



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ Η/Υ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΚΑΙ ΔΙΚΤΥΩΝ

ΧΡΟΝΟΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ ΔΕΔΟΜΕΝΩΝ ΣΕ ΠΟΛΛΑΠΛΑ

ΚΑΝΑΛΙΑ ΕΚΠΟΜΠΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΠΑΣΧΟΥ ΒΑΣΙΛΕΙΟΥ

Επιβλέπων Καθηγητής :

Δρ. Κατσαρός Δημήτριος

Διδάσκων ΠΔ 407/80

Τμήμα Μηχανικών Η/Υ

Τηλεπικοινωνιών και Δικτύων

Πανεπιστήμιο Θεσσαλίας

Βόλος ,Σεπτέμβριος 2007



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΒΙΒΛΙΟΘΗΚΗ & ΚΕΝΤΡΟ ΠΛΗΡΟΦΟΡΗΣΗΣ
ΕΙΔΙΚΗ ΣΥΛΛΟΓΗ «ΓΚΡΙΖΑ ΒΙΒΛΙΟΓΡΑΦΙΑ»**

Αριθ. Εισ.: 6011/1
Ημερ. Εισ.: 05-11-2007
Δωρεά: Συγγραφέα
Ταξιθετικός Κωδικός: ΠΤ – ΜΗΥΤΔ
2007
ΠΑΣ



ΧΡΟΝΟΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ ΔΕΔΟΜΕΝΩΝ ΣΕ ΠΟΛΛΑΠΛΑ ΚΑΝΑΛΙΑ ΕΠΟΜΠΗΣ

Η πτυχιακή αυτή εργασία υποβάλλεται κατ' απαίτηση του πανεπιστημίου Θεσσαλίας για την απόκτηση του διπλώματος του Μηχανικού Η/Υ Τηλεπικοινωνιών και Δικτύων

Από τον

Πάσχο Βασιλείο

Επιβλέπων Καθηγητής:

Δρ. Κατσαρός Δημήτριος

Διδάσκων ΠΔ 407/80

Τμήμα Μηχανικών Η/Υ

Τηλεπικοινωνιών και Δικτύων

Πανεπιστήμιο Θεσσαλίας

Επιτροπή εξέτασης

1ος Εξεταστής

Δρ.Μποζάνης Παναγιώτης

Επίκουρος καθηγητής

Τμήμα Μηχανικών Η/Υ

Τηλεπικοινωνιών και Δικτύων

Πανεπιστήμιο Θεσσαλίας

2ος Εξεταστής

Δρ. Κατσαρός Δημήτριος

Διδάσκων ΠΔ 407/80

Τμήμα Μηχανικών Η/Υ

Τηλεπικοινωνιών και Δικτύων

Πανεπιστήμιο Θεσσαλίας

Σεπτέμβριος, 2007

Ευχαριστίες

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω όλους όσους συνέβαλαν στην ολοκλήρωση της διπλωματικής μου εργασίας .

Πάνω από όλα τον επιβλέποντα καθηγητή μου κ.Δημήτριο Κατσαρό καθώς και τον δεύτερο υπεύθυνο επιβλέποντα κ.Παναγιώτη Μποζάνη για την πολύτιμη βοήθεια την αμέριστη συμπαράσταση και την δημιουργική συνεργασία που αναπτύχθηκε καθ'όλη τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας.

Δεν θα μπορούσα ακόμα να παραλείψω όλους τους καθηγητές του τμήματος Μηχανικών Η/Υ Τηλεπικοινωνιών και Δικτύων για τη συμβολή τους στις γνώσεις που απέκτησα κατά τη διάρκεια των σπουδών μου καθώς επίσης σε όλους όσους ήταν κοντά μου τόσα χρόνια στηρίζοντας με τόσο υλικά όσο και ηθικά.

Εισαγωγή

Με την ταχύτατη ανάπτυξη των ασύρματων επικοινωνιών δίνεται στους χρήστες η δυνατότητα να έχουν πρόσβαση σε πληροφορίες οπουδήποτε βρίσκονται μέσω των κινητών τους συσκευών όπως laptops , palmtops ,έξυπνα τηλέφωνα κ.α.

Σε αυτού του είδους τις ασύρματες επικοινωνίες η εκπομπή δεδομένων(broadcasting) είναι ένας αποδοτικός τρόπος για ταυτόχρονη διάδοση δεδομένων σε ένα μεγάλο αριθμό χρηστών. Με αυτό τον τρόπο οι χρήστες σε κάποιο κυψελωτό δίκτυο μπορούν να απολαμβάνουν υπηρεσίες όπως ενημέρωση για τον καιρό ,νέα σχετικά με την κίνηση στους δρόμους μέσω εφαρμογών που απαιτούν συνεχή εκπομπή δεδομένων προς στους χρήστες οι οποίοι πρέπει να έχουν πρόσβαση σε αυτά οποιαδήποτε χρονική στιγμή.

Σε ένα τέτοιο σενάριο ο ένας εξυπηρετητής(server) στο σταθμό βάσης (base station) εκπέμπει περιοδικά τα δεδομένα από ένα συγκεκριμένο σύνολο πάνω από ένα ασύρματο κανάλι τη στιγμή που οι χρήστες συντονίζονται σε αυτό το “ κοινόχρηστο ” κανάλι περιμένοντας ο καθένας μέχρι τα επιθυμητά δεδομένα να εκπεμφθούν.Ο server με βάση το πρόγραμμα εκπομπής(broadcast schedule) του αποφασίζει ποίο αντικείμενο θα εκπέμψει σε κάθε χρονική στιγμή .

Οι χρήστες απο τη μεριά τους αφού συντονιστούν στο σωστό κανάλι περιμένουν τα δεδομένα που χρειάζονται μέχρι αυτά να εκπεμφθούν από το server, ωστόσο κάτι τέτοιο πέραν του ότι επιβαρύνει το χρόνο έναρξης της επικοινωνίας ταυτόχρονα δημιουργεί σπατάλη ενέργειας για τη συσκευή του χρήστη κάτι το οποίο είναι αρκετά σημαντικό όταν μιλάμε για κινητές συσκευές οι οποίες έχουν περιορισμένη διαθέσιμη ενέργεια.

Περιεχόμενα

Ευχαριστίες	1
Εισαγωγή.....	2
Περιεχόμενα	3
Κεφάλαιο 1 Εισαγωγή:.....	5
1.1 Ορισμός του Προβλήματος.....	5
1.2 Περίληψη Εργασίας.....	6
1.3 Βασικές έννοιες και Ορολογία.....	7
1.4 Διάρθρωση της Εργασίας	5
Κεφάλαιο 2 Κινητός Υπολογισμός	8
2.1 Ορισμός Κινητού Υπολογισμού	8
2.2 Δομή και Λειτουργία Ασύρματου Δικτύου	9
2.3 Χαρακτηριστικά Κινητού Περιβάλλοντος	11
2.4 Καταστάσεις Λειτουργίας Χρηστών.....	13
Κεφάλαιο 3 Μετάδοση Δεδομένων	15
3.1 Γενικό Μοντέλο Εμπομπής	15
3.2 Αρχιτεκτονικές Εμπομπής Δεδομένων	16
3.3 Σχολιασμός Μεθόδων Μετάδοσης	19
3.4 Ασύρματο Περιβάλλον Επικοινωνίας	19
3.5 Access και Tuning Time	20
3.6 Προγράμματα Εμπομπής	21
3.7 Δίσκοι Εμπομπής	23
3.8 Πολλαπλά Κανάλια Εμπομπής (Multiple channels)	24
Κεφάλαιο 4 Παρουσίαση Προβλήματος	26
4.1 Ορισμός Προβλήματος	26
4.2 Σχετικές Μελέτες	27
4.2.1 Προσεγγιστικός Αλγόριθμος Greedy	27
4.2.2 Προσεγγιστικός Αλγόριθμος Growing Segments	27
4.3 Η παρούσα εργασία	28
Κεφάλαιο 5 Broadcast Αλγόριθμοι	29
5.1 Αλγόριθμος Δυναμικού Προγραμματισμού(DP)	29
5.1.1 Ορισμός Μεταβλητών του Συστήματος	30
5.1.2 Ορισμός Μέσης Αναμενόμενης Καθυστερήσης	31
5.1.3 Δύο Χρήσιμα Πορίσματα	32
5.1.4 Ορισμός DP αλγορίθμου	33
5.1.5 Σχηματική Απεικόνιση Λύσης	35
5.2 Επέκταση αλγορίθμου DP για κανάλια μεταβλητού bandwidth.....	36
5.2.1 Ορισμός Επιπλέον Μεταβλητών του Συστήματος	37
5.2.2 Ορισμός Μέσης Αναμενόμενης Καθυστερήσης	37
5.2.3 Ορισμός VB-DP Αλγορίθμου	40
5.2.4 Σχόλια σχετικά με τον VB-DP Αλγόριθμο	43

5.3 Broadcast Αλγόριθμος του Ζυγού(Balance Algorithm).....	44
5.3.1 Ορισμοί Μεταβλητών και Παραδοχές	44
5.3.2 Ορισμός Αλγορίθμου Ζυγού(Balance Algorithm)	45
5.2.3.1 Συνάρτηση Partition	45
5.2.3.2 Ο Αλγόριθμος σε Ψευδοκώδικα	48
5.3.3 Παράδειγμα Εφαρμογής Αλγορίθμου.....	50
Κεφάλαιο 6 Απόδοση Αλγορίθμων	59
6.1 Σχεδιασμός Προσομοίωσης	59
6.1.1 Σύστημα και Παράμετροι Προσομοίωσης	59
6.2 Προσομοιώσεις και Αποτελέσματα	61
6.2.1 Παράμετροι που Μεταβάλλονται	61
6.2.2 Μετρήσεις ,Επιδόσεις και Σύγκριση Αλγορίθμων	62
6.2.3 Σχολιασμός Αποτελεσμάτων	68
Κεφάλαιο 7 Αποτίμηση Εργασίας και Συμπεράσματα	69
7.1 Αποτίμηση Εργασίας	69
Αναφορές – Βιβλιογραφία	71

Κεφάλαιο 1

Εισαγωγή

1.1 Ορισμός του Προβλήματος

Η εκπομπή δεδομένων (broadcast) είναι μια αποδοτική και κλιμακούμενη τεχνική μετάδοσης δεδομένων σε έναν απεριόριστο αριθμό χρηστών που είναι συντονισμένοι σε κάποιο κανάλι. Η κυκλική εκπομπή δεδομένων σε ένα κανάλι είναι μια βασική τεχνική προγραμματισμού εκπομπής (schedule technique) δεδομένων γνωστή και ως ομοιόμορφο scheduling (flat scheduling). Όταν πολλαπλά κανάλια είναι διαθέσιμα χρειάζεται να υπάρχει μια συγκεκριμένη τεχνική ανάθεσης των δεδομένων στα κανάλια. Η κατανομή των δεδομένων μεταξύ των καναλιών με βάση τις πιθανότητες ζήτησης του καθενός είναι μια τεχνική ανάθεσης γνωστή και ως επικλινής ανάθεση (skewed allocation). Σε όλα τα παραπάνω λαμβάνουμε υπόψη και έναν ακόμα σημαντικό παράγοντα που εμπεριέχεται σε όλες τις ασύρματες επικοινωνίες, την μεταβλητότητα του εύρους ζώνης (bandwidth variation) μεταξύ των καναλιών μετάδοσης.

Σε αυτή την εργασία το πρόβλημα που καλούμαστε να λύσουμε είναι αυτό της εκπομπής δεδομένων σε πολλαπλά κανάλια, διαφορετικού εύρους ζώνης το καθένα, θεωρώντας ότι η ανάθεση των δεδομένων στα κανάλια είναι skewed και ότι υπάρχει ομοιόμορφο (flat) scheduling σε κάθε κανάλι με αντικειμενικό στόχο την ελαχιστοποίηση του μέσου χρόνου αναμονής για κάθε χρήστη.

1.2 Περίληψη εργασίας

Στόχος της παρούσας εργασίας είναι να λύσει με αποδοτικό τρόπο το πρόβλημα της ανάθεσης δεδομένων ,με διαφορετική πιθανότητα ζήτησης το καθένα , σε κανάλια διαφορετικού εύρους ζώνης .Το πόσο αποδοτικό είναι το scheduling scheme φαίνεται από το μέσο χρόνο αναμονής για κάθε χρήστη. Αρχικά μελετήθηκε ένας αλγόριθμος δυναμικού προγραμματισμού (DP algorithm) ο οποίος παρουσίαζε μια λύση του προβλήματος θεωρώντας ότι όλα τα κανάλια είχαν το ίδιο εύρος ζώνης , στη συνέχεια προσπαθήσαμε να τον επεκτείνουμε να λαμβάνοντας υπ'όψιν την πιθανή μεταβλητότητα του εύρους ζώνης στα κανάλια ,επιχειρήσαμε να τον προσομοιώσουμε θεωρώντας κάποιο υποθετικό σύστημα εκπομπής δεδομένων και παρακάτω θα σας παρουσιάσουμε τα αποτελέσματα των μετρήσεων.Στη συνέχεια συγκρίναμε τον αλγόριθμο που κατασκευάσαμε με κάποιον άλλο αλγόριθμο ανάθεσης δεδομένων σε κανάλια μεταβλητού εύρους ζώνης ο οποίος στηρίζεται στη λογική κάποιου είδους συσταδοποίησης .Πλήρης περιγραφή των αλγορίθμων βρίσκεται στη συνέχεια της εργασίας όπως επίσης και αναλυτική παρουσίαση σύγκριση και σχολιασμός των αποτελεσμάτων που προέκυψαν από τη σύγκριση των δύο μεθόδων.

Θα πρέπει να επισημάνουμε ότι σε όλη μας τη μελέτη έχουμε θεωρήσει ότι τα δεδομένα που εκπέμπονται είναι δεδομένα ίδιου μεγέθους .

1.3 Βασικές έννοιες και ορολογία

- **Κυκλική εκπομπή δεδομένων(cyclically broadcasting):**Περιοδική εκπομπή δεδομένων διαδοχικά ένα ένα με την ίδια σειρά κάθε φορά.
- **Flat scheduling:** Είναι ο προγραμματισμός για εκπομπή των δεδομένων ένα ένα με την ίδια σειρά κάθε φορά με ίδια συχνότητα εκπομπής για κάθε αντικείμενο.
- **Skewed allocation:** Είναι η ανάθεση δεδομένων στα κανάλια με διαφορετικές πιθανότητες ζήτησης.
- **Variant Bandwidth:** Είναι η παρουσία στο σύστημά μας καναλιών με διαφορετικό εύρο ζώνης
- **Broadcasting Scheme :** Είναι μία καθορισμένη διάταξη εκπομπής των δεδομένων σε κάθε κανάλι η οποία έχει προκύψει από τη εφαρμογή κάποιου αλγορίθμου στα δεδομένα ενός προβλήματος.
- **Κινητός Υπολογισμός :** Είναι η δυνατότητα που παρέχεται στους χρήστες να συμμετέχουν στη ροή πληροφορίας μέσα απο κατάλληλες συσκευές οπουδήποτε και αν βρίσκονται.

Λέξεις Κλειδιά:Wireless Communication ,data broadcast ,multiple channels , probability demand ,skewed allocation ,flat scheduling ,average waiting time ,dynamic programming

Κεφάλαιο 2

Κινητός Υπολογισμός

2.1 Ορισμός Κινητού Υπολογισμού

Με την ανάπτυξη της τεχνολογίας στις ασύρματες επικοινωνίες οι κινητοί χρήστες μπορούν να έχουν πρόσβαση σε πληροφορίες οπουδήποτε και οποτεδήποτε χρησιμοποιώντας διάφορες ασύρματες κινητές συσκευές όπως φορητοί υπολογιστές ,υπολογιστές χειρός ,έξυπνα τηλέφωνα κ.α.

Αυτή ακριβώς είναι και η έννοια του κινητού υπολογισμού,έτσι το προκύπτον περιβάλλον ονομάζεται *περιβάλλον κινητού υπολογισμού* (mobile computing environment) και ο τομέας που ασχολείται και ερευνά γι'αυτό *κινητός υπολογισμός*.

Η έρευνα στον τομέα του κινητού υπολογισμού είναι αρκετά σημαντική και ενδιαφέρουσα και συνεχώς αναπτύσσεται προσπαθώντας να καλύψει όλο και περισσότερες ανάγκες των χρηστών οι οποίοι με τη σειρά τους γίνονται περισσότερο απαιτητικοί. Οι προσπάθειες της έρευνας κινούνται κυρίως προς τα δύο κατευθύνσεις η μία είναι προς την βελτίωση του τεχνολογικού μέσου και η άλλη προς την ανάπτυξη εφαρμογών και προτύπων που θα υποστηρίζουν την επικοινωνία χρηστών μέσα απο το εκάστοτε μέσο μετάδοσης .

Ήδη αναπτύσσονται καινούργια πρωτόκολλα για τις ασύρματες μεταδόσεις που προσφέρουν μεγαλύτερη ταχύτητα , διαθεσιμότητα , ασφάλεια και είναι ικανά

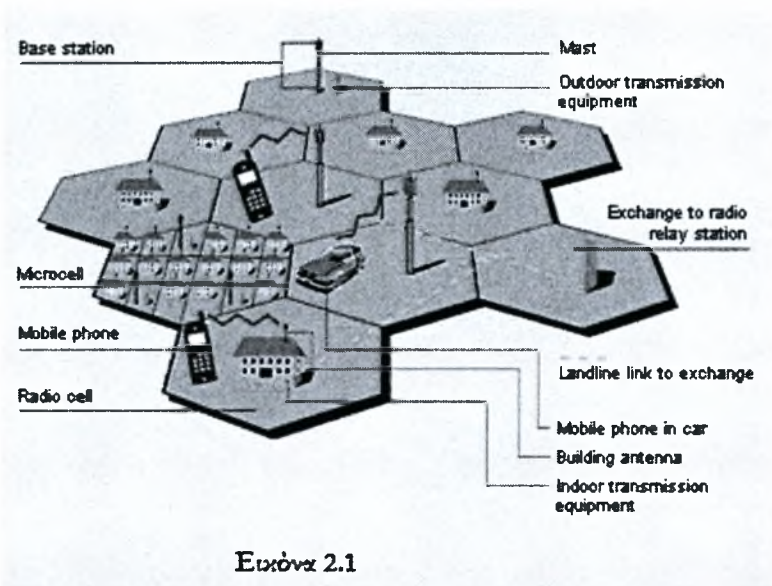
να εξυπηρετήσουν ολοένα και περισσότερους χρήστες καλύπτοντας και ολοένα και μεγαλύτερες περιοχές.

Αλλά και στον τομέα της δημιουργίας συστημάτων που λειτουργούν σε περιβάλλοντα κινητού υπολογισμού υπάρχουν πολλές προκλήσεις κυρίως λόγω του ότι οι παράμετροι που πρέπει να ληφθούν υπ' όψιν ποικίλουν ,είναι δύσκολο να καθοριστούν και συχνά μεταβάλλονται καθώς και λόγω του ότι οι δυσκολίες που πρέπει να αντιμετωπιστούν είναι σημαντικές (μεταβλητότητα εύρους ζώνης ,συχνές αποσυνδέσεις απο το δίκτυο ,περιορισμένοι υπολογιστικοί πόροι με βάση τις δυνατότητες του υλικού και περιορισμένη διαθέσιμη ενέργεια).

2.2 Δομή και Λειτουργία Ασύρματου Δικτύου

Η ανάπτυξη και η διάδοση των ασύρματων δικτύων ήταν ταχύτατη και ειδικά τα τελευταία χρόνια .Κάποια από τα σημαντικότερα πλεονεκτήματα τους είναι ότι παρέχουν την δυνατότητα φθηνής και γρήγορης εγκατάστασης και επιτρέπουν στο χρήστη να κινείται όσο αυτός είναι συνδεδεμένος στο δίκτυο.

Μέσα σε ένα τέτοιο δίκτυο θεωρούμε τους σταθερούς κόμβους (κεραίες ,σταθμοί βάσεις) και τους κινούμενους (χρήστες με ασύρματες συσκευές,που κινούνται μέσα στο δίκτυο).Οι σταθεροί κόμβοι είναι συνδεδεμένοι μεταξύ τους μέσω ενός σταθερού ,συνήθως ενσύρματου δικτύου ,ηψηλής ταχύτητας (της τάξης των Mbps ,Gbps).Οι κόμβοι αυτοί συνήθως είναι μεγάλες κεραίες που μοιράζουν το δίκτυο στα όρια κάποιας περιοχής και ονομάζονται σταθμοί βάσης (base station).



Μία από τις πιο διαδεδομένες μορφές ασυρμάτων δικτύων είναι τα κυψελωτά δίκτυα (cellular networks). Σε αυτά η περιοχή που καλύπτει ένας σταθμός βάσης ονομάζεται κυψέλη (cell). Οι χρήστες κινούμενοι σε κάποια περιοχή μπορεί να εξυπηρετούνται κάθε φορά από διαφορετικό σταθμό βάσης μεταβαίνοντας από κυψέλη σε κυψέλη. Τέτοια δίκτυα είναι συνήθως τα δίκτυα κινητής τηλεφωνίας (GSM networks)

Άλλα είδη ασύρματων δικτύων είναι τα εξής:

- **Ασύρματα τοπικά δίκτυα (WLANs).** Τα δίκτυα αυτά χρησιμοποιούνται για την κάλυψη μικρών περιοχών, δεν υπάρχει η θεώρηση των κυψελών και οι κινητοί χρήστες έχουν πρόσβαση στο δίκτυο μέσω κάποιας συσκευής που ονομάζεται **access point**.
- **WiFi δίκτυα:** Τα Wi-Fi δίκτυα είναι ευρέως χρησιμοποιούμενα ασύρματα δίκτυα σε συστήματα υπολογιστών και επιτρέπουν τη σύνδεση τους στο διαδίκτυο ή σε άλλες συσκευές που υποστηρίζουν παρόμοιες λειτουργίες

Ο χρήστης που κινείται σε ένα ασύρματο δίκτυο μπορεί να βρεθεί σε διάφορες καταστάσεις . Αρχικά ο χρήστης μπορεί να βρίσκεται εκτός περιοχής κάλυψης του δικτύου, στη συνέχεια ο χρήστης προσπαθεί να ανιχνεύσει κάποιο δίκτυο στην περιοχή του και εφόσον αυτό συμβεί προσπαθεί να συντονιστεί σε κάποιο από τα διαθέσιμα δίκτυα ανταλλάσσοντας δεδομένα. Από τη στιγμή που του επιτραπεί η σύνδεση ,ο χρήστης επικοινωνεί μέσω του σταθμού βάσης(BS) ή κάποιου access point με τους άλλους χρήστες του δικτύου. Ο χρήστης μπορεί ακόμα να βρεθεί και σε μία ενδιάμεση κατάσταση από τη στιγμή που μπορεί να είναι στα όρια της περιοχής κάλυψης ενός σταθμού και κινούμενος να μεταβεί σε περιοχή κάλυψης ενός γειτονικού σταθμού βάσης .Σε αυτή τη μεταβατική κατάσταση βρίσκονται πολλές φορές οι χρήστες που κινούνται μέσα σε κάποιο κυψελωτό δίκτυο. Σε αυτή την περίπτωση ο χρήστης σταδιακά αρχίζει να αντιλαμβάνεται σημαντική εξασθένηση του σήματος από το σταθμό βάσης της κυψέλης που βρίσκεται ενώ ταυτόχρονα αρχίζει να αντιλαμβάνεται σήμα από το γειτονικό σταθμό βάσης .Σε κάποια χρονική στιγμή ο χρήστης θα είναι συνδεδεμένος και στους δύο BS και στη συνέχεια συνδέεται στο δεύτερο BS και αποσυνδέεται από τον πρώτο.Το φαινόμενο αυτό στο οποίο ο χρήστης βρίσκεται σε μία μεταβατική κατάσταση ονομάζεται handoff.

2.3 Χαρακτηριστικά Κινητού Περιβάλλοντος

Τα κύρια χαρακτηριστικά ενός κινητού υπολογιστικού περιβάλλοντος είναι το μέσο μετάδοσης ,η κινητικότητα των χρηστών μέσα σε αυτό και η φορητότητα των συσκευών.

1. Μέσο μετάδοσης (Transmission Medium)

Σε ένα ασύρματο περιβάλλον μετάδοσης το μέσο που χρησιμοποιείται για τη μετάδοση δεδομένων είναι ο αέρας.Αυτό από τι μία εξασφαλίζει μία μεγάλη ευελιξία στην εγκατάσταση του δικτύου

,χαμηλό κόστος όπως επίσης και δυνατότητα για κάλυψη πολύ μεγαλύτερων περιοχών από την άλλη όμως σε ένα τέτοιο περιβάλλον υπάρχει πολλές φορές διακύμανση του εύρους ζώνης καθώς η λήψη του σήματος πολλές φορές εξαρτάται και από εξωτερικούς παράγοντες (μορφολογία φυσικού περιβάλλοντος ,φαινόμενο **fading ,multi-path**) κάτι το οποίο μπορεί να προκαλέσει συχνές αποσυνδέσεις από το δίκτυο κατάσταση μη επιθυμητή για τον χρήστη.Σε ένα τέτοιο περιβάλλον για την επικοινωνία του χρήστη με το σταθμό βάσης θεωρούμε δύο ειδών κανάλια.Το καθοδικό (**downlink**) κανάλι μέσα από το οποίο μεταδίδονται οι πληροφορίες από το σταθμό βάσης προς το χρήστη και το ανοδικό (**uplink**) κανάλι μέσα από το οποίο μεταφέρονται δεδομένα από το χρήστη προς το σταθμό βάσης.Επειδή οι πληροφορίες που χρειάζεται να πάρει κάποιος χρήστης από το σταθμό βάσης είναι περισσότερες από αυτές που πρέπει να στείλει το εύρος ζώνης του uplink καναλιού είναι το πολύ ίσο με αυτό του downlink καναλιού, ωστόσο συνήθως το downlink κανάλι έχει μεγαλύτερο εύρος ζώνης από το uplink

2. Κινητικότητα (Mobility)

Ένας από τους πιο σημαντικούς αλλά και αστάθμητους παράγοντες σε ασύρματες μεταδόσεις είναι οι κινητικότητας των χρηστών.Αυτό από τη μία είναι ένα από τα πιο σημαντικά πλεονεκτήματα που παρέχουν οι ασύρματες επικοινωνίες στους χρήστες έναντι των ενσύρματων δικτύων ,ωστόσο αυτός ο παράγοντας κάνει το περιβάλλον να μεταβάλλεται σε κάθε χρονική στιγμή κάτι το οποίο δυσκολεύει τη μελέτη του ,ενώ ταυτόχρονα μεταβάλλει σημαντικά την απόδοση του δικτύου.Λόγω αυτού του παράγοντα το σύστημα ξοδεύει επιπλέον ενέργεια στην προσπάθεια του για συνεχή εντοπισμό των κινούμενων χρηστών. Από τη μεριά των χρηστών επηρεάζεται η συνδεσιμότητα τους στο δίκτυο λόγω απώλειας σήματος είτε εξαιτίας του φαινομένου

handoff είτε εξαιτίας μη καλής λήψης σήματος απο εξωτερικά εμπόδια του περιβάλλοντος .

3. Φορητότητα (Portability)

Τα χαρακτηριστικά των συσκευών που χρησιμοποιούνται σε τέτοια περιβάλλοντα μετάδοσης υπόκεινται αρκετούς περιορισμούς .Οι συσκευές αυτές πρέπει από τη μία να είναι σχετικά μικρές και ελαφριές ώστε να είναι εύκολα μεταφέρσιμες και από την άλλη να διαθέτουν επαρκή υπολογιστική ισχύ αλλά και μνήμη ώστε να μπορούν να τρέξουν εφαρμογές που χρειάζεται ο χρήστης για την επικοινωνία του .Ωστόσο το μέγεθος τους θέτει ένα περιορισμό στη δυνατότητα αύξησης της επίδοσης του υλικού της συσκευής κάτι το οποίο πρέπει να λαμβάνεται πάντα υπ'όψιν .Τέλος η ενέργεια την οποία έχει στη διάθεση της μία τέτοια συσκευή είναι αρκετά περιορισμένη και πρέπει ακόμα και ο σχεδιασμός τους συστήματος και του περιβάλλοντος επικοινωνίας να είναι τέτοιος ώστε να μειώνει κατά το δυνατόν περισσότερο την απαίτηση για σπατάλη ενέργειας από τη μεριά του χρήστη.

2.4 Καταστάσεις Λειτουργίας Χρηστών

Ανάλογα με τις ανάγκες επικοινωνίας των χρηστών αλλά και ανάλογα της κατάστασης του περιβάλλοντος κάθε χρονική στιγμή οι χρήστες μπορούν να βρεθούν σε δύο διαφορετικές καταστάσεις .

- **Ενεργή Κατάσταση(active mode):** Είναι η κατάσταση στην οποία ο χρήστης επικοινωνεί με το δίκτυο στέλνοντας ή λαμβάνοντας

δεδομένα. Η κατάσταση αυτή είναι η κατάσταση πλήρους κατανάλωσης ενέργειας.

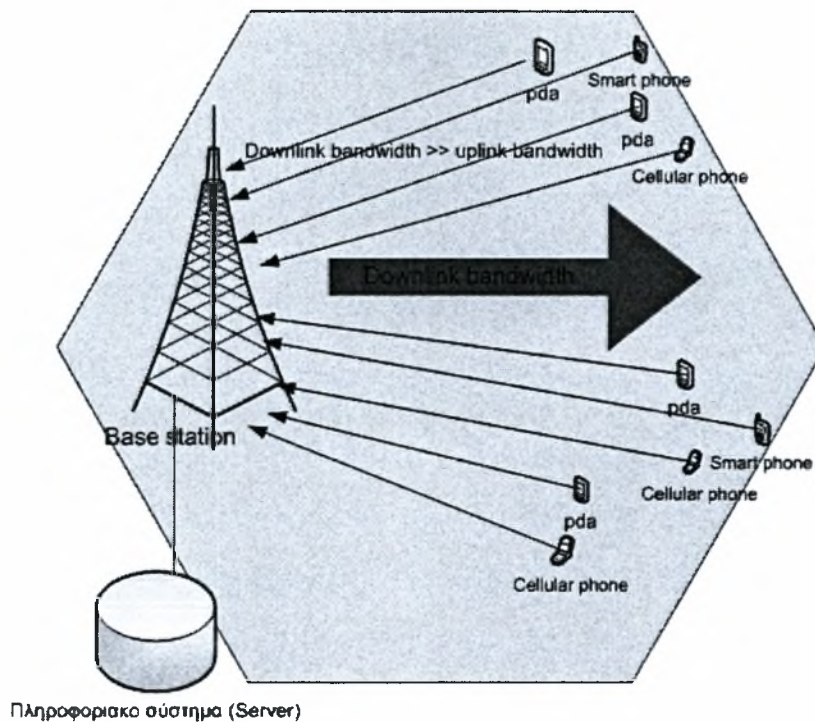
- **Κατάσταση Εξοικονόμησης Ενέργειας (doze mode):** Είναι η κατάσταση στην οποία ο χρήστης βρίσκεται σε κάποιου είδους “λίθαργο” ,εκτελεί μειωμένες λειτουργίες ώστε να εξοικονομεί ενέργεια.

Κεφάλαιο 3

Μετάδοση Δεδομένων

3.1 Γενικό Μοντέλο Εκπομπής

Το γενικό μοντέλο εκπομπής αντιστοιχεί σε ένα θεωρητικό μοντέλο ασύρματου δικτύου στο οποίο θεωρούμε ότι υπάρχει μία κυψέλη στο οποίο αντιστοιχεί ένας σταθμός βάσης με έναν server που εξυπηρετεί τις αιτήσεις των χρηστών, κάποιοι χρήστες με ασύρματες συσκευές που συνδέονται στο υπάρχων δίκτυο. Θεωρούμε ότι ο αριθμός των χρηστών είναι πολύ μεγαλύτερος από τον αριθμό των servers καθώς επίσης και ότι το εύρος ζώνης του καθοδικού καναλιού είναι πολύ μεγαλύτερο από αυτό του καθοδικού καναλιού (downlink bandwidth \gg uplink bandwidth).



Εικόνα 3.1

3.2 Αρχιτεκτονικές Εκπομπής Δεδομένων

Στα ασύρματα υπολογιστικά περιβάλλοντα ξεχωρίζουμε τρία είδη μετάδοσης δεδομένων, τη μετάδοση *σημείο προς σημείο* (*pure pull* ή *point to point*), την *καθαρή εκπομπή* (*pure push* ή *broadcast*) και την *υβριδική και κατ' απαίτηση εκπομπή* (*hybrid & on demand broadcast*).

Αναλυτικότερα:

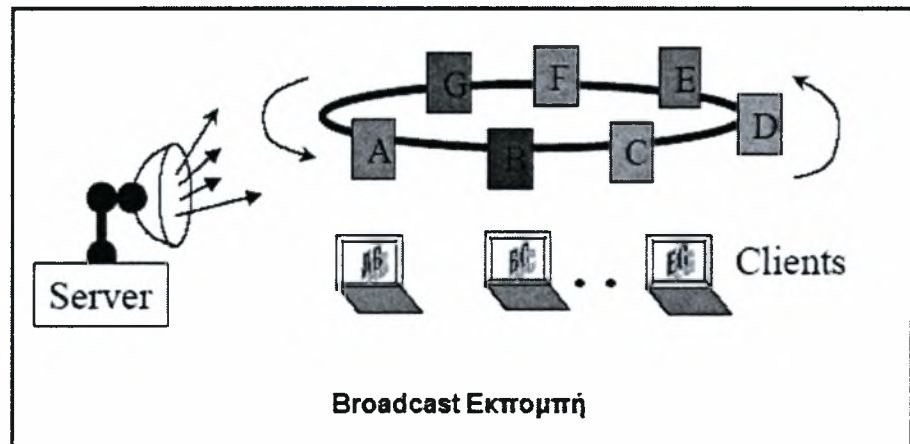
- **Point to Point**

Σε αυτού του είδους τη μετάδοση κάθε φορά που κάποιος χρήστης θελήσει να κάνει μία αίτηση για κάποια δεδομένα από τον server

δημιουργείται ένα κανάλι εκπομπής μεταξύ τους .Ο server αφού λάβει την αίτηση και ολοκληρώσει τη διαδικασία υπολογισμού των δεδομένων στη συνέχεια στέλνει τα δεδομένα πίσω στο χρήστη και το κανάλι εκπομπής κλείνει.

- **Broadcast**

Σε αυτού του είδους τη μετάδοση τα δεδομένα στέλνονται ταυτόχρονα σε όλους τους χρήστες που βρίσκονται μέσα στην περιοχή εκπομπής (**broadcast area**) πχ σε μία κυψέλη. Έτσι ο server δεν περιμένει κάποια αίτηση από τους clients της περιοχής του αλλά περιοδικά εκπέμπει ένα σύνολο δεδομένων σε κάποια κανάλια μετάδοσης (**broadcast channels**) στα οποία έχουν πρόσβαση όλοι οι χρήστες που βρίσκονται μέσα στην περιοχή εκπομπής .Η σειρά και ο τρόπος μετάδοσης των δεδομένων σε κάποια μετάδοση broadcast ονομάζεται πρόγραμμα εκπομπής (**broadcasting schedule**).



Εικόνα 3.2

- **Hybrid and On Demand Broadcast**

Σε αυτού του είδους τη μετάδοση ο κάθε client στέλνει την αίτησή του και ο server αφού λάβει όλες τις αιτήσεις προσαρμόζει την broadcast εκπομπή του τις ομαδοποιεί και τις εκπέμπει. Ο server μπορεί να προσαρμόσει την εκπομπή δεδομένων του με βάση τη δημοτικότητα (**popularity**) των δεδομένων και τις ανάγκες των χρηστών.

3.3 Σχολιασμός Μεθόδων Μετάδοσης

Στην **point-to-point** μετάδοση σπαταλάται σημαντικό εύρος ζώνης καθώς εκπέμπεται το ίδιο αντικείμενο πολλές φορές στο κανάλι (για διαφορετικούς χρήστες) όταν υπάρχει επικάλυψη στα δεδομένα που αυτοί ζητούν. Έτσι ένα σύστημα που χρησιμοποιεί **point-to-point** μετάδοση δεν μπορεί να κλιμακωθεί σε μεγάλο αριθμό κινητών χρηστών. Από την άλλη η **broadcast** μετάδοση χρησιμοποιεί πιο αποδοτικά το διαθέσιμο εύρος ζώνης (**bandwidth**) καθώς επεξεργάζεται τις αιτήσεις των χρηστών και εκπέμπει τα δεδομένα που έχουν ζητήσει οι χρήστες ακολουθιακά και περιοδικά με διαφορετική πιθανόν συχνότητα εκπομπής για το κάθε ένα. Αυτό όμως δίνει ένα μειονέκτημα σε αυτού του είδους τις μεταδόσεις καθώς αυξάνει τον χρόνο αναμονής του χρήστη στο κανάλι. Ωστόσο είναι σημαντική η εξοικονόμηση ενέργειας που επιτυγχάνεται με την **broadcast** μετάδοση κάτι το οποίο οφείλεται στο γεγονός ότι οι συσκευές των κινητών χρηστών καταναλώνουν περισσότερη ενέργεια όταν στέλνουν δεδομένα από ότι όταν λαμβάνουν.

3.4 Ασύμμετρο Περιβάλλον επικοινωνίας

Όσον αφορά το εύρος ζώνης στα ασύρματα περιβάλλοντα μετάδοσης το περιβάλλον που κυρίως συναντούμε είναι το *ασύμμετρο* περιβάλλον επικοινωνίας. Έτσι σε πολλές υπάρχουσες αλλά και αναπτυσσόμενες εφαρμογές η χωρητικότητα του καθοδικού(**downlink**) καναλίου επικοινωνίας από τους servers προς τους clients είναι πολύ μεγαλύτερη από τη χωρητικότητα του καναλίου από τους clients προς τους servers.

Η ασυμμετρία επικοινωνίας μπορεί να προκύψει για δύο λόγους:

- Οι περιορισμοί στο εύρος ζώνης του φυσικού μέσου επικοινωνίας.
Σκεφτείτε για παράδειγμα ότι στατικοί servers διαθέτουν ισχυρούς αναμεταδότες ,ενώ οι κινητοί χρήστες έχουν μικρή ή καθόλου δυνατότητα μετάδοσης
- Εξαιτίας του προτύπου ροής πληροφορίας στην εφαρμογή .Όπως γίνεται σε ένα σύστημα ανάκτησης πληροφορίας, όπου ο αριθμός των χρηστών είναι πολύ μεγαλύτερος από τον αριθμό των servers, είναι ασύμμετρο επειδή δεν υπάρχει αρκετή χωρητικότητα (είτε στο δίκτυο είτε στους servers)για να εξυπηρετηθούν όλες οι ταυτόχρονες αιτήσεις που μπορεί να συμβούν.

3.5 Access και Tuning Time

Για τη μελέτη των προβλημάτων μετάδοσης δύο είναι τα μεγέθη που θα μας απασχολήσουν περισσότερο.

- **Access Time ή Latency(Χρόνος Πρόσβασης)**

Access Time είναι ο χρόνος από τη στιγμή που ένας χρήστης εκδηλώνει ενδιαφέρον για κάποιο αντικείμενο στέλνοντας μία αίτηση για αυτό στο server μέχρι τη στιγμή που το αποκτά.

- **Tuning Time(Χρόνος Συντονισμού)**

Tuning Time είναι ο χρόνος κατά τον οποίο ο χρήστης παραμένει σε ενεργή κατάσταση ώντας συντονισμένος στο κανάλι .Ο χρόνος αυτός προσδιορίζει και την κατανάλωση ενέργειας για την απόκτηση των δεδομένων

Θα πρέπει εδώ να ορίσουμε και μία επιπλέον έννοια τον *κύκλο εκπομπής*.

Ο *κύκλος εκπομπής (broadcast cycle)* είναι στην ουσία το μέγεθος του προγράμματος εκπομπής (*broadcast schedule*) και ισούται με το πλήθος των δεδομένων του προγράμματος επί το χρόνο μετάδοσης του κάθε δεδομένου. Ο χρόνος μετάδοσης ορίζουμε να μετριέται σε μονάδες χρόνου που τις ονομάζουμε **tics**.

Στόχος μας είναι η ελαχιστοποίηση του *broadcast cycle* και κατά συνέπεια η ελαχιστοποίηση των παραπάνω χρόνων.

3.6 Προγράμματα Εκπομπής

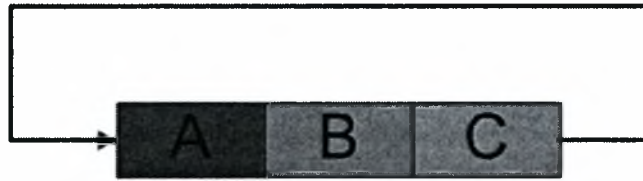
Για την εκπομπή των δεδομένων από το server στους χρήστες διακρίνουμε τρεις διαφορετικούς τρόπους, τρία διαφορετικά προγράμματα εκπομπής (*schedules*).

- **Ομοιόμορφο Πρόγραμμα Εκπομπής (flat schedule)**

Όταν έχουμε *flat scheduling* τα δεδομένα εκπέμπονται ακολουθιακά και κυκλικά. Έτσι όλα τα δεδομένα σε κάποιο κανάλι εκπέμπονται με την ίδια συχνότητα.

Π.χ. Αν το σύνολο των δεδομένων που θέλω να εκπέμψω είναι
 $P = \{A, B, C\}$

Αυτά θα εκπεμφθούν ως εξής :



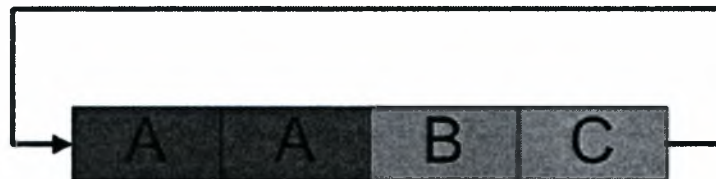
flat scheduling

Εικόνα 3.6.1

- **Κυρτό Πρόγραμμα Εκπομπής (skewed schedule)**

Στο *skewed scheduling* κάποιο από τα δεδομένα παρουσιάζει μεγαλύτερη συχνότητα εμφάνισης, πολλαπλάσια από τα άλλα. Θεωρούμε έτσι ότι δημιουργείται μία υποακολουθία ιδίων αντικειμένων η οποία εκπέμπεται διαδοχικά χωρίς να παρεμβαίνει κάποιο από τα άλλα όμοια της υποακολουθίας αντικείμενα.

Πχ. Αν στο προηγούμενο σύνολο αντικειμένων θεωρήσω ότι το A εκπέμπεται δύο φορές πιο συχνά σε σχέση με τα B και C τότε το πρόγραμμα εκπομπής θα έχει ως εξής:



skewed scheduling

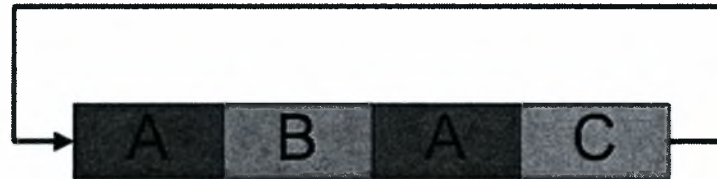
Εικόνα 3.6.2

- **Πολλαπλών Δίσκων (Multi Disk scheduling)**

Αυτό το είδος broadcasting είναι το ίδιο με το προηγούμενο με την διαφορά ότι η υποακολουθία όμοιων αντικειμένων που δημιουργείται δεν εκπέμπεται συνεχόμενη αλλά μπορούν ενδιάμεσα να

περεμβληθούν μη όμοια αντικείμενα σε σχέση με αυτά της υποακολουθίας .

Πχ Θεωρούμε το προηγούμενο παράδειγμα τώρα το πρόγραμμα εκπομπής μπορεί να είναι το εξής



Multi-Disk scheduling

Εικόνα 3.6.3

Θα πρέπει να σημειώσουμε ότι σε όλων των ειδών τις μεταδόσεις τα δεδομένα εκπέμπονται πάντα κυκλικά (**broadcast cycle**) έτσι αναλόγως το **schedule** που χρησιμοποιούμε τα δεδομένα έχουν συγκεκριμένη συχνότητα εμφάνισης.

3.7 Δίσκοι Εκπομπής (Broadcast Disks)

Οι δίσκοι εκπομπής είναι στην ουσία μία μοντελοποίηση των διαθέσιμων καναλιών εκπομπής ,παρέχουν ένα κομψό τρόπο για τη δημιουργία προγράμματος εκπομπής δημιουργώντας μία εναέρια μνήμη ,ωστόσο δεν μας δίνουν τη μεθοδολογία για την επιλογή των παραμέτρων του συστήματος.

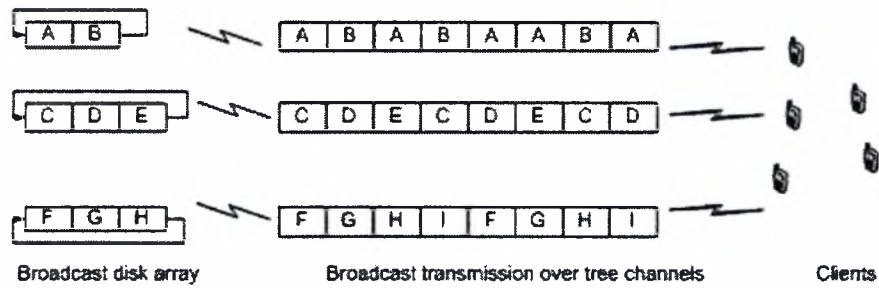
Η μοντελοποίηση των δίσκων εκπομπής γίνεται ως εξής:

1. Διατάσσουμε τα αντικείμενα από το πιο δημοφιλές (**hottest**) στο λιγότερο δημοφιλές. Κάθε αντικείμενο έχει μία συγκεκριμένη πιθανότητα ζήτησης (**demand probability**) από τους χρήστες έτσι η διάταξη τους γίνεται με βάση αυτήν.
2. Στη συνέχεια διαμερίζουμε τη λίστα των διατεταγμένων δεδομένων σε πολλαπλές διαμερίσεις (**partitions**) όπου συνήθως η κάθε διαμέριση έχει αντικείμενα με παρόμοιες πιθανότητες ζήτησης. Οι διαμερίσεις αυτές ονομάζονται **Δίσκοι**.
3. Επιλέγουμε τις σχετικές συχνότητες εκπομπής του κάθε Δίσκου.

Ο αριθμός των δίσκων καθορίζει τον αριθμό των διαφορετικών συχνοτήτων με τις οποίες θα εκπεμφθούν τα δεδομένα ενώ οι σχετικές συχνότητες καθορίζουν το μέγεθος του **broadcast cycle** και τον ρυθμό άφιξης της κάθε σελίδας. Συνήθως οι γρήγοροι δίσκοι έχουν λίγες σελίδες.

3.8 Πολλαπλά Κανάλια Εκπομπής(Multiple Channels)

Η ανάγκη για περαιτέρω μείωση του μέσου χρόνου αναμονής του χρήστη οδήγησε στην ανάπτυξη της ιδέας για χρησιμοποίηση παραπάνω του ενός καναλιών σχετικά με τη μετάδοση των δεδομένων κάτι το οποίο μπορεί να οδηγήσει σε σημαντική μείωση του **broadcast cycle**. Με τον όρο πολλαπλά κανάλια εννοούμε είτε το λογικό διαχωρισμό ενός μεγάλου καναλιού μεγάλου **bandwidth** ,είτε την πραγματική χρησιμοποίηση πολλών καναλιών διαφορετικού **bandwidth** το καθένα για την παράλληλη εκπομπή των δεδομένων προς τους χρήστες .



Εικόνα 3.8

Με αυτό τον τρόπο μπορεί μία εφαρμογή να αποκτήσει επιπλέον κανάλια ώστε να μπορεί να εξυπηρετήσει μεγαλύτερο πληθυσμό χρηστών, επίσης η χρήση πολλαπλών καναλιών βοηθάει σε περίπτωση που υπάρχουν χρήστες με ετερογενείς δυνατότητες, δηλαδή ο καθένας συντονίζεται με τον server για τη μεταφορά δεδομένων σε διαφορετική συχνότητα. Υπάρχει ακόμα καλύτερη απόδοση και μεγαλύτερη ανοχή σε σφάλματα όταν υπάρχουν αρκετές αποσυνδέσεις αφού αν για κάποιο λόγο γίνει αποσύνδεση από κάποιο κανάλι δεν χάνεται όλος ο **broadcast** κύκλος αλλά μόνο τα δεδομένα που εκπέμπονταν από το συγκεκριμένο κανάλι.

Κεφάλαιο 4

Παρουσίαση Προβλήματος

4.1 Ορισμός Προβλήματος

Σε όλες τις broadcast μεταδόσεις ένας server στον σταθμό βάσης εκπέμπει επαναλαμβανόμενα τα δεδομένα από ένα σύνολο ασύρματων καναλιών ενώ την ίδια στιγμή οι clients συντονίζονται σε ένα κοινόχρηστο κανάλι περιμένοντας για το επιθυμητό αντικείμενο. Ο server ακολουθεί ένα **broadcast schedule** προκειμένου να αποφασίσει ποίο από τα αντικείμενα του συνόλου δεδομένων πρέπει να μετάδοθει σε οποιαδήποτε χρονική στιγμή. Ένα αποδοτικό broadcast schedule ελαχιστοποιεί τον μέσο χρόνο αναμονής του χρήστη δηλαδή το χρόνο που σπαταλάται από ένα χρήστη πριν λάβει το δεδομένο που χρειάζεται. Ο χρόνος αναμονής για τον χρήστη αυξάνει με το μέγεθος του συνόλου των αντικειμένων που είναι προς μετάδοση από τον server. Στην πραγματικότητα ο client πρέπει να περιμένει τη μετάδοση πολλών αντικειμένων που δεν τον ενδιαφέρουν μέχρι να αρχίσει να λαμβάνει αυτά που χρειάζεται. Η χρήση πολλαπλών καναλιών μπορεί να βελτιώσει την αποδοτικότητα του συστήματος με τη δημιουργία μικρότερων προγραμμάτων εκπομπής σε κάθε κανάλι. Έτσι κάθε χρήστης μπορεί να έχει πρόσβαση είτε σε ένα μοναδικό κανάλι είτε σε πολλά κανάλια διαθέσιμα την ίδια χρονική στιγμή. Στόχος μας είναι η δημιουργία ενός αποδοτικού **broadcast scheme** για τη μετάδοση δεδομένων σε πολλαπλά

κανάλια διαφορετικού εύρους ζώνης. Το πόσο αποδοτικό είναι ένα **broadcast scheme** φαίνεται από το κατά πόσο ελαχιστοποιείται ο μέσος χρόνος αναμονής για τον χρήστη προκειμένου να λάβει τα δεδομένα που έχει ζητήσει. Στην παρούσα εργασία θεωρούμε ότι σε όλες τις μεταδόσεις χρησιμοποιείται **flat broadcasting schedule** και ότι τα δεδομένα που εκπέμπονται είναι όλα του ίδιου μεγέθους.

4.2 Σχετικές Μελέτες

Στην θεωρία των Δίσκων Εκπομπής (broadcast disks) έχουν αναπτυχθεί μερικοί αλγόριθμοι οι οποίοι καλούνται να επιλύσουν το πρόβλημα της διαμέρισης του συνόλου των δεδομένων σε υποσύνολα στοιχείων τα οποία στη συνέχεια τοποθετούνται για εκπομπή στα διαθέσιμα κανάλια.

4.2.1 Προσεγγιστικός Αλγόριθμος Greedy

Ο *Greedy* αλγόριθμος βασίζεται στην τεχνική του δυναμικού προγραμματισμού προσπαθεί να βρει το καλύτερο σημείο διαμέρισης διασπώντας διαδοχικά το αρχικό σύνολο. Ο αλγόριθμος αυτός τρέχει σε χρόνο $O(N*K)$.

4.2.2 Προσεγγιστικός Αλγόριθμος Growing Segments

Η μέθοδος *Growing Segments* ξεκινά με μία αρχική “ελάχιστη” διαμέριση αναθέτοντας ένα αντικείμενο σε κάθε δίσκο. Κατόπιν επαυξάνει κάθε τμήμα της διαμέρισης περιλαμβάνοντας τόσα αντικείμενα όσα καθορίζονται από κάποια παράμετρο *increment* και υπολογίζει ποια από αυτές τις επαυξήσεις δίνει τη μεγαλύτερη ελάττωση της μέσης καθυστέρησης. Κατόπιν εφαρμόζεται το ίδιο στην διαμέριση που προέκυψε μέχρι να καλύψει όλα τα αντικείμενα.

Η παράμετρος **increment** είναι πολύ σημαντική, όσο μεγαλύτερη είναι η τιμή της τόσο μικρότερη είναι η πολυπλοκότητα του χρόνου εκτέλεσης του αλγορίθμου αλλά και μικρότερη η ποιότητα του παραγόμενου προγράμματος εκπομπής.

Οι δύο παραπάνω αλγόριθμοι αναφέρονται σε ανάθεση δεδομένων με διαφορετικές πιθανότητες ζήτησης σε πολλαπλά κανάλια ίδιου εύρους ζώνης.

Όσον αφορά προβλήματα στα οποία τα έχουμε πολλαπλά κανάλια με διαφορετικό εύρος ζώνης έχουμε να αναφέρουμε τον αλγόριθμο *TOSA* ο οποίος δεν κρίνεται ιδιαίτερα αποδοτικός και ο αλγόριθμος *adaptive algorithm (AP)*.

4.3 Η Παρούσα Εργασία

Η υλοποίηση της παρούσας εργασίας έχει στηριχτεί στην μελέτη και επέκταση του αλγορίθμου DP (dynamic programming) του οποίου η λύση αναφέρεται για ένα σύστημα στο οποίο όλα τα κανάλια έχουν το ίδιο εύρος ζώνης. Εμείς προσαρμόσαμε τη λειτουργία του αλγορίθμου σε σύστημα στο οποίο το εύρος ζώνης των πολλαπλών καναλιών είναι μεταβλητό στη συνέχεια ο αλγόριθμος που δημιουργήσαμε συγκρίνεται με έναν άλλο αλγόριθμο ο οποίος λύνει το ίδιο πρόβλημα με μία διαφορετική τεχνική. Αναλυτική παρουσίαση των παραπάνω αλγορίθμων ακολουθεί παρακάτω.

Κεφάλαιο 5

Broadcast Αλγόριθμοι

5.1 Αλγόριθμος Δυναμικού Προγραμματισμού (DP)

Ο αλγόριθμος Δυναμικού Προγραμματισμού (*Dynamic Programming, DP*), όπως έχουμε προαναφέρει δίνει λύση στο πρόβλημα της ανάθεσης δεδομένων με διαφορετική πιθανότητα ζήτησης σε πολλαπλά κανάλια **ίδιου** εύρους ζώνης. Ο αλγόριθμος δουλεύει έτσι ούτως ώστε τα δεδομένα με μεγαλύτερη πιθανότητα ζήτησης να βρίσκονται σε κανάλια με λιγότερα αντικείμενα. Αυτό σημαίνει ότι τα συγκεκριμένα κανάλια θα έχουν μικρότερο **broadcasting cycle** από ότι τα άλλα και επομένως κάθε αντικείμενο που εκπέμπεται μέσα από αυτά τα κανάλια θα εκπέμπεται με μεγαλύτερη συχνότητα από ότι αντικείμενα που βρίσκονται σε κανάλια που εκπέμπουν περισσότερα αντικείμενα. Αυτό έχει σαν αποτέλεσμα την μείωση του χρόνου αναμονής από τη μεριά του χρήστη καθώς με αυτόν τον τρόπο τα πιο δημοφιλή αντικείμενα εκπέμπονται πιο συχνά.

5.1.1 Ορισμός Μεταβλητών του Συστήματος

Σε αυτή την ενότητα ορίζουμε τις μεταβλητές του συστήματος που θα μελετήσουμε και θα προσομοιώσουμε και όλες τις παραδοχές που θα κάνουμε.

Θεωρούμε ένα σύνολο από K πανομοιότυπα κανάλια και ένα σύνολο $D = \{d_1, d_2, \dots, d_N\}$ από δεδομένα-αντικείμενα. Κάθε δεδομένο d_i χαρακτηρίζεται από μία *πιθανότητα* p_i και ένα *μήκος* z_i με $1 \leq i \leq N$. Η πιθανότητα p_i εκφράζει την πιθανότητα ζήτησης του δεδομένου d_i που έχει ζητηθεί από τους clients και δεν μεταβάλλεται με το πέρασμα του χρόνου.

Προφανώς ισχύει ότι η κατανομή των πιθανοτήτων ζήτησης στα δεδομένα κάθε φορά είναι τέτοια ώστε να ισχύει $\sum_{i=1}^N p_i = 1$. Το μήκος z_i είναι ένας ακέραιος αριθμός που μετράει πόσες χρονικές μονάδες (ή *tics*) χρειάζονται ώστε να μεταδοθεί το αντικείμενο d_i σε ένα οποιοδήποτε κανάλι. Στην περίπτωση που όλα τα μήκη των δεδομένων είναι τα ίδια ισχύει δηλαδή $z_i = z$ για κάθε $1 \leq i \leq N$ τότε τα μήκη αυτά ονομάζονται *ομοιόμορφα (uniform lengths)* και θεωρούμε ότι είναι ίσα με τη μονάδα $z = 1$. Σε αντίθετη περίπτωση όπου τα μήκη δεν είναι ίδια τότε τα μήκη ονομάζονται *ανομοιόμορφα (non-uniform lengths)*.

Τα αντικείμενα αυτά πρέπει να διαμοιραστούν σε K σύνολα (groups) G_1, G_2, \dots, G_K . Το group G_j περιέχει τα δεδομένα που ανατίθενται στο κανάλι j με $1 \leq j \leq K$. Η πληθικότητα του συνόλου G_j δηλώνεται ως N_j ενώ το άθροισμα των μηκών των αντικειμένων του συνόλου δηλώνεται ως Z_j και επομένως ισχύει ότι $Z_j = \sum_{d_i \in G_j} z_i$. Πρέπει να παρατηρήσουμε ότι από τη στιγμή που τα αντικείμενα ενός οποιοδήποτε συνόλου G_j εκπέμπονται

κυκλικά ,με βάση τον ορισμό του flat scheduling ,τότε το Z_j δηλώνει την περίοδο του προγράμματος εκπομπής στο κανάλι j .Είναι επίσης προφανές ότι στην *uniform* περίπτωση $Z_j = N_j$ για κάθε $1 \leq j \leq K$

Στη μελέτη που κάναμε θεωρούμε το πρόβλημα στο οποίο όλα τα δεδομένα που εκπέμπονται έχουν το ίδιο μήκος αναφερόμαστε λοιπόν στην *uniform* περίπτωση του προβλήματος.

5.1.2 Ορισμός Μέσης Αναμενόμενης Καθυστέρησης

Δεδομένου ότι το αντικείμενο d_i ανατίθεται στο κανάλι j και θεωρώντας ότι οι clients μπορούν να ξεκινήσουν να ακούν οποιαδήποτε χρονική στιγμή με την ίδια πιθανότητα τότε η αναμενόμενη καθυστέρηση για τον client (*client expected delay*) ώστε αυτός να παραλάβει το αντικείμενο d_i είναι ίση με τη μισή περίοδο δηλαδή ίση με $\frac{Z_j}{2}$.Έτσι θεωρώντας ότι ένας μηχανισμός ευρετηρίασης (indexing) επιτρέπει στους clients να γνωρίζουν προκαταβολικά τα περιεχόμενα των καναλιών η **μέση αναμενόμενη καθυστέρηση** (*average expected delay (AED)*) για όλα τα κανάλια θα είναι :

$$AED = \frac{1}{2} \sum_{j=1}^K Z_j \sum_{d_i \in G_j} p_i \quad (1)$$

Έτσι δοθέντων K καναλιών και ενός συνόλου D, N αντικειμένων όπου κάθε αντικείμενο d_i συνοδεύεται από κάποια πιθανότητα p_i και από το μήκος του z_i το *K-uniform Allocation Problem* έχει να κάνει με τη διαμέριση (partitioning) του συνόλου D σε K υποσύνολα (groups) G_1, G_2, \dots, G_K ξένα μεταξύ τους έτσι ώστε να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

AED που δίνεται στην (1). Θα πρέπει να ξαναθυμηθούμε ότι σε αυτό το πρόβλημα (*K-uniform Allocation Problem*) ισχύει ότι $Z_j = N_j$.

5.1.3 Δύο Χρήσιμα Πορίσματα

Πόρισμα 1: Έστω G_h και G_j είναι δύο groups που αναφέρονται σε μία βέλτιστη λύση, ανάθεση. Έστω και ότι d_i και d_k είναι αντικείμενα με $d_i \in G_h$ και $d_k \in G_j$. Αν $N_h < N_j$ τότε $p_i \geq p_k$. Κατά τον ίδιο τρόπο αν $p_i > p_k$ τότε ισχύει $N_h \leq N_j$.

Με λίγα λόγια το παραπάνω πόρισμα λέει ότι τα πιο δημοφιλή αντικείμενα ανατίθενται στα λιγότερο “φορτωμένα κανάλια” ώστε να εμπέμπονται πιο συχνά.

Πόρισμα 2: Έστω d_1, d_2, \dots, d_N τα N αντικείμενα του συνόλου D με $p_i \geq p_k$ για κάθε $i < k$. Υπάρχει μία βέλτιστη λύση για την ανάθεση των αντικειμένων μεταξύ των K διαθέσιμων groups G_1, G_2, \dots, G_K όπου κάθε σύνολο αποτελείται από διαδοχικά αντικείμενα

Από δω και πέρα κατά τον ίδιο τρόπο θεωρούμε ότι τα δεδομένα είναι ταξινομημένα κατά φθίνουσα τιμή πιθανότητας και οι βέλτιστες λύσεις – αναθέσεις θα πρέπει να αναζητώνται μέσα στην κλάση κάποιας τμηματοποίησης (segmentation).

Με τον όρο segmentation εννοούμε μία διαμέριση της μορφής G_1, \dots, G_K τέτοια ώστε αν $d_i \in G_j$ και $d_k \in G_h$ τότε $d_i \in G_h$ κάθε φορά που ισχύει $i \leq h \leq k$.

Επίσης μια τμηματοποίηση της μορφής

$\underbrace{d_1, \dots, d_{B_1}}_{G_1}, \underbrace{d_{B_1+1}, \dots, d_{B_2}}_{G_2}, \dots, \underbrace{d_{B_{K-1}+1}, \dots, d_N}_{G_K}$ μπορεί για περισσότερη

ευκολία να συμβολιστεί ως μία πλειάδα με $K-1$ στοιχεία ($(K-1)$ -tuple) $(B_1, B_2, \dots, B_{K-1})$ όπου το B_j είναι ο δείκτης του τελευταίου αντικειμένου που ανήκει στο group G_j . Παρατηρούμε ότι από τη δήλωση της πλειάδας λείπει ο δείκτης B_K ο οποίος παραλείπεται σκοπίμως γιατί σε κάθε περίπτωση αυτός είναι ίσος με N . Επίσης από δω και στο εξής το B_{K-1} θα αναφέρεται ως το τελικό άκρο (**final border**) της λύσης. Τέλος η πληθικότητα του G_j δηλαδή ο αριθμός N_j των αντικειμένων του group G_j είναι ίση με $N_j = B_j - B_{j-1}$ όπου $B_0 = 0$ και $B_K = N$.

Με τον παράνω τρόπο καταφέρνουμε να αναφερόμαστε σε κάποια τμηματοποίηση μέσω μίας μόνο πλειάδας κάτι πολύ πιο απλό και εύχρηστο.

5.1.4 Ορισμός DP Αλγορίθμου

Ο αλγόριθμος δυναμικού προγραμματισμού (DP) λέει το εξής:

Δοθέντων δύο οποιονδήποτε ακεραίων $n \leq N$ και $k \leq K$, ορίζουμε ως $OPT_{n,k}$ την βέλτιστη (*optimal*) λύση για την ομαδοποίηση των αντικειμένων d_1, d_2, \dots, d_N , που είναι ταξινομημένα σε φθίνουσα σειρά με βάση τις πιθανότητες ($p_i \geq p_{i+1} \forall i \in [1, N]$), σε k ομάδες και έστω ότι $opt_{n,k}$ είναι το αντίστοιχο κόστος. Θεωρούμε ακόμα ως $C_{i,h}$ το κόστος για την ανάθεση των διαδοχικών αντικειμένων d_i, \dots, d_h σε ένα group – κανάλι. Με βάση τα παραπάνω και τον ορισμό της μέσης αναμενόμενης

καθυστερήσης (AED βλ. τυπο (1)) το κόστος αυτό θα είναι ίσο με

$$C_{i,h} = \frac{1}{2}(h-i+1) \sum_{q=i}^h p_q \cdot \text{Ως εκ' τούτου } opt_{n,l} = C_{l,n} \text{ για κάθε } n.$$

Όταν το $k > l$ η ακόλουθη επαναληπτική εξίσωση υπολογίζει το κόστος της βέλτιστης λύσης.

$$opt_{n,k} = \min_{l \in \{1,2,\dots,n-1\}} \{opt_{l,k-1} + C_{l+1,n}\} \quad (2)$$

Ο DP αλγόριθμος είναι η υλοποίηση δυναμικού προγραμματισμού μέσα από την επαναληπτική εξίσωση (2). Προκειμένου να βρούμε την βέλτιστη λύση $OPT_{n,k}$ θεωρούμε τον πίνακα M διαστάσεων $K \times N$ με $M_{k,n} =$

$opt_{n,k}$. Τα στοιχεία του πίνακα M υπολογίζονται γραμμή γραμμή με βάση

την σχέση (2). Είναι προφανές ότι ο $M_{K,N}$ περιέχει το κόστος μιας βέλτιστης λύσης για το ***K-Uniform Allocation Problem***. Προκειμένου

να δημιουργήσουμε μια βέλτιστη ανάθεση ένας δεύτερος πίνακας F

χρησιμοποιείται ώστε να αποθηκεύεται πληροφορία σχετικά με τα τελικά άκρα (**final borders**) των segmentations που σχετίζονται με τα στοιχεία

του πίνακα M . Στην εξίσωση (2) η τιμή του l που ελαχιστοποιεί το

άθροισμα στα δεξιά είναι το τελικό άκρο της λύσης $OPT_{n,k}$ και

αποθηκεύεται στο στοιχείο $F_{k,n}$. Με αυτό τον τρόπο η βέλτιστη

τμηματοποίηση (*segmentation*) δίνεται από τη βέλτιστη λύση

$OPT_{N,K} = (B_1, B_2, \dots, B_{K-1})$ όπου ξεκινώντας από το ότι $B_K = N$ η τιμή

του B_k είναι ίση με $F_{k+1, B_{k+1}}$ για $k=1, \dots, K-1$.

Εκτιμώντας την πολυπλοκότητα του αλγορίθμου παρατηρούμε ότι χρειάζονται

$O(N)$ συγκρίσεις για να συμπληρωθεί ένα στοιχείο του πίνακα M που σημαίνει ότι

χρειάζονται $O(N^2)$ συγκρίσεις για να συμπληρωθεί μία σειρά. Έτσι από τη στιγμή

που έχουμε K γραμμές η πολυπλοκότητα του αλγορίθμου DP είναι $O(N^2K)$.

5.1.5 Σχηματική Απεικόνιση της Λύσης

Αναλύοντας την εξίσωση (2) έχουμε το εξής:

$$opt_{n,k} = \min_{l \in \{1,2,\dots,n-1\}} \{opt_{l,k-1} + C_{l+1,n}\} = \min \begin{Bmatrix} opt_{1,k-1} + C_{2,n} \\ opt_{2,k-1} + C_{3,n} \\ \vdots \\ opt_{n-1,k-1} + C_{n,n} \end{Bmatrix} \quad (2)$$

Ο υπολογισμός του βέλτιστου κόστους μπορεί να υπολογιστεί υλοποιώντας μία αντίστοιχη συνάρτηση $opt(n,k)$. Όπως παρατηρούμε η συνάρτηση $opt(n,k)$ καλείται τον εαυτό της αναδρομικά προκειμένου να υπολογίσει το κόστος στις βέλτιστης λύσης. Η τιμή της συνάρτησης $opt(n,k)$ μετά τον υπολογισμό αποθηκεύεται στο στοιχείο $M_{k,n}$ του πίνακα M . Με αυτό τον τρόπο ο πίνακας συμπληρώνεται γραμμή γραμμή όπως έχουμε προαναφέρει. Έτσι κάθε φορά που πρέπει να υπολογιστεί το στοιχείο $M_{k,n}$ του πίνακα M , δηλαδή η τιμή της συνάρτησης $opt(n,k)$ με $n \leq N$ και $k \leq K$, όπως παρατηρούμε από τη σχέση (2) στον υπολογισμό αυτό συμμετέχουν όλες οι προηγούμενες τιμές του M που βρίσκονται πάνω και δεξιά από το στοιχείο $M_{k,n}$ του πίνακα. Σχηματική απεικόνιση του υπολογισμού της λύσης φαίνεται στην παρακάτω εικόνα

$M_{1,1}=\text{opt}_{1,1}=C_{1,1}$	$M_{1,2}=\text{opt}_{2,1}=C_{1,2}$	$M_{1,N}=\text{opt}_{N,1}=C_{1,N}$
$M_{2,1}=\text{opt}_{1,2}$	$M_{2,2}=\text{opt}_{2,2}$	$M_{1,N}=\text{opt}_{N,1}=C_{1,N}$
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	$M_{k,n}=\text{opt}_{k,n}$	⋮
⋮	⋮	⋮	⋮	⋮
$M_{1,1}=\text{opt}_{1,1}=C_{1,1}$	$M_{1,2}=\text{opt}_{2,1}=C_{1,2}$	$M_{1,N}=\text{opt}_{N,1}=C_{1,N}$

Σχηματική απεικόνιση Υπολογισμού Στοιχείου $M(k,n)$

Εικόνα 5.15

5.2 Επέκταση Αλγορίθμου DP για κανάλια μεταβλητού bandwidth (Variant Bandwidth-DP ή VB-DP)

Αντικειμενικός στόχος της παρούσας εργασίας ήταν όπως προαναφέραμε η επίλυση του προβλήματος της ανάθεσης δεδομένων σε πολλαπλά κανάλια **μεταβλητού bandwidth**. Έτσι στηριζόμενοι στην λειτουργία του απλού αλγορίθμου δυναμικού προγραμματισμού προσπαθήσαμε να επεκτείνουμε τον εν λόγω αλγόριθμο ώστε να λύνει το πρόβλημα στην περίπτωση που έχουμε κανάλια με διαφορετικό **bandwidth**.

5.2.1 Ορισμός Επιπλέον Μεταβλητών Συστήματος

Στο συγκεκριμένο πρόβλημα πρέπει να ενσωματώσουμε συν τοις άλλοις την επιπλέον πληροφορία που σχετίζεται με το *bandwidth* των καναλιών. Έτσι πέρα από τις μεταβλητές που έχουμε ορίσει στην ενότητα 5.1.1 ορίζουμε επιπλέον την μεταβλητή $b_j, j \leq K$ που δηλώνει την τιμή του διαθέσιμου εύρους ζώνης του καναλιού j .

Θα πρέπει σε αυτό το σημείο να επισημάνουμε το εξής. Από τη στιγμή που το **bandwidth** των καναλιών είναι διαφορετικό από το ένα στο άλλο, το μήκος του χρονικής μονάδας (**tic** ή **time slot**) σε κάθε κανάλι θα είναι διαφορετικό. Με άλλα λόγια από τη στιγμή που το **bandwidth** των καναλιών είναι διαφορετικό, ένα αντικείμενο (**data item**) σε κάθε κανάλι έχει διαφορετικό κύκλο εκπομπής, όπου η διάρκεια του κύκλου εκπομπής (**broadcasting cycle**) για ένα αντικείμενο είναι το αντίστοιχο του διαστήματος εκπομπής (**broadcasting interval**) για κάθε αντικείμενο.

5.2.2 Ορισμός Μέσης Αναμενόμενης Καθυστερήσης

Σε αυτή την περίπτωση για τον υπολογισμό της μέσης αναμενόμενης καθυστέρησης πρέπει να λάβουμε υπ' όψιν τον παράγοντα b_j

Λήμμα 1: Δοθέντος ενός καναλιού c_j με bandwidth b_j ο χρόνος αναμονής

για κάθε χρήστη είναι $t_j = \frac{N_j z}{2b_j}$ όπου N_j ο αριθμός των αντικειμένων

που έχουν ανατεθεί στο κανάλι j , z το μέγεθος των αντικειμένων (στην περίπτωση μας αφού αναφερόμαστε σε uniform πρόβλημα $z=1$).

Απόδειξη: Αφού θεωρούμε ότι όλα τα δεδομένα έχουν το ίδιο μέγεθος ($z=1$), το συνολικό μέγεθος των δεδομένων που εκπέμπονται από ένα

κανάλι c_j είναι όσο και το πλήθος τους N_j , έτσι ο broadcast κύκλος σε αυτή την περίπτωση ισούται με το πλήθος $\frac{N_j}{b_j}$. Έτσι αν υποθέσουμε ότι ο χρόνος αναμονής του χρήστη μέχρι αυτός να λάβει το αντικείμενο που ζήτησε είναι ομοιόμορφα κατανομημένος στο διάστημα $[0, \frac{N_j}{b_j}]$ τότε η μέση τιμή του χρόνου αναμονής θα είναι η μέση τιμή της κατανομής δηλαδή $\frac{N_j}{2b_j}$.

Λήμμα 2: Ο μέσος χρόνος αναμονής του χρήστη για όλα τα δεδομένα σε ένα πρόγραμμα εκπομπής δίνεται από τον τύπο:

$$t = \frac{1}{2} \sum_{j=1}^K N_j \sum_{d_i \in G_j} \frac{p_i}{b_j} \quad (3)$$

Ο τύπος αυτός είναι παρόμοιος με τον τύπο (1) μόνο που εδώ για τον υπολογισμό της καθυστέρησης λαμβάνουμε υπ'όψιν το bandwidth b_j

Απόδειξη: Ο μέσος χρόνος αναμονής υπολογίζεται ως η μέση τιμή του χρόνου αναμονής t_j .

$$t = E[t_j] = \sum_{j=1}^K P_j t_j = \sum_{j=1}^K P_j \frac{N_j}{2b_j} = \frac{1}{2} \sum_{j=1}^K \frac{P_j N_j}{b_j}, \quad \text{όπου } P_j \text{ είναι η}$$

αθροιστική πιθανότητα ζήτησης των αντικειμένων που βρίσκονται στο

$$\text{κανάλι } j, \text{ δηλαδή } P_j = \sum_{d_i \in G_j} p_i.$$

Από τη σχέση (3) για ευκολία ορίζουμε μια συνάρτηση κόστους ως εξής:

$$\text{cost} = \sum_{j=1}^K \text{cost}(j) = \sum_{j=1}^K \frac{N_j P_j}{b_j}, \text{ δηλαδή } \text{cost}(j) = \frac{N_j P_j}{b_j} \quad (4)$$

Λήμμα 3: Έστω μια βέλτιστη λύση S , τότε για δύο οποιαδήποτε κανάλια c_i και c_j με bandwidth b_i και b_j αντίστοιχα ισχύει ότι $F_i < F_j$ αν και μόνο αν $b_i < b_j$, όπου $F_i = N_i P_i$, (το F_i ονομάζεται *product value*)

Απόδειξη: Η απόδειξη μας περιλαμβάνει 2 μέρη, αρχικά θα αποδείξουμε το “αν” κομμάτι χρησιμοποιώντας την μέθοδο της “εις άτοπον απαγωγή”. Θεωρούμε μια βέλτιστη λύση στην οποία υπάρχουν δύο κανάλια c_i και c_j με $b_i > b_j$. Αν θεωρήσουμε ότι ισχύει $F_i < F_j$ μπορούμε να ανταλλάξουμε αντικείμενα μεταξύ των καναλιών c_i και c_j .

Αν υπολογίσουμε το συνολικό κόστος, έτσι όπως έχει οριστεί από τη συνάρτηση κόστους από τη σχέση (4), πριν την εναλλαγή των αντικειμένων έχουμε:

$$\text{cost}' = \sum_{k=1, k \neq i, j}^K \frac{N_k P_k}{b_k} + \frac{N_i P_i}{b_i} + \frac{N_j P_j}{b_j} = \sum_{k=1, k \neq i, j}^K \frac{F_k}{b_k} + \frac{F_i}{b_i} + \frac{F_j}{b_j}$$

Από την άλλη το συνολικό κόστος μετά την εναλλαγή των αντικειμένων θα είναι:

$$\text{cost}'' = \sum_{k=1, k \neq i, j}^K \frac{N_k P_k}{b_k} + \frac{N_j P_j}{b_j} + \frac{N_i P_i}{b_i} = \sum_{k=1, k \neq i, j}^K \frac{F_k}{b_k} + \frac{F_j}{b_j} + \frac{F_i}{b_i}$$

Έτσι η μετάβολη που θα έχουμε στο κόστος με την εναλλαγή των αντικειμένων θα είναι:

$$\text{cost}'' - \text{cost}' = \frac{(F_j - F_i)}{b_i} + \frac{(F_i - F_j)}{b_j} = F_j - F_i \left(\frac{1}{b_i} - \frac{1}{b_j} \right) < 0.$$

Καταλήξαμε σε άτοπο! Η διαφορά αυτή έπρεπε να είναι θετική καθώς το κόστος $cost'$ είναι το κόστος της βέλτιστης λύσης και επομένως θα έπρεπε να είναι πάντα μικρότερο από το κόστος οποιασδήποτε άλλης ανάθεσης δηλαδή $cost'' > cost'$.

Το υπόλοιπο της απόδειξης αποδεικνύεται τοιουτοτρόπως.

Το **λήμμα 3** δηλώνει ότι σε μία βέλτιστη λύση τα αντικείμενα σε ένα κανάλι μεγαλύτερου bandwidth σχετίζονται με ένα μεγαλύτερο product value F_x το οποίο εξαρτάται από το πλήθος των αντικειμένων που βρίσκονται στο κανάλι x (N_x) και από την αθροιστική πιθανότητα ζήτησης των αντικειμένων (P_x). Έτσι η ίδια διάταξη που αντιστοιχεί στις τιμές του bandwidth μεταξύ δύο καναλιών ισχύει και για τις αντίστοιχες product value τους και αντίστροφα.

5.2.3 Ορισμός VB-DP Αλγορίθμου

Επεκτείνουμε τον ορισμό του DP-αλγορίθμου ορίζοντας με αυτόν τον τρόπο το νέο αλγόριθμο VB-DP ως εξής :

Δοθέντων δύο οποιονδήποτε ακεραίων $n \leq N$ και $k \leq K$, ορίζουμε ως $OPT_{n,k}$ την βέλτιστη (*optimal*) λύση για την ομαδοποίηση των αντικειμένων d_1, d_2, \dots, d_n , που είναι ταξινομημένα σε φθίνουσα σειρά με βάση τις πιθανότητες ($p_i \geq p_{i+1} \forall i \in [1, N]$), σε k ομάδες και έστω ότι $opt_{n,k}$ είναι το αντίστοιχο κόστος. Θεωρούμε ακόμα ως $C_{i,h,j}$ το κόστος για την ανάθεση των διαδοχικών αντικειμένων d_i, \dots, d_h στο *group* – κανάλι j . Με βάση τα παραπάνω και τον ορισμό της μέσης αναμενόμενης

καθυστέρησης (AED βλ. τυπο (3)) το κόστος αυτό θα είναι ίσο με

$$C_{i,h,j} = \frac{1}{2} \frac{(h-i+1)}{b_j} \sum_{q=i}^h P_q \cdot$$

$$\Omega \text{ εκ' τούτου } opt_{n,l} = \min_{j \in \{1, \dots, K\}} \{C_{1,n,j}\} \text{ για κάθε } n. (5)$$

Εμείς θεωρούμε ότι κατά την εφαρμογή του αλγορίθμου τα κανάλια είναι ήδη ταξινομημένα κατά φθίνουσα τιμή εύρους ζώνης δηλαδή αν $i < j$ τότε $b_i > b_j$ για κάθε και το αντίστροφο $l < i, j < K$.

Αν λάβουμε υπ' όψιν το παραπάνω τότε η σχέση (5) γίνεται

$opt_{n,l} = C_{1,n,l}$ και αυτό γιατί το κανάλι με το μεγαλύτερο bandwidth είναι το πρώτο κανάλι c_l και κατά συνέπεια το b_l είναι αυτό που ελαχιστοποιεί την (5) σε οποιαδήποτε περίπτωση.

Όταν το $k > l$ η ακόλουθη επαναληπτική εξίσωση υπολογίζει το κόστος της βέλτιστης λύσης.

$$opt_{n,k} = \min_{l \in \{1, 2, \dots, n-1\}} \{opt_{l,k-1} + C_{l+1,n,k}\} \quad (6)$$

Όπως και πριν ,προκειμένου να βρούμε την βέλτιστη λύση $OPT_{n,k}$

θεωρούμε τον πίνακα M διαστάσεων $K \times N$ με $M_{k,n} = opt_{n,k}$. Τα στοιχεία του πίνακα M υπολογίζονται γραμμή γραμμή με βάση την σχέση (6).

Είναι προφανές ότι ο $M_{K,N}$ περιέχει το κόστος μίας βέλτιστης λύσης για το ***K-Uniform Allocation Problem***. Προκειμένου να δημιουργήσουμε μια βέλτιστη ανάθεση ένας δεύτερος πίνακας F χρησιμοποιείται ώστε να αποθηκεύεται πληροφορία σχετικά με τα τελικά άκρα των segmentations που σχετίζονται με τα στοιχεία του πίνακα M . Στην εξίσωση (5) η τιμή του l που ελαχιστοποιεί το άθροισμα στα δεξιά είναι το τελικό άκρο της λύσης $OPT_{n,k}$ και αποθηκεύεται στο στοιχείο $F_{k,n}$. Με αυτό τον τρόπο η

βέλτιστη τμηματοποίηση (*segmentation*) δίνεται από τη βέλτιστη λύση $OPT_{N,K} = (B_1, B_2, \dots, B_{K-1})$ όπου ξεκινώντας από το ότι $B_K = N$ η τιμή του B_k είναι ίση με $F_{k+1, B_{k+1}}$ για $k=1, \dots, K-1$

Εκτιμώντας την πολυπλοκότητα του αλγορίθμου VB-DP παρατηρούμε ότι και σε αυτόν χρειάζονται $O(N)$ συγκρίσεις για να συμπληρωθεί ένα στοιχείο του πίνακα M που σημαίνει ότι χρειάζονται $O(N^2)$ συγκρίσεις για να συμπληρωθεί μία σειρά. Έτσι από τη στιγμή που έχουμε K γραμμές η πολυπλοκότητα του αλγορίθμου VB-DP είναι $O(N^2K)$.

Τέλος αφού ο η ανάθεση των αντικειμένων στα κανάλια έχει προκύψει μετά από την εφαρμογή του αλγορίθμου δυναμικού προγραμματισμού, πρέπει να επαληθευτεί ότι μεταξύ των καναλιών και των αντικειμένων ισχύει το **λήμμα 3**

Αυτό περιγράφεται με τον παρακάτω ψευδοκώδικα:

```

n=1;

while(n<K)

{

    if( $F_n < F_{n+1}$ ) /*  $F_i = N_i P_i$  */

    {

        Μετακίνησε το αντικείμενο με την μεγαλύτερη πιθανότητα
ζήτησης από το κανάλι  $c_{n-1}$  στο κανάλι  $c_n$ 

        n=0;

    }

    n++;

```

5.2.4 Σχόλια σχετικά με τον VB-DP Αλγόριθμο

Ο αλγόριθμος αυτός λαμβάνοντας υπ' όψιν του το bandwidth των καναλιών τις πιθανότητες ζήτησης των δεδομένων που έχουμε προς αποστολή δημιουργεί μία διαμέριση των δεδομένων που θέλουμε να εκπεμφθούν από το σύστημά μας με τέτοιο τρόπο ώστε τα δεδομένα με τις υψηλότερες πιθανότητες ζήτησης να τοποθετούνται σε κανάλια με πιο μεγάλο εύρος ζώνης κάτι το οποίο συνεπάγεται ότι ο broadcast κύκλος τους είναι μικρότερος από ότι κανάλια που διαθέτουν μικρότερο bandwidth και άρα η συχνότητα εκπομπής τους είναι μεγαλύτερη. Με αυτό τον τρόπο ο αλγόριθμος δυναμικού προγραμματισμού *VB-DP* προσπαθεί να ελαχιστοποιήσει τον μέσο χρόνο αναμονής για τον κάθε χρήστη.

5.3 Broadcast Αλγόριθμος του Ζυγού(Balance Algorithm)

Ο αλγόριθμος αυτός δίνει μία άλλη λύση στο πρόβλημα τις ανάθεσης δεδομένων προς εκπομπή σε κανάλια μεταβλητού εύρους ζώνης εφαρμόζοντας στα δεδομένα μια τεχνική προσπάθειας εξισορρόπησης λαμβάνοντας υπ'όψιν μια συνάρτηση κόστους. Η λειτουργία του αλγορίθμου μοιάζει αρκετά με τη λειτουργία ενός ζυγού όταν προσπαθεί να ισοροπήσει ,στον οποίο τα δεδομένα που βρίσκονται στη μία μεριά του πρέπει να έχουν λίγο μεγαλύτερο ή τουλάχιστον ίσο “βάρος” με αυτά που βρίσκονται στην άλλη μεριά του ζυγού.Στη δεδομένη περίπτωση το ρόλο του βάρους τον παίζει η τιμή που επιστρέφεται απο μία συνάρτηση κόστους.

5.3.1 Ορισμοί Μεταβλητών και Παραδοχές

Και σε αυτό τον αλγόριθμο ισχύουν οι παραδοχές που έχουμε κάνει και πριν σχετικά με την κατάσταση του συστήματος ότι δηλαδή τα αντικείμενα προς εκπομπή είναι ταξινομημένα κατά φθίνουσα σειρά με βάση την πιθανότητα ζήτησης τους ($p_i \geq p_{i+1} \forall i \in [1, N]$) και ότι τα κανάλια είναι ταξινομημένα κατά φθίνουσα τιμή εύρους ζώνης ($b_i \geq b_{i+1} \forall i \in [1, K]$).

Επίσης από τη στιγμή που στο πρόβλημα μας υπάρχει το στοιχείο της διαφορετικότητας του bandwidth μεταξύ των καναλιών θα ισχύει και το **λήμμα 3** που αναφέρθηκε στην παράγραφο 5.2.3. Τέλος όσον αφορά τα σύμβολα και τις μεταβλητές που χρησιμοποιούμε σε αυτό τον αλγόριθμο είναι ίδια με όσα έχουμε ορίσει παραπάνω στην παράγραφο 5.1.1

5.3.2 Ορισμός Αλγορίθμου Ζυγού(Balance Algorithm)

Ο αλγόριθμος αυτός ξεκινά αναθέτοντας όλα τα δεδομένα ,ταξινομημένα κατα φθίνουσα σειρά, στον κανάλι c_1 που είναι αυτό με το μεγαλύτερο bandwidth. Στη συνέχεια καλείται η συνάρτηση $partition(D_x)$ η οποία δέχεται σαν παράμετρο ένα σύνολο δεδομένων ,στην προκειμένη περίπτωση το σύνολο D_1 των δεδομένων που έχουν ανατεθεί στο c_1 . Η συνάρτηση αυτή διαχωρίζει το σύνολο δεδομένων που δέχεται σαν παράμετρο ,το D_1 δηλαδή ,σε δύο υποσύνολα ξένα μεταξύ τους D_1^* και D_1^{**} ($D_1^* \cap D_1^{**} = \emptyset$ και $D_1^* \cup D_1^{**} = D_1$).Στη συνέχεια το σύνολο D_1 αφαιρείται από το κανάλι c_1 και το σύνολο D_1^* ανατίθεται στο c_1 ενώ το D_1^{**} στο c_2 . Στο επόμενο βήμα το σύνολο των αντικειμένων που βρίσκονται στο κανάλι c_2 , D_2 αποτελεί τη παράμετρο για την κλήση ξανά της συνάρτησης $partition(D_x)$ κ.ο.κ. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να φτάσουμε στον κανάλι c_{K-1} οπότε καλείται η $partition(D_x)$ με παράμετρο D_{K-1} . Με το πέρας όλης αυτής της διαδικασίας έχουν ανατεθεί σε όλα τα κανάλια αντικείμενα.

5.3.2.1 Συνάρτηση *Partition*

Η συνάρτηση $partition(D_x)$ έχει σαν λειτουργία τον διαμερισμό του συνόλου D_x ,που δέχεται σαν όρισμα ,σε δύο επιμέρους υποσύνολα ξένα μεταξύ τους D_x^* , D_x^{**} ,τα οποία επιστρέφει , των οποίων η ένωση μας δίνει το D_x .Για να γίνει ο διαχωρισμός αυτός η συνάρτηση αυτή χρησιμοποιεί τον ορισμό της συνάρτησης κόστους για κάποιο κανάλι όπως αυτή έχει

οριστεί παραπάνω δηλαδή $cost(j) = \frac{N_j P_i}{b_j}$ (σχέση 4, παρ. 5.2.2). Ακόμα

λαμβάνεται υπ'όψιν αυτό που έχει αποδειχθεί στο **λήμμα 3** (παρ 5.2.2) ότι

δηλαδή τα κανάλια με μεγαλύτερο bandwidth b_j έχουν μεγαλύτερη product value F_j από αυτά με μικρότερο bandwidth (δηλαδή $b_i < b_j \Leftrightarrow F_i < F_j$)

Αρχικά λοιπόν όλα τα στοιχεία του συνόλου D_x εισάγονται στο κανάλι c_x . Τη δεδομένη χρονική στιγμή, δεδομένου ότι το πλήθος των στοιχείων του D_x είναι N_x , η συνάρτηση κόστους για το κανάλι c_x έχει την τιμή

$$\text{cost}(x) = \frac{N_x P_x}{b_x}, \text{ ενώ από την άλλη } \text{cost}(x+1) \text{ είναι } 0 \text{ καθώς το } c_{x+1} \text{ είναι}$$

στη δεδομένη φάση άδειο.

Στο επόμενο βήμα εξετάζουμε την ανισότητα $\text{cost}(x) > \text{cost}(x+1)$ (6) και αν η ανισότητα είναι αληθής (στη δεδομένη περίπτωση είναι) μετακινούμε το στοιχείο με την μικρότερη πιθανότητα ζήτησης από το κανάλι c_x στο κανάλι c_{x+1}

Στην συνέχεια ξαναεξετάζουμε την ανισότητα (6) με τις νέες τιμές των $\text{cost}(x)$ και $\text{cost}(x+1)$ και αν η ανισότητα είναι αληθής επαναλαμβάνουμε τη διαδικασία μέχρι να αλλάξει η φορά της ανισότητας. Με αυτό τον τρόπο καθώς επαναλαμβάνεται η διαδικασία αυτή το κανάλι c_{x+1} αποκτά όλο ένα και περισσότερα στοιχεία ενώ αντίθετα ο αριθμός των στοιχείων του c_x μειώνεται σταδιακά, αυτό έχει σαν αποτέλεσμα την σταδιακή αύξηση της τιμής τη συνάρτησης $\text{cost}(x+1)$ και ταυτόχρονα την μείωση της τιμής της συνάρτησης $\text{cost}(x)$. Η λειτουργία της συνάρτησης αυτής μοιάζει με τη λειτουργία ενός ζυγού στον οποίο αρχικά μόνο το ένα μέρος του διαθέτει αντικείμενα και ο ζυγός γέρνει απόλυτα προς τη μεριά που βρίσκονται τα αντικείμενα. Καθώς μετακινούμε στοιχεία από τη μία μεριά του ζυγού προς την άλλη η ισορροπία του ζυγού αρχίζει και αλλάζει και αρχίζει και γέρνει όλο και λιγότερο προς την μεριά που έγερνε αρχικά

Όταν κάποια στιγμή η ανισότητα (6) πάψει να ισχύει δηλαδή $cost(x) < cost(x+1)$ επιστρέφουμε στο c_x το τελευταίο στοιχείο του που είχαμε μεταφέρει στο c_{x+1} έτσι η ανισότητα γίνεται ξανά $cost(x) > cost(x+1)$. Στην ουσία αυτό που επιθυμούμε είναι η συνάρτηση κόστους για το κανάλι c_x να είναι οριακά μεγαλύτερη από αυτή του c_{x+1} .

Πρέπει να επισημάνουμε ότι επειδή μας ενδιαφέρει τα στοιχεία στα κανάλια να είναι πάντα ταξινομημένα κατά φθίνουσα τιμή πιθανότητας πρέπει η απομάκρυνση των στοιχείων από το c_x να είναι FIFO γι'αυτό απομακρύνουμε το τελευταίο στοιχείο του κάθε φορά αυτό δηλαδή που έχει τη μικρότερη πιθανότητα ζήτησης από όλα τα στοιχεία που ανήκουν στο c_x και το κανάλι c_{x+1} γεμίζει και αυτό με FIFO σειρά.

Με λίγα λόγια η συνάρτηση $partition(D_x)$ προσπαθεί να διαχωρίσει το σύνολο εισόδου D_x σε δύο σύνολα. Κατά το διαχωρισμό αυτό υπάρχει ένα στοιχείο p που μπορούμε να το ονομάζουμε στοιχείο διαχωρισμού τέτοιο ώστε στο κανάλι c_x να έχει τα στοιχεία $d_{x,1}, \dots, d_{x,p}$ και το c_{x+1} να έχει τα στοιχεία $d_{x,p+1}, \dots, d_{x,N_x}$

και το οποίο χαρακτηρίζεται ως κρίσιμο σημείο γιατί καθορίζει τη φορά της ανισότητας (6), είναι το στοιχείο που αν ανήκε στο κανάλι c_{x+1} θα άλλαζε τη φορά της ανισότητας δηλαδή θα ίσχυε $cost(x) < cost(x+1)$.

5.3.2.2 Ο Αλγόριθμος σε Ψευδοκώδικα

Αλγόριθμος Balance

Input: Το σύνολο αντικειμένων D , σύνολο καναλιών K με διαφορετικό *bandwidth*.

Output: K ζένα μεταξύ τους υποσύνολα του συνόλου αντικειμένων D , $\{D_i, 1 \leq i \leq K\}$

Begin

1. Ταξινόμηση αντικειμένων του συνόλου D με φθίνουσα σειρά με βάση τις πιθανότητες ζήτησης $p_i, 1 \leq i \leq N$
2. Ταξινόμηση καναλιών με φθίνουσα σειρά με βάση το *bandwidth* των καναλιών $b_j, 1 \leq j \leq K$
3. Τοποθετούμε όλα τα στοιχεία του D στο κανάλι c_1 που είναι και αυτό με το μεγαλύτερο *bandwidth*.
4. **for**($i=1; i < K; i++$)
5. $\{D', D''\} = \text{Partition}(D_i);$
6. Αντικατάσταση του συνόλου D_i από το D' στο κανάλι c_i και ανάθεση των στοιχείων του D'' στο c_{i+1}

end

Συνάτηση Partition(D_x)

Input: Ένα σύνολο αντικειμένων D_x του οποίου τα στοιχεία $d_{x,1}, d_{x,2}, \dots, d_{x,N_x}$ είναι διατεταγμένα με φθίνουσα τιμή πιθανότητας $p_{x,i}$
 $1 \leq i \leq N_x$

Output: δύο ζένα μεταξύ τους υποσύνολα του συνόλου αντικειμένων D_x και D_x'' , όπου $D_x' + D_x'' = D_x$

Begin

1. $D_x' = \{d_{x,1}, d_{x,2}, \dots, d_{x,N_x}\}$

2. **do**

3. $cost(x) = \frac{N_x P_x}{b_x};$

$$cost(x+1) = \frac{N_{x+1} P_{x+1}}{b_{x+1}};$$

Μετακίνησε το στοιχείο d_{x,N_x} από το σύνολο D_x' στο D_x''

$N_x = N_x - 1; N_{x+1} = N_x + 1; /*συμβατικά το N_x και το N_{x+1} είναι οι πληθάρηθμοι των συνόλων D_x' και D_x'' αντίστοιχα*/$

4. **while**($cost(x) > cost(x+1)$)

Μετακίνησε το τελευταίο στοιχείο που μεταφέρθηκε από το c_x στο c_{x+1} πίσω στο c_x

$$N_x = N_x + 1; N_{x+1} = N_{x+1} - 1;$$

Επιστροφή συνόλων D_x' και D_x''

end

5.3.3 Παράδειγμα Εφαρμογής Αλγορίθμου

Έστω ότι στο σύστημα μας έχουμε προς εκπομπή $N=6$ σε σύνολο δεδομένα και ότι τα διαθέσιμα κανάλια είναι $K=3$. Επίσης οι πιθανότητες των δεδομένων και το bandwidth των καναλιών φαίνονται παρακάτω:

1. Data Probabilities

p1	p2	p3	p4	p5	p6
0.4	0.3	0.1	0.09	0.07	0.04

2. Channel Bandwidth

b1	b2	b3
12	6	3

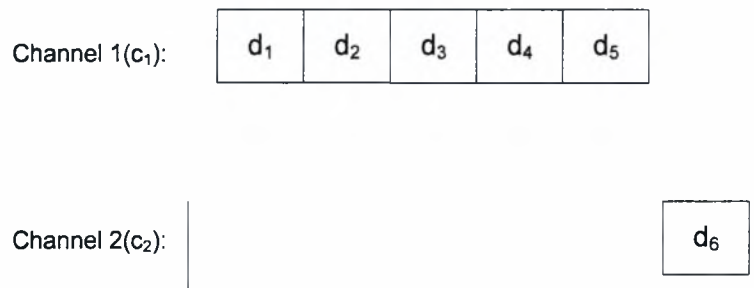
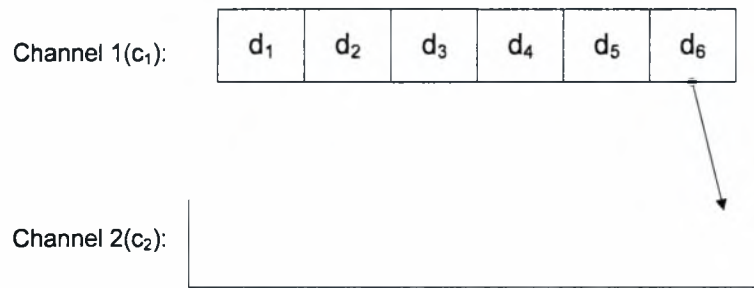
Αρχικά όλα τα αντικείμενα ανατίθενται στο κανάλι c_1 και τα υπόλοιπα κανάλια είναι άδεια.

Channel 1(c_1):

d_1	d_2	d_3	d_4	d_5	d_6
-------	-------	-------	-------	-------	-------

Εικόνα 5.3.3.1

Στην συνέχεια μετακινούμε το στοιχείο με τη μικρότερη πιθανότητα στο επόμενο κανάλι δηλαδή το d_6 το τοποθετούμε στο c_2 και συγκρίνουμε τα κόστη που προφανώς είναι $cost(1) > cost(2)$



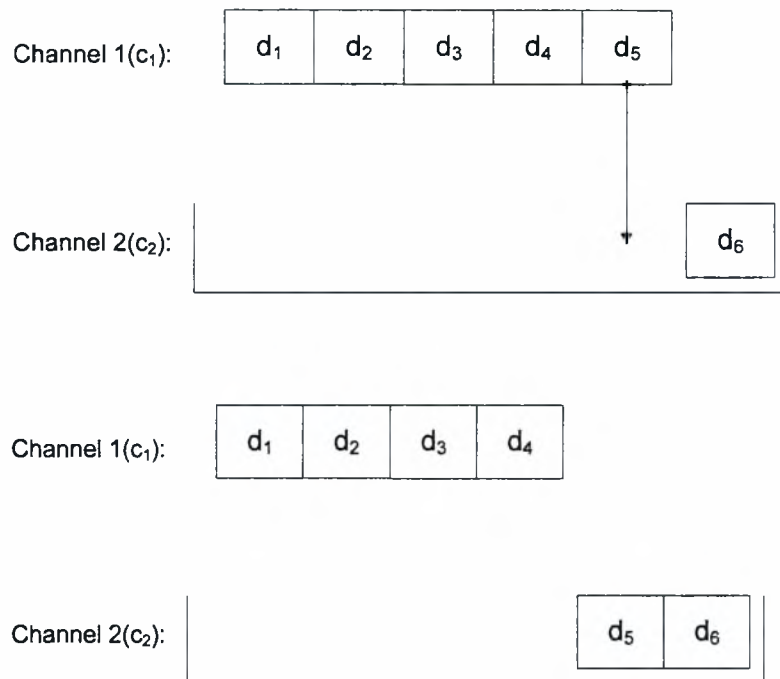
Εικόνα 5.3.3.2

Υπολογίζουμε τώρα τα κόστη που προκύπτουν από την ανάθεση του d_6 στο c_2

$$\text{cost}(1) = \frac{5 * (0.4 + 0.3 + 0.1 + 0.09 + 0.07)}{12} = 0.4$$

$$\text{cost}(2) = \frac{1 * 0.04}{6} = 0.0067$$

$\text{cost}(1) > \text{cost}(2)$, άρα μετακινούμε και το στοιχείο d_5 στο c_2



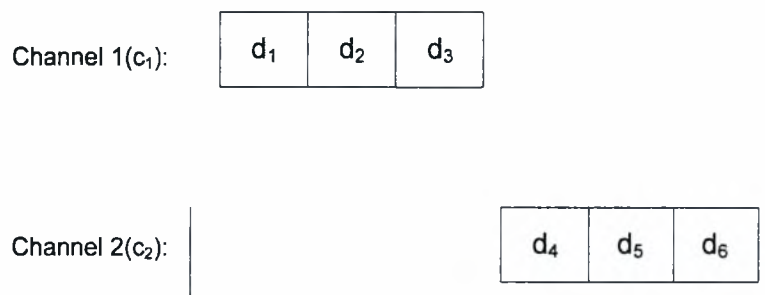
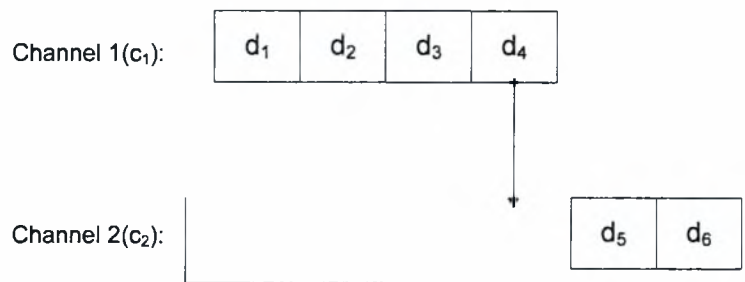
Εικόνα 5.3.3.3

Υπολογίζουμε ξανά τα κόστη που προκύπτουν από την ανάθεση του d_5 στο c_2

$$\text{cost}(1) = \frac{4 * (0.4 + 0.3 + 0.1 + 0.09)}{12} = 0.296$$

$$\text{cost}(2) = \frac{2 * (0.07 + 0.04)}{6} = 0.037$$

$\text{cost}(1) > \text{cost}(2)$, άρα μετακινούμε και το στοιχείο d_4 στο c_2



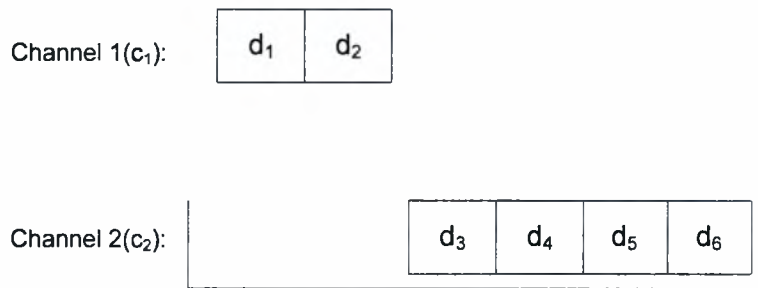
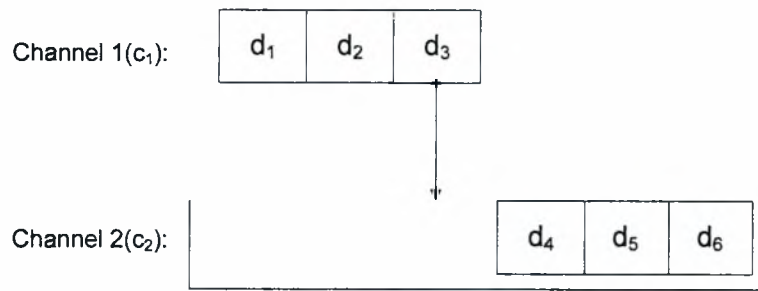
Εικόνα 5.3.3.4

Υπολογίζουμε ξανά τα κόστη που προκύπτουν μετά την ανάθεση του d_4 στο c_2

$$cost(1) = \frac{3 * (0.4 + 0.3 + 0.1)}{12} = 0.2$$

$$cost(2) = \frac{3 * (0.09 + 0.07 + 0.04)}{6} = 0.1$$

Και πάλι $cost(1) > cost(2)$, άρα μετακινούμε και το στοιχείο d_3 στο c_2



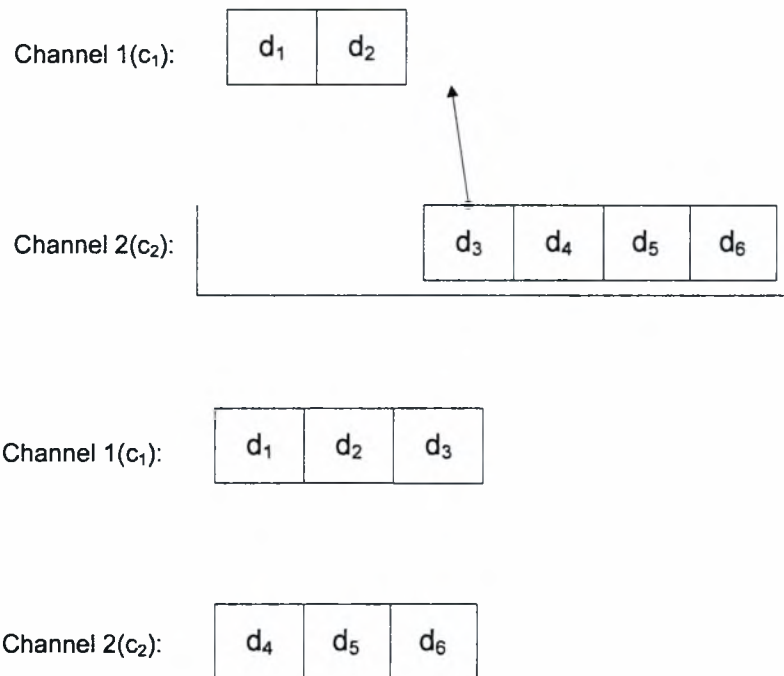
Εικόνα 5.3.3.5

Υπολογίζουμε ξανά τα κόστη που προκύπτουν μετά την ανάθεση του d_3 στο c_2

$$\text{cost}(1) = \frac{2 * (0.4 + 0.3)}{12} = 0.12$$

$$\text{cost}(2) = \frac{4 * (0.1 + 0.09 + 0.07 + 0.04)}{6} = 0.2$$

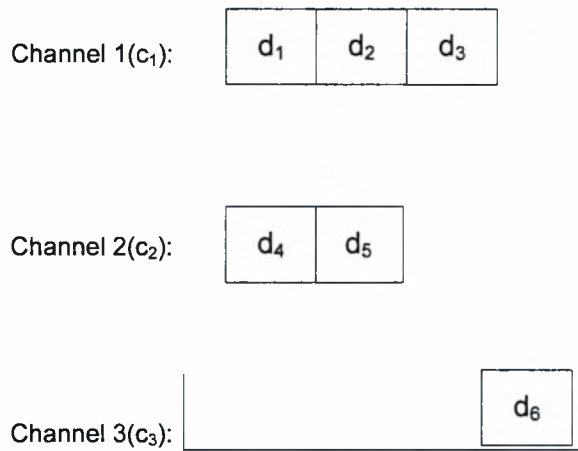
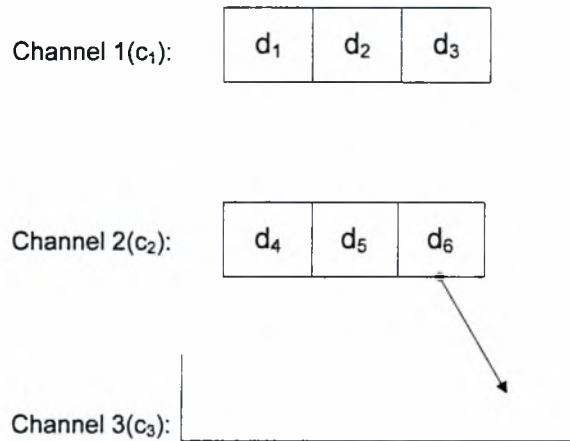
Σε αυτό το σημείο βλέπουμε ότι $\text{cost}(1) < \text{cost}(2)$ και γι'αυτό επιστρέφουμε το τελευταίο στοιχείο που μεταφέραμε πίσω στο c_1



Εικόνα 5.3.3.6

Σε αυτό το στάδιο έχουμε τελειώσει με τη διαμέριση των στοιχείων που ανήκουν στο κανάλι c_1 . Στο επόμενο βήμα συνεχίζουμε με το κανάλι c_2 , καλώντας την *partition*(D_2)

Αρχικά μεταφέρουμε το στοιχείο d_6 στο c_3



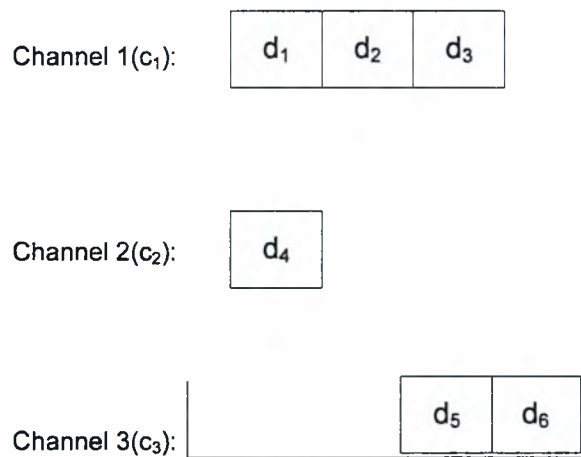
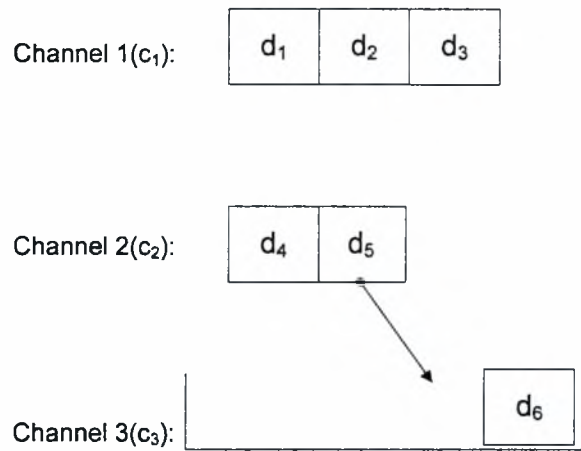
Εικόνα 5.3.3.7

Υπολογίζοντας τα κόστη $cost(2)$, $cost(3)$ έχουμε

$$cost(2) = \frac{2 * (0.09 + 0.07)}{6} = 0.053$$

$$cost(3) = \frac{1 * 0.04}{3} = 0.013$$

Αφού $cost(2) > cost(3)$ μεταφέρουμε το στοιχείο d_5 στο c_3



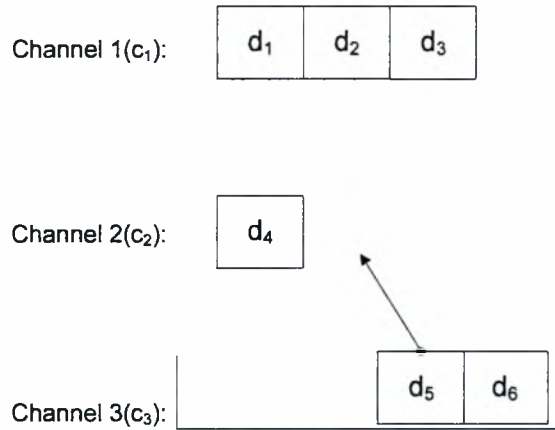
Εικόνα 5.3.3.8

Υπολογίζοντας τα κόστη $cost(2)$, $cost(3)$ έχουμε

$$cost(2) = \frac{1 * (0.09)}{6} = 0.015$$

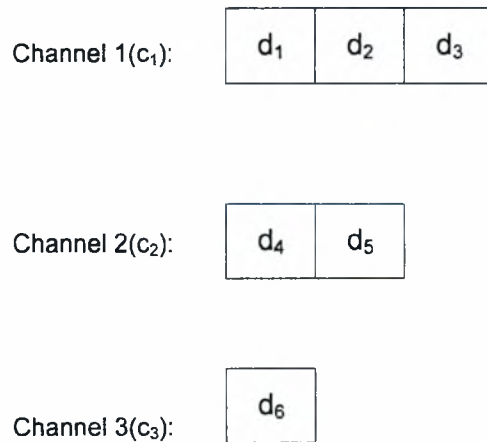
$$cost(3) = \frac{2 * (0.07 + 0.04)}{3} = 0.073$$

Σε αυτό το σημείο βλέπουμε ότι $cost(2) < cost(3)$ και γι'αυτό επιστρέφουμε το τελευταίο στοιχείο που μεταφέραμε το d_5 δηλαδή πίσω στο c_1



Εικόνα 5.3.3.9

Έτσι η τελική ανάθεση των αντικειμένων στα κανάλια μετά το πέρας του αλγορίθμου είναι η εξής.



Εικόνα 5.3.3.10

Κεφάλαιο 6

Απόδοση Αλγορίθμων

6.1 Σχεδιασμός Προσομοίωσης

Σε αυτό το σημείο πρέπει να αναφέρουμε τις παραμέτρους και τις μεταβλητές του συστήματος προσομοίωσης που θεωρήσαμε προκειμένου να προσομοιώσουμε την εκπομπή δεδομένων με βάση τους παραπάνω αλγορίθμους, να μετρήσουμε την απόδοση καθενός και να τους συγκρίνουμε με βάση τα αποτελέσματα των προσομοιώσεων που προέκυψαν.

6.1.1 Σύστημα και Παράμετροι Προσομοίωσης

Σε αυτό το μέρος παρουσιάζουμε το πρόγραμμα προσομοίωσης και τις παραδοχές που έχουμε κάνει:

1. Το πρόγραμμα αρχικά παράγει N σε πλήθος τιμές πιθανότητας που αντιστοιχούν σε κάθε αντικείμενο. Οι πιθανότητες αυτές είναι οι πιθανότητες ζήτησης των αντικειμένων από τους χρήστες του συστήματος και ακολουθούν τη *zipfian* κατανομή η οποία χαρακτηρίζεται από μία συγκεκριμένη μεταβλητή που χαρακτηρίζει το *skewness*, που μας δείχνει το μέτρο της ασυμμετρίας της κατανομής πιθανότητας.

2. Στη συνέχεια με βάση τον αριθμό των δεδομένων N , τον αριθμό των καναλιών K αλλά και τις πιθανότητες που έχουν προκύψει από το προηγούμενο στάδιο εφαρμόζεται ένας από τους δύο αλγόριθμους που προαναφέραμε και προκύπτει η βέλτιστη ανάθεση των δεδομένων ανάλογα με το ποιος αλγόριθμος έχει εφαρμοστεί κάθε φορά
3. Στο τρίτο στάδιο γίνεται η προσομοίωση της εκπομπής των δεδομένων και για το σύστημα μας έχουμε ορίσει χρόνο λειτουργίας του συστήματος μετάδοσης **10000** χρονομονάδες (**timeslots**). Το σύστημα εκπομπής λειτουργεί ως εξής:
 - Ο client αποφασίζει σε ποίο χρόνο θα ξυπνήσει με βάση κάποια εκθετική κατανομή
 - Ο client αποφασίζει ποίο δεδομένο θα ζητήσει με βάση κάποια zipfian κατανομή και ταυτόχρονα στέλει την αίτηση του στο server.
 - Αφού ο server λάβει την αίτηση ο client περιμένει μέχρι ο server να του στείλει πίσω το αντικείμενο που ζήτησε. Το πρόγραμμα υπολογίζει το χρόνο καθυστέρησης του client μέχρις ότου το αντικείμενο εκπνευθεί από τον server.
 - Στη συνέχεια ο client αποφασίζει αν θα κοιμηθεί ή όχι ή θα κάνει αίτηση για άλλο αντικείμενο με βάση την διωνυμική κατανομή
 - Η ίδια διαδικασία επαναλαμβάνεται μέχρις ότου εξαντληθεί ο χρόνος προσομοίωσης
 - Μετά το πέρας της προσομοίωσης υπολογίζεται ο συνολικός χρόνος αναμονής για κάποιον χρήστη.

6.2 Προσομοιώσεις και Αποτελέσματα

Σε αυτό το σημείο αναφέρουμε τις προσομοιώσεις και τα αποτελέσματα που προέκυψαν. Αυτό που αποτελεί κύριο στοιχείο της μελέτης μας και μας έχει απασχολήσει καθόλη τη διάρκεια της εργασίας είναι η ελαχιστοποίηση του χρόνου αναμονής του χρήστη από τη στιγμή που κάνει την αίτηση προς τον server για κάποια δεδομένα μέχρι και την στιγμή που ο χρήστης παραλάβει τα δεδομένα που ζήτησε.

Αντιλαμβάνεται κανείς ότι το κύριο στοιχείο των μετρήσεων είναι πώς επηρεάζεται ο χρόνος αυτός κατά τη διάρκεια που οι παράμετροι του συστήματος μεταβάλλονται.

6.2.1 Παράμετροι που Μεταβάλλονται

Οι παράμετροι που θεωρούμε ότι μεταβάλλονται στο σύστημα μας είναι οι εξής:

1. Ο αριθμός των δεδομένων προς μετάδοση N
2. Ο αριθμός των πολλαπλών καναλιών K
3. Μεταβλητότητα του bandwidth a . Ο συντελεστής αυτός μας δείχνει την κλιμάκωση του bandwidth μεταξύ των καναλιών, δηλαδή $b_{j+1} = a * b_j, 1 \leq j \leq K$ και δεδομένου ότι τα κανάλια είναι ταξινομημένα κατά φθίνουσα τιμή bandwidth θα είναι $a < 0$.
4. Παράμετρος της zipfian συνάρτησης θ (*skew*). Ο τύπος της zipfian συνάρτησης που μας δίνει της πιθανότητες των δεδομένων

$$\text{είναι } p_i(\theta) = \frac{\left(\frac{1}{i}\right)^\theta}{\sum_{i=1}^N \left(\frac{1}{i}\right)^\theta}, 1 \leq i \leq N. \text{ Σημειώνουμε ότι όταν } \theta=0$$

έχουμε ομοιόμορφη κατανομή καθώς όλες οι πιθανότητες είναι ίδιες .

6.2.2 Μετρήσεις , Επιδόσεις και Σύγκριση Αλγορίθμων

Στις μετρήσεις που κάνουμε θεωρούμε όλες τις παραπάνω παραμέτρους σταθερές εκτός από μία την οποία μεταβάλλουμε παρατηρώντας πώς επηρεάζεται ο μέσος χρόνος αναμονής και βλέποντας κάθε φορά ποιός από τους δύο αλγορίθμους είναι πιο αποδοτικός.

Όλες οι μετρήσεις που κάναμε και η σχηματική απεικόνιση των αποτελεσμάτων φαίνεται παρακάτω:

1. Μέσος χρόνος αναμονής vs αριθμός των δεδομένων

Σταθερές Παράμετροι:

$$\theta=0.8, K=5, a=\frac{3}{5}=0.6, b_1=50\text{kbps}$$

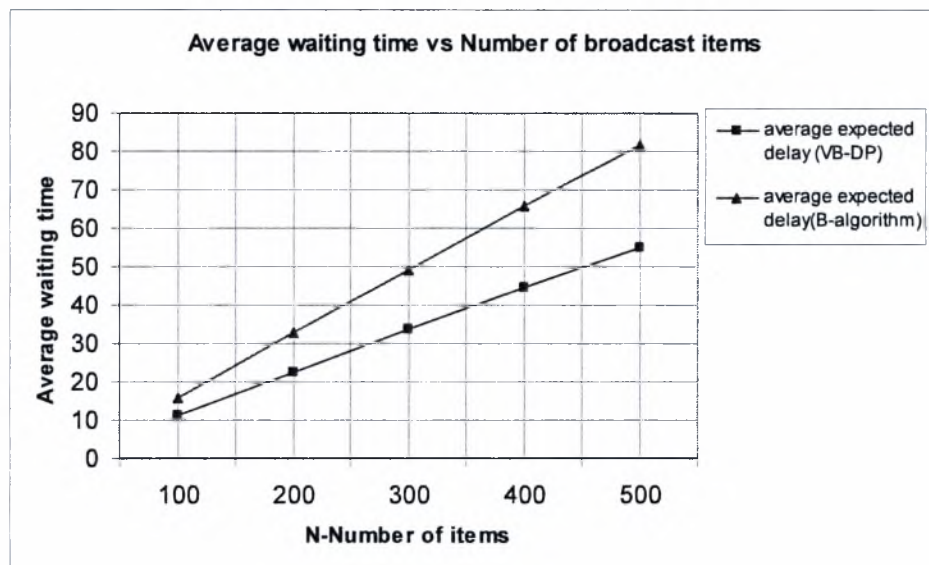
Με εφαρμογή του τύπου $b_{j+1}=a*b_j$ βρίσκουμε το bandwidth που αντιστοιχεί σε κάθε κανάλι.

$$b_1=50, b_2=30, b_3=18, b_4=10.8, b_5=6.48$$

Μεταβλητή παράμετρος:

Θεωρούμε ότι ο αριθμός των δεδομένων προς εκπομπή μεταβάλλεται και χρησιμοποιούμε τις τιμές

$$N=(100, 200, 300, 400, 500)$$



Εικόνα 6.2.2.1

2. Μέσος χρόνος αναμονής vs με τον αριθμό των καναλιών

Σταθερές Παράμετροι:

$$\theta=0.8, N=120, a=\frac{4}{5}=0.8, b_1=50\text{kbps}$$

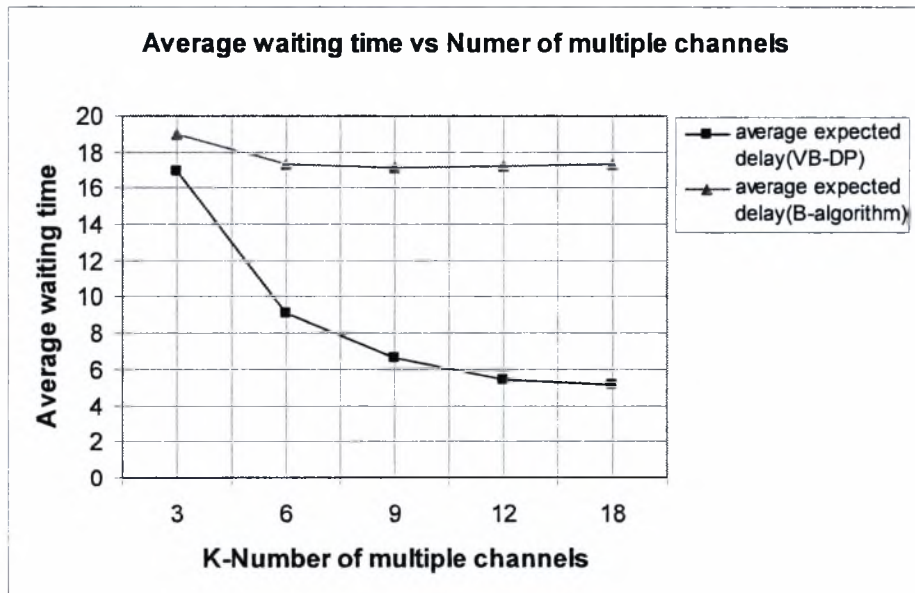
Με εφαρμογή του τύπου $b_{j+1}=a*b_j$ βρίσκουμε το bandwidth που αντιστοιχεί σε κάθε κανάλι.

$$b_1=50, b_2=40, b_3=32, b_4=25.6, b_5=20.48, b_6=16.3840, \\ b_7=13.1072, b_8=10.4857, b_9=8.3886, b_{10}=6.7108, \\ b_{11}=5.3687, b_{12}=4.2949, b_{13}=3.4359, b_{14}=2.7487, b_{15}=2.1990, \\ b_{16}=1.7592, b_{17}=1.4073, b_{18}=1.1259$$

Μεταβλητή παράμετρος:

Θεωρούμε ότι μεταβάλλεται ο αριθμός των πολλαπλών καναλιών και χρησιμοποιούμε τις τιμές:

$$K=(3,6,9,12,18)$$



Εικόνα 6.2.2.2

3. Μέσος χρόνος αναμονής vs μεταβλητότητα bandwidth

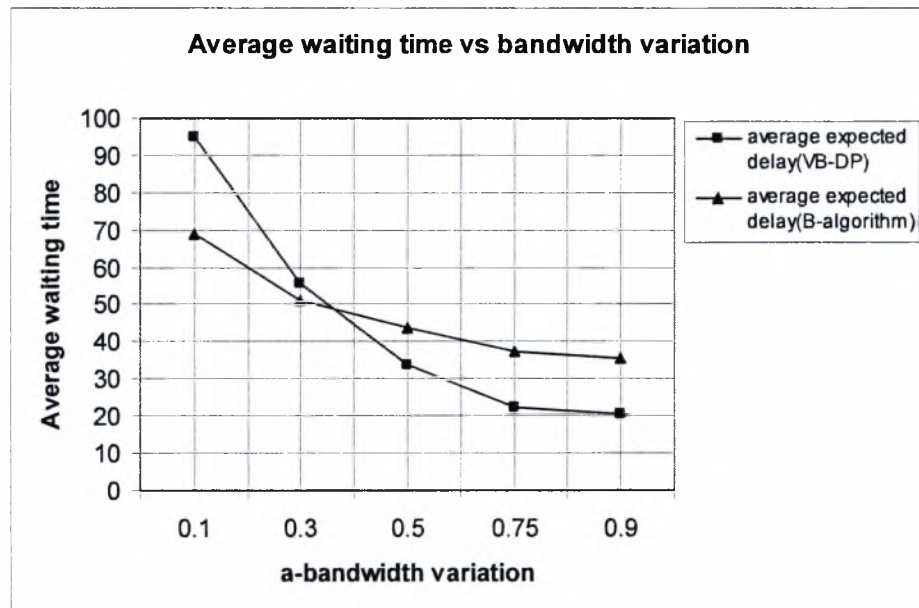
Σταθερές Παράμετροι:

$$\theta=0.8, N=250, K=5, b_1=50\text{kbps}$$

Με εφαρμογή του τύπου $b_{j+1}=a * b_j$ βρίσκουμε το bandwidth που αντιστοιχεί σε κάθε κανάλι.

Μεταβλητή παράμετρος:

Θεωρούμε ότι μεταβάλλεται ο αριθμός a δηλαδή τη μεταβλητότητα του bandwidth και χρησιμοποιούμε τις τιμές:
 $a=(0.9, 0.75, 0.5, 0.3, 0.1)$



Εικόνα 6.2.2.3

4. Μέσος χρόνος αναμονής vs skewness με (a=0.25)

Σταθερές Παράμετροι:

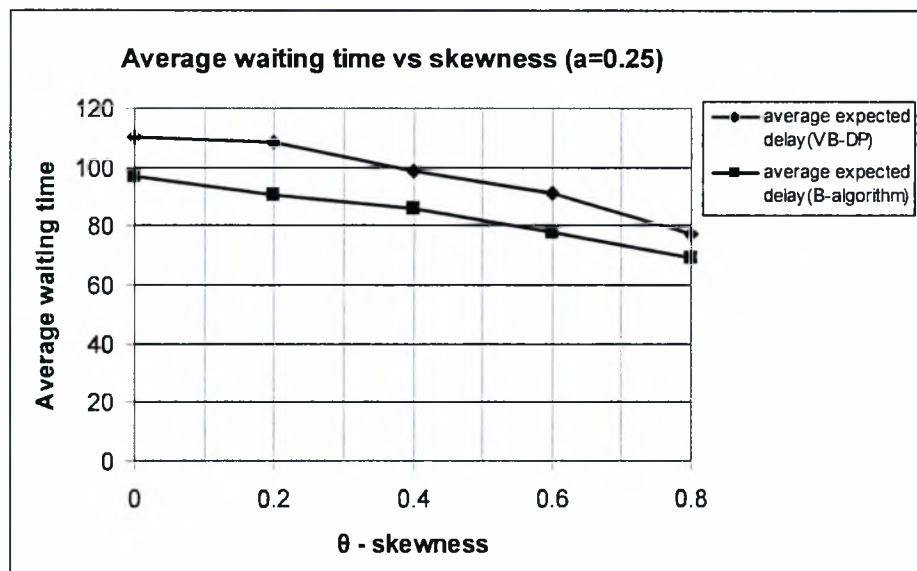
$$N=320, K=8, b_1=50\text{kbps}, a=0.25$$

Με εφαρμογή του τύπου $b_{j+1}=a * b_j$ βρίσκουμε το bandwidth που αντιστοιχεί σε κάθε κανάλι.

Μεταβλητή παράμετρος:

Θεωρούμε ότι μεταβάλλεται η παράμετρος skewness – θ που χαρακτηρίζει το μέτρο της ασυμμετρίας των πιθανοτήτων ζήτησης των δεδομένων εκπομπής.

$$\theta=(0, 0.2, 0.4, 0.6, 0.8)$$



Εικόνα 6.2.2.4

5. Μέσος χρόνος αναμονής vs skewness με (a=0.75)

Σταθερές Παράμετροι:

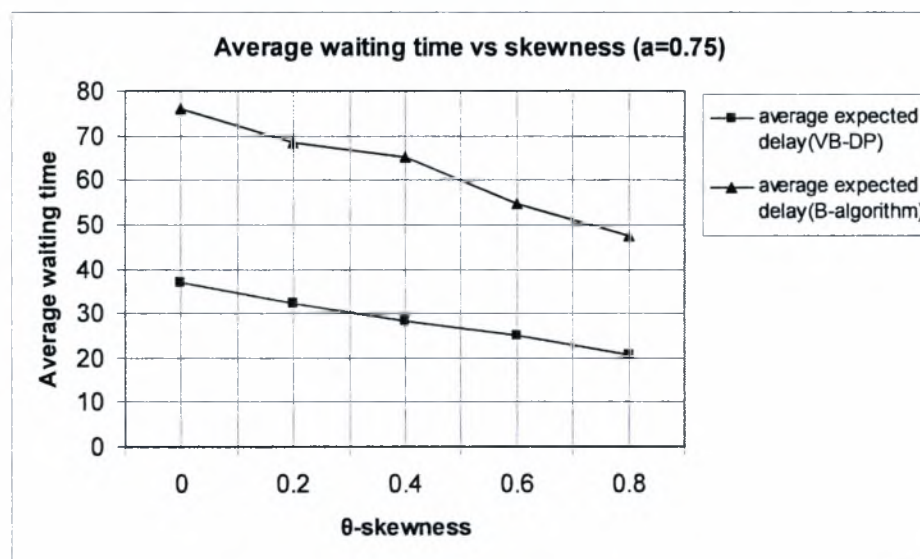
$$N=320, K=8, b_1=50\text{kbps}, a=0.75$$

Με εφαρμογή του τύπου $b_{j+1}=a * b_j$ βρίσκουμε το bandwidth που αντιστοιχεί σε κάθε κανάλι.

Μεταβλητή παράμετρος:

Θεωρούμε ότι μεταβάλλεται η παράμετρος skewness – θ που χαρακτηρίζει το μέτρο της ασυμμετρίας των πιθανοτήτων ζήτησης των δεδομένων εκπομπής.

$$\theta=(0, 0.2, 0.4, 0.6, 0.8)$$



Εικόνα 6.2.2.5

6.2.3 Σχολιασμός Αποτελεσμάτων

Όπως παρατηρούμε από τα αποτελέσματα των μετρήσεων τα οποία αποτυπώνονται στις παραπάνω γραφικές παραστάσεις ο αλγόριθμος δυναμικού προγραμματισμού VB-DP (variable-bandwidth dynamic programming algorithm) είναι πιο αποδοτικός σε σχέση με τον αλγόριθμο του ζυγού (balance- algorithm) σχεδόν σε όλες τις περιπτώσεις των μετρήσεων. Παρατηρούμε επίσης ότι η παράμετρος a (bandwidth variation) παίζει σημαντικό ρόλο σχετικά με το πόσο αποδοτικότερος είναι ο ένας αλγόριθμος έναντι του άλλου, έτσι αν παρατηρήσουμε στη συγκεκριμένη γραφική παράσταση (Εικόνα 6.2.2.3) όσο η τιμή του a κυμαίνεται σε χαμηλά επίπεδα ο αλγόριθμος του ζυγού είναι πιο αποδοτικός από τον αλγόριθμο δυναμικού προγραμματισμού ωστόσο καθώς το a αυξάνεται ο αλγόριθμος δυναμικού προγραμματισμού γίνεται πάλι αποδοτικότερος του αλγορίθμου ζυγού. Επίσης στην περίπτωση που μεταβάλλεται η παράμετρος θ έχουμε πάρει μετρήσεις για την περίπτωση που το $a=0.25$ και για όταν $a=0.75$. Παρατηρούμε ότι η παράμετρος a επηρεάζει την αποδοτικότητα του αλγορίθμου VB-DP σε σχέση με τον balance-algorithm και όταν το $a=0.25$ ο balance αλγόριθμος είναι πιο αποδοτικός από τον VB-DP ενώ το αντίθετο συμβαίνει όταν $a=0.75$. Μετά από αυτό μπορούμε να συμπεράνουμε ότι η μόνη παράμετρος που επηρεάζει την υπεροχή του VB-DP έναντι του balance αλγορίθμου είναι η παράμετρος που μας δίνει τη μεταβλητότητα στο bandwidth των καναλιών a όπου οι μικρές τιμές του a ευνοούν τον balance αλγόριθμο ενώ οι μεγαλύτερες τον αλγόριθμο VB-DP.

Κεφάλαιο 7

Αποτίμηση Εργασίας και Συμπεράσματα

7.1 Αποτίμηση Εργασίας

Στην εργασία αυτή προσπαθήσαμε να παρουσιάσουμε το πρόβλημα της εκπομπής δεδομένων σε πολλαπλά κανάλια μεταβλητού εύρους ζώνης. Στη μελέτη μας θεωρήσαμε ότι τα δεδομένα προς εκπομπή χαρακτηρίζονται από διαφορετικές (skewed) πιθανότητες ζήτησης και ότι η εκπομπή των δεδομένων γίνεται με ομοιόμορφο χρονοπρογραμματισμό (flat scheduling). Στην προσπάθεια μας να επιλύσουμε αυτό το πρόβλημα αναπτύξαμε δύο αλγόριθμους, τον αλγόριθμο δυναμικού προγραμματισμού για πολλαπλά κανάλια (VB-DP) και τον αλγόριθμο του ζυγού (balance algorithm). Στα πειράματα-προσομοιώσεις που κάναμε θεωρήσαμε ως παραμέτρους που επηρεάζουν την επίδοση των αλγορίθμων τον αριθμό των προς εκπομπή δεδομένων N , τον αριθμό των καναλιών K , την παράμετρο θ (skewness) της zipfian κατανομής από την οποία προκύπτουν οι τιμές των πιθανοτήτων ζήτησης των δεδομένων, τη μεταβλητότητα του bandwidth a . Κατά τη διάρκεια των προσομοιώσεων και εξετάζοντας τα αποτελέσματα που προέκυψαν είδαμε ότι ο αλγόριθμος VB-DP λύνει πιο αποδοτικά το πρόβλημα της ανάθεσης δεδομένων σε πολλαπλά κανάλια μεταβλητού εύρους ζώνης σε σχέση με τον αλγόριθμο του ζυγού σχεδόν σε όλες τις περιπτώσεις. Ο μόνος παράγοντας που προκαλεί κάποια μεταβολή στην αποδοτικότητα του αλγορίθμου VB-DP σε σχέση με τον balance αλγόριθμο είναι η

παράμετρος α όπου παρατηρήσαμε ότι όταν η τιμή του α είναι μικρή κυρίως κάτω από 0.3 ο αλγόριθμος VB-DP κρίνεται λιγότερο αποδοτικός σε σχέση με τον balance αλγόριθμο ενώ για μεγαλύτερη τιμή του α ο VB-DP συμπεριφέρεται καλύτερα από τον balance αλγόριθμο.

Αναφορές - Βιβλιογραφία

- [1] Optimal Skewed Data Allocation on Multiple Channels with Flat Broadcast per Channel
Elia Ardizzoni ,Alan A.Bertossi ,M.Cristina Pinotti ,Member ,IEEE Computer Society ,Shashank Ramaprasad ,Romeo Rizzi , and Madhusudana V.S. Shashanka
- [2] Variant Bandwidth Channel Allocation in the Data Broadcasting Environment
Chung-Hua Chu,Hao-Ping Hung ang Ming-Syan Chen
- [3] W.G. Yee, "Efficient Data Allocation for Broadcast Disk Arrays," Technical Report GIT-CC-02-20, Georgia Inst. of Technology, 2001.
- [4] S.-C. Lo and A.L.P. Chen, "Optimal Index and Data Allocation in Multiple Broadcast Channels," Proc. 16th IEEE Int'l Conf. Data Eng. (ICDE), Feb. 2000.
- [5] Dependent data broadcasting for unordered queries in a multiple channel mobile environment. *IEEE Trans. on Knowledge and Data Engineering*, 16(6), Jun. 2004.
J.-L. Huang and M.-S. Chen.
- [6] Wireless Information Highways
Δ.Κατσαρός ,Α.Νανόπουλος ,Ι.Μανωλόπουλος 2004
- [7] S. Acharya, R. Alonso, M. J. Franklin, and S. B. Zdonik. Broadcast disks: Data management for asymmetric communications environments. In *Proceedings of the 1995 ACM International Conference on Management of Data*, May 1995.
- [8] B. Zheng, X. Wu, and D. L. Lee. Tosa: A nearoptimal scheduling algorithm for multi-channel data broadcast. In *Proceedings of the 6th International Conference on Mobile Data Management (MDM-05)*, May 2005.

- [9] D. L. L. Jianliang Xu and B. Li. On bandwidth allocation for data dissemination in cellular mobile networks. In *ACM/Kluwer Journal of Wireless Networks (WINET), Special Issue on Advances in Mobile and Wireless*, volume 9 of 2, pages 103–116, 2003.
- [10] C. Kenyon, N. Schabanel, and N. Young, “Polynomial Time Approximation Scheme for Data Broadcast,” Proc. ACM Symp. Theory of Computing (STOC), pp. 659-666, 2000.
- [11] W.G. Yee, S. Navathe, E. Omiecinski, and C. Jermaine, “Efficient Data Allocation over Multiple Channels at Broadcast Servers,” IEEE Trans. Computers, vol. 51, no. 10, pp. 1231-1236, Oct. 2002.

