη κίνηση να συσσωρεύεται στην κάτω- και η απαίτηση των χρηστών για άμεση ανταπόκριση των αιτημάτων τους από το σύστημα.

Στο παρόν κεφάλαιο γίνεται λόγος για το σύστημα CDMA/HDR, το οποίο ανταποκρίνεται στις σύγχρονες ανάγκες των κινητών χρηστών. Επίσης, αναφέρονται τα κυριότερα χαρακτηριστικά του ασύρματου περιβάλλοντος επικοινωνίας και περιγράφεται το μοντέλο του συστήματος που χρησιμοποιείται στην παρούσα διπλωματική εργασία. Ορίζονται οι έννοιες «αλγόριθμοι χρονοδρομολόγησης» και «κριτήρια ποιότητας υπηρεσίας» (Quality of Service - QoS) που χρησιμοποιούνται, όταν εξυπηρετούνται υπηρεσίες μη πραγματικού και πραγματικού χρόνου.

1.1 Η Τεχνολογία CDMA/HDR ([6], [7], [8])

Η τεχνολογία CDMA/HDR ή CDMA2000 1xEV-DO ή IS-856 ήταν προϊόν εργασίας κυρίως της Qualcomm Inc και είναι από τις κύριες τεχνολογίες πρόσβασης στο μέσο που υποστηρίζονται από τα 3G ασύρματα συστήματα. Η ονομασία 1xEV-DO αντιπροσωπεύει την προέλευση και το κύριο χαρακτηριστικό της: 1xRTT (Radio Transmission Technology – Τεχνολογία Ραδιο-Μεταφοράς) EVolution to Data Only (Εξέλιξη σε Δεδομένα Μόνο). Το HDR είναι ένα μοντέλο, το οποίο δίνει έμφαση στη ζεύξη καθόδου πακέτων δεδομένων που απασχολεί ένα μεμονωμένο φέρον συχνότητας του συστήματος. Η κύρια ιδέα του 1xEV-DO είναι ότι χρησιμοποιείται ξεχωριστό κανάλι για τη μεταφορά των δεδομένων, αντί να υπάρχει ανταγωνισμός στην εκχώρηση εύρους ζώνης ανάμεσα σε δεδομένα και φωνή, όπως συμβαίνει με το IS-95. Επιπρόσθετα, χρησιμοποιείται η τεχνολογία CDMA για τη μετάδοση δεδομένων, δεδομένου ότι προτιμάται από πολλούς κατασκευαστές και φορείς για τη δυνατότητα της ταυτόχρονης εκπομπής δύο ή και περισσότερων χρηστών στο ίδιο εύρος συχνοτήτων. Επίσης, υπάρχει διαχωρισμός στη συχνότητα που χρησιμοποιείται στην ζεύξη ανόδου (uplink ή reverse link) και καθόδου (downlink ή forward link). Διακρίνονται οι εξής εκδοχές του CDMA/HDR:

- η EV-DO Release 0 (EV-DO Rel.0)
- η EV-DO Revision A (EV-DO Rev.A)
- η EV-DO Gold Multicast
- η EV-DO Platinum Multicast

Η EV-DO Rel.0 προσφέρει μέγιστο ρυθμό μετάδοσης στην κάτω ζεύξη 2.4 Mbps και αντίστοιχα 153 Kbps στην πάνω ζεύξη. Βελτιστοποιείται για μετάδοση δεδομένων υψηλών ταχυτήτων και υποστηρίζει άριστα εφαρμογές πακέτων βέλτιστης προσπάθειας, ενώ χαρακτηρίζεται από υψηλή χωρητικότητα δικτύου και χαμηλότερο κόστος για τους λειτουργούς του δικτύου.

Όσον αφορά την ποιότητα υπηρεσίας, προσφέρεται σε δύο μορφές, τη βασισμένη σε χρήστη και τη βασισμένη σε εφαρμογή. Η πρώτη δίνει προτεραιότητα σε πακέτα που ανήκουν σε χρήστες υψηλής προτεραιότητας σύμφωνα με κάποιο τιμολογιακό διαχωρισμό, ενώ η δεύτερη δίνει προτεραιότητα σε εφαρμογές υψηλότερης προτεραιότητας ανάλογα με την ευαισθησία τους στην καθυστέρηση.

Το Gold Multicast αποτελεί αναβάθμιση σε λογισμικό του EV-DO Rel.0. Επιτρέπει υπηρεσίες πολυεκπομπής και μονοεκπομπής στο ίδιο φέρον EV-DO των 1.25 MHz σε οποιοδήποτε επιθυμητό ποσοστό κάνοντας το σύστημα πιο ευέλικτο κατά την εκχώρηση πόρων.

Το EV-DO Rev.A αυξάνει το μέγιστο ρυθμό μετάδοσης στα 3.1 Mbps στην κάτω ζεύξη, ενώ το αυξάνει στα 1.8 Mbps στην άνω ζεύξη. Προσφέρει, επίσης, υψηλότερη χωρητικότητα τομέα μέσα στο ίδιο φέρον των 1.25 MHz. Το φυσικό του στρώμα και ο βελτιωμένος βαθμός ποιότητας υπηρεσίας (Quality of Service – QoS) βελτιώνουν την απόδοση των ευαίσθητων σε καθυστέρηση εφαρμογών.

Το EV-DO Rev.A επιχειρεί βελτιώσεις στην εναέρια ζεύξη, ώστε να μειωθούν οι καθυστερήσεις μετάδοσης και να είναι δυνατός ο έλεγχος των υπηρεσιών με διάφορες και ισχυρές μεθόδους. Έτσι είναι δυνατός εκτός από τις βασισμένες σε χρήστη και εφαρμογή μεθόδους και ένας νέος βαθμός ποιότητας υπηρεσίας «βασισμένος στη ροή». Συνεπώς, υπάρχει μηχανισμός ανάθεσης προτεραιότητας εντός μιας κατηγορίας εφαρμογών ίδιας βαθμίδας ως προς τη προτεραιότητα βασισμένη σε εφαρμογή.

Το Platinum Multicast είναι μια οικονομικά συμφέρουσα μέθοδος ταυτόχρονης εκπομπής καναλιών βίντεο και ήχου σε μια εικονική απεριόριστη βάση

συνδρομητών. έτσι παρέχεται μεγαλύτερη χωρητικότητα, που σημαίνει περισσότερα κανάλια ή καλύτερη ανάλυση του περιεχομένου του ίδιου αριθμού καναλιών, που συνεπάγεται και αναλογικά χαμηλότερο κόστος παράδοσης πάνω από το εναέριο μέσο.

1.2 Το Ασύρματο Περιβάλλον

Δεδομένης της κινητής φύσης των χρηστών, τα ασύρματα δίκτυα κερδίζουν όλο και περισσότερο έδαφος στον αγώνα με τα ενσύρματα. Ωστόσο, τα σχήματα εκχώρησης πόρων και χρονοδρομολόγησης που χρησιμοποιούνται στα ενσύρματα δίκτυα δεν μπορούν να μεταφερθούν απευθείας στα ασύρματα δίκτυα λόγω κάποιων ιδιαίτερων χαρακτηριστικών που παρουσιάζουν, όπως το περιορισμένο εύρος ζώνης, χρονικά και τοπικά μεταβαλλόμενες συνθήκες διαύλου και απόδοση εξαρτώμενη από αυτές.

Η διάδοση στο εναέριο μέσο εξαρτάται χονδρικά από τρεις σχεδόν ανεξάρτητους μεταξύ τους παράγοντες, τις απώλειες διαδρομής, την αργή λογαριθμοκανονική σκίαση και την ταχεία πολυδίοδη εξασθένιση. Ο πρώτος παράγοντας εξαρτάται από την κινητικότητα των χρηστών, ενώ οι άλλοι δύο είναι χρονικά εξαρτώμενοι. Ομοίως χρονικά εξαρτώμενη είναι επομένως και η ποιότητα υπηρεσίας των χρηστών. Για υπηρεσίες φωνής, καλύτερη υπηρεσία σημαίνει καλύτερη ποιότητα φωνής, ενώ για υπηρεσίες δεδομένων σημαίνει εν γένει υψηλότερους ρυθμούς μετάδοσης.

Όσον αφορά τις αποσβέσεις και τις παρεμβολές, αυτές αντιμετωπίζονται με τη χρήση τεχνικών μέτρησης και ελέγχου του διαύλου και τελικά καταπίεση της παρεμβολής και εξομάλυνση.

Παραμένει όμως το πρόβλημα της περιορισμένης χωρητικότητας. Εδώ εμπλέκονται οι αλγόριθμοι χρονοδρομολόγησης και κατανομής πόρων στους χρήστες του ασύρματου δικτύου, η οποία θα πρέπει να γίνεται κατά τρόπο πλήρως αποδοτικό, ώστε να εξυπηρετούνται όσο το δυνατόν περισσότεροι χρήστες και με υψηλή ποιότητα υπηρεσίας.

1.3 Αλγόριθμοι Χρονοδρομολόγησης

Με τον όρο αλγόριθμος χρονοδρομολόγησης εννοούμε μια πολιτική ή ένα σχήμα χρονισμού για την επιλογή από το σταθμό βάσης του χρήστη i που θα εξυπηρετηθεί κάθε χρονική σχισμή k (time slot) αλλά και για τον καθορισμό του ποσού των δεδομένων που θα λάβει αυτός, την ταχύτητα μετάδοσής των και την ισχύ της κεραίας του σταθμού βάσης. Με άλλα λόγια, με τον όρο πολιτική δρομολόγησης εννοείται ένα σύνολο κανόνων που διέπουν τη διαδικασία επιλογής με βάση συγκεκριμένα κριτήρια. Έστω Q τυχαία πολιτική χρονοδρομολόγησης χρηστών. Τότε, αν τη χρονική σχισμή k επιλέγεται ο χρήστης i του συστήματος για εξυπηρέτηση, ισχύει $i_Q(k) = i$. Έστω ακόμη $\{U_i^k\}$ η στοχαστική διαδικασία που αφορά το χρήστη i , όπου ως U_i^k ορίζεται η συνάρτηση χρησιμοποίησής του, η οποία εκφράζει το επίπεδο ευχαρίστησης-απόδοσης του χρήστη i τη χρονική θυρίδα k . Τη χρονική σχισμή k μια πολιτική δρομολόγησης Q επιλέγει το χρήστη $i = Q(\vec{U}, k) \in \{1, 2, \dots, N\}$ και του προσφέρει $U_{Q(\vec{U}, k)}^k$ πόρους (δηλ. U_i^k). Λόγω του γεγονότος ότι η κατάσταση του διαύλου ενός χρήστη είναι χρονικά συσχετισμένη, εξαιτίας της σκίασης για παράδειγμα, η απόδοση ενός χρήστη, καθώς και η ευχαρίστησή του σε σχέση με την ποιότητα υπηρεσίας του είναι και αυτή χρονικά συσχετισμένη. Επιπρόσθετα, η απόδοση διαφορετικών χρηστών μπορεί επίσης να σχετίζεται. Για παράδειγμα, υψηλή ενδοκυβελική παρεμβολή συνεπάγεται αυτομάτως μείωση της απόδοσης όλων των χρηστών.

Ο στόχος των αλγορίθμων χρονοδρομολόγησης είναι διπλός και μάλιστα πολλές φορές οι αλγόριθμοι επιχειρούν να συγκεράσουν αντικρουόμενους στόχους. Από τη μια πλευρά πρέπει να επιτύχουν μέγιστο βαθμό χρησιμοποίησης των ασύρματων πόρων και από την άλλη να εξασφαλίσουν υψηλή ποιότητα υπηρεσίας στους κινητούς χρήστες. Η χρησιμοποίηση του φάσματος πρέπει να είναι μέγιστη, αφού το διατιθέμενο εύρος ζώνης στις ασύρματες επικοινωνίες όχι μόνο είναι περιορισμένο, αλλά και η τελική ρυθμαπόδοση που μπορεί να επιτευχθεί είναι χρονικά εξαρτώμενη, ειδικά αν είναι επιθυμητή η επίτευξη υψηλών ρυθμών μετάδοσης.

Οι διάφοροι όμως κινητοί χρήστες δεν μπορούν να επιτύχουν την ίδια απόδοση, διότι είναι πολύ διαφορετικές οι συνθήκες του διαύλου για τον καθένα. Γι' αυτό έχουν προταθεί οι *οπορτουνιστικές (ευκαιριακές) πολιτικές χρονοδρομολόγησης*

(opportunistic scheduling policies). Μέσω αυτών, οι χρονικά και χωρικά μεταβαλλόμενοι δίαυλοι μπορούν να αξιοποιηθούν αποτελεσματικά προκειμένου να αυξηθεί η συνολική απόδοση του συστήματος. Η βασική ιδέα είναι ότι σε ένα σύστημα πολλών χρηστών είναι πολύ πιθανό να υπάρχει ένας με σηματοθορυβικό λόγο σχετικά μεγαλύτερο από τους υπόλοιπους χρήστες (relatively best) και έτσι η επιλογή του τελικά να οδηγήσει σε αύξηση της απόδοσης του συστήματος.

Ένας αλγόριθμος όμως με σκοπό απλά τη μεγιστοποίηση της απόδοσης του συστήματος καταλήγει σε άνιση μεταχείριση των χρηστών, ειδικά στην περίπτωση των χρηστών που απέχουν πολύ από το σταθμό βάσης, αφού κοντά στο σταθμό βάσης επιτυγχάνεται κατά κύριο λόγο καλύτερος σηματοθορυβικός λόγος (near-far effect). Αντίθετα, πρέπει να ληφθεί μέριμνα για θέματα εξασφάλισης κάποιων ελάχιστων απαιτήσεων υπηρεσίας για κάθε κινητό χρήστη, ώστε να αποφευχθεί ο αποκλεισμός ορισμένων κινητών χρηστών με χειρότερες συνθήκες διαύλου σε σχέση με τους υπόλοιπους.

Αυτό σημαίνει ότι ο αντίστοιχος δρομολογητής πρέπει να δέχεται πληροφορία αφενός για το ρυθμό μετάδοσης που θα επιτευχθεί αν επιλεγεί ο συγκεκριμένος χρήστης και αφετέρου πληροφορία σχετική με τη συνολική επίδοση της υπηρεσίας του, ώστε να είναι δίκαιη και ικανοποιητική η εξυπηρέτηση όλων των χρηστών από το σύστημα.

1.4 Μοντέλο Συστήματος

Θεωρείται μοντέλο χρονοδρομολόγησης μεταβλητής κατάστασης διαύλου βασισμένο στο σύστημα CDMA2000 1xEV-DO. Η μελέτη εστιάζεται σε μια κυψέλη και στη ζεύξη καθόδου. Θεωρείται σύστημα με χρονική πολυπλεξία και συγκεκριμένα με διαίρεση του χρόνου σε χρονικές σχισμές ή χρονοθυρίδες (time slots) διάρκειας 1.67 msec. Ένα τέτοιο σύστημα μπορεί να διαθέτει περισσότερους του ενός διαύλους, αλλά κατά τη διάρκεια μιας χρονοσχισμής μόνο ένας χρήστης απασχολεί ένα μόνο δίαυλο. Εδώ η χρονοδρομολόγηση επικεντρώνεται σε ένα μόνο δίαυλο. Γίνεται η παραδοχή ότι κάθε χρονική στιγμή όλοι οι χρήστες του συστήματος έχουν πακέτα προς εξυπηρέτηση. Επίσης, ο σταθμός βάσης εκπέμπει πάντα με τη μέγιστη ισχύ του.

Κάθε χρονική σχισμή γίνεται ανταλλαγή μηνυμάτων μεταξύ κινητού χρήστη και σταθμού βάσης μέσω του διαύλου σηματοδοσίας είτε σε λειτουργίες αναζήτησης-αναφοράς θέσης του κινητού χρήστη είτε σε λειτουργίες εκχώρησης εύρους ζώνης. Κάθε κινητό τερματικό μετρά το λόγο σήματος προς θόρυβο συν παρεμβολή (SINR) των σημάτων που δέχεται από τους κοντινούς σταθμούς βάσης. Δεδομένης της κινητικότητας του χρήστη και του συνεχώς μεταβαλλόμενου ασύρματου περιβάλλοντος, αυτό το SINR είναι δυνατόν να μεταβάλλεται πολύ συχνά, ακόμα και κάθε χρονική στιγμή. Γι' αυτό απαιτείται μηχανισμός ταχείας και συχνής ενημέρωσης του SINR ή του ρυθμού μετάδοσης που υποστηρίζεται κάθε χρονική σχισμή. Θεωρούμε παρ' όλ' αυτά ότι παραμένει σταθερό κατά τη διάρκεια μιας χρονικής σχισμής, κάτι που δεν απέχει ιδιαίτερα από την πραγματικότητα, αφού η διάρκειά της είναι αρκετά μικρή.

Ο δρομολογητής βρίσκεται στο σταθμό βάσης και επιλέγει τον προς εξυπηρέτηση χρήστη στην αρχή της χρονοθυρίδας. Στο σύστημα EV-DO γνωρίζει τους ρυθμούς μετάδοσης που μπορούν να επιτευχθούν, διότι κάθε χρήστης μετράει ένα πιλοτικό σηματοθορυβικό λόγο και δίνει αναφορά πίσω στο σταθμό βάσης με ένα μήνυμα Ελέγχου του Ρυθμού Μετάδοσης (Data Rate Control - DRC). Ο SINR και ο ρυθμός μετάδοσης συνδέονται μεταξύ τους με αναλογική σχέση. Ουσιαστικά ο χρήστης κάνει επιλογή του σταθμού βάσης με το υψηλότερο SINR. Επιπρόσθετα, στην περίπτωση δεδομένων που δεν είναι ευαίσθητα στην καθυστέρηση μπορούν να χρησιμοποιηθούν τεχνικές κωδικοποίησης διόρθωσης λαθών.

Πάντως, οι χρονικά μεταβαλλόμενοι διάυλοι των χρηστών θεωρούνται στοχαστικές διαδικασίες -έστω $\rho_i(t)$ - με σταθερές μέσες τιμές $\bar{\rho}_i$ και δίνονται από τη σχέση:

$$\rho_i(t) = G_i(t) \frac{P_{max}}{I_i + Wn_0} \quad (1.1)$$

όπου $G_i(t)$ το κέρδος του διαύλου χρήστη, I_i η ενδοκυψελική παρεμβολή που υφίσταται ο χρήστης i , W το εύρος ζώνης απλωμένου φάσματος, n_0 η μονόπλευρη φασματική πυκνότητα ισχύος του προσθετικού λευκού θορύβου τύπου Gauss (AWGN) και P_{max} η μέγιστη ισχύς εκπομπής του σταθμού βάσης. Θεωρώντας γραμμική σχέση μεταξύ του στιγμιαίου SINR και του ρυθμού μετάδοσης που μπορεί

να επιτευχθεί, ο ρυθμός μετάδοσης που αντιλαμβάνεται ο χρήστης $i = 1, 2, \dots, N$ δίνεται από τη σχέση:

$$r_i(t) = \frac{W}{\gamma_i} \rho_i(t) \quad (1.2)$$

όπου γ_i ο απαιτούμενος λόγος της ενέργειας ψηφίου προς τη φασματική ένταση παρεμβολής $\frac{E_b}{n_0}$.

Γίνεται η παραδοχή ότι οι διαδικασίες $\rho_i(t)$ είναι στατικές με την ευρεία έννοια και εργοδικές για όλους τους χρήστες και ότι πρόκειται για ακολουθίες ταυτόσημα και ανεξάρτητα κατανεμημένες (identically and independently distributed - i.i.d.). Η τυχαία μεταβλητή της Rayleigh εξασθένισης (fading) του διαύλου $\rho_i(t)$ θεωρείται ότι ακολουθεί εκθετική κατανομή με καθορισμένη μέση τιμή $E[\rho_i(t)] = \bar{\rho}_i$,

συναρτήσεις πυκνότητας πιθανότητας (density functions) $f_{\rho_i}(\rho_i) = \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i}$ και

συναρτήσεις κατανομής συσσώρευσης (cumulative distribution functions)

$$F_{\rho_i}(\rho_i) = 1 - e^{-\frac{\rho_i}{\bar{\rho}_i}}.$$

Η εξασθένιση Rayleigh που υποτίθεται για τους διαύλους περιγράφει το φαινόμενο των ταχέων διαλείψεων, κάτι το οποίο βρίσκεται κοντά στην πραγματικότητα για την περίπτωση του πυκνά οικοδομημένου περιβάλλοντος του κινητού χρήστη εντός αστικής περιοχής.

Τέλος, οι χρήστες του συστήματος μπορεί να ενδιαφέρονται για υπηρεσίες πραγματικού ή μη χρόνου. Στο εξής, ο αριθμός (σύνολο) των χρηστών υπηρεσίας πραγματικού χρόνου συμβολίζεται με N_{RT} (S_{RT}) και ο αριθμός (σύνολο) των χρηστών υπηρεσίας μη πραγματικού χρόνου με N_{NRT} (S_{NRT}), ενώ ο αριθμός (σύνολο) όλων των χρηστών στο σύστημα θα συμβολίζεται με $N = N_{RT} + N_{NRT}$ (S).

1.5 Κριτήρια Ποιότητας Υπηρεσιών

Οι σύγχρονες υπηρεσίες διακρίνονται σε πραγματικού (Real Time services – RT services) και μη (Non Real Time services – NRT services) χρόνου, ανάλογα με την ανοχή τους σε καθυστέρηση. Έτσι οι υπηρεσίες πραγματικού χρόνου, π.χ. υπηρεσίες χαμηλής καθυστέρησης για συνομιλία και ανταλλαγή μηνυμάτων βίντεο, είναι ευαίσθητες σε καθυστέρηση, ενώ οι υπηρεσίες μη πραγματικού χρόνου, όπως π.χ. κινητό ηλεκτρονικό ταχυδρομείο, όχι. Κάθε υπηρεσία χαρακτηρίζεται από το βαθμό της ποιότητας (Quality of Service - QoS) με την οποία πρέπει να παρέχεται αυτή, προκειμένου ο χρήστης να λαμβάνει ικανοποιητική εξυπηρέτηση. Στο εξής, ο βαθμός της ποιότητας υπηρεσίας θεωρείται ότι εκφράζεται μέσω των κριτηρίων ποιότητας υπηρεσίας.

1.5.1 Υπηρεσίες μη πραγματικού χρόνου (Non Real Time services)

Εν γένει, τα κριτήρια ποιότητας υπηρεσίας σχετίζονται με την ικανοποίηση κάποιων ελάχιστων απαιτήσεων υπηρεσίας από το δρομολογητή και με θέματα δικαιοσύνης μεταξύ των χρηστών σε μακροχρόνιο διάστημα. Αυτά έχουν να κάνουν με:

- όρια ρυθμαπόδοσης (throughput) μακροπρόθεσμα
- δικαιοσύνη μεταξύ των χρηστών μακροπρόθεσμα

Πρακτικά, όσον αφορά τα όρια ρυθμαπόδοσης, απαιτείται η επίτευξη ενός ελάχιστου ρυθμού μετάδοσης, προκειμένου να εξασφαλιστεί ότι κάθε χρήστης έλαβε αρκετό όγκο δεδομένων για την ικανοποίηση της υπηρεσίας του. Μέτρα της απόδοσης μπορούν να θεωρηθούν η άμεση μέτρηση της απόδοσης (σε bits/sec) ή η χρηματική αξία της απόδοσης (σε €/sec) ή η απόδοση μείον την κατανάλωση ισχύος. Εδώ θεωρείται το πρώτο. Μαθηματικά, το κριτήριο της ελάχιστης απόδοσης εκφράζεται ως εξής:

$$E\left(U_i I_{\{Q(\bar{v})=i\}}\right) \geq C_i \quad (1.3)$$

όπου U_i η απόδοση του χρήστη i , $I_{\{Q(\vec{U})=i\}} = \begin{cases} 1, & \text{αν } Q(\vec{U})=i \\ 0, & \text{αλλιώς} \end{cases}$ και C_i η ελάχιστη

απαίτηση του χρήστη i σε μέση απόδοση.

Η δικαιοσύνη μεταξύ των χρηστών μπορεί να είναι χρονική (temporal) ή χρηστική (utilitarian). Χρονική δικαιοσύνη σημαίνει ότι οι χρήστες έχουν εξυπηρετηθεί επί ίσο αριθμό χρονοθυρίδων, δηλαδή ίσες φορές, ή σε συγκεκριμένο ποσοστό ο καθένας σε σχέση με το συνολικό χρόνο εξυπηρέτησής τους, ορισμένο από τις απαιτήσεις των χρηστών σύμφωνα με κάποιο βάρος. Η μαθηματική έκφραση της χρονικής δικαιοσύνης είναι η εξής:

$$P\{Q(\vec{U})=i\} \geq f_i \quad (1.4)$$

όπου $Q(\vec{U})$ η πολιτική δρομολόγησης που χρησιμοποιείται, i ο χρήστης που επιλέγεται για εξυπηρέτηση και f_i η ελάχιστη απαίτηση σε χρονικές θυρίδες του χρήστη i . Η χρηστική δικαιοσύνη εκφράζει δικαιοσύνη σε όρους απόδοσης των χρηστών. Πρακτικά σημαίνει ότι κάθε χρήστης έχει λάβει ένα συγκεκριμένο κλάσμα της συνολικής απόδοσης του συστήματος. Η μαθηματική απόδοση του κριτηρίου χρηστικής δικαιοσύνης είναι η εξής:

$$E\left(U_i I_{\{Q(\vec{U})=i\}}\right) \geq a_i E(U_i) \quad (1.5)$$

όπου U_i η απόδοση του χρήστη i , $I_{\{Q(\vec{U})=i\}} = \begin{cases} 1, & \text{αν } Q(\vec{U})=i \\ 0, & \text{αλλιώς} \end{cases}$ και a_i το ελάχιστο

κλάσμα της συνολικής μέσης απόδοσης που πρέπει να λάβει ο χρήστης i .

1.5.2 Υπηρεσίες πραγματικού χρόνου (Real Time Services [12], [17])

Οι υπηρεσίες πραγματικού χρόνου διαφοροποιούνται από αυτές μη πραγματικού χρόνου από την ανοχή τους στην καθυστέρηση, όπως αναφέρθηκε στην υποενότητα 1.5.1. Επιπρόσθετα, οι χρήστες υπηρεσιών πραγματικού χρόνου πρέπει να τηρούν αυστηρούς βραχυπρόθεσμους περιορισμούς. Τα κριτήρια ποιότητας υπηρεσιών πραγματικού χρόνου συνοψίζονται στα ακόλουθα:

- όρια ανοχής καθυστέρησης στην εξυπηρέτηση των χρηστών

- όρια όγκου ληφθέντων δεδομένων σε βραχυπρόθεσμα διαστήματα
- εγγυήσεις ελάχιστων ρυθμών μετάδοσης δεδομένων

Όσον αφορά τις απαιτήσεις σε καθυστέρηση, ο χρήστης υπηρεσίας πραγματικού χρόνου πρέπει να εξυπηρετηθεί σύντομα από το δίκτυο, αλλιώς η επικοινωνία του παρουσιάζει διαλείψεις, με αποτέλεσμα η ποιότητά αυτής να είναι χαμηλή. Μαθηματικά ο περιορισμός καθυστέρησης για τις υπηρεσίες πραγματικού χρόνου εκφράζεται με μια άνω φραγμένη πιθανότητα, την *πιθανότητα αστοχίας καθυστέρησης*, και αποδίδεται ως εξής:

$$P[d_i(t) > T_i] \leq g_i^1 \quad (1.6)$$

όπου $d_i(t)$ η καθυστέρηση εξυπηρέτησης του χρήστη i τη χρονική στιγμή t , T_i το κατώφλι μέγιστης καθυστέρησης που μπορεί να ανεχτεί ο χρήστης i , g_i^1 το άνω όριο της πιθανότητας καθυστέρησης του χρήστη i . Έτσι, όσο μικρότερες είναι οι τιμές της παραπάνω πιθανότητας αστοχίας, τόσο μεγαλύτερη είναι η ευχαρίστηση των χρηστών υπηρεσίας πραγματικού χρόνου. Αν ο παραπάνω περιορισμός εφαρμοστεί στην αρχή κάθε χρονοθυρίδας, εξασφαλίζεται ότι η καθυστέρηση εξυπηρέτησης του χρήστη υπηρεσίας πραγματικού χρόνου ξεπερνά σπανιότητα το κατώφλι T (σύμφωνα πάντα με το όριο g_i^1).

Δεύτερον, κάθε χρήστης υπηρεσίας πρέπει να έχει λάβει αρκετό όγκο δεδομένων σε δεδομένα μικρά χρονικά διαστήματα, για να απολαμβάνει υψηλής ποιότητας, χωρίς διακοπές, υπηρεσία. Πρακτικά, το δεύτερο κριτήριο ποιότητας που στο εξής αναφέρεται ως *περιορισμός δικαιοσύνης βραχυπρόθεσμης απόδοσης* (short term throughput fairness constraint) ή απλά *περιορισμός βραχυπρόθεσμης απόδοσης* πρέπει να φροντίζει για την κατάλληλη ανάθεση πόρων του συστήματος σε κάθε έναν από τους χρήστες του εντός μικρών μη επικαλυπτόμενων χρονικών διαστημάτων. Η μαθηματική απόδοση του περιορισμού δικαιοσύνης βραχυπρόθεσμης απόδοσης δίνεται μέσω μιας άνω φραγμένης πιθανότητας, της *πιθανότητας αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης*, και είναι η ακόλουθη:

$$P[b_i(t) < B_i]_{W_i} \leq g_i^2 \quad (1.7)$$

όπου $b_i(t)$ η τυχαία μεταβλητή του γεγονότος ο χρήστης i να λαμβάνει ποσότητα δεδομένων b_i μέσα στο χρονικό διάστημα παρατήρησης W_i , B_i το ελάχιστο ποσό

δεδομένων που πρέπει να έχει λάβει ο χρήστης i μέσα στο χρονικό διάστημα παρατήρησης W_i για να εξυπηρετηθεί ικανοποιητικά και g_i^2 το άνω όριο της αντίστοιχης πιθανότητας αστοχίας. Συνεπώς, όσο μικρότερη είναι η παραπάνω πιθανότητα αστοχίας, τόσο μεγαλύτερος είναι ο βαθμός της ικανοποίησης του χρήστη i . Δεδομένου του απαιτούμενου ρυθμού μετάδοσης στη ζεύξη καθόδου του χρήστη i R_i , το κατώφλι μονάδων δεδομένων B_i μπορεί να εκτιμηθεί από τη σχέση:

$$B_i = R_i \cdot W_i \cdot T_{slot} \quad (1.8)$$

όπου T_{slot} η χρονική διάρκεια μιας χρονοθυρίδας.

Το τρίτο κριτήριο βασίζεται στο γεγονός ότι οι υπηρεσίες πραγματικού χρόνου απαιτούν ένα συγκεκριμένο σταθερό ρυθμό μετάδοσης για την ικανοποίηση της υπηρεσίας τους. Έτσι, αν R_i ο ελάχιστος απαιτούμενος ρυθμός μετάδοσης και r_i ο ρυθμός μετάδοσης που επιτυγχάνει ο χρήστης i , τότε το τρίτο κριτήριο εκφράζεται μαθηματικά:

$$r_i \geq R_i \quad (1.9)$$

Αξίζει να σημειωθεί ότι το κριτήριο της σχέσης (1.7) ουσιαστικά αποτελεί αντίστοιχη έκφραση του δεύτερου κριτηρίου, αλλά μακροπρόθεσμα, οπότε η εξασφάλιση του κριτηρίου της σχέσης (1.5) συνεπάγεται και την εξασφάλιση του τρίτου. Η εγγύηση ελάχιστου ποσού δεδομένων σε συγκεκριμένο χρονικό διάστημα ισοδυναμεί με εγγύηση ελάχιστου μέσου ρυθμού μετάδοσης για μακροπρόθεσμο διάστημα. Εφόσον το χρονικό διάστημα παρατήρησης παραμένει σταθερό, ο μέσος ρυθμός μετάδοσης σε ένα διάστημα αντικατοπτρίζει το μέσο ρυθμό μετάδοσης σε μεγάλο χρονικό διάστημα, αν διαιρεθεί με τον αριθμό των μη επικαλυπτόμενων διαστημάτων παρατήρησης που περιέχονται στο μεγάλο χρονικό διάστημα.

Η εγγύηση των απαιτήσεων καθυστέρησης και βραχυπρόθεσμης απόδοσης για όλους τους χρήστες υπηρεσίας πραγματικού χρόνου έγκειται στην ελαστικότητα του δρομολογητή να αλλάζει τις προτεραιότητες των χρηστών ως προς την επιλογή τους για εξυπηρέτηση σύμφωνα με τις αντίστοιχες επιδόσεις τους σε καθυστέρηση και βραχυπρόθεσμη απόδοση.

2

ΚΥΡΙΟΙ ΑΛΓΟΡΙΘΜΟΙ

ΧΡΟΝΟΔΡΟΜΟΛΟΓΗΣΗΣ ΧΡΗΣΤΩΝ

ΥΠΗΡΕΣΙΩΝ ΜΗ ΠΡΑΓΜΑΤΙΚΟΥ ΧΡΟΝΟΥ

Παρακάτω αναλύονται οι βασικοί αλγόριθμοι χρονοδρομολόγησης που χαρακτηρίζονται ομορτυνιστικοί. Στόχος τους είναι, όπως προαναφέρθηκε, η μεγιστοποίηση της απόδοσης του συστήματος με παράλληλη εξυπηρέτηση κάποιων κριτηρίων ποιότητας υπηρεσίας.

2.1 Αλγόριθμος Μεγιστοποίησης της Απόδοσης του Συστήματος (MAX [5])

2.1.1 Μορφοποίηση του προβλήματος

Ο συγκεκριμένος αλγόριθμος, γνωστός και ως «άπληστος» (greedy), στοχεύει καθαρά στην εκμετάλλευση των καλύτερων συνθηκών διαύλου των μεμονωμένων χρηστών κάθε χρονική στιγμή για την αύξηση της μέσης ρυθμαπόδοσης του συστήματος. Το μαθηματικό πρόβλημα είναι το παρακάτω:

$$\underset{Q \in \Theta}{\text{maximize}} \sum_{i=1}^N E \left(U_{Q(\bar{v})} I_{Q(k)=i} \right) \quad (2.1)$$

όπου Q η πολιτική δρομολόγησης, Θ το σύνολο όλων των στατικών πολιτικών δρομολόγησης, $I_{Q(k)=i} = \begin{cases} 1, & \text{όταν } Q(k) = i \\ 0, & \text{αλλιώς} \end{cases}$ η συνάρτηση ενδείκτης.

2.1.2 Πολιτική Δρομολόγησης

Ορίζεται η πολιτική i_{MAX}^* :

$$i_{MAX}^*(k) = \underset{i \in S}{\text{arg max}} \{U_i(k)\} \quad (2.2)$$

όπου ως $U_i(k)$ είθισται να χρησιμοποιείται το $\rho_i(k)$, το οποίο είναι ο λόγος σήματος προς θόρυβο συν παρεμβολή του χρήστη i τη χρονική στιγμή k . Η συγκεκριμένη πολιτική κάνει βέλτιστη αξιοποίηση των πόρων του δικτύου, επιτυγχάνοντας τη μέγιστη δυνατή μέση ρυθμαπόδοση του συστήματος. Καταλήγει, όμως, σε άνιση μεταχείριση των χρηστών και εκτός αυτού και αδυναμία εξυπηρέτησης όλων, όπως θα φανεί σε επόμενα κεφάλαια.

2.2 Αλγόριθμος Αναλογικής Δικαιοσύνης (Proportional Fairness – PF [2],[5])

2.2.1 Μορφοποίηση του προβλήματος [19]

Οι αλγόριθμοι αυτής της κατηγορίας στηρίζουν την επιλογή του επόμενου προς εξυπηρέτηση χρήστη σε μια εξισορρόπηση ανάμεσα στους στιγμιαίους πιθανούς ρυθμούς μετάδοσης και στη δικαιοσύνη. Η επιλογή του χρήστη με τον υψηλότερο ρυθμό μετάδοσης κάθε χρονοθυρίδα καταλήγει στην απομόνωση και μη εξυπηρέτηση των χρηστών με χαμηλότερους στιγμιαίους σηματοθορυβικούς λόγους SINR. Η δίκαιη κατανομή πόρων θα ελαττώσει τη συνολική ρυθμαπόδοση σε σχέση με τη μέγιστη δυνατή, αλλά θα παρέχει αποδεκτά επίπεδα εξυπηρέτησης στους χρήστες με δυσμενείς συνθήκες διαύλου. Σε κάθε χρονοθυρίδα επιλέγεται ο χρήστης με τις καλύτερες στιγμιαίες συνθήκες διαύλου συγκριτικά με μια στατιστική μέτρηση της μέχρι τώρα ληφθείσας υπηρεσίας. Το μαθηματικό πρόβλημα βελτιστοποίησης, υπό τη θεώρηση ότι οι συναρτήσεις ευχαρίστησης των χρηστών είναι λογαριθμικές συναρτήσεις της ταχύτητας μετάδοσης δεδομένων που τελικά επιτυγχάνουν, είναι το παρακάτω:

$$\underset{Q \in \Theta}{\text{maximize}} \left\{ E \left[\sum_{i \in S} \log(R_i(t)) I_{i_{PF}(t)=i} \right] \right\} \quad (2.3)$$

όπου $I_{Q(k)=i} = \begin{cases} 1, & \text{όταν } Q(k) = i \\ 0, & \text{αλλιώς} \end{cases}$ η συνάρτηση ενδείκτης του ενδεχομένου $Q(k) = i$,

Q η πολιτική δρομολόγησης και Θ το σύνολο των πολιτικών δρομολόγησης.

2.2.2 Πολιτική Δρομολόγησης

Ορίζεται η πολιτική i_{PF}^* :

$$i_{PF}^*(k) = \underset{i \in S}{\text{arg max}} \left\{ \frac{DRC_i(k)}{R_i^*(k)} \right\} \quad (2.4)$$

όπου $DRC_i(k)$ ο ρυθμός που αιτείται από τον κινητό χρήστη i τη χρονοθυρίδα k , $R_i^*(k)$ ο μέσος ρυθμός μετάδοσης του χρήστη i μέχρι τη χρονοθυρίδα k .

Ο όρος $R_i^*(k)$ εκφράζει τη δικαιοσύνη μεταξύ των χρηστών και δίνεται από τον τύπο:

$$R_i^*(k+1) = \left(1 - \frac{1}{k_c}\right) R_i^*(k) + \frac{1}{k_c} R_i'(k) \quad (2.5)$$

όπου $R_i'(k)$ ο στιγμιαίος ρυθμός μετάδοσης του χρήστη i τη χρονοθυρίδα k , που σημαίνει ότι τις χρονοθυρίδες που ο χρήστης i δεν εξυπηρετείται ο όρος $R_i'(k)$ ισούται με μηδέν, k_c θετική σταθερά σχετιζόμενη με το μέγιστο χρονικό διάστημα κατά το οποίο ένας χρήστης του συστήματος δεν εξυπηρετείται. Προσεκτικότερη παρατήρηση της σχέσης (2.5) καταλήγει στα εξής συμπεράσματα:

- ο πρώτος όρος δείχνει την επίδραση των προηγούμενων ρυθμών μετάδοσης που δέχτηκε ή όχι ο χρήστης, ενώ ο δεύτερος ορίζει την επίδραση της τρέχουσας επιλογής του δρομολογητή στο συγκεκριμένο χρήστη.
- η παράμετρος k_c παίζει ρόλο στη σύγκλιση του ρυθμού $R_i^*(k)$. Έτσι παρατηρείται ότι για μεγάλες τιμές της k_c ο ρυθμός $R_i^*(k)$ παραμένει αμετάβλητος ουσιαστικά.
- μεγάλες τιμές της k_c υποδεικνύουν ότι ο δρομολογητής μπορεί να περιμένει για μεγαλύτερα χρονικά διαστήματα να βελτιωθεί η κατάσταση του διαύλου του κάθε χρήστη
- μεγάλες τιμές της k_c επίσης δείχνουν ότι τα προς μετάδοση πακέτα ενός συγκεκριμένου χρήστη είναι δυνατό να καθυστερήσουν κατά k_c χρονοθυρίδες, όταν οι συνθήκες του διαύλου του συγκεκριμένου χρήστη χειροτερεύουν απότομα.

Ως γνωστόν, η κατάσταση του διαύλου αποτελεί στοχαστική διαδικασία και μάλιστα οι δίαυλοι όλων των χρηστών έχουν την ίδια κατανομή, αλλά διαφορετικές μέσες τιμές εν γένει. Ο αλγόριθμος PF προσπαθεί να εξαλείψει την επίδραση των διαφορετικών μέσων τιμών, για να επέλθει δικαιοσύνη ανάμεσα στους χρήστες. Έτσι, κάθε χρήστης δέχεται εξυπηρέτηση επί ίσο αριθμό χρονοθυρίδων, υπό την υπόθεση

ότι οι διάυλοι των χρηστών αποτελούν ανεξάρτητες και όμοια κατανεμημένες μεταβλητές [5] με τους υπόλοιπους και εν τέλει λαμβάνει μέρος των συνολικών πόρων του συστήματος αναλογικά με τη μέση κατάσταση του διαύλου του.

Τέλος, ο αλγόριθμος PF, χρησιμοποιούμενος σε συστήματα HDR (High Data Rate) παρέχει δικαιοσύνη κατά την ακόλουθη έννοια. Αν χρησιμοποιηθεί άλλος αλγόριθμος χρονοδρομολόγησης για να αυξήσει τη ρυθμαπόδοση ενός συγκεκριμένου χρήστη κατά $x\%$ σε σχέση με αυτήν που επιτυγχάνει με τον PF, το άθροισμα όλων των επί τοις εκατό μειώσεων της ρυθμαπόδοσης των υπόλοιπων χρηστών θα είναι μεγαλύτερο από $x\%$. Υπό αυτήν την έννοια ο αλγόριθμος PF είναι βέλτιστος. Η ιδιότητά του αυτή είναι γνωστή και ως *κριτήριο αναλογικής δικαιοσύνης* (proportional fairness criteria).

2.3 Αλγόριθμος Σχετικά Καλύτερου (Relatively Best – RB [5])

2.3.1 Μορφοποίηση του προβλήματος

Ο συγκεκριμένος αλγόριθμος επιλέγει το χρήστη με τις στιγμιαίες καλύτερες συνθήκες διαύλου συγκριτικά με τη μέση κατάσταση του διαύλου του. Το μαθηματικό πρόβλημα στη γενικευμένη μορφή του είναι το παρακάτω [20]:

$$\underset{Q \in \Theta}{\text{maximize}} \left\{ \sum_{i=1}^N E \left[\log \rho_i I_{Q(k)=i} \right] \right\} \quad (2.6)$$

όπου Q η πολιτική δρομολόγησης, Θ το σύνολο όλων των στατικών πολιτικών δρομολόγησης, N το σύνολο των χρηστών του συστήματος, ρ_i η τυχαία μεταβλητή που περιγράφει τη στιγμιαία κατάσταση του διαύλου του χρήστη i και $I_{Q(k)=i}$ η συνάρτηση ενδείκτης του ενδεχομένου $Q(k) = i$.

2.3.2 Πολιτική Δρομολόγησης

Η αντίστοιχη πολιτική δρομολόγησης i_{RB}^* είναι η εξής:

$$i_{RB}^*(k) = \arg \max_{i \in S} \left(\frac{\rho_i(k) - \bar{\rho}_i}{c_i} \right) \quad (2.7)$$

όπου $\rho_i(k)$ ο λόγος σήματος προς θόρυβο συν παρεμβολή του χρήστη i τη χρονοθυρίδα k , $\bar{\rho}_i$ η μέση κατάσταση του διαύλου του χρήστη i και c_i θετικές σταθερές ελέγχου.

Η πολιτική (2.7) λαμβάνει υπόψιν τις στιγμιαίες και μέσες καταστάσεις των διαύλων των χρηστών, ώστε να εξυπηρετείται ο χρήστης με τις «σχετικά καλύτερες» συνθήκες διαύλου. Ο δρομολογητής επιλέγει το χρήστη i που έχει καλή ποιότητα διαύλου σε σχέση με την αναμενόμενη ποιότητα, που εκφράζει η μέση τιμή $\bar{\rho}_i$ και ταυτόχρονα προσπαθεί να επιτύχει τις απαιτήσεις δικαιοσύνης όσον αφορά το πλήθος των χρονοθυρίδων εξυπηρέτησης κάθε χρήστη.

Η ανωτέρω πολιτική προϋποθέτει πλήρη γνώση της κατάστασης του διαύλου του χρήστη, αφού προϋποτίθεται η γνώση της μέσης κατάστασης του διαύλου. Επίσης, παρατηρείται κανονικοποίηση της διαφοράς $\rho_i(k) - \bar{\rho}_i$ ως προς τις παραμέτρους ελέγχου. Για την περίπτωση που αυτές είναι ίσες για όλους τους χρήστες η πολιτική Relatively Best κυμαίνεται μεταξύ της μεγιστοποίησης της μέσης απόδοσης του συστήματος, όπως ο MAX, και της μακροπρόθεσμης δικαιοσύνης, όπως ο PF ([5]). Η ταχύτητα τότε της επίτευξης δικαιοσύνης μεταξύ των χρηστών εξαρτάται από την τιμή των παραμέτρων c_i .

Για την περίπτωση όπου τα κλάσματα του χρόνου που εκχωρούνται σε κάθε χρήστη i είναι ίσα με $1/N$, η πολιτική (2.7) εκφυλίζεται στη μορφή $i_{RB}^{**}(k) = \arg \max_{i \in N} \{ \log \rho_i - \log \bar{\rho}_i \}$, οπότε θεωρείται βέλτιστη για την επίλυση του προβλήματος (2.6) και μεγιστοποιεί τη ρυθμαπόδοση του συστήματος για $r_i(k) = \log \rho_i$ ([20]).

2.4 Σχήμα Δρομολόγησης Χρονικής Δικαιοσύνης (Temporal fairness Scheduling Scheme - TF [4])

2.4.1 Μορφοποίηση του προβλήματος

Ο πόρος του συστήματος που μοιράζεται στους χρήστες είναι ο χρόνος, δηλαδή το πλήθος των χρονοθυρίδων που εκχωρούνται σε κάθε χρήστη για εξυπηρέτηση. Άρα απαιτείται η εκχώρηση συγκεκριμένου κλάσματος των συνολικών χρονοθυρίδων σε κάθε χρήστη. Έστω f_i το ελάχιστο κλάσμα του χρόνου που πρέπει να εκχωρηθεί στο χρήστη i , όπου $f_i \geq 0$, $\sum_{i=1}^N f_i \leq 1$ και N ο αριθμός των χρηστών στην κυψέλη. Τα f_i είναι προκαθορισμένα και θεωρούνται προκαθορισμένες σταθερές δικαιοσύνης. Η τιμή του f_i δείχνει το ελάχιστο ποσό του χρόνου, κατά το οποίο πρέπει να διακινεί πληροφορία ο χρήστης i και καθορίζεται από την κλάση του χρήστη, την τιμή που πληρώνει για την ασύρματη υπηρεσία ή τις συνθήκες του διαύλου. Ο δρομολογητής επιλέγει τότε ποια χρονική σχισμή πρέπει να εκχωρηθεί σε κάθε χρήστη i , δεδομένων των χρονικών του απαιτήσεων f_i . Μαθηματικά το πρόβλημα εκφράζεται ως εξής:

$$\left. \begin{array}{l} \underset{Q \in \Theta}{\text{maximize}} E\left(U_{Q(\bar{u})}\right) \\ \text{έτσι, ώστε } P\{Q(\bar{U})=i\} \geq f_i \quad i \in N \end{array} \right\} \quad (2.8)$$

όπου Q η πολιτική δρομολόγησης και Θ το σύνολο όλων των στατικών πολιτικών δρομολόγησης.

2.4.2 Πολιτική Δρομολόγησης

Ορίζεται η πολιτική i_{TF}^* :

$$i_{TF}^*(k) = \underset{i \in S}{\text{arg max}} (U_i + v_i^*) \quad (2.9)$$

όπου v_i^* παράμετροι τέτοιες, ώστε

$$\alpha) \min_i (v_i^*) = 0$$

β) Για όλα τα i , $P\{i_{TF}^*(\bar{U}) = i\} \geq f_i$, και

γ) Για όλα τα i , αν $P\{i_{TF}^*(\bar{U}) = i\} > f_i$, τότε $v_i^* = 0$.

Οι τιμές των v_i^* καθορίζονται από την κατανομή του διανύσματος \bar{U} και τις τιμές των f_i . Στην πράξη, η κατανομή της απόδοσης \bar{U} είναι άγνωστη και έτσι είναι απαραίτητη η εκτίμηση των παραμέτρων v_i^* . Μια εκτίμηση αυτών δίνεται στο [3], όπου οι παράμετροι v_i^* υπολογίζονται από τη σχέση:

$$v_i^{k+1} = v_i^k - a^k \left(I_{\{i_{TF}^*(k)=i\}} - f_i \right) \quad (2.10)$$

όπου για παράδειγμα $a^k = \frac{1}{k}$. Έτσι, αν τεθεί $v_i^1 = 0$ σαν αρχική συνθήκη ή κάποια άλλη εκτίμηση βάσει στατιστικών μετρήσεων, υπολογίζονται οι παράμετροι v_i^* . Σημειώνεται ότι η ακολουθία $\{v_i^k\}$ συγκλίνει στο v_i^* με πιθανότητα 1.

Αποδεικνύεται ότι η ανωτέρω πολιτική αποτελεί λύση του προηγούμενου προβλήματος, δηλαδή μεγιστοποιεί τη μέση απόδοση του συστήματος κάτω από χρονικούς περιορισμούς δικαιοσύνης.

Επίσης αποδεικνύεται ότι αν οι τιμές απόδοσης-ευχαρίστησης U_i είναι μεταξύ τους ανεξάρτητες, τότε:

$$E\left(U_i I_{\{i_{TF}^*(\bar{U})=i\}}\right) \geq P\{i_{TF}^*(\bar{U}) = i\} E\{U_i\} \geq f_i E\{U_i\} \quad (2.11)$$

όπου $I_A = \begin{cases} 1, & \text{αν συμβαίνει το } A \\ 0, & \text{αλλιώς} \end{cases}$ η συνάρτηση ενδείκτης του ενδεχομένου A ,

$E\left(U_i I_{\{i_{TF}^*(\bar{U})=i\}}\right)$ η μέση απόδοση του χρήστη i χρησιμοποιώντας την ομορτυνιστική πολιτική και $P\{i_{TF}^*(\bar{U}) = i\} E\{U_i\}$ η μέση απόδοση του χρήστη i χρησιμοποιώντας μια μη ομορτυνιστική πολιτική και $P\{i_{TF}^*(\bar{U}) = i\}$ το κλάσμα των πόρων του συστήματος, εδώ του χρόνου, που εκχωρείται στο χρήστη i . Η συνθήκη (2.11) εξασφαλίζει ότι, εφόσον υπάρχει ανεξαρτησία στις τιμές απόδοσης των χρηστών, η

οπορτουνιστική δρομολόγηση δεν είναι χειρότερη από κάποια μη οπορτουνιστική πολιτική που εκχωρεί το ίδιο μέρος της πηγής στο χρήστη, όσον αφορά τη μέση απόδοση κάθε χρήστη. Συνεπώς, κάθε χρήστης αντιλαμβάνεται καλύτερη μέση απόδοση και ταυτόχρονα πετυχαίνει ένα ελάχιστο επίπεδο απόδοσης – εκείνο μιας μη οπορτουνιστικής πολιτικής.

2.5 Σχήμα Δρομολόγησης Χρηστικής Δικαιοσύνης (Utilitarian Fairness Scheduling Scheme - UF [4])

Η συγκεκριμένη πολιτική δρομολόγησης έχει σκοπό να εξασφαλίσει πως όλοι οι χρήστες του συστήματος λαμβάνουν τουλάχιστον ένα συγκεκριμένο ποσοστό της συνολικής μέσης απόδοσης του συστήματος.

2.5.1 Μορφοποίηση του προβλήματος

Ισχύει ότι $E\left(U_i I_{\{Q(\bar{v})=i\}}\right)$ είναι η μέση απόδοση του χρήστη i κάνοντας χρήση της πολιτικής Q και $E\left(U_{Q(\bar{v})}\right) = \sum_{i=1}^N E\left(U_i I_{\{Q(\bar{v})=i\}}\right)$ είναι η συνολική απόδοση του συστήματος χρησιμοποιώντας την πολιτική Q . Έστω a_i το ελάχιστο κλάσμα της συνολικής μέσης απόδοσης που απαιτείται από το χρήστη i , όπου $a_i \geq 0$ για όλα τα i , και $\sum_i a_i \leq 1$. Τότε το μαθηματικό πρόβλημα είναι το εξής:

$$\left. \begin{array}{l} \text{maximize } E\left(U_{Q(\bar{v})}\right) \\ \text{έτσι, ώστε } E\left(U_i I_{\{Q(\bar{v})=i\}}\right) \geq a_i E\left(U_{Q(\bar{v})}\right) \quad i \in N \end{array} \right\} \quad (2.12)$$

όπου Θ το σύνολο όλων των στατικών πολιτικών δρομολόγησης. Τα a_i είναι προκαθορισμένες παράμετροι δικαιοσύνης και $\varepsilon' = \sum_i a_i$ είναι μια ρυθμιστική παράμετρος, της οποίας όσο μικρότερη είναι η τιμή τόσο μεγαλύτερη είναι η πιθανότητα να βελτιωθεί η απόδοση του συστήματος.

Μια ειδική περίπτωση του ανωτέρω ομορτιστικού προβλήματος είναι η θεώρηση της μεγιστοποίησης της ελάχιστης σταθμισμένης απόδοσης των χρηστών,

δηλαδή $\frac{E\left(U_i I_{\{Q(\bar{v})=i\}}\right)}{\beta_i}$, όπου β_i προκαθορισμένα θετικά βάρη. Αυτό είναι

ισοδύναμο με την περίπτωση όπου τίθεται $a_i = \frac{\beta_i}{\sum_j \beta_j}$ στο πρόβλημα χρηστικής

δικαιοσύνης. Σε αυτήν την περίπτωση, επειδή $\sum_i a_i = 1$ όλοι οι ανισοτικοί περιορισμοί είναι ενεργοί για κάθε δυνατή λύση, δηλαδή οι ανισοτικοί περιορισμοί ικανοποιούνται με ισότητα.

2.5.2 Πολιτική Δρομολόγησης

Ορίζεται η πολιτική i_{UF}^* ως εξής:

$$i_{UF}^*(k) = \arg \max_{i \in S} \left((\kappa + v_i^*) U_i \right) \quad (2.13)$$

όπου $\kappa = 1 - \sum_{i=1}^N a_i v_i^*$ και τα v_i^* είναι πραγματικές παράμετροι, τέτοιες, ώστε:

α) $\min(v_i^*) = 0$

β) Για όλα τα i , $E\left(U_i I_{\{i_{UF}^*(\bar{v})=i\}}\right) \geq a_i E\left(U_{i_{UF}^*(\bar{v})}\right)$

γ) Για όλα τα i , αν $E\left(U_i I_{\{i_{UF}^*(\bar{v})=i\}}\right) > a_i E\left(U_{i_{UF}^*(\bar{v})}\right)$, τότε $v_i^* = 0$.

Αποδεικνύεται ότι η πολιτική δρομολόγησης (2.13) είναι λύση του προβλήματος χρηστικής δικαιοσύνης, όπως αυτό τέθηκε νωρίτερα. Όμοια με τις v_i^* στο σχήμα δρομολόγησης χρονικής δικαιοσύνης η παράμετρος v_i^* θεωρείται ρυθμιστική για την ικανοποίηση των περιορισμών χρηστικής δικαιοσύνης. Η βέλτιστη πολιτική δρομολόγησης πάντα επιλέγει το σχετικά καλύτερο χρήστη να εξυπηρετηθεί. Σε αυτήν την περίπτωση καλύτερος χρήστης είναι ο i , αν $(\kappa + v_i^*) U_i = \max_j (\kappa + v_j^*) U_j$, όπου το κ είναι σταθερά για όλους τους χρήστες. Όπως πριν, αν $v_i^* > 0$, τότε ο

χρήστης i είναι «άτυχος» και η μέση απόδοσή του ισούται με την ελάχιστη απαιτούμενη, δηλαδή $E\left(U_i I_{\{i_{UF}^*(\bar{U})=i\}}\right) = a_i E\left(U_{i_{UF}^*(\bar{U})}\right)$.

Κάθε πολιτική που ικανοποιεί τον περιορισμό του προβλήματος χρηστικής δικαιοσύνης έχει την ιδιότητα:

$$\frac{a_i}{1-\varepsilon'+a_j} \leq \frac{E\left(U_i I_{\{i_{UF}^*(\bar{U})=i\}}\right)}{E\left(U_j I_{\{i_{UF}^*(\bar{U})=j\}}\right)} \leq \frac{1-\varepsilon'+a_i}{a_j} \quad (2.14)$$

για $i, j = 1, 2, \dots, N$. Με άλλα λόγια ο περιορισμός χρηστικής δικαιοσύνης ελέγχει τη μέγιστη μείωση των τιμών απόδοσης μεταξύ των χρηστών.

Ο περιορισμός χρηστικής δικαιοσύνης εξασφαλίζει ότι ένας χρήστης δέχεται τουλάχιστον ένα συγκεκριμένο μέρος της συνολικής απόδοσης και έτσι σε κάποιες περιπτώσεις είναι καταλληλότερος από τον περιορισμό χρονικής δικαιοσύνης. Ωστόσο ένα σημαντικό μειονέκτημα του αλγορίθμου χρηστικής δικαιοσύνης είναι η επίπτωση ενός χρήστη με χαμηλής ποιότητας διάυλο στη συνολική απόδοση του συστήματος. Από τον περιορισμό του προβλήματος (2.12) παρατηρείται

$$E\left(U_{i_{UF}^*(\bar{U})}\right) \leq \frac{E(U_i)}{a_i}.$$

Έτσι, ένας χρήστης με κακές συνθήκες διαύλου (πολύ μικρή

τιμή $E(U_i)$) και με μεγάλη τιμή a_i , θα υποβάθμιζε σημαντικά τη συνολική απόδοση του συστήματος, διότι πρέπει να του εκχωρηθεί ένα σημαντικό ποσοστό του συνόλου των χρονοθυρίδων, προκειμένου να ικανοποιήσει τις απαιτήσεις του σε δικαιοσύνη. Μια λύση θα μπορούσε να είναι μια προσαρμοστική στρατηγική κατώφλιου. Συγκεκριμένα, αν:

$$\frac{E\left(U_i I_{\{i_{UF}^*(\bar{U})=i\}}\right)}{P\{i_{UF}^*(\bar{U})=i\} E\left(U_{i_{UF}^*(\bar{U})}\right)} \leq B \quad (2.15)$$

όπου B ένα προκαθορισμένο κατώφλι, τότε είναι δυνατή η μείωση των τιμών των a_i , επειδή ο χρήστης i δεν μπορεί να χρησιμοποιήσει αποδοτικά το λιγοστό φάσμα.

2.6 Σχήμα δρομολόγησης εγγύησης ελάχιστης απόδοσης (Minimum Performance Guarantee Scheduling Scheme - MP [4])

Εδώ στόχος είναι η μεγιστοποίηση της απόδοσης του συστήματος με την εξασφάλιση ελάχιστου επιπέδου απόδοσης για κάθε χρήστη. Έτσι, αποφεύγεται ο αποκλεισμός των χρηστών που έχουν δυσμενή κατάσταση διαύλου κατά το μεγαλύτερο μέρος του χρόνου που είναι ενεργοί, ενώ παράλληλα οι πόροι που απομένουν μετά την εξασφάλιση της ελάχιστης ρυθμαπόδοσης κατανέμονται στους χρήστες με τη στιγμιαία καλύτερη κατάσταση διαύλου με σκοπό τη μεγιστοποίηση της ρυθμαπόδοσης του συστήματος.

2.6.1 Μορφοποίηση του προβλήματος

Έστω ότι κάθε χρήστης έχει ελάχιστες απαιτήσεις απόδοσης C_i και $\vec{C} = \{C_1, C_2, \dots, C_N\}$ το διάνυσμα των απαιτήσεων σε απόδοση. Το πρόβλημα τίθεται μαθηματικά ως εξής:

$$\left. \begin{array}{l} \text{maximize } E\left(U_{Q(\vec{v})}\right) \\ \text{έτσι, ώστε } E\left(U_i I_{\{Q(\vec{v})=i\}}\right) \geq C_i \quad i \in N \end{array} \right\} \quad (2.16)$$

όπου Q η πολιτική δρομολόγησης, Θ είναι το σύνολο όλων των στατικών πολιτικών.

Εδώ εκτός από το πρόβλημα του καθορισμού της πολιτικής, η οποία μεγιστοποιεί τη συνολική απόδοση του συστήματος υπό κάποιες απαιτήσεις δεδομένου βαθμού ποιότητας (QoS), ανακύπτει και το πρόβλημα της επαλήθευσης ότι το \vec{C} είναι πραγματοποιήσιμο διάνυσμα απαιτήσεων (feasible). Στα προβλήματα χρονικής και χρηστικής δικαιοσύνης οι λύσεις εξασφαλίζονται ως πραγματοποιήσιμες από τις συνθήκες $\sum_i f_i$ και $\sum_i a_i$ αντίστοιχα. Παρ' όλ' αυτά, υπάρχουν ρυθμίσεις με τις οποίες το πρόβλημα ξεπερνάται, π.χ. $C_i = \psi_i E(U_i)$, όπου $\psi_i \geq 0$ και $\sum_i \psi_i \leq 1$. Μια τέτοια ρύθμιση ικανοποιείται από μια μη ομορτονιστική πολιτική δρομολόγησης, στην οποία ο χρήστης i επιλέγεται να εξυπηρετηθεί σε δεδομένη χρονοθυρίδα με

πιθανότητα ψ_i , οπότε είναι πραγματοποιήσιμο και για ομορτυνιστικές πολιτικές δρομολόγησης.

2.6.2 Πολιτική Δρομολόγησης

Υποθέτουμε ότι διάλυμα απαιτήσεων είναι πραγματοποιήσιμο (feasible). Ορίζεται η πολιτική $i_{MP}^*(k)$ ως εξής:

$$i_{MP}^*(k) = \arg \max_{i \in S} (\alpha_i^* U_i) \quad (2.17)$$

όπου τα α_i^* είναι πραγματικές παράμετροι που ικανοποιούν τις παρακάτω συνθήκες:

α) $\min_i (\alpha_i^*) = 1$

β) Για όλα τα i , $E \left(U_i I_{\{i_{MP}^*(\bar{v})=i\}} \right) \geq C_i$

γ) Για όλα τα i , αν $E \left(U_i I_{\{i_{MP}^*(\bar{v})=i\}} \right) > C_i$, τότε $\alpha_i^* = 1$.

Αποδεικνύεται ότι η παραπάνω πολιτική είναι λύση του προηγούμενου προβλήματος εξασφάλισης ελάχιστων απαιτήσεων απόδοσης. Η παράμετρος $\bar{\alpha}$ ρυθμίζει τις τιμές απόδοσης των χρηστών και η πολιτική δρομολόγησης (2.17) επιλέγει το σχετικά καλύτερο χρήστη για τον οποίο ισχύει $\alpha_i^* U_i = \max_j \alpha_j^* U_j$. Αν ο συντελεστής ελέγχου για ένα χρήστη είναι μεγαλύτερος από τη μονάδα, τότε ο χρήστης είναι «άτυχος» και δέχεται μόνο μια μέση τιμή απόδοσης που ισούται με τις ελάχιστες απαιτήσεις του σε απόδοση.

Η ανωτέρω ομορτυνιστική πολιτική επικρατεί των μη ομορτυνιστικών. Θεωρώντας μια μη ομορτυνιστική πολιτική που εκχωρεί στο χρήστη i ποσοστό ψ_i της πηγής (χρονοθυρίδες), όπου $\sum_i \psi_i = 1$ και ο χρήστης i παίρνει μια μέση τιμή απόδοσης $\psi_i E(U_i)$. Τότε $C_i = \psi_i E(U_i)$ για όλα τα i , οπότε είναι πραγματοποιήσιμο. Η ομορτυνιστική πολιτική πάντα καταλήγει σε μη χειρότερες τιμές απόδοσης σε σχέση με μια μη ομορτυνιστική πολιτική, θεωρώντας αμελητέο κόστος σηματοδότησης. Σημειώνεται, τέλος ότι μια πολιτική δρομολόγησης της μορφής

$i_{MP}^*(k) = \arg \max_i (a_i U_i)$ είναι βέλτιστη και για άλλους τύπους προβλημάτων δρομολόγησης.

Οι προαναφερθέντες αλγόριθμοι δρομολόγησης σχεδιάστηκαν, για να εξυπηρετήσουν με διαφορετικά μακροπρόθεσμα κριτήρια δικαιοσύνης τους χρήστες υπηρεσιών μη πραγματικού χρόνου καθυστέρησης εξυπηρέτησης. Έτσι, το κριτήριο με το οποίο δρομολογούνται οι χρήστες είναι η μεγιστοποίηση της συνολικής ρυθμαπόδοσης και η επίτευξη δικαιοσύνης ανάμεσα στους χρήστες. Αντίθετα, οι παραπάνω αλγόριθμοι δεν προσφέρονται για υψηλών απαιτήσεων υπηρεσίες πραγματικού χρόνου, οπότε δημιουργήθηκε η ανάγκη σχεδιασμού νέων αλγορίθμων δρομολόγησης.

3

ΚΥΡΙΟΙ ΑΛΓΟΡΙΘΜΟΙ

ΧΡΟΝΟΔΡΟΜΟΛΟΓΗΣΗΣ ΧΡΗΣΤΩΝ

ΥΠΗΡΕΣΙΩΝ ΠΡΑΓΜΑΤΙΚΟΥ ΧΡΟΝΟΥ

Οι αλγόριθμοι που περιγράφηκαν στο προηγούμενο κεφάλαιο προσπαθούν να επιτύχουν είτε κριτήρια δικαιοσύνης είτε ελάχιστης εξυπηρέτησης των χρηστών του συστήματος και έχουν σχεδιαστεί για υπηρεσίες μη πραγματικού χρόνου. Αδυνατούν όμως να υποστηρίξουν υπηρεσίες πραγματικού χρόνου, καθώς απαιτούν διαφορετικά κριτήρια ποιότητας υπηρεσίας, όπως αυτά αποδίδονται στην ενότητα 1.5.3. Για το λόγο αυτό προτάθηκαν νέοι γενιά αλγόριθμοι που υποστηρίζουν υπηρεσίες ευαίσθητες στην καθυστέρηση.

3.1 Ο αλγόριθμος «Πρώτα η Τροποποιημένη Μέγιστη Σταθμισμένη Καθυστέρηση (Modified Largest Weighted Delay First – M-LWDF)

Ο συγκεκριμένος αλγόριθμος επιλέγει τον επόμενο προς εξυπηρέτηση χρήστη υπηρεσίας λαμβάνοντας υπόψιν την καθυστέρηση που έχει υποστεί το προς αποστολή πακέτο στην ουρά του μέχρι την παρούσα χρονοθυρίδα, καθώς και την κατάσταση του διαύλου του χρήστη. Μαθηματικά ο τρόπος με τον οποίο επιλέγεται ο εκάστοτε προς εξυπηρέτηση χρήστης αποδίδεται ως εξής:

$$i_{MLDWF}^*(k) = \max_{i \in S} \{ \delta_i d_i(k) \rho_i(k) \} \quad (3.1)$$

όπου i_{MLDWF}^* η πολιτική δρομολόγησης, δ_i θετική σταθερά για το χρήστη i , $d_i(k)$ η καθυστέρηση εξυπηρέτησης του χρήστη i τη χρονοθυρίδα k και $\rho_i(k)$ η κατάσταση του διαύλου του χρήστη τη χρονοθυρίδα k . Ως $d_i(k)$ χρησιμοποιείται ο αριθμός των χρονοθυρίδων που παρήλθαν από την τελευταία εξυπηρέτηση του χρήστη i και ως $\rho_i(k)$ ο σηματοθορυβικός λόγος που επιτυγχάνει ο χρήστης i .

Η επιλογή των παραμέτρων δ_i επιτρέπει τον έλεγχο της κατανομής της καθυστέρησης των πακέτων για διαφορετικούς χρήστες. Αύξηση της παραμέτρου δ_i του χρήστη i διατηρώντας σταθερές τις αντίστοιχες δ_j παραμέτρους των υπόλοιπων χρηστών μειώνει την καθυστέρηση των πακέτων για τη συγκεκριμένη ροή i με κόστος την αύξηση των καθυστερήσεων των λοιπών ροών. Έτσι είναι δυνατή η μορφοποίηση της κατανομής των καθυστερήσεων μεταξύ των χρηστών.

Αποδεικνύεται στο [13] ότι ο αλγόριθμος M-LWDF είναι βέλτιστος υπό την έννοια ότι μπορεί να διατηρήσει όλες τις ουρές σταθερές, αν αυτό είναι εφικτό με οποιονδήποτε αλγόριθμο. Επιπλέον, παραμένει βέλτιστος σε απόδοση, αν η καθυστέρηση $d_i(k)$ όλων ή μερικών χρηστών αντικατασταθεί από το μήκος της ουράς των δεδομένων. Από αυτή του την ιδιότητα απορρέει ότι η μακροπρόθεσμη ρυθμαπόδοση του συστήματος είναι ίση με το ρυθμό άφιξης πακέτων στο σύστημα.

Χαρακτηριστικό του αλγορίθμου M-LWDF είναι ότι η απόφαση δρομολόγησης έγκειται τόσο στις στιγμιαίες συνθήκες του διαύλου του χρήστη όσο και στην κατάσταση των ουρών. Επιπλέον, η συγκεκριμένη πολιτική κρίνεται εύκολη στην

υλοποίηση, καθώς ο δρομολογητής πρέπει απλώς να σημαδεύει τα εισερχόμενα πακέτα όλων των χρηστών ή να διατηρεί πληροφορία για το στιγμιαίο μήκος των ουρών τους.

➤ Περιορισμοί καθυστέρησης

Όσον αφορά την τήρηση των κριτηρίων εγγύησης ποιότητας υπηρεσιών πραγματικού χρόνου, ο αλγόριθμος M-LWDF δεν εγγυάται ότι ικανοποιεί τους περιορισμούς καθυστέρησης από μόνος του, αλλά εξαρτάται από την επιλογή των παραμέτρων δ_i . Μια ορθή επιλογή για τους περιορισμούς καθυστέρησης δίνεται από την παρακάτω σχέση:

$$\delta_i = \frac{a_i}{\bar{\rho}_i} \quad (3.2)$$

όπου:

$$a_i = -\frac{\log g_i^l}{T_i} \quad (3.3)$$

και g_i^l το μέγιστο της πιθανότητας αστοχίας καθυστέρησης του χρήστη i , T_i το κατώφλι καθυστέρησης του χρήστη i και $\bar{\rho}_i$ η μέση κατάσταση του διαύλου του χρήστη i .

Η παράμετρος a_i της σχέσης (3.3) ενσωματώνει τις απαιτήσεις QoS των χρηστών και παρέχει διαφοροποιημένους βαθμούς ποιότητας υπηρεσίας μεταξύ τους. Η τιμή της a_i , όπως φαίνεται και από τον τύπο υπολογισμού της (3.3), διαμορφώνεται με βάση τις τιμές του κατωφλίου καθυστέρησης και του μεγίστου της πιθανότητας αστοχίας καθυστέρησης, τα οποία αντανακλούν διαφορετικούς βαθμούς ποιότητας υπηρεσίας, ανάλογα με τις τιμές που τους ανατίθενται.

Επιπρόσθετα, η επιλογή του προς εξυπηρέτηση χρήστη βασίζεται στο μέγιστο του γινομένου $a_i \cdot d_i \cdot \left(\frac{\rho_i(k)}{\bar{\rho}_i} \right)$. Όσο μεγαλύτερη είναι η καθυστέρηση εξυπηρέτησης του χρήστη i ή η ποιότητα του διαύλου σε σχέση με τη μέση τιμή του ή όσο υψηλότερη είναι η απαίτηση του βαθμού εξυπηρέτησής του, τόσο πιθανότερο είναι να επιλεγεί ο χρήστης i για εξυπηρέτηση. Συνεπώς, η πολιτική φιλοξενεί ικανοποιητικά όλους τους χρήστες, οι οποίοι μπορούν να εξυπηρετηθούν από

οποιαδήποτε άλλη πολιτική δρομολόγησης τηρώντας παράλληλα τις απαιτήσεις τους σε καθυστέρηση.

➤ Περιορισμοί ρυθμού μετάδοσης

Όσον αφορά την εγγύηση ενός ελάχιστου ρυθμού μετάδοσης r_i για κάθε χρήστη, αυτή παρουσιάζει ενδιαφέρον για τους χρήστες υπηρεσιών μη πραγματικού χρόνου και είναι δυνατή με τον έλεγχο μιας ουράς συμβόλων. Έτσι, με μικρές αλλαγές στην υπό εξέταση πολιτική, μπορούμε να επιτύχουμε τον παραπάνω στόχο. Αναλυτικότερα, αν για κάθε χρήστη i θεωρείται ένα εικονικό δοχείο και τα σύμβολα φτάνουν στο δοχείο με σταθερό ρυθμό r_i , η μόνη τροποποίηση που χρειάζεται η πολιτική M-LWDF είναι ότι ως $d_i(k)$ θεωρείται πλέον η καθυστέρηση του συμβόλου που περιμένει μεγαλύτερο χρονικό διάστημα στο δοχείο i . Στο τέλος κάθε χρονοθυρίδας που εξυπηρετείται ο χρήστης i , αφαιρείται από το δοχείο της πραγματικής ουράς του ο αριθμός συμβόλων που αντιστοιχεί στο πραγματικό ποσό δεδομένων που διακινήθηκε. Από τη στιγμή που τα σύμβολα καταφτάνουν με σταθερό ρυθμό, ισχύει $d_i(k) = \frac{\text{αριθμός συμβόλων στο δοχείο } i}{r_i}$. Σε ειδικές

περιπτώσεις, όπως η επιλογή προς εξυπηρέτηση ενός χρήστη που δεν έχει αρκετά δεδομένα να διακινήσει, το δοχείο μειώνεται σαν να υπήρχαν αρκετά δεδομένα, ενώ στην περίπτωση που ο προς εξυπηρέτηση χρήστης δεν έχει καθόλου δεδομένα, το δοχείο μειώνεται κατά το αντίστοιχο ποσό και ο δρομολογητής επιλέγει το δεύτερο καλύτερο χρήστη να εξυπηρετήσει. Αν πάλι, ένα δοχείο περιέχει λιγότερα δεδομένα από αυτά που πρέπει να απομακρυνθούν, το περιεχόμενό του τίθεται ίσο με μηδέν.

Σε κάθε περίπτωση απαιτείται μόνο η πληροφορία για τον αριθμό των συμβόλων σε κάθε εικονικό δοχείο, κάτι πολύ απλό στην υλοποίηση, η οποία μπορεί να επιτευχθεί μέσω ενός μετρητή σε λογισμικό. Αποδεικνύεται ότι, αν οι ουρές των συμβόλων είναι σταθερές, τότε η πραγματική απόδοση κάθε χρήστη είναι τουλάχιστον r_i . Άρα, η πολιτική εγγυάται ελάχιστη ρυθμαπόδοση r_i για κάθε χρήστη.

Επιπρόσθετα, αυτή η πολιτική παρέχει απομόνωση μεταξύ των χρηστών. Αφού η δρομολόγηση βασίζεται σε ουρές συμβόλων ή πακέτων, μεγάλη εκρηκτικότητα κίνησης ενός συγκεκριμένου χρήστη δεν επηρεάζει την ελάχιστη ρυθμαπόδοση των υπόλοιπων χρηστών.

Οι παράμετροι δ_i επιτρέπουν τον έλεγχο της χρονικής κλίμακας στην οποία παρέχεται η ελάχιστη ρυθμαπόδοση. Μεγαλύτερες τιμές των γ_i οδηγούν σε ταχύτερη επίτευξη του ελάχιστου ρυθμού μετάδοσης και άρα επίτευξη βραχυπρόθεσμης δικαιοσύνης σε ρυθμαπόδοση. Στενωπός της τεχνικής στο συγκεκριμένο θέμα είναι η επιλογή των κατάλληλων τιμών των παραμέτρων δ_i .

Τέλος, η πολιτική δρομολόγησης M-LWDF παρέχει ευελιξία ως προς την επιλογή του είδους των απαιτήσεων ποιότητας υπηρεσίας, δηλαδή ο σταθμός βάσης μπορεί να επιλέξει ανάμεσα σε αυστηρές απαιτήσεις καθυστέρησης και εγγύηση ελάχιστου ρυθμού μετάδοσης με βάση την εκάστοτε υπηρεσία ενός χρήστη με κατάλληλη κανονικοποίηση των αντίστοιχων παραμέτρων $d_i(k)$. Στην πρώτη περίπτωση ως $d_i(k)$ χρησιμοποιείται η καθυστέρηση του επικεφαλής πακέτου στην ουρά αναμονής του χρήστη i , ενώ στη δεύτερη περίπτωση η μεγαλύτερη καθυστέρηση συμβόλου στο δοχείο i , σύμφωνα με τις παραπάνω ορολογίες. Σε κάθε περίπτωση όμως δεν είναι δυνατή η εγγύηση των δύο κριτηρίων ταυτόχρονα.

3.2 Ο αλγόριθμος Εκθετικής Πολιτικής Δρομολόγησης (The Exponential Rule – EXP)

Ο συγκεκριμένος αλγόριθμος ανήκει επίσης στην οικογένεια των πορτουνιστικών αλγορίθμων χρονοδρομολόγησης. Διακρίνονται δύο εκδοχές της εκθετικής πολιτικής δρομολόγησης, η πολιτική εκθετικού μήκους ουράς (Exponential Queue Length Rule – EXP-Q) και η πολιτική εκθετικού χρόνου αναμονής (Exponential Waiting Time Rule – EXP-W). Η πολιτική EXP-Q χρησιμοποιείται για την περίπτωση που η κίνηση κάθε χρήστη αντιπροσωπεύεται από ένα δοχείο και ως $Q_i(k)$ θεωρείται η μεγαλύτερη καθυστέρηση μεταξύ των πακέτων (συμβόλων) του δοχείου. Αντίστοιχα, η πολιτική EXP-W ενδείκνυται για τις περιπτώσεις όπου η κίνηση κάθε χρήστη αντιπροσωπεύεται από μια ουρά αναμονής, οπότε και μας ενδιαφέρει η καθυστέρηση $d_i(k)$ του επικεφαλής πακέτου της ροής. Οι αντίστοιχες μαθηματικές εκφράσεις είναι οι ακόλουθες:

$$i_{EXP-Q}^*(k) = \arg \max_{i \in S} \left\{ \delta_i \cdot \rho_i(k) \cdot \exp \left(\frac{a_i Q_i(k)}{\beta + [\bar{Q}_i(k)]^\eta} \right) \right\} \quad (\text{EXP-Q}) \quad (3.4)$$

$$i_{EXP-W}^*(k) = \arg \max_{i \in S} \left\{ \delta_i \cdot \rho_i(k) \cdot \exp \left(\frac{a_i d_i(k)}{\beta + [\bar{d}_i(k)]^\eta} \right) \right\} \quad (\text{EXP-W}) \quad (3.5)$$

όπου δ_i , a_i , β θετικές σταθερές και $\eta \in (0,1)$ σταθερό, $\rho_i(k) \equiv \rho_i^{m(k)}$ ο ρυθμός μετάδοσης που αντιστοιχεί στην κατάσταση m του διαύλου τη χρονοθυρίδα k , $\bar{Q}_i(k) \doteq \frac{1}{N} \sum_{i=1}^N a_i Q_i(k)$ και $\bar{d}_i(k) = \frac{1}{N} \sum_{i=1}^N a_i d_i(k)$, N ο αριθμός των χρηστών στο σύστημα. Παρακάτω, η μελέτη εστιάζεται στην πολιτική EXP-W, καθώς τα συμπεράσματα και για την πολιτική EXP-Q προκύπτουν ομοίως.

Η επιλογή των δ_i δίνεται, όπως και στον M-LWDF, από τη σχέση (3.2). Οι τιμές των a_i , $i = 1, 2, \dots, N$, καθορίζουν την ταλάντευση μεταξύ της μείωσης των σταθμισμένων καθυστερήσεων και της αναλογικής δικαιοσύνης για μικρές τιμές των καθυστερήσεων. Για μικρές ανοχές σε καθυστέρηση, τιμές που αποδίδουν ικανοποιητικά στην πράξη δίνονται, όπως και στον M-LDWF, από τη σχέση (3.3).

Η επιλογή των β δεν επηρεάζει ουσιαστικά τον αλγόριθμο, καθώς είναι ίδιες για όλους τους χρήστες και συνήθως λαμβάνουν την τιμή 1. Αντίστοιχα, η τιμή της παραμέτρου η που χρησιμοποιείται και στη βιβλιογραφία για προσομοιώσεις είναι 0.5.

Ο αλγόριθμος EXP είναι βέλτιστος σε ρυθμαπόδοση (throughput optimal), ανεξάρτητα από την επιλογή των παραμέτρων a_i , δ_i , β , $\eta \in (0,1)$ για $i \in N$ ([14]), δηλαδή διατηρεί σταθερές τις ουρές στο σταθμό βάσης, αν αυτό είναι δυνατό από οποιονδήποτε άλλο αλγόριθμο και ταυτόχρονα μεγιστοποιεί τη συνολική ρυθμαπόδοσή του. Συνεπώς, η μακροπρόθεσμη ρυθμαπόδοση είναι ίση με το ρυθμό άφιξης πακέτων στο σύστημα.

Επιπρόσθετα, η πολιτική δρομολόγησης EXP προσπαθεί να εξισώσει τις σταθμισμένες καθυστερήσεις $a_i d_i(k)$ όλων των χρηστών του συστήματος, όταν εμφανίζουν σημαντικές διαφορές. Αν ένας χρήστης βιώνει μεγαλύτερη καθυστέρηση από τους υπόλοιπους και μάλιστα η διαφορά της καθυστέρησής του σε σχέση με τους

υπόλοιπους χρήστες είναι μεγαλύτερη από $[\bar{d}_i(k)]^n$, τότε ο εκθετικός όρος λαμβάνει μεγάλες τιμές και «παραμερίζει» τις συνθήκες που επικρατούν στο διάυλο του συγκεκριμένου χρήστη, αρκεί αυτές να μην οδηγούν βεβαίως σε μηδενικό ρυθμό μετάδοσης, δίνοντας υψηλότερη προτεραιότητα σε αυτό. Από την άλλη μεριά, για μικρές τιμές διαφορών καθυστέρησης, ο εκθετικός όρος είναι κοντά στη μονάδα και η πολιτική τείνει σε μια πολιτική αναλογικής δρομολόγησης (PF). Άρα η πολιτική δρομολόγησης EXP ταλαντεύεται ανάμεσα σε μια πολιτική αναλογικής δρομολόγησης και μια πολιτική που εξισορροπεί τις καθυστερήσεις των χρηστών του συστήματος. Ο παράγοντας β είναι παρών στον παρονομαστή του εκθέτη των σχέσεων (3.4) και (3.5) για την αποφυγή απειρισμού του όρου, όταν οι καθυστερήσεις είναι μικρές.

➤ Περιορισμοί καθυστέρησης

Το κριτήριο ποιότητας για το συγκεκριμένο αλγόριθμο είναι της μορφής (1.6). Ο αλγόριθμος EXP, όπως τονίστηκε στην παραπάνω παράγραφο, επιδιώκοντας την εξίσωση των σταθμισμένων καθυστερήσεων όλων των χρηστών υπηρεσιών πραγματικού χρόνου του συστήματος, επιχειρεί τη διατήρηση αυτών σε επίπεδα χαμηλότερα από το κατώφλι T_i για κάθε χρήστη. Αξίζει να σημειωθεί ότι η χρονοδρομολόγηση που πραγματοποιεί ο EXP συμπεριφέρεται δίκαια σε όλους τους χρήστες ως προς την καθυστέρηση, δηλαδή όχι μόνο επιδιώκει χαμηλή καθυστέρηση, αλλά το επιδιώκει για όλους τους χρήστες υπηρεσίας πραγματικού χρόνου. Έτσι δεν εμφανίζονται φαινόμενα πολύ μικρής καθυστέρησης για ένα χρήστη και ιδιαίτερα αυξημένης για κάποιον άλλο.

➤ Περιορισμοί ελάχιστου ρυθμού μετάδοσης

Για την εξυπηρέτηση των χρηστών υπηρεσίας μη πραγματικού χρόνου που τυχόν υπάρχουν στο σύστημα ο αλγόριθμος δρομολόγησης EXP εγγυάται ένα ελάχιστο ρυθμό μετάδοσης δεδομένων, ο οποίος αναφέρεται σε μακροπρόθεσμα διαστήματα. Αυτό υλοποιείται με τη βοήθεια ενός δοχείου συμβόλων όμοιου με αυτό που περιγράφηκε στην ενότητα 3.1 για το χειρισμό των χρηστών μη πραγματικού χρόνου από τον αλγόριθμο M-LWDF.

3.3 Ο αλγόριθμος Οπορτουνιστικής Δρομολόγησης Εκχώρησης Βασισμένης στην Πιθανότητα Πρόσβασης (Access Probability based Assignment Opportunistic Scheduling – APAOS)

Ο αλγόριθμος APAOS βασίζεται σε μια σημαντική παρατήρηση που εισάγεται στο [5]. Για δεδομένο αριθμό χρηστών N στο σύστημα και για σταθερές μέσες τιμές ποιότητας του διαύλου αυτών, η πιθανότητα εξυπηρέτησης των χρηστών υπό τις πολιτικές (2.2), (2.4) και (2.7) είναι σταθερή και σχετική με την πιθανότητα των χρηστών να έχουν καλύτερες συνθήκες διαύλου την επόμενη χρονοθυρίδα. Στόχος του APAOS είναι να αναγκάσει τη διαδικασία επιλογής των προς εξυπηρέτηση χρηστών να καταλήξει στα ίδια αποτελέσματα σαν να εξυπηρετούσε όλους τους χρήστες με αυτές τις σταθερές πιθανότητες, προκειμένου να επιτύχει υψηλές επιδόσεις ρυθμαπόδοσης και ταυτόχρονα να ικανοποιήσει τις απαιτήσεις ποιότητας υπηρεσίας των χρηστών. Η μαθηματική έκφραση του αλγορίθμου είναι:

$$i_{APAOS}^*(k) = \arg \max_{i \in S} \left\{ \frac{\rho_i(k)}{\bar{\rho}_i} + \ln(B_i) \right\} \quad (3.6)$$

όπου i_{APAOS}^* η πολιτική δρομολόγησης που χρησιμοποιείται, $\rho_i(k)$ ο λόγος σήματος προς θόρυβο συν παρεμβολή (SINR) του χρήστη i τη χρονοθυρίδα k , $\bar{\rho}_i$ η μέση τιμή SINR του χρήστη i που χαρακτηρίζει την κατάσταση του διαύλου του, B_i κατάλληλα ρυθμισμένο βάρος του χρήστη i ανάλογα με τη βαρύτητα εξυπηρέτησης που πρέπει να δοθεί στους χρήστες.

Γίνεται φανερό ότι σε αντίθεση με τους μέχρι τώρα εισαχθέντες αλγορίθμους, όπου ο πόρος του συστήματος ήταν ο χρόνος, ο επιπλέον πόρος που μοιράζεται στους χρήστες από το συγκεκριμένο αλγόριθμο είναι η πιθανότητα εξυπηρέτησης. Υπό αυτές τις συνθήκες, το σύστημα μπορεί να εξυπηρετήσει χρήστες υπηρεσίας πραγματικού ή μη χρόνου, στηριζόμενο στις πιθανότητες εξυπηρέτησης που επιβάλλονται για τον κάθε χρήστη και δίνοντας προτεραιότητα στους χρήστες υπηρεσίας πραγματικού χρόνου.

Οι τιμές των πιθανοτήτων είναι τέτοιες, ώστε αν η επιλογή των χρηστών βασιζόταν μόνο σε αυτές, τηρούνται οι αντίστοιχες απαιτήσεις ποιότητας υπηρεσίας πραγματικού χρόνου, όπως αυτές εκφράζονται από τη σχέση (1.6). Ο

μετασχηματισμός των απαιτήσεων των χρηστών σε σταθερές πιθανότητες εξυπηρέτησης την επόμενη χρονοθυρίδα επιτρέπει τον υπολογισμό των βαρών B_i , τα οποία και οδηγούν στην ικανοποίηση των αντίστοιχων περιορισμών που θέτει η εκάστοτε υπηρεσία πραγματικού και μη χρόνου.

Οι πιθανότητες εξυπηρέτησης υπολογίζονται ως ακολούθως. Έστω x_i τυχαία μεταβλητή του γεγονότος ο χρήστης i υπηρεσίας πραγματικού χρόνου με πιθανότητα εξυπηρέτησης την επόμενη χρονοθυρίδα $p_{RT,i}$ να εξυπηρετηθεί ακριβώς μετά από k χρονοθυρίδες από την τελευταία του εξυπηρέτηση, χωρίς να έχει καμιά πρόσβαση στους πόρους του συστήματος τις προηγούμενες $k-1$ συναπτές χρονοθυρίδες. Η πιθανότητα του γεγονότος $x_i = k$ δίνεται από τη σχέση:

$$P[x_i = k] = p_{RT,i} (1 - p_{RT,i})^{k-1}, \quad k = 1, 2, \dots, \infty \quad (3.7)$$

Άρα:

$$P[x_i > T_i] = \sum_{m=T_i+1}^{\infty} P[x_i = m] = (1 - p_{RT,i})^{T_i}, \quad i = 1, 2, \dots, N_{RT} \text{ και } T_i \geq 0 \quad (3.8)$$

Από τις (1.6) και (3.8) προκύπτει ότι:

$$(1 - p_{RT,i})^{T_i} = g_i^I \quad (3.9)$$

$$\Rightarrow p_{RT,i} = 1 - \sqrt[T_i]{g_i^I} \quad (3.10)$$

Αν p_i η πιθανότητα εξυπηρέτησης του χρήστη i , P_{RT} η πιθανότητα εξυπηρέτησης όλων των χρηστών υπηρεσίας πραγματικού χρόνου και P_{NRT} η πιθανότητα εξυπηρέτησης όλων των χρηστών υπηρεσίας μη πραγματικού χρόνου και δεδομένου ότι για όλες τις πιθανότητες ισχύει $\sum_{i=1}^N p_i = 1$, θα είναι:

$$P_{NRT} = 1 - P_{RT} \quad (3.11)$$

Όσον αφορά την πιθανότητα εξυπηρέτησης του χρήστη υπηρεσίας j μη πραγματικού χρόνου $p_{NRT,j}$, αυτή μπορεί να προκύψει ανάλογα με τις απαιτήσεις της συγκεκριμένης υπηρεσίας. Έτσι, αν το ζητούμενο για την υπηρεσία μη πραγματικού χρόνου είναι η μακροπρόθεσμη δικαιοσύνη πρόσβασης, η πιθανότητα $p_{NRT,j}$ είναι:

$$p_{NRT,j} = \frac{P_{NRT}}{N_{NRT}} \quad (3.12)$$

Αν πάλι το ζητούμενο είναι η μεγιστοποίηση της ρυθμαπόδοσης, αγνοώντας τη δικαιοσύνη μεταξύ των χρηστών υπηρεσίας μη πραγματικού χρόνου, η αντίστοιχη πιθανότητα πρόσβασης $p_{NRT,j}$ δίνεται ακολούθως:

$$p_{NRT,j} = \frac{\bar{\rho}_j P_{NRT}}{\sum_{j=1, j \in S_{NRT}} \bar{\rho}_j} \quad (3.13)$$

Γνωρίζοντας πλέον τις αναμενόμενες πιθανότητες εξυπηρέτησης για κάθε χρήστη $p_{RT,i}$ και $p_{NRT,j}$ υπολογίζεται η πιθανότητα εξυπηρέτησης του χρήστη i με την υπόθεση κατανομής Rayleigh για το διάυλο:

$$P[i_{APAOS}^*(k) = i] = \int_0^\infty \prod_{m=1, m \neq i}^N F_{\rho_m} \left(\bar{\rho}_m \left(\frac{\rho_i}{\bar{\rho}_i} + \ln \left(\frac{B_i}{B_m} \right) \right) \right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i \quad (3.14)$$

όπου $F_{\rho_i}(\rho_i) = 1 - e^{-\frac{\rho_i}{\bar{\rho}_i}}$ η συνάρτηση κατανομής Rayleigh. Με τη βοήθεια της (3.14) υπολογίζονται και τα βάρη B_i :

$$-p_{RT,i} + \int_0^\infty \prod_{\substack{m=1 \\ m \neq i}}^N \left(1 - \frac{B_m}{B_i} e^{-\frac{\rho_i}{\bar{\rho}_i}} \right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i = 0 \quad (3.15)$$

$$-p_{NRT,j} + \int_0^\infty \prod_{\substack{m=1 \\ m \neq j}}^N \left(1 - \frac{B_m}{B_j} e^{-\frac{\rho_j}{\bar{\rho}_j}} \right) \frac{e^{-\frac{\rho_j}{\bar{\rho}_j}}}{\bar{\rho}_j} d\rho_j = 0 \quad (3.16)$$

οι οποίες αποδεικνύονται στο παράρτημα 3.1.

Παρατηρείται από τη σχέση (3.6) ότι επιλέγεται ο χρήστης i για τον οποίο ισχύει

$\rho_i < \bar{\rho}_i \left(\frac{\rho_j}{\bar{\rho}_j} + \ln \left(\frac{B_j}{B_i} \right) \right)$. Άρα θα πρέπει επίσης να ισχύει:

$$\bar{\rho}_i \left(\frac{\rho_j}{\bar{\rho}_j} + \ln \left(\frac{B_j}{B_i} \right) \right) \geq 0 \Rightarrow \begin{cases} \bar{\rho}_i = 0 & \text{ή} \\ \bar{\rho}_i > 0 \text{ και } \rho_j > \bar{\rho}_j \ln \left(\frac{B_i}{B_j} \right) & \text{ή} \\ \rho_j = \bar{\rho}_j \ln \left(\frac{B_i}{B_j} \right) & \text{ή} \\ \bar{\rho}_i < 0 \text{ και } \rho_j < \bar{\rho}_j \ln \left(\frac{B_i}{B_j} \right) & \text{(αδύνατο)} \end{cases} \quad (3.17)$$

και συνεπώς:

$$\rho_{0,i} = \bar{\rho}_i \ln \left(\frac{B_j}{B_i} \right) \quad (3.18)$$

Η σχέση (3.18) είναι απαραίτητη προϋπόθεση για να είναι σωστός ο ορισμός της πολιτικής της σχέσης (3.6). Αυτό σημαίνει και ότι τα ολοκληρώματα των σχέσεων (3.15) και (3.16) θα έπρεπε να υπολογίζονται από $\rho_{0,i}$ έως άπειρο αντί του διαστήματος από μηδέν έως άπειρο. Κάτι τέτοιο όμως θα καθιστούσε εξαιρετικά δύσκολη τη διαδικασία ανάδειξης των βαρών B , χωρίς να υφίστανται σημαντικές διαφορές στην πράξη. Σε κάθε περίπτωση, ο υπολογισμός των βαρών B που αποτελούν βέλτιστες λύσεις του συστήματος $N \times N$ (3.15) έως (3.16) συνιστά τη σημαντικότερη δυσκολία του παρόντος αλγορίθμου.

Τέλος, αξίζει να σημειωθεί ότι ο αλγόριθμος APAOS επιχειρεί να ξεπεράσει τη δυσκολία των αλγορίθμων της μορφής $i^*(k) = \arg \max_i \{w_i(k) \rho_i(k)\}$, όπου w_i κατάλληλο βάρος, οι οποίοι αδυνατούν να παρέχουν αυστηρούς περιορισμούς σε καθυστέρηση εξυπηρέτησης των χρηστών υπηρεσίας πραγματικού χρόνου. Αυτό επιχειρείται μέσω του υπολογισμού εξαρχής των βαρών B για τους χρήστες υπό την υπόθεση διαύλων κατανομής Rayleigh. Επίσης, παρατηρείται ότι δεν είναι οι απόλυτες τιμές των βαρών που παίζουν σημαντικό ρόλο όσο οι σχετικές τιμές τους, γι' αυτό και υπολογίζονται οι κατάλληλοι λόγοι βαρών για κάθε ζεύγος χρηστών στο σύστημα, οι οποίοι και οδηγούν το δρομολογητή στο επιθυμητό αποτέλεσμα.

3.4 Ο αλγόριθμος Οπορτουνιστικής Δρομολόγησης βασισμένης σε Ελαστική Βραχυπρόθεσμη Δικαιοσύνη (Elastic Short-term Fairness based opportunistic Scheduling – ESFS)

Ο αλγόριθμος ESFS ανήκει σε μια νέα γενιά αλγορίθμων χρονοδρομολόγησης, η οποία εστιάζει κάθε χρονική στιγμή στο μέρος των πόρων του συστήματος (εδώ πρόκειται για το χρόνο εξυπηρέτησης) που έχει λάβει ο χρήστης εντός σύντομων χρονικών διαστημάτων, ώστε να μπορεί να χαρακτηρίσει την ικανοποίηση ή όχι των κριτηρίων της ποιότητας υπηρεσίας του. Με άλλα λόγια, χρησιμοποιεί πληροφορία εξυπηρέτησης εντός ενός προκαθορισμένου χρονικού διαστήματος ενδιαφέροντος, το οποίο καλείται χρονικό παράθυρο παρατήρησης, προκειμένου να παρέχει διαφορετικούς περιορισμούς ελαστικής βραχυπρόθεσμης δικαιοσύνης σε χρήστες με διαφορετικές απαιτήσεις σε ποιότητα υπηρεσίας.

3.4.1 Μορφοποίηση του προβλήματος

Το πρόβλημα που καλείται να επιλύσει ο αλγόριθμος ESFS είναι αφενός η μεγιστοποίηση της μέσης απόδοσης του συστήματος και αφετέρου η εγγύηση ελάχιστου αριθμού χρονοθυρίδων εξυπηρέτησης σε βραχυπρόθεσμα όρια για κάθε χρήστη. Το πρόβλημα έχει τη μαθηματική μορφή:

$$\left. \begin{aligned} & \underset{Q \in \Theta}{\text{maximize}} \left\{ \sum_{i=1}^N E \left[r_i(k) I_{Q(k)=i} \right] \right\} \forall k \\ & \text{έτσι ώστε } P \left[S_i(k, W_i) < \phi_i W_i \right] \leq g_i^3 \end{aligned} \right\} \quad (3.19)$$

όπου Q η πολιτική δρομολόγησης, Θ το σύνολο των πολιτικών δρομολόγησης, $S_i(k, W_i)$ ο αριθμός των χρονοθυρίδων τις οποίες εξυπηρετήθηκε ο χρήστης i κατά το χρονικό διάστημα από τη χρονοθυρίδα $k - W_i + 1$ μέχρι τη χρονοθυρίδα k , W_i το χρονικό παράθυρο παρατήρησης στο οποίο επιβάλλονται οι βραχυπρόθεσμες απαιτήσεις ποιότητας υπηρεσίας και εξαρτάται από την εκάστοτε υπηρεσία (πραγματικού χρόνου), ϕ_i κατάλληλο βάρος που ανατίθεται στο χρήστη i και g_i^3 το άνω όριο της αντίστοιχης πιθανότητας αστοχίας.

Ουσιαστικά η συνθήκη του προβλήματος (3.19) εκφράζει την πιθανότητα αστοχίας ο χρήστης i να εξυπηρετηθεί τουλάχιστον τόσες φορές εντός του παραθύρου παρατήρησης όσες ορίζει η δικαιοσύνη εξυπηρέτησής του στο ίδιο διάστημα, η διατήρηση της οποίας εκφράζεται μέσω του όρου $\phi_i W_i$. Η πιθανότητα αυτή αναφέρεται στο εξής ως *πιθανότητα αστοχίας βραχυπρόθεσμης χρονικής δικαιοσύνης*.

Αφού η χρήση ελαστικής βραχυπρόθεσμης δικαιοσύνης παρέχει στατιστικά όρια στις πιθανότητες αστοχίας βραχυπρόθεσμης χρονικής δικαιοσύνης, αξιοποιείται ο τρόπος με τον οποίο τα όρια αυτά επηρεάζουν τις καθυστερήσεις των χρηστών και τους περιορισμούς σε μακροπρόθεσμη δικαιοσύνη. Αυτό φαίνεται σαφέστερα μέσω των παρακάτω προτάσεων, οι αποδείξεις των οποίων παραλείπονται, αφού τίθενται στο [17].

Πρόταση 1

Το άνω όριο της πιθανότητας αστοχίας βραχυπρόθεσμης χρονικής δικαιοσύνης ενός χρήστη είναι επίσης άνω όριο της πιθανότητας του γεγονότος η καθυστέρηση εξυπηρέτησης του χρήστη να είναι μεγαλύτερη από $W_i - \phi_i W_i$. Μαθηματικά:

$$\text{Αν } P[S_i(k, W_i) < \phi_i W_i] \leq g_i^2 \text{ τότε } P[d_i > W_i - \phi_i W_i] \leq g_i^2 \quad (3.20)$$

Πρόταση 2

Αν η πιθανότητα αστοχίας βραχυπρόθεσμης χρονικής δικαιοσύνης κάθε χρήστη είναι μηδενική, τότε η απόλυτη μακροπρόθεσμη δικαιοσύνη που επιτυγχάνεται για όλους τους χρήστες για τους οποίους έχει γίνει η υπόθεση ότι έχουν πάντα δεδομένα προς μετάδοση στις ουρές τους (continuously backlogged) για κάθε χρονικό διάστημα T (σε χρονοθυρίδες) μεγαλύτερο από το μέγιστο παράθυρο παρατήρησης $W_{max} = \max_{i \in N} \{W_i\}$. Μαθηματικά:

$$\text{Αν } P[S_i(k, W_i) < \phi_i W_i] = 0, \text{ τότε } \forall i, j \left| \frac{\hat{S}_i(k, k+T)}{\phi_i} - \frac{\hat{S}_j(k, k+T)}{\phi_j} \right| = 0 \quad (3.21)$$

όπου $\hat{S}_i(k, k+T)$ η υπηρεσία σε χρονοθυρίδες που έλαβε ο χρήστης i κατά το διάστημα από τη χρονοθυρίδα k μέχρι τη χρονοθυρίδα $k+T$ και η συνθήκη

$$\left| \frac{\hat{S}_i(k, k+T)}{\phi_i} - \frac{\hat{S}_j(k, k+T)}{\phi_j} \right| = 0 \quad (3.22)$$

εκφράζει τη μακροπρόθεσμη δικαιοσύνη μεταξύ των χρηστών, όταν εφαρμόζεται για κάθε ζεύγος χρηστών i, j και για κάθε χρονοθυρίδα k .

3.4.2 Πολιτική Δρομολόγησης

Η πολιτική δρομολόγησης ESFS αναθέτει σε κάθε χρήστη i του συστήματος ένα χρονικό παράθυρο παρατήρησης W_i έτσι, ώστε να διατηρεί πληροφορία για τον αριθμό των χρονοθυρίδων που εξυπηρετήθηκε ο χρήστης εντός του W_i . Η πιθανότητα εξυπηρέτησης του χρήστη i την επόμενη χρονική σχισμή αυξάνεται, όταν κατά το παρελθόν χρονικό διάστημα W_i ο χρήστης i έχει εξυπηρετηθεί κατά κλάσμα του W_i μικρότερο από αυτό που ορίζει η δικαιοσύνη που εφαρμόζεται και η οποία είναι ανάλογη συγκεκριμένων βαρών ϕ_i . Η μαθηματική έκφραση της πολιτικής δρομολόγησης ESFS είναι:

$$i_{ESFS}^*(k) = \arg \min_{i \in S} \hat{U}_i(k) \quad (3.23)$$

όπου η μη αρνητική συνάρτηση $\hat{U}_i(k)$ ορίζεται ως ακολούθως:

$$\hat{U}_i(k) = \max \{ U_i^*(k) - F_i(\Delta S_i(k, W_i)), 0 \} \quad (3.24)$$

όπου $\Delta S_i(k, W_i) = 1 - \frac{S_i(k, W_i)}{W_i \phi_i}$ η κανονικοποιημένη ως προς $W_i \phi_i$ απόκλιση των χρονοθυρίδων που εξυπηρετήθηκε ο χρήστης i κατά το διάστημα $k - W_i + 1$ έως k από τον αντίστοιχο αριθμό χρονοθυρίδων που πρέπει να λάβει σύμφωνα με τις απαιτήσεις της υπηρεσίας του, $F_i(x)$ προκαθορισμένη συνάρτηση σύμφωνα με τις απαιτήσεις υπηρεσίας κάθε χρήστη και του είδους δικαιοσύνης που επιβάλλουν οι απαιτήσεις της εκάστοτε κλάσης υπηρεσίας (ανάλογα με την ευαισθησία της στην καθυστέρηση) και $U_i^*(k) = U_{max}^* - \rho_i(k)$, όπου $U_{max}^* = 2\theta$.

Η παράμετρος $\Delta S_i(k, W_i)$ ουσιαστικά εκφράζει την αστοχία της δικαιοσύνης κατά το παρελθόν χρονικό διάστημα μήκους W_i και εξαρτάται από το μήκος αυτού.

Για μεγάλες τιμές W_i αναμένονται μικρές τιμές $\Delta S_i(k, W_i)$ κατά μέσο όρο και έτσι ο δρομολογητής χρειάζεται να επέμβει λιγότερες φορές για την αποκατάσταση της δικαιοσύνης εξυπηρέτησης εντός του κάθε συρόμενου παραθύρου μήκους W_i χρονοθυρίδων. Από την άλλη πλευρά, για μικρές τιμές W_i η παράμετρος $\Delta S_i(k, W_i)$ αντανακλά με μεγαλύτερη ευαισθησία την έλλειψη υπηρεσίας του χρήστη i .

Η πολιτική δρομολόγησης (3.23) στοχεύει στην εισαγωγή της πληροφορίας εξυπηρέτησης εντός του W_i μέσα στη νέα συνάρτηση κόστους $\hat{U}_i(k)$. Γι' αυτό, η $\hat{U}_i(k)$ περιέχει εκτός από τη συνάρτηση κόστους U_i^* που εξαρτάται από τις συνθήκες του διαύλου του χρήστη, και τη συνάρτηση $F(x)$ που εξαρτάται από την παράμετρο $\Delta S_i(k, W_i)$, η οποία εισάγει τη δικαιοσύνη μεταξύ των χρηστών όσον αφορά τον αριθμό των χρονοθυρίδων εξυπηρέτησης εντός του W_i . Έτσι, όταν ο χρήστης i αργεί να εξυπηρετηθεί, η παράμετρος $\Delta S_i(k, W_i)$ αυξάνεται, οπότε ο αντίστοιχος όρος $F_i(\Delta S_i(k, W_i))$ μπορεί να αναιρέσει την επίδραση του διαύλου του χρήστη στο βαθμό προτεραιότητας για την εξυπηρέτησή τους. Στην ακραία περίπτωση κατά την οποία η δικαιοσύνη εξυπηρέτησης εντός του W_i ασκεί επίδραση στη συνάρτηση κόστους, δηλαδή όταν $U_i^*(k) = F_i(\Delta S_i(k, W_i))$, ο χρήστης i εξυπηρετείται ανεξάρτητα από τις συνθήκες του διαύλου του.

Η συνάρτηση $F(x)$ είναι η ακόλουθη:

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{U_{max}^*}{e^{TH} - 1} (e^{Tx} - 1) & 0 \leq x \leq H \\ U_{max}^* & x > H \end{cases} \quad (3.25)$$

όπου $0 < H < 1$, $T > 0$ και U_{max}^* η μέγιστη τιμή της συνάρτησης κόστους $U_i^*(k)$ που αντανακλά τις συνθήκες του διαύλου του χρήστη. Οι παράμετροι H , T επιτρέπουν τον έλεγχο της ταχύτητας της επίπτωσης της παραμέτρου $\Delta S_i(k, W_i)$ στη συνάρτηση $F(x)$. Γενικά, μεγάλες τιμές αυτών οδηγούν σε περισσότερο ομορτυνιστικό δρομολογητή, ενώ μικρότερες τιμές αυτών καταλήγουν σε δρομολογητή πιο κοντά σε

αυτόν κυκλικής επαναφοράς (round robin), προκειμένου να τηρεί τις απαιτήσεις σε καθυστέρηση και δικαιοσύνη.

Συγκεκριμένα, η παράμετρος H αντανακλά την έλλειψη υπηρεσίας που μπορεί να ανεχτεί ένας χρήστης. Μικρές τιμές της H σημαίνουν μικρή ανοχή στην απόκλιση της ληφθείσας υπηρεσίας από τη δέουσα και οδηγούν σε μεγάλες τιμές της συνάρτησης $F(x)$, που σημαίνει ότι η δρομολόγηση του συγκεκριμένου χρήστη τη συγκεκριμένη χρονική σχισμή βασίζεται κυρίως στην τήρηση της δικαιοσύνης εξυπηρέτησης.

Η παράμετρος T , από την άλλη, ελέγχει την ευαισθησία του δρομολογητή στην παραβίαση της χρονικής δικαιοσύνης. Μεγάλες τιμές της παραμέτρου T σημαίνουν αναίσθητο δρομολογητή σε μικρές παραβιάσεις της δικαιοσύνης. Πάντως, αναθέτοντας διαφορετικές τιμές H , T σε διαφορετικές κλάσεις χρηστών είναι δυνατή η ύπαρξη δρομολογητή με διαφορετικό βαθμό ικανοποίησης συγκεκριμένων παραγόντων καθυστέρησης και δικαιοσύνης ανά χρήστη.

Τέλος, είναι φανερό πως ο αλγόριθμος ESFS ικανοποιεί τις απαιτήσεις ποιότητας υπηρεσίας πραγματικού χρόνου σε βραχυπρόθεσμη καθυστέρηση, ενώ ταυτόχρονα δεν τις επιβάλλει στους χρήστες υπηρεσιών με χαλαρότερες απαιτήσεις. Σαν αποτέλεσμα, αυξάνεται η μέση ρυθμαπόδοση του συστήματος και γίνεται καλύτερη αξιοποίηση των πόρων του συστήματος.

ΠΑΡΑΡΤΗΜΑ 3.1

Μαθηματική Απόδειξη της σχέσης (3.15)

Η πιθανότητα εξυπηρέτησης του χρήστη i τη χρονοθυρίδα k δίνεται από τη σχέση (3.14):

$$\begin{aligned}
 P\left[i_{APAO}^*(k) = i\right] &= \int_0^\infty \prod_{m=1, m \neq i}^N F_{\rho_m} \left(\bar{\rho}_m \left(\frac{\rho_i}{\bar{\rho}_i} + \ln \left(\frac{B_i}{B_m} \right) \right) \right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i = \\
 &= \int_0^\infty \prod_{m=1, m \neq i}^N \left(1 - e^{-\frac{\bar{\rho}_m}{\bar{\rho}_i} \left(\frac{\rho_i}{\bar{\rho}_i} + \ln \left(\frac{B_i}{B_m} \right) \right)} \right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i = \\
 &= \int_0^\infty \prod_{m=1, m \neq i}^N \left(1 - e^{-\frac{\rho_i}{\bar{\rho}_i}} \cdot e^{-\ln \left(\frac{B_i}{B_m} \right)} \right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i = \\
 &= \int_0^\infty \prod_{m=1, m \neq i}^N \left(1 - \frac{B_m}{B_i} e^{-\frac{\rho_i}{\bar{\rho}_i}} \right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i \quad \text{Ο.Ε.Δ.}
 \end{aligned}$$

Άρα, αν $p_{RT,i}$ η καθορισμένη πιθανότητα εξυπηρέτησης των χρηστών υπηρεσίας πραγματικού χρόνου με βάση τα κριτήρια της ποιότητας υπηρεσίας τους, τα βάρη B θα υπολογίζονται από τη σχέση:

$$-p_{RT,i} + \int_0^\infty \prod_{m=1, m \neq i}^N \left(1 - \frac{B_m}{B_i} e^{-\frac{\rho_i}{\bar{\rho}_i}} \right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i = 0 \quad (3.26)$$

που συμπίπτει με τη σχέση (3.15). Ομοίως γίνεται και η απόδειξη της σχέσης (3.16) για τους χρήστες υπηρεσίας μη πραγματικού χρόνου.

4

ΤΟ ΚΡΙΤΗΡΙΟ ΤΗΣ ΒΡΑΧΥΠΡΟΘΕΣΜΗΣ ΡΥΘΜΑΠΟΔΟΣΗΣ

Τα σχήματα δρομολόγησης που περιγράφηκαν στο κεφάλαιο 3 ικανοποιούν τις απαιτήσεις των χρηστών σε καθυστέρηση. Ωστόσο, οι υπηρεσίες πραγματικού χρόνου έχουν αυξημένες απαιτήσεις, τόσο σε επίπεδο καθυστέρησης εξυπηρέτησης, όσο και σε επίπεδο όγκου δεδομένων που πρέπει να έχει ληφθεί κατά το πρόσφατο παρελθόν, ώστε να θεωρείται ότι οι χρήστες των αντίστοιχων υπηρεσιών χαίρουν υψηλής ποιότητας εξυπηρέτησης. Για παράδειγμα, οι υπηρεσίες φωνής ανέχονται μέγιστη μονόδρομη καθυστέρηση *100-150 msec* και απαιτούν ρυθμό μετάδοσης *64 Kbps*.

Στο παρόν κεφάλαιο αποδεικνύεται ότι το κριτήριο ποιότητας υπηρεσίας που εκφράζεται μέσω της πιθανότητας αστοχίας καθυστέρησης της σχέσης (1.6) είναι ελλιπές ως προς την εξασφάλιση της ποιότητας υπηρεσίας πραγματικού χρόνου, οπότε είναι αναγκαία η εξασφάλιση επίσης χαμηλής πιθανότητας αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης που δίνεται από τη σχέση (1.7). Αποδεικνύεται ότι η εγγύηση χαμηλής πιθανότητας αστοχίας καθυστέρησης δεν εξασφαλίζει και χαμηλή πιθανότητα αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης, οπότε η υποβάθμιση της ποιότητας υπηρεσίας είναι αναπόφευκτη.

4.1 Η Έννοια της Βραχυπρόθεσμης Απόδοσης

Η βραχυπρόθεσμη απόδοση έχει την έννοια της απόδοσης (του όγκου δεδομένων και κατ' επέκταση του ρυθμού μετάδοσης) που επιτυγχάνει ένας χρήστης του συστήματος υπολογισμένης σε σύντομα χρονικά διαστήματα. Υπονοεί ότι ο χρήστης δεν πρέπει να μένει χωρίς ικανοποιητική εξυπηρέτηση σε όρους ληφθέντων δεδομένων για μεγάλα χρονικά διαστήματα. Η αντίστοιχη πιθανότητα αστοχίας, όσον αφορά τη διατήρηση των δεδομένων που λαμβάνονται σε επίπεδα κατώτερα από το κατώφλι B_i μονάδων δεδομένων που ορίζουν οι απαιτήσεις της εκάστοτε υπηρεσίας, δίνεται από τη σχέση (1.7). Η τήρηση του κριτηρίου βραχυπρόθεσμης απόδοσης υπαγορεύει τη διατήρηση της συγκεκριμένης πιθανότητας σε πολύ χαμηλά επίπεδα.

Όταν ο χρήστης υπηρεσίας πραγματικού χρόνου i έχει λάβει στην αρχή μιας χρονοθυρίδας λιγότερα δεδομένα από το αντίστοιχο κατώφλι του B_i στο χρονικό διάστημα W_i , η προτεραιότητά του για εξυπηρέτηση πρέπει να αυξηθεί, προκειμένου να μην παραβιαστούν οι απαιτήσεις του για βραχυπρόθεσμη απόδοση. Από την άλλη μεριά, όταν ένας χρήστης έχει λάβει περισσότερα δεδομένα από B_i στο χρονικό διάστημα W_i , τότε η προτεραιότητά του πρέπει να είναι χαμηλότερη από άλλων χρηστών.

Το προηγούμενο επιχείρημα βασίζεται στην παρατήρηση ότι όταν ο χρήστης υπηρεσίας πραγματικού χρόνου i λαμβάνει σε κάθε συρόμενο (χρονικά κυλιόμενο) χρονικό διάστημα παρατήρησης W_i ποσό δεδομένων μεγαλύτερο από το προκαθορισμένο κατώφλι του B_i , τότε και η ρυθμαπόδοσή του θα είναι μεγαλύτερη από τον απαιτούμενο σταθερό ρυθμό μετάδοσης R_i . Όσο όμως τηρούνται οι απαιτήσεις ποιότητας υπηρεσίας, όταν επιτυγχάνεται ο ρυθμός μετάδοσης R_i , επιπλέον βελτίωση στη ρυθμαπόδοση (δηλαδή μέση ρυθμαπόδοση μεγαλύτερη από R_i) δε βελτιώνει περαιτέρω το βαθμό ικανοποίησης του χρήστη i . Η επιλογή ενός άλλου χρήστη βελτιώνει την ποιότητα υπηρεσίας του, ενώ η επιλογή του χρήστη i , παρ' ότι αυξάνει το ρυθμό μετάδοσής του, αφού εξυπηρετείται περισσότερες φορές, η βελτίωση στην ποιότητα υπηρεσίας του δεν είναι αισθητή.

Επιπλέον, μειώνοντας τις προτεραιότητες εκείνων των χρηστών υπηρεσίας πραγματικού χρόνου που λαμβάνουν περισσότερους πόρους εξαιτίας του προσωρινά

καλού περιβάλλοντος μετάδοσης αποκτάται το πλεονέκτημα της επανεκχώρησης των περισσών πόρων του συστήματος σε χρήστες με προσωρινά άσχημες συνθήκες διαύλου. Διαφορετικά, οι περισσοί πόροι του συστήματος θα μπορούσαν να ανατεθούν σε χρήστες υπηρεσίας μη πραγματικού χρόνου.

4.2 Μαθηματική Απόδειξη Ανεπάρκειας του Κριτηρίου Καθυστερήσης [17]

Εκτός από τη διαίσθηση περί ανεπαρκούς έκφρασης του βαθμού ποιότητας υπηρεσίας από τη χαμηλή καθυστέρηση εξυπηρέτησης των χρηστών του συστήματος, είναι απαραίτητη και η μαθηματική υπόσταση αυτής, προκειμένου να θεωρηθεί έγκυρη. Γι' αυτό παρατίθεται ο υπολογισμός της βραχυπρόθεσμης καθυστέρησης εξυπηρέτησης και ρυθμαπόδοσης υπό κάποια γενική πολιτική δρομολόγησης.

Οι πολιτικές δρομολόγησης που εξετάζονται στη συνέχεια είναι η πολιτική μεγιστοποίησης της απόδοσης του συστήματος (MAX) και η πολιτική αναλογικής δικαιοσύνης (PF) υπό τη θεώρηση συστήματος με χρήστες υπηρεσιών μόνο πραγματικού χρόνου ($S \equiv S_{RT}$). Ως γνωστόν, οι δύο αυτοί αλγόριθμοι εκχωρούν πόρους στους χρήστες του δικτύου, επιτυγχάνοντας υψηλή ρυθμαπόδοση, ο μεν MAX τη μέγιστη δυνατή για το σύστημα με βάση τις στιγμιαίες συνθήκες διαύλου και ο δε PF τη μέγιστη δυνατή, ώστε να έχουν λάβει όλοι ένα δίκαιο μέρος από τους πόρους του συστήματος. Λαμβάνοντας υπόψιν λοιπόν το σενάριο, για το οποίο οι δύο αυτοί αλγόριθμοι καλύπτουν τις απαιτήσεις σε καθυστέρηση εξυπηρέτησης, θα έπρεπε με τους υψηλούς ρυθμούς μετάδοσης που πετυχαίνουν να πληρούν τις βραχυπρόθεσμες απαιτήσεις σε ρυθμαπόδοση. Κάτι τέτοιο όμως, όπως αποδεικνύεται στη συνέχεια, δε συμβαίνει.

Έστω Q η πολιτική δρομολόγησης που χρησιμοποιείται και $P[i_Q^*(k) = i]$ η πιθανότητα εξυπηρέτησης του χρήστη i τη χρονοθυρίδα k . Έστω ακόμη ότι οι χρήστες έχουν ανεξάρτητες συνθήκες διαύλου μεταξύ τους και ότι ο δρομολογητής κατέχει πλήρως την κατάσταση των διαύλων των χρηστών προτού προβεί στην επιλογή του επόμενου προς εξυπηρέτηση χρήστη στην αρχή κάθε χρονοθυρίδας k ,

οπότε η πιθανότητα εξυπηρέτησης του χρήστη i κάθε χρονοθυρίδα ισούται με την πιθανότητα εξυπηρέτησής του κάθε χρονική στιγμή.

Αν η πιθανότητα εξυπηρέτησης του χρήστη i κάθε χρονική στιγμή είναι σταθερή υπό την πολιτική δρομολόγησης που χρησιμοποιείται, κάτι που ισχύει για τις πολιτικές MAX και PF, όπως αποδεικνύεται στην ενότητα 4.3[17] και αναφέρεται στο [5], η αντίστοιχη καθυστέρηση εξυπηρέτησης ακολουθεί γεωμετρική κατανομή, οπότε για την πιθανότητα αστοχίας καθυστέρησης εξυπηρέτησης του χρήστη i υπηρεσίας πραγματικού χρόνου ισχύει:

$$P[d_i > T_i] = \left(1 - P[i_Q^*(k) = i]\right)^{T_i} \quad (4.1)$$

Επιπρόσθετα, υπό την υπόθεση ότι ο ρυθμός μετάδοσης του χρήστη i , όταν εξυπηρετείται σε μια χρονοθυρίδα εντός του χρονικού παραθύρου παρατήρησης μήκους W_i χρονοθυρίδων, ισούται με το μέσο ρυθμό μετάδοσής του $\bar{R}_{Q,i}$ υπό την πολιτική δρομολόγησης Q , η πιθανότητα αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης δίνεται από τη σχέση:

$$P[b_i < B'_i(m, \bar{R}_{Q,i})]_{W_i} = \sum_{n=0}^{m-1} \binom{W_i}{n} \left(P[i_Q^*(k) = i]\right)^n \left(1 - P[i_Q^*(k) = i]\right)^{W_i-n} \quad (4.2)$$

όπου b_i η τυχαία μεταβλητή του γεγονότος ο χρήστης i να λαμβάνει b_i μονάδες δεδομένων μέχρι την κατά σειρά $m = 0, 1, \dots, W_i$ χρονοθυρίδα εντός W_i διαδοχικών χρονοθυρίδων, $B'_i(m, \bar{R}_{Q,i}, T_{slot}) = m \cdot \bar{R}_{Q,i} \cdot T_{slot}$ το μέγιστο ποσό δεδομένων που μπορεί να έχει λάβει ο χρήστης i εξυπηρετούμενος $m = 0, 1, \dots, W_i$ χρονοσχισμές εντός του χρονικού διαστήματος μήκους W_i χρονικών σχισμών (συμπεριλαμβανομένης της τρέχουσας χρονοθυρίδας). Εδώ πρέπει να σημειωθεί ότι λόγω της προηγούμενης υπόθεσης είναι πιθανή η υπερεκτίμηση των επιδόσεων των εκάστοτε αλγορίθμων όσον αφορά την εγγύηση επίτευξης πιθανοτικής βραχυπρόθεσμης απόδοσης.

Αξίζει να σημειωθεί ότι η πιθανότητα (4.2) δε διαφέρει από την (1.7), παρόλο που φαίνεται το κατώφλι $B'_i(m, \bar{R}_{Q,i}, T_{slot})$ να υπολογίζει τις μονάδες δεδομένων που θα λάμβανε ο χρήστης i , αν εξυπηρετούνταν m όχι απαραίτητα διαδοχικές χρονοθυρίδες, αλλά εντός του χρονικού κυλιόμενου διαστήματος W_i (αφού θεωρείται ότι εξυπηρετείται με το μέσο ρυθμό μετάδοσης $\bar{R}_{Q,i}$), ενώ το κατώφλι B_i της σχέσης

(1.7) δείχνει το ελάχιστο ποσό δεδομένων που πρέπει να έχει λάβει ο χρήστης i καθ' όλο το διάστημα διάρκειας W_i χρονοθυρίδων. Αλλά, η μορφή της (4.2) βοηθάει στη ροή της μαθηματικής απόδειξης, καθώς η πληροφορία που μπορούμε να δεχτούμε ως γνωστή αφορά τον αριθμό των χρονοθυρίδων που εξυπηρετήθηκε ο χρήστης i .

Από τις σχέσεις πάντως (4.1) και (4.2) προκύπτει ότι η πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης και η πιθανότητα αστοχίας καθυστέρησης εξυπηρέτησης υπό την υπόθεση ρυθμού μετάδοσης σε κάθε χρονοθυρίδα ίσο με το μέσο ρυθμό του χρήστη, εξαρτώνται από τον όρο $(1 - P[i_Q^*(k) = i])$. Παρ' όλ' αυτά, η διατήρηση χαμηλών τιμών για τον όρο αυτό συνεπάγεται χαμηλή πιθανότητα αστοχίας καθυστέρησης, αλλά δε συνεπάγεται απαραίτητα χαμηλή πιθανότητα αστοχίας βραχυπρόθεσμης ρυθμαπόδοσης, αφού η τελευταία σύμφωνα με τη σχέση (4.2) εξαρτάται επίσης και από τον όρο $P[i_Q^*(k) = i]$, ο οποίος αυξάνεται με τη μείωση της διαφοράς $(1 - P[i_Q^*(k) = i])$. Άρα, λοιπόν, αυτή είναι μια πρώτη αίσθηση ότι η εγγύηση του κριτηρίου (1.6) δεν αρκεί για τη διατήρηση υψηλής ποιότητας υπηρεσίας πραγματικού χρόνου.

4.3 Οι Πιθανότητες Αστοχίας για τις Πολιτικές MAX και PF [17]

Στη συνέχεια μελετάται η συμπεριφορά των αλγορίθμων MAX και PF ως προς τα κριτήρια (1.6) και (1.7). Πρόκειται για αλγορίθμους που εστιάζουν στο ρυθμό μετάδοσης, ο πρώτος στη μεγιστοποίησή του και ο δεύτερος στη δίκαιη διανομή του σε μακροπρόθεσμα διαστήματα με στόχο τη μεγιστοποίηση της ρυθμαπόδοσης κατά το δυνατόν. Θεωρείται η περίπτωση που οι δύο αλγόριθμοι δίνουν χαμηλές πιθανότητες αστοχίας καθυστέρησης εξυπηρέτησης.

4.3.1 Υπολογισμός πιθανοτήτων αστοχίας του αλγορίθμου MAX

Όπως αναφέρεται στην ενότητα 2.1, η πολιτική δρομολόγησης MAX εκφράζεται ως εξής:

$$i_{MAX}^*(k) = \arg \max_{i \in S} \{\rho_i(k)\} \quad (4.3)$$

όπου i_{MAX}^* η πολιτική δρομολόγησης και ο σηματοθορυβικός λόγος που επιτυγχάνει ο χρήστης i τη χρονοθυρίδα k , οπότε η πιθανότητα εξυπηρέτησης του χρήστη i τη χρονοθυρίδα k εκφράζεται, σε συμφωνία με τη συνάρτηση $F_{\rho_i}(\rho_i)$ της ενότητας 1.4, ως εξής:

$$P[i_{MAX}^*(k) = i] = \int_0^\infty \prod_{j=1, j \neq i}^N \left(1 - e^{-\frac{\rho_i}{\bar{\rho}_j}}\right) e^{-\frac{\rho_i}{\bar{\rho}_i}} d\rho_i \quad (4.4)$$

Ο ασυμπτωτικός ρυθμός μετάδοσης του χρήστη i δίνεται από τη σχέση:

$$\bar{R}_{MAX,i} = \frac{W}{\gamma_i} \int_0^\infty \rho_i \prod_{j=1, j \neq i}^N \left(1 - e^{-\frac{\rho_i}{\bar{\rho}_j}}\right) e^{-\frac{\rho_i}{\bar{\rho}_i}} d\rho_i \quad (4.5)$$

σε αντιστοιχία με τη σχέση (1.2).

Έστω τώρα ότι χρησιμοποιούνται δύο κλάσεις χρηστών, σύμφωνα με τις συνθήκες του διαύλου τους. Έστω N_g (S_g) ο αριθμός (σύνολο) των χρηστών με ευνοϊκές συνθήκες διαύλου («καλοί» χρήστες) και μέσο SINR $\bar{\rho}_g$ και $N_b = N - N_g$ (S_b) ο αριθμός (σύνολο) των χρηστών με δυσμενείς συνθήκες διαύλου («κακοί» χρήστες) και μέσο SINR ίσο με $\bar{\rho}_b$ τέτοιο, ώστε $\bar{\rho}_g = a \cdot \bar{\rho}_b$, όπου $a \geq 1$. Η πιθανότητα εξυπηρέτησης ενός «καλού» χρήστη $i \in S_g$ προκύπτει από την (4.4) με χρήση των θεωρημάτων «binomial theorem» και «distributed law of products» (βλ. Παράρτημα 4.1):

$$P[i_{MAX}^*(k) = i]_{i \in S_g} = \sum_{n=0}^{N_g-1} \sum_{m=0}^{N-N_g} \binom{N_g-1}{n} \binom{N-N_g}{m} \frac{(-1)^{n+m}}{n+am+1} \quad (4.6)$$

επομένως και από τη σχέση (4.5) προκύπτει:

$$\bar{R}_{MAX,i \in S_g} = \sum_{n=0}^{N_g-1} \sum_{m=0}^{N-N_g} \binom{N_g-1}{n} \binom{N-N_g}{m} \frac{(-1)^{n+m}}{(n+am+1)^2} \frac{W \cdot \bar{\rho}_g}{\gamma_i} \quad (4.7)$$

Από τη σχέση (4.6) προκύπτει ότι για δεδομένο αριθμό ενεργών χρηστών στο σύστημα και δεδομένο λόγο a «καλών» προς «κακούς» χρήστες η πιθανότητα εξυπηρέτησης ενός χρήστη την επόμενη χρονοθυρίδα είναι σταθερή υπό την πολιτική MAX. Σε συμφωνία με τις σχέσεις (4.1) και (4.2) οι πιθανότητες αστοχίας

καθυστέρησης και βραχυπρόθεσμης δικαιοσύνης ρυθμαπόδοσης δίνονται από τις σχέσεις:

$$P[d_i > T_i]_{i \in S_g} = \left(1 - P[i_{MAX}^*(k) = i]_{i \in S_g}\right)^{T_i} \quad (4.8)$$

$$P[b_i < B'_i(m, \bar{R}_{Q,i})]_{W_i} = \sum_{n=0}^{m-1} \binom{W_i}{n} \left(P[i_{MAX}^*(k) = i]_{i \in S_g}\right)^n \left(1 - P[i_{MAX}^*(k) = i]_{i \in S_g}\right)^{W_i-n} \quad (4.9)$$

με το $P[i_{MAX}^*(k) = i]_{i \in S_g}$ να δίνεται από τη σχέση (4.6).

Αντίστοιχα συμπεράσματα προκύπτουν και για τους «κακούς» χρήστες ακολουθώντας όμοια ανάλυση.

4.3.2 Υπολογισμός πιθανοτήτων αστοχίας του αλγορίθμου PF

Η πολιτική δρομολόγησης PF, όπως αναλύεται στην ενότητα 2.2, εκφράζεται από την παρακάτω σχέση, η οποία επαναλαμβάνεται εδώ σε μια ισοδύναμη μορφή για λόγους ευκολίας:

$$i_{PF}^*(k) = \arg \max_{i \in S} \left\{ \frac{\rho_i(k)}{\rho_i^*(k)} \right\} \quad (4.10)$$

όπου $\rho_i^*(k+1) = (1-\beta)\rho_i^*(k) + \lambda\rho'_i(k)$ και $\rho'_i(k) = 0$ για τις περιπτώσεις που ο χρήστης i δεν εξυπηρετείται, $\beta > 0$ και $\lambda > 0$ μικρή θετική σταθερά. Αποδεικνύεται στο [19] ότι υπό την υπόθεση Rayleigh εξασθένισης (Rayleigh fading) διαύλου, όπως συμβαίνει εδώ, η μέση ρυθμαπόδοση $\bar{R}_{PF,i}$ του χρήστη i μετά από παρέλευση αρκετού χρόνου ισούται με μια τιμή ανάλογη προς το μέσο ρυθμό μετάδοσης που θα μπορούσε να επιτύχει με τις συνθήκες του διαύλου του, αν ήταν μόνος του στο σύστημα. Συγκεκριμένα, το κέρδος δρομολόγησης κάθε χρήστη $G(N)$ είναι καθορισμένο για συγκεκριμένο αριθμό ενεργών χρηστών, άρα:

$$\bar{R}_{PF,i} = \frac{G(N)W}{N} \frac{\bar{\rho}_i}{\gamma_i} \quad (4.11)$$

όπου $G(N) = \sum_{j=1}^N \frac{I}{j}$. Εξαιτίας αυτής της παρατήρησης, όταν το σύστημα βρίσκεται σε σταθερή κατάσταση, δηλαδή μετά τη σύγκλιση του $\rho_i^*(k)$, η πολιτική PF μετασχηματίζεται στην ακόλουθη:

$$i_{PF}^*(k) = \arg \max_{i \in S} \left\{ \frac{\rho_i(k)}{\bar{\rho}_i \frac{G(N)}{N}} \right\} \quad (4.12)$$

Σύμφωνα με την πολιτική (4.12) η πιθανότητα εξυπηρέτησης του χρήστη i σε μια χρονοθυρίδα μπορεί να υπολογιστεί ως:

$$P[i_{PF}^*(k) = i] = \int_0^\infty \prod_{j=1, j \neq i}^N F_{\rho_j} \left(\frac{\bar{\rho}_j}{\bar{\rho}_i} \rho_i \right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i \quad (4.13)$$

το οποίο εύκολα καταλήγει στο:

$$P[i_{PF}^*(k) = i] = \sum_{n=1}^N \binom{N}{n} \frac{(-1)^{n+1}}{N} = \frac{1}{N} \quad (4.14)$$

Άρα παρατηρείται ότι η πιθανότητα επιλογής ενός χρήστη για εξυπηρέτηση την επόμενη χρονοσχισμή εξαρτάται μόνο από τον αριθμό των ενεργών χρηστών στο σύστημα, γεγονός που φανερώνει τη δικαιοσύνη σε μακροπρόθεσμο επίπεδο που εγγυάται η πολιτική PF.

Από τις σχέσεις (4.1) και (4.2) οι πιθανότητες αστοχίας καθυστέρησης και βραχυπρόθεσμης δικαιοσύνης ρυθμαπόδοσης δίνονται από τις σχέσεις:

$$P[d_i > T_i]_{i \in S} = \left(1 - P[i_{PF}^*(k) = i]_{i \in S} \right)^{T_i} \quad (4.15)$$

$$P[b_i < B'_i(m, \bar{R}_{Q,i})]_{W_i} = \sum_{n=0}^{m-1} \binom{W_i}{n} \left(P[i_{PF}^*(k) = i]_{i \in S} \right)^n \left(1 - P[i_{PF}^*(k) = i]_{i \in S} \right)^{W_i - n} \quad (4.16)$$

με το $P[i_{PF}^*(k) = i]_{i \in S}$ να δίνεται από τη σχέση (4.14).

4.3.3 Γραφική προσέγγιση

Στη συγκεκριμένη ενότητα παρουσιάζεται διαγραμματικά η ανεπάρκεια του κριτηρίου (1.6) για την εγγύηση του κριτηρίου (1.7), δηλαδή της ελάχιστης βραχυπρόθεσμης ρυθμαπόδοσης για τους αλγορίθμους MAX και PF, για τους οποίους παρατέθηκαν οι αντίστοιχες πιθανότητες αστοχίας στις ενότητες 4.3.1 και 4.3.2.

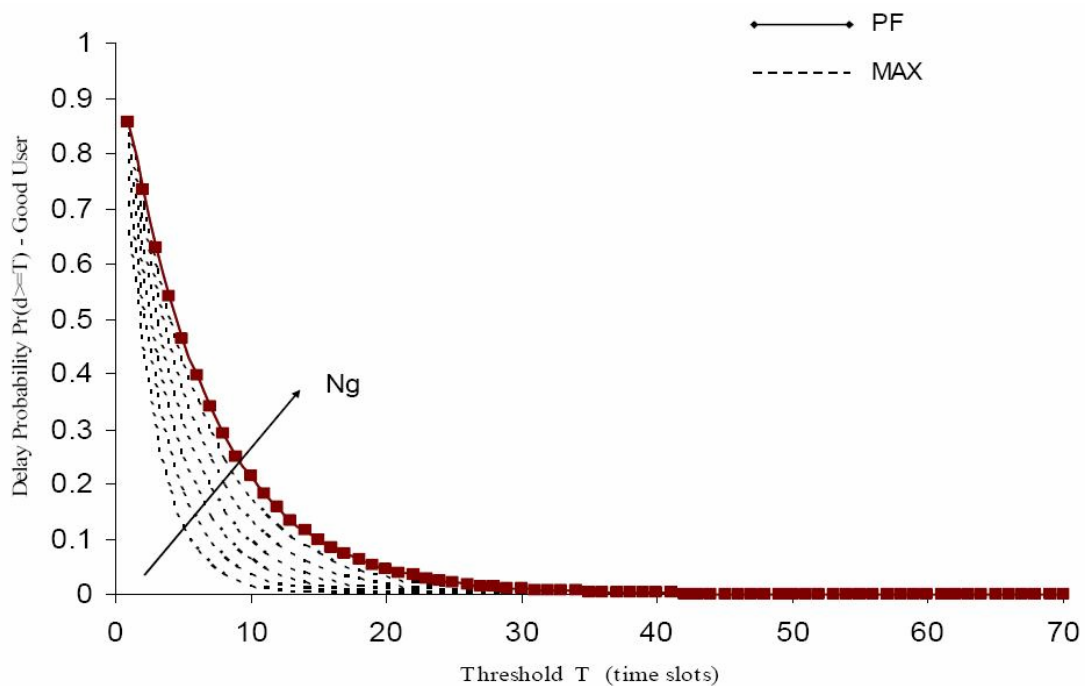
Πρόκειται για τα σχήματα 4.1 έως 4.4, στα οποία παρουσιάζονται οι ιδιότητες καθυστέρησης εξυπηρέτησης και βραχυπρόθεσμης ρυθμαπόδοσης. Η διακεκομμένη γραμμή αντιστοιχεί στην πολιτική δρομολόγησης MAX και η συνεχής στην PF. Θεωρείται ότι στο σύστημα υπάρχουν $N = N_{RT} = 7$ ενεργοί χρήστες, οι οποίοι διαχωρίζονται σε δύο κλάσεις, τους «καλούς» και τους «κακούς» χρήστες ανάλογα με τις συνθήκες του διαύλου τους. Παρέχονται αποτελέσματα για μεταβαλλόμενες τιμές του αριθμού N_g των «καλών» χρηστών από 1 έως 7. Η κατεύθυνση του βέλους αντιστοιχεί στη φορά αύξησης της παραμέτρου N_g . Επίσης θεωρήθηκε ότι η μέση τιμή του SINR του διαύλου των «καλών» χρηστών είναι $\bar{\rho}_g = 6dB$ και των «κακών» χρηστών $\bar{\rho}_b = 3dB$.

Τα σχήματα 4.1 και 4.2 δείχνουν την πιθανότητα αστοχίας καθυστέρησης των «καλών» και «κακών» χρηστών του συστήματος αντίστοιχα συναρτήσει του κατωφλίου καθυστέρησης T_i χρονοθυρίδων για διάφορες τιμές του αριθμού N_g των «καλών» χρηστών. Από αυτά σημειώνεται ότι στην πολιτική PF όλοι οι χρήστες έχουν την ίδια πιθανότητα πρόσβασης στους πόρους του συστήματος κάθε χρονική θυρίδα ανεξάρτητα από την κατάσταση του διαύλου τους (την κλάση τους) ή τη σύσταση του συστήματος, οπότε επιτυγχάνεται η δικαιοσύνη στα πλαίσια της ταχύτητας εξυπηρέτησης. Επίσης, οι πιθανότητες αστοχίας καθυστέρησης είναι ιδιαίτερα χαμηλές για κατώφλια μεγαλύτερα από 30 χρονοθυρίδες, δηλαδή μεγαλύτερα από 50.1 msec. Άρα σύμφωνα με τη μέγιστη ανεκτή μονόδρομη καθυστέρηση στα 100-150 msec οι απαιτήσεις καθυστέρησης των υπηρεσιών φωνής τηρούνται.

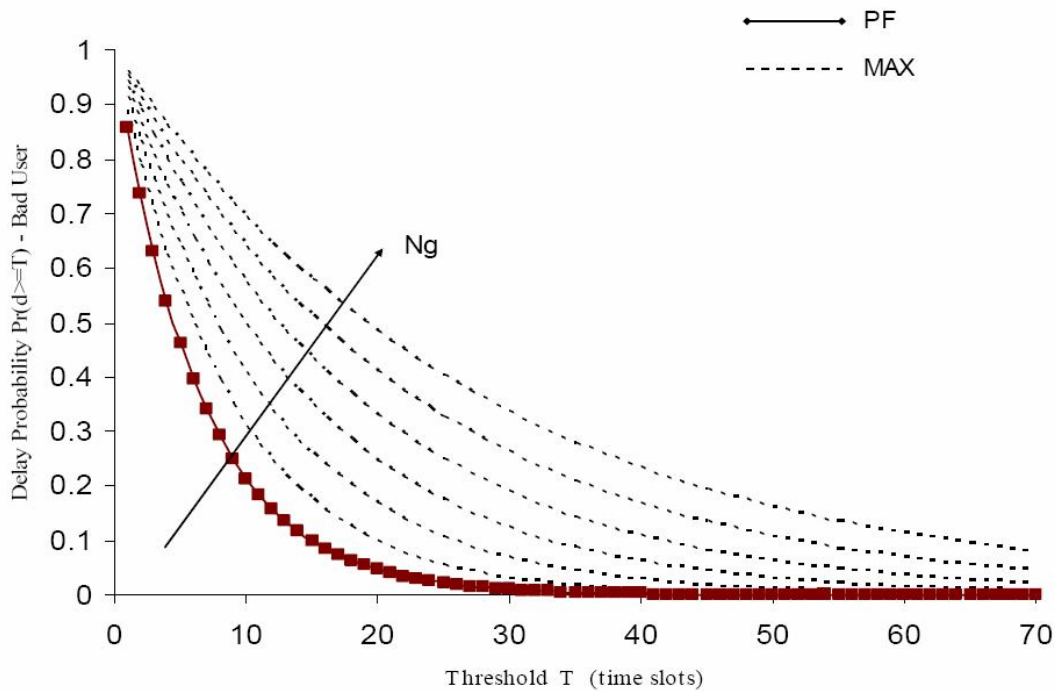
Από την άλλη, από το σχήμα 4.1 ο MAX αλγόριθμος ευνοεί τους «καλούς» χρήστες εξυπηρετώντας τους με μικρότερες πιθανότητες αστοχίας καθυστέρησης, ενώ μεγάλες πιθανότητες αστοχίας καθυστέρησης παρατηρούνται για τους «κακούς»

χρήστες στο σχήμα 4.2. Οι πιθανότητες αστοχίας καθυστέρησης διατηρούνται χαμηλές για κατώφλια μεγαλύτερα από 30 χρονοθυρίδες για τους «καλούς» χρήστες για οποιαδήποτε από τις τιμές N_g που θεωρούνται, ενώ σε περιπτώσεις μικρού πλήθους «καλών» χρηστών οι αντίστοιχες πιθανότητες είναι χαμηλές και για το κατώφλι μόλις 20 χρονοθυρίδων.

Ο εξοστρακισμός των «κακών» χρηστών είναι εντονότερος όσο αυξάνεται ο αριθμός N_g , διότι τότε το σύστημα, το οποίο προτιμά τους «καλούς» χρήστες, έχει περισσότερους τέτοιους για να επιλέξει. Αντίθετα, η πιθανότητα αστοχίας καθυστέρησης των «κακών» χρηστών μειώνεται όσο λιγότεροι είναι οι «καλοί» χρήστες, διότι εξυπηρετούνται χωρίς να υπάρχει μεγάλος ανταγωνισμός μεταξύ τους και περισσεύουν πόροι και για τους «κακούς» χρήστες, που ενίοτε έχουν σχετικά ευνοϊκές στιγμιαίες συνθήκες διαύλου. Οι πιθανότητες αστοχίας καθυστέρησης για τους «κακούς» χρήστες είναι μικρές μόνο για τιμές κατωφλίου μεγαλύτερες των 70 χρονοθυρίδων.



Σχήμα 4. 1: Πιθανότητα αστοχίας καθυστέρησης για τους «καλούς» χρήστες συναρτήσει του κατωφλίου καθυστέρησης.



Σχήμα 4. 2: Πιθανότητα αστοχίας καθυστέρησης για τους «κακούς» χρήστες συναρτήσει του κατωφλίου καθυστέρησης.

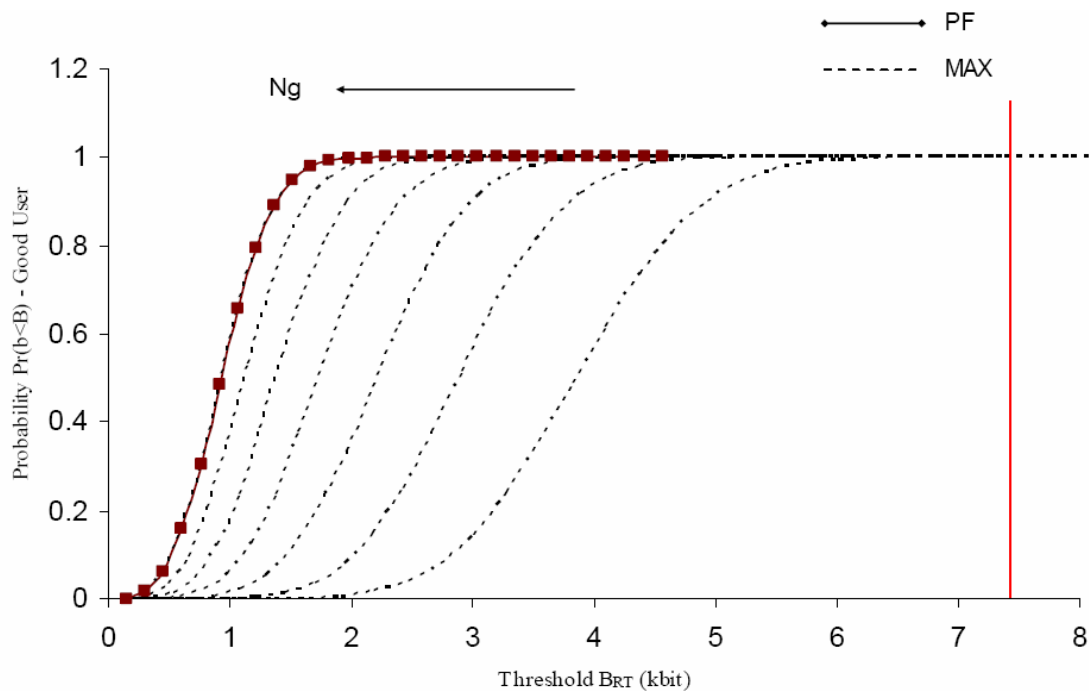
Παρόμοια γενικά συμπεράσματα προκύπτουν και για την περίπτωση των απαιτήσεων σε ρυθμαπόδοση, όπως φανερώνουν τα σχήματα 4.3 και 4.4, στα οποία παρουσιάζεται η πιθανότητα αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης συναρτήσει του κατωφλίου ληφθέντων δεδομένων B_i (σε *Kbit*) για τους «καλούς» και τους «κακούς» χρήστες αντίστοιχα, όταν θεωρηθεί χρονικό παράθυρο παρατήρησης W_i 70 χρονοθυρίδων. Ο PF παρουσιάζει και πάλι δίκαιη συμπεριφορά μακροπρόθεσμα ανεξαρτήτως της κατάστασης του διαύλου ή του ποσοστού των «καλών» χρηστών στο σύστημα. Οι πιθανότητες αστοχίας, όμως, είναι σχεδόν ίσες με τη μονάδα από το κατώφλι του 1 *Kbit*.

Ο MAX, αντίθετα, δίνει μεγάλες πιθανότητες αστοχίας στους «κακούς» χρήστες για πολύ μικρότερα κατώφλια δεδομένων σε σχέση με τους «καλούς». Όσο λιγότεροι είναι οι «καλοί» χρήστες στο σύστημα, τόσο πιο αργά, δηλαδή για μεγαλύτερα κατώφλια δεδομένων, και με μικρότερη κλίση, επομένως, φτάνει τη μονάδα η πιθανότητα αστοχίας των «καλών» χρηστών. Σαν αποτέλεσμα, οι πιθανότητες αστοχίας των «καλών» χρηστών είναι καλύτερες από αυτές που δίνει ο PF. Έτσι, για την περίπτωση ενός μόνο «καλού» χρήστη η πιθανότητα αστοχίας βραχυπρόθεσμης

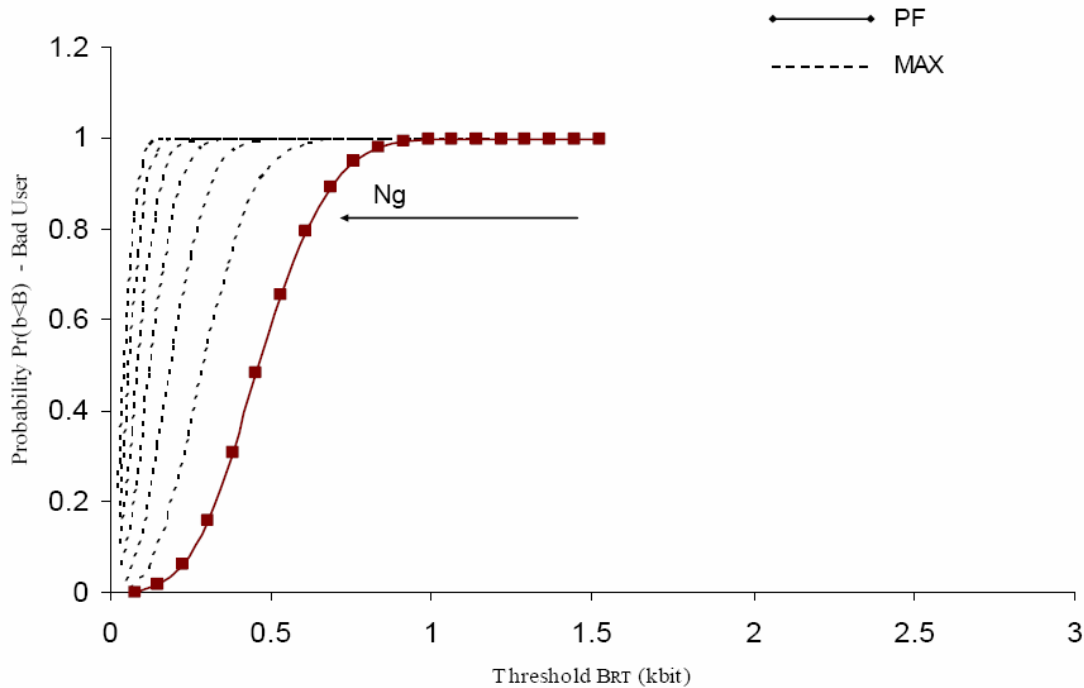
ρυθμαπόδοσης είναι χαμηλή μέχρι και το κατώφλι των 3 Kbits, ενώ όσο αυξάνεται η μεταβλητή N_g , οι πιθανότητες είναι υψηλές από το 1 Kbit κιόλας.

Όσον αφορά τους «κακούς» χρήστες, συμβαίνει το αντίθετο. Όσο αυξάνεται ο αριθμός των «καλών» χρηστών στο σύστημα, τόσο γρηγορότερα αυξάνεται η πιθανότητα καθυστέρησης, αλλά αυτή τη φορά τα αποτελέσματα είναι σε κάθε περίπτωση χειρότερα από αυτά του αλγορίθμου PF. Έτσι, για κατώφλια B_i πολύ μικρότερα του 1 Kbit η πιθανότητα αστοχίας βραχυπρόθεσμης ρυθμαπόδοσης γίνεται 1.

Η ουσία, πάντως, είναι ότι κανένας από τους δύο αλγορίθμους δεν τηρεί τις βραχυπρόθεσμες απαιτήσεις απόδοσης, καθώς η αντίστοιχη πιθανότητα αστοχίας είναι σχεδόν ίση με τη μονάδα για κατώφλια δεδομένων πολύ μικρά.



Σχήμα 4. 3: Πιθανότητα αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης για τους «καλούς» χρήστες συναρτήσει του κατωφλίου ληφθέντων δεδομένων.



Σχήμα 4. 4: Πιθανότητα αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης για τους «κακούς» χρήστες συναρτήσει του κατώφλιου ληφθέντων δεδομένων.

Στη συνέχεια, η ανάλυση επικεντρώνεται στην περίπτωση των χρηστών υπηρεσίας πραγματικού χρόνου, παραδείγματος χάριν σε χρήστες υπηρεσιών φωνής. Όταν τα κατώφλια καθυστέρησης είναι μεγαλύτερα από 70 χρονοθυρίδες, που αντιστοιχούν σε κατώφλια των $70 \cdot 1.67m\ sec = 116.9m\ sec \approx 120m\ sec$ παρατηρείται ότι οι πιθανότητες αστοχίας της καθυστέρησης εξυπηρέτησης λαμβάνουν εξαιρετικά μικρές τιμές για τους χρήστες με ευνοϊκές συνθήκες διαύλου, οπότε ικανοποιούνται οι περιορισμοί καθυστέρησης.

Από την άλλη μεριά, στην αρχή κεφαλαίου αναφέρθηκε ότι οι υπηρεσίες φωνής απαιτούν σταθερό ρυθμό μετάδοσης $64\ Kbps$, οπότε έρχεται στο προσκήνιο ο περιορισμός βραχυπρόθεσμης δικαιοσύνης ρυθμαπόδοσης. Για ρυθμό μετάδοσης $64\ Kbps$ και χρονικό παράθυρο παρατήρησης 70 χρονοθυρίδων, το κατώφλι μονάδων δεδομένων $B = 64 \cdot 70 \cdot 1.67\ bits = 7481.6\ bits$ υπολογίζεται μέσω της σχέσης (1.8) και αντιστοιχεί στην κόκκινη κάθετη γραμμή του σχήματος 4.3. Παρατηρείται, λοιπόν, ότι ακόμα και για τους «καλούς» χρήστες, για τους οποίους πληρούνται οι απαιτήσεις καθυστέρησης, δεν ικανοποιούνται οι απαιτήσεις σε βραχυπρόθεσμο επιτυγχανόμενο ρυθμό μετάδοσης. Συγκεκριμένα, σύμφωνα με το σχήμα 4.3 για το κατώφλι των $B = 7481.6\ bits$ η πιθανότητα αστοχίας δικαιοσύνης βραχυπρόθεσμης απόδοσης είναι σχεδόν ίση με τη μονάδα. Άρα ούτε καν ο αλγόριθμος MAX, ο

οποίος επιτυγχάνει μέγιστη ρυθμαπόδοση του συστήματος δεν ικανοποιεί τις βραχυπρόθεσμες απαιτήσεις απόδοσης.

ΠΑΡΑΡΤΗΜΑ 4.1

Μαθηματική Απόδειξη της σχέσης (4.6)

Εδώ παρατίθεται η απόδειξη της σχέσης (4.6) που δίνει την πιθανότητα εξυπηρέτησης την επόμενη χρονοσχισμή του χρήστη i , ο οποίος χαρακτηρίζεται από ευνοϊκές συνθήκες διαύλου. Όπως αναφέρεται και στην ενότητα 4.3.1, η σχέση (4.6) αποδεικνύεται με τη βοήθεια του ακόλουθου νόμου:

$$(a+b)^N = \sum_{i=0}^N \binom{N}{i} b^i a^{N-i}, \quad N \in \mathbb{N}^* \text{ και } a, b \in \mathbb{R}^* \quad (4.17)$$

Θεωρώντας, λοιπόν ότι ο χρήστης $i \in S_g$, σύμφωνα με τη σχέση (4.4) έχουμε:

$$\begin{aligned} P\left[i_{MAX}^* = i\right]_{i \in S_g} &= \int_0^\infty \prod_{j=1, j \neq i}^N \left(1 - e^{-\frac{\rho_i}{\bar{\rho}_j}}\right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i = \\ &= \int_0^\infty \prod_{n=1, n \neq i}^{N_g} \left(1 - e^{-\frac{\rho_i}{\bar{\rho}_n}}\right) \prod_{m=1}^{N-N_g} \left(1 - e^{-\frac{\rho_i}{\bar{\rho}_m}}\right) \frac{e^{-\frac{\rho_i}{\bar{\rho}_i}}}{\bar{\rho}_i} d\rho_i = \quad (4.21) \\ &= \int_0^\infty \sum_{n=0}^{N_g-1} \binom{N_g-1}{n} (-1)^n e^{-n\frac{\rho_i}{\bar{\rho}_g}} \cdot \sum_{m=0}^{N-N_g} \binom{N-N_g}{m} (-1)^m e^{-m\frac{\rho_i}{\bar{\rho}_b}} \cdot \frac{e^{-\frac{\rho_i}{\bar{\rho}_g}}}{\bar{\rho}_g} d\rho_i = \\ &= \sum_{n=0}^{N_g-1} \sum_{m=0}^{N-N_g} \binom{N_g-1}{n} \binom{N-N_g}{m} (-1)^{n+m} \int_0^\infty \frac{e^{-\left(\frac{n}{\bar{\rho}_g} + \frac{m}{\bar{\rho}_b} + \frac{1}{\bar{\rho}_g}\right)\rho_i}}{\bar{\rho}_g} d\rho_i = \\ &= \sum_{n=0}^{N_g-1} \sum_{m=0}^{N-N_g} \binom{N_g-1}{n} \binom{N-N_g}{m} (-1)^{n+m} \left[-\frac{1}{\left(\frac{n}{\bar{\rho}_g} + \frac{m}{\bar{\rho}_b} + \frac{1}{\bar{\rho}_g}\right) \bar{\rho}_g} e^{-\left(\frac{n}{\bar{\rho}_g} + \frac{m}{\bar{\rho}_b} + \frac{1}{\bar{\rho}_g}\right)\rho_i} \right]_0^\infty = \\ &= \sum_{n=0}^{N_g-1} \sum_{m=0}^{N-N_g} \binom{N_g-1}{n} \binom{N-N_g}{m} (-1)^{n+m} \left(-\frac{1}{n+1+m\frac{\bar{\rho}_g}{\bar{\rho}_b}} (0-1) \right) \frac{\bar{\rho}_g}{\bar{\rho}_b} = \\ &= \sum_{n=0}^{N_g-1} \sum_{m=0}^{N-N_g} \binom{N_g-1}{n} \binom{N-N_g}{m} \frac{(-1)^{n+m}}{n+am+1}. \quad (\text{O.E.}\Delta) \end{aligned}$$

Η απόδειξη για την περίπτωση που ο χρήστης i χαρακτηρίζεται από δυσχερείς συνθήκες διαύλου ($i \in S_b$) προκύπτει ακολουθώντας παρόμοια διαδικασία. Επίσης αντίστοιχη είναι η απόδειξη της σχέσης (4.14) για την πιθανότητα εξυπηρέτησης ενός χρήστη την επόμενη χρονοθυρίδα στον αλγόριθμο PF.

5

Ο ΑΛΓΟΡΙΘΜΟΣ USTF

Στο παρόν κεφάλαιο περιγράφεται μια νέα πολιτική δρομολόγησης που εισάγεται στο [17] και έρχεται να δώσει λύση στο πρόβλημα που αναλύεται στο κεφάλαιο 3 περί ανεπαρκούς εξυπηρέτησης των χρηστών υπηρεσίας πραγματικού χρόνου από τους ευαίσθητους σε καθυστέρηση αλγορίθμους. Κύριο μέλημα του νέου αλγορίθμου είναι η τήρηση των κριτηρίων ποιότητας υπηρεσίας πραγματικού χρόνου στα πλαίσια της απαιτούμενης βραχυπρόθεσμης ρυθμαπόδοσης, οπότε και να τηρούνται οι απαιτήσεις μικρής καθυστέρησης εξυπηρέτησης και να διατηρούν οι χρήστες τους απαιτούμενους ρυθμούς μετάδοσης δεδομένων. Η πολιτική αυτή ονομάζεται Πλαίσιο Εργασίας για την επίτευξη Βραχυπρόθεσμων Περιορισμών Δεδομένων βασισμένο σε Συναρτήσεις Χρησιμοποίησης (Utility based Short-term Throughput driven Framework - USTF).

5.1 Το Μαθηματικό Πρόβλημα

Το μαθηματικό πρόβλημα που καλείται να επιλύσει ο αλγόριθμος USTF, είναι η μεγιστοποίηση της απόδοσης του συστήματος με την εγγύηση χαμηλής πιθανότητας

αστοχίας δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης. Η έκφρασή του δίνεται μαθηματικά από τη σχέση:

$$\left. \begin{aligned} & \underset{Q \in \Theta}{\text{maximize}} \left\{ \sum_{i=1}^N E \left[U_i^*(k) I_{Q(k)=i} \right] \right\} \forall k \\ & \text{έτσι ώστε } P \left[b_i(k, W_i) < B_i \right]_{W_i} \leq g_i^2 \end{aligned} \right\} \quad (5.1)$$

όπου Q η πολιτική δρομολόγησης, Θ το σύνολο των πολιτικών δρομολόγησης, $U_i^*(k)$ η συνάρτηση χρησιμοποίησης εξαρτώμενη από τις συνθήκες του διαύλου για το χρήστη i τη χρονοθυρίδα k , $b_i(k, W_i)$ η ποσότητα δεδομένων που έχει λάβει ο χρήστης i κατά το διάστημα $k - W_i + 1$ έως k , B_i το αντίστοιχο κατώφλι δεδομένων για το διάστημα W_i και g_i^2 το άνω όριο της πιθανότητας αστοχίας βραχυπρόθεσμης ρυθμαπόδοσης. Αξίζει να σημειωθεί ότι δε λαμβάνεται πλέον υπόψιν το κριτήριο της καθυστέρησης εξυπηρέτησης, το οποίο είναι ανεπαρκές, αλλά θεωρείται ότι θα ικανοποιείται με την εγγύηση της συνθήκης του προβλήματος (5.1). Οι συναρτήσεις χρησιμοποίησης $U_i^*(k)$ πρέπει να είναι φραγμένες (άνω και κάτω) έτσι, ώστε η πολιτική δρομολόγησης που θα επιλύει το πρόβλημα (5.1) να είναι εφαρμόσιμη και να μην επηρεάζεται η αποδοτικότητα του συστήματος από τυχόν μακροπρόθεσμες απαιτήσεις που αυτές περιέχουν.

5.2 Η Πολιτική Δρομολόγησης

Η πολιτική USTF ενσωματώνει, όπως μαρτυρά και η ονομασία της, τις απαιτήσεις σε βραχυπρόθεσμη απόδοση στη συνάρτηση κόστους. Ορίζεται, λοιπόν, η πολιτική δρομολόγησης i_{USTF}^* του χρήστη i υπηρεσίας πραγματικού χρόνου τη χρονοθυρίδα k :

$$i_{USTF}^*(k) = \underset{i \in S}{\text{arg min}} \hat{U}_i(k) \quad (5.2)$$

όπου η μη αρνητική συνάρτηση $\hat{U}_i(k)$ δίνεται από τη σχέση:

$$\hat{U}_i(k) = U_i^*(k) + \begin{cases} (-1) \cdot I_i(k) \cdot U_i(k), & \text{αν } b_i(k, W_i) \leq (Tr + 1) B_i \\ G_i(k) \cdot [U_{max}^*(k) - U_i^*(k)], & \text{αν } b_i(k, W_i) > (Tr + 1) B_i \end{cases} \quad (5.3)$$

όπου $I_i(k) \in [0,1]$ και $G_i(k) \in [0,1]$ κανονικοποιημένες παράμετροι που δείχνουν πόσο μεγάλη ανάγκη έχουν οι χρήστες υπηρεσιών πραγματικού χρόνου για πρόσβαση στους πόρους του συστήματος τη χρονοθυρίδα k , ανάλογα με την ποσότητα δεδομένων που έχουν λάβει κατά το χρονικό διάστημα παρατήρησης, Tr μια παράμετρος πυροδότησης του συστήματος που καθορίζει το βαθμό πρόνοιας του συστήματος, $U_i^*(k)$ η συνάρτηση κόστους του χρήστη i εξαρτώμενη από τις συνθήκες του διαύλου και τυχόν απαιτήσεις που θέτει η υπηρεσία του και U_{max}^* η μέγιστη τιμή της.

Σύμφωνα με τον ορισμό της πολιτικής USTF (σχέσεις (5.2), (5.3)), η κατανομή των ληφθέντων δεδομένων για κάθε χρήστη παίζει καθοριστικό ρόλο στη διαμόρφωση της αντίστοιχης συνάρτησης κόστους του. Όταν τα δεδομένα που έχει λάβει ο χρήστης i δεν ισοκατανέμονται στο χρόνο και συνεπώς η ποσότητα δεδομένων που έχουν ληφθεί εντός του χρονικού παραθύρου W_i του εκάστοτε χρήστη i είναι μικρότερη από το κατώφλι B_i που ορίζει η υπηρεσία του, μειώνεται η συνάρτηση κόστους του και συνεπώς αυξάνεται η πιθανότητα εξυπηρέτησής του την επόμενη χρονοθυρίδα.

Συγκεκριμένα, η παράμετρος $I_i(k)$ είναι υπεύθυνη για τη μείωση της συνάρτησης κόστους του χρήστη i σε συνθήκες ελλιπούς εξυπηρέτησης εντός παραθύρου W_i σχισμών πριν από την τρέχουσα χρονοθυρίδα, ακόμα και όταν η (δυσμενής) κατάσταση του διαύλου του δεν το επιτρέπει. Άρα, η παράμετρος $I_i(k)$ δείχνει την ανάγκη του χρήστη i να προσεγγίσει τους πόρους του δικτύου και δίνεται από τη σχέση:

$$I_i(k) = 1 - \sum_{n=1}^{W_i-1} \frac{\omega_i(n, W_i) b'_i(k-n)}{(Tr+1)B_i}, \text{ αν } b_i(k, W_i) \leq (Tr+1)B_i \quad (5.4)$$

όπου:

$b'_i(k)$ η πραγματική ποσότητα δεδομένων που έχει λάβει ο χρήστης i τη χρονοθυρίδα

k , τέτοια, ώστε $b_i(k, W_i) = \sum_{n=1}^{W_i-1} b'_i(k-n)$,

$\omega_i(k, W_i)$ ένα βάρος που ανατίθεται σε καθεμιά χρονοθυρίδα από αυτές που απαρτίζουν το τελευταίο χρονικό διάστημα παρατήρησης από $k - W_i + 1$ μέχρι $k - 1$

που καθορίζει τη σημασία της ποσότητας δεδομένων που έχουν ληφθεί τη συγκεκριμένη χρονοθυρίδα για την παράμετρο $I_i(k)$. Τα συγκεκριμένα βάρη ρυθμίζονται έτσι, ώστε η ποσότητα των δεδομένων που έχει ληφθεί κατά τις πιο απομακρυσμένες χρονοθυρίδες να είναι σημαντικότερη σε σχέση με την αντίστοιχη ποσότητα για τις τωρινές χρονοθυρίδες, καθώς οι πιο απομακρυσμένες είναι εκείνες που επηρεάζουν περισσότερο την τήρηση ή όχι των βραχυπρόθεσμων περιορισμών ρυθμαπόδοσης του χρήστη. Μια ορθή επιλογή για τα βάρη $\omega_i(k, W_i)$ δίνεται στην ακόλουθη σχέση:

$$\omega_i(n, W_i) = \frac{W_i - n}{W_i - 1} \quad (5.5)$$

Έτσι, για παράδειγμα, για χρονικό παράθυρο W_i και τρέχουσα χρονοθυρίδα k τα αντίστοιχα βάρη $\omega_i(n, W_i)$ τίθενται ως εξής:

χρονοθυρίδα	$k - W_i + 1$	$k - W_i + 2$	$k - W_i + 3$...	$k - 2$	$k - 1$
βάρος $\omega_i(n, W_i)$	$1 / (W_i - 1)$	$2 / (W_i - 1)$	$3 / (W_i - 1)$...	$(W_i - 2) / (W_i - 1)$	$(W_i - 1) / (W_i - 1)$

Η επιλογή για τα $\omega_i(n, W_i)$ σύμφωνα με τη σχέση (5.5) σημαίνει ότι ακόμα και αν δύο χρήστες i, j υπηρεσίας πραγματικού χρόνου με τις ίδιες συνθήκες διαύλου, δηλαδή $U_i^*(k) = U_j^*(k)$, έχουν λάβει την ίδια ποσότητα δεδομένων $b_i(k, W_i) = b_j(k, W_j)$ κατά το ίδιο χρονικό διάστημα παρατήρησης $W_i = W_j$, αλλά ο χρήστης i έχει εξυπηρετηθεί σε πιο πρόσφατες χρονοθυρίδες από το χρήστη j , τότε θα προτιμηθεί ο χρήστης j για εξυπηρέτηση, καθώς θα ισχύει $I_i(k) < I_j(k)$, το οποίο κατά τη σχέση (5.3) οδηγεί σε μείωση της συνάρτησης $\hat{U}_i(k)$.

Όσον αφορά τώρα την περίπτωση που ο χρήστης i έχει λάβει περισσότερα δεδομένα από αυτά που ορίζει το κατώφλι B_i κατά τις προηγούμενες $W_i - 1$ χρονοθυρίδες, δηλαδή $b_i(k, W_i) > (Tr + 1)B_i$, τότε ο χρήστης i δεν έχει άμεση ανάγκη

για εξυπηρέτηση. Σε αυτήν την περίπτωση φαίνεται ο ρόλος της παραμέτρου $G_i(k)$, η οποία αναλαμβάνει να σε αυτό το σημείο να αυξήσει την τιμή της συνάρτησης $\hat{U}_i(k)$, ώστε να λάβουν προτεραιότητα χρήστες με ανεκπλήρωτες τις απαιτήσεις τους σε βραχυπρόθεσμη απόδοση. Η $G_i(k)$ δίνεται τότε από τη σχέση:

$$G_i(k) = \frac{\sum_{n=1}^{W_i-1} \omega_i(n, W_i) \cdot b'_i(k-n)}{\sum_{n=1}^{W_i-1} \omega_i(n, W_i) \cdot R_i^{max} \cdot T_{slot}}, \text{ αν } b_i(k, W_i) > (Tr + I)B_i \quad (5.6)$$

όπου R_i^{max} ο μέγιστος ρυθμός μετάδοσης του χρήστη i που μπορεί να φιλοξενήσει το σύστημα, $T_{slot} = 1.67 \text{ msec}$ η χρονική διάρκεια μιας χρονοθυρίδας. Ο παρονομαστής της σχέσης (5.6) χρησιμοποιείται για λόγους κανονικοποίησης, προκειμένου η μέγιστη τιμή της παραμέτρου $G_i(k)$ να είναι 1. Η παράμετρος $G_i(k)$ δείχνει τη χαμηλή ανάγκη του χρήστη i να εξυπηρετηθεί την επόμενη χρονική σχισμή. Άρα, αν ο χρήστης i έχει λάβει δεδομένα κοντά στη μέγιστη τιμή που μπορεί να λάβει, τότε η παράμετρος $G_i(k)$ αυξάνεται, με αποτέλεσμα να αυξάνεται και η συνάρτηση $\hat{U}_i(k)$ και να δίνεται προτεραιότητα εξυπηρέτησης σε άλλους χρήστες.

Ακόμη, η παράμετρος πυροδότησης που υπεισέρχεται στις συνθήκες της σχέσης (5.3) Tr έχει το ρόλο του βαθμού πρόνοιας που λαμβάνει το δίκτυο. Έτσι, η παράμετρος Tr αυξάνει την τιμή του κατωφλίου B_i , με το οποίο συγκρίνεται κάθε χρονοθυρίδα η ποσότητα των δεδομένων που έχει λάβει ο χρήστης i κατά τις προηγούμενες διαδοχικές $W_i - 1$ χρονοθυρίδες, εξαναγκάζοντας κατά κάποιο τρόπο την ταχύτερη τήρηση των απαιτήσεων βραχυπρόθεσμης απόδοσης και εξασφαλίζοντας εγκυρότερα τον απαιτούμενο ρυθμό μετάδοσης (και λίγο παραπάνω από αυτόν). Πρόκειται ουσιαστικά για δικλείδα ασφαλείας. Μεγάλες τιμές της Tr καταλήγουν σε περισσότερο προνοητικό αλγόριθμο και κατά συνέπεια σε μικρότερες πιθανότητες αστοχίας της δικαιοσύνης βραχυπρόθεσμης απόδοσης. Σε αντίθεση, το τίμημα μια τέτοιας επιλογής είναι η μείωση της συνολικής απόδοσης του συστήματος.

Αξίζει να σημειωθεί ότι ο αλγόριθμος USTF, όχι μόνο αυξάνει την πιθανότητα εξυπηρέτησης των χρηστών που δεν έχουν εξυπηρετηθεί επιμελώς μέσω της παραμέτρου $I_i(k)$, αλλά επίσης μειώνει την προτεραιότητα εξυπηρέτησης των

χρηστών που έχαιραν προνομιακής μεταχείρισης κατά το πρόσφατο παρελθόν μέσω της παραμέτρου $G_i(k)$. Συνεπώς, ο αλγόριθμος συνολικά μετριάζει τον οπορτουνιστικό του χαρακτήρα, σύμφωνα με τον οποίο θα έπρεπε να δίνεται μεγαλύτερη βαρύτητα στις συνθήκες του διαύλου των χρηστών. Αντιθέτως, ο USTF λαμβάνει υπόψη του μεν τις συνθήκες του διαύλου, αλλά με περιορισμένη ισχύ ως προς την επιλογή του χρήστη που θα λάβει εξυπηρέτηση, κυρίως όσον αφορά τους χρήστες υπηρεσιών πραγματικού χρόνου. Κατ' αυτόν τον τρόπο επιτυγχάνονται οι επιθυμητοί ρυθμοί μετάδοσης για όλους τους χρήστες υπηρεσίας πραγματικού χρόνου, αφήνοντας το παιχνίδι της μεγιστοποίησης της απόδοσης στους χρήστες υπηρεσίας μη πραγματικού χρόνου.

Ένα σημείο που παρουσιάζει ενδιαφέρον στον αλγόριθμο USTF είναι το γεγονός ότι η συνάρτηση κόστους $\hat{U}_i(k)$ χαρακτηρίζεται από αθροιστική σχέση ανάμεσα στη συνάρτηση που εκφράζει την κατάσταση του διαύλου του χρήστη και την προτεραιότητα εξυπηρέτησης με βάση το βαθμό εκπλήρωσης των κριτηρίων βραχυπρόθεσμης απόδοσης. Αυτό κάνει ευκολότερη τη συνύπαρξη σε ένα σύστημα χρηστών υπηρεσιών τόσο πραγματικού όσο και μη πραγματικού χρόνου, που αποτελεί και βασική επιδίωξη των σύγχρονων επικοινωνιών, καθώς δίνει την ευελιξία στο διαχειριστή να θεωρήσει μια αντίστοιχη συνάρτηση κόστους για τους χρήστες υπηρεσίας μη πραγματικού χρόνου.

Τέλος, αξίζει να σημειωθεί ότι η γενικότερη ιδέα της αντιστοίχισης των απαιτήσεων της ποιότητας υπηρεσίας των χρηστών υπηρεσιών πραγματικού χρόνου στην πολιτική δρομολόγησης αποσυσχετίζει τις διαδικασίες αντίδρασης του δρομολογητή, όταν παραβιάζονται οι μεν ή δε απαιτήσεις. Αυτό σημαίνει πως τυχόν μακροπρόθεσμες απαιτήσεις ενσωματωμένες στη συνάρτηση κόστους της πολιτικής δρομολόγησης –κατάλληλα ορισμένης– δεν επηρεάζουν την αποδοτικότητα του συστήματος, υπό την προϋπόθεση ότι είναι άνω και κάτω φραγμένη.

6

ΠΡΟΣΟΜΟΙΩΣΕΙΣ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

Μετά τη θεωρητική ανάλυση των προηγούμενων κεφαλαίων κρίνεται σκόπιμη η προσομοίωση ορισμένων από τους αλγόριθμους που περιγράφηκαν, προκειμένου να εκτιμηθούν οι επιδόσεις τους και να εξακριβωθεί ότι όντως πληρούν τις «υποσχέσεις» τους για την τήρηση των συνθηκών που θέτουν. Τα αποτελέσματα παρέχονται μέσω διαγραμμάτων, τα οποία επιτρέπουν τη σύγκριση μεταξύ των υπό εξέταση αλγόριθμων, όσον αφορά τα κριτήρια ποιότητας ενδιαφέροντος. Οι αλγόριθμοι που προσομοιώθηκαν είναι οι MAX, M-LWDF, EXP και USTF.

6.1 Το Μοντέλο του Συστήματος Προσομοίωσης

Θεωρήθηκε το μοντέλο όπως περιγράφεται στην παράγραφο 1.4. Το σύστημα συνίσταται από $N=7$ ενεργούς χρήστες, οι οποίοι έχουν πάντα δεδομένα προς μετάδοση στις ουρές τους. Η μέση τιμή του σηματοθορυβικού λόγου SINR του διαύλου κάθε χρήστη i θεωρείται ίση με $\bar{\rho}_i = 6dB$. Προκειμένου να ληφθεί υπόψη και η επίδραση των μεταβαλλόμενων καταστάσεων των διαύλων των χρηστών στη ρυθμαπόδοση του συστήματος, οι διάυλοι μοντελοποιούνται μέσω του μοντέλου

αλυσίδας Markov 8-καταστάσεων ([18]) ως διάλοιο Rayleigh εξασθένισης. Ο πίνακας 6.1 παρουσιάζει τις στιγμιαίες τιμές SINR και τις αντίστοιχες στιγμιαίου ρυθμού μετάδοσης, στην οποία βρίσκεται ο χρήστης i κάθε χρονοθυρίδα ανάλογα με την κατάσταση του διαύλου του. Το γεγονός ότι χρησιμοποιήθηκαν οι τιμές του συγκεκριμένου πίνακα για το στιγμιαίο ρυθμό μετάδοσης αντί των τιμών που θα προέκυπταν από τη σχέση (1.2) οδηγεί σε πιο ρεαλιστικά αποτελέσματα, αφού το ίδιο συμβαίνει και στο σύστημα CDMA 1xEV-DO.

SNR \geq (dB)	-12.5	-9.5	-8.5	-6.5	-5.7	-4.0	-1.0	1.3	3.0	7.2	9.5
Ρυθμός Μετάδοσης (Kbps)	38.4	76.8	102.6	153.6	204.8	307.2	614.4	921.6	1228.8	1843.2	2457.6

Πίνακας 6. 1

Ο επιθυμητός ρυθμός μετάδοσης για τις υπηρεσίες πραγματικού χρόνου θεωρείται ίσος με $64 Kbps$, δηλαδή θεωρείται ότι πρόκειται για υπηρεσίες φωνής. Τα όρια των πιθανοτήτων αστοχίας καθυστέρησης και δικαιοσύνης βραχυπρόθεσμης ρυθμαπόδοσης θεωρούνται ίσα και λαμβάνουν την τιμή $g_i^1 = g_i^2 = 0.005$, δηλαδή επιδιώκονται σχεδόν μηδενικά όρια για τις πιθανότητες αστοχίας. Το κατώφλι καθυστέρησης T_i είναι μεταβλητό και λαμβάνει διαδοχικά τις τιμές 20, 50 και 100 χρονοθυρίδες, που αντιστοιχούν σε 33.4, 83.5 και 167 msec. Το χρονικό παράθυρο παρατήρησης της τήρησης των απαιτήσεων σε βραχυπρόθεσμη ρυθμαπόδοση W_i μεταβάλλεται παράλληλα με το κατώφλι καθυστέρησης T_i έτσι, ώστε να λαμβάνουν την ίδια τιμή, οπότε προκύπτουν τα αντίστοιχα ελάχιστα κατώφλια ληφθέντων δεδομένων B_i 2137.6, 5344, 10688 bits. Για τον αλγόριθμο δρομολόγησης ESFS θεωρήθηκε ότι τα βάρη $\phi_i = \frac{1}{N}$ και για τις παραμέτρους T , H δόθηκαν οι τιμές $T = 0.1$ και $H = 0.3$, οι οποίες σύμφωνα με το [9] αντιστοιχούν σε αυστηρή δικαιοσύνη βραχυπρόθεσμης απόδοσης. Επίσης, για τον αλγόριθμο USTF τέθηκε $Tr = 0.3$, $U_i^*(k) = U_{max}^* - \rho_i(k)$ και $U_{max}^* = 20$.

Επίσης, θεωρήθηκαν δύο σενάρια υλοποίησης του συστήματος. Στο πρώτο (Σ1) υπάρχουν μόνο χρήστες υπηρεσιών πραγματικού χρόνου, ενώ στο δεύτερο (Σ2) συνυπάρχουν με χρήστες υπηρεσιών μη πραγματικού χρόνου. Ο αριθμός των

χρηστών υπηρεσιών πραγματικού χρόνου N_{RT} στο Σ2 λαμβάνει τις τιμές $N_{RT} = 1, 3, 5$. Για την περίπτωση των χρηστών υπηρεσιών μη πραγματικού χρόνου οι συναρτήσεις κόστους των αντίστοιχων αλγορίθμων υπολογίζονται ως ακολούθως.

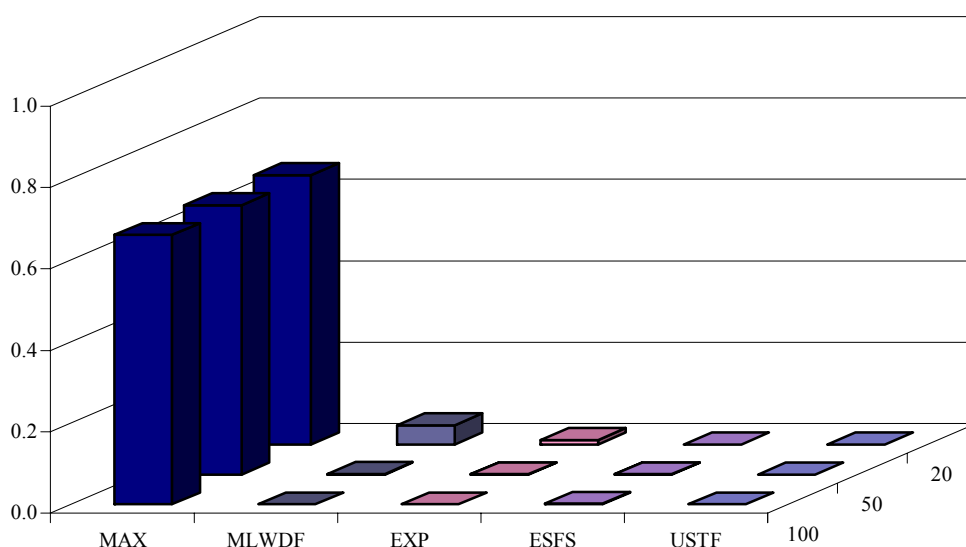
- για τον αλγόριθμο MAX, ο οποίος δεν κάνει διάκριση ανάμεσα σε υπηρεσίες πραγματικού ή μη χρόνου, η συνάρτηση κόστους είναι η ίδια και για τις δύο κλάσεις χρηστών.
- στον αλγόριθμο M-LWDF η συνάρτηση κόστους υπολογίζεται κατ' αντιστοιχία με τη συνάρτηση κόστους που προκύπτει από τη σχέση (3.1) ως
$$U_{i,MLWDF}^*(k) = \frac{\rho_i(k)}{\bar{\rho}_i}.$$
- για τον αλγόριθμο EXP η συνάρτηση κόστους των χρηστών υπηρεσιών μη πραγματικού χρόνου είναι η ίδια με την αντίστοιχη συνάρτηση στον αλγόριθμο M-LWDF.
- για τον αλγόριθμο ESFS, θεωρείται καθαρά ομορτουριστική δρομολόγηση των χρηστών υπηρεσιών μη πραγματικού χρόνου, οπότε η συνάρτηση κόστους είναι
$$\hat{U}_{i,ESFS}(k) \equiv U_{i,ESFS}^*(k) = 20 - \rho_i(k).$$
- τέλος, στον αλγόριθμο USTF το ίδιο με τον αλγόριθμο ESFS.

Η προσομοίωση διαρκεί επί 20.000 χρονοθυρίδες, δηλαδή αρκετό χρονικό διάστημα, για να αξιολογηθούν και τυχόν μακροπρόθεσμες επιδόσεις. Η πληροφορία που συλλέγεται στο τέλος κάθε προσομοίωσης αφορά την πιθανότητα αστοχίας καθυστέρησης εξυπηρέτησης, την πιθανότητα αστοχίας δικαιοσύνης βραχυπρόθεσμης απόδοσης και το μέσο ρυθμό μετάδοσης κατά τη διάρκεια της προσομοίωσης ανά χρήστη.

6.2 Τα Αποτελέσματα της Προσομοίωσης για το Σ1

Στην παρούσα ενότητα παρουσιάζονται τα αποτελέσματα των προσομοιώσεων του Σ1 σε συγκριτικά διαγράμματα, που επιτρέπουν την αξιολόγηση των αλγορίθμων ξεχωριστά, αλλά και τη σύγκριση μεταξύ τους. Στόχος είναι να διαπιστωθεί και να εξηγηθεί ότι οι αλγόριθμοι που τηρούν μόνο τις απαιτήσεις χαμηλής καθυστέρησης

εξυπηρέτησης δεν ικανοποιούν πλήρως το χρήστη, καθώς δεν τηρούν και τις απαιτήσεις σε βραχυπρόθεσμη απόδοση. Επιπλέον, αναμένεται ότι ο αλγόριθμος MAX είναι ακατάλληλος για τον προγραμματισμό εξυπηρέτησης χρηστών υπηρεσιών πραγματικού χρόνου, αφού δεν περιλαμβάνει κάποιο από τα κριτήρια της ποιότητας υπηρεσίας τους.

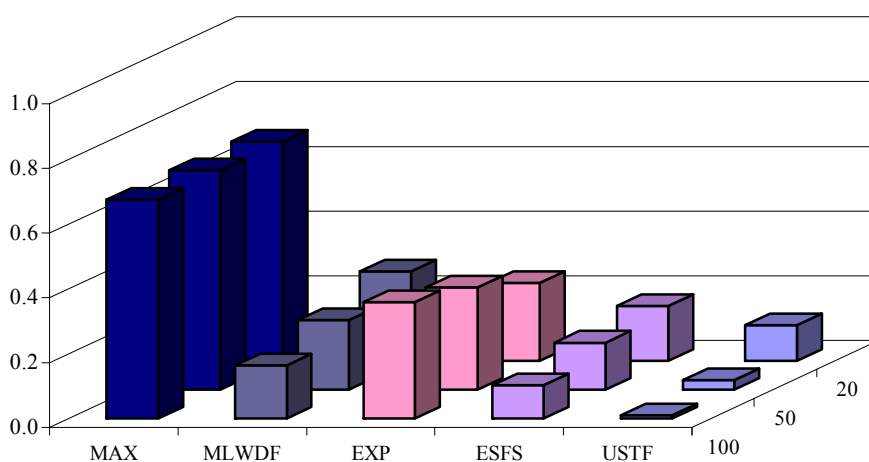


Σχήμα 6. 1: Πιθανότητα αστοχίας καθυστέρησης εξυπηρέτησης συναρτήσει του κατωφλίου καθυστέρησης T_i ανά αλγόριθμο σε σύστημα όπου υπάρχουν μόνο RT χρήστες.

Το σχήμα 6.1 παρουσιάζει τη μέση πιθανότητα αστοχίας της καθυστέρησης εξυπηρέτησης συναρτήσει του μεγέθους της μέγιστης επιτρεπτής καθυστέρησης (T_i) για κάθε αλγόριθμο. Είναι φανερό, όπως και αναμενόμενο, πως οι αλγόριθμοι που χαρακτηρίζονται ως «ευαίσθητοι» στην καθυστέρηση (M-LWDF, EXP) εκπληρώνουν τις απαιτήσεις σε καθυστέρηση με τον EXP να δίνει καλύτερα αποτελέσματα σε μικρότερες τιμές κατωφλίου. Επιπρόσθετα, οι νεότεροι ESFS και USTF καταλήγουν σε άριστες επιδόσεις σε σχέση με τους περιορισμούς καθυστέρησης. Οι αντίστοιχες πιθανότητες αστοχίας είναι μηδενικές για τις τιμές κατωφλίου T_i 50 και 100 χρονοθυρίδων στον ESFS και σχεδόν μηδενικές για $T_i = 20$ χρονοθυρίδες, ενώ στον USTF είναι μηδενικές για όλες τις τιμές κατωφλίου.

Αντίθετα, ο «άπληστος» αλγόριθμος δρομολόγησης MAX δίνει πιθανότητες αστοχίας καθυστέρησης απαγορευτικές για μια υπηρεσία πραγματικού χρόνου. Παρατηρείται, δε, ότι ο αλγόριθμος MAX δεν εξαρτάται από την επιλογή του κατωφλίου καθυστέρησης, αφού δε λαμβάνει καμία μέριμνα γι' αυτό. Επίσης, προκύπτει από το σχήμα 6.1 ότι οι τιμές των πιθανοτήτων αστοχίας βελτιώνονται στην πορεία της ιστορίας, δηλαδή από τον παλαιότερο προς τους νεότερους αλγορίθμους, καθώς αυτός ήταν και ο λόγος που οδήγησε στο σχεδιασμό τους μέχρι και τον ESFS.

Ακολούθως, το σχήμα 6.2 δείχνει τη μέση τιμή της πιθανότητας αστοχίας της δικαιοσύνης βραχυπρόθεσμης απόδοσης συναρτήσει του μεγέθους του παραθύρου παρατήρησης W_i ανά αλγόριθμο. Μεγαλύτερες τιμές της συγκεκριμένης πιθανότητας αστοχίας αναδεικνύουν μεγαλύτερη αδυναμία του δρομολογητή να εκπληρώσει τους περιορισμούς του ελάχιστου απαιτούμενου αριθμού δεδομένων που λαμβάνονται εντός μικρών χρονικών διαστημάτων των υπηρεσιών πραγματικού χρόνου.

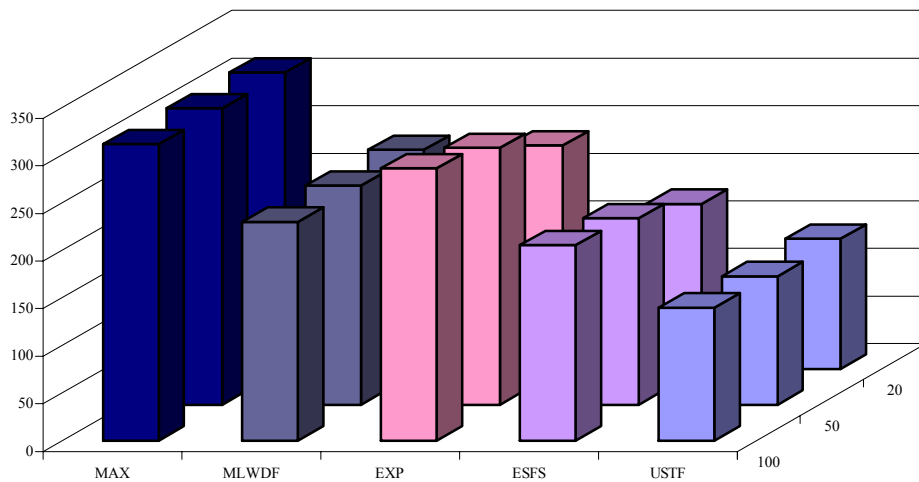


Σχήμα 6. 2: Πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης συναρτήσει του μεγέθους του παραθύρου παρατήρησης W_i ανά αλγόριθμο, όταν υπάρχουν μόνο RT χρήστες στο σύστημα.

Από το σχήμα 6.2 προκύπτει ότι μόνο ο αλγόριθμος USTF ικανοποιεί τις απαιτήσεις βραχυπρόθεσμης απόδοσης και οι επιδόσεις του βελτιώνονται όσο αυξάνεται το παράθυρο παρατήρησης, πράγμα λογικό, αφού χαλαρώνει το διάστημα επιβολής των απαιτήσεων. Η αύξηση του W_i σημαίνει ότι δείχνεται μεγαλύτερη

ανοχή στην υπέρβαση των ορίων ρυθμαπόδοσης σε μικρά χρονικά διαστήματα, οπότε η υπηρεσία που υποστηρίζεται γίνεται χαλαρότερη από πλευράς σταθερότητας του ρυθμού μετάδοσης. Πάντως, ο USTF υποστηρίζει ικανοποιητικά υπηρεσίες είτε χαλαρές είτε αυστηρές σε ρυθμούς μετάδοσης. Η μέση πιθανότητα αστοχίας των βραχυπρόθεσμων στόχων σε απόδοση για τον USTF κυμαίνεται από 0.11 στο στενότερο παράθυρο W_i μέχρι τιμές τάξης του 0.01 για τις επόμενες τιμές του W_i , δηλαδή παρέχει ικανοποιητική ποιότητα υπηρεσίας. Αυτό, σε συνδυασμό με την ικανοποίηση των απαιτήσεων καθυστέρησης καθιστά τον USTF ιδιαίτερα αποτελεσματικό αλγόριθμο, όσον αφορά την εξυπηρέτηση χρηστών με υψηλές απαιτήσεις σε ποιότητα επικοινωνίας, καθώς καταφέρνει να προσφέρει υπηρεσία και γρήγορα και αποδοτικά.

Εκείνο που φαίνεται αντικρουόμενο στα διαγράμματα 6.1 και 6.2 είναι ότι οι αλγόριθμοι M-LWDF, EXP και ESFS, οι οποίοι κατά το σχήμα 6.1 ικανοποιούν τις απαιτήσεις καθυστέρησης, δεν ικανοποιούν τις απαιτήσεις βραχυπρόθεσμης ρυθμαπόδοσης. Επεκτείνοντας την παρατήρηση στο σχήμα 6.3, όπου δίνεται η μέση απόδοση των χρηστών συναρτήσει του κατωφλίου T_i (ή και του παραθύρου W_i μιας και είναι ίσα), διαπιστώνεται ότι οι συγκεκριμένοι αλγόριθμοι παρουσιάζουν υψηλότερη μέση ρυθμαπόδοση από τον αλγόριθμο USTF, ο οποίος ικανοποιεί και τα δύο κριτήρια. Επίσης, ο αλγόριθμος MAX, που δεν ικανοποιεί κανένα κριτήριο, παρουσιάζει τη μεγαλύτερη μέση απόδοση από όλους τους άλλους αλγορίθμους. Άρα, μπορεί να ειπωθεί ότι η μεγιστοποίηση της απόδοσης του συστήματος αυτή καθεαυτή δεν είναι ο αυτοσκοπός, αλλά ούτε και σημαίνει μεγαλύτερη ευχαρίστηση των χρηστών με βάση την ποιότητα της υπηρεσίας που λαμβάνουν, ιδιαίτερα όταν υπηρεσίες με διαφορετικά χαρακτηριστικά πρέπει να συνυπάρχουν στο ίδιο ασύρματο δίκτυο (όπως άλλωστε φανερώνει και ο σχεδιασμός των αλγορίθμων του κεφαλαίου 2 για υπηρεσίες ακόμα μη πραγματικού χρόνου).

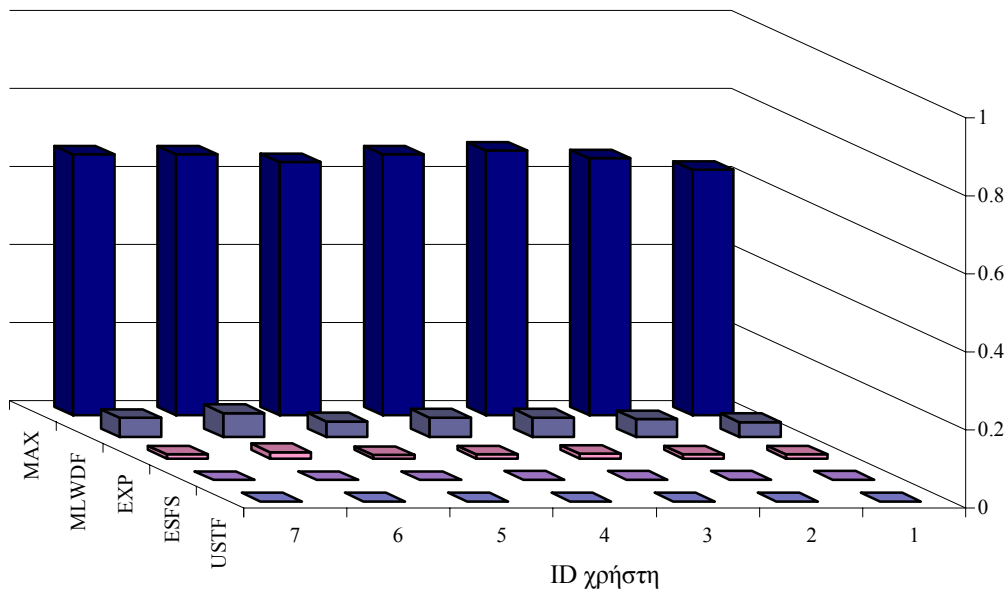


Σχήμα 6. 3: Μέση ρυθμαπόδοση σε (Kbps) κάθε χρήστη συναρτήσει του κατωφλίου καθυστέρησης T_i ανά αλγόριθμο, όταν υπάρχουν μόνο RT χρήστες στο σύστημα.

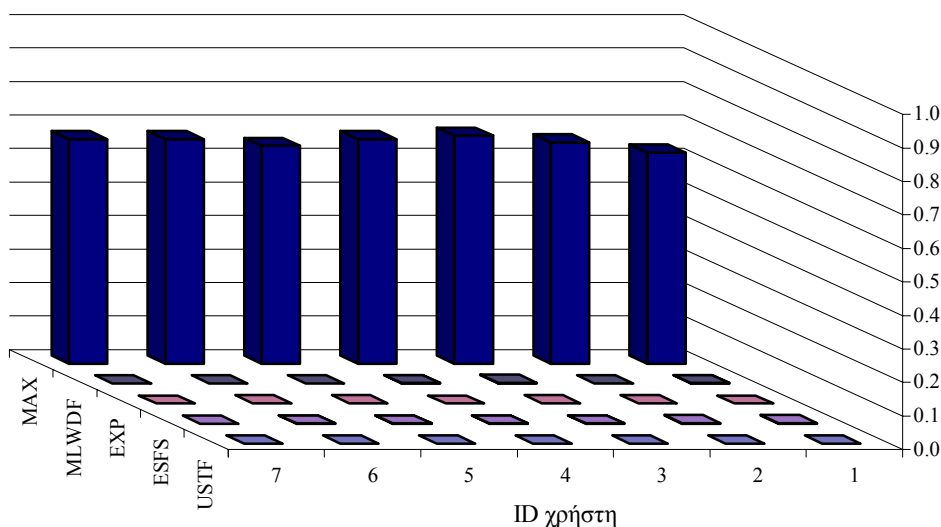
Το σημαντικό, λοιπόν, συμπέρασμα που προκύπτει από το σχήμα 6.3 είναι ότι οι αλγόριθμοι MAX, M-LWDF, EXP και ESFS χάριν της υψηλής απόδοσης μετά από παρέλευση αρκετού χρόνου εστιάζουν περισσότερο στον ομορτυνιστικό τους χαρακτήρα, χωρίς να φροντίζουν για την τήρηση σταθερού ρυθμού μετάδοσης. Με άλλα λόγια, οι χρήστες εξυπηρετούνται με περιορισμό την καθυστέρηση, αλλά η ποσότητα των δεδομένων που λαμβάνονται δεν είναι ίσα κατανεμημένη σε μικρά διαστήματα παρατήρησης. Έτσι, ο μεγαλύτερος μέσος ρυθμός μετάδοσης είναι αποτέλεσμα μικρών ρυθμών μετάδοσης με βάση τη στιγμιαία τιμή του $\rho_i(k)$ σε κάποια διαστήματα και μεγαλύτερων ρυθμών μετάδοσης σε κάποια άλλα. Ο USTF, αντιθέτως, φροντίζει για την ισοκατανομή των δεδομένων που λαμβάνονται εντός του μικρού χρονικού παραθύρου και επιτυγχάνει χαμηλότερο μέσο ρυθμό μετάδοσης από τους υπόλοιπους αλγορίθμους, αλλά αρκετό, ώστε να ικανοποιούνται οι απαιτήσεις των υπηρεσιών πραγματικού χρόνου, και μάλιστα μεγαλύτερο από τον απαιτούμενο των 64 Kbps.

Επιπρόσθετα, από το σχήμα 6.3 προκύπτει ότι ο USTF παρουσιάζει σταθερότητα ως προς το μέσο ρυθμό μετάδοσης των χρηστών, ο οποίος είναι ανεξάρτητος από την τιμή του μεγέθους του παραθύρου W_i . Αυτή η ιδιότητα είναι ιδιαίτερως σημαντική, διότι αποδεικνύεται ότι σε κάθε κατώφλι επιτυγχάνεται σταθερός ρυθμός, κάτι πολύ

σημαντικό για υπηρεσίες πραγματικού χρόνου. Επίσης, ο ρυθμός μετάδοσης είναι σταθερός για τον MAX, ο οποίος δεν κάνει διάκριση σε υπηρεσίες, αλλά και για τον M-LWDF, ο οποίος όμως δεν τηρεί τις βραχυπρόθεσμες απαιτήσεις απόδοσης.



Σχήμα 6. 4: Πιθανότητα αστοχίας καθυστέρησης ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης T_i 20 χρονοθυρίδων.



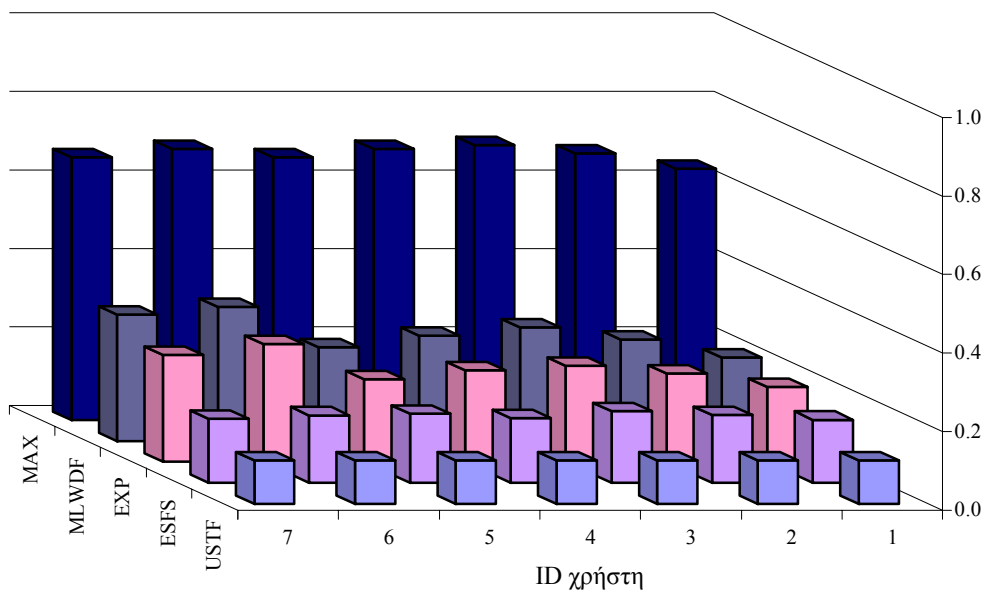
Σχήμα 6. 5: Πιθανότητα αστοχίας καθυστέρησης ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης T_i 50 χρονοθυρίδων.

Τα σχήματα 6.4 και 6.5 παρουσιάζουν τη μεταβολή των πιθανοτήτων αστοχίας καθυστέρησης ανά χρήστη του συστήματος για διαφορετικό κατώφλι T_i το καθένα. Οι ευαίσθητοι σε καθυστέρηση αλγόριθμοι παρουσιάζουν πολύ μικρές πιθανότητες αστοχίας καθυστέρησης για όλους τους χρήστες, οπότε τους συμπεριφέρονται δίκαια ως προς το συγκεκριμένο θέμα. Ο αλγόριθμος MAX επίσης έχει ίδια συμπεριφορά για τους χρήστες του συστήματος. Κάτι παρόμοιο δεν αναμένεται στην περίπτωση που κάποιοι από τους χρήστες είχαν δυσμενέστερες συνθήκες διαύλου από τους υπόλοιπους, καθώς τότε οι συγκεκριμένοι χρήστες θα κατέληγαν πιθανώς σε αποκλεισμό.

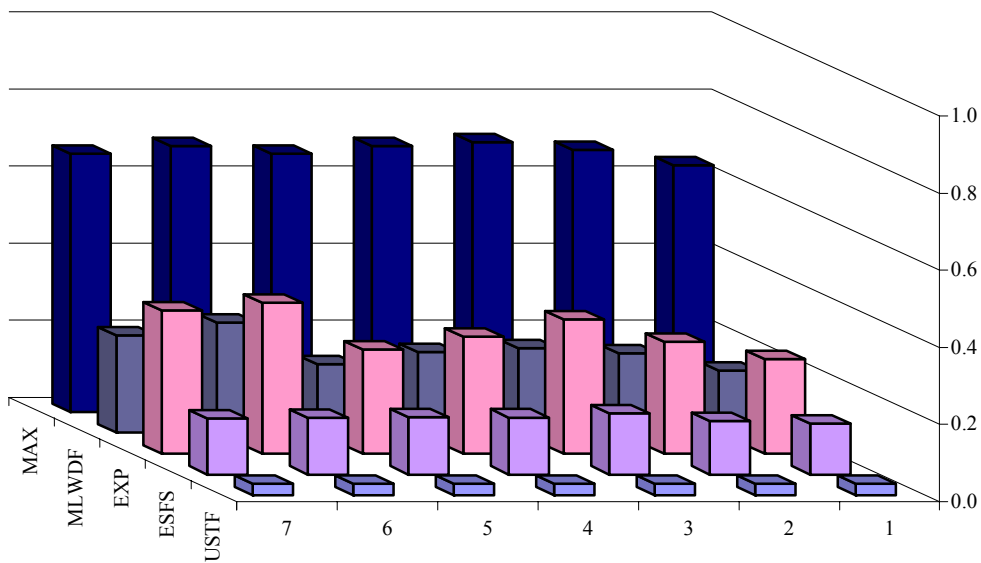
Αντίστοιχα, τα σχήματα 6.6 και 6.7 δείχνουν την πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης ανά χρήστη του συστήματος για διαφορετικό παράθυρο W_i το καθένα. Εδώ οι αλγόριθμοι M-LWDF και EXP χαρακτηρίζονται από ανομοιόμορφη συμπεριφορά για τις αντίστοιχες πιθανότητες αστοχίας. Η ομοιομορφία που χαρακτηρίζει τον ESFS στο συγκεκριμένο θέμα οφείλεται στο γεγονός ότι και αυτός, όπως ο USTF, κάνει ισοκατανομή σε κάποιο χρονικό πλαίσιο, αλλά για τον αριθμό των χρονοθυρίδων εξυπηρέτησης. Έτσι, δεδομένου ότι οι χρήστες έχουν δίαυλο με ίδια μέση τιμή $\bar{\rho}_i$ και λαμβάνουν ίδιο ποσοστό

εξυπηρέτησης ως προς τον αριθμό των χρονοθυρίδων, είναι λογικό να έχουν και την ίδια πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης.

Επίσης, ο USTF διατηρεί τις πιθανότητες αστοχίας βραχυπρόθεσμης ρυθμαπόδοσης το ίδιο χαμηλές για όλους τους χρήστες, όπως άλλωστε δηλώνουν και οι σχέσεις ορισμού του. Από τη στιγμή που εξασφαλίζει ίδιο σταθερό ρυθμό μετάδοσης για όλους τους χρήστες, η διατήρηση ίδιων πιθανοτήτων αστοχίας είναι η άλλη όψη αυτού του νομίσματος.



Σχήμα 6. 6: Πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης ανά χρήστη και ανά αλγόριθμο για παράθυρο παρατήρησης $W_i = 20$ χρονοθυρίδων, όταν όλοι οι χρήστες είναι RT.

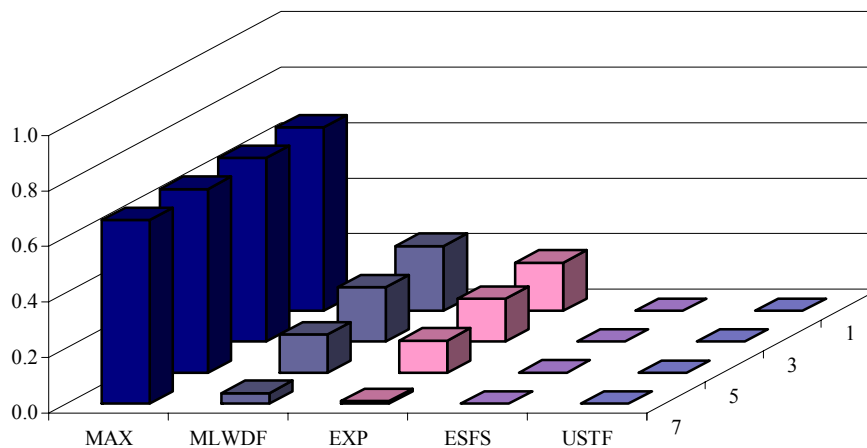


Σχήμα 6. 7: Πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης ανά χρήστη και ανά αλγόριθμο για παράθυρο παρατήρησης $W_i = 50$ χρονοθυρίδων, όταν όλοι οι χρήστες είναι RT.

6.3 Τα Αποτελέσματα της Προσομοίωσης για το Σ2

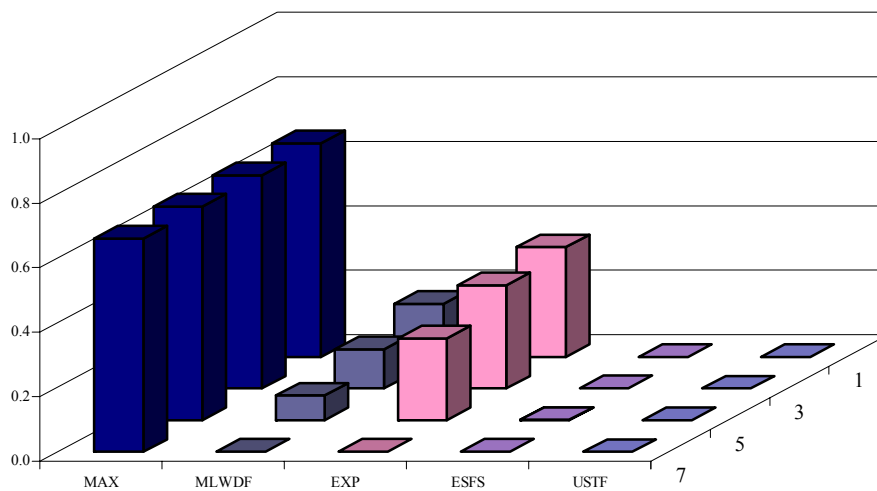
Αφού ολοκληρώθηκε η μελέτη του συστήματος που περιείχε μόνο χρήστες υπηρεσιών πραγματικού χρόνου (Real-Time services - RT), ακολουθεί η ανάλυση των αποτελεσμάτων του συστήματος που περιέχει επίσης χρήστες υπηρεσίας μη πραγματικού χρόνου (Non Real-Time services - NRT), ώστε να αξιολογηθεί η συνύπαρξή τους υπό τη σκέπη καθενός από τους παραπάνω αλγορίθμους δρομολόγησης.

Τα σχήματα 6.8 και 6.9 δείχνουν τη μέση πιθανότητα αστοχίας των χρηστών υπηρεσίας πραγματικού χρόνου υπό μεταβλητό πλήθος αυτών για τιμή κατωφλίου T_i ίση με 20 και 50 χρονοσχιμές αντίστοιχα. Διαπιστώνεται ότι οι αλγόριθμοι MAX, M-LWDF και EXP δεν καλύπτουν ικανοποιητικά τις απαιτήσεις καθυστέρησης, όταν υπεισέρχονται και χρήστες υπηρεσίας μη πραγματικού χρόνου στο σύστημα.



Σχήμα 6. 8: Πιθανότητα αστοχίας καθυστέρησης συναρτήσει του αριθμού των χρηστών RT ανά αλγόριθμο για κατώφλι 20 χρονοθυρίδων.

Συγκεκριμένα, ο MAX έχει τα ίδια αποτελέσματα με το Σ1, διότι μοναδικό κριτήριο της επιλογής του επόμενου χρήστη για εξυπηρέτηση είναι η μεγιστοποίηση της απόδοσης και όχι κάποιο κριτήριο ποιότητας υπηρεσίας. Για τους αλγορίθμους M-LWDF και EXP η υπέρβαση του ορίου καθυστέρησης αυξάνεται με τη μείωση του αριθμού N_{RT} των χρηστών υπηρεσίας πραγματικού χρόνου. Αυτό οφείλεται στο γεγονός ότι οι συναρτήσεις κόστους για τους NRT χρήστες μεγιστοποιούνται (όπως είναι επιθυμητό στις εν λόγω πολιτικές) ευκολότερα από αυτές των RT χρηστών με βάση τις παραμέτρους των αντίστοιχων πολιτικών. Αντίθετα, εξ ορισμού η συνάρτηση κόστους για τους NRT χρήστες στους αλγορίθμους ESFS και USTF ελαχιστοποιείται (επιθυμητό για αυτές πολιτικές) δυσκολότερα, δίνοντας κατ' αυτόν τον τρόπο προτεραιότητα στους χρήστες υπηρεσίας πραγματικού χρόνου, ασχέτως του αριθμού τους στο σύστημα.



Σχήμα 6. 9: Πιθανότητα αστοχίας καθυστέρησης συναρτήσει του αριθμού των χρηστών RT ανά αλγόριθμο για κατώφλι 50 χρονοθυρίδων.

Σαν αποτέλεσμα, οι αλγόριθμοι M-LWDF και EXP παρουσιάζουν ακόμα χειρότερα αποτελέσματα ως προς τη βραχυπρόθεσμη ρυθμαπόδοση, όπως δείχνουν τα σχήματα 6.10 και 6.11. Μάλιστα, δίνουν παρόμοιες τιμές με αυτές που παρέχει ο αναίσθητος στην καθυστέρηση MAX αλγόριθμος σε βαθμό ο EXP να τις ξεπερνάει για παράθυρο W_i μήκους 50 χρονοθυρίδων, οπότε πρόκειται για πλήρη αποκλεισμό των χρηστών υπηρεσιών πραγματικού χρόνου σε αυτήν την περίπτωση.

Και όμως μια πρώτη ματιά στα σχήματα 6.16 και 6.17 της μέσης ρυθμαπόδοσης του συστήματος δείχνει ότι οι ρυθμοί μετάδοσης του μέσου χρήστη είναι υψηλοί για τους συγκεκριμένους αλγορίθμους, οπότε θα έπρεπε να επιτυγχάνονται οι απαιτήσεις σε ρυθμούς μετάδοσης. Η παρατήρηση και των υπόλοιπων σχημάτων φανερώνει ότι οι υψηλοί ρυθμοί μετάδοσης αντιστοιχούν μόνο στους χρήστες υπηρεσιών μη πραγματικού χρόνου.

Στα σχήματα 6.12 και 6.13 δίνεται η μέση ρυθμαπόδοση των RT χρηστών συναρτήσει του πλήθους των RT χρηστών ανά αλγόριθμο και για διαφορετικό κατώφλι. Τα σχήματα 6.14 και 6.15 είναι αντίστοιχα, αλλά για τους NRT χρήστες. Επίσης στα σχήματα 6.18 έως 6.22 είναι συγκριτικά της μέσης ρυθμαπόδοσης RT και

NRT χρηστών για διαφορετικό πλήθος N_{RT} των RT χρηστών και για διαφορετικά κατώφλια καθυστέρησης.

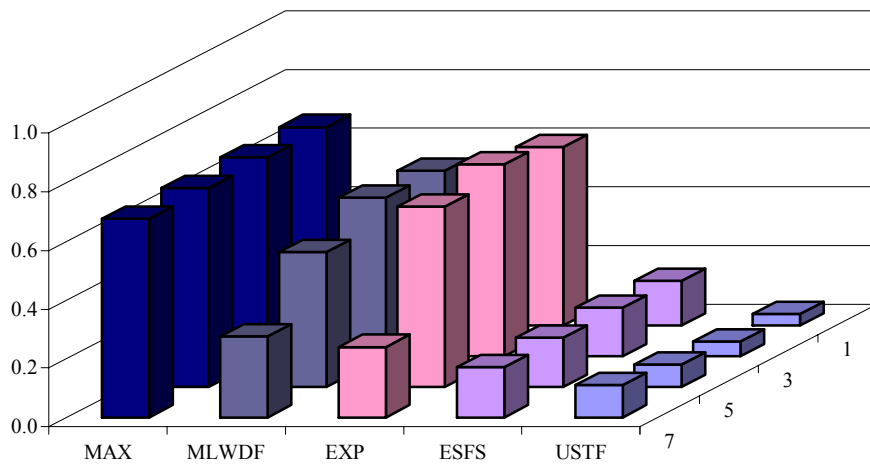
Από αυτά τα σχήματα φαίνεται πως οι αλγόριθμοι M-LWDF και EXP πετυχαίνουν πολύ μικρή μέση ρυθμαπόδοση για τους RT χρήστες και αντίστοιχα πολύ μεγάλη για τους NRT χρήστες. Ειδικά στο κατώφλι των 50 χρονοθυρίδων και όσο λιγοστεύουν οι RT χρήστες στο σύστημα, ο ρυθμός απόδοσής τους είναι ανεπίτρεπτα χαμηλός.

Συνεπώς, οι αλγόριθμοι M-LWDF και EXP είναι εντελώς ακατάλληλοι για τη χρονοδρομολόγηση σε ένα πραγματικό σύστημα ενοποιημένων υπηρεσιών, το οποίο φιλοξενεί τόσο υπηρεσίες πραγματικού χρόνου όσο και χαμηλότερων απαιτήσεων υπηρεσίες, καθώς δεν τηρούν καμιά από τις απαιτήσεις των RT χρηστών για επαρκή ποιότητα της υπηρεσίας τους.

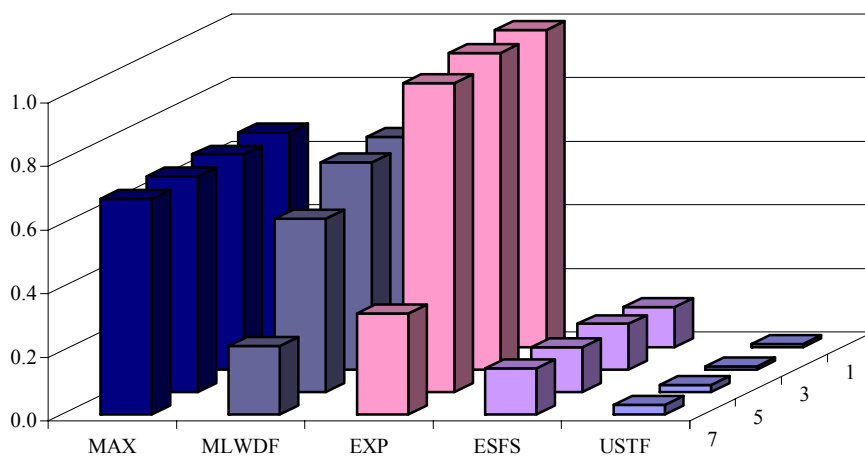
Ο αλγόριθμος ESFS σύμφωνα με τα σχήματα 6.8 και 6.9 ικανοποιεί πλήρως τις απαιτήσεις μικρής καθυστέρησης εξυπηρέτησης. Αυτό οφείλεται στην ενδογενή μορφή του αλγορίθμου, ο οποίος φροντίζει να κάνει δίκαιη διανομή των χρονοθυρίδων ανά χρήστη εντός του παραθύρου παρατήρησης, δηλαδή ισοκατανέμει τις χρονοθυρίδες ενός παραθύρου παρατήρησης στους χρήστες για εξυπηρέτηση. Παράλληλα, αυτό οδηγεί και σε αρκετά αυστηρότερα κριτήρια καθυστέρησης. Ακόμα, σύμφωνα με τα σχήματα 6.10 και 6.11 ο ESFS δεν κατορθώνει παρ' όλα αυτά να εκπληρώσει τις προσδοκίες ενός χρήστη αυστηρής RT υπηρεσίας σε βραχυπρόθεσμη απόδοση, καθώς διακρίνονται αρκετά υψηλές πιθανότητες αστοχίας ως προς τη συναφή απαίτηση (γύρω στο 0.15).

Το γεγονός ότι οι τιμές αυτές είναι αρκετά μικρότερες από τις αντίστοιχες τιμές για τους αλγόριθμους M-LWDF και EXP δικαιολογείται μέσα από τα σχήματα 6.12 έως 6.14 και 6.18 έως 6.23. Σε αυτά παρατηρείται ότι η μέση ρυθμαπόδοση ενός RT χρήστη είναι αρκετά υψηλότερη σε σχέση με αυτήν που πετυχαίνει ο USTF και μάλιστα για την περίπτωση ενός μόνο RT χρήστη στο σύστημα ξεπερνάει και τη μέση ρυθμαπόδοση του MAX. Για την αποφυγή σύγχυσης, η μέση ρυθμαπόδοση του συστήματος συνολικά λαμβάνει την υψηλότερη τιμή στην περίπτωση του MAX αλγορίθμου. Επιστρέφοντας στην ανάλυση του ESFS, οι χρήστες NRT πετυχαίνουν πολύ μικρό μέσο ρυθμό μετάδοσης, κάτι που οπτικοποιείται καλύτερα στα σχήματα 6.18 έως και 6.23. Άρα, λοιπόν, ο αλγόριθμος ESFS καταλήγει σε σχεδόν αποκλεισμό

των χρηστών υπηρεσιών μη πραγματικού χρόνου και επίσης δεν ενδείκνυται για τη δρομολόγηση των χρηστών σε ένα σύστημα και με τις δύο κλάσεις χρηστών.



Σχήμα 6. 10: Μέση πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης ανά αλγόριθμο συναρτήσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης $W_i = 20$ χρονοθυρίδων.



Σχήμα 6. 11: Μέση πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης ανά αλγόριθμο συναρτήσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης $W_i = 50$ χρονοθυρίδων.

Τέλος, ο USTF ικανοποιεί πλήρως τις απαιτήσεις για χαμηλή καθυστέρηση εξυπηρέτησης των RT χρηστών, πετυχαίνοντας τις χαμηλότερες πιθανότητες

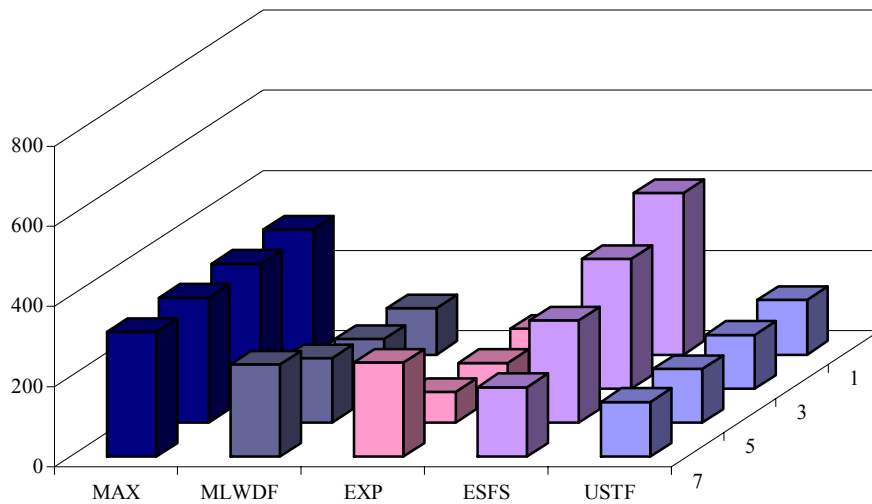
αστοχίας και διατηρώντας τις μηδενικές. Επίσης, τα σχήματα 6.10 και 6.11 δείχνουν ότι εκπληρώνει και τις απαιτήσεις τους για ελάχιστη ποσότητα δεδομένων και κατ' επέκταση ρυθμό μετάδοσης ανά χρονικό παράθυρο παρατήρησης και μάλιστα τόσο καλύτερα όσο μικρότερος είναι ο αριθμός των RT χρηστών που βρίσκονται στο σύστημα. Αυτό οφείλεται στην απαίτηση του αλγορίθμου για κατάλληλη κατανομή της ποσότητας δεδομένων εξυπηρέτησης στο χρόνο για όλους τους RT χρήστες. Συνεπώς, αντεπεξέρχεται στις απαιτήσεις των υπηρεσιών πραγματικού χρόνου, όταν βρίσκονται και NRT χρήστες στο σύστημα ταυτόχρονα. Με απλά λόγια, τα πειραματικά αποτελέσματα επιβεβαιώνουν τις σχεδιαστικές επιλογές της προτεινόμενης πολιτικής χρονισμού των χρηστών και κυρίως εκείνης που αφορά την απεικόνιση στο σύστημα όχι μόνο του ποσού της πληροφορίας που έχει λάβει ο εκάστοτε χρήστης εντός μικρών χρονικών παραθύρων αλλά ταυτόχρονα και της επί μέρους κατανομής της ανά χρονοσχιμή.

Όσον αφορά τώρα τους NRT χρήστες, τα σχήματα 6.14 και 6.15 δείχνουν ότι πετυχαίνουν ικανοποιητική ρυθμαπόδοση και παρατηρώντας τα σχήματα 6.18 έως 6.23 διαπιστώνεται ότι είναι και μεγαλύτερη από αυτή των RT χρηστών. Ουσιαστικά, λοιπόν, ο USTF φροντίζει να δώσει στους RT χρήστες τους απαραίτητους πόρους για να ικανοποιήσουν τις απαιτήσεις καθυστέρησης και βραχυπρόθεσμης ρυθμαπόδοσης και στη συνέχεια κατανέμει τους διαθέσιμους πόρους του συστήματος στους NRT χρήστες, ώστε να μεγιστοποιηθεί η μέση ρυθμαπόδοσή τους και κατ' επέκταση η συνολική του συστήματος, αλλά και σε εκείνους τους RT χρήστες, των οποίων οι βραχυπρόθεσμες απαιτήσεις δεδομένων δεν ικανοποιούνται.

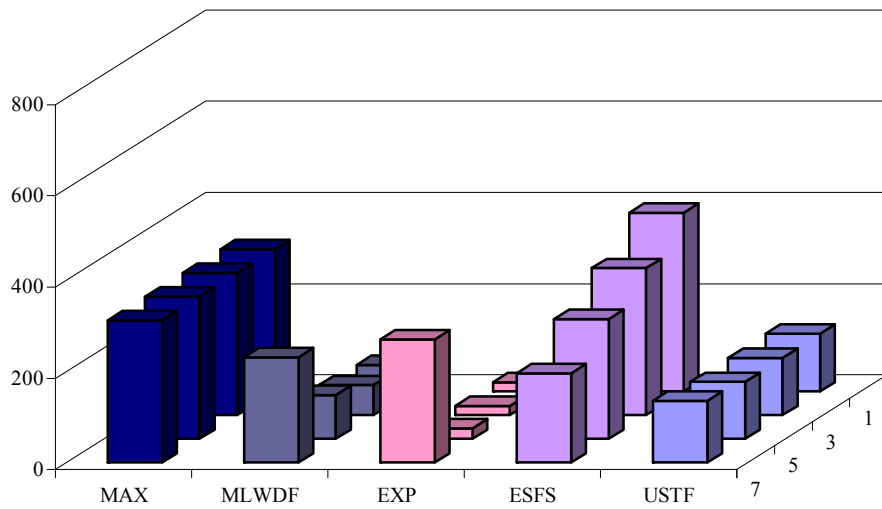
Με άλλα λόγια ο USTF δίνει προτεραιότητα στις απαιτήσεις των RT χρηστών, χωρίς να εξοστρακίζει τους NRT χρήστες, ενώ παράλληλα αυξάνει κατά το δυνατόν την αξιοποίηση των πόρων του δικτύου. Ικανοποιεί τις απαιτήσεις των RT χρηστών, αλλά αφού το κατορθώσει δίνει προτεραιότητα στους NRT χρήστες, αντί να συνεχίζει να εξυπηρετεί τους RT χρήστες.

Επίσης, από τα σχήματα 6.16 και 6.17 προκύπτει ότι η μέση ρυθμαπόδοση του συστήματος δεν είναι πολύ μικρότερη από αυτήν των υπολοίπων αλγορίθμων, που σημαίνει ότι διατηρεί σε ικανοποιητικά επίπεδα τη μέση ρυθμαπόδοση του συστήματος, κάνοντας απλά καλύτερη διανομή των πόρων και πιο δίκαιη μεταξύ RT και NRT χρηστών. Η μέση ρυθμαπόδοση του συστήματος για τον USTF αυξάνεται, όταν αυξάνεται το πλήθος των NRT χρηστών, που σημαίνει ότι είναι ευκολότερο να

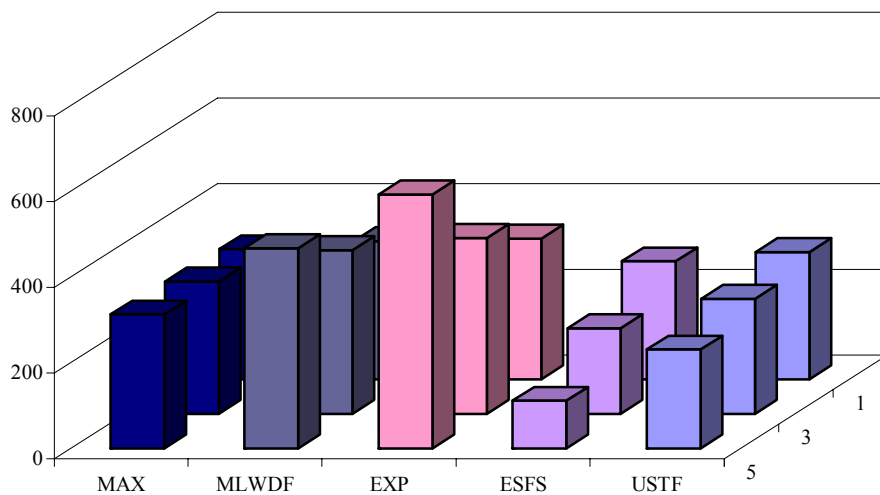
διαχειριστεί τις απαιτήσεις των λιγότερων RT χρηστών και να αναθέσει τους λοιπούς πόρους στους NRT χρήστες με στόχο τη μεγιστοποίηση της συνολικής ρυθμαπόδοσης. Με άλλα λόγια, με μικρότερο πλήθος RT χρηστών είναι μικρότερος ο αριθμός των χρηστών με τη μέση σταθερή μικρότερη ρυθμαπόδοση και μεγαλύτερος ο αριθμός των χρηστών με μεγαλύτερη ρυθμαπόδοση.



Σχήμα 6. 12: Μέση ρυθμαπόδοση σε (Kbps) των RT χρηστών για κατώφλι καθυστέρησης $T_i = 20$ χρονοθυρίδες συναρτήσει του πλήθους τους.

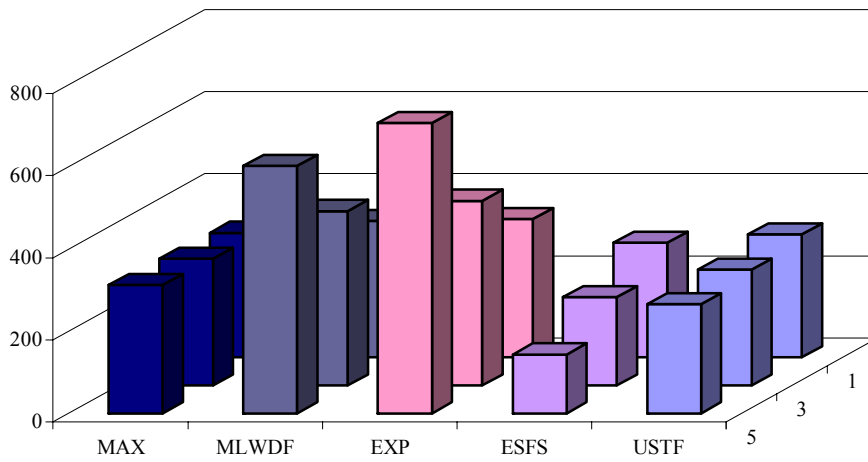


Σχήμα 6. 13: Μέση ρυθμαπόδοση σε (Kbps) των RT χρηστών για κατώφλι $T_i = 50$ χρονοθυρίδες συναρτήσει του πλήθους τους.

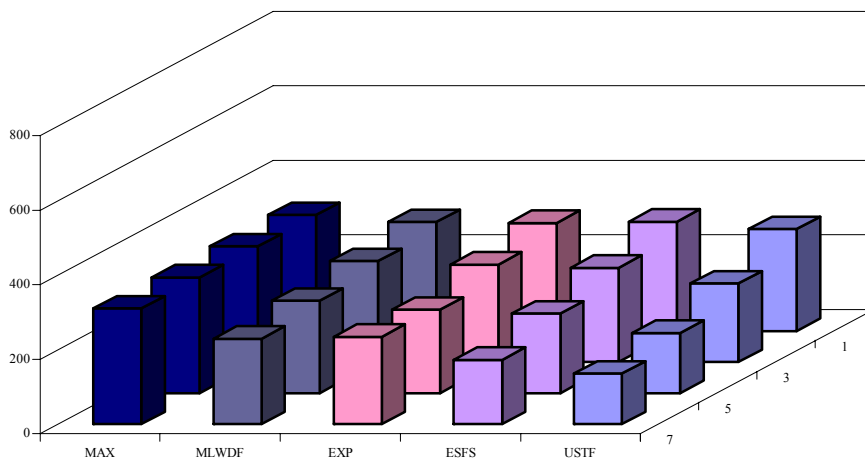


Σχήμα 6. 14: Η μέση ρυθμαπόδοση σε (Kbps) των NRT χρηστών του συστήματος συναρτήσει του αριθμού των RT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες.

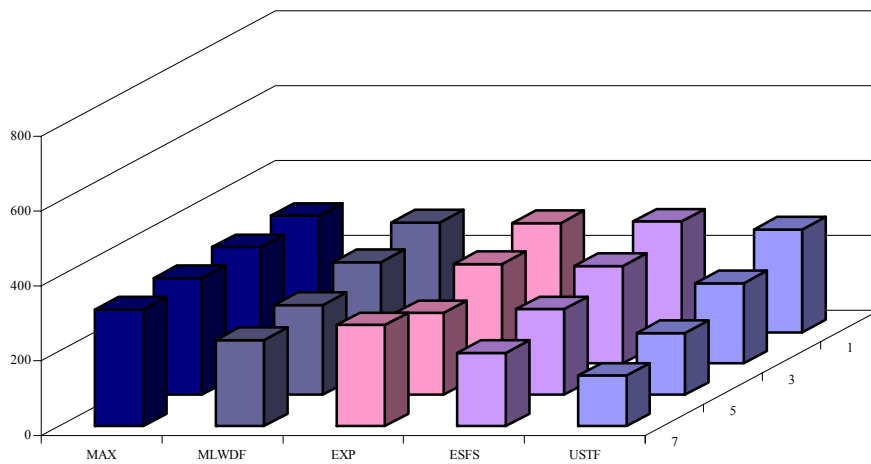
Μέση Ρυθμαπόδοση των NRT χρηστών ($T_i = 50$ χρονοθυρίδες)



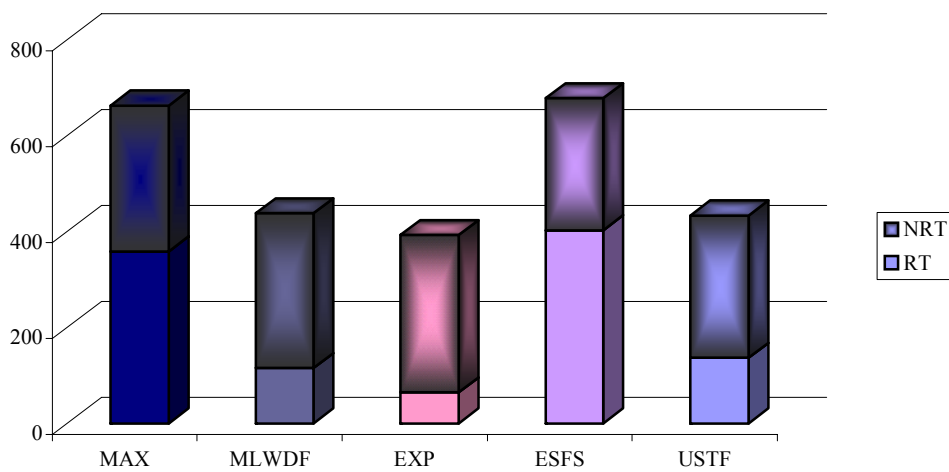
Σχήμα 6. 15: Η μέση ρυθμαπόδοση σε (Kbps) των NRT χρηστών του συστήματος συναρτήσει του πλήθους των RT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες.



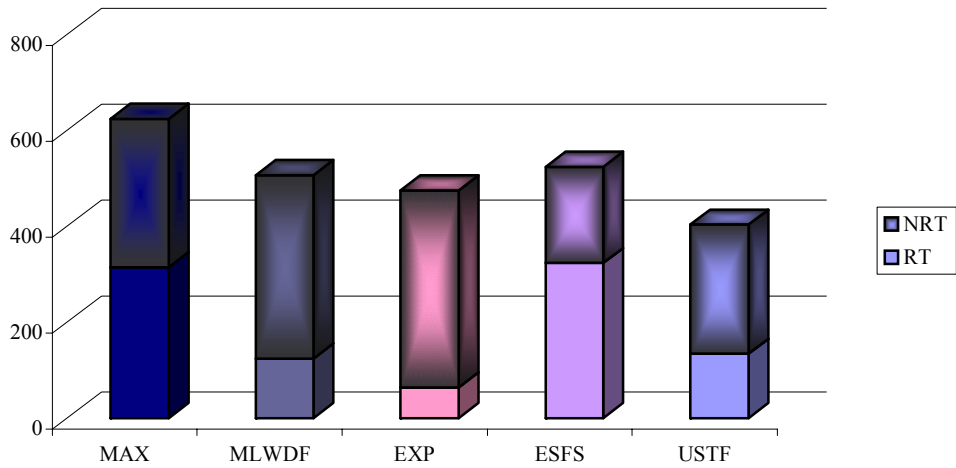
Σχήμα 6. 16: : Η μέση ρυθμαπόδοση σε (Kbps) του συστήματος συναρτήσει του αριθμού των RT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες.



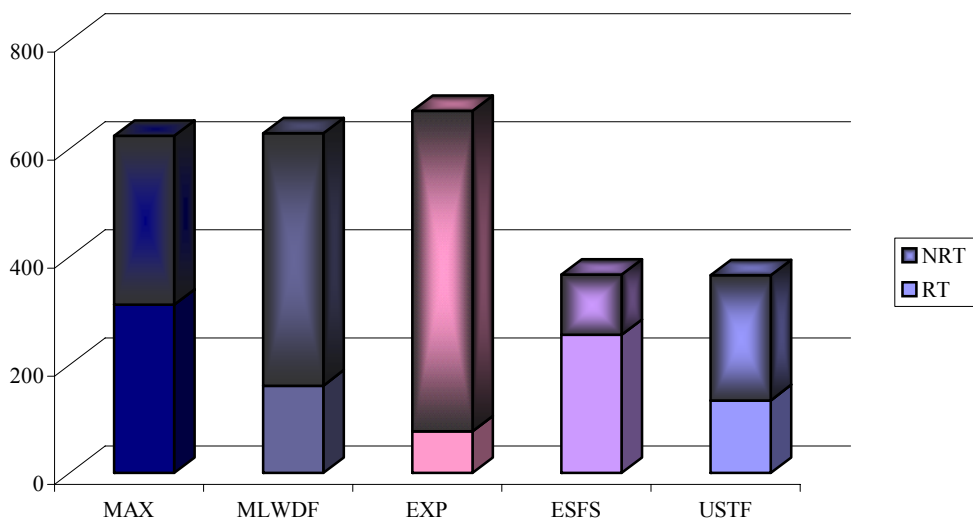
Σχήμα 6. 17: Η μέση ρυθμαπόδοση σε (Kbps) του συστήματος συναρτῆσει του αριθμού των RT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες.



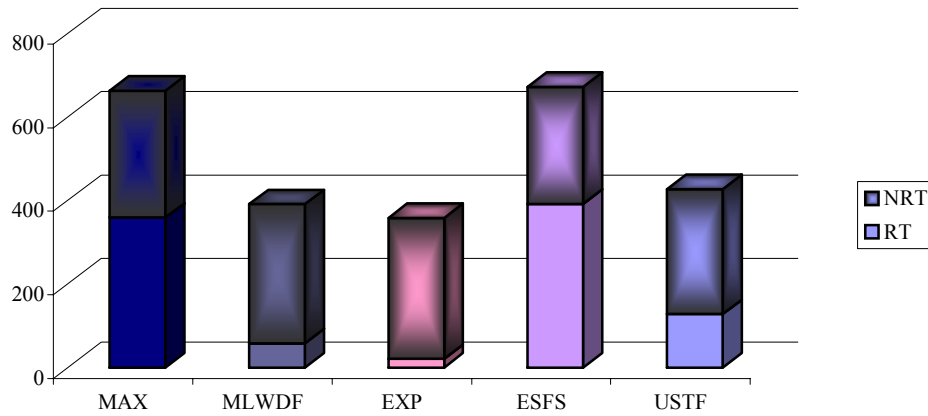
Σχήμα 6. 18: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες και 1 RT χρήστη.



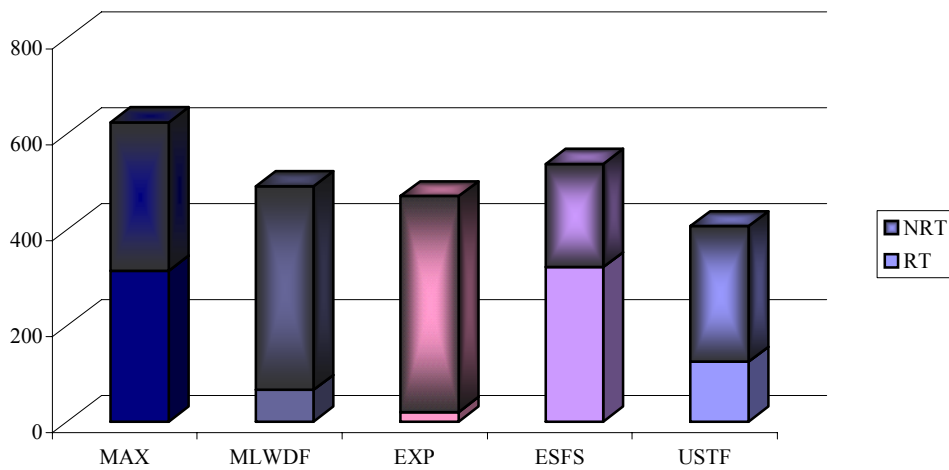
Σχήμα 6. 19: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες και 3 RT χρήστες.



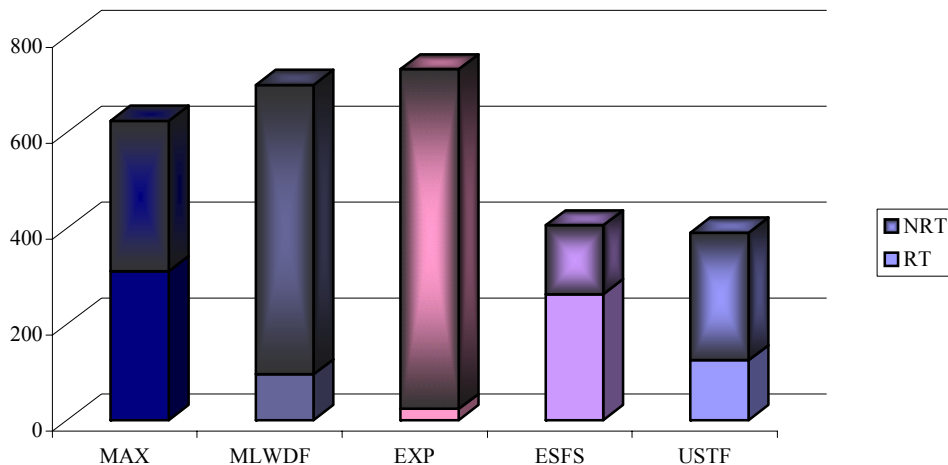
Σχήμα 6. 20: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 20$ χρονοθυρίδες και 5 RT χρήστες.



Σχήμα 6. 21: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες και 1 RT χρήστη.



Σχήμα 6. 22: Η μέση ρυθμαπόδοση σε (Kbps) των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες και 3 RT χρήστες.



Σχήμα 6. 23: Η μέση ρυθμαπόδοση των RT και NRT χρηστών ανά αλγόριθμο για κατώφλι $T_i = 50$ χρονοθυρίδες και 5 RT χρήστες.

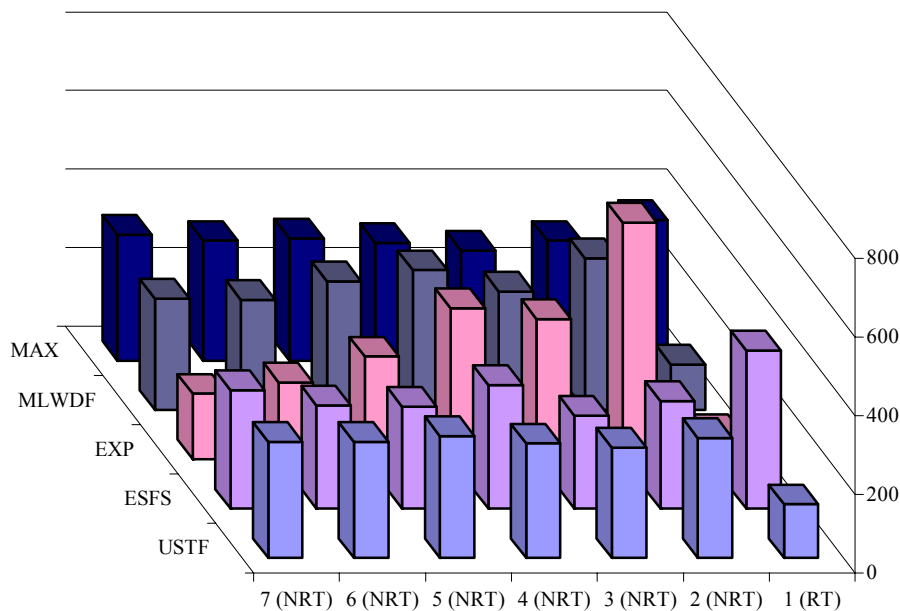
Για την πληρότητα των παραπάνω συμπερασμάτων σχετικά με τη μέση ρυθμαπόδοση των χρηστών παρουσιάζεται στα σχήματα 6.24 έως 6.26 η μέση απόδοση κάθε χρήστη του συστήματος για κάθε αλγόριθμο, όταν το κατώφλι καθυστέρησης και το παράθυρο παρατήρησης είναι 20 χρονοθυρίδες για τις περιπτώσεις που στο σύστημα υπάρχουν ένας, τρεις και πέντε RT χρήστες αντίστοιχα. Από τα σχήματα προκύπτει κατ' αρχάς ότι ο MAX αλγόριθμος συμπεριφέρεται σχεδόν το ίδιο σε όλους τους χρήστες, είτε αυτοί είναι χρήστες RT είτε NRT. Οι τυχόν διαφοροποιήσεις οφείλονται προφανώς στις εκάστοτε στιγμιαίες διαφοροποιήσεις της κατάστασης του διαύλου των αντίστοιχων χρηστών.

Ακόμη, όσον αφορά τους αλγόριθμους που είναι ευαίσθητοι σε καθυστέρηση, παρατηρείται ότι συμπεριφέρονται κατά τον ίδιο τρόπο σε σχέση με το ρυθμό μετάδοσης που πετυχαίνουν σε όλους τους RT χρήστες, ενώ ο ρυθμός μετάδοσης των NRT χρηστών διαφοροποιείται έντονα. Αυτή η παρατήρηση είναι σημαντική, αν αναλογιστεί κανείς ότι όλοι οι χρήστες έχουν την ίδια μέση τιμή SINR που χαρακτηρίζει το δίαυλο της επικοινωνίας τους. Συνεπώς, ανάμεσα στους NRT χρήστες εξυπηρετείται κάθε φορά εκείνος με τον καλύτερο χρονικά δίαυλο, χωρίς να τηρούνται κριτήρια δικαιοσύνης μεταξύ τους.

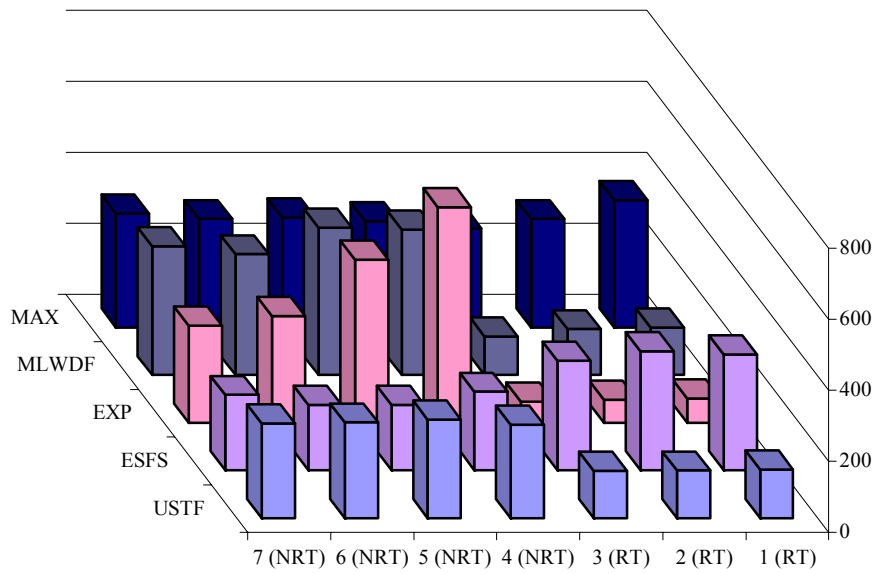
Ο αλγόριθμος ESFS, από την άλλη, χειρίζεται σχεδόν ισότιμα τους NRT χρήστες από πλευράς ρυθμού μετάδοσης, αλλά παρουσιάζει μικρές διακυμάνσεις ανά τους NRT χρήστες, όταν έχουν αριθμητική υπεροχή στο σύστημα, οι οποίες αποδίδονται στις στιγμιαίες διακυμάνσεις των διαύλων των χρηστών.

Τέλος, ο αλγόριθμος USTF παρουσιάζει πλήρως δίκαιη κατανομή του μέσου ρυθμού μετάδοσης ανά χρήστη σε κάθε κλάση χρηστών. Μάλιστα αυτό το πετυχαίνει για όλους τους λόγους N_{RT} / N_{NRT} που δοκιμάζονται. Τονίζεται ότι το συγκεκριμένο αποτέλεσμα αντεπεξέρχεται στις προσδοκίες του διαχειριστή, καθώς οι χρήστες έχουν την ίδια μέση κατάσταση διαύλου, άρα αναμένεται και σχεδόν ίδιος ρυθμός μετάδοσης για όλους τους χρήστες. Ο αλγόριθμος USTF θα μπορούσε να χαρακτηριστεί προσανατολισμένος στο ρυθμό μετάδοσης, αφού πληροί όλα τα κριτήρια υπηρεσίας που σχετίζονται με αυτόν για RT ή NRT χρήστες.

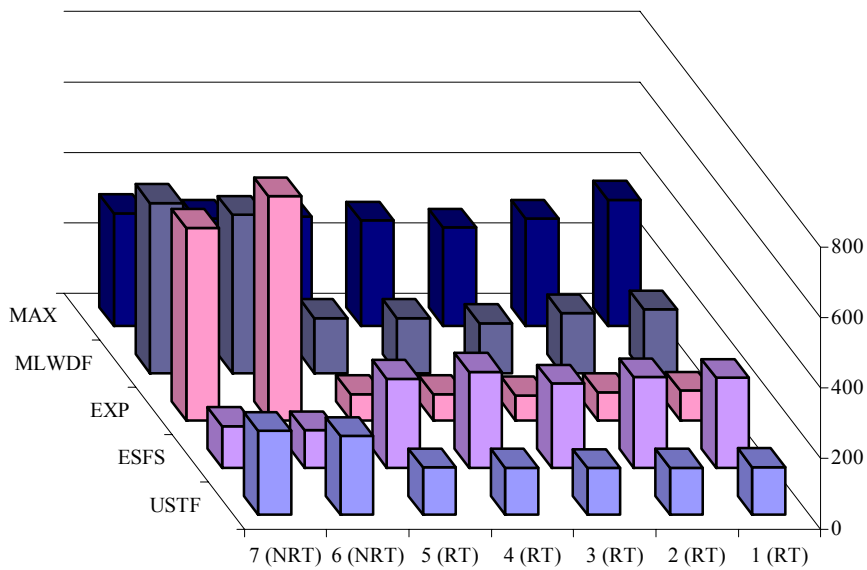
Μάλιστα φαίνεται και εδώ ότι ο αλγόριθμος USTF εκχωρεί τους απαιτούμενους πόρους του συστήματος στους RT χρήστες για την εξυπηρέτηση των αναγκών τους και αναθέτει τους υπόλοιπους διαθέσιμους σε NRT υπηρεσίες. Γι' αυτό και η ρυθμαπόδοση των NRT χρηστών είναι υψηλότερη από αυτή των RT. Επίσης, οι MLWDF και EXP στοχεύουν σε αύξηση της μέσης ρυθμαπόδοσης του συστήματος, χωρίς να καταφέρνουν να ικανοποιήσουν τους RT χρήστες, ενώ ο ESFS υπερευνοεί τους RT χρήστες, χωρίς και πάλι να καταφέρνει να εκπληρώσει όλες τις προσδοκίες τους.



Σχήμα 6. 24: Η απόδοση (σε Kbps) ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης και παράθυρο παρατήρησης 20 χρονοθυρίδων και για σύστημα με 1 RT χρήστη.



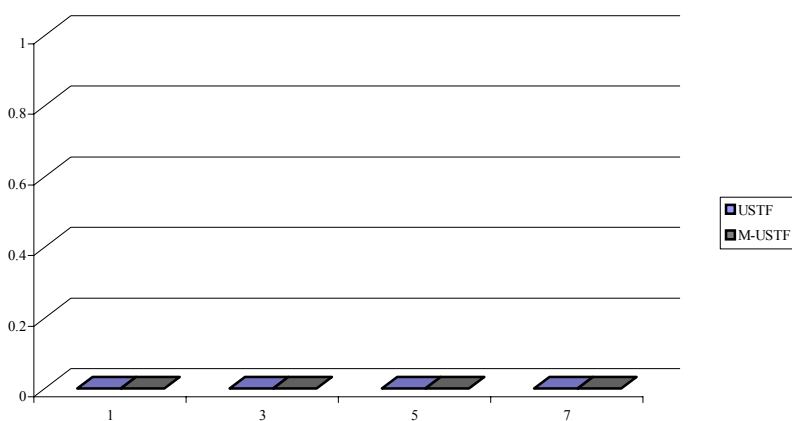
Σχήμα 6. 25: Η απόδοση (σε Kbps) ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης και παράθυρο παρατήρησης 20 χρονοθυρίδων και για σύστημα με 3 RT χρήστες.



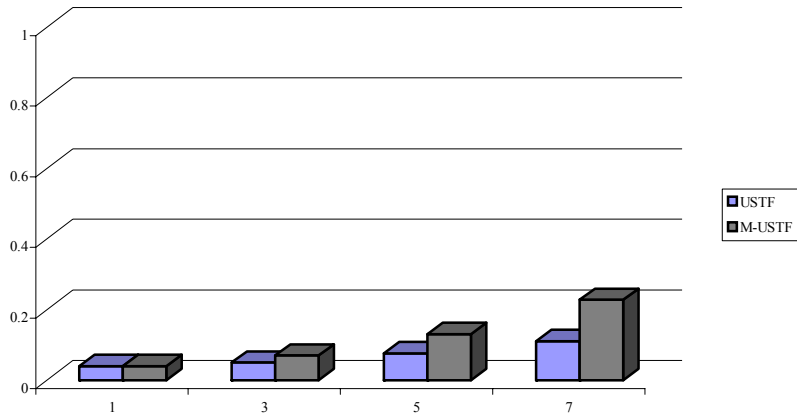
Σχήμα 6. 26: Η απόδοση (σε Kbps) ανά χρήστη και ανά αλγόριθμο για κατώφλι καθυστέρησης και παράθυρο παρατήρησης 20 χρονοθυρίδων και για σύστημα με 5 RT χρήστες.

6.4 Η Σημασία της «Τιμωρίας» στον Αλγόριθμο USTF

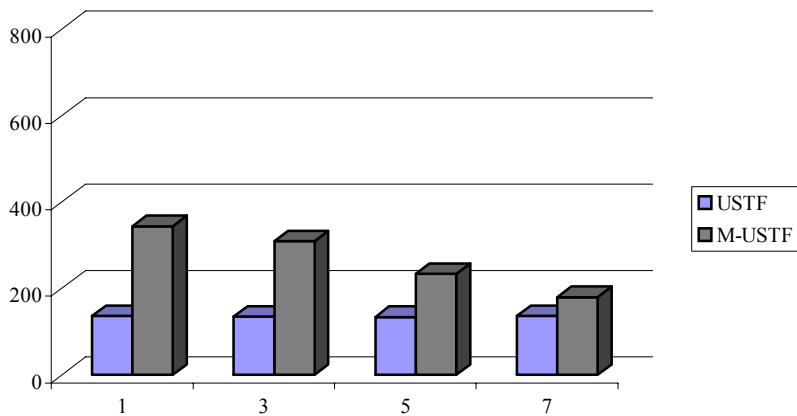
Στη συνέχεια ο αλγόριθμος USTF συγκρίνεται με μια τροποποιημένη μορφή του, που ονομάζουμε Τροποποιημένο – USTF (Modified USTF – M-USTF). Ο M-USTF διαφέρει από τον USTF στο γεγονός ότι δεν «τιμωρεί» τους χρήστες που έχουν ικανοποιήσει ήδη τις απαιτήσεις τους σε βραχυπρόθεσμη ρυθμαπόδοση, οπότε δεν υφίσταται η παράμετρος $G_i(k)$ και η σχέση ορισμού της περιλαμβάνει μόνο την άνω συνθήκη της σχέσης (5.3). Το σενάριο που θεωρήθηκε συμπίπτει το σενάριο Σ2. Στόχος της συγκεκριμένης ενότητας είναι να εξηγηθεί ο λόγος που ο αλγόριθμος USTF κατορθώνει να τηρήσει τις απαιτήσεις των RT χρηστών, χωρίς να αποκλείει την εξυπηρέτηση των NRT χρηστών.



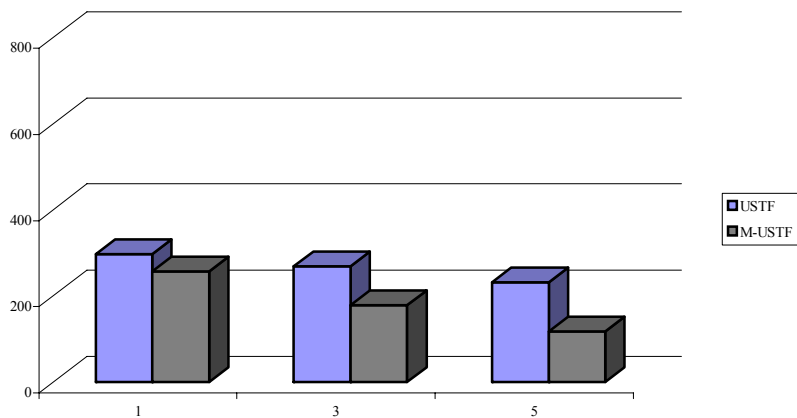
Σχήμα 6. 27: Η πιθανότητα αστοχίας καθυστέρησης εξυπηρέτησης συναρτήσει του πλήθους των RT χρηστών για κατώφλι καθυστέρησης T_i 20 χρονοθυρίδων.



Σχήμα 6. 28: Η πιθανότητα αστοχίας βραχυπρόθεσμης απόδοσης συναρτῆσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης W_i 20 χρονοθυρίδων.

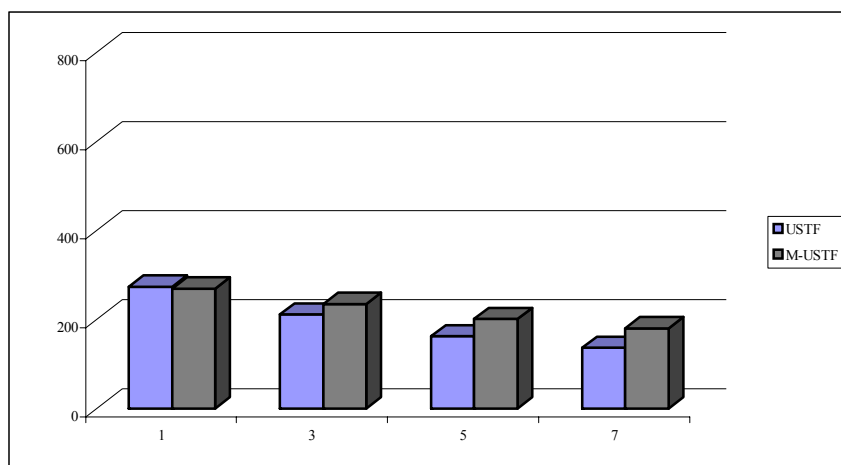


Σχήμα 6. 29: Η μέση ρυθμαπόδοση (σε Kbps) των RT χρηστών συναρτῆσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης W_i 20 χρονοθυρίδων.



Σχήμα 6. 30: Η μέση ρυθμαπόδοση (σε Kbps) των NRT χρηστών συναρτῆσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης W_i 20 χρονοθυρίδων.

Το σχήμα 6.27 φανερώνει ότι και οι δύο μορφές του USTF τηρούν τις απαιτήσεις καθυστέρησης για κάθε πλήθος RT χρηστών στο σύστημα. Το σχήμα 6.28 όμως δείχνει ότι δεν καλύπτονται οι απαιτήσεις σε βραχυπρόθεσμη απόδοση, καθώς γενικά ο M-USTF δίνει μεγαλύτερες πιθανότητες αστοχίας. Τα σχήματα 6.29 και 6.30 δείχνουν ότι οι RT χρήστες πετυχαίνουν μεγαλύτερη μέση ρυθμαπόδοση στον M-USTF σε σχέση με τον USTF και αντίστοιχα οι NRT χρήστες υπό τον M-USTF μικρότερη σε σχέση με τον USTF. Τα σχήματα 6.29 και 6.30 που δείχνουν τη μέση ρυθμαπόδοση σε συνδυασμό με το σχήμα 6.28 καταλήγουν στο συμπέρασμα ότι ο M-USTF δεν κατανέμει κατάλληλα τους πόρους στους χρήστες του συστήματος, ώστε να ικανοποιούνται οι απαιτήσεις ποιότητας μιας υπηρεσίας πραγματικού χρόνου, παρ' ότι τηρεί τις απαιτήσεις καθυστέρησης και εκχωρεί αρκετούς πόρους στους RT χρήστες. Συνεπώς, δε διαφέρει σημαντικά από τους αλγορίθμους που έλαβαν μέρος στη μελέτη της ενότητας 6.3.



Σχήμα 6. 31: Η μέση ρυθμαπόδοση (σε Kbps) του συστήματος συναρτήσει του πλήθους των RT χρηστών για παράθυρο παρατήρησης $W; 20$ χρονοθυρίδων.

Τώρα μπορεί να εξηγηθεί ο λόγος για τον οποίο ικανοποιεί τα κριτήρια ποιότητας ο αλγόριθμος USTF, αν αναλογιστούμε ότι η μόνη διαφορά με τον M-USTF είναι ο όρος $G_i(k)$. Ανατρέχοντας στο κεφάλαιο 5, ανακαλείται πως ο συγκεκριμένος όρος φανερώνει τη χαμηλή ανάγκη ενός RT χρήστη να εξυπηρετηθεί, δεδομένου ότι έχει λάβει αρκετή ποσότητα δεδομένων κατά τις προηγούμενες χρονοθυρίδες, ώστε να ικανοποιούνται οι απαιτήσεις της υπηρεσίας του κατά το

τελευταίο χρονικό διάστημα διάρκειας W_i χρονοθυρίδων. Ο όρος αυτός φροντίζει για τη διατήρηση σταθερού ρυθμού μετάδοσης σε όλους τους RT χρήστες και αποτελεί τη σημαντικότερη καινοτομία του USTF αλγορίθμου.

Θα μπορούσε να χαρακτηριστεί ως όρος «τιμωρίας» για αυτούς τους χρήστες, ο οποίος μειώνει την προτεραιότητα εξυπηρέτησής των, επιτρέποντας σε χρήστες που έχουν εξυπηρετηθεί ελλιπώς ή λιγότερο κατά τις τελευταίες W_i χρονοθυρίδες να εξυπηρετηθούν, ακόμα και αν έχουν ελαφρώς δυσμενέστερες συνθήκες διαύλου.

Αυτό φαίνεται από το σχήμα 6.31, όπου η μέση ρυθμαπόδοση του συστήματος υπό τον USTF αλγόριθμο είναι ελαφρώς μικρότερη από αυτή του M-USTF. Η εξυπηρέτηση των RT χρηστών που χαρακτηρίζονται από μικρή ποσότητα δεδομένων εξυπηρέτησης εντός του διαστήματος W_i χρονοθυρίδων μειώνει τη ρυθμαπόδοση του συστήματος. Το γεγονός όμως ότι η διαφορά της ρυθμαπόδοσης του συστήματος από ένα αλγόριθμο που δίνει μεγαλύτερη βαρύτητα στις συνθήκες του διαύλου των RT χρηστών είναι μικρή οφείλεται στο γεγονός ότι αφού τηρηθούν τα κριτήρια ποιότητας για τους RT χρήστες, παίρνουν σειρά οι NRT, με σκοπό τη μεγιστοποίηση της συνολικής μέσης ρυθμαπόδοσης.

Συμπερασματικά, μέσω της σύγκρισης των αλγορίθμων MAX, M-LWDF, EXP και ESFS αποδεικνύεται πως ο καλύτερος από πλευράς μέσης ρυθμαπόδοσης αλγόριθμος δεν είναι ο καλύτερος αλγόριθμος εξυπηρέτησης γενικά, όταν τεθούν και απαιτήσεις υπηρεσιών πραγματικού χρόνου. Επίσης, η τήρηση μιας απαίτησης μιας τέτοιας υπηρεσίας δε συνεπάγεται και την τήρηση όλων των απαιτήσεων αυτής. Έτσι, τα παραπάνω σχήματα και παρατηρήσεις αποδεικνύουν και πειραματικά το συμπέρασμα του κεφαλαίου 4 περί αδυναμίας του κριτηρίου καθυστέρησης εξυπηρέτησης να εκπληρώσει όλες τις απαιτήσεις μιας υπηρεσίας πραγματικού χρόνου. Συγκεκριμένα, οι αλγόριθμοι M-LWDF, EXP και ESFS που δημιουργήθηκαν με σκοπό να τηρήσουν μόνο το συγκεκριμένο κριτήριο δεν απέδωσαν καρπούς στο θέμα της απαίτησης για ελάχιστο ρυθμό μετάδοσης. Συνεπώς, ήταν αναγκαία η άφιξη ενός νέου αλγορίθμου, όπως ο USTF, ο οποίος θα πληρούσε και το κριτήριο της βραχυπρόθεσμης ρυθμαπόδοσης, ώστε να δέχονται έγκαιρα όλοι οι χρήστες υπηρεσιών πραγματικού χρόνου τα δεδομένα που τους αντιστοιχούν για την πλήρη κάλυψη των αναγκών τους και ο οποίος σύμφωνα με τα αποτελέσματα αυτού του κεφαλαίου αντεπεξέρχεται πλήρως στις ανωτέρω απαιτήσεις.

7

ΣΥΜΠΕΡΑΣΜΑΤΑ

Στη συγκεκριμένη διπλωματική εργασία θεωρήθηκε και αναλύθηκε το πρόβλημα της διαχείρισης των πόρων ενός συστήματος CDMA/HDR στη ζεύξη καθόδου και της αποδοτικής κατανομής τους στους χρήστες που περιλαμβάνει. Δεδομένου ότι ένα τέτοιο σύστημα προσανατολίζεται να υποστηρίζει πλέον και υπηρεσίες πραγματικού χρόνου, μελετήθηκαν συγκεκριμένα κριτήρια που εκφράζουν την ποιότητα της υπηρεσίας αυτών. Η αποτελεσματική τους εξυπηρέτηση υπαγορεύει τη μελέτη των συγκεκριμένων κριτηρίων εντός σύντομων χρονικών διαστημάτων όσον αφορά την καθυστέρηση εξυπηρέτησης, καθώς και την ποσότητα των δεδομένων που λαμβάνονται βραχυπρόθεσμα σύμφωνα με τους ρυθμούς μετάδοσης που επιτυγχάνονται. Συγκεκριμένα, η αποδοτική κατανομή των πόρων εντός μικρών χρονικών διαστημάτων είναι καίρια για την ικανοποιητική εξυπηρέτηση των χρηστών υπηρεσιών πραγματικού χρόνου, αλλά και θεμελιώδες ζήτημα όσον αφορά τη βελτιστοποίηση της συνολικής ρυθμαπόδοσης του συστήματος.

Έτσι, παρουσιάστηκαν και αναλύθηκαν βασικοί υπάρχοντες αλγόριθμοι δρομολόγησης χρηστών υπηρεσιών μη πραγματικού χρόνου αρχικά. Αφού σημειώθηκε ότι αυτοί δεν καλύπτουν τις ανάγκες των υπηρεσιών πραγματικού χρόνου, παρουσιάστηκαν διάφοροι αλγόριθμοι ευαίσθητοι στην καθυστέρηση της

εξυπηρέτησης των χρηστών υπηρεσίας πραγματικού χρόνου, οι οποίοι διατείνονται πως εξυπηρετούν αποδοτικά τις υπηρεσίες πραγματικού χρόνου. Μέσω, όμως, αναλυτικής παρουσίασης και μελέτης θεμελιωδών αλγορίθμων, οι οποίοι τηρώντας τις απαιτήσεις της καθυστέρησης εξυπηρέτησης και επιτυγχάνοντας υψηλούς μέσους ρυθμούς μετάδοσης, φάνηκαν αδύναμοι να τηρήσουν τις απαιτήσεις βραχυπρόθεσμης ρυθμαπόδοσης, αποδείχτηκε πως το κριτήριο καθυστέρησης είναι ανεπαρκές ως προς την έκφραση των απαιτήσεων της ποιότητας μιας υπηρεσίας πραγματικού χρόνου.

Βάσει της αίσθησης ότι το κριτήριο βραχυπρόθεσμης ρυθμαπόδοσης παρέχει επίσης και χαμηλή καθυστέρηση εξυπηρέτησης στους χρήστες υπηρεσίας πραγματικού χρόνου, παρουσιάστηκε μια νέα πολιτική δρομολόγησης, η οποία εστιάζει στο συγκεκριμένο κριτήριο, εισάγοντάς το στη συνάρτηση χρησιμοποίησής της. Η συγκεκριμένη πολιτική στοχεύει στην ελαχιστοποίηση των πιθανοτήτων αστοχίας βραχυπρόθεσμης ρυθμαπόδοσης για όλους τους χρήστες υπηρεσίας πραγματικού χρόνου, ενώ παράλληλα δεν αποκλείει τους χρήστες υπηρεσιών μη πραγματικού χρόνου, οι οποίοι, πάραυτα, δεν υπόκεινται στις αυστηρές απαιτήσεις της έτερης κλάσης υπηρεσιών, αλλά εξυπηρετούνται με στόχο τη μεγιστοποίηση της ρυθμαπόδοσης του συστήματος.

Συνεπώς, η μελέτη καταλήγει στην προσομοίωση ορισμένων από τους ευαίσθητους στην καθυστέρηση εξυπηρέτησης αλγορίθμους, από την οποία προκύπτουν συγκριτικά διαγράμματα της απόδοσής των. Από αυτά, επαληθεύεται πως ο νέος αλγόριθμος όντως διατηρεί χαμηλές πιθανότητες αστοχίας βραχυπρόθεσμης ρυθμαπόδοσης και καθυστέρησης εξυπηρέτησης για τους χρήστες υπηρεσίας πραγματικού χρόνου έναντι των προηγούμενων, ενώ εξυπηρετεί αποδοτικά και τις υπηρεσίες μη πραγματικού χρόνου, οπότε και εστιάζει στη μεγιστοποίηση της ρυθμαπόδοσης του συστήματος.

Η περαιτέρω μελέτη του αλγορίθμου σε πολυπλοκότερα συστήματα με περισσότερες κλάσεις χρηστών και υπηρεσιών και η μετέπειτα τυχόν βελτιστοποίησή του είναι ζητήματα ιδιαίτερα ενδιαφέροντα που καλύπτουν μέρος της παρούσας ερευνητικής προσπάθειας.

8

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and S. Viterbi, "CDMA/HDR: A Bandwidth Efficient High Speed Wireless Data Service for Nomadic Users," *Communications Magazine, IEEE* , vol.38, no.7, pp.70-77, Jul 2000.
- [2] A. Jalali, R. Padovani, and R. Pankaj, "Data Throughput of CDMAHDR A High Efficiency-High Data Rate Personal Communication Wireless System," *in Proc. VTC*, pp. 1854-1858, 2000.
- [3] Xin Liu, Edwin K.P. Chong and Ness B. Shroff, "Opportunistic Transmission Scheduling with Resource-Sharing Constraints in Wireless Networks," *Journal on Selected Areas in Communications*, vol.19, no 10, pp.2053-2064, October 2001.
- [4] Xin Liu, Edwin K.P. Chong and Ness B. Shroff, "A Framework for Opportunistic Scheduling in Wireless Networks," *Computer Networks*, vol. 41, pp. 451–474, March 2003.
- [5] Timotheos Kastrinogiannis and Symeon Papavassiliou, "Probabilistic Short-term Delay and Throughput Requirements of Multimedia Services in High Throughput Wireless Networks," *IEEE Sharnoff Symposium*, 2007.
- [6] Qualcomm Inc. website for white papers (1xEVDO), "QCOM 3G Overview," 2004.
- [7] Vivek P. Mhatre, "CDMA 2000: 1xEVDO and 1xEVDV, An Overview," *Talk for EE 647*, 2004.
- [8] Qualcomm Inc. website for white papers (1xEVDO), "1xEVDO Overview," 2004.
- [9] Timotheos Kastrinogiannis and Symeon Papavassiliou, "Satisfying Elastic Short Term Fairness in High Throughput Wireless Communication Systems with Multimedia Services," *in Proc. of IEEE ICC*, June 2007.

- [10] Μ. Θεολόγου, "Δίκτυα Κινητών και Προσωπικών Επικοινωνιών," Πανεπιστημιακές Εκδόσεις Ε.Μ.Π., Αθήνα, 2006.
- [11] T. Kastrinogiannis, D. Soulios, K. Kastrinogiannis, S. Papavassiliou, "A Utility Based Framework for Efficient Downlink Resource Allocation in WCDMA Wireless Networks with Multimedia Services," *IEEE PIMRC*, 2007.
- [12] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing Quality of Service over a Shared Wireless Link," *Communications Magazine, IEEE*, vol.39, no.2pp.150-154, Feb 2001.
- [13] M. Andrews et al., "CDMA Data QoS Scheduling on the Forward Link with Variable Channel Conditions," *Bell Labs Tech. Memo.*, Apr. 2000.
- [14] S. Shakkottai, A. Stolyar, "Scheduling for Multiple Flows Sharing a Time-Varying Channel: The Exponential Rule," *Bell Laboratories Technical Report*, December, 2000.
- [15] S. Shakkottai, A. Stolyar, "Scheduling Algorithms for a Mixture of Real-Time and Non-Real-Time Data in HDR," *Bell Laboratories Technical Report*, August, 2000.
- [16] Timotheos Kastrinogiannis and Symeon Papavassiliou, "Probabilistic Resource Allocation under Opportunistic Scheduling in Wireless Networks with Multimedia Services," *ISCC*, 2006.
- [17] Timotheos Kastrinogiannis and Symeon Papavassiliou, "Utility based Short-term Throughput Driven Scheduling Approach for Efficient Resource Allocation in CDMA Wireless Networks," *Wireless Personal Communications*, 2007.
- [18] H. Wang and N. Moayeri, "Finite-state Markov channel – a useful model for radio communication channels," *IEEE Trans. on Vehic. Tech.*, vol. 44, pp. 163–171, March 1995.
- [19] H. J. Kushner and P. A. Whiting, "Convergence of proportional-fair sharing algorithms under general conditions," *IEEE Trans. on Wireless Commun.*, vol.3, no.4, pp. 1250- 1259, July 2004.
- [20] F. Berggren and R. Jäntti, "Asymptotically Fair Transmission Scheduling Over Fading Channels," *IEEE Trans. on Wireless Commun.*, vol.3, no.1, pp. 326- 336, January 2004.

Ευρετήριο Συντομογραφιών

APAOS	Access Probability based Assignment Opportunistic Scheduling	45
CDMA	Code Division Multiple Access	13
ESFS	Elastic Short-term Fairness based opportunistic Scheduling	49
EXP	Exponential Rule	42
EXP-Q	Exponential Queue Length Rule	42
EXP-W	Exponential Waiting Time Rule	42
EV-DO	EVolution to Data Only	13
HDR	High Data Rate	13
MAX	Maximum Throughput Scheduling Algorithm	25
M-LWDF	Modified Largest Weighted Delay First	39
MP	Minimum Performance Guarantee Scheduling Scheme	35
NRT	Non Real Time	20, 21
PF	Proportional Fairness	26
QoS	Quality of Service	22
RB	Relatively Best	28
RT	Real Time	20, 21
TF	Temporal fairness Scheduling Scheme	30
UF	Utilitarian Fairness Scheduling Scheme	32
USTF	Utility based Short-term Throughput driven Framework	72
M-USTF	Modified USTF	102