

Δρομολόγηση μέσω Συνεργατικής Μάθησης σε Οπορτουνιστικά Δίκτυα

Γεώργιος Λυπόπουλος

Διπλωματική Εργασία

Επιβλέπων: Ευάγγελος Παπαπέτρου

Ιωάννινα, Μάρτιος , 2024



**ΤΜΗΜΑ ΜΗΧ. Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ**

ΙΩΑΝΝΙΝΩΝ

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF IOANNINA**

Ευχαριστίες

Με την ολοκλήρωση αυτής της διπλωματικής εργασίας, ολοκληρώνονται οι προπτυχιακές μου σπουδές στο Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής του Πανεπιστημίου Ιωαννίνων.

Θέλω ιδιαίτερα να εκφράσω τις ευχαριστίες μου προς τον καθηγητή μου και επιβλέποντα της εργασίας, κ. Ευάγγελο Παπαπέτρου, για την ακαδημαϊκή του καθοδήγηση κατά τη διάρκεια όλων των σταδίων της έρευνας, παρέχοντας εύστοχες και εποικοδομητικές συμβουλές.

Τέλος, θέλω να εκφράσω την ευγνωμοσύνη μου προς την οικογένεια μου για τη συνεχή στήριξη της.

Περίληψη

Η παρούσα διπλωματική εργασία αναδεικνύει την ανάπτυξη πρωτοποριακών αλγορίθμων δρομολόγησης στο πλαίσιο οπορτουνιστικών δικτύων, αξιοποιώντας την ισχύ της μηχανικής μάθησης για την ενίσχυση της αποτελεσματικότητας και της αξιοπιστίας στη μεταφορά πληροφοριών. Αυτή η έρευνα στοχεύει να προσφέρει λύσεις στο πρόβλημα της επιβάρυνσης του δικτύου λόγω υπερβολικής αναπαραγωγής πακέτων, βελτιώνοντας συνεπώς τη συνολική διαχείριση δεδομένων.

Η μεθοδολογία της εργασίας περιλαμβάνει την εκπόνηση και την εφαρμογή σύνθετων αλγορίθμων, οι οποίοι ενσωματώνουν εξελιγμένες τεχνικές μηχανικής μάθησης για την ακριβή πρόβλεψη της συμπεριφοράς του δικτύου και την αποτελεσματική αξιολόγηση των δρομολογήσεων. Μέσω της εφαρμογής αυτών των τεχνολογιών, ελπίζουμε στη βελτίωση της επικοινωνίας μεταξύ των κόμβων, προσφέροντας μειωμένους χρόνους απόκρισης και αυξημένη αποδοτικότητα στη μεταφορά δεδομένων.

Τα ευρήματα της εργασίας υπογραμμίζουν τη σημασία της ενσωμάτωσης τεχνολογιών μηχανικής μάθησης στην επίλυση περίπλοκων προβλημάτων δικτυακής δρομολόγησης, ανοίγοντας νέους ορίζοντες για την έρευνα και την ανάπτυξη στον τομέα της ψηφιακής επικοινωνίας. Επιπλέον, η διατριβή παρέχει μια στέρεη βάση για μελλοντικές έρευνες, ενθαρρύνοντας την περαιτέρω ανάπτυξη καινοτόμων λύσεων στην επίλυση των προκλήσεων που αντιμετωπίζουν τα οπορτουνιστικά δίκτυα.

Abstract

The present thesis highlights the development of innovative routing algorithms within opportunistic networks, leveraging the power of machine learning to enhance the efficiency and reliability of information transfer. This research aims to offer solutions to the problem of network congestion due to excessive packet replication, thereby improving the overall data management.

The methodology of the work includes the formulation and application of complex algorithms that incorporate advanced machine learning techniques for accurate prediction of network behavior and effective evaluation of routing decisions. Through the implementation of these technologies, we hope to improve communication between nodes, offering reduced response times and increased efficiency in data transfer.

The findings of the study underscore the importance of integrating machine learning technologies in solving complex network routing problems, opening new horizons for research and development in the field of digital communication. Moreover, the dissertation provides a solid foundation for future research, encouraging further development of innovative solutions to the challenges faced by opportunistic networks.

Πίνακας Περιεχομένων

1.	Εισαγωγή	6
1.1.	Βασική Περιγραφή Οπορτουνιστικών Δικτύων και Πρακτικές εφαρμογές.....	6
1.2.	Εξέλιξη Αλγορίθμων Δρομολόγησης: Το Παρελθόν και το Παρόν	10
1.3.	Προκλήσεις που Αντιμετωπίζουν τα Οπορτουνιστικά Δίκτυα.....	12
1.4.	Αντικείμενο της διπλωματικής εργασίας.....	13
2.	Σχετική Βιβλιογραφία	15
2.1.	Μοντελοποίηση και Βασικοί ορισμοί των Οπορτουνιστικών δικτύων.....	15
2.2.	Πρωτόκολλα Δρομολόγησης στα Οπορτουνιστικά Δίκτυα.....	19
2.3.	Μηχανική μάθηση στα Οπορτουνιστικά Δίκτυα	28
3.	Συνεργατική μάθηση στη δρομολόγηση	30
3.1.	Συνεργατική Μάθηση στα Οπορτουνιστικά Δίκτυα	30
3.2.	Ο Αλγόριθμος F-CBR	32
4.	Πειραματική Αξιολόγηση	36
4.1.	Παράμετροι και εργαλεία προσομοίωσης.....	36
4.2.	Αποτελέσματα Πειραμάτων	39
5.	Συμπεράσματα	49

1.

Εισαγωγή

1.1. Βασική Περιγραφή Οπορουνιστικών Δικτύων και Πρακτικές εφαρμογές

Τα Οπορουνιστικά Δίκτυα[1] είναι ένα είδος δικτύων με καθυστέρηση (DTN) και αποτελούν μια προηγμένη έκδοση των κλασικών κινητών ασύρματων δικτύων ad hoc (MANETs). Στα MANETs, ήταν απαραίτητο να υπάρχει μια συνεχής διαδρομή από τον αποστολέα στον παραλήπτη πριν από τη μετάδοση μηνυμάτων. Ωστόσο, λόγω της κινητικότητας των συσκευών, της απενεργοποίησης ή της έλλειψης ενέργειας και των ασταθών ασύρματων συνδέσεων, αυτή η διαδρομή δεν εξασφαλίζεται πάντα.

Τα Οπορουνιστικά Δίκτυα λύνουν αυτό το πρόβλημα εκμεταλλευόμενα την κινητικότητα, επιτρέποντας στους κόμβους να μεταφέρουν μηνύματα οι ίδιοι καθώς κινούνται. Τα μηνύματα παραμένουν αποθηκευμένα μέχρι να εντοπιστεί ένας κατάλληλος επόμενος κόμβος για την προώθησή τους. Αυτό αντιπροσωπεύει ένα νέο μοντέλο Store-Carry-and-Forward . Τα Οπορουνιστικά Δίκτυα είναι επίσης γνωστά ως Δίκτυα Τσέπης, καθώς οι χρήστες λειτουργούν ως δρομολογητές αλλά και ως τελικοί χρήστες του συστήματος.

Τα Οπορουνιστικά Δίκτυα είναι χρήσιμα σε πολλές περιπτώσεις όπου οι συμβατικές μέθοδοι επικοινωνίας μπορεί να αποτύχουν ή να μην είναι εφικτές.

Κάποιες από τις πιθανές εφαρμογές περιλαμβάνουν:

1) Σε Καταστάσεις Εκτάκτου Ανάγκης:

- Καταστάσεις καταστροφής ή έκτακτης ανάγκης, όπως φυσικές καταστροφές, όπου οι υποδομές επικοινωνιών μπορεί να έχουν καταστραφεί ή να είναι ανεπαρκείς.

2) Σε Απομονωμένες Περιοχές:

- Σε απομονωμένες ή αγροτικές περιοχές όπου οι συνήθεις υποδομές επικοινωνιών είναι περιορισμένες ή ανύπαρκτες.

3) Σε Περιοχές με Υψηλή Κινητικότητα:

- Σε περιβάλλοντα όπου οι συσκευές κινούνται συχνά και είναι δύσκολο να διατηρηθεί σταθερή σύνδεση.

4) Σε Έρευνα και Παρακολούθηση Περιβάλλοντος:

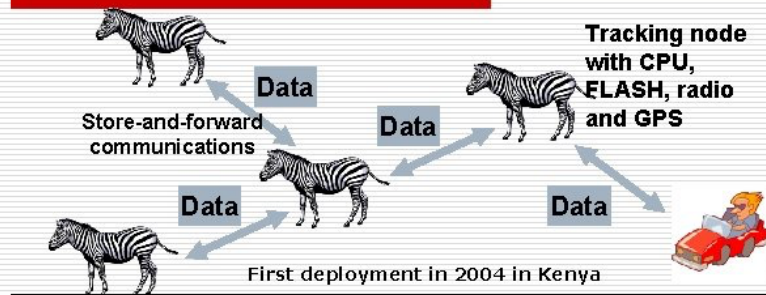
- Για τη συλλογή δεδομένων από αισθητήρες που είναι κινητοί ή εγκατεστημένοι σε κινητά αντικείμενα, όπως ζώα στη φύση.

5) Σε Εφαρμογές Κοινωνικής Δικτύωσης:

- Σε περιβάλλοντα όπου οι άνθρωποι θέλουν να επικοινωνήσουν ή να μοιραστούν πληροφορίες χωρίς την ύπαρξη σταθερής σύνδεσης, όπως σε μεγάλες εκδηλώσεις ή φεστιβάλ.

Έχοντας προηγουμένως αναφερθεί σε μερικές βασικές πιθανές χρήσεις των Οπορτουνιστικών Δικτύων, είναι καιρός να επιστρέψουμε στην πραγματικότητα και να ρίξουμε μια ματιά σε πραγματικές εφαρμογές αυτών των δικτύων.

ZebraNet (a real life application)



Sensor Network Attributes	ZebraNet	Other Sensor Networks
Node mobility	Highly mobile	Static or moderate mobile
Communication range	Miles	Meters
Sensing frequency	Constant sensing	Sporadic sensing
Sensing device power	Hundreds of mW	Tens of mW

<http://www.princeton.edu/~mrm/zebranet.html>

Παρακάτω παραθέτω ορισμένα παραδείγματα πραγματικών εφαρμογών Οπορτουνιστικών Δικτύων:

1) Saratoga [2]: Το Saratoga είναι ένα πρωτόκολλο που αναπτύχθηκε από την SSTL (Surrey Satellite Technology Ltd) για επικοινωνίες μεταξύ πλανητών. Το Saratoga χρησιμοποιήθηκε από τη NASA για πειράματα, με στόχο τη μεταφορά εικόνων που είχαν καταγραφεί από αισθητήρες δορυφόρων προς τη Γη.

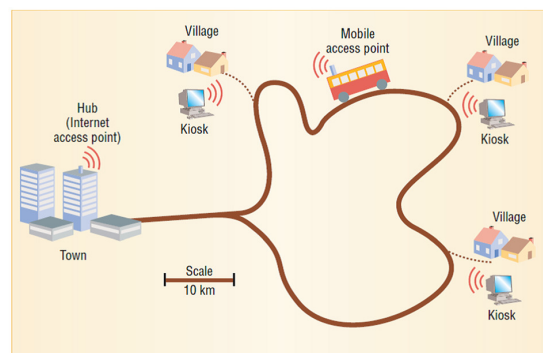
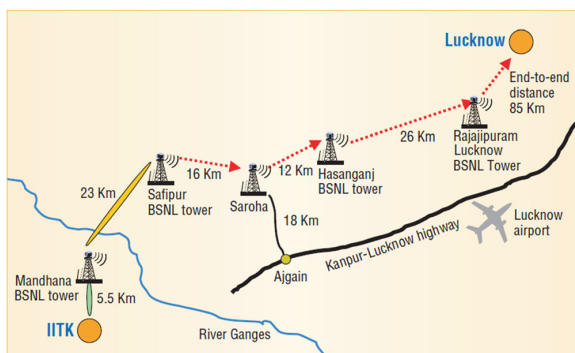
2) SWIM [3]: Το Shared Wireless Info Model (SWIM) είναι ένα Opportunistic Network-based Framework και σχεδιάστηκε για την παρακολούθηση των φαλαινών. Σε αυτό το framework, χρησιμοποιούνται αισθητήρες που είναι συνδεδεμένοι στην πλάτη των φαλαινών. Αυτοί οι αισθητήρες κινούνται μαζί με τα ζώα και ανιχνεύουν δεδομένα που σχετίζονται με τις φυσικές τους δραστηριότητες.

3) Lasso [4]: Το Lasso είναι μια εφαρμογή που βασίζεται σε Οπορτουνιστικά Δίκτυα και χρησιμοποιεί την τεχνολογία Bluetooth για να δημιουργήσει συνδέσεις μεταξύ διαφορετικών συσκευών. Ο βασικός στόχος του Lasso είναι να βοηθήσει ομάδες ανθρώπων, όπως τουρίστες

που βρίσκονται σε μια άγνωστη πόλη. Συνδέοντας τις συσκευές τους, κάθε μέλος της ομάδας μπορεί να γνωρίζει την ακριβή τοποθεσία των άλλων.

4) ZebraNet [5]: Το ZebraNet αποτελεί ένα Opportunistic Network-based Framework και επικεντρώνεται στην παρακολούθηση της άγριας ζωής, ειδικότερα της ζέβρας, και υλοποιήθηκε από το Πανεπιστήμιο Princeton στο Κένυα. Στο Framework αυτό, οι αισθητήρες δεμένοι στον λαιμό της ζέβρας επιτρέπουν σε αυτήν να κινείται ελεύθερα στη ζούγκλα ενώ συλλέγουν δεδομένα σχετικά με τη θέση και της συμπεριφορά της.

5) DakNet [6]: Αναπτύχθηκε από ερευνητές του MIT Media Lab και εφαρμόστηκε σε διάφορα μέρη του Καμπούτζη και της Ινδίας, το DakNet παρέχει ασύρματη συνδεσιμότητα με βάση την ασύγχρονη επικοινωνία. Ο στόχος του είναι να παρέχει οικονομική συνδεσιμότητα σε αγροτικά χωριά στην Ινδία, όπου η εγκατάσταση κανονικής πρόσβασης στο Διαδίκτυο δεν είναι οικονομικά προσιτή. Σε κάθε χωριό, υπάρχουν κιόσκια εξοπλισμένα με ψηφιακή αποθήκευση και ασύρματες επικοινωνίες μικρής εμβέλειας. Κινητές συσκευές πρόσβασης (MAPs), τοποθετημένες σε λεωφορεία, μοτοσικλέτες κ.λπ., ανταλλάσσουν ασύρματα δεδομένα με τα κιόσκια.



1.2. Εξέλιξη Αλγορίθμων Δρομολόγησης: Το Παρελθόν και το Παρόν

Έχοντας ήδη αναφέρει τα βασικά χαρακτηριστικά ενός Οπορτουνιστικού δικτύου καθώς και βασικές εφαρμογές αυτών ήρθε η στιγμή να αναλύσουμε ένα από τα σημαντικότερα ζητήματα που παρουσιάζουν τα εν λόγω δίκτυα και αυτό δεν είναι άλλο από τον τρόπο με τον οποία διαδίδονται τα πακέτα μέσα σε αυτό. Από την πρώτη στιγμή δημιουργίας τους, μίζων θέμα αποτελούσε η υλοποίηση αποδοτικών αλγορίθμων δρομολόγησης οι οποίοι θα παρείχαν αποδοτική μετάδοση των πακέτων.

Αρχικά, πρέπει να παρουσιαστούν οι δύο βασικές προσεγγίσεις στις οποίες ομαδοποιούνται τα πρωτόκολλα δρομολόγησης. Αυτές είναι η προσέγγιση Single-Copy και η προσέγγιση Multi-Copy. Κατ' αρχάς, με την προσέγγιση Single-Copy αναφερόμαστε σε πρωτόκολλα τα οποία επιτρέπουν την ύπαρξη μόνο ενός αντιγράφου ενός πακέτου μέσα στο δίκτυο. Αντιθέτως, τα πρωτόκολλα Multi-Copy επιτρέπουν την ύπαρξη πολλαπλών αντιγράφων ενός πακέτου μέσα στο δίκτυο την ίδια χρονική στιγμή. Χάρη σε αυτή την πολλαπλή ύπαρξη αντιγράφων, τα πρωτόκολλα Multi-Copy εμφανίζουν υψηλότερες πιθανότητες παράδοσης ενός πακέτου στον τελικό προορισμό, καθιστώντας τα αποδοτικότερα σε σχέση με τα πρωτόκολλα Single-Copy. Γι' αυτό το λόγο, στην συγκεκριμένη διπλωματική εργασία, επικεντρωθήκαμε στην ανάλυση μόνο των πρωτοκόλλων Multi-Copy.

Ένα πρωταρχικό πρωτόκολλο δρομολόγησης ,που ενσωματώνεται στα Multi-Copy πρωτόκολλα, για την μετάδοση πακέτων στα οπορτουνιστικά δίκτυα αποτέλεσε το Epidemic Routing[7] που χρησιμοποιεί την ιδέα της επιδημικής εξάπλωσης των μηνυμάτων. Σύμφωνα με αυτό, κάθε κόμβος που λαμβάνει ένα μήνυμα το αποθηκεύει και προσπαθεί να το μεταδώσει σε άλλους κόμβους κατά τη διάρκεια των συναντήσεών του. Το μήνυμα

εξαπλώνεται επιδημικά, χωρίς να υπάρχει εγγυημένο μονοπάτι από τον αποστολέα στον παραλήπτη. Ωστόσο το συγκεκριμένο πρωτόκολλο δρομολόγησης παρουσιάζει κάποια βασικά προβλήματα όπως την υπέρ-Αναπαραγωγή πακέτων , καθώς δεν υπάρχει κάποιο κριτήριο με το οποίο μεταδίδονται τα πακέτα από κόμβο σε κόμβο. Επιπλέον παρουσιάζει υπερβολική κατανάλωση ενέργειας λόγω της συνεχόμενης και ακατάπαυστης ανταλλαγής πακέτων ανά των κόμβων.

Τα προβλήματα που αντιμετωπίσαμε οδήγησαν στην ανάγκη ανάπτυξης αλγορίθμων δρομολόγησης που έχουν ως κύριο στόχο τη χρήση πληροφοριών που παρέχονται από το ίδιο το δίκτυο. Αυτές οι πληροφορίες προέρχονται από τους κόμβους και αφορούν τη χρησιμότητα τους για την παράδοση των πακέτων.

Συνοπτικά, οι αλγόριθμοι αυτοί επιδιώκουν να λάβουν υπόψη τη χρησιμότητα κάθε κόμβου ως προς τη δρομολόγηση πακέτων. Αντί να βασίζονται αποκλειστικά σε σταθερά μονοπάτια ή επιδημική διάδοση, οι αλγόριθμοι αυτοί αξιοποιούν πληροφορίες για την τρέχουσα κατάσταση του κάθε κόμβου. Τα πρωτόκολλα δρομολόγησης που περιγράψαμε μόλις τώρα μπορούν να χωριστούν σε δυο πολύ βασικές κατηγορίες την Spray-based και την Dynamic-Replication. Μιλώντας για την πρώτη κατηγορίας μπορούμε να πούμε ότι οι αλγόριθμοι που ανήκουν σε αυτήν προκαθορίζουν τον αριθμό αντιγράφων κάθε πακέτου που θα δημιουργηθεί μέσα στο δίκτυο και στην συνέχεια με την χρήση των πληροφοριών που δέχονται από άλλους κόμβους λαμβάνουν την απόφαση για το αν θα μεταδώσουν το πακέτο σε αυτούς η όχι. Τα πλεονεκτήματα που παρουσιάζει η συγκεκριμένη προσέγγιση είναι η ισορροπία ανάμεσα στην αποδοτικότητα ως προς την παράδοση των πακέτων και τον αριθμό των αντιγράφων που θα δημιουργηθούν μέσα στο δίκτυο λύνοντας το πρόβλημα που παρουσίαζε το πρωτόκολλο Epidemic-Routing. Ωστόσο εξακολουθούν να υπάρχουν προβλήματα που καθιστούν την συγκεκριμένη προσέγγιση ελλιπής. Στην άλλη πλευρά, αναφερόμενοι στην προσέγγιση Dynamic-Replication, εξετάζουμε αλγορίθμους οι οποίοι ακολουθούν έναν πιο ελαστικό τρόπο μετάδοσης των πακέτων μέσα στο δίκτυο. Ειδικότερα, αναφερόμαστε σε

αλγορίθμους οι οποίοι, σε αντίθεση με εκείνους της Spray-based προσέγγισης, δεν προκαθορίζουν τον αριθμό αντιγράφων που θα δημιουργηθούν μέσα στο δίκτυο. Αντ' αυτού, οι εν λόγω αλγόριθμοι επικεντρώνονται στη χρήση του ιστορικού των κόμβων, και συγκεκριμένα στις βασικές πληροφορίες που έχουν ανταλλαχθεί μεταξύ τους, με σκοπό τον εντοπισμό των πλέον κατάλληλων κόμβων για τη μετάδοση των αντιγράφων. Η εν λόγω προσέγγιση στοχεύει στη σημαντική βελτίωση των υπάρχουσων μεθόδων, επιτυγχάνοντας καλύτερες επιδόσεις στην ικανότητα παράδοσης πακέτων στους τελικούς προορισμούς. Ωστόσο, έχει την τάση να δημιουργεί πολλαπλά αντίγραφα πακέτων εντός του δικτύου.

1.3. Προκλήσεις που Αντιμετωπίζουν τα Οπορτουνιστικά Δίκτυα

Όπως αναφέραμε παραπάνω, τα Οπορτουνιστικά Δίκτυα έχουν κάποια προβλήματα που δεν έχουν ακόμα επιλυθεί [7][8]. Αυτά τα προβλήματα πρέπει να θεωρούνται ως προκλήσεις, ώστε οι μελλοντικοί ερευνητές να αφιερώσουν τις προσπάθειές τους στην αντιμετώπιση αυτών των εμποδίων για μια ομαλή και αποτελεσματική μετάδοση δεδομένων στα Οπορτουνιστικά Δίκτυα. Παρακάτω παρουσιάζονται μερικές από τις προκλήσεις που αντιμετωπίζουν:

1) Υπερπαραγωγή Αντιγράφων : Η παραγωγή μεγάλου αριθμού αντιγράφων αποτελεί το σημαντικότερο πρόβλημα στα Οπορτουνιστικά Δίκτυα, δημιουργώντας σημαντικές προκλήσεις σε θέματα ενέργειας και αποθηκευτικού χώρου. Συγκεκριμένα, καθώς οι κόμβοι λειτουργούν ως δρομολογητές εντός του δικτύου, αναλαμβάνουν την αποθήκευση των αντιγράφων μέχρι την παράδοσή τους σε επόμενο κόμβο. Αυτό οδηγεί σε ένα μείζον πρόβλημα σχετικά με τον αποθηκευτικό χώρο που διαθέτει κάθε κόμβος.

Επιπροσθέτως, η ενέργεια που απαιτείται για τη μεταφορά των αντιγράφων σε έναν επόμενο πιθανό κόμβο αποτελεί επίσης έναν σημαντικό παράγοντα ανησυχίας, καθώς επιβαρύνει τους πόρους του δικτύου.

2) High Mobility: Οι κόμβοι στα Οπορτουνιστικά Δίκτυα κινούνται συνεχώς, και εξαιτίας αυτής της υψηλής κινητικότητας, δεν διατηρούν σταθερή επαφή με άλλους κόμβους για μεγάλα χρονικά διαστήματα, με αποτέλεσμα η σύνδεσή τους να διακόπτεται.

3) Privacy: Κατά τη μετάδοση δεδομένων, όλοι οι μεσολαβητικοί κόμβοι αποθηκεύουν το μήνυμα που προωθούν στον επόμενο κόμβο για να εξασφαλίσουν την ασφαλή παράδοση. Κάθε κόμβος διατηρεί το μήνυμα μέχρι να λάβει επιβεβαίωση για την επιτυχημένη παράδοση του.

Αυτή η πρακτική μπορεί να οδηγήσει σε παραβίαση του απορρήτου, καθώς κακόβουλοι κόμβοι ενδέχεται να καταχραστούν τα δεδομένα που προωθούν ως μεσολαβητές.

4) High Delay: Τα Οπορτουνιστικά Δίκτυα αντιμετωπίζουν δυσκολία στη μείωση της καθυστέρησης λόγω των ασταθών συνδέσεων. Παρά την αστάθεια, οι κόμβοι προσπαθούν να εξασφαλίσουν επιτυχημένη παράδοση. Ωστόσο, η υψηλή καθυστέρηση πρέπει να θεωρείται πρόκληση με στόχο την ελαχιστοποίησή της.

1.4. Αντικείμενο της διπλωματικής εργασίας

Μετά από μια εκτενή ανάλυση των βασικών αρχών των οπορτουνιστικών δικτύων και των προκλήσεων που αντιμετωπίζουν, ήρθε η κατάλληλη στιγμή να αναφερθούμε στο αντικείμενο που

παρουσιάζει η συγκεκριμένη διπλωματική εργασία. Έχοντας ήδη κάνει μια ιστορική αναδρομή όσο αναφορά τα πρωτόκολλα δρομολόγησης που έχουν επικρατήσει μέχρι στιγμής ανά καιρούς, η θέση της συγκεκριμένης διπλωματικής εργασίας έρχεται να επεκτείνει αλγόριθμους που έχουν ήδη υλοποιηθεί στο παρελθόν.

Μέχρι στιγμής, έχουν υλοποιηθεί διάφορες προσπάθειες με σκοπό τη μείωση του αριθμού των αντιτύπων που παράγονται εντός του δικτύου κατά τη χρήση ενός αλγορίθμου Dynamic-Replication. Μεταξύ αυτών, οι προσεγγίσεις Compare & Replicate (CNR), Delegation Forwarding (DF), και Coordinating Replication Decisions (COORD) οι οποίες κατέδειξαν μείωση στον αριθμό των αντιγράφων, παρόλα αυτά δεν κατάφεραν να επιλύσουν πλήρως το πρόβλημα. Εν συνεχεία, μια εξαιρετικά σημαντική προσέγγιση που προτάθηκε ήταν η εφαρμογή τεχνικών μηχανικής μάθησης, ειδικότερα η προσέγγιση Cluster based Replication (CbR), η οποία επέδειξε υψηλές αποδόσεις και ιδιαίτερα καλά αποτελέσματα όσον αφορά τη μείωση της παραγωγής αντιγράφων.

Στην παρούσα διπλωματική εργασία, προτείνουμε μια πρωτοποριακή προσέγγιση για τη δρομολόγηση σε ομορτυνιστικά δίκτυα, με κύριο στόχο την επέκταση των τεχνικών μηχανικής μάθησης. Συγκεκριμένα, προτείνουμε την εφαρμογή του Federated Learning μεταξύ των κόμβων του δικτύου. Αυτή η καινοτομία αναμένεται να βελτιώσει σημαντικά τη διαχείριση των πληροφοριών που λαμβάνει κάθε κόμβος κατά την επικοινωνία του με τους υπόλοιπους κόμβους του δικτύου, ενισχύοντας τις αποφάσεις σχετικά με τη μετάδοση αντιγράφων. Με την εν λόγω τροποποίηση, ευελπιστούμε στην αποτελεσματική αντιμετώπιση των προβλημάτων που έχουν εντοπιστεί, δίνοντας ιδιαίτερη έμφαση στην πρόκληση της υπερπαραγωγής πακέτων.

Για την αξιολόγηση της προτεινόμενης προσέγγισης, διενεργούνται προσομοιώσεις υπό πραγματικές συνθήκες. Κατά τη διάρκεια αυτής της διαδικασίας, πραγματοποιείται σύγκριση μεταξύ των τροποποιημένων αλγορίθμων και των υφιστάμενων, με σκοπό την απόδειξη του βαθμού στον οποίο η συνεργατική μάθηση συμβάλλει στη βελτίωση της δρομολόγησης στα ομορτυνιστικά δίκτυα.

2.

Σχετική Βιβλιογραφία

Σε αυτό το κεφάλαιο, παρουσιάζονται οι βασικές έννοιες που σχετίζονται με το θέμα της διπλωματικής εργασίας. Αναλύονται διάφορες προσεγγίσεις γύρω από τα πρωτόκολλα δρομολόγησης στα ομορπονιστικά δίκτυα, ενώ ταυτόχρονα παρουσιάζονται και γνωστοί αλγόριθμοι που τις υλοποιούν. Τέλος, εξηγείται πώς η συνεργατική μάθηση ενσωματώνεται στους αλγορίθμους δρομολόγησης, παρέχοντας στον αναγνώστη βασικές πληροφορίες για την προσέγγιση που θα περιγράφει στη συνέχεια. Αυτό το κεφάλαιο αποτελεί τη θεμελιώδη βάση για την κατανόηση του περιεχομένου της εργασίας.

2.1. Μοντελοποίηση και Βασικοί ορισμοί των Ομορπονιστικών δικτύων

Στα προηγούμενα κεφάλαια αναφέρθηκε ότι οι κόμβοι στα ομορπονιστικά δίκτυα εκμεταλλεύονται την υψηλή κινητικότητα τους για τη μεταφορά πληροφοριών που έχουν στην διάθεση τους. Ωστόσο, δεν έχει ακόμα προσδιοριστεί ο τρόπος με τον οποίο αυτή η μεταφορά πληροφορίας λαμβάνει χώρα. Συγκεκριμένα, ένας κόμβος που κινείται εντός του δικτύου επικοινωνεί και ανταλλάζει πληροφορίες με κόμβους που συναντά στην διαδρομή του. Αυτή η επικοινωνία πραγματοποιείται μέσω της δημιουργίας επαφής μεταξύ των κόμβων. Ας εξετάσουμε λοιπόν πιο αναλυτικά τον ορισμό και τη φύση της επαφής.

Επαφή(Contact)

Στα ομορτυνιστικά δίκτυα, η έννοια του "link" μεταξύ δύο κόμβων δεν είναι τόσο καθορισμένη όπως στα παραδοσιακά δίκτυα, όπου οι συνδέσεις μπορεί να είναι σταθερές και προκαθορισμένες. Στα ομορτυνιστικά δίκτυα, η επικοινωνία μεταξύ των κόμβων συχνά βασίζεται σε προσωρινές ευκαιρίες επαφής (contacts) που προκύπτουν όταν οι κινητοί κόμβοι βρίσκονται σε φυσική εγγύτητα.

Συνεπώς, αντί να μιλάμε για "links" που συνδέουν σταθερά δύο κόμβους, χρησιμοποιείται η έννοια της "επαφής" (contact) για να περιγράψει τις περιοδικές ευκαιρίες επικοινωνίας μεταξύ των κινητών κόμβων. Κατά αυτόν τον τρόπο, αναδεικνύεται η ανάγκη για εφαρμογή πρωτοκόλλων και μηχανισμών που εκμεταλλεύονται αυτές τις προσωρινές συναντήσεις για τη μετάδοση δεδομένων.

Ένα από τα κρίσιμα ζητήματα για την κατανόηση της ανταλλαγής δεδομένων μεταξύ των κόμβων κατά τη διάρκεια της επαφής τους είναι ο τρόπος με τον οποίο αυτοί συνεργάζονται και επικοινωνούν για να ανταλλάξουν πληροφορίες με αποτελεσματικότητα και συνοχή. Σε αυτό το πλαίσιο, εισάγεται και ορίζεται η έννοια του Anti-Entropy Session, το οποίο παρέχει στους κόμβους του δικτύου τη δυνατότητα, κατά τη διάρκεια της επαφής τους, να ανταλλάσσουν μηνύματα που θα τους οδηγήσουν στην βέλτιστη δυνατή ανταλλαγή πληροφοριών. Ας αναλύσουμε λοιπόν περαιτέρω τον ορισμό και τη φύση του Anti-Entropy Session.

Anti-Entropy Sessions

Κάθε κόμβος[7] σε ένα ομορτυνιστικό δίκτυο διαθέτει έναν αποθηκευτικό χώρο(buffer) για να κρατά δημιουργημένα ή ληφθέντα πακέτα που πρόκειται να σταλθούν σε άλλες κόμβους.

Για να διαχειριστεί αποτελεσματικά αυτά τα μηνύματα, κάθε κόμβος έχει ένα "summary Vector," δηλαδή μια σύνοψη πακέτων τα οποία έχει στην κατοχή του.

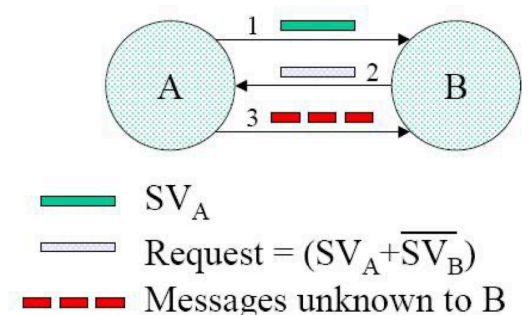
Όταν δύο κόμβοι βρίσκονται σε επαφή (Contact), ο κόμβος με το χαμηλότερο αναγνωριστικό ξεκινά ένα "Anti-Entropy Session" με τον κόμβο που έχει το υψηλότερο αναγνωριστικό. Κατά τη διάρκεια αυτής της συνεδρίας, τα μηνύματα προωθούνται. Για να αποφευχθούν περιττές συνεδρίες, κάθε κόμβος διατηρεί μια λίστα με τους κόμβους που έχει επικοινωνήσει πρόσφατα, ώστε να μην επαναλαμβάνει τη διαδικασία με συσκευές με τις οποίες έχει ήδη συνδεθεί μέσα σε ένα καθορισμένο χρονικό διάστημα.

Ωστόσο το κύριο πρόβλημα που έρχεται να λύσει το Anti-Entropy Session είναι η μη δημιουργία διπλοτύπων σε ένα κόμβο. Αυτό επιτυγχάνεται με την παρακάτω διαδικασία:

❑1: αποστολή μηνύματος Summary Vector (σύνοψη μηνυμάτων που ο A κατέχει)

❑2: απάντηση με Request (σύνοψη μηνυμάτων που ο επιθυμεί να λάβει)

❑3: μεταφορά δεδομένων από τον A στον B



Τέλος, ένα άλλο ζήτημα που απαιτεί περαιτέρω ανάλυση είναι ο τρόπος με τον οποίο, κατά τη διάρκεια ενός Anti-Entropy Session μεταξύ δύο κόμβων, γίνεται η επιλογή των πακέτων που θα μεταδοθούν. Όπως έχουμε ήδη αναφέρει στους αλγορίθμους Dynamic Replication, οι κόμβοι ανταλλάσσουν αντίγραφα πακέτων βάσει κάποιων βασικών κριτηρίων. Σε αυτό το στάδιο, εισάγεται η έννοια του utility value, η οποία παρέχει στους κόμβους τη δυνατότητα να δημιουργούν κριτήρια βασισμένα σε αριθμητικές συγκρίσεις. Ας αναλύσουμε λοιπόν περαιτέρω τον ορισμό και τη φύση του Utility Value.

Utility Values

Τα Utility Values αξιολογούν τους κόμβους του δικτύου βάσει της δυνατότητας τους να παραδώσουν ένα πακέτο στον τελικό του προορισμό. Τα Utility values [8][9] προκύπτουν από το connectivity profile ενός κόμβου και μπορούν να βασίζονται σε διάφορα χαρακτηριστικά, όπως ο ρυθμός επαφής, ο χρόνος μεταξύ διαδοχικών επαφών και κοινωνικά χαρακτηριστικά του κόμβου. Υπάρχουν δύο μεγάλες κατηγορίες Utility με βάση τον προορισμό:

1. Εξαρτημένα από τον προορισμό:

Σε αυτήν την περίπτωση, κάθε κόμβος διατηρεί ένα utility value για κάθε πιθανό προορισμό.

2. Ανεξάρτητα από τον προορισμό:

Σε αυτήν την περίπτωση, ο κάθε κόμβος χρησιμοποιεί ένα μόνο utility value για να ποσοτικοποιήσει την ικανότητά του να αλληλεπιδρά με άλλους κόμβους.

Κατά αυτόν τον τρόπο, τα Utility Values παρέχουν ένα πλαίσιο για να κατανοήσουμε και να αξιολογήσουμε τη σημασία και την

αποτελεσματικότητα των κόμβων σε ένα δίκτυο, είτε αυτή η αξιολόγηση επικεντρώνεται σε συγκεκριμένους προορισμούς είτε είναι γενικότερη και ανεξάρτητη. Τέλος, υπάρχουν διάφοροι αλγόριθμοι που εφαρμόζουν διαφορετικές μεθοδολογίες με στόχο τον καθορισμό των utility values των κόμβων.

2.2. Πρωτόκολλα Δρομολόγησης στα Οπορτουνιστικά Δίκτυα

Τα πρωτόκολλα δρομολόγησης στα κινητά οπορτουνιστικά δίκτυα μπορούν να χωριστούν σε δύο βασικές κατηγορίες: τα single-copy και τα multi-copy. Στα single-copy πρωτόκολλα, χρησιμοποιείται ένα μόνο αντίγραφο για κάθε πακέτο, ενώ στα multi-copy πρωτόκολλα υπάρχει πληθώρα αντιγράφων του ίδιου πακέτου την ίδια χρονική στιγμή στο δίκτυο. Τα multi-copy πρωτόκολλα έχουν το πλεονέκτημα της αυξημένης πιθανότητας εντοπισμού του κόμβου προορισμού για κάθε πακέτο. Συγκεκριμένα εφόσον υπάρχουν πολλά αντίγραφα ενός πακέτου στο δίκτυο, η πιθανότητα ενός από αυτά να φτάσει γρήγορα και αποτελεσματικά στον προορισμό είναι υψηλότερη. Αυτό μπορεί να οδηγήσει σε βελτιωμένη αξιοπιστία και ταχύτερους χρόνους παράδοσης σε σύγκριση με τα single-copy πρωτόκολλα. Επομένως, τα multi-copy πρωτόκολλα είναι πιο αποτελεσματικά όσον αφορά την παράδοση στον προορισμό λόγω της πολυπληθούς ύπαρξης αντιγράφων, προσφέροντας μεγαλύτερη ευκρίνεια.

Όπως έχουμε ήδη αναφέρει σε προηγούμενο κεφάλαιο η κατηγορία των multi-copy πρωτοκόλλων χωρίζεται σε 2 διαφορετικές προσεγγίσεις την Spray-based και την Dynamic replication προσέγγιση.

Στην περίπτωση της Spray-based προσέγγισης μιλάμε για αλγορίθμους στους οποίους οι κομβοί που παράγουν ένα πακέτο προκαθορίζουν και

τον μέγιστο αριθμό αντιγράφων(L) που θα δημιουργηθούν για το συγκεκριμένο πακέτο. Στην συνέχεια με την χρήση της spray process τα αντίγραφα αυτά διανέμονται σε άλλους κόμβους του δικτύου όταν έρθουν σε επαφή(contact).

Η επιλογή για το αν ο κόμβος που ήρθαμε σε επαφή πάρει το αντίγραφο του πακέτου μπορεί να γίνει είτε τυφλά, δηλαδή κάθε κόμβος είναι κατάλληλος να δεχτεί το πακέτο, είτε με την χρήση μιας utility value με την οποία εκφράζεται η καταλληλότητα του συγκεκριμένου κόμβου να προωθήσει το πακέτο στο τελικό προορισμό.

Ειδικότερα έστω ότι ένας κόμβος A (με utility value $U_A(d)$ για τον προορισμό d) κουβαλάει ένα πακέτο με προορισμό τον d και συναντάει (contact) τον κόμβο B (με utility value $U_B(d)$). Το A θα δημιουργήσει αντίγραφο και θα το στείλει στον B μόνο αν ισχύει ότι :

$$U_B(d) > U_A(d) + U_{th} \quad (1)$$

Όπου το U_{th} είναι μια παράμετρος η οποία εξασφαλίζει ότι ο νέος κόμβος που θα δεχτεί το πακέτο p θα έχει utility value μεγαλύτερο κατά ένα ορισμένο threshold. Με αποτέλεσμα ο ίδιος να αποτελεί ικανότερος ως προς την μεταφορά του πακέτο στον τελικό προορισμό.

Ας αναφερθούμε τώρα στη μετρική με την οποία ορίζουμε ένα utility:

Όπως αναφερθήκαμε και παραπάνω τα utility values μπορεί να είναι είτε destination dependent είτε destination independent. Στο συγκεκριμένο παράδειγμα μιλάμε για ένα utility value το οποίο ανήκει στην πρώτη κατηγορία καθώς κάθε κόμβος έχει ένα ξεχωριστό utility value για κάθε destination(d). Αν υποθέταμε ότι βρισκόμαστε στην δεύτερη κατηγορία τότε κάθε κόμβος θα είχε μονάχα ένα utility value και η σύγκριση που

γίνεται στην (1) θα γινόταν με αυτά τα 2 μοναδικά utility value που έχει κάθε κομβός.

Η προσέγγιση Spray-based έχει το πλεονέκτημα να επιτρέπει την αποτελεσματική ισορροπία μεταξύ της απόδοσης παράδοσης και του ρυθμού δημιουργίας αντιγράφων. Αυτό επιτυγχάνεται με τον προκαθορισμό της παραμέτρου L.

Ωστόσο, η σημαντική αδυναμία αυτής της προσέγγισης είναι η ανάγκη να προκαθοριστεί η βέλτιστη τιμή για το L. Αυτό είναι κρίσιμο, καθώς εξαρτάται από τις ειδικές ιδιότητες του δικτύου.

Η αδυναμία πρόβλεψης της βέλτιστης τιμής για την παράμετρο L μας ώθησε στην ανάγκη δημιουργίας πιο ευέλικτων προσεγγίσεων για πρωτόκολλα δρομολόγησης. Έχοντας πλέον ως στόχο την δημιουργία πρωτοκόλλων που είναι ανεξάρτητα από τις ειδικές ιδιότητες του δικτύου δημιουργήθηκε η εναλλακτική στρατηγική , γνωστή ως "Dynamic Replication".

Στην συγκεκριμένη προσέγγιση κάθε κόμβος που φέρει ένα πακέτο ακολουθεί μια utility-based προσέγγιση και δημιουργεί δυναμικά ένα αντίγραφο με βάση το utility value του κόμβου με τον οποίον έρχεται σε επικοινωνία. Αυτό σημαίνει ότι ο κάθε κόμβος εξετάζει το utility value του κόμβου που συναντά και αποφασίζει αν τελικά θα παραδώσει αντίγραφο του πακέτου σε αυτόν. Αυτή η ευελιξία επιτρέπει την προσαρμογή του αριθμού των αντιγράφων σύμφωνα με τις συνθήκες του δικτύου και χωρίς την ανάγκη προκαθορισμένων τιμών.

Έχοντας αναλύσει σε βάθος την Dynamic Replication προσέγγιση ήρθε η στιγμή να παρουσιάσουμε μερικούς βασικούς αλγόριθμους που εντάσσονται σε αυτήν.

Θα ξεκινήσουμε αναλύοντας έναν από τους βασικούς αλγόριθμους και συγκεκριμένα τον "*Compare & Replicate (CNR)* [10],[11],[12]". Στην

συγκεκριμένη υλοποίηση στην περίπτωση όπου δυο κόμβοι A και B, έρθουν σε επαφή (contact) τότε ο A στέλνει αντιγραφο πακέτου στο B αν ισχύει η σχέση (1).

Επομένως πλέον ο αριθμός των αντιγράφων που θα δημιουργηθούν δεν εξαρτάται από μια προκαθορισμένη τιμή L αλλά από την τοπολογία του δικτύου. Ωστόσο ο συγκεκριμένος αλγόριθμος όπως αποδείχτηκε παρουσιάζει κάποια βασικά προβλήματα τα οποία αδυνατεί να περιορίσει. Ένα από τα σημαντικότερα προβλήματα το οποίο αντιμετωπίζει ο συγκεκριμένος αλγόριθμος είναι ότι πολύ συχνά καταλήγουμε στη δημιουργία υπερβολικών αντιγράφων το οποίο από μόνο του μπορεί να δημιουργήσει τρομερά προβλήματα σε θέματα αποθήκευσης αλλά και ενέργειας που χρειάζονται για την μεταφορά όλων αυτών των αντιγράφων από τους κόμβους του δικτύου.

Στην ανάγκη μας λοιπόν να περιορίσουμε την υπέρ-αναπαραγωγή αντιγράφων που παρουσιάζει ο CNR στρέψαμε την προσοχή μας σε αλγόριθμους ικανούς να χρησιμοποιούν πληροφορίες που αντλούν από το ίδιο το δίκτυο με σκοπό την μείωση του πλήθους αχρειάστων αντιγράφων. Ένας τέτοιου είδους αλγόριθμος αποτελεί ο Delegation Forwarding (DF) [13].

Η DF προσέγγιση ήρθε για να αντιμετωπίσει το πρόβλημα της δημιουργίας υπερβολικών αντιγράφων κάνοντας χρήση του ιστορικού επαφών (contacts) κάθε κόμβου. Πιο συγκεκριμένα έστω την περίπτωση όπου ο κόμβος A κουβαλάει ένα πακέτο p με προορισμό το d, και έρχεται σε επαφή με τον κόμβο B. Ο κόμβος A θα δημιουργήσει αντίγραφο και θα το παραδώσει στο B μονάχα αν ισχύει ότι:

$$U_B(d) > \tau_A^p \quad (= \max\{U_k(d)\}) \quad (3) \quad k \in N_A \quad (2)$$

Όπου ως NA ορίζουμε την λίστα όλων των κόμβων με τους οποίους έχει

έρθει σε επαφή ο A. Ως τ_A^p ορίζουμε το μέγιστο utility value που έχει

καταγράψει ο κόμβος A από κομβους που δέχτηκαν το πακέτο p

Επομένως με αυτόν τον τρόπο δεν επιτρέπουμε στους κόμβους να δημιουργούν αντίγραφα πακέτων για κόμβους οι οποίοι έχουν μικρότερο utility value σε σχέση με κόμβους που έχουν ήδη το πακέτο μέσα στο δίκτυο.

Ένας άλλος πολύ σημαντικός αλγόριθμος που εντάσσεται και αυτός στην ίδια κατηγορία αλγορίθμων που ανήκει και ο DF είναι ο Coordinating Replication Decisions (COORD)[14] αλγόριθμος.

Ο COORD λειτουργεί με παρόμοιο τρόπο με τον DF, αλλά η κύρια διαφορά ανάμεσά τους είναι ότι ενώ ο DF χρησιμοποιεί μόνο το ιστορικό του κόμβου που ενδέχεται να στείλει το πακέτο, ο COORD χρησιμοποιεί το ιστορικό και των δυο κόμβων που εμπλέκονται στην επαφή. Συγκεκριμένα, στον COORD κατά την επαφή δύο κόμβων ανταλλάσσονται οι μέγιστες utility values που έχουν καταγράψει οι κόμβοι για τον συγκεκριμένο προορισμό. Με αυτόν τον τρόπο, κάθε κόμβος αποκτά μια πιο εκτεταμένη εικόνα του συνολικού δικτύου, προσφέροντας έτσι βελτιωμένη λειτουργία και απόδοση.

Οι αλγόριθμοι DF και COORD παρόλο που παρέχουν μια σημαντική βελτίωση στην απόδοση δεν καταφέρνουν να λύσουν εξολοκλήρου το πρόβλημα της δημιουργίας υπερβολικών αντιγράφων μέσα στο δίκτυο. Αν παρατηρήσουμε λίγο πιο προσεκτικά τις σχέσεις (1)-(2) θα παρατηρήσουμε ότι ως πρωταρχικό σκοπό τους αυτοί οι αλγόριθμοι έχουν την βελτίωση του utility value όσο ένα πακέτο πλησιάζει στον τελικό κόμβο προορισμό. Ωστόσο, αυτό προκαλεί [15] μια σημαντική

πρόκληση, καθώς εξαρτάται άμεσα από την τιμή που θα ορίσουμε για την μεταβλητή δU_{min} . Αν επιλέξουμε ένα μικρό δU_{min} ενδέχεται να παρουσιάσουν πάλι μεγάλο αριθμό αντιγράφων.

Από την άλλη πλευρά αν γίνει επιλογή μιας μεγάλης τιμής για την μεταβλητή δU_{min} τότε ενδέχεται να απορριφθούν αρκετοί κόμβοι με τελικό αποτέλεσμα την μη παράδοση του πακέτου στον τελικό κόμβο.

Για να μπορέσουμε να ορίσουμε[15] μια ιδανική τιμή για την μεταβλητή δU_{min} θα πρέπει να πάρουμε υπόψη κάποιους βασικούς παράγοντες όπως την utility function, network dynamics, την απόσταση μεταξύ του κόμβου, που έχει στην κατοχή του το πακέτο, και τον τελικό κομβο(Destination).

Βάση των παραπάνω παρατηρήσεων αντιλαμβανόμαστε ότι οι δυο προσεγγίσεις από μόνες τους δεν αρκούν για να λύσουν το πρόβλημα της δημιουργίας υπερβολικών αντιγράφων. Έχοντας πλέον αναλύσει τα προβλήματα που παρουσιάζονται στις ήδη υπάρχουσες υλοποιήσεις, μια νέα προσέγγιση έρχεται να βελτίωση την αποδοτικότητα των αλγορίθμων δρομολόγησης. Αυτή η προσέγγιση ονομάζεται "Cluster based Replication (CbR)"[15]. Ο CbR δεν λειτουργεί αυτόνομα ως ένας αλγόριθμος, αλλά αποτελεί έναν μηχανισμό που προστίθεται για να συμπληρώσει τις ήδη υπάρχουσες προσεγγίσεις CNR, DF και COORD. Συγκεκριμένα, ο CbR τροποποιεί τη διαδικασία λήψης αποφάσεων για τη μετάδοση ενός αντιγράφου σε έναν κόμβο. Αντί να συγκρίνουμε δύο τιμές, όπως στις σχέσεις (1)-(2), λαμβάνουμε υπόψη τις συστάδες στις οποίες ανήκουν αυτές οι δύο τιμές.

Διαδικασίες που ορίζουν την CbR προσέγγιση

- ❖ **Data Collection and Training:** Η διαδικασία εκπαίδευσης επιτρέπει σε κάθε κόμβο να συλλέξει αρκετά δείγματα από utility values, έτσι ώστε να μπορεί να αναγνωρίσει συστάδες από αυτές τις τιμές. Κατά τη διάρκεια της training περιόδου, ο κόμβος χρησιμοποιεί τη διαδικασία λήψης αποφάσεων του βασικού αλγορίθμου, δηλαδή CnR, DF ή COORD. Στο τέλος της training περιόδου, ο κόμβος εφαρμόζει μια τεχνική για να εντοπίσει τις συστάδες των utility values (K-Means Algorithm) [17].
- ❖ **Update:** Αυτή η φάση ξεκινά μετά την ολοκλήρωση της training περιόδου. Καθώς το δίκτυο εξελίσσεται, κάθε κόμβος συνεχίζει να καταγράφει νέες utility values μέσω των επαφών (contact) του. Αυτές οι νέες καταγραφές εμπλουτίζουν την αντίληψή του για το πώς κατανέμονται οι utility values στο δίκτυο. Η διαδικασία ενημέρωσης στοχεύει στο να ανανεώσει ανάλογα το αποτέλεσμα της ομαδοποίησης, δηλαδή τον τρόπο με τον οποίο τα utility values ομαδοποιούνται.
- ❖ **Decision Making:** Αυτή η φάση είναι η διαδικασία με την οποία παίρνετε η απόφαση για το αν θα επιτραπεί η αντιγραφή, η οποία εκμεταλλεύεται τις συστάδες που έχουν σχηματιστεί στις προηγούμενες φάσεις. Η διαδικασία αυτή ξεκινά μετά την ολοκλήρωση της training περιόδου και λειτουργεί ταυτόχρονα με τη διαδικασία ενημέρωσης. Σε αντίθεση με τις άλλες δύο διαδικασίες, η υλοποίησή της διαφέρει για κάθε έκδοση του CbR.

Τεχνική υλοποίηση του CbR

A. Detecting Clusters of Utility values

Ο CbR[15] ξεκινά με μια training περίοδο στην οποία κάθε κόμβος του δικτύου συλλέγει πληροφορίες τις οποίες λαμβάνει από κόμβους με τους οποίους έρχεται σε επαφή. Πιο συγκεκριμένα κάθε κόμβος διατηρεί μια λίστα από utility values για κάθε προορισμό(destination) για τον οποίον έχει στην κατοχή του ένα πακέτο προς παράδοση. Όλα αυτά που αναφέραμε συμβαίνουν στην περίπτωση όπου έχουμε να κάνουμε με utility τα οποία είναι εξαρτώμενα από τον προορισμό(destination). Σε περίπτωση που μιλάμε για utility τα οποία είναι ανεξάρτητα από αυτόν τότε αντί να διατηρούμε διαφορετικές λίστες για κάθε διαφορετικό προορισμό πλέον έχουμε μια μοναδική λίστα η οποία διατηρεί κάθε utility το οποίο στάλθηκε από κόμβους επαφής. Κάτι που πρέπει να σημειωθεί είναι ότι η training περίοδος δεν αποφέρει παράπανω επιβάρυνση στον αλγόριθμο. Ένα πρόβλημα του οποίου η επίλυση ήταν ζωτικής σημασίας για την συγκεκριμένη προσέγγιση ήταν η χρονική στιγμή κατά την οποία θα τελειώνει η training περίοδος. Μας ενδιαφέρει η διάρκεια της training περιόδου να μας επιτρέπει την απόκτηση επαρκούς αριθμού δειγμάτων από utility values αλλά ταυτόχρονα μας ενδιαφέρει και η γρήγορη έξοδος από αυτήν καθώς αυτό σημαίνει και γρήγορη έναρξη της διαδικασίας αντιγραφής βασισμένης σε συστάδες. Για την επίλυση του συγκεκριμένου προβλήματος ορίστηκε ένα threshold το οποίο ορίζει τον ελάχιστο αριθμό των utility values που πρέπει να έχουν καταγεγραμμοί στην λίστα. Με λίγα λόγια ορίσαμε τον ελάχιστο αριθμό από utility values που θα πρέπει να έχει το δείγμα μας. Επομένως όταν ο αριθμός από utility values ξεπεράσει αυτό το threshold τότε ολοκληρώνεται και η training περίοδος.

Μια σημαντική λεπτομέρεια όπου θα πρέπει να θίξουμε είναι ότι κάθε διαφορετικός προορισμός αποκτά utility values με διαφορετικό ρυθμό και αυτό έχει ως συνέπεια τον διαφορετικό χρόνο που απαιτείται για να βγεί από την training περίοδο.

Καθώς βγαίνουμε από την training περίοδο επόμενο βήμα αποτελεί η υλοποίηση ενός αλγορίθμου συσταδοποίησης ο οποίος θα εφαρμόζεται στα utility values που έχουν καταγραφεί. Ο αλγόριθμος που χρησιμοποιήθηκε είναι ο K-Means[17] σε συνδιασμό με το Silhouette criterion[18] με το οποίο ορίζεται ο αριθμός των clusters που απαιτούνται για κάθε λίστα η οποία διατηρεί utility values , δηλαδή πιο συγκεκριμένα για κάθε διαφορετικό προορισμό(destination) για τον οποίον έχουν καταγράψει utility values.

B. Updating the Clustering Result

Εφόσον μιλάμε για Opportunistic networks μπορούμε να υποθέσουμε ότι η κινητικότητα των κόμβων κατά την πάροδο του χρόνου θα δημιουργεί αισθητές αλλαγές στο ίδιο το δίκτυο κάτι το οποίο δεν μπορεί να καταγράφει μόνο κατά την training περίοδο. Για αυτόν τον λόγο δημιουργήθηκε μια ρουτίνα[15] η οποία τρέχει μετά το πέρας της training περιόδου και ανανεώνει τα αποτελέσματα των αρχικών cluster. Η ρουτίνα αυτή χρησιμοποιεί μια τεχνική που ονομάζεται LVQ[19]. Με πολύ απλά λόγια κάθε φορά που ένας κόμβος καταγράφει μια νέα utility value επιλέγει το cluster στο οποίο θα την ενσωματώσει και ταυτόχρονα ανανεώνει το κέντρο του συγκεκριμένου cluster.

C. Utilizing Clusters on Replication Decision Making

Έχοντας ολοκληρωθεί η training περίοδος, πλέον κάθε κόμβος είναι σε θέση να χρησιμοποιήσει τις συστάδες που έχει δημιουργήσει με σκοπό η απόφαση για την αντιγραφή ενός πακέτου να βασίζεται σε αυτές . Έχει

ήδη αναφερθεί ότι ο CbR[15] δεν αποτελεί από μόνος του έναν συνολικό αλγόριθμο ο οποίος παίρνει αποφάσεις για τις αντιγραφές των πακέτων αλλά αντιθέτως αποτελεί μηχανισμό ο οποίος σε συνεργασία με τους αλγόριθμους CNR,DF,COORD καταλήγουν στις τελικές αποφάσεις για το ποια αντίγραφα πακέτων αξίζει να δημιουργηθούν ή όχι.

Μια σημαντική λεπτομέρεια που πρέπει να ειπωθεί είναι πως εφόσον μιλάμε για διαφορετικούς αλγόριθμους που τρέχουν κάτω από τον μηχανισμό που παρέχει ο CbR, τότε αναγκάστηκα μιλάμε και για διαφορετικά κριτήρια με τα οποία αυτοί λαμβάνουν αποφάσεις για τις αντιγραφές των πακέτων (1) - (2), επομένως μιλάμε και για διαφορετικές υλοποιήσεις βάση των συγκεκριμένων κριτηρίων[15].

2.3. Μηχανική μάθηση στα Οπορτουνιστικά Δίκτυα

Αφού διαπιστώσαμε[15] ότι οι προσεγγίσεις CNR, DF και COORD δεν είναι ικανές να αντιμετωπίσουν αποτελεσματικά το πρόβλημα της υπερ-αναπαραγωγής ενός πακέτου στο δίκτυο, καταλήξαμε στο συμπέρασμα ότι θα ήταν χρήσιμο να εκμεταλλευτούμε το ιστορικό των επαφών ενός κόμβου και να δημιουργήσουμε συστάδες βασισμένες στα utility values που έχει καταγράψει. Καθώς γνωρίζουμε ότι το utility value ενός κόμβου καθορίζει την καταλληλότητά του για την προώθηση/παράδοση ενός πακέτου, η δημιουργία αυτών των συστάδων βασισμένων στα utility values είναι σαν να δημιουργούμε ομάδες κόμβων με διαφορετικές ικανότητες παράδοσης. Η κεντρική ιδέα πίσω από αυτήν την προσέγγιση είναι να παραδίδονται αντίγραφα μόνο σε κόμβους που ανήκουν σε ομάδες με αυξημένη ικανότητα παράδοσης, με στόχο τη μείωση των περιττών αντιγραφών πακέτων εντός του δικτύου. Για τον λόγο αυτό δημιουργήθηκε η Cluster Based Replication (CbR) προσέγγιση. Για την

υλοποίηση της CbR προσέγγισης έρχεται να μας βοηθήσει η μηχανική μάθηση παρέχοντας μας κατάλληλους αλγόριθμους που θα μας βοηθήσουν στην προσπάθεια μας να ομαδοποιήσουμε τα δεδομένα μας. Πιο συγκεκριμένα μιλάμε για αλγόριθμους ομαδοποίησης μη επιβλεπόμενης μάθησης οι οποίοι είναι σε θέση να δημιουργήσουν συστάδες από utility values οι οποίες θα παίζουν καθοριστικό ρόλο στην τελική απόφαση μετάδοσης ενός αντιγράφου μέσα στο δίκτυο. Βάση της προσέγγισης που περιγράψαμε κάθε κόμβος θα επεξεργάζεται ατομικά τα δεδομένα που αποκτά μέσω των επαφών του με άλλους κόμβους και θα δημιουργεί δικές του ατομικές συστάδες. Αυτό παρέχει την δυνατότητα σε κάθε κόμβο να δρα αυτόνομα καθώς δεν εξαρτάται από τις αποφάσεις που λαμβάνουν οι υπόλοιποι κόμβοι του δικτύου.

Στην συγκεκριμένη διπλωματική εργασία παρατηρήσαμε ότι παρόλο που η CbR προσέγγιση παρουσιάζει θετικά αποτελέσματα, ένας κόμβος προσδιορίζει τις συστάδες από utility χρησιμοποιώντας μονάχα πληροφορίες που ανέκτησε ο ίδιος από το δίκτυο. Πίστευουμε πως οι κόμβοι μπορούν να ανταλλάξουν πληροφορίες ώστε να έχουν βαθύτερη και ακριβέστερη εκτίμηση για της συστάδες από utility που υπάρχουν στο δίκτυο. Η προσέγγιση που παρουσιάζουμε λοιπόν εντάσσεται υπό την αιγίδα μιας συγκεκριμένης κατηγορίας την μηχανικής μάθησης που ονομάζεται συνεργατική μάθηση(Federated learning)[16]. Πλέον οι κόμβοι του δικτύου θα μπορούν να συνεργαστούν για να λάβουν ακριβέστερες αποφάσεις.

3. Συνεργατική μάθηση στη δρομολόγηση

3.1. Συνεργατική Μάθηση στα Οπορτουνιστικά Δίκτυα

Στην συγκεκριμένη διπλωματική εργασία στόχος μας ήταν η δημιουργία μια νέας προσέγγισης και συγκεκριμένα της Federated Learning CbR (F-CbR). Η προσέγγιση F-CbR ενσωματώνει την τεχνολογία της συνεργατικής μάθησης, προσφέροντας στους κόμβους του δικτύου τη δυνατότητα να συνεργάζονται με τρόπο που επιτρέπει την ανταλλαγή χρήσιμων πληροφοριών. Μέσω αυτής της συνεργασίας, κάθε κόμβος αποκτά την ικανότητα να εκπαιδεύει τοπικά ένα μοντέλο, το οποίο στη συνέχεια τον βοηθά στη δημιουργία των κατάλληλων συστάδων. Αυτές οι συστάδες χρησιμοποιούνται για την ομαδοποίηση των κόμβων του δικτύου σε ομάδες, βάσει της ικανότητας τους να μεταφέρουν δεδομένα προς έναν τελικό προορισμό. Η αφετηρία για την ανάπτυξη της συγκεκριμένης προσέγγισης ήταν η μελέτη του τρόπου με τον οποίο, στην προσέγγιση CbR, οι κόμβοι εκπαιδεύουν το τοπικό τους μοντέλο. Στην προσέγγιση CbR, η διαδικασία ομαδοποίησης των κόμβων σε συστάδες πραγματοποιείται αποκλειστικά με βάση τα utility values που έχει καταγράψει ο κάθε κόμβος από το δίκτυο. Για αυτόν τον λόγο, στη νέα μας προσέγγιση, κάθε κόμβος θα λαμβάνει υπόψη και τα utility values που προέρχονται από άλλους κόμβους του δικτύου, μέσω της συνεργατικής μάθησης, επιτρέποντας έτσι μια πιο πλούσια και διαδραστική διαδικασία εκμάθησης και ανταλλαγής πληροφοριών.

Η F-CbR προσέγγιση θα βασίζεται κατα κύριο λόγο στην προσέγγιση CbR που περιγράψαμε στο προηγούμενο κεφάλαιο. Το πλάνο μας είναι να κληρονομήσουμε μεγάλα κομμάτια της CbR προσέγγισης κάνοντας αλλαγές στις δομές μέσω των οποίων οι κόμβοι διατηρούν τα utility values που έχουν ανακτήσει από το δίκτυο, καθώς επίσης και στον τρόπο με τον οποίο ο μηχανισμός χρησιμοποιεί αυτά τα utility values για να δημιουργήσει τις συστάδες που μας ενδιαφέρουν.

Συγκεκριμένα, οι κόμβοι πλέον δεν θα μεταφέρουν μόνον πληροφορίες για τις ικανότητες τους (utility value) να παραδίδουν πακέτα σε διάφορους προορισμούς, αλλά θα μοιράζονται και πληροφορίες που έχουν να κάνουν με τις ικανότητες των κόμβων με τους οποίους έχουν έλθει ήδη σε επαφή. Ως εκ τούτου, αποτελεί κρίσιμη ανάγκη η δημιουργία πρόσθετων δομών αποθήκευσης για τα utility values, καθώς και η ανάπτυξη μηχανισμών που θα είναι υπεύθυνοι για την ανάκτηση και χρήση των εν λόγω τιμών.

Η άποψή μας προς τη Federated learning CBR προσέγγιση (F-CBR) είναι αισιόδοξη, καθώς πιστεύουμε ότι μπορεί να ενισχύσει σημαντικά την απόδοση της δρομολόγησης σε συνεργασία με τους υφιστάμενους αλγόριθμους, όπως αναφέρθηκαν προηγουμένως (CNR, DF, COORD).

Συγκεκριμένα, η προσέγγιση αυτή επιτρέπει σε κάθε κόμβο να εμβαθύνει στην κατανόησή του για το δίκτυο. Μέσω της συλλογής δειγμάτων από utility values, κάθε κόμβος δημιουργεί συστάδες που αντιπροσωπεύουν με μεγαλύτερη ευκρίνεια τη δομή του δικτύου.

Αυτή η εμπλοκή του σε πιο λεπτομερή δεδομένα επιτρέπει στον κόμβο να λαμβάνει ενημερωμένες αποφάσεις για το πώς να διαχειρίζεται τα δεδομένα του με σκοπό να μειώσει την παραγωγή περιττών αντιγράφων.

Στην ουσία, η προσέγγιση αυτή ενισχύει τη συνεργατική διαδικασία μεταξύ των κόμβων, βοηθώντας τους να λαμβάνουν πιο έξυπνες και ενημερωμένες αποφάσεις για τη δρομολόγηση σε πραγματικό χρόνο.

3.2. Ο Αλγόριθμος F-CBR

Όπως έχουμε ήδη αναφέρει, ο αλγόριθμος F-CbR κληρονομεί τους βασικούς μηχανισμούς του αρχικού αλγορίθμου CbR. Εν συνεχεία, θα παρουσιαστούν αναλυτικά οι καινοτομίες που εισάγει η πρόταση μας, τόσο σε ό,τι αφορά τον τρόπο με τον οποίο κάθε κόμβος αποθηκεύει αλλά και μοιράζεται τα δεδομένα που συλλέγει από το δίκτυο, όσο και στη μεθοδολογία που ακολουθείται για την παραγωγή των συστάδων.

A. Detecting Clusters of Utility values

Όπως προαναφέρθηκε και στην αρχική προσέγγιση του Cluster based Routing (CbR), κατά τη διάρκεια της training περιόδου, κάθε κόμβος συλλέγει τις utility values που λαμβάνει από άλλους κόμβους με τους οποίους έρχεται σε επαφή. Η καινοτομία στην εξέλιξη της προσέγγισης αυτής είναι ότι, κατά την επαφή δυο κόμβων, ο κάθε κόμβος δεν μεταφέρει απλώς το δικό του utility value για κάθε συγκεκριμένο προορισμό. Αντίθετα, μεταφέρει επίσης μια λίστα με τα utility values που έχει καταγράψει για τον εν λόγω προορισμό από προηγούμενες επαφές του.

Αυτό οδηγεί στην ανάγκη κάθε κόμβου να διαχειρίζεται δύο διαφορετικές λίστες για τα utility values. Η πρώτη λίστα, γνωστή ως "OneHopList," περιλαμβάνει τα utility values από τους κόμβους με τους οποίους έρχεστε σε άμεση επαφή ενώ η δεύτερη λίστα, γνωστή ως "TwoHopList," περιλαμβάνει τα utility values που ανήκουν σε κόμβους με τους οποίους δεν έχει έρθει σε άμεση επαφή αλλά έχουν έρθει οι κόμβοι επαφές του. Ας θυμηθούμε λοιπόν ότι κάθε utility-based αλγόριθμος ,

συμπεριλαμβανομένου και των CNR, DF και COORD , κατά την διάρκεια μιας επαφής δυο κόμβων ανταλλάσσουν τα utility values τους. Πλέον κάθε κόμβος πρέπει να μεταφέρει το utility value του αλλά και την OneHopList του για κάθε διαφορετικό προορισμό για τον οποίο θέλει να ζητήσει ένα αντίγραφο πακέτου. Σε αυτό το σημείο μη θέλοντας αυτή η μεταφορά της OneHopList από κόμβο σε κόμβο να επιβαρύνει σημαντικά την λειτουργία του αλγορίθμου λόγω υπερβολικών τιμών που θα έπρεπε να μεταφερθούν, καταλήξαμε στο συμπέρασμα πως δεν είναι απαραίτητη η μεταφορά ολόκληρης της λίστας αλλά μονάχα 3 στατιστικών δεδομένων τα οποία θα την περιέγραφαν επαρκώς.

Συγκεκριμένα ο κάθε κόμβος αντί να μεταφέρει ολόκληρη την OneHopList του για ένα προορισμό(destination) θα μεταφέρει τον αριθμό των utility values που υπάρχουν μέσα σε αυτή (length of sample) , το mean των utility values καθώς και το standard deviation των συγκεκριμένων τιμών.

Με αυτά τα 3 μονάχα στατιστικά δεδομένα ο κόμβος που τα δέχεται μπορεί να παράξει δεδομένα τα οποία θα διατηρεί στην TwoHopList που αντιστοιχεί στο συγκεκριμένο προορισμό (Destination). Τα παραγόμενα utility values θα έχουν σημαντικές ομοιότητες με τα αυθεντικά δεδομένα(utility values) που είναι αποθηκευμένα μέσα στην OneHopList του κομβού που τα μετέφερε. Μια σημαντική παρατήρηση που πρέπει να γίνει είναι πως εφόσον κάθε κόμβος μεταφέρει τα utility values του για κάθε ένα πακέτο του οποίου ζητάει αντιγράφο και όχι για κάθε διαφορετικό destination τότε πρέπει ο κόμβος που τα δέχεται να καταγράφει στις δομές του μονάχα μια φορά τα δεδομένα , ακόμα και στην περίπτωση ύπαρξης πακέτων με τον ίδιο προορισμό, μειώνοντας έτσι τον θόρυβο που θα προέκυπται στα δείγματα τους.

Σε αυτό το σημείο όπως και στην αρχική προσέγγιση του CbR έρχεται το δίλλημα για το πότε πρέπει να ολοκληρωθεί η training περίοδος. Όπως αναφέραμε και παραπάνω αυτή αποτελεί μια αρκετά σημαντική απόφαση καθώς επηρεάζει κατά πολύ την αποδοτικότητα του αλγορίθμου μας. Στην συγκεκριμένη προσέγγιση δημιουργούνται διαφορά σενάρια κατα τα οποία μπορεί να ολοκληρώνεται η training

περίοδος. Ένα πρώτο σενάριο αποτελούσε να προκαθοριστεί μια σταθερά (threshold) η οποία θα όριζε τον ελάχιστο αριθμό από utility values τα οποία θα έπρεπε να είχαν αποθηκευτεί στην OneHopList λίστα μας καθώς ταυτόχρονα και ένα ακόμα threshold το οποίο θα όριζε τον ελάχιστο αριθμό από utility values τα οποία θα έπρεπε να είχαν αποθηκευτεί στην TwoHopList λίστα. Με αλλά λογά η training περίοδος θα ολοκληρωνόταν όταν ο αριθμός των utility values που είχαν αποθηκευτεί στις δύο δομές ξεπερνούσε τα αντίστοιχα threshold.

Ωστόσο στο συγκεκριμένο σενάριο ενδέχεται η μια εκ των δυο λίστα μας να δέχεται δεδομένα με ρυθμό κατα πολύ μεγαλύτερο από την άλλη με αποτέλεσμα την ανάγκη “φρεναρίσματος” της διαδικασίας μέχρι να ξεπεραστεί και το threshold που απομένει. Αυτό θα είχε ως αποτέλεσμα να καθυστερήσει αρκετά η ολοκλήρωση της training περιόδου με συνέπεια και την καθυστέρηση της έναρξης της διαδικασίας αντιγραφής βασισμένης σε συστάδες. Ένα δεύτερο σενάριο έρχεται να διατηρήσει την αρχική συνθήκη κατά την οποία ο CbR ολοκληρώνει την training περίοδο και αυτή δεν είναι άλλη από τον ορισμό μιας μονάχα σταθεράς(threshold) , η οποία ορίζει τον κατώτατο αριθμό από utility values τα οποία θα πρέπει να διατηρεί η OneHopList. Καθώς η γρήγορη έξοδος από την training περίοδο αποτελεί μια σημαντική παράμετρο για την αποδοτικότητα του αλγορίθμου το δεύτερο σενάριο φαίνεται ικανοποιητικό προσδίδοντας μας ένα αρκετά ικανοποιητικό δείγμα αλλά και ταυτόχρονα μια γρήγορη έξοδο.

Ένα ουσιαστικό στοιχείο που επισημαίνεται και στην προσέγγιση του CbR είναι ότι στην περίπτωση όπου τα utility values είναι εξαρτώμενα από τον προορισμό(destination) τότε κατά συνέπεια εφόσον για κάθε διαφορετικό προορισμό διατηρούνται και διαφορετικές δομές για την αποθήκευση των utility values τότε είναι λογικό πως οι training περίοδοι θα ολοκληρώνονται και σε διαφορετικές χρονικές στιγμές.

Ας δούμε τώρα πως συμπεριφέρεται η συγκεκριμένη προσέγγιση κατά την ολοκλήρωση της training περιόδου. Κάθε κόμβος υλοποιεί έναν αλγόριθμο συσταδοποίησης στο δείγμα που έχει καταγράψει. Η βασική διαφορά σε αυτό το σημείο σε σχέση με την αρχική προσέγγιση του CbR

είναι ότι πλέον το δείγμα μας δεν περιορίζεται στην OneHopList αλλά στο συνδιασμό της με την TwoHopList όποτε είναι αναγκαία η σύμπτυξη τους σε μια κοινή λίστα η οποία θα αποτελεί και το τελικό δείγμα. Ο αλγόριθμος συσταδοποίησης που χρησιμοποιήθηκε είναι παρόμοιος με της αρχικής προσέγγισης και συγκεκριμένα είναι ο K-Means[17] σε συνδιασμό με το Silhouette criterion[18].

B. Updating the Clustering Result

Όπως περιγράψαμε και παραπάνω στην αρχική προσέγγιση του CbR είναι αναγκαίο μετά την ολοκλήρωση της training περιόδου η υπαρκτή ενός μηχανισμού ανανεώσεις των αρχικών clusters. Στην συγκεκριμένη προσέγγιση κληρονομείται ο μηχανισμός LVQ[19]. Η διαφορά ανάμεσα στις δυο προσεγγίσεις είναι πως στην αρχική τα clusters ανανεώνονταν μονάχα από τα νέα utility values που ανήκαν σε κόμβους επαφής. Ενώ στην νέα προσέγγιση είναι απαραίτητο τα clusters να ανανεώνονται και από τα παραγόμενα utility values που αναφέραμε παραπάνω. Δηλαδή το LVQ δεν δέχεται μονάχα το utility value του κόμβου επαφής αλλά και την λίστα που περιέχει τα παραγόμενα utility values που παράχθηκαν βάση των 3 στατιστικών δεδομένων που μας μετέφερε ο συγκεκριμένος κόμβος. Μια πολύ σημαντική λεπτομέρεια είναι πως κατά το τέλος της training περιόδου, η λίστα OneHopList δεν σταματάει να ανανεώνεται καθώς παραμένει απαραίτητη για την μεταφορά των 3 στατιστικών δεδομένων σε άλλους κόμβους. Με την ανανέωση της την διατηρούμε συνεχώς ενημερωμένη ως προς την δομή του δικτύου καθιστώντας την κατάλληλη για μεταφορά στατιστικών δεδομένων μεγαλύτερης ευκρίνειας ως προς την πραγματική εικόνα του δικτύου έχοντας ως αποτέλεσμα την καλύτερη ανανέωση των αρχικών clusters.

4.

Πειραματική

Αξιολόγηση

4.1. Παράμετροι και εργαλεία προσομοίωσης

Για να μπορέσω να αξιολογήσω την απόδοση του F-CBR έγινε χρήση του προσομοιωτή Adyton[20]. Ο Adyton παρέχει μια πληθώρα από πρωτόκολλα δρομολόγησης και είναι σε θέση να επεξεργάζεται έναν μεγάλο αριθμό από γνωστά contact traces που πηγάζουν από πραγματικά δίκτυα[21]. Για την συγκεκριμένη αξιολόγηση έγιναν χρήση traces που αντιπροσωπεύουν ομορτυνιστικά δίκτυα διαφορετικής κλίμακας. Πιο συγκεκριμένα χρησιμοποιήθηκαν 2 conference traces, ονόματι Infocom'05 [22] and Sigcomm'09 [23]. Επιπλέον χρησιμοποιήθηκαν 2 traces από διάφορα Campuses στα οποία οι συμμετέχοντες, φοιτητές και μέλη ΔΕΠ, μετακινούνται σε μια μεγάλη περιοχή. Τα συγκεκριμένα traces είναι τα MIT Reality [24] και το Milano pmtr [25]. Τέλος χρησιμοποιήθηκε το Cambridge upmc [26] το οποίο είναι ένα city-level trace το οποίο συλλέχθηκε στο Κέιμπριτζ του Ηνωμένου Βασιλείου.

TABLE I
PROPERTIES OF REAL-WORLD OPPORTUNISTIC TRACES

Trace Name	# Nodes	Duration (days)	Area
Infocom '05[29]	41	3	conference
Sigcomm '09[21]	76	3.7	conference
MIT Reality[20]	97	283	campus
Milano pmtr[30]	44	18.9	campus
Cambridge upmc[31]	52	11.4	city

Όπως και ο CbR, έτσι και ο F-CbR έχει την δυνατότητα προσαρμογής σε σχεδόν οποιαδήποτε προτεινόμενη μετρική χρησιμότητας (utility metric). Εννοείται ότι η επιλογή της μετρικής χρησιμότητας έχει άμεσο αντίκτυπο στα αποτελέσματα και, κατά επέκταση, στα κέρδη του F-CbR. Για να αξιολογήσουμε την απόδοση του F-CbR, έγινε χρήση ενός συνόλου από τέσσερις ευρέως αναγνωρισμένες μετρικές χρησιμότητας. Λεπτομερέστερα, εφαρμόζουμε τις ακόλουθες μετρικές χρησιμότητας.

- DestEnc [13] : Η DestEnc μετρική εξαρτάται από τον προορισμό και καταγράφει τον συνολικό αριθμό των επαφών με έναν συγκεκριμένο κόμβο.
- LTS [27], [7]: Η LTS μετρική εξαρτάται από τον προορισμό η οποία λαμβάνει τιμές από το σύνολο $[0, 1]$. Είναι αντιστρόφως ανάλογη με το χρόνο που έχει περάσει από την τελευταία επαφή ενός κόμβου με κάποιον κόμβο προορισμό.
- Prophet [28], [12] : Είναι μια μετρική που εξαρτάται από τον προορισμό. Η συγκεκριμένη μετρική προτείνεται στον γνωστό αλγόριθμο PROPHET. Η μετρική έχει τη μεταβατική ιδιότητα, δηλαδή καταγράφει την καταλληλότητα ενός κόμβου να παραδώσει ένα μήνυμα στο προορισμό όχι μόνο άμεσα αλλά και έμμεσα.
- SPM [29] : Η Social Pressure μετρική εξαρτάται από τον προορισμό και αποτυπώνει τη φιλία μεταξύ των κόμβων του δικτύου. Εξαρτάται από τη συχνότητα, τη διάρκεια και την τακτικότητα των επαφών που έχει καταγράψει ο κόμβος.

Όσον αφορά τώρα τις παραμέτρους που επιλέχτηκαν για την διαδικασία της ομαδοποιήσεως των δεδομένων, ορίστηκε μια μικρή τιμή K_{max} ίση με

4 η οποία αντιστοιχεί στον μέγιστο αριθμό συστάδων που θα μπορούν να ομαδοποιηθούν τα δεδομένα μας. Επιπλέον χρησιμοποιήθηκε μια περίοδος εκπαίδευσης ίση με 50 δείγματα, καθώς επίσης ορίστηκε παράμετρος ίση με 450 δείγματα που αντιστοιχεί στο μέγεθος της λίστας TwoHopList. Τέλος ο ρυθμός μάθησης LVQ ανέρχεται στο 0.05, δηλαδή η απόσταση μεταξύ μιας νέας προστιθέμενης τιμής και του κέντρου της συστάδας της μειώνεται κατά 5% μετακινώντας το κέντρο προς τη νέα τιμή.

Σε κάθε πείραμα που διεξάχθηκε χρησιμοποιήθηκε ένα traffic load ίσο με 1000 πακέτα. Η επιλογή ζεύγους/προορισμού επιλέγεται τυχαία για κάθε πακέτο, ενώ ο χρόνος δημιουργίας του επιλέγεται με την χρήση ομοιόμορφης πιθανότητας στο διάστημα κατά το οποίο τόσο η πηγή όσο και ο προορισμός είναι ενεργοί κόμβοι στο δίκτυο. Κάθε πακέτο χαρακτηρίζεται από μια TTL(Time-to-Live) παράμετρο της οποίας η τιμή ορίστηκε ίση με το 20% της συνολικής διάρκειας του Trace που χρησιμοποιήθηκε. Για να μπορέσουμε να μειώσουμε το Statical bias που ενδεχόμενος παρουσιαζόταν με σκοπό την παρακολούθηση του δικτύου μας σε κατάσταση ισορροπίας , χρησιμοποιήθηκε μια warm-up και μια cool-down περίοδο κατά της οποίες δεν παράγονται πακέτα. Η συνολική διάρκεια και των δυο περιόδων είναι ίση με το 20% την συνολικής διάρκειας του Trace.

4.2. Αποτελέσματα Πειραμάτων

Πρώτο Πείραμα

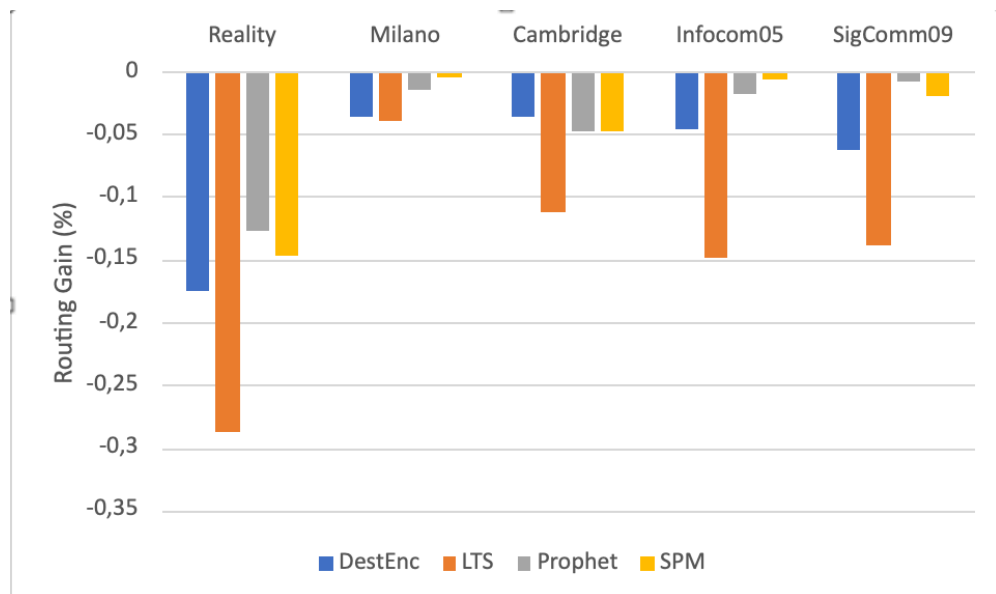
Στην πρώτη μας πείραμα, εξετάσαμε την λειτουργία του F-CbR σε συνδυασμό με τον αλγόριθμο DF για όλα τα contact Traces, χρησιμοποιώντας κάθε φορά διαφορετική utility metric. Σε αυτό το πείραμα, υποθέτουμε ότι κάθε κόμβος στο δίκτυο έχει ένα απεριόριστο buffer. Για να μετρήσουμε την εξοικονόμηση στις μεταδόσεις με την χρήση του F-CbR, χρησιμοποιήσαμε τη ποσότητα Routing Gain (RG), δηλαδή το ποσοστό των μεταδόσεων που εξοικονομούνται. Το ποσοστό αυτό υπολογίστηκε με την χρήση του τύπου $(1 - \text{TFCbR} / \text{TCbR})\%$, όπου TFCbR είναι οι μεταδόσεις για κάθε πακέτο που παραδίδεται από τον F-CbR όταν λειτουργεί με τον DF, ενώ το TCbR αφορά τις μεταδόσεις ανά παραδοθέν πακέτο για την έκδοση του CbR με τον DF. Στην εικόνα 1(α) μπορούμε να δούμε τα σχετικά διαγράμματα. Τα αποτελέσματα δείχνουν αρνητικά ποσοστά στο Routing Gain, υποδηλώνοντας αύξηση στις μεταδόσεις με τη χρήση του F-CbR σε σύγκριση με τον CbR. Αυτό υποδεικνύει ότι ο F-CbR μπορεί να μην ήταν τόσο αποτελεσματικός όσο ο CbR στη μείωση των αναγκαίων μεταδόσεων για την παράδοση πακέτων. Αυτή η παρατήρηση μπορεί να βοηθήσει στην περαιτέρω ανάπτυξη του F-CbR ή στην κατάλληλη επιλογή μετρικών για ειδικές συνθήκες του δικτύου.

Ας συνεχίσουμε λοιπόν με την μελέτη μιας άλλης ποσότητας που θα μας βοηθήσει στην αξιολόγηση του αλγορίθμου μας. Συγκεκριμένα υπολογίσαμε την ποσότητα delivery rate change. Η συγκεκριμένη ποσότητα υπολογίστηκε με την χρήση του τύπου $(\text{DFCbR} / \text{DFCbR} - 1)$, όπου DFCbR είναι το ποσοστό παράδοσης του F-CbR όταν λειτουργεί με τον DF, ενώ το DCbR είναι το ποσοστό παράδοσης του CbR με την χρήση του DF. Από τα διαγράμματα τις εικόνας 1 (β) φαίνεται πως υπάρχει

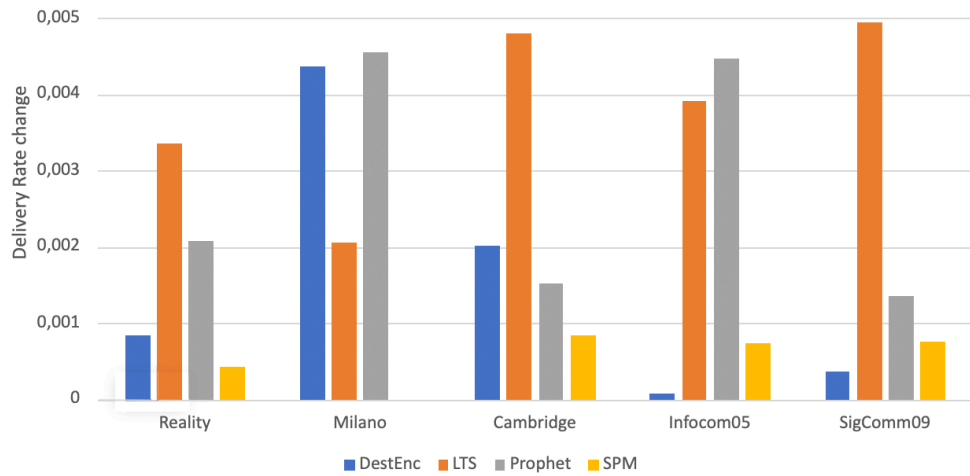
μικρή αλλά θετική βελτίωση στο ποσοστό παράδοσης πακέτων για όλες τις utility metrics και για όλα τα contact traces.

Η βελτίωση φαίνεται να φτάνει σε ποσοστά μέχρι και 0,5% ενώ κυρίως φαίνεται να κυμαίνεται κοντά στο 0,1%.

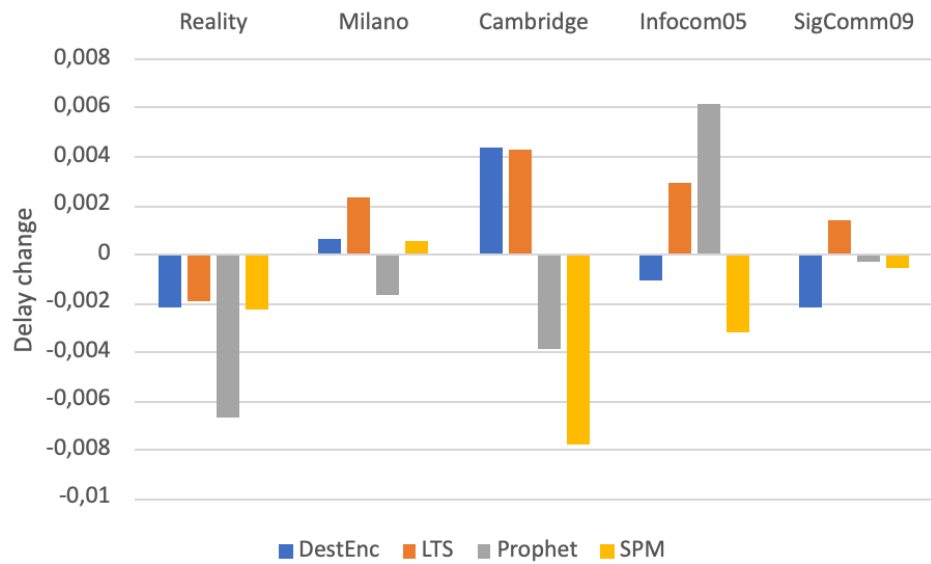
Όπως αναμενόταν, τα αυξημένα επίπεδα αντιγραφής στο F-CbR έχουν επίδραση και στην καθυστέρηση παράδοσης. Η Εικόνα 1 (γ) αποκαλύπτει το Delay Change, το οποίο ορίζεται με τον ίδιο τρόπο που το κάνει και το Delivery Rate Change. Από τα διαγράμματα, διαπιστώνουμε ότι στην περίπτωση του Trace Reality υφίσταται μια ήπια μείωση στην καθυστέρηση. Η αντιγραφή μπορεί να φανταστεί ως διαδικασία που διανέμει αντίγραφα στον προορισμό μέσω διάφορων διαδρομών, κάτι που σημαίνει ότι η αύξηση της αντιγραφής δημιουργεί νέες διαδρομές. Αυτό βελτιώνει τις πιθανότητες εύρεσης του προορισμού και επομένως μειώνει την καθυστέρηση παράδοσης του πακέτου. Ωστόσο, υπάρχουν περιπτώσεις όπου η καθυστέρηση αυξάνεται παρά την παραγωγή περισσότερων αντιγράφων. Σε τέτοιες περιπτώσεις, μπορούμε να υποθέσουμε ότι ο ρόλος της μετρικής χρησιμότητας που χρησιμοποιείται είναι καθοριστικός.



Εικόνα 1(α)



Εικόνα 1(β)



Εικόνα 1 (γ)

Δεύτερο Πείραμα

Στο επόμενο πείραμα, επικεντρωθήκαμε αποκλειστικά στο σύνολο δεδομένων που πηγάει από το MIT Reality Trace, χρησιμοποιώντας το πλήρες φάσμα των διαθέσιμων Utility Metrics. Προσπαθήσαμε να αξιολογήσουμε την απόδοση του μοντέλου F-CbR, διατηρώντας σταθερό το μέγεθος των δειγμάτων που επεξεργάζεται η OneHopList(50), ενώ

παράλληλα προσαρμόζαμε δυναμικά το μέγεθος των δειγμάτων για την TwoHopList. Αυτή η μεθοδολογία υλοποιήθηκε σε πέντε διαφορετικές αναλογίες, οι οποίες απεικονίζονται λεπτομερώς στον επόμενο πίνακα :

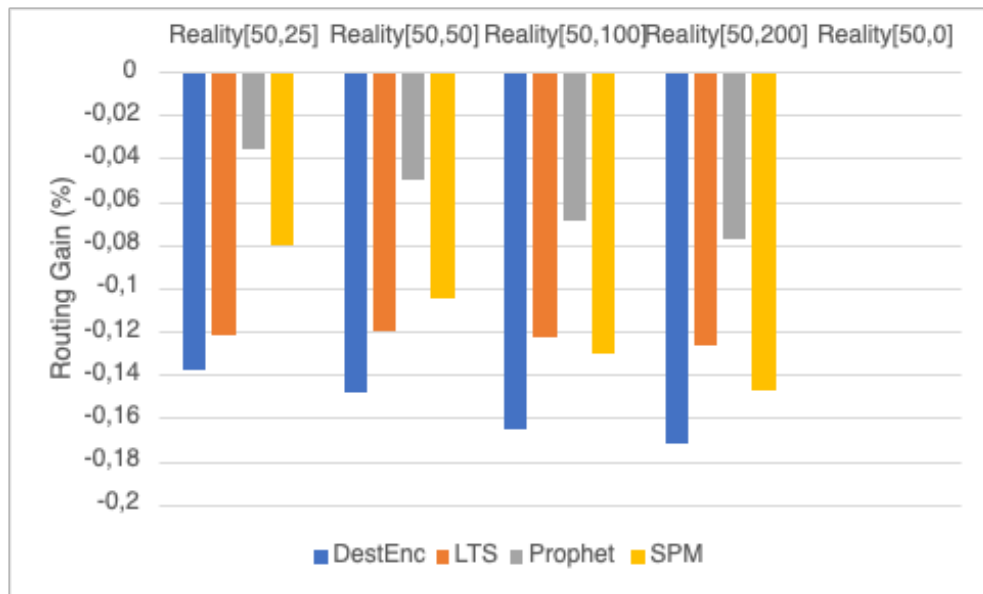
OneHopList	TwoHopList
50	0
50	25
50	50
50	100
50	200

Στο εν λόγω πείραμα, επαναχρησιμοποιήθηκαν οι τρεις κρίσιμες μετρικές που είχαν καθοριστεί στο προηγούμενο πείραμα: Routing Gain, Delivery Rate Change και Delay Change. Για την εκτίμηση των συγκεκριμένων μετρικών, πραγματοποιήθηκε συγκριτική ανάλυση μεταξύ της προσέγγισης F-CbR, η οποία υιοθετεί τον υποκείμενο αλγόριθμο DF, και της προσέγγισης CbR που επίσης εφαρμόζει τον αλγόριθμο DF. Εντός του πλαισίου του πειράματος, ο αλγόριθμος CbR χρησιμοποιεί ένα δείγμα μεγέθους 50.

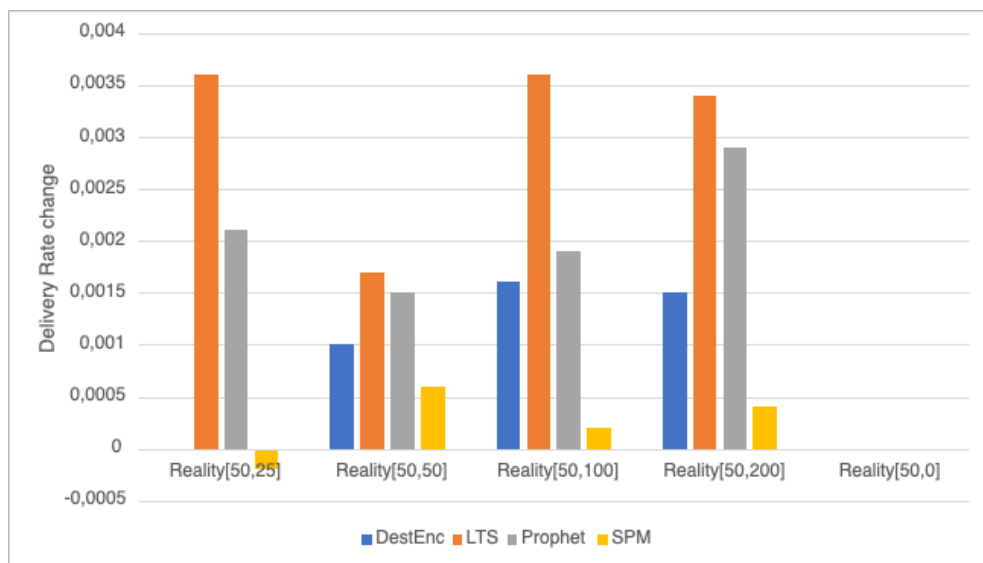
Στην Εικόνα 2(α), είναι εμφανή τα διαγράμματα που αφορούν το Routing Gain. Από την ανάλυση των διαγραμμάτων παρατηρούμε ότι, καθώς αυξάνεται το μέγεθος των δειγμάτων που λαμβάνει η TwoHopList, σημειώνεται αύξηση στις απαιτούμενες μεταδόσεις από τον αλγόριθμο F-CbR. Όταν ορίζεται το μέγεθος δειγμάτων της TwoHopList σε μηδέν, παρατηρείται παρόμοια συμπεριφορά με εκείνη του αλγορίθμου CbR, γεγονός που είναι αναμενόμενο. Αυτό συμβαίνει επειδή, με την αφαίρεση των δειγμάτων από την TwoHopCbR, ο αλγόριθμος F-CbR καθίσταται ισοδύναμος με τον απλό CbR.

Στην εικόνα 2(β), παρουσιάζονται τα διαγράμματα που αφορούν το Delivery Rate Change. Από την ανάλυση των διαγραμμάτων προκύπτει ότι ο αλγόριθμος F-CbR σημειώνει μια ελαφρά αλλά θετική βελτίωση στο ποσοστό παράδοσης πακέτων, για όλες τις αναλογίες δειγμάτων μεταξύ των δύο λιστών, OneHopList και TwoHopList. Από την ανάλυση του διαγράμματος, διαπιστώνεται ότι οι μετρικές LTS και Prophet ξεχωρίζουν για τη θετική επίδραση που έχουν στην αλλαγή του ποσοστού παράδοσης πακέτων. Ειδικότερα, οι δύο αυτές μετρικές φαίνεται να επιτυγχάνουν υψηλότερη βελτίωση στο ποσοστό παράδοσης σε σύγκριση με τις μετρικές DestEnc και SPM, κάτι που ενδέχεται να υποδηλώνει μια πιο αποτελεσματική διαχείριση των δεδομένων από τους κόμβους όταν χρησιμοποιούνται αυτές οι μετρικές. Το γεγονός αυτό μπορεί να ερμηνευτεί ως ένδειξη της υπεροχής τους στην ακρίβεια και την αποδοτικότητα της πρόβλεψης και της αξιολόγησης των διαδρομών παράδοσης πακέτων. Όταν ορίζεται το μέγεθος δειγμάτων της TwoHopList σε μηδέν, όπως και στην περίπτωση του Routing Gain, παρατηρείται παρόμοια συμπεριφορά με εκείνη του αλγορίθμου CbR, γεγονός που είναι αναμενόμενο.

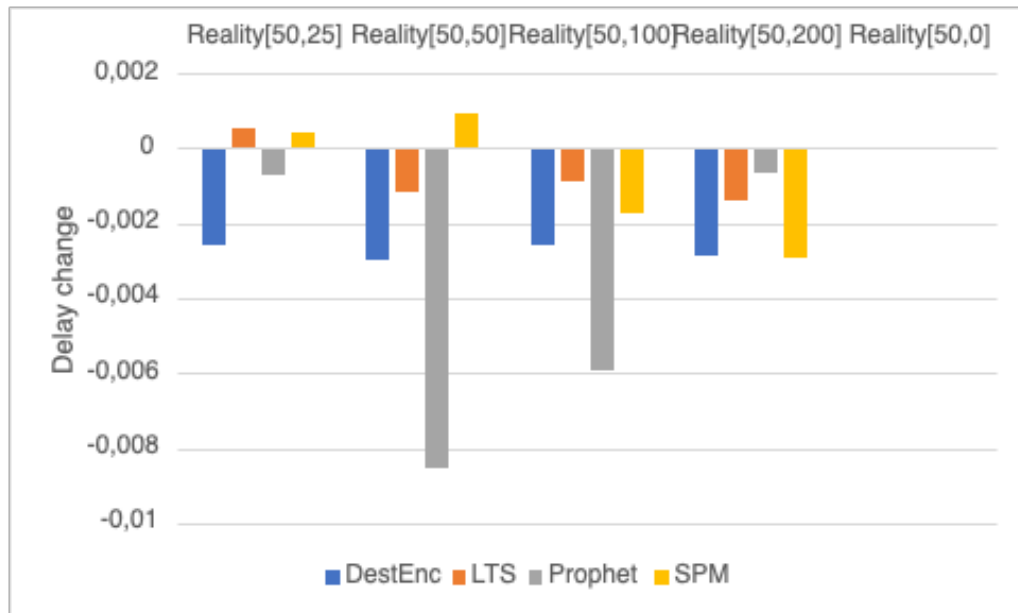
Στην εικόνα 2(γ) διακρίνουμε τα διαγράμματα που αφορούν το Delay Change. Αναλύοντας τα δεδομένα, παρατηρούμε ότι ο αλγόριθμος F-CbR επιφέρει μια ελαφρά μείωση στον χρόνο καθυστέρησης. Ειδικότερα, οι μετρικές LTS και SPM δείχνουν μικρότερη βελτίωση στον χρόνο καθυστέρησης συγκριτικά με τις μετρικές DestEnc και Prophet. Παρατηρούμε ότι ο αυξανόμενος αριθμός δειγμάτων στην TwoHopList δεν συνδέεται απαραίτητα με θετική επίδραση στη μείωση του χρόνου καθυστέρησης. Όταν το μέγεθος των δειγμάτων για την TwoHopList ορίζεται σε μηδέν, ο F-CbR δείχνει να συμπεριφέρεται παρόμοια με τον CbR αλγόριθμο.



Εικόνα 2(α)



Εικόνα 2(β)



Εικόνα 2(γ)

Τρίτο Πείραμα

Στο τρίτο και τελευταίο μας πείραμα, επικεντρωθήκαμε και πάλι στα δεδομένα που πηγάζουν από το MIT Reality trace, χρησιμοποιώντας το πλήρες φάσμα των διαθέσιμων Utility Metrics. Σε αυτή μας την προσπάθεια να αξιολογήσουμε την συμπεριφορά και την αποδοτικότητα του F-CbR κρίναμε απαραίτητο να εξετάσουμε την ανταπόκριση του αλγορίθμου κάτω από συνθήκες όπου ο αριθμός των δειγμάτων παρέμενε σταθερός(150 δείγματα), ενώ μεταβάλλαμε δυναμικά την αναλογία των δειγμάτων μεταξύ των λιστών OneHopList και TwoHopList. Αυτή η μεθοδολογία υλοποιήθηκε σε τέσσερις διαφορετικές αναλογίες, οι οποίες απεικονίζονται λεπτομερώς στον επόμενο πίνακα :

OneHopList	TwoHopList
25	125
50	100
75	75
100	50

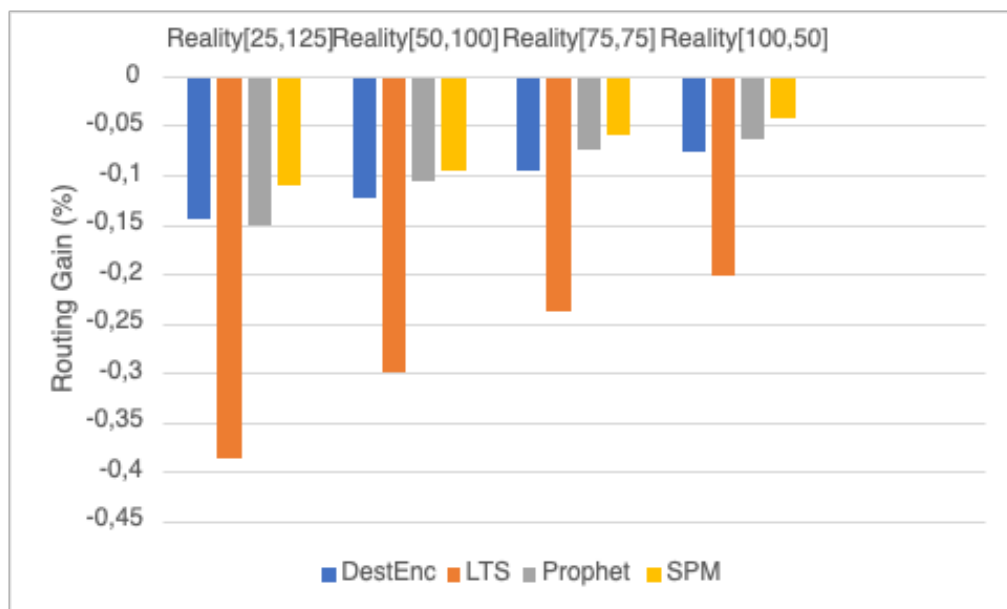
Στο εν λόγω πείραμα, χρησιμοποιήθηκαν οι τρεις κρίσιμες μετρικές που είχαν καθοριστεί στα προηγούμενα πειράματα: Routing Gain, Delivery Rate Change και Delay Change. Για την εκτίμηση των συγκεκριμένων μετρικών, πραγματοποιήθηκε συγκριτική ανάλυση μεταξύ της προσέγγισης F-CbR, η οποία υιοθετεί τον υποκείμενο αλγόριθμο DF, και της προσέγγισης CbR που επίσης εφαρμόζει τον αλγόριθμο DF. Στο πλαίσιο του εν λόγω πειράματος, ο αλγόριθμος CbR υιοθετεί ως μέγεθος για το δείγμα του, την ίδια τιμή που καθορίζεται για το μέγεθος της λίστας OneHopList, όπως αυτή ορίζεται από τον αλγόριθμο F-CbR. Ως εκ τούτου, για καθεμία από τις τέσσερις διαφορετικές αναλογίες μεταξύ των δύο λιστών, ο CbR λαμβάνει ως μέγεθος του δείγματος, το μέγεθος που έχει καθοριστεί για την OneHopList.

Στην εικόνα 3(α) διακρίνουμε τα διαγράμματα που αφορούν το Routing Gain. Απο την ανάλυση των διαγραμμάτων μπορούμε να παρατηρούμε ότι με την αύξηση του ποσοστού των δειγμάτων που ανατίθενται στην TwoHopList και την ταυτόχρονη μείωση του ποσοστού των δειγμάτων για την OneHopList, υπάρχει αντίστοιχη αύξηση στον αριθμό των απαιτούμενων μεταδόσεων από τον αλγόριθμο F-CbR. Αυτό το φαινόμενο είναι συνεπές και επαναλαμβανόμενο σε όλες τις utility metrics, με την LTS να εμφανίζει τα μεγαλύτερα ποσοστά αύξησης.

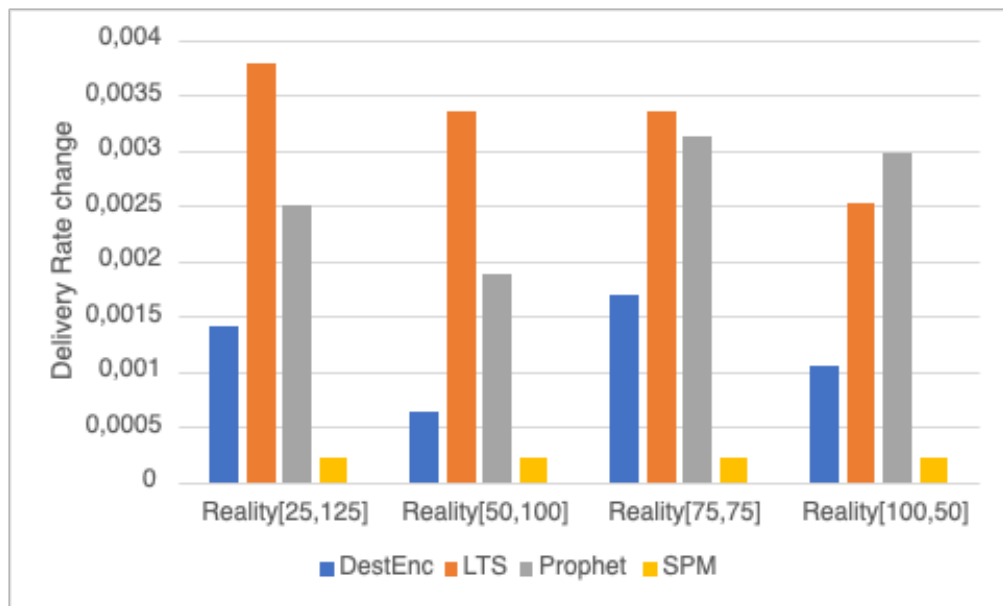
Στην εικόνα 3(β), όπου παρατίθενται τα διαγράμματα σχετικά με το Delivery Ratio Change, διαπιστώνεται ότι ο αλγόριθμος F-CbR καταγράφει μια σταθερή και θετική εξέλιξη στο ποσοστό επιτυχούς παράδοσης πακέτων σε όλες τις δοκιμασμένες αναλογίες δειγμάτων μεταξύ των δύο λιστών, δηλαδή της OneHopList και της TwoHopList. Αυτή η μείωση είναι πιο έντονη καθώς αυξάνεται το ποσοστό των δειγμάτων που εκχωρούνται στην TwoHopList, ενώ παράλληλα μειώνεται το ποσοστό των δειγμάτων για την OneHopList. Είναι ιδιαίτερα ενδιαφέρον ότι οι μετρικές Prophet και LTS διακρίνονται για την ευνοϊκή επίδρασή τους στη βελτίωση του ποσοστού παράδοσης, καθώς αυτές οι μετρικές σημειώνουν αισθητά υψηλότερες βελτιώσεις

στο ποσοστό παράδοσης σε σχέση με τις υπόλοιπες μετρικές, όπως η DestEnc και η SPM.

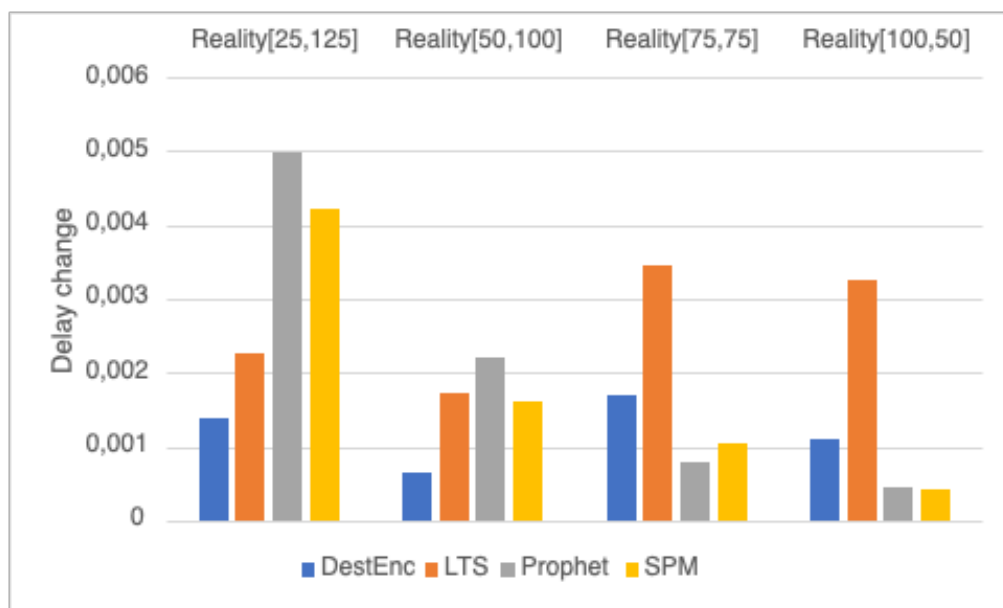
Στην εικόνα 3(γ), όπου παρουσιάζονται τα διαγράμματα που αναφέρονται στο Delay Change, παρατηρείται ότι ο αλγόριθμος F-CbR καταγράφει μια σημαντική μείωση στον χρόνο καθυστέρησης των πακέτων. Αυτή η μείωση είναι πιο έντονη καθώς αυξάνεται το ποσοστό των δειγμάτων που εκχωρούνται στην TwoHopList, ενώ παράλληλα μειώνεται το ποσοστό των δειγμάτων για την OneHopList. Συγκεκριμένα, στις περιπτώσεις όπου εφαρμόστηκαν οι μετρικές Prophet και SPM, παρατηρούνται οι μεγαλύτερες μειώσεις στον χρόνο καθυστέρησης σε σύγκριση με τις άλλες δύο μετρικές. Αυτό υποδηλώνει ότι η ρύθμιση των ποσοστών δειγμάτων μεταξύ των δύο λιστών, καθώς και η επιλογή της κατάλληλης μετρικής, μπορεί να έχει σημαντική επίδραση στην βελτίωση της απόδοσης του αλγορίθμου F-CbR όσον αφορά τον χρόνο καθυστέρησης.



Εικόνα 3(α)



Εικόνα 3(β)



Εικόνα 3(γ)

5.

Συμπεράσματα

Η διπλωματική εργασία αυτή παρουσιάζει μια σημαντική καινοτομία στον τομέα των ομορτυνιστικών δικτύων, μέσω της ανάπτυξης του αλγορίθμου Federated Learning Cluster-Based Routing (F-CbR), ο οποίος ενσωματώνει την έννοια της συνεργατικής μάθησης για την βελτιστοποίηση της διαδικασίας δρομολόγησης. Η βασική καινοτομία της εργασίας εστιάζεται στην ικανότητα του F-CbR να αξιοποιεί τα δεδομένα και τις πληροφορίες που προκύπτουν από την αλληλεπίδραση των κόμβων του δικτύου, με σκοπό την αυτόματη βελτίωση της απόδοσης της δρομολόγησης μέσω της μάθησης από τη συνεργασία των κόμβων.

Η εφαρμογή της συνεργατικής μάθησης στη δρομολόγηση επιτρέπει στον αλγόριθμο F-CbR να προσαρμόζεται στις συνεχώς μεταβαλλόμενες συνθήκες του δικτύου, ενισχύοντας την αποδοτικότητα και την αξιοπιστία της παράδοσης δεδομένων. Ο F-CbR καθιστά δυνατή την ανάπτυξη ενός πιο ευέλικτου, αποδοτικού και αυτοματοποιημένου συστήματος δρομολόγησης, το οποίο μπορεί να ανταποκρίνεται δυναμικά στις αλλαγές του δικτύου, ενισχύοντας την αποτελεσματικότητα της επικοινωνίας στα ομορτυνιστικά δίκτυα. Αυτή η πρωτοποριακή προσέγγιση τονίζει την αξία της ένταξης των προηγμένων τεχνολογιών συνεργατικής μάθησης στη βελτίωση των δικτυακών υποδομών, καθορίζοντας νέους δρόμους για την έρευνα και ανάπτυξη στον τομέα των ομορτυνιστικών δικτύων.

Με βάση τα αποτελέσματα των πειραμάτων, φαίνεται ότι ο F-CbR εμφανίζει θετικά αποτελέσματα, με ταυτόχρονη βελτίωση της ικανότητας παράδοσης και της καθυστέρησης. Ωστόσο, απαιτείται περαιτέρω διερεύνηση των παραμέτρων του αλγορίθμου, ώστε να εξεταστεί εάν μπορούν να επιτευχθούν ακόμη καλύτερες επιδόσεις, φτάνοντας σε σημείο να επιλύσουν το κύριο πρόβλημα που αντιμετωπίζουν τα ομορτυνιστικά δίκτυα, δηλαδή την παραγωγή μεγάλου αριθμού αντιγράφων.

Βιβλιογραφία

- [1] C. Boldrini, K. Lee, M. Önen, J. Ott, and E. Pagani, "Opportunistic networks," *Comput. Commun.*, vol. 48, no. 14, pp. 1–4, Mar 2014
- [2] Wood, L., Eddy, W. M., Ivancic, W., McKim, J., & Jackson, C. (2007, September). Saratoga: a Delay-Tolerant Networking convergence layer with efficient link utilization. In 2007 International Workshop on Satellite and Space Communications (pp. 168-172). IEEE.
- [3] Small, T., & Haas, Z. J. (2003, June). The shared wireless infostation model: a new ad hoc networking paradigm (or where there is a whale, there is a way). In Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing (pp. 233-244)
- [4] Saloni, M., Julien, C., Murphy, A. L., & Picco, G. P. (2017, September). Lasso: A device-to-device group monitoring service for smart cities. In 2017 International Smart Cities Conference (ISC2) (pp. 1-6). IEEE.
- [5] Martonosi, M. (2004). The princeton ZebraNet project: sensor networks for wildlife tracking. Princeton University, 2-7
- [6] Pentland A, Fletcher R, Hasson A. DakNet: rethinking connectivity in developing nations. *Computertomographie*. 2004; 37(1): 78-83.
- [7] R. Ramanathan, R. Hansen, P. Basu, R. R. Hain and R. Krishnan, "Prioritized epidemic routing for opportunistic networks", *In Proceedings of the 1st international MobiSys workshop on Mobile opportunistic networking*, pp. 62-66, 2007.
- [8] S. C. Nelson, M. Bakht, and R. Kravets, "Encounter-based routing in DTNs," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, 2009, pp. 846–854
- [9] Trifunovic, S., Kouyoumdjieva, S. T., Distl, B., Pajevic, L., Karlsson, G., & Plattner, B. (2017). A decade of research in opportunistic networks: challenges, relevance, and future directions. *IEEE Communications Magazine*, 55(1), 168-173.
- [10] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, "Diversity of forwarding paths in pocket switched networks," in *Proc. ACM SIGCOMM Internet Meas. Conf. (IMC)*, 2007, pp. 161–174
- [11] W. Moreira, P. Mendes, and S. Sargento, "Opportunistic routing based on daily routines," in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, June 2012.
- [12] A. Lindgren, A. Doria, E. Davies, and S. Grasic, "Probabilistic Routing Protocol for Intermittently Connected Networks," RFC 6693 (Experimental), RFC Editor, Fremont, CA, USA, pp. 1–113, Aug. 2012. [Online]. Available: <https://www.rfc-editor.org/rfc/rfc6693.txt>
- [13] V. Erramilli, M. Crovella, A. Chaintreau, and C. Diot, "Delegation Forwarding," in *Proc. ACM Int. Conf. Mobile Ad Hoc Netw. and Comp. (MobiHoc)*, 2008, pp. 251–260.
- [14] N. Papanikos and E. Papapetrou, "Coordinating Replication Decisions in Multi-copy Routing for Opportunistic Networks," in *Proc. IEEE Int. Conf. Wireless and Mobile Comput., Netw. and Commun. (WiMob)*, 2014, pp. 8–13
- [15] E. Papapetrou, A. Likas. (2018) Cluster-based Replication: a Forwarding Strategy for Mobile Opportunistic Networks. In 2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM), IEEE, 14-19

- [16] Bharati, S., Mondal, M., Podder, P., & Prasath, V. B. (2022). Federated learning: Applications, challenges and future directions. *International Journal of Hybrid Intelligent Systems*, 18(1-2), 19-35.
- [17] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1988.
- [18] L. Kaufman and P. Rousseeuw, *Finding Groups in Data: an introduction to cluster analysis*. Hoboken, NJ, USA: Wiley, 1990.
- [19] T. Kohonen, "Learning vector quantization," in *The Handbook of Brain Theory and Neural Networks*, 1st ed., M. A. Arbib, Ed. Cambridge, MA, USA: MIT Press, 1995.
- [20] N. Papanikos, D.-G. Akestoridis, and E. Papapetrou, "CRAWDAD toolset tools/simulate/uo/adyton (v. 2016-04-21)," Downloaded from <http://crawdad.org/tools/simulate/uo/adyton/20160421>, Apr. 2016.
- [21] T. H. David Kotz. CRAWDAD: A Community Resource for Archiving Wireless Data At Dartmouth. [Online]. Available: <http://crawdad.cs.dartmouth.edu>
- [22] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD data set cambridge/haggle (v. 2006-01-31)," Downloaded from <http://crawdad.org/cambridge/haggle/>, Jan. 2006.
- [23] A.-K. Pietilainen, "CRAWDAD data set thlab/sigcomm2009 (v. 2012- 07-15)," Downloaded from <http://crawdad.org/thlab/sigcomm2009/>, Jul. 2012.
- [24] N. Eagle and A. S. Pentland, "CRAWDAD data set mit/reality 07-01)," Downloaded from <http://crawdad.org/mit/reality/>, Jul. 2005.
- [25] S. Gaito, E. Pagani, and G. P. Rossi, "Strangers help friends to communicate in opportunistic networks," *Computer Networks*, vol. 55, no. 2, pp. 374 – 385, 2011.
- [26] J. Leguay, A. Lindgren, and T. Friedman, "CRAWDAD data set upmc/content (v. 2006-11-17)," Downloaded from <http://crawdad.org/upmc/content/>, Nov. 2006.
- [27] H. Dubois-Ferriere, M. Grossglauser, and M. Vetterli, "Age matters: efficient route discovery in mobile ad hoc networks using encounter ages," in *Proc. ACM Mobile Ad Hoc Netw. and Comput. (MobiHoc)*, 2003, pp. 257–266.
- [28] A. Lindgren, A. Doria, and O. Schele'n, "Probabilistic routing in inter- mittently connected networks," *ACM SIGMOBILE Mobile Comput. and Commun. Rev.*, vol. 7, no. 3, pp. 19–20, 2003.
- [29] E. Bulut and B. K. Szymanski, "Exploiting friendship relations for efficient routing in mobile social networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 12, pp. 2254–2265, 2012.