



# ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ, ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

## Τομογραφία και Ουδετερότητα Δικτύου

### ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

**ΔΕΣΠΟΙΝΑΣ ΓΚΑΤΖΙΟΥΡΑ**

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

Αθήνα, Οκτώβριος 2019





Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Τομέας Επικοινωνιών, Ηλεκτρονικής και Συστημάτων Πληροφορικής

## Τομογραφία και Ουδετερότητα Δικτύου

### ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

**ΔΕΣΠΟΙΝΑΣ ΓΚΑΤΖΙΟΥΡΑ**

**Επιβλέπων:** Συμεών Παπαβασιλείου

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 21η Οκτωβρίου 2019.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....

Σ. Παπαβασιλείου

Καθηγητής Ε.Μ.Π.

.....

Θ. Βαρβαρίγου

Καθηγήτρια Ε.Μ.Π.

.....

Ι. Ρουσσάκη

Επ. Καθηγήτρια Ε.Μ.Π.

Αθήνα, Οκτώβριος 2019

(Υπογραφή)

.....

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

Copyright ©–All rights reserved Δέσποινα Γκατζιούρα, 2019.

Με επιφύλαξη παντός δικαιώματος.

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



---

## Περίληψη

---

Σε μία εποχή που το Διαδίκτυο εισχωρεί σε όλο το φάσμα της ανθρώπινης καθημερινότητας, από τη διασκέδαση μέχρι την εργασία και από την αναγκαιότητα μέχρι την υπερβολή, που ο όγκος των παραγόμενων δεδομένων διαρκώς διαστέλλεται και η ανάγκη για εύκολη και γρήγορη πρόσβαση στη γνώση καθίσταται επιτακτική, εγείρονται σημαντικά ζητήματα αναφορικά με τον τρόπο με τον οποίο τα διάφορα δίκτυα που απαρτίζουν το Διαδίκτυο διαχειρίζονται την πληροφορία που τα διαπερνά. Προκύπτουν θέματα ασφάλειας, ιδιωτικότητας και δίκαιης αντιμετώπισης των διαφόρων κινήσεων δεδομένων, καθώς και ζητήματα ποιότητας των παρεχόμενων υπηρεσιών, διαφάνειας και συμμόρφωσης με τις εκάστοτε συμβάσεις.

Έτσι φτάνουμε στην ανάγκη ορισμού της *Ουδετερότητας Δικτύου*, η οποία σχετίζεται με την ισότιμη αντιμετώπιση όλων των κινήσεων δεδομένων από κάποιο δίκτυο, δίχως διακρίσεις λόγω περιεχομένου, προέλευσης ή προορισμού. Με σκοπό να εξασφαλιστούν τα παραπάνω αναπτύσσονται τεχνικές *Τομογραφίας Δικτύου*, οι οποίες προσφέρουν τη δυνατότητα συναγωγής της τοπολογίας και των εσωτερικών χαρακτηριστικών ενός δικτύου, με γνώση μόνον ακραίων μετρήσεων, δηλαδή μετρήσεων που πραγματοποιούνται στα άκρα του δικτύου αυτού.

Στην παρούσα διπλωματική έγινε εκτενής μελέτη ποικίλων τεχνικών Τομογραφίας Δικτύου, καθώς και ορισμένων προσεγγίσεων της Ουδετερότητας Δικτύου. Εν συνεχεία, υλοποιήθηκε ο αλγόριθμος *Urgma*, ο οποίος ανήκει στην οικογένεια των *Αλγορίθμων Ο-*

μαδοποίησης και επιτρέπει την ανακατασκευή της δενδρικής τοπολογίας και τον υπολογισμό των χαρακτηριστικών των ακμών αυτής, όταν τα χαρακτηριστικά αυτά ικανοποιούν την υπερμετρική ιδιότητα. Εν προκειμένω, τα χαρακτηριστικά που μελετήθηκαν είναι η διακύμανση των καθυστερήσεων και οι απώλειες των ακμών. Τέλος, έγινε χρήση του συστήματος δοκιμών που υποστηρίζει το *jFed*, για την κατασκευή δενδρικών, υπερμετρικών τοπολογιών και τη συλλογή ακραίων μετρήσεων, οι οποίες χορηγήθηκαν στον Upsilon, οδηγώντας σε πλήρη και απόλυτα ακριβή ανακατασκευή των τοπολογιών, καθώς και σε πολύ ακριβείς εκτιμήσεις των χαρακτηριστικών όλων των ακμών των τοπολογιών αυτών.

**Λέξεις Κλειδιά:** Τομογραφία Δικτύου, Ουδετερότητα Δικτύου, μετρικές εύρους ζώνης, ανακατασκευή τοπολογίας, εκτίμηση μετρικών απόδοσης ζεύξεων, μετρήσεις απ'άκρη σ'άκρη, υπερμετρικά δένδρα, *jFed*, UPGMA.

---

## *Abstract*

---

In an era where the Internet is penetrating into the whole spectrum of humans' everyday life -from entertainment to work and from necessity to exaggeration-, when a vast amount of data is constantly being generated and the information-centric societies render the quick and easy access to knowledge crucial, important questions are being raised relating to the way information is being handled by the various networks that compose the Internet. Security, privacy and fair treatment of data traffic issues, as well as issues regarding Quality of Service (QoS), transparency and compliance with various agreements, arise.

As a result, the need to define *Network Neutrality*, as the equal treatment of all data traffic that goes through a network by the network itself, without discrimination based on its content, origin or destination, occurs. In order for all the above mentioned issues to be ensured, *Network Tomography* techniques are being developed. These techniques enable the inference of the topology and the internal characteristics of a network, based only on end-to-end measurements.

In the framework of this diploma thesis, various techniques in the field of Network Tomography, along with different approaches concerning Network Neutrality, have been studied. Thereafter the clustering algorithm, *Upgma*, was implemented, allowing the inference of ultrametric topologies and their internal edges' characteristics, more con-

cretely, their losses and jitter. At last, *jFed* testbed was used in order to construct ultrametric, tree topologies and realize end-to-end measurements, which were later provided to the Upgma algorithm, allowing it to fully reconstruct the topologies, as well as accurately estimate the metrics of all their edges.

**Keywords:** Network Tomography, Network Neutrality, bandwidth metrics, topology inference, link performance parameters estimation, end-to-end measurements, ultrametric trees, jFed, UPGMA.

---

## Ευχαριστίες

---

Πρώτα από όλα, θα ήθελα να ευχαριστήσω τον καθηγητή μου, κύριο Συμεών Παπαβασιλείου, ο οποίος με τη σιγουριά του, την ηρεμία του και την εμπιστοσύνη που μου έδειξε, με ώθησε να δουλέψω, με αυτοπεποίθηση και θετική διάθεση. Διότι είναι δύσκολο σε έναν τομέα τόσο ανταγωνιστικό και απαιτητικό, να βρεις ανθρώπους που θα σου δώσουν το θάρρος να πιστέψεις στον εαυτό σου, να δοκιμάσεις και να πειραματιστείς.

Εν συνεχεία, θα ήθελα να ευχαριστήσω θερμά τον Αναπληρωτή Καθηγητή του Τμήματος Πληροφορικής του Ιόνιου Παν/μίου Βασίλη Καρυώτη, καθώς και τον διδακτορικό φοιτητή Γρηγόρη Κακαάβα, για τον χρόνο που μου αφιέρωσαν, τη βοήθειά που απλόχερα μού προσέφεραν, και φυσικά, το φιλικό και ευχάριστο κλίμα που δημιούργησαν και το οποίο πλαισίωσε τη συνεργασία μας, προσδίδοντάς της μία πορεία ομαλή και εξελικτική.

Θα ήθελα ακόμη να ευχαριστήσω τους γονείς μου, για τα εφόδια που μού εξασφάλισαν και τη στήριξη που μού προσέφεραν, καθώς και τον αδερφό μου, Δημήτρη, για την παρέα, τις συζητήσεις και τις γαστρονομικές εξερευνήσεις.

Έπειτα, δε θα γινόταν να μην ευχαριστήσω τις φίλες και τους φίλους μου, με τις/τους οποίες/ους μοιράστηκα τις ομορφότερες, αλλά και τις δυσκολότερες, στιγμές της φοιτητικής μου ζωής, καθιστώντας την, τελικά, ένα μοναδικό μίγμα συζητήσεων, συναισθημάτων, αναζητήσεων και μεταβολών. Τέλος, θα ήθελα να ευχαριστήσω τους/τις συμφοιτητές/τριες μου, για την πεντάχρονη αυτή συμβίωση στα αμφιθέατρα του Πολυτεχνείου.



---

## Περιεχόμενα

---

|  |           |
|--|-----------|
| Περίληψη   | 5         |
| Abstract   | 7         |
| Ευχαριστίες  | 9         |
| Περιεχόμενα  | 13        |
| <b>1 Εισαγωγή</b>  | <b>15</b> |
| 1.1 Πρόλογος . . . . .   | 15        |
| 1.2 Αντικείμενο της Διπλωματικής . . . . .   | 16        |
| 1.3 Δομή της Διπλωματικής . . . . .  | 16        |
| <b>2 Τομογραφία Δικτύου και Βασικοί Ορισμοί</b>  | <b>17</b> |
| 2.1 Γενικά . . . . .   | 17        |
| 2.2 Ορισμός του προβλήματος . . . . .  | 18        |
| 2.3 Παρουσίαση τομογραφικών μεθόδων . . . . .  | 20        |
| 2.3.1 Αλγόριθμοι για την εύρεση των συμφορημένων ακμών του δικτύου .   | 21        |
| 2.3.2 Αλγόριθμοι για την εύρεση της τοπολογίας του δικτύου . . . . .   | 26        |
| 2.3.3 Αλγόριθμοι για την παρακολούθηση δικτύων μέσω της επιλογής ενός<br>συνόλου μονοπατιών - monitoring paths . . . . . | 41        |

|          |  |           |
|----------|--|-----------|
| 2.4      | Παρουσίαση των σημαντικότερων μετρικών ενός δικτύου . . . . .  | 44        |
| 2.4.1    | Χωρητικότητα . . . . .   | 45        |
| 2.4.2    | Διαθέσιμο εύρος ζώνης και μέση χρησιμοποίηση . . . . .   | 46        |
| 2.4.3    | TCP throughput και bulk transfer capacity . . . . .  | 47        |
| 2.5      | Παρουσίαση των διαφόρων τεχνικών/εργαλείων για την εκτίμηση των μετρί-<br>κών ενός δικτύου . . . . .                               | 48        |
| 2.5.1    | Τεχνικές για την εκτίμηση των μετρικών ενός δικτύου . . . . .  | 48        |
| 2.5.2    | Εργαλεία για την εκτίμηση των μετρικών ενός δικτύου . . . . .  | 56        |
| <b>3</b> | <b>Ουδετερότητα Δικτύου</b>  | <b>61</b> |
| 3.1      | Γενικά . . . . .   | 61        |
| 3.2      | Παρουσίαση των διαφορετικών προσεγγίσεων του προβλήματος της Ουδετε-<br>ρότητας Δικτύου . . . . .                                  | 63        |
| 3.2.1    | Κατηγοριοποίηση των δικτυακών κινήσεων ανάλογα με τις απαιτήσεις<br>τους σε εύρος ζώνης . . . . .                                  | 63        |
| 3.2.2    | Μοντελοποίηση των ουδέτερων και μη δικτύων με χρήση Θεωρίας<br>Παιγνίων . . . . .  | 66        |
| 3.2.3    | Ανίχνευση παραβιάσεων της ουδετερότητας σε δίκτυα και εντοπισμός<br>των μη-ουδέτερων ακμών βάσει εξωτερικών παρατηρήσεων . . . . . | 70        |
| 3.2.4    | Διαφάνεια στη θέση της ουδετερότητας . . . . .   | 73        |
| <b>4</b> | <b>Παρουσίαση του μοντέλου και του πειραματικού σχήματος</b>   | <b>77</b> |
| 4.1      | Περιγραφή και μοντελοποίηση του προβλήματος . . . . .  | 77        |
| 4.1.1    | Προσθετικοί και υπερμετρικοί χώροι . . . . .   | 78        |
| 4.1.2    | Κατασκευή του πίνακα αποστάσεων βάσει των απωλειών και των κα-<br>θυστερήσεων των ακμών . . . . .                                  | 79        |
| 4.2      | Κατασκευή τοπολογιών και συλλογή μετρήσεων μέσω του jFed . . . . .   | 81        |
| <b>5</b> | <b>Παρουσίαση του αλγορίθμου και των αποτελεσμάτων</b>   | <b>83</b> |
| 5.1      | Unweighted Pair Group Method with Arithmetic mean - UPGMA . . . . .  | 83        |



---

|          |   |            |
|----------|---|------------|
| 5.2      | Αποτελέσματα και σχολιασμός . . . . .                 | 85         |
| <b>6</b> | <b>Επίλογος</b>                                       | <b>89</b>  |
| <b>7</b> | <b>Ανοιχτά προβλήματα και συναφής βιβλιογραφία</b>    | <b>91</b>  |
|          | <b>Βιβλιογραφία</b>                                   | <b>95</b>  |
|          | <b>Παράρτημα</b>                                      | <b>102</b> |
| A        | Τοπολογίες . . . . .                                  | 102        |
| B        | Πηγαίος κώδικας . . . . .                             | 122        |
| C        | Θεμελιώδη στοιχεία των Markov Random Fields . . . . . | 132        |



## Κεφάλαιο 1

---

### Εισαγωγή

---

#### 1.1 Πρόλογος

Η έκρηξη του μεγέθους του Διαδικτύου που λαμβάνει χώρα τις τελευταίες δεκαετίες ανοίγει το δρόμο σε μία κοινωνία και έναν τρόπο ζωής που εστιάζουν στην εύκολη και άμεση πρόσβαση σε μία πληθώρα πληροφοριών.

Μέσα σε αυτό το πλαίσιο, καλούμαστε να βρούμε τρόπους ώστε, όχι μόνο να καταστήσουμε την πρόσβαση αυτή δυνατή, αλλά και να εξασφαλίσουμε την ισότητα, τη δικαιοσύνη, την εγκυρότητα, την ασφάλεια και την ιδιωτικότητα όλων των εμπλεκομένων και ιδίως των τελικών χρηστών που δε διαθέτουν τεχνικές γνώσεις σχετικές με τις υπηρεσίες που λαμβάνουν.

Βέβαια, η ακατάπαυστη κλιμάκωση του μεγέθους του Διαδικτύου, ο αποκεντρωμένος χαρακτήρας του, τα αντικρουόμενα συμφέροντα των εμπλεκομένων, η επιτακτική ανάγκη για ιδιωτικότητα και οι τεράστιες απαιτήσεις σε εύρος ζώνης (τεράστια, και συνεχώς αυξανόμενη, ποικιλία διαφορετικών τύπων κινήσεων δεδομένων, με διαφορετικές απαιτήσεις και ευαισθησίες / πεπερασμένο εύρος ζώνης) καθιστούν την εξασφάλιση όσων αναφέρθηκαν παραπάνω ιδιαίτερα δύσκολη και απαιτούν την εύρεση και το σχεδιασμό νέων μεθόδων, εργαλείων και κανονισμών.

Έτσι οδηγούμαστε στα, σχετικά νέα και ανθίζοντα, πεδία της *Τομογραφίας και Ουδέτε-*

ρότητας Δικτύου, τα οποία παρέχουν καλύτερη κατανόηση της δομής και των χαρακτηριστικών που ενυπάρχουν στο εσωτερικό των δικτύων, προσφέροντας, παράλληλα, καινοτόμες τεχνικές για την εξασφάλιση της ομαλής και έννομης λειτουργίας του Διαδικτύου και των υπηρεσιών που αυτό παρέχει.

## 1.2 Αντικείμενο της Διπλωματικής

Το αντικείμενο της παρούσας διπλωματικής συντίθεται τόσο από τη μελέτη των διαφόρων τεχνικών Τομογραφίας Δικτύου, καθώς και των διαφόρων προσεγγίσεων της Ουδετερότητας Δικτύου, όσο και από την εφαρμογή μίας τεχνικής Τομογραφίας Δικτύου, μέσω της διεξαγωγής πειραμάτων στο εργαλείο jFed και της υλοποίησης του αλγορίθμου UPGMA.

## 1.3 Δομή της Διπλωματικής

Στο κεφάλαιο **2** παρουσιάζονται και αναλύονται ποικίλλες τεχνικές Τομογραφίας Δικτύου, μετρικές που χρησιμοποιούνται για τον χαρακτηρισμό των ακμών των δικτύων (χωρητικότητα, εύρος ζώνης κλπ.), καθώς και μέθοδοι και εργαλεία που επιτρέπουν την εκτίμηση των μετρικών αυτών. Στο κεφάλαιο **3** παρουσιάζονται και αναλύονται διάφορες όψεις του προβλήματος της Ουδετερότητας Δικτύου. Στο κεφάλαιο **4** περιγράφεται το μοντέλο και η διαδικασία που ακολουθήθηκε για τη διεξαγωγή των πειραμάτων, καθώς και το εργαλείο που χρησιμοποιήθηκε. Στο κεφάλαιο **5** παρουσιάζεται ο αλγόριθμος που υλοποιήθηκε και τα αποτελέσματα που παράχθηκαν. Στο κεφάλαιο **6** εκτίθενται τα συμπεράσματα που προκύπτουν από την παρούσα διπλωματική και τέλος, στο κεφάλαιο **7** παρουσιάζονται ορισμένα ανοιχτά προβλήματα που εμπίπτουν στο πεδίο της Τομογραφίας και της Ουδετερότητας Δικτύου, καθώς και συναφής βιβλιογραφία.

## Κεφάλαιο 2

---

### Τομογραφία Δικτύου και Βασικοί Ορισμοί

---

#### 2.1 Γενικά

Τα τελευταία χρόνια το Διαδίκτυο έχει μετατραπεί σε ένα μαζικό, κατανομημένο δίκτυο, με δισεκατομμύρια χρήστες, του οποίου το μέγεθος συνεχώς αυξάνεται, καθώς διάφορες δραστηριότητες, όπως το ηλεκτρονικό εμπόριο, εξαπλώνονται.

Η ανομοιογενής και ακαθόριστη δομή του καθιστά ιδιαίτερες απαιτητικές ορισμένες λειτουργίες, όπως η δυναμική δρομολόγηση, η ανίχνευση ανωμαλιών ή/και κακόβουλων κινήσεων, η διασφάλιση της ποιότητας παροχής υπηρεσιών στους τελικούς χρήστες, η διασφάλιση της τήρησης των συμφωνιών μεταξύ των παρόχων υπηρεσιών ή/και των παρόχων υπηρεσιών με τους τελικούς χρήστες (SLAs), κ.α.

Επιπλέον, τα προηγούμενα σε συνδυασμό με την αποκεντρωμένη φύση του καθιστούν αδύνατη τη δημιουργία ενός συντονισμένου πλαισίου για την πραγματοποίηση εξειδικευμένων μετρήσεων. Ένα τέτοιο πλαίσιο θα έπρεπε να βασίζεται στη συνεργασία των διαφόρων εξυπηρετητών και δρομολογητών για τη συλλογή και τη διανομή σημαντικών στατιστικών δεδομένων του δικτύου (π.χ., traffic rates, link delays, dropped packet rates). Ο λόγος που καθιστά τη συνεργασία αυτή ανέφικτη είναι ότι κάτι τέτοιο θα εισήγαγε υπερβολικές επιπλέον απαιτήσεις τόσο στο κομμάτι των υπολογισμών και της επικοινωνίας μεταξύ των συσκευών του δικτύου για τον υπολογισμό και την ανταλλαγή των παραπάνω δεδομένων,

όσο και στο κομμάτι του χρησιμοποιούμενου λογισμικού και υλισμικού.

Επιπροσθέτως, στο σημείο αυτό προκύπτουν θέματα ιδιωτικότητας και προστασίας δεδομένων που θεωρούνται εμπιστευτικά, καθώς επίσης προκύπτει και η ανάγκη για πολιτικές τεχνικής και οικονομικής διαχείρισης των μεγαλύτερων ή μικρότερων ιδιωτικών δικτύων που απαρτίζουν το Διαδίκτυο.

Συνεπώς, κατ'αυτόν τον τρόπο, το πρόβλημα μετατρέπεται σε ένα αντίστροφο πρόβλημα, δηλαδή ένα πρόβλημα του οποίου οι βασικές πτυχές (τα εσωτερικά χαρακτηριστικά) δεν είναι άμεσα παρατηρήσιμες και συνεπώς απαιτούνται εργαλεία/τεχνικές για τον καθορισμό μοντέλων, τα οποία βασιζόμενα σε ακραίες (end-to-end) μετρήσεις συνάγουν την τοπολογία του δικτύου, καθώς και τις ζητούμενες παραμέτρους/χαρακτηριστικά (inverse problems, inference problems). Οι τεχνικές αυτές αποτελούν *τεχνικές τομογραφίας δικτύου* και δεν απαιτούν γνώσεις για την εσωτερική τοπολογία του δικτύου ούτε και πραγματοποίηση μετρήσεων από το σύνολο των εσωτερικών κόμβων.

Προβλήματα συγγενικά με το πρόβλημα της τομογραφίας δικτύου είναι το πρόβλημα της τομογραφικής ανακατασκευής εικόνας και της ανακατασκευής φυλογενετικού δέντρου. Στην περίπτωση της τομογραφικής ανακατασκευής εικόνας, η υπό-μελέτη φυσική ποσότητα είναι απρόσιτη, συνεπώς η εκτίμηση της κατανομής της μπορεί να επιτευχθεί μόνο μέσω της πληροφορίας που εξάγεται από τις προβολές/τομές της σε διάφορες κατευθύνσεις [1]. Όσον αφορά στα φυλογενετικά δέντρα, η κατασκευή τους αποσκοπεί στην ανακατασκευή της εξελικτικής ιστορίας των ειδών, μέσω της αναπαράστασης των συγγενικών σχέσεων μεταξύ αυτών [2].

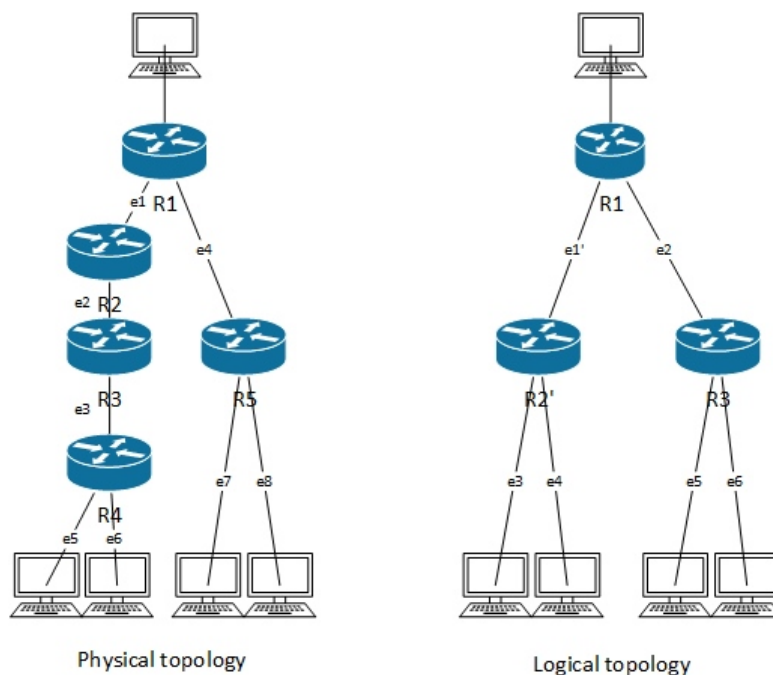
## 2.2 Ορισμός του προβλήματος

Το Διαδίκτυο μοντελοποιείται ως ένας γράφος, άλλοτε κατευθυνόμενος και άλλοτε όχι, ανάλογα με τα χαρακτηριστικά λειτουργίας του δικτύου, τα οποία αναγκάζουν και την αντίστοιχη τεχνική να προσαρμοστεί και να κάνει διαφορετική θεώρηση για το μοντέλο. Οι κόμβοι του γράφου αντιπροσωπεύουν τα στοιχεία του Διαδικτύου (τελικοί χρήστες, δρομολογητές, μεταγωγείς) και οι ακμές του τις λογικές συνδέσεις μεταξύ των στοιχείων αυτών.

Οι λογικές συνδέσεις μπορεί να αποτελούνται από μία ή και περισσότερες φυσικές συνδέσεις, όπως στην περίπτωση που συνδέονται μεταξύ τους διαδοχικά στοιχεία χωρίς να δημιουργούνται διακλαδώσεις (βλ. Εικόνα 2.1). Ένα σύνολο διαδοχικών ακμών αποτελεί ένα μονοπάτι του δικτύου.

Η εύρεση της τοπολογίας του δικτύου μπορεί να πραγματοποιηθεί είτε με τη συνεργασία των εσωτερικών του κόμβων, με χρήση της εντολής traceroute, είτε χωρίς αυτή, με την αποστολή διερευνητικών πακέτων (probes) προς ορισμένους τελικούς κόμβους, δηλαδή κόμβους δίχως παιδιά (φύλλα του δέντρου).

Επιπλέον, δεδομένης της τοπολογίας, με χρήση διαφόρων πιθανοτικών αλγορίθμων και στατιστικών μοντέλων, οδηγούμαστε στον υπολογισμό των ζητούμενων χαρακτηριστικών του δικτύου.



Εικόνα 2.1: Σύγκριση μίας φυσικής τοπολογίας (αριστερά) με την αντίστοιχη της λογική τοπολογία (δεξιά). Η λογική τοπολογία προκύπτει από αντικατάσταση των τριών δρομολογητών R2, R3, R4 από έναν (R2') και των τριών φυσικών ακμών e1, e2, e3 από μία λογική (e1').

## 2.3 Παρουσίαση τομογραφικών μεθόδων

Οι τομογραφικές μέθοδοι, όπως παρουσιάζονται στην παρούσα ενότητα, διαρθρώνονται στις ακόλουθες τρεις (3) βασικές κατηγορίες:

- Μέθοδοι που αποσκοπούν στη *συναγωγή των συμφορημένων ακμών ενός δικτύου* (σε μία δεδομένη χρονική στιγμή ή σε ένα χρονικό διάστημα), γνωρίζοντας τον γράφο του δικτύου και τα συμφορημένα μονοπάτια του (με ή χωρίς υπόθεση για ανεξαρτησία των ακμών).
- Μέθοδοι που αποσκοπούν στη *συναγωγή της τοπολογίας ενός δικτύου* διαθέτοντας μόνο μετρήσεις προερχόμενες από τα άκρα του (ακραίες μετρήσεις), είτε με χρήση στατιστικών μεθόδων (MLE, MPLE, MCMC), είτε με χρήση συσσωρευτικών αλγορίθμων. Οι συσσωρευτικοί αλγόριθμοι μπορεί να βασίζονται είτε στην ανακατασκευή της τοπολογίας μέσω της εύρεσης των κόμβων με τη μεγαλύτερη ομοιότητα (αδέρφια) εντός ενός συνόλου τελικών χρηστών με DFS διάταξη, είτε στην κατασκευή πινάκων αποστάσεων, όπου οι αποστάσεις και τα μήκη των ακμών θα υπολογίζονται βάσει προσθετικών μετρικών και θα αντιστοιχούν σε κάποια χαρακτηριστικά των ακμών, όπως οι απώλειες και οι καθυστερήσεις, είτε στην κατασκευή πινάκων μετάβασης, θεωρώντας τον γράφο ως ένα Markov Random Field. Στις δύο τελευταίες περιπτώσεις, η χρήση των μετρικών για την περιγραφή των μηκών των ακμών επιτρέπει και τη *συναγωγή των επιμέρους χαρακτηριστικών τους*.
- Μέθοδοι που αποσκοπούν στη *συναγωγή των χαρακτηριστικών όλων των μονοπατιών ενός δικτύου*, μέσω της επιλογής ενός συνόλου γραμμικά ανεξάρτητων μονοπατιών (monitoring paths) και της παρακολούθησης των επιθυμητών χαρακτηριστικών στα μονοπάτια του συνόλου αυτού.



### 2.3.1 Αλγόριθμοι για την εύρεση των συμφορημένων ακμών του δικτύου

#### Boolean Inference και Bayesian Inference αλγόριθμοι και ο αλγόριθμος Congestion Probability Inference

Στο [3] παρουσιάζονται τρεις αλγόριθμοι, οι *Sparsity*, *Bayesian - Independence* και *Bayesian - Correlation*, καθώς, επίσης, προτείνεται και ένας νέος αλγόριθμος, ο *Congestion Probability Inference*.

Οι τρεις πρώτοι αλγόριθμοι ανήκουν στην κατηγορία των *Boolean Inference* αλγορίθμων. Για την υλοποίηση των αλγορίθμων αυτών το δίκτυο μοντελοποιείται ως ένας κατευθυνόμενος, ακυκλικός γράφος. Επιπλέον, ο χρόνος διαιρείται σε ίσα διαστήματα και πραγματοποιούνται πειράματα που περιλαμβάνουν ένα πεπερασμένο σύνολο τέτοιων διαστημάτων. Η κατάσταση συμφόρησης (congestion status) κάθε ακμής και κάθε μονοπατιού μοντελοποιούνται ως στατικές τυχαίες διαδικασίες, όπως φαίνεται στη συνέχεια.

$$X_{e_i}(t) = \begin{cases} 1, & \text{if link } e_i \text{ is congested during time interval } t \\ 0, & \text{otherwise.} \end{cases}$$

$$Y_{p_i}(t) = \begin{cases} 1, & \text{if path } p_i \text{ is congested during time interval } t \\ 0, & \text{otherwise.} \end{cases}$$

Μία ακμή (μονοπάτι) είναι καλή (καλό) κατά τη διάρκεια ενός χρονικού διαστήματος, εάν ο ρυθμός απώλειας πακέτων σε αυτήν (αυτό) είναι μικρότερος από κάποιο κατώφλι. Διαφορετικά, στην ακμή (στο μονοπάτι) υπάρχει συμφόρηση. Επίσης, οι αλγόριθμοι που ανήκουν στη συγκεκριμένη κατηγορία λειτουργούν υπό ορισμένες συνθήκες και υποθέσεις, όπως η αναγνωρισιμότητα, η διαχωρισιμότητα και η ανεξαρτησία των ακμών.

**Συνθήκη 2.3.1. Αναγνωρισιμότητα:** Δύο οποιεσδήποτε ακμές δεν διασχίζονται από τα ίδια ακριβώς μονοπάτια.

**Υπόθεση 2.3.2. Διαχωρισιμότητα:** Ένα μονοπάτι είναι καλό αν και μόνο αν όλες οι ακμές που διασχίζει είναι καλές.

**Υπόθεση 2.3.3. Ανεξαρτησία:** Όλες οι ακμές είναι ανεξάρτητες.

Η αναγνωρισιμότητα των ακμών εξασφαλίζει τη δυνατότητα κατασκευής εξισώσεων με μοναδική λύση, επειδή επιτρέπει τον προσδιορισμό ενός συνόλου μονοπατιών που θα παρουσιάζει συμφύρηση, όταν και μόνο όταν μία συγκεκριμένη (μοναδική) ακμή παρουσιάζει συμφύρηση. Εάν δύο ακμές διασχίζονταν από τα ίδια ακριβώς μονοπάτια δεν θα ήταν εφικτό, σε περίπτωση συμφύρησης του συνόλου αυτών των μονοπατιών, να προσδιοριστεί σε ποια ακμή εκ των δύο οφείλεται η συμφύρηση (ούτε θα μπορούσε να βρεθεί κάποιο άλλο σύνολο μονοπατιών, η συμφύρηση του οποίου θα διαφοροποιούσε τις ακμές, εφόσον, βάσει της υπόθεσης, δεν υπάρχει μονοπάτι που να διασχίζει τη μία και όχι την άλλη). Όσο για τη διαχωρισιμότητα, είναι προφανές ότι εάν μία ακμή παρουσιάζει συμφύρηση, τότε όλα τα μονοπάτια που τη διασχίζουν θα παρουσιάζουν επίσης συμφύρηση.

Δεδομένων των παραπάνω ορίζεται το ακόλουθο πρόβλημα: δοθέντος του γράφου του δικτύου, ενός χρονικού διαστήματος  $t$  και του συνόλου των συμφορημένων μονοπατιών, ζητούμε να συνάγουμε το σύνολο των συμφορημένων ακμών, κατά τη διάρκεια του διαστήματος αυτού. Το πρόβλημα αυτό είναι ένα κακώς ορισμένο πρόβλημα, δηλαδή δεν υπάρχει αλγόριθμος που να μπορεί να το λύσει ακριβώς για κάθε είσοδο.

Έχοντας ορίσει το πρόβλημα μένει η παρουσίαση των προαναφερθέντων αλγορίθμων.

- *Sparsity*: Η ιδέα στην οποία βασίζεται ο αλγόριθμος αυτός είναι ότι λίγες συμφορημένες ακμές είναι υπεύθυνες για πολλά συμφορημένα μονοπάτια. Διαισθητικά αυτό σημαίνει ότι σε όσα περισσότερα συμφορημένα μονοπάτια ανήκει μία ακμή, τόσο πιθανότερο είναι να σημειωθεί ως συμφορημένη.
- *Bayesian - Independence*: Ο αλγόριθμος αυτός μετατρέπει το πρόβλημα του Boolean Inference σε πρόβλημα Εκτίμησης Μέγιστης Πιθανοφάνειας (Bayesian Inference): από όλες τις πιθανές λύσεις προσπαθεί να επιλέξει αυτή που συμβαίνει με τη μεγαλύτερη πιθανότητα.

- *Bayesian - Correlation*: Ο αλγόριθμος αυτός αποτελεί βελτιωμένη έκδοση του προηγούμενου δίχως την απαίτηση για ανεξαρτησία των ακμών, αφού ορίζει σύνολα συσχέτισης, τα οποία περιέχουν τις ακμές που είναι πιθανό να εμφανίζουν συσχέτιση (για παράδειγμα, επειδή ανήκουν στο ίδιο Αυτόνομο Σύστημα).
- *Congestion Probability Inference*: Ο αλγόριθμος αυτός θέτει προς επίλυση ένα πρόβλημα ευκολότερο του Boolean Inference: την πιθανότητα κάθε ακμή να είναι συμφορημένη κατά τη διάρκεια ενός συνόλου διαδοχικών χρονικών διαστημάτων και όχι ενός μόνο χρονικού διαστήματος (στις παραπάνω τυχαίες μεταβλητές το χρονικό διάστημα  $t$  αντικαθίσταται από ένα σύνολο χρονικών διαστημάτων  $\pi$ , που ονομάζεται στιγμιότυπο). Επίσης, ομαδοποιεί τις ακμές σε σύνολα συσχέτισης, τα οποία, όπως προαναφέρθηκε, αποτελούνται από ακμές που μπορεί να παρουσιάζουν συσχέτιση μεταξύ τους αλλά όχι με ακμές που ανήκουν σε διαφορετικά σύνολα (π.χ. ακμές που αντιστοιχούν σε συνδέσεις εντός του ίδιου Αυτόνομου Συστήματος θεωρείται ότι αποτελούν ένα σύνολο συσχέτισης). Συνεπώς, δεν απαιτεί την ανεξαρτησία των ακμών και μπορεί να θεωρηθεί ως ένα καλά ορισμένο πρόβλημα.

Χρησιμοποιούνται δύο ειδών τοπολογίες για τον έλεγχο των αλγορίθμων, οι Sparse και Brite. Η δεύτερη τοπολογία μοντελοποιεί μία πλήρη τοπολογία επιπέδου Αυτόνομου Συστήματος, ενώ η πρώτη χρησιμοποιεί τα αποτελέσματα της εντολής traceroute για την κατασκευή μίας τοπολογίας επιπέδου IP. Για το λόγο αυτό η τοπολογία Brite οδηγεί στη δημιουργία ενός περισσότερο ‘πυκνού’ γράφου.

Συμπερασματικά, ύστερα από δοκιμές των παραπάνω αλγορίθμων, οι 2 πρώτοι λειτουργούν το ίδιο καλά ή κακά, αναλόγως την τοπολογία, δίνοντας καλύτερα αποτελέσματα όταν εφαρμόζονται στις Brite τοπολογίες, ενώ ο τρίτος αλγόριθμος λειτουργεί καλύτερα και από τους δύο πρώτους, αλλά και πάλι με τους περιορισμούς που θέτει το πρόβλημα του Boolean Inference. Πληροφορίες για την επίδοση του τέταρτου αλγορίθμου θα δοθούν στη συνέχεια.

Στο [4] παρουσιάζεται αναλυτικά ο αλγόριθμος *Congestion Probability Inference*, ο οποίος, με είσοδο ακραίες μετρήσεις σε ένα στιγμιότυπο χρόνου  $\pi$  (ο χρόνος διαιρείται σε ίσα

χρονικά διαστήματα που ονομάζονται στιγμιότυπα), υπολογίζει την πιθανότητα κάθε ακμή να είναι συμφορημένη στο στιγμιότυπο αυτό, δίχως την απαίτηση για ανεξαρτησία των ακμών, αφού ορίζονται σύνολα συσχέτισης. Η συσχέτιση των ακμών οφείλεται είτε σε χρήση κοινών πόρων του δικτύου που παρουσιάζουν συμφόρηση, είτε σε κάποιες κινήσεις δεδομένων που εμπλέκουν ορισμένες ακμές, όπως, π.χ., μια Επίθεση Κατανεμημένης Άρνησης Υπηρεσιών (DDoS Attack).

Ο παραπάνω αλγόριθμος στηρίζεται στην αναγνώριση όλων των ασυσχέτιστων μονοπατιών και έπειτα, όλων των ασυσχέτιστων ζευγών μονοπατιών, και στον σχηματισμό γραμμικά ανεξάρτητων εξισώσεων, που συνδέουν την πιθανότητα ενός μονοπατιού (ζεύγους μονοπατιών) να είναι καλό (καλά), με την πιθανότητα όλες οι ακμές του μονοπατιού αυτού (των ζευγών αυτών) να είναι καλές (ιδιότητα της διαχωρισιμότητας).

Για παράδειγμα, εάν θεωρήσουμε ένα μονοπάτι  $P_1$  που αποτελείται από δύο ασυσχέτιστες ακμές  $e_1$  και  $e_2$ , τότε η πιθανότητα το μονοπάτι αυτό να είναι καλό, δίνεται από την ακόλουθη σχέση:

$$\begin{aligned} \mathbb{P}(Y_{P_1} = 0) &= \mathbb{P}(X_{e_1} = 0)\mathbb{P}(X_{e_2} = 0) \Leftrightarrow \\ \log \mathbb{P}(Y_{P_1} = 0) &= \log (\mathbb{P}(X_{e_1} = 0)\mathbb{P}(X_{e_2} = 0)) \\ &= \log \mathbb{P}(X_{e_1} = 0) + \log \mathbb{P}(X_{e_2} = 0) \Leftrightarrow \\ & \quad y_1 = x_1 + x_2 \end{aligned} \tag{2.1}$$

Ανάλογα με τον βαθμό συσχέτισης του προβλήματος, ο βαθμός του συστήματος γραμμικά ανεξάρτητων εξισώσεων που θα προκύψει θα είναι ίσος ή μικρότερος του πλήθους των ακμών του γράφου, οπότε και θα έχουμε μοναδική ή πολλαπλές, αντίστοιχα, λύσεις. Στη δεύτερη περίπτωση επιλέγεται η λύση που ελαχιστοποιεί το απόλυτο σφάλμα.

Για τον έλεγχο αυτού του αλγορίθμου κατασκευάστηκαν τοπολογίες τόσο με χρήση του Brite Topology Generator, όσο και στην πλατφόρμα δοκιμών PlanetLab. Τα αποτελέσματα επιβεβαίωσαν την ακρίβεια του παραπάνω αλγορίθμου, ακόμη και σε περιπτώσεις όπου δεν είναι γνωστά όλα τα μοτίβα συσχέτισης. Επίσης, έπειτα από σύγκριση με κάποιον αλγόριθμο που απαιτεί την ανεξαρτησία των ακμών, φάνηκε πως ο Congestion Probability Inference

έχει καλύτερη επίδοση, ακόμη και σε περιπτώσεις όπου τα επίπεδα συμφόρησης του δικτύου είναι χαμηλά και οι συμφορημένες ακμές παρουσιάζουν μικρή συσχέτιση.

### Ο αλγόριθμος Congestion Probability Computation

Στο [5] παρουσιάζεται ο αλγόριθμος *Congestion Probability Computation*, ο οποίος δοθέντος του γράφου του δικτύου και του συνόλου των συμφορημένων μονοπατιών σε ένα στιγμιότυπο  $\pi$ , υπολογίζει την πιθανότητα συμφόρησης κάθε συνόλου ακμών, δηλαδή την πιθανότητα όλες οι ακμές που αντιστοιχούν στο παραπάνω σύνολο, να είναι συμφορημένες. Ουσιαστικά μάς πληροφορεί σχετικά με τη συχνότητα με την οποία κάθε σύνολο ακμών παρουσιάζει συμφόρηση σε ένα μεγάλο χρονικό διάστημα, σε αντίθεση με την πληροφορία που δίνουν οι Boolean Inference αλγόριθμοι, που συνίσταται στην εύρεση του συνόλου ακμών που παρουσιάζει συμφόρηση σε κάθε συγκεκριμένη χρονική στιγμή.

Προκειμένου να υπολογιστεί η πιθανότητα συμφόρησης κάθε συνόλου ακμών, αρκεί να υπολογιστεί η πιθανότητα συμφόρησης κάθε υποσυνόλου συσχέτισης  $E$  το οποίο υπάρχει πιθανότητα να είναι συμφορημένο, ή ισοδύναμα, για κάθε τέτοιο υποσύνολο συσχέτισης, να υπολογιστεί η πιθανότητα όλες οι ακμές σε αυτό να είναι καλές. Ως υποσύνολο συσχέτισης (correlation subset) ορίζεται ένα μη κενό υποσύνολο ενός συνόλου συσχέτισης.

Για τον υπολογισμό των παραπάνω πιθανοτήτων προτείνεται ένας αλγόριθμος, ο οποίος σχηματίζει τόσες γραμμικά ανεξάρτητες εξισώσεις όσες είναι δυνατό να σχηματιστούν, εφαρμόζοντας την ακόλουθη εξίσωση (2.2) σε διαφορετικά σύνολα μονοπατιών.

$$\begin{aligned} \mathbb{P}\left(\bigcap_{p \in P} Y_p = 0\right) &= \mathbb{P}\left(\bigcap_{e \in \text{Links}(P)} X_e = 0\right) \\ &= \prod_{C \in C^*} \mathbb{P}\left(\bigcap_{e \in \text{Links}(P) \cap C} X_e = 0\right) \end{aligned} \quad (2.2)$$

Στη συνέχεια επεξηγούνται οι συμβολισμοί της ανωτέρω εξίσωσης: το  $P$  αναφέρεται σε κάποιο μονοπάτι, το  $\text{Links}(P)$  αναφέρεται στις ακμές που διασχίζονται από το  $P$  και το  $C^*$  αναφέρεται στα υποσύνολα συσχέτισης.

Η είσοδος του αλγορίθμου είναι μια διάταξη όλων των πιθανά συμφορημένων υποσυνόλων συσχέτισης, τα οποία μπορεί να εμφανίζονται στις εξισώσεις που προκύπτουν από

την εφαρμογή της (2.2) στα διάφορα σύνολα μονοπατιών. Η έξοδος είναι μια διάταξη των μονοπατιών στα οποία εφαρμόζοντας την (2.2) σχηματίζεται ένα σύστημα εξισώσεων. Ο αλγόριθμος αυτός σχηματίζει έναν πίνακα  $R$ , κάθε γραμμή του οποίου αντιστοιχεί σε ένα σύνολο μονοπατιών και κάθε στήλη σε ένα υποσύνολο συσχέτισης. Εάν το υποσύνολο συσχέτισης ανήκει στο σύνολο μονοπατιών, η τιμή του συγκεκριμένου στοιχείου του πίνακα είναι 1, διαφορετικά 0. Ο αλγόριθμος κατασκευάζει τον πίνακα  $N$ , που είναι ο μηδενοχώρος του  $R$ . Στη συνέχεια, επιλέγει το υποσύνολο συσχέτισης που είναι περισσότερο πιθανό να σχηματίζει κάποιο σύνολο μονοπατιών, που αν προστεθεί στο σύστημα εξισώσεων, θα αυξήσει το βαθμό του συστήματος (χρήση της συνάρτησης *SortbyHammingWeight*). Έπειτα, επιλέγει το σύνολο που μπορεί να σχηματιστεί μόνο από το επιλεγθέν υποσύνολο συσχέτισης και όχι από τα συμπληρωματικά του και ελέγχει εάν όντως η προσθήκη του αυξάνει το βαθμό του συστήματος και αν ναι, το προσθέτει και ενημερώνει το μηδενοχώρο  $N$  του  $R$ . Η διαδικασία αυτή επαναλαμβάνεται έως ότου ο πίνακας  $N$  δεν έχει άλλες στήλες ή δεν υπάρχουν άλλα σύνολα μονοπατιών που να μπορούν να προστεθούν στο σύστημα εξισώσεων.

Επιπλέον, έγινε εκτέλεση του αλγορίθμου στις τοπολογίες Brite και Sparse και σύγκρισή του με τους αλγορίθμους *Probability Congestion Inference* και *Bayesian-Independence*. Τα αποτελέσματα έδειξαν ότι στις Brite τοπολογίες, οι δύο πρώτοι αλγόριθμοι έχουν αντίστοιχες επιδόσεις, ανεξάρτητα από το είδος της συμφόρησης και την ανεξαρτησία ή όχι των ακμών, ενώ στις Sparse τοπολογίες, ο αλγόριθμος *Congestion Probability Computation* έχει μεγαλύτερη ακρίβεια και απαιτεί τον λιγότερο χρόνο για την επίτευξή της, σε σύγκριση με τους δύο άλλους αλγορίθμους.

### 2.3.2 Αλγόριθμοι για την εύρεση της τοπολογίας του δικτύου

#### Αλγόριθμοι βασιζόμενοι στην Εκτίμηση Μέγιστης Πιθανοφάνειας

Στο [6] παρουσιάζεται ένας νέος τρόπος για την πραγματοποίηση μετρήσεων στο δίκτυο, μέσω των οποίων μπορούμε να συνάγουμε την τοπολογία του: τα *sandwich probes*. Τα διε-

ρευνητικά αυτά πακέτα, ονομάζονται έτσι διότι αποτελούνται από δυο μικρά πακέτα, με κοινό προορισμό, χωρισμένα από ένα μεγαλύτερο πακέτο, με διαφορετικό προορισμό. Η ιδέα είναι ότι το τρίτο πακέτο θα εισέρχεται σε ουρά αναμονής πίσω από το δεύτερο, με αποτέλεσμα να μεγαλώνει το διάστημα που χωρίζει την άφιξη του πρώτου από αυτή του τρίτου πακέτου στον κοινό τους προορισμό, σε κάθε ακμή, που αποτελεί μέρος του κοινού μονοπατιού από την πηγή προς τους δύο προορισμούς. Συνεπώς, το διάστημα αυτό σχετίζεται με το εύρος ζώνης του μοιραζόμενου μονοπατιού.

Η παραπάνω μέθοδος βασίζεται σε διαφορές καθυστερήσεων στους τελικούς κόμβους και όχι στις καθυστερήσεις καθ'αυτές, με αποτέλεσμα να μην απαιτεί τον συγχρονισμό των ρολογιών της πηγής και των προορισμών, κάτι που θα ήταν ιδιαίτερα κοστοβόρο. Επίσης, η μέθοδος αυτή υπονοεί ότι η άγνωστη τοπολογία είναι ένας γράφος με δενδρική δομή.

Με χρήση της παραπάνω μεθόδου μπορεί να υπολογιστεί η στατιστική πιθανοφάνεια κάθε πιθανής τοπολογίας και στη συνέχεια, με χρήση κάποιου εκτιμητή μέγιστης πιθανοφάνειας, να επιλεγεί η τοπολογία που την μεγιστοποιεί. Η πρόκληση, τώρα, έγκειται στην εύρεση μίας μεθόδου για αποδοτική αναζήτηση στο χώρο των πιθανών λύσεων. Για το σκοπό αυτό προτείνεται μία *Markov Chain Monte Carlo* διαδικασία, η οποία πραγματοποιεί γρήγορη αναζήτηση στον χώρο των λύσεων, επικεντρωνόμενη στις περιοχές με τη μέγιστη πιθανοφάνεια, επιδιώκοντας, έτσι, να αναγνωρίσει την ολικά βέλτιστη λύση. Επιπλέον, προστίθεται και η έννοια της *ποινης*, η οποία αποφεύγει την αναγνώριση ως βέλτιστων, λύσεων που αντιστοιχούν σε δέντρα με περισσότερες ακμές και άρα περισσότερους βαθμούς ελευθερίας (*overfitting problem*).

Αρχικά, πραγματοποιούνται οι ακραίες μετρήσεις με τη χρήση των διερευνητικών πακέτων που παρουσιάστηκαν προηγουμένως. Στη συνέχεια, εκτιμώνται οι ακραίες μετρικές για κάθε ζεύγος προορισμών, βάσει των μετρήσεων που προέκυψαν. Τέλος, οι μετρικές αυτές χρησιμοποιούνται για την εκτίμηση της τοπολογίας του δικτύου, με χρήση κάποιου αλγορίθμου.

Προκειμένου να προσδιοριστούν οι απαραίτητες μετρικές πραγματοποιούνται  $N(N - 1)$  μετρήσεις. Η μετρική που χρησιμοποιείται είναι το σύνολο των δειγματικών μέσων των μετρήσεων, ανά ζεύγη προορισμού,  $x = \{x_{ij}\}$ . Στη συνέχεια παρουσιάζεται το υποσύνολο

των μετρήσεων, που αντιστοιχεί στη μέτρηση  $K = k$ :

$$\{\delta(k) : r_1(k) = i, r_2(k) = j\} \quad (2.3)$$

όπου το  $\delta(k)$  αντιστοιχεί στη διαφορά των χρόνων άφιξης των δύο μικρών πακέτων στον κοινό τους προορισμό, το  $r_1$  αντιστοιχεί στον προορισμό των δύο μικρών πακέτων και το  $r_2$  στον προορισμό του μεγάλου πακέτου. Οι παραπάνω μετρήσεις θεωρείται ότι είναι ανεξάρτητες και έχουν την ίδια κατανομή. Επίσης, ορίζεται και το σύνολο των αντίστοιχων διακυμάνσεων  $\sigma = \{\sigma_{ij}\}$ .

Στη συνέχεια, ορίζεται το σύνολο όλων των πιθανών τοπολογιών (δέντρων)  $\mathcal{F}$ , το οποίο ονομάζεται δάσος. Ο στόχος του προβλήματος είναι η εύρεση του δέντρου  $T^*$ , που αντιστοιχεί στην πραγματική τοπολογία, αναζητώντας μεταξύ όλων των δέντρων του δάσους  $\mathcal{F}$ , βάσει μόνο των δεδομένων  $x$  και καμίας άλλης πληροφορίας.

Ακολούθως είναι αναγκαία η υιοθέτηση κάποιου πιθανοτικού μοντέλου για το σύνολο των μετρικών  $\{x_{ij}\}$ . Συνεπώς, ορίζονται τα εξής: η θεωρητική διαφορά καθυστέρησης  $\mu_l$  που αντιστοιχεί σε κάθε ακμή  $e_l$  του δέντρου και το κοινό μονοπάτι που ακολουθείται από την πηγή προς τους προορισμούς  $i, j$ ,  $S_{i,j}$ . Τότε, η θεωρητική τιμή του  $x_{ij}$  δίνεται από τον τύπο  $\gamma_l \equiv \sum_{l \in S_{i,j}} \mu_l$ . Η μετρική ακολουθεί Γκαουσιανή κατανομή με μέση τιμή  $\gamma_l$  και διασπορά  $\sigma^2$  (σύμφωνα με το Κεντρικό Οριακό Θεώρημα):

$$x_{ij} \sim \mathcal{N}(\gamma_{ij}, \sigma_{ij}^2) \quad (2.4)$$

Στη συνέχεια, ορίζεται η πυκνότητα πιθανότητας των δεδομένων  $x$  ως  $p(x|T, \mu(T))$ , όπου  $\mu(T) = \{\mu_l\}$  και  $l$  αντιστοιχεί σε όλες τις εσωτερικές ακμές του δικτύου (δεν συμπεριλαμβάνονται οι τελικές ακμές που καταλήγουν στους προορισμούς/φύλλα του δέντρου). Θεωρώντας κάποιο δέντρο  $T$ , βάσει των δεδομένων  $x$ , μπορεί να υπολογιστεί το  $\mu(T)$ . Επιλέγοντας τα  $\{\mu_l\}$  που μεγιστοποιούν την  $p(x|T, \mu(T))$  και ικανοποιούν τη σχέση  $\{\mu_l \geq 0\}$ , μπορεί να οριστεί η πιθανοφάνεια του  $T$ :

$$L(x|T) \equiv \log p(x|T, \hat{\mu}(T)) \quad (2.5)$$



όπου  $\hat{\mu}(T)$  αντιστοιχεί στην εκτίμηση μέγιστης πιθανοφάνειας του  $\mu(T)$ .

Εισάγοντας και την *ποινή*, το κριτήριο πιθανοφάνειας μετατρέπεται στο ακόλουθο:

$$L_\lambda(x|T) = \log p(x|T, \hat{\mu}(T)) - \lambda n(T) \quad (2.6)$$

όπου  $n(T)$  αντιστοιχεί στο πλήθος των ακμών του  $T$  και  $\lambda \geq 0$  είναι μία παράμετρος που επιλέγεται από το χρήστη και στοχεύει στην εξισορρόπηση μεταξύ του πόσο καλά προσαρμόζεται η λύση στα δεδομένα του προβλήματος και του πλήθους των ακμών που αυτή περιέχει.

Το δέντρο με τη μέγιστη πιθανοφάνεια, με ύπαρξη *ποινής*, ορίζεται ως ακολούθως:

$$\hat{T}_\lambda \equiv \max_{T \in \mathcal{F}} L_\lambda(x|T) \quad (2.7)$$

Ο καθορισμός μιας ντετερμινιστικής στρατηγικής βελτιστοποίησης με χρήση της εξίσωσης (2.7) για την εύρεση του  $\hat{T}_\lambda$  είναι ιδιαίτερος δύσκολος, διότι ο χώρος των πιθανών λύσεων είναι τεράστιος. Παρόλα αυτά, οι λύσεις που έχουν μεγάλη πιθανοφάνεια είναι λίγες και χρησιμοποιώντας την ύστερη κατανομή,  $p(T, \mu|x)$ , ως οδηγό, η αναζήτηση περιορίζεται στις περιοχές που έχουν τη μεγαλύτερη πιθανοφάνεια. Η ιδέα της χρήσης της ύστερης κατανομής ως οδηγού προέκυψε από τον επαναπροσδιορισμό του Κριτηρίου Μέγιστης Πιθανοφάνειας με ύπαρξη *ποινής* ως εξής:

$$\exp L_\lambda(x|T) = p(x|T, \mu(T)) \exp(-\lambda n(T)) \propto p(T, \mu|x) \quad (2.8)$$

δεδομένου ότι υιοθετούμε ως πρότερη  $p(T) \propto \exp(-\lambda n(T))$ . Παράγοντας, λοιπόν, τυχαία δείγματα  $(T, \mu)$  από την ύστερη κατανομή, αυτά θα είναι συγκεντρωμένα στις περιοχές υψηλότερης πιθανοφάνειας.

Η διαδικασία παραγωγής των δειγμάτων αυτών αποτελεί μία στοχαστική διαδικασία, διαρθρωμένη με τρία βήματα: βήμα γέννησης, βήμα θανάτου και  $\mu$ -βήμα. Η διαδικασία αυτή πραγματοποιείται με χρήση του αλγορίθμου reversible jump Markov Chain Monte Carlo. Το βήμα γέννησης επιλέγει έναν κόμβο με περισσότερα από δύο παιδιά, κρατάει δύο από αυτά και εισάγει έναν επιπλέον κόμβο, ως το νέο πατέρα των παιδιών αυτών. Το βήμα θανάτου είναι το ακριβώς αντίστροφο (για να αποφευχθεί η πιθανότητα να κολλήσει ο αλγόριθμος

σε κάποια ανεπιθύμητη κατάσταση), δηλαδή επιλέγεται ένας κόμβος με δύο παιδιά και διαγράφεται. Το  $\mu$ -βήμα επιλέγει με τυχαίο τρόπο μία ακμή και αλλάζει την τιμή του  $\mu_l$  αυτής.

Η πιθανότητα να οδηγηθεί ο αλγόριθμος από μία κατάσταση  $s_0 = (T_0, \mu_0)$  σε μία κατάσταση  $s_1 = (T_1, \mu_1)$  είναι η ακόλουθη:

$$\min \left\{ 1, \frac{p(T_1, \mu_1 | x) q(s_0 | s_1)}{p(T_0, \mu_0 | x) q(s_1 | s_0)} \times \mathfrak{J}_{f(s_1, s_0)} \right\} \quad (2.9)$$

όπου  $q(s_j | s_i)$  είναι η πιθανότητα να προταθεί η κίνηση από την κατάσταση  $s_i$  στην  $s_j$  και  $\mathfrak{J}_{f(s_j, s_i)}$  είναι ο Ιακωβιανός πίνακας του μετασχηματισμού  $f$  που αντιστοιχίζει τις παραμέτρους της κατάστασης  $s_i$  σε αυτές της  $s_j$  (είναι πιθανό η διάσταση του χώρου των παραμέτρων να αλλάζει, για αυτό και εισάγεται ο Ιακωβιανός πίνακας).

Τέλος, δημιουργήθηκε το εργαλείο *nettom*, το οποίο χρησιμοποιήθηκε για τη διεξαγωγή πειράματος, με σκοπό τον έλεγχο των παραπάνω αλγορίθμων. Το πείραμα διεξήχθη έξι φορές, σε διαφορετικές μέρες και ώρες. Στα τέσσερα από τα έξι πειράματα ο αλγόριθμος κατάφερε να συμπεράνει την τοπολογία, αποτυγχάνοντας να αναγνωρίσει μόνο μία σύνδεση. Η αποτυχία αυτή οφείλεται είτε στην πολύ υψηλή ταχύτητα της σύνδεσης (η οποία, συνεπώς, δε δημιουργεί ουρές αναμονής), είτε στο ότι η σύνδεση αυτή πραγματοποιεί συστηματική κατανομή φόρτου, βάσει του προοιμίου (prefix) των πακέτων.

### Αλγόριθμοι ψευδο - πιθανοφάνειας

Στο [7] προτείνεται μία εναλλακτική προσέγγιση του προβλήματος εύρεσης της λύσης με τη μέγιστη πιθανοφάνεια, η οποία περιορίζει τον υπολογιστικό φόρτο μεθόδων μεγάλης ακρίβειας, όπως αυτή που προτάθηκε στο [6]. Η προσέγγιση αυτή ανήκει στην ομάδα των *pseudo-likelihood* προσεγγίσεων.

Το μοντέλο που χρησιμοποιείται στην συγκεκριμένη τεχνική είναι το ακόλουθο:

$$\mathbf{Y} = \mathbf{A}\mathbf{X} \quad (2.10)$$

όπου  $X$  διάνυσμα διάστασης  $J$  που αναπαριστά τις δυναμικές παραμέτρους του δικτύου (π.χ. καθυστερήσεις συνδέσεων, μετρήσεις στη ροή της κίνησης),  $Y$  διάνυσμα διάστασης  $I$  που αναπαριστά τις ακραίες μετρήσεις και  $A$  πίνακας δρομολόγησης διαστάσεων

$I \times J$ . Συνήθως, ισχύει ότι  $I \ll J$ , επομένως ο  $A$  δεν είναι πλήρους διάστασης και συνεπώς απαιτούνται ορισμένοι περιορισμοί, προκειμένου να εξασφαλιστεί η αναγνωριστικότητα του μοντέλου. Ένας περιορισμός είναι η θεώρηση της ανεξαρτησίας των γραμμών του  $X$ , κάτι που στην πραγματικότητα δεν ικανοποιείται πάντοτε, αλλά αποτελεί μια καλή πρώτη προσέγγιση. Μια επόμενη θεώρηση είναι ότι τα στοιχεία του  $X$  αποτελούν τυχαίες μεταβλητές και  $X \sim f_j(\theta_j)$ ,  $j = 1, \dots, J$ . Επομένως, η παράμετρος του μοντέλου είναι η  $\theta = (\theta_1, \dots, \theta_J)$ .

Η κύρια ιδέα της συγκεκριμένης τεχνικής είναι η αποσύνθεση του αρχικού μοντέλου σε μία σειρά απλούστερων υπο-προβλημάτων ( $Y^S = A^S X^S$ , όπου  $S$  το σύνολο των υποπροβλημάτων), μέσω της επιλογής ζευγών γραμμών του πίνακα  $A$  και η εύρεση των επιμέρους οριακών πιθανοφανειών, το γινόμενο των οποίων θα δώσει τη συνάρτηση ψευδο-πιθανοφάνειας (pseudo - likelihood). Η μεγιστοποίηση του λογαρίθμου της pseudo - likelihood συνάρτησης ( $L^p$ ) δίνει την εκτίμηση μέγιστης ψευδο-πιθανοφάνειας ( $MPL$ ) της παραμέτρου  $\theta$ . Η εύρεση της  $MPL$  γίνεται με χρήση κάποιου αλγορίθμου *pseudo - EM* (Expectation Maximization).

Η παραπάνω μέθοδος εφαρμόζεται τόσο στο πρόβλημα της συναγωγής της κατανομής που ακολουθούν οι καθυστερήσεις των συνδέσεων με χρήση *multicast - probing* όσο και στο πρόβλημα της συναγωγής του πίνακα Πηγής - Προορισμού (Origin-Destination - OD - matrix).

Στο πρώτο πρόβλημα παρατηρήσιμες είναι οι καθυστερήσεις επιπέδου-μονοπατιού ( $Y$ ), ενώ ζητούμενες είναι οι καθυστερήσεις επιπέδου-ακμής ( $X$ ), οι οποίες λαμβάνουν τιμές από ένα διακριτό σύνολο  $\{0, q, 2q, \dots, mq, \infty\}$ , όπου  $q$  το εύρος του κάθε συνόλου και  $m$  μία σταθερά, με πιθανότητες  $\{\theta_{j0}, \theta_{j1}, \dots, \theta_{jm}, \theta_{j\infty}\}$ . Το μέγεθος  $q$  αντιπροσωπεύει το συμβιβασμό μεταξύ της ακρίβειας της εκτίμησης και του κόστους των υπολογισμών. Σε ένα *multicast* δέντρο με  $I$  τελικούς κόμβους μπορούν να σχηματιστούν  $I(I - 1)/2$  υποδέντρα, ο συνδυασμός των οποίων, αν θεωρηθούν ανεξάρτητα, επιτρέπει την ανάκτηση των καθυστερήσεων των συνδέσεων. Ο σχηματισμός των υποδέντρων είναι ισοδύναμος με την αποσύνθεση του προβλήματος μέσω της επιλογής γραμμών του  $A$ , η οποία αναφέρθηκε προηγουμένως. Συνεπώς, το πρόβλημα μπορεί να λυθεί με χρήση της pseudo - likelihood μεθόδου. Για την αξιολόγηση της επίδοσης της pseudo - likelihood μεθόδου χρησιμοποιείται μικρό δέντρο,

ώστε να μπορεί να εφαρμοστεί τόσο η *MPLE*, όσο και η *MLE* μέθοδος. Τα αποτελέσματα δείχνουν πως οι δύο μέθοδοι έχουν παρόμοιες επιδόσεις, ενώ η *MPLE* μέθοδος είναι πιο εύρωστη (*robust*) από την *MLE*.

Στο δεύτερο πρόβλημα, παρατηρήσιμες είναι οι μετρήσεις στις κινήσεις των συνδέσεων, ενώ ζητούμενες είναι οι *OD* μεταβλητές. Οι ζητούμενες μεταβλητές θεωρείται ότι είναι ανεξάρτητες και ισόνομα κατανεμημένες (independent and identically distributed - i.i.d.) και ότι ακολουθούν κατανομή *Poisson* σε μία γενική τοπολογία δικτύου. Για τη μείωση της υπολογιστικής πολυπλοκότητας του *EM* αλγορίθμου μπορεί να χρησιμοποιηθεί το μοντέλο της Κανονικής Κατανομής, ως προσέγγιση του μοντέλου της Κατανομής *Poisson*. Στην περίπτωση αυτή οι μεταβλητές  $X$  θα πρέπει, επιπλέον, να ικανοποιούν τη σχέση Μέσου - Διακύμανσης (mean - variance relationship, επομένως θα ισχύει το εξής:  $X_j \sim N(\lambda_j, \phi\lambda_j^c)$  και  $Y \sim N(A\lambda, A\Sigma A')$ , όπου  $\lambda = (\lambda_1, \dots, \lambda_J)$ ,  $\Sigma = \phi \text{diag}(\lambda_1^c, \dots, \lambda_J^c)$ , όπου  $\phi$  θετικό βαθμωτό μέγεθος και  $c$  σταθερά δύναμη. Συνεπώς, η παράμετρος του συνολικού μοντέλου θα είναι η  $\theta = (\phi, \lambda)$ . Διαφορετικά, καλή προσέγγιση θα αποτελούσε η Κατανομή Νόμου Δύναμης, η οποία περιγράφει τη σχέση Μέσου - Διακύμανσης. Η συνολική πολυπλοκότητα του *pseudo EM* αλγορίθμου είναι  $\mathbf{O}(N_e^{3.5})$ , όπου  $N_e$  το πλήθος των ακραίων κόμβων, η οποία είναι σημαντικά μικρότερη από την πολυπλοκότητα  $\mathbf{O}(N_e^5)$  του *EM* αλγορίθμου. Επιπλέον, όπως και πριν γίνεται αξιολόγηση και σύγκριση των επιδόσεων των μεθόδων *MLE* και *MPLE*. Από τα πειράματα προκύπτει ότι σε μικρά δίκτυα οι δύο αυτές μέθοδοι έχουν αντίστοιχες επιδόσεις, ενώ σε μεγαλύτερα δίκτυα η *MPLE* επιταχύνει σημαντικά τους υπολογισμούς έναντι της *MLE*, παραμένοντας, παράλληλα, ικανοποιητικά ακριβής.

Στη συνέχεια της παρούσας εργασίας παρουσιάζεται και αναλύεται η διαδικασία που ακολουθούν οι συσσωρευτικοί αλγόριθμοι, οι οποίοι, βασιζόμενοι σε μία 'άπληστη' διαδικασία, υπολογίζουν το *ALT* δέντρο (Agglomerative Likelihood Tree). Αλγόριθμοι τέτοιου τύπου θα παρουσιαστούν αναλυτικότερα παρακάτω.

Έπειτα, παρουσιάζεται η διαδικασία εύρεσης της λύσης με τη μέγιστη πιθανοφάνεια (δηλαδή του *MLT*), η οποία εισάγει μία 'παγκόσμια' προσέγγιση του προβλήματος αυτού, αφού λαμβάνει υπόψιν της όλες τις διαθέσιμες μετρήσεις (σε αντίθεση με την συσσωρευτική μέθοδο, που χρησιμοποιεί υποσύνολα των διαθέσιμων μετρήσεων) γεγονός που οδηγεί σε

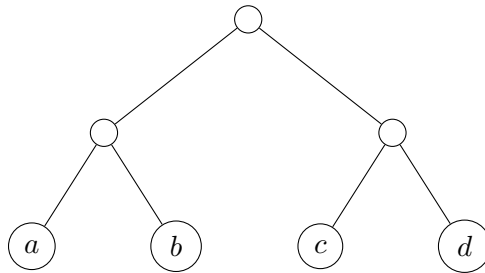
περισσότερο ακριβή αποτελέσματα, αλλά και σε μεγαλύτερο υπολογιστικό κόστος. Για τη μείωση του υπολογιστικού αυτού κόστους, προτείνεται η χρήση της *MCMC* μεθόδου, η οποία παρουσιάστηκε προηγουμένως. Επιπλέον, για τη δειγματοληψία των τιμών που παράγονται μέσω της *MCMC* μεθόδου, έτσι ώστε να ακολουθούν την επιθυμητή κατανομή, προτείνεται η χρήση του αλγορίθμου *Metropolis-Hastings*, η βελτίωση της ακρίβειας και της επίδοσης του οποίου επιτυγχάνεται με χρήση του *ALT* ως αρχή της Μαρκοβιανής αλυσίδας.

### Συσσωρευτικοί αλγόριθμοι - *agglomerative algorithms*

**Αλγόριθμοι βασιζόμενοι στη διάταξη των τελικών χρηστών βάσει της Κατά Βάθος αναζήτησης - DFS** Σκοπός της παρούσας τεχνικής, η οποία παρουσιάζεται στο [8], είναι η μείωση του αριθμού των μετρήσεων και συνεπώς του αριθμού των ανά ζεύγη διερευνητικών πακέτων που πρέπει να αποσταλούν, με σκοπό την εύρεση της τοπολογίας του δικτύου. Αυτό μπορεί να συμβεί μέσω των ομοιοτήτων,  $s_{i,j}$ , που παρουσιάζουν ορισμένα ζεύγη κόμβων (κοινά μονοπάτια, συσχέτιση καθυστερήσεων συνδέσεων κ.α.), ενώ άλλα όχι, γεγονός που οδηγεί στο συμπέρασμα ότι το να ληφθούν υπόψιν μόνο οι μετρήσεις που αφορούν έναν κόμβο και τους γείτονές του αρκεί για την ανακατασκευή της τοπολογίας του δικτύου. Πράγματι, κατ'αυτόν τον τρόπο, ο αριθμός των μετρήσεων μειώνεται στις  $O(lN \log(N))$  ανά ζεύγη μετρήσεις, για ένα ισορροπημένο  $l$ -ικό δέντρο.

Η διάταξη των τελικών χρηστών μιας δενδρικής τοπολογίας θεωρείται DFS διάταξη εάν υπάρχει Κατά Βάθος αναζήτηση στη δενδρική αυτή τοπολογία που δίνει αυτή τη διάταξη. Με χρήση κάποιας DFS διάταξης κατασκευάζεται ο πίνακας μοιραζόμενου μονοπατιού  $P$ , ο οποίος αναδεικνύει το πλήθος των κοινών δρομολογητών μεταξύ δύο τελικών χρηστών. Για παράδειγμα, η εξίσωση (2.11) παρουσιάζει τον πίνακα  $P$  που αντιστοιχεί σε μία *DFS* διάταξη  $\{a, b, c, d\}$  της τοπολογίας της εικόνας (2.3) που φαίνεται ακολούθως:

$$P^{DFS} = \begin{bmatrix} & a & b & c & d \\ a & 2 & 2 & 1 & 1 \\ b & 2 & 2 & 1 & 1 \\ c & 1 & 1 & 2 & 2 \\ d & 1 & 1 & 2 & 2 \end{bmatrix} \quad (2.11)$$

Εικόνα 2.2: Ο πίνακας μοιραζόμενου μονοπατιού μιας *DFS* διάταξης

Εικόνα 2.3: Μία απλή λογική τοπολογία δικτύου

Εάν ο πίνακας  $P$  κατασκευάζεται για κάποια *DFS* διάταξη τότε τα στοιχεία του,  $p_{i,j}$ , έχουν την ακόλουθη ιδιότητα:  $p_{i,i+1} \geq p_{i,i+k}$ : για όλα τα  $k = \{1, 2, \dots, N - i\}$ ,  $i = \{1, 2, \dots, N - 1\}$ , όπου  $N$  το πλήθος των τελικών χρηστών.

*Συνθήκη 1:* Λέγεται ότι οι πίνακες  $P$  και  $S$  (πίνακας ομοιότητας) ικανοποιούν τη συνθήκη της *Μονοτονίας* όταν ισχύουν τα ακόλουθα: {εάν  $p_{i,j} \geq p_{j,k}$ , τότε  $s_{i,j} \geq s_{j,k}$ , για όλους τους τελικούς χρήστες  $i, j, k$ }.

Στη συνέχεια προτείνεται ο αλγόριθμος *OrderedTopo*, ο οποίος, προκειμένου να ανακατασκευάσει την τοπολογία του δικτύου, χρειάζεται ως είσοδο μία *DFS* διάταξη των τελικών χρηστών και  $N - 1$  διερευνητικά πακέτα (τις ανά ζεύγη ομοιότητες  $s_{i,i+1}$ , για  $i = \{1, 2, \dots, N - 1\}$ ). Ο αλγόριθμος αυτός εντοπίζει τους κόμβους με τη μεγαλύτερη ομοιότητα, δηλαδή τους κόμβους που έχουν το μεγαλύτερο μοιραζόμενο μονοπάτι σε σχέση με όλους τους άλλους, οι οποίοι, κατ'επέκταση, αποτελούν αδερφούς κόμβους (*siblings*) και δημιουργεί έναν νέο κόμβο, ως πατέρα των δύο προηγούμενων, κάνοντας τις απαραίτη-

τες αλλαγές για τη σωστή προσθήκη του νέου κόμβου στο δέντρο (ανανέωση ακμών και κόμβων του αρχικού δέντρου). Το δέντρο που επιστρέφει ο αλγόριθμος έπειτα από  $N - 1$  επαναλήψεις αποτελεί την εκτιμώμενη τοπολογία του δικτύου.

*Συνθήκη 2:* Οι πίνακες  $P$  και  $S$  ικανοποιούν τη συνθήκη  $\delta$ -περιθωρίου, εάν για όλους τους χρήστες  $i, j, k$ , για τις παρατηρούμενες ομοιότητες, ισχύει ότι  $s_{i,j} \geq s_{j,k} + \delta, \delta > 0$ , αν και μόνο αν ισχύει ότι  $p_{i,j} \geq p_{j,k}$ . Το περιθώριο  $\delta$  ορίζεται ως η ελάχιστη καθυστέρηση λόγω ουράς αναμονής που προκαλείται από κάποιον δρομολογητή στο εν λόγω δίκτυο.

Ενώ ο παραπάνω αλγόριθμος απαιτεί τη γνώση της  $DFS$  διάταξης, ακολούθως προτείνεται ο αλγόριθμος *MarginOrder*, ο οποίος με είσοδο το μη διατεταγμένο σύνολο των τελικών χρηστών, τον πίνακα ομοιότητας  $S$  και το περιθώριο  $\delta$ , κατασκευάζει τη  $DFS$  διάταξη των τελικών χρηστών του δικτύου. Αρχικά, υπολογίζει τις ανά ζεύγη ομοιότητες  $s_{1,j}$ , για όλα τα  $j = \{2, \dots, N\}$  και τις ταξινομεί. Με αυτόν τον τρόπο βρίσκει τη μερική διάταξη  $\pi$  των τελικών χρηστών ως προς τον κόμβο  $x_1$  και, έπειτα, δημιουργεί το σύνολο  $I = \{i : s_{1,\pi_i} - s_{1,\pi_{i+1}} \geq \delta\}$ . Τέλος, εκ των στοιχείων του  $I$  επιλέγει ως στοιχείο διχοτόμησης  $i^*$  εκείνο που ικανοποιεί τη σχέση  $i^* = \operatorname{argmin}_{i \in I} |\pi_i - \frac{N}{2}|$ . Η ίδια διαδικασία εφαρμόζεται στα υποσύνολα που προκύπτουν, εφόσον αυτά διαθέτουν περισσότερα από 2 στοιχεία. Ουσιαστικά ο παραπάνω αλγόριθμος συνάγει σταδιακά τη διάταξη, υιοθετώντας μία λογική διαίρει και βασιλευε, εντοπίζοντας τους κόμβους που σίγουρα δεν είναι αδέρφια και πάντα υπό το πρίσμα της *Συνθήκης 1*.

**Πλαίσιο για το σχεδιασμό αλγορίθμων βάσει προσθετικών μετρικών - additive metrics** Στο [9] παρουσιάζονται μέθοδοι για την κατασκευή additive metrics και την εκτίμηση του μήκους των μοιραζόμενων μονοπατιών, καθώς και αποδοτικοί αλγόριθμοι (πολυωνυμικής πολυπλοκότητας) για την εύρεση της τοπολογίας του δικτύου. Επιπλέον, παρουσιάζονται και τρόποι διαχείρισης δυναμικών δικτύων (δικτύων όπου δυναμικά εισέρχονται ή/και εξέρχονται κόμβοι), καθώς η έως τώρα μελέτη ήταν εστιασμένη σε στατικά δίκτυα (stationary networks).

Για τον αλγόριθμο που προτείνεται στη συγκεκριμένη εργασία, το δίκτυο μοντελοποιείται ως ένας κατευθυνόμενος γράφος με δενδρική δομή,  $T(s, D)$ , όπου  $D$  το σύνολο των τελικών κόμβων που είναι προσβάσιμοι από την πηγή  $s$  και  $s \cup D \subset V$ , όπου  $V$  το σύνολο των κόμβων του δικτύου. Επίσης, εισάγεται ο ακόλουθος συμβολισμός: κάθε κόμβος  $k \in V$  έχει γονιό  $f(k) \in V$  και ένα σύνολο παιδιών  $c(k) = \{j \in V : f(j) = k\}$ .

Επιπρόσθετα, ορίζονται τα δύο είδη διερευνητικών πακέτων που αποστέλλονται προς απόκτηση των ακραίων μετρήσεων: τα πακέτα πολυ-εκπομπής (multicast probing packets) και μονο-εκπομπής (unicast probing packets). Κατά το multicast-probing, όταν ένας εσωτερικός κόμβος δεχτεί ένα πακέτο από τον γονιό του θα στείλει αντίγραφα του πακέτου αυτού σε όλα του τα παιδιά. Επομένως, τα πακέτα που θα φτάσουν στους τελικούς κόμβους - προορισμούς θα έχουν την ίδια δικτυακή εμπειρία (απώλειες, καθυστερήσεις κλπ.) στις κοινές συνδέσεις. Από την άλλη, κατά το unicast-probing αποστέλλεται μία σειρά από διαδοχικά *unicast* πακέτα ( $1 \times k$  string), ένα για κάθε προορισμό. Εφόσον τα  $k$  αυτά πακέτα αποστέλλονται με πολύ μικρή χρονική διαφορά, θεωρείται ότι έχουν την ίδια δικτυακή εμπειρία.

Στη συνέχεια, ορίζεται το σύνολο  $Z_e$  των μεταβλητών που περιγράφουν τις καταστάσεις των συνδέσεων, για όλες τις συνδέσεις/ακμές του δικτύου. Το σύνολο αυτό παίρνει τιμές στο σύνολο  $Z$  και η κατανομή του παραμετροποιείται από τη συνάρτηση  $\theta_e$ , η οποία μπορεί να αντιπροσωπεύει τις καθυστερήσεις, τη χρησιμοποίηση των συνδέσεων κ.α. Η παραμετροποίηση γίνεται ως εξής:  $\mathbb{P}(Z_e = z) = \theta_e(z), \forall z \in Z$ . Επίσης, ορίζεται το σύνολο  $X$  στο οποίο παίρνουν τιμές οι μεταβλητές  $X_k$ , οι οποίες περιγράφουν το αποτέλεσμα της μέτρησης στον κόμβο  $k$ . Συνεπώς, ισχύει το εξής:  $X_k = g(X_{f(k)}, Z_{e_k})$ , όπου  $e \in E$  κάποια ακμή της δενδρικής δομής (του δικτύου).

Επεκτείνοντας όσα έχουν ειπωθεί έως τώρα, θεωρώντας πως ισχύουν οι συνθήκες της *Ανεξαρτησίας* των ακμών και της *Στασιμότητας* κατά τη διάρκεια της περιόδου των μετρήσεων, οι μεταβλητές  $X_k$  σχηματίζουν ένα *Markov Random Field* (MRF) (βλ. Παράρτημα "). Πιο συγκεκριμένα,  $\forall k \in V$ , η υπό συνθήκη κατανομή της  $X_k$  δεδομένων άλλων μεταβλητών ( $X_j : j \neq k$ ) στο  $T(s, D)$  είναι η ίδια με την υπό συνθήκη κατανομή της  $X_k$  δεδομένων, μόνο, των μεταβλητών των γειτόνων του  $k$ , ( $X_j : j \in f(k) \cup c(k)$ ).



Μία μετρική  $d$  ορίζεται ως additive στο  $T(s, D)$ , εάν ισχύουν τα ακόλουθα:

1.  $0 < d(e) < \infty, \forall e \in E$
2.  $d(i, j) = \sum_{e \in P(i, j)} d(e), \forall i, j \in V$ .

Στη συνέχεια, ορίζονται οι μεταβλητές κατάστασης των συνδέσεων, ώστε να εκφράζουν τις απώλειες, τη χρησιμοποίηση και την καθυστέρηση αυτών, καθώς και οι αντίστοιχες μετρικές.

1. *Απώλειες των συνδέσεων:* Θεωρώντας τη μεταβλητή κατάσταση  $Z_e$  ως μία τυχαία *Bernoulli* μεταβλητή, ισχύει ότι:

$$\begin{aligned} \mathbb{P}(Z_e = 1) &= \alpha_e, \text{ το πακέτο διασχίζει την σύνδεση } e \\ \mathbb{P}(Z_e = 0) &= 1 - \alpha_e = \bar{\alpha}_e, \text{ το πακέτο χάνεται στη σύνδεση} \end{aligned} \quad (2.12)$$

Το αποτέλεσμα στην προκειμένη περίπτωση είναι:  $L_k = L_{f(k)} \cdot Z_{e_k} = \prod_{e \in P(s, k)} Z_e$  και η μετρική μπορεί να οριστεί ως  $d_l(e) = -\log(\alpha_e), \forall e \in E$ , εάν  $0 < \alpha_e < 1$ .

2. *Χρησιμοποίηση των συνδέσεων:* Θεωρώντας τη μεταβλητή κατάσταση  $Z_e$  ως μία τυχαία *Bernoulli* μεταβλητή, ισχύει ότι:

$$\begin{aligned} \mathbb{P}(Z_e = 1) &= \gamma_e, \text{ το πακέτο δεν υφίσταται καμία καθυστέρηση στη σύνδεση } e \\ \mathbb{P}(Z_e = 0) &= 1 - \gamma_e = \bar{\gamma}_e, \text{ το πακέτο υφίσταται καθυστέρηση στη σύνδεση} \end{aligned} \quad (2.13)$$

Το αποτέλεσμα στην προκειμένη περίπτωση είναι:  $U_k = U_{f(k)} \cdot Z_{e_k} = \prod_{e \in P(s, k)} Z_e$  και η μετρική ορίζεται ως  $d_u(e) = -\log(\gamma_e), \forall e \in E$ , εάν  $0 < \gamma_e < 1$ .

3. *Καθυστέρηση των συνδέσεων:*

$$\begin{aligned} \theta_e &= \text{var}[Z_e], \text{ ή} \\ \theta_e(i) &= \mathbb{P}(Z_e = i), i \in Z \end{aligned} \quad (2.14)$$

Το αποτέλεσμα στην προκειμένη περίπτωση είναι:  $T_k = T_{f(k)} + Z_{e_k} = \sum_{e \in P(s, k)} Z_e$  και η μετρική ορίζεται ως  $d_v(e) = \text{var}[Z_e]$ .

Έπειτα, ορίζονται τα ακόλουθα:

1. Οι αποστάσεις μεταξύ των τελικών κόμβων:  $d(U^2) = \{d(i, j) : i, j \in U\}$ , όπου  $U = s \cup D$ .
2. Το μήκος του μονοπατιού από την πηγή  $s$  στον τελικό κόμβο  $i$ :  $\rho(i) = d(s, i)$ , όπου  $d$  additive metric. Γενικεύοντας  $\forall i \in D$ , το μήκος ορίζεται ως  $\rho(s, D) = \{\rho(i)\}$ .
3. Το μήκος του κοινού μονοπατιού από την πηγή προς τους τελικούς κόμβους  $i, j$ :  $\rho(i, j) = d(s, ij)$ , όπου  $ij$  ο τελευταίος κοινός πρόγονος των  $i, j$  από την πηγή προς τους  $i, j$ . Γενικεύοντας  $\forall i, j \in D$ , το μήκος ορίζεται ως  $\rho(s, D^2) = \{\rho(i, j)\}$ .

Δεδομένων όσων ορίστηκαν έως τώρα, κατασκευάζονται οι εκτιμητές των μηκών  $\rho(i)$  και  $\rho(i, j)$ , με χρήση των τριών παραπάνω μετρικών. Η εκτίμηση του μήκους,  $\rho(i, j)$ , μπορεί να χρησιμοποιηθεί για την ανάκτηση των παραμέτρων που περιγράφουν την επίδοση της ακμής  $ij$ . Οι εκτιμητές παρουσιάζονται εν συνεχεία:

1.  $\hat{\rho}_l(i) = -\log \bar{L}_i$ ,  $\hat{\rho}_l(i, j) = -\log \bar{L}_i \bar{L}_j / \bar{L}_{ij}$ , όπου

$$\bar{L}_i = \frac{1}{n} \sum_{t=1}^n L_i^{(t)} \quad (2.15)$$

$$\bar{L}_{ij} = \frac{1}{n} \sum_{t=1}^n L_i^{(t)} L_j^{(t)} \quad (2.16)$$

Οι απώλειες  $L_i$  της σύνδεσης  $i$  προκύπτουν ως εξής:

$$L_i^{(t)} = \begin{cases} 1, & T_i^{(t)} < \infty \\ 0, & T_i^{(t)} = \infty \end{cases} \quad (2.17)$$

όπου  $T_i^{(t)}$  είναι η μετρούμενη καθυστέρηση του  $t$ -οστού probe από την πηγή  $s$  στον κόμβο  $i$ .

2.  $\hat{\rho}_u(i) = -\log \bar{U}_i$ ,  $\hat{\rho}_u(i, j) = -\log \bar{U}_i \bar{U}_j / \bar{U}_{ij}$ , όπου

$$\bar{U}_i = \frac{1}{n} \sum_{t=1}^n U_i^{(t)} \quad (2.18)$$

$$\bar{U}_{ij} = \frac{1}{n} \sum_{t=1}^n U_i^{(t)} U_j^{(t)} \quad (2.19)$$

Η χρησιμοποίηση  $U_i$  της σύνδεσης  $i$  προκύπτει ως εξής:

$$U_i^{(t)} = \begin{cases} 1, & T_i^{(t)} - T_i^{min} \leq \epsilon \\ 0, & T_i^{(t)} - T_i^{min} > \epsilon \end{cases} \quad (2.20)$$

όπου  $T_i^{min} = \min_t T_i^{(t)}$  και  $\epsilon$  μία μικρή τιμή για την κάλυψη πιθανών σφαλμάτων μέτρησης.

3.  $\hat{\rho}_v(i) = \hat{v}ar(T_i)$ ,  $\hat{\rho}_v(i, j) = \hat{c}ov(T_i, T_j)$ , όπου

$$\hat{v}ar(T_i) = \frac{1}{n-1} \sum_{t=1}^n (T_i^{(t)} - \bar{T}_i)^2 \quad (2.21)$$

$$\hat{c}ov(T_i, T_j) = \frac{1}{n-1} \sum_{t=1}^n (T_i^{(t)} - \bar{T}_i)(T_j^{(t)} - \bar{T}_j) \quad (2.22)$$

Η καθυστέρηση  $T_i$  της σύνδεσης  $i$  προκύπτει ως εξής:

$$\bar{T}_i = \frac{1}{n} \sum_{t=1}^n T_i^{(t)}. \quad (2.23)$$

Στη συνέχεια προτείνονται αντίστοιχες μετρικές βασιζόμενες σε unicast-probing. Συγκεκριμένα, για ένα ζεύγος διαδοχικών πακέτων  $\{a, b\}$  ορίζουμε τον οριακό ρυθμό επιτυχίας της ακμής  $e$ ,  $\alpha_e = \mathbb{P}(Z_e^x = 1)$ ,  $x = \{a, b\}$ , καθώς και τον υπό συνθήκη ρυθμό επιτυχίας της  $e$ ,  $\beta_e = \mathbb{P}(Z_e^b = 1 | Z_e^a = 1)$ . Τότε θα ισχύει ότι  $0 < \alpha_e < \beta_e \leq 1$  και συνεπώς  $0 < \frac{\alpha_e}{\beta_e} < 1$  και μπορεί να κατασκευαστεί η μετρική  $d_l^i(e) = -\log \frac{\alpha_e}{\beta_e}$ .

Έπειτα παρουσιάζεται ο αλγόριθμος *RNJ*, ο οποίος είναι παρόμοιος με τον αλγόριθμο *OrderedTopo* που παρουσιάστηκε προηγουμένως, με τη μόνη διαφορά ότι, εφόσον δε διαθέτει

(ούτε υπολογίζει) κάποια διάταξη των τελικών κόμβων, προκειμένου να βρει τους κόμβους με το μεγαλύτερο μοιραζόμενο μονοπάτι υπολογίζει και συγκρίνει τα μήκη των μοιραζόμενων μονοπατιών όλων των δυνατών ζευγών τελικών κόμβων. Αυτό οδηγεί σε πολυπλοκότητα  $\mathcal{O}(N^2 \log N)$ , για δίκτυο με  $N$  κόμβους-προορισμούς.

Ο παραπάνω αλγόριθμος λειτουργεί υπό την συνθήκη της Στασιμότητας του δικτύου, η οποία, ιδίως στην περίπτωση του unicast-probing, είναι πιθανό να μην ικανοποιείται. Για το λόγο αυτό προτείνονται αλγόριθμοι που χειρίζονται την προσθήκη και τη διαγραφή κόμβων.

Η διαδικασία  $AddNode(T, j, k, \Delta)$  για την προσθήκη κόμβων δέχεται ως είσοδο το δέντρο  $T$ , τον κόμβο προς εισαγωγή  $j$ , τον κόμβο  $k$  βάσει του οποίου θα γίνει η προσθήκη του  $j$  και την παράμετρο  $\Delta \leq \min_{e \in Ed}(e)$ . Αρχικά, ελέγχεται εάν ο κόμβος  $k$  είναι φύλλο του  $T$  και αν ναι, τότε ο νέος κόμβος θα είναι αδερφός αυτού. Διαφορετικά, εντοπίζεται ο κόμβος - παιδί  $d_{i^*}$  του  $k$  που μεγιστοποιεί το  $\hat{\rho}(j, d_i)$  και ελέγχεται η διαφορά  $d$  των εκτιμήσεων του μήκους του κοινού μονοπατιού δύο κόμβων - παιδιών  $(d_1, d_2)$  του  $k$  και του  $d_{i^*}$  με τον  $j$  ( $d = \hat{\rho}(d_1, d_2) - \hat{\rho}(j, d_{i^*})$ ). Εάν  $d \geq \frac{\Delta}{2}$ , τότε ο  $j$  τοποθετείται στο ίδιο επίπεδο με τον  $k$ . Εάν  $|d| < \frac{\Delta}{2}$ , τότε ο  $j$  θα είναι παιδί του  $k$  στο νέο δέντρο και τέλος, αν  $\hat{\rho}(j, d_{i^*}) - \hat{\rho}(d_1, d_2) \geq \frac{\Delta}{2}$ , τότε ο  $j$  θα είναι αδερφός ή παιδί κάποιου απογόνου  $c_{i^*}$  του  $k$  στο νέο δέντρο, οπότε και καλείται αναδρομικά η συνάρτηση  $AddNode(T, j, c_{i^*}, \Delta)$ .

Η διαδικασία διαγραφής κόμβου  $j$  είναι σχετικά απλή. Αρχικά, διαγράφεται ο  $j$  και η ακμή  $(f(j), j)$ . Έπειτα, εάν ο  $f(j)$ , μετά τη διαγραφή του  $j$  απομένει με μόνο ένα παιδί, τότε ο  $f(j)$  διαγράφεται και το παιδί αυτού συνδέεται με τον  $f(f(j))$ , ώστε να διατηρείται η ιδιότητα σύμφωνα με την οποία κάθε κόμβος έχει τουλάχιστον δύο παιδιά.

**Χρήση πινάκων μετάβασης για την ανάκτηση των παραμέτρων που περιγράφουν την επίδοση των ακμών** Όπως και πριν, το δίκτυο μοντελοποιείται ως ένας κατευθυνόμενος γράφος με δενδρική δομή,  $T(V, E)$ , οι καταστάσεις,  $S_k$ , των ακμών παρουσιάζουν χωρική ανεξαρτησία και το σύνολο των τυχαίων μεταβλητών,  $X_V = (X_k, k \in V)$ , που υποδηλώνουν την έχβαση του πακέτου στον κόμβο  $k$ , σχηματίζουν ένα *Markov Random Field* (MRF).

Στη γενική περίπτωση, όπου οι παράμετροι που περιγράφουν την επίδοση των ακμών αποτελούν διανυσματικά μεγέθη, μπορεί να μην είναι δυνατόν αυτές να προσδιοριστούν μοναδικά, βάσει των μηκών των ακμών, υπό οποιαδήποτε προσθετική μετρική, όπως αυτές παρουσιάστηκαν στην υποενότητα 2.3.2. Στην περίπτωση αυτή, μπορεί να γίνει ανάκτηση των παραμέτρων αυτών από τους πίνακες μετάβασης των ακμών (διότι οι παράμετροι των ακμών αποτελούν τις εισόδους των πινάκων μετάβασης), χρησιμοποιώντας μάλιστα, μόνο την κατεύθυνση από τον γονιό προς το παιδί ( $f(k) \rightarrow k$ ) (forward link transition matrix).

Για κάθε ακμή ( $(f(k), k)$ ) μίας λογικής, multicast, δενδρικής τοπολογίας, ορίζουμε ως forward link transition matrix,  $P_{f(k)k}$ , τον  $\mathcal{M} \times \mathcal{M}$  πίνακα (όπου  $|\mathcal{X}| = \mathcal{M}$ ) με εισόδους:

$$P_{ij}(x, y) = \mathbb{P}(X_j = y | X_i = x), x, y \in \mathcal{X}. \quad (2.24)$$

Εάν οι πίνακες μετάβασης των ακμών είναι αντιστρέψιμοι, ώστε  $\det(P_{ij}) > 0$ , όχι ίσοι με κάποιον πίνακα μετάθεσης<sup>1</sup>, ώστε  $\det(P_{ij}) < 1$ , και υπάρχει κόμβος  $k \in V$  με θετική περιθωριακή κατανομή (marginal distribution), τότε επιβεβαιώνεται ότι η ακόλουθη μετρική ( $d(l_k)$ ) είναι προσθετική:

$$d(l_k) = -\log|P_{f(k)k}| - \log|P_{kf(k)}|, l_k = (f(k), k) \in E, |P_{ij}| = \det(P_{ij}). \quad (2.25)$$

Εάν ισχύουν όσα αναφέρθηκαν παραπάνω για τους forward link transition matrices και επιπλέον ισχύει ότι οι τελευταίοι είναι ανακατασκευάσιμοι από τις γραμμές τους<sup>2</sup>, τότε στο [10] παρουσιάζεται αλγόριθμος για την ανάκτηση των πινάκων αυτών, μέσω των οποίων, ακολούθως, ανακτώνται οι παράμετροι της επίδοσης των ακμών της δενδρικής τοπολογίας.

### 2.3.3 Αλγόριθμοι για την παρακολούθηση δικτύων μέσω της επιλογής ενός συνόλου μονοπατιών - monitoring paths

<sup>1</sup>Ως πίνακας μετάθεσης (permutation matrix) ορίζεται ένας πίνακας που έχει ένα το πολύ στοιχείο σε κάθε γραμμή και στήλη ίσο με 1.

<sup>2</sup>Ένας πίνακας είναι ανακατασκευάσιμος από τις γραμμές του όταν μπορεί να προσδιοριστεί μοναδικά από το μη-διατεταγμένο σύνολο των γραμμών του.

Στην εργασία [11] παρουσιάζεται ένας τρόπος παρακολούθησης δικτύων, με επιλογή ενός συνόλου-βάσης  $k$  γραμμικά ανεξάρτητων μονοπατιών. Το σύνολο αυτό χρησιμεύει στην περιγραφή του συνόλου των  $\mathbf{O}(n^2)$  μονοπατιών (σε δίκτυο με  $n$  τελικούς χρήστες) διότι κάθε end-to-end μονοπάτι μπορεί να γραφεί ως γραμμικός συνδυασμός των μονοπατιών του συνόλου-βάσης. Έτσι, παρακολουθώντας τους ρυθμούς απωλειών και τις καθυστερήσεις μόνο των  $k$  αυτών μονοπατιών, μπορούν να συναχθούν οι αντίστοιχες μετρικές για όλα τα υπόλοιπα μονοπάτια. Κατ'αυτόν τον τρόπο μειώνεται ο αριθμός των απαιτούμενων μετρήσεων, καθιστώντας την παρακολούθηση του δικτύου μία διαδικασία περισσότερο αποδοτική και με δυνατότητες κλιμάκωσης. Σκοπός της διαδικασίας αυτής είναι ο εντοπισμός μη λειτουργικών μονοπατιών και διαταραχών στην απόδοση των δικτύων και ανάκαμψη από αυτές.

Στο σημείο αυτό πρέπει να διευκρινιστεί ότι το ενδιαφέρον της παρούσας εργασίας δεν είναι στραμμένο τόσο στη συναγωγή των χαρακτηριστικών μεμονωμένων ακμών, αλλά ολόκληρων μονοπατιών.

Σε πρώτη φάση ορίζεται ένα αλγεβρικό μοντέλο, εφαρμόσιμο σε κάθε τοπολογία δικτύου. Ένα μονοπάτι ορίζεται ως ένα διάνυσμα - στήλη  $v \in \{0, 1\}^s$ , όπου  $s$  αντιστοιχεί στο πλήθος των  $\Pi$  συνδέσεων και του οποίου το  $j$ -οστό στοιχείο  $v_j$  ισούται με 1 εάν η ακμή  $j$  ανήκει στο μονοπάτι, ενώ διαφορετικά ισούται με 0. Επίσης, ορίζεται η πιθανότητα  $l_j$  το πακέτο να χαθεί στην ακμή  $j$ , καθώς και ο ρυθμός απωλειών  $p$  ενός μονοπατιού ως ακολούθως:

$$1 - p = \prod_{j=1}^s (1 - l_j)^{v_j} \quad (2.26)$$

Λογαριθμίζοντας την εξίσωση (2.26) και ορίζοντας  $b_i = \log(1 - p_i)$ ,  $x_i = \log(1 - l_i)$  και τον πίνακα  $G \in \{0, 1\}^{r \times s}$ , κάθε γραμμή του οποίου αντιστοιχεί σε ένα μονοπάτι και  $G_{ij} = 1$  αν το μονοπάτι  $i$  περιέχει την ακμή  $j$ , διαφορετικά  $G_{ij} = 0$ , προκύπτει το ακόλουθο μοντέλο:

$$Gx = b \quad (2.27)$$

Στη συνέχεια, πραγματοποιώντας row reduction βρίσκουμε το row space και επομένως το βαθμό του πίνακα  $G$ . Εάν δύο ή περισσότερες ακμές εμφανίζονται μαζί σε όλα τα μονοπάτια στα οποία συμμετέχουν, ουσιαστικά είναι εξαρτημένες, με αποτέλεσμα ο  $G$  να μην είναι πλήρους βαθμού και το σύστημα να μην έχει μοναδική λύση για αυτές. Οι ακμές αυτές

χαρακτηρίζονται ως μη αναγνωρίσιμες. Στη συνέχεια, μπορούν να διαχωριστούν τα αναγνωρίσιμα από τα μη αναγνωρίσιμα στοιχεία ως εξής:  $x = x_G + x_N$ , όπου  $x_G \in \mathcal{R}(G^T)$  η προβολή του  $G$  στον χώρο των γραμμών (διάστασης  $k$ ) και  $x_N$  η προβολή του  $G$  στον μηδενικό χώρο. Ότι πληροφορία μπορεί να δώσει το σύστημα (2.27) περιέχεται στο  $x_G$ , το οποίο προκειμένου να προσδιοριστεί πλήρως απαιτεί μόνο  $k$  ανεξάρτητες εξισώσεις. Επομένως, λαμβάνοντας μετρήσεις για  $k$  το πλήθος μονοπάτια, μέσω του νέου συστήματος  $b = Gx = Gx_G + Gx_N = Gx_G$  καθίσταται δυνατός ο υπολογισμός όλων των στοιχείων του  $b$  και κατ'επέκταση και των υπολοίπων στοιχείων του  $x$ .

Η επιλογή των  $k$  μονοπατιών γίνεται με χρήση  $QR$  παραγοντοποίησης, για την εύρεση του βαθμού του πίνακα  $G$ .

Το Διαδίκτυο φαίνεται να ακολουθεί κατανομή Νόμου-Δύναμης (scale-free network) [12], [13], δηλαδή  $P(k) \sim k^{-\gamma}$ , όπου  $P(k)$  το κλάσμα των κόμβων που διαθέτουν  $k$  συνδέσεις με άλλους κόμβους. Αυτό σημαίνει ότι είναι λίγοι οι κόμβοι που έχουν τις περισσότερες συνδέσεις, με αποτέλεσμα οι γράφοι που αντιστοιχούν σε πραγματικά δίκτυα να μην είναι ιδιαιτέρως 'πυκνοί' (όπως ένα clique, στο οποίο κάθε κόμβος συνδέεται με όλους τους άλλους και επομένως όλα τα  $\mathbf{O}(n^2)$  μονοπάτια, για γράφο  $n$  κόμβων, είναι μεταξύ τους ανεξάρτητα). Επιπλέον, το διαδίκτυο διαθέτει ιεραρχία επιπέδου Αυτόνομων Συστημάτων (AS). Βάσει των παραπάνω, μπορεί να θεωρηθεί ότι  $k \sim \mathbf{O}(n \log n)$ , γεγονός που αποτελεί ένδειξη της δυνατότητας κλιμάκωσης της προτεινόμενης διαδικασίας. Η θεώρηση αυτή ελέγχεται με χρήση γραμμικής παλινδρόμησης (linear regression) τόσο σε πραγματικές τοπολογίες, όσο και σε συνθετικές τοπολογίες, που κατασκευάζονται βάσει ιεραρχικού μοντέλου ή μοντέλου *Barabasi-Albert* ή *Waxman*. Ομοίως, κατανομή Νόμου-Δύναμης ακολουθεί και η χρησιμοποίηση των ακμών (λίγες ακμές ανήκουν στα συντομότερα μονοπάτια που ενώνουν την πλειοψηφία των ζευγών - κόμβων), συνεπώς αντίστοιχη θεώρηση για το μέγεθος του  $k$  μπορεί να γίνει και εντός ενός AS.

Στη συνέχεια παρουσιάζεται η διαδικασία που ακολουθείται για τη διαγραφή μονοπατιών. Εάν το μονοπάτι προς διαγραφή δεν ανήκει στον  $\bar{G} \in \mathbb{R}^{k \times s}$  (πίνακας πλήρους βαθμού που προέκυψε από την  $QR$  παραγοντοποίηση του  $G$ ), η διαγραφή του είναι πολύ απλή. Διαφορετικά, απαιτείται έλεγχος για το αν ο πίνακας  $G$  περιέχει μονοπάτι γραμμικά εξαρτημένο

από το μονοπάτι προς διαγραφή (δηλαδή με μη μηδενικά στοιχεία στην κατεύθυνση του προς διαγραφή μονοπατιού). Στην περίπτωση αυτή θα πρέπει να γίνει προσθήκη του μονοπατιού αυτού στον  $\bar{G}$  (έτσι προκύπτει ο  $\bar{G}'$ ) και ανανέωση του πίνακα  $R$ . Ο αλγόριθμος αυτός έχει πολυπλοκότητα  $\mathbf{O}(k^2)$ .

Έπειτα, μελετάται το πρόβλημα της ομοιόμορφης κατανομής του φόρτου. Προκειμένου ο φόρτος των μετρήσεων να κατανέμεται ομοιόμορφα μεταξύ των τελικών κόμβων, πριν την εφαρμογή του Αλγορίθμου για την επιλογή των  $k$  μονοπατιών προς παρακολούθηση, πραγματοποιείται τυχαία αναδιάταξη των μονοπατιών - γραμμών του  $G$ .

Επιπλέον, εφόσον η εύρεση της τοπολογίας γίνεται με χρήση *traceroute*, είναι πιθανό να υπάρξουν ελλιπείς πληροφορίες δρομολόγησης (π.χ. λόγω κόμβων που δεν απαντάνε). Για το λόγο αυτό προστίθενται ακμές εκεί που η φυσιολογική διαδρομή παύει να είναι διαθέσιμη, ενώ για μονοπάτια που είναι εξ'ολοκλήρου μη ανιχνεύσιμα, προστίθεται μία απευθείας ακμή από την πηγή στον προορισμό. Η προκύπτουσα κατ'αυτόν τον τρόπο τοπολογία είναι ισοδύναμη αυτής που θα προέκυπτε εάν υπήρχαν ολοκληρωμένες πληροφορίες δρομολόγησης. Επίσης, τα ψευδώνυμα<sup>3</sup> (aliases) αγνοούνται, διότι δεν επηρεάζουν την ακρίβεια των εκτιμήσεων των ζητούμενων χαρακτηριστικών, απλώς εισάγουν την ανάγκη για μερικές επιπλέον μετρήσεις.

## 2.4 Παρουσίαση των σημαντικότερων μετρικών ενός δικτύου

Η ανάγκη για βελτιστοποίηση της απόδοσης μεταφοράς από άκρη σ' άκρη ενός δικτύου, της δρομολόγησης μέσω ενός επικαλυπτικού δικτύου, καθώς και της peer-to-peer κατανομής αρχείων, οδήγησε στην ανάγκη εύρεσης βέλτιστων μονοπατιών, δηλαδή μονοπατιών αποτελούμενων από ακμές μεγάλης χωρητικότητας ή για την ακρίβεια, μεγάλης διαθέσιμης χωρητικότητας, προς αποφυγήν συμφόρησης, μείωση του χρόνου μετάδοσης και βελτίωση της απόδοσης των εφαρμογών - διεργασιών που χρησιμοποιούν το δίκτυο [14].

<sup>3</sup>Τα ψευδώνυμα είναι εναλλακτικά ονόματα που χρησιμοποιούνται για τον ίδιο τομέα.



Προκειμένου να διευκολυνθεί ο χαρακτηρισμός των ακμών ενός δικτύου και κατ'επέκταση η κατάσταση κατάλληλων μονοπατιών, απαιτείται ο ορισμός μετρικών για την περιγραφή της κατάστασης των ακμών. Οι μετρικές αυτές, όπως ορίζονται στο [14], παρουσιάζονται ακολούθως.

### 2.4.1 Χωρητικότητα

Προκειμένου να οριστεί η *χωρητικότητα* (capacity) ενός hop (απόσταση επανάληψης), πρέπει πρώτα να οριστεί ο *ρυθμός μετάδοσης* μίας ακμής επιπέδου ζεύξης δεδομένων<sup>4</sup>. Ο ρυθμός μετάδοσης αντιστοιχεί στον ρυθμό (bit rate) με τον οποίο μία ακμή επιπέδου 2 μπορεί να μεταφέρει δεδομένα. Ο ρυθμός αυτός εξαρτάται τόσο από το υποκείμενο μέσο μετάδοσης, όσο και από το hardware (υλισμικό) του ηλεκτρονικού ή οπτικού αποστολέα/παραλήπτη. Στη συνέχεια, πρέπει να υπογραμμιστεί ότι στο επόμενο επίπεδο<sup>5</sup> ο ρυθμός μετάδοσης είναι μικρότερος από τον ονομαστικό (δηλαδή τον ρυθμό μετάδοσης του επιπέδου 2 όπως ορίστηκε προηγουμένως), λόγω του overhead (επίφορτος) που οφείλεται στην ενθλάκωση του Επιπέδου 2 (layer 2 encapsulation). Τώρα καθίσταται δυνατός ο ορισμός της χωρητικότητας.

**Ορισμός 2.4.1.** Η χωρητικότητα  $C_i$  ενός hop  $i$  ορίζεται ως ο μέγιστος δυνατός ρυθμός μετάδοσης δεδομένων στο Επίπεδο 3 στο hop αυτό. Όμως, ο μέγιστος δυνατός ρυθμός μετάδοσης στο Επίπεδο 3 προκύπτει από πακέτα μεγέθους MTU<sup>6</sup>, επομένως η χωρητικότητα ενός hop ορίζεται ως ο ρυθμός μετάδοσης, μετρημένος στο Επίπεδο 3, στον οποίο το hop αυτό μπορεί να μεταφέρει IP πακέτα μεγέθους MTU.

**Ορισμός 2.4.2.** Η ακμή ενός μονοπατιού με τη μικρότερη χωρητικότητα ονομάζεται *narrow link*.

<sup>4</sup>Επίπεδου 2 στο μοντέλο OSI (Layer 2 link ή Data Link Layer).

<sup>5</sup>Επίπεδο 3 ή Επίπεδο Δικτύου του μοντέλου OSI (Layer 3 ή IP Layer).

<sup>6</sup>Maximum Transmission Unit. Η μονάδα μέγιστης μετάδοσης σε μία συναλλαγή Επιπέδου 3.

**Ορισμός 2.4.3.** Η χωρητικότητα απ'άκρη σ'άκρη ενός μονοπατιού (end-to-end capacity of a path) ισούται με τη χωρητικότητα του narrow link του μονοπατιού αυτού.

### 2.4.2 Διαθέσιμο εύρος ζώνης και μέση χρησιμοποίηση

Αυτό όμως που πραγματικά ενδιαφέρει κάποιον αλγόριθμο επιλογής βέλτιστου μονοπατιού είναι η διαθέσιμη χωρητικότητα του μονοπατιού αυτού, η οποία εξαρτάται όχι μόνο από την υποκείμενη τεχνολογία του αποστολέα/παραλήπτη και το μέσο μετάδοσης, αλλά και από τον φόρτο δεδομένων των ακμών του μονοπατιού αυτού. Έτσι προκύπτει η ανάγκη για τον ορισμό του διαθέσιμου εύρους ζώνης (available bandwidth).

**Ορισμός 2.4.4.** Το διαθέσιμο εύρος ζώνης μίας ακμής ορίζεται ως η διαθέσιμη χωρητικότητα της ακμής αυτής κατά τη διάρκεια μίας δεδομένης χρονικής περιόδου.

Ουσιαστικά συνδέεται με τη μέση χρησιμοποίηση της ακμής (average utilization) κατά τη διάρκεια της χρονικής αυτής περιόδου, με την ακόλουθη σχέση:

$$A_i = (1 - u_i)C_i, \quad (2.28)$$

όπου  $A_i$  το διαθέσιμο εύρος ζώνης,  $C_i$  η χωρητικότητα του hop  $i$  και  $u_i$  η μέση χρησιμοποίηση της ακμής, η οποία δίνεται από την επόμενη σχέση:

$$u_i = \frac{1}{\tau} \int_{t-\tau}^t u(x) dx, \quad (2.29)$$

όπου  $u(x)$  το στιγμιαίο διαθέσιμο εύρος ζώνης της ακμής σε χρόνο  $x$ , το οποίο λαμβάνει τιμές 1 ή 0 (ανάλογα αν η ακμή χρησιμοποιείται ή όχι τη δεδομένη χρονική στιγμή  $x$ ), και  $\tau$  ο χρόνος ολοκλήρωσης.

**Ορισμός 2.4.5.** Η ακμή ενός μονοπατιού με το μικρότερο διαθέσιμο εύρος ζώνης ονομάζεται *tight link*.

**Ορισμός 2.4.6.** Το διαθέσιμο εύρος ζώνης απ'άκρη σ'άκρη ενός μονοπατιού (end-to-end available bandwidth) ορίζεται ως το διαθέσιμο εύρος ζώνης του tight link του μονοπατιού αυτού.

### 2.4.3 TCP throughput και bulk transfer capacity

Σε αντίθεση με το εύρος ζώνης, το οποίο αντιστοιχεί στον μέγιστο αριθμό των πακέτων που θεωρητικά μπορούν να αποστέλλονται μέσω μίας σύνδεσης, υπάρχει και μία ακόμη μετρική, φαινομενικά παρόμοια, στην πραγματικότητα όμως πολύ διαφορετική, το throughput (διεκπεραιωτικότητα) [15].

**Ορισμός 2.4.7.** Ως throughput ορίζεται ο αριθμός των πακέτων που επιτυχημένα αποστέλλονται μέσω μίας σύνδεσης, υπολογισμένος είτε στη μονάδα του χρόνου είτε σε ένα μικρό χρονικό διάστημα [15].

Δεδομένου ότι το TCP αποτελεί το κυρίαρχο πρωτόκολλο μεταφοράς στο Διαδίκτυο, μεταφέροντας περίπου το 90% των κινήσεων δεδομένων, θα ήταν ιδιαίτερος χρήσιμος ο ορισμός μίας μετρικής, η οποία να αντιπροσωπεύει το throughput που μπορεί να επιτευχθεί από μία TCP σύνδεση. Η μετρική αυτή ονομάζεται Bulk Transfer Capacity (BTC) και ορίζεται ακολούθως.

**Ορισμός 2.4.8.** Ως BTC ορίζεται το μέγιστο throughput που μπορεί να επιτευχθεί από μία και μόνη TCP σύνδεση.

Το BTC αφορά μόνο TCP συνδέσεις. Για αυτό το λόγο, πέρα από την υποκείμενη τεχνολογία, εξαρτάται και από διάφορα χαρακτηριστικά του TCP πρωτοκόλλου, όπως τα μεγέθη των buffers του TCP socket και τον αργό μηχανισμό εκκίνησης της σύνδεσης TCP (TCP handshake), αλλά και από τον αριθμό των ανταγωνιστικών TCP συνδέσεων, τον τύπο του cross traffic, την συμφόρηση του αντίστροφου μονοπατιού, τον Round Trip Time (RTT) κ.α. Για αυτό και, για ορισμένους τύπους κινήσεων δεδομένων (όπως μία τυπική φόρτωση ιστοσελίδας), αποτελεί αποτελεσματικότερη μετρική από το διαθέσιμο εύρος ζώνης, με το οποίο δεν πρέπει να συγχέεται.

## 2.5 Παρουσίαση των διαφόρων τεχνικών/εργαλείων για την εκτίμηση των μετρικών ενός δικτύου

Στη συνέχεια θα παρουσιαστούν ορισμένες τεχνικές και εργαλεία που επιτρέπουν την εκτίμηση των διαφόρων μετρικών (χωρητικότητα, εύρος ζώνης, bulk transfer capacity), είτε σε μία μεμονωμένη ακμή, είτε σε ένα end-to-end μονοπάτι του δικτύου.

### 2.5.1 Τεχνικές για την εκτίμηση των μετρικών ενός δικτύου

#### Variable Packet Size Probing (VPS)

Η τεχνική αυτή έχει ως στόχο τη μέτρηση της χωρητικότητας κάθε hop κατά μήκος ενός μονοπατιού [14]. Βασίζεται στον υπολογισμό του RTT από την πηγή σε κάθε hop του μονοπατιού ως συνάρτηση του μεγέθους του διερευνητικού πακέτου, εκμεταλλευόμενο το πεδίο TTL της επικεφαλίδας IP του πακέτου, το οποίο το αναγκάζει να λήξει σε κάποιο συγκεκριμένο hop. Έπειτα, ο δρομολογητής του hop αυτού αποστέλλει ICMP time-exceeded πακέτα πίσω στην πηγή, τα οποία χρησιμοποιούνται για τον υπολογισμό του RTT μέχρι το hop αυτό.

Προκειμένου να επεξηγηθεί η τεχνική αυτή απαιτείται ο ορισμός των τριών ακόλουθων ειδών καθυστέρησης:

- *Καθυστέρηση μετάδοσης (serialization delay)*: είναι ο χρόνος που απαιτείται για να μεταδοθεί ένα πακέτο μεγέθους  $L$  μέσω μίας ζεύξης με ρυθμό μετάδοσης  $C$ . Ουσιαστικά, είναι ο χρόνος που απαιτείται ώστε να εισάγει ο πομπός την πληροφορία στο μέσο μετάδοσης και ισούται με  $L/C$ .
- *Καθυστέρηση διάδοσης (propagation delay)*: είναι ο χρόνος που απαιτείται ώστε κάθε *bit* του πακέτου να διασχίσει μία ζεύξη και να φτάσει στον προορισμό και είναι ανεξάρτητη του μεγέθους του πακέτου.

- *Καθυστέρηση αναμονής* (queuing delay): είναι η καθυστέρηση που προκύπτει εάν υπάρξει συμφόρηση στις πόρτες εισόδου ή εξόδου των buffers των δρομολογητών ή των μεταγωγέων που μεσολαβούν στο εκάστοτε μονοπάτι.

Η τεχνική *VPS* βασίζεται στην αποστολή πολλαπλών διερευνητικών πακέτων δεδομένου μεγέθους, από την πηγή προς κάθε συσκευή Επιπέδου 3 κατά μήκος του μονοπατιού. Υποθέτει ότι τουλάχιστον ένα από τα πακέτα αυτά μαζί με τα ICMP replies που παράγονται κατά τη λήξη του *TTL*, δε θα υποστεί καθυστέρηση αναμονής. Έτσι, επιλέγοντας τον μικρότερο RTT αυτός θα εξαρτάται μόνο από την καθυστέρηση μετάδοσης (η οποία εξαρτάται από το μέγεθος  $L$  του πακέτου) και από την καθυστέρηση διάδοσης (ανεξάρτητη του  $L$ ). Συνεπώς, ο ελάχιστος RTT ( $T_i(L)$ ) για ένα δεδομένο πακέτο μεγέθους  $L$  μέχρι το hop  $i$  θα δίνεται από τον ακόλουθο τύπο:

$$T_i(L) = \alpha + \sum_{k=1}^i \frac{L}{C_k} = \alpha + \beta_i L \quad (2.30)$$

Βάσει των παραπάνω, ο όρος  $\beta_i$  αντιστοιχεί την καθυστέρηση μετάδοσης και αποτελεί την κλίση της ευθείας  $T_i(L)$ . Επομένως, η χωρητικότητα μέχρι το  $i$ -οστό hop του μονοπατιού δίνεται από τον τύπο:

$$C_i = \frac{1}{\beta_i - \beta_{i-1}} \quad (2.31)$$

Η τεχνική αυτή μπορεί να οδηγήσει σε σημαντική υποτίμηση της χωρητικότητας εάν στο μονοπάτι συμπεριλαμβάνονται μεταγωγείς Επιπέδου 2, διότι οι συσκευές αυτές δεν παράγουν ICMP TTL-expired replies.

### Packet Pair/Train Dispersion Probing (PPTD)

Η τεχνική αυτή έχει ως στόχο τη μέτρηση της end-to-end χωρητικότητας ενός μονοπατιού [14]. Βασίζεται στην αποστολή πολλών ζευγών ίδιου μεγέθους πακέτων από την πηγή στον προορισμό, ο διαχωρισμός<sup>7</sup> (dispersion) των οποίων σχετίζεται με τη χωρητικότητα

<sup>7</sup> Δηλαδή η χρονική απόσταση μεταξύ του τελευταίου bit του πρώτου και του δεύτερου πακέτου.

του μονοπατιού μέσω του ακόλουθου τύπου:

$$\Delta_R = \max_{i=0,\dots,H} \left( \frac{L}{C_i} \right) = \frac{L}{\min_{i=0,\dots,H} (C_i)} = \frac{L}{C}, \quad (2.32)$$

όπου  $\Delta_R$  το συνολικό dispersion έως ότου το ζεύγος πακέτων φτάσει στον προορισμό,  $C_i$  η χωρητικότητα κάθε hop  $(0,\dots,H)$ ,  $L$  το μέγεθος κάθε πακέτου του ζεύγους και  $C$  η χωρητικότητα του μονοπατιού<sup>8</sup>.

Η τεχνική αυτή μπορεί, παρουσία cross-traffic, να οδηγήσει είτε σε υποτίμηση (τα πακέτα του cross-traffic μεσολαβούν των πακέτων του ζεύγους, αυξάνοντας το dispersion αυτών) είτε σε υπερτίμηση της χωρητικότητας (τα πακέτα του cross-traffic καθυστερούν το πρώτο πακέτο του ζεύγους περισσότερο από το δεύτερο, σε κάποια ακμή που ακολουθεί την narrow στο μονοπάτι).

Διάφορες τεχνικές που κάνουν χρήση ζευγών πακέτων έχουν προταθεί έως τώρα, όμως, τελικά, η έρευνα στράφηκε σε μία παραλλαγή της προηγούμενης μεθόδου, η οποία περιλαμβάνει την αποστολή πολλών διαδοχικών πακέτων (packet trains), όπου το dispersion ορίζεται ως η χρονική απόσταση μεταξύ του τελευταίου bit του πρώτου και του τελευταίου πακέτου του packet train ( $\Delta_R(N)$ ). Από το dispersion αυτό μπορεί να υπολογιστεί το dispersion rate ως εξής:

$$D = \frac{(N-1)L}{\Delta_R(N)} \quad (2.33)$$

Απουσία cross-traffic, το dispersion rate ισούται με την χωρητικότητα του μονοπατιού. Παρουσία, όμως, cross-traffic μπορεί να υπολογιστεί μονάχα ένα μέσο dispersion rate (Average Dispersion Rate - ADR) για το οποίο ισχύει το ακόλουθο:

$$A < ADR \leq C, \quad (2.34)$$

όπου  $A$  το διαθέσιμο εύρος ζώνης και  $C$  η χωρητικότητα του μονοπατιού.

Ορισμένα πλεονεκτήματα των τεχνικών που στηρίζονται στην αποστολή packet trains αντί για packet pairs αποτελούν το γεγονός ότι αποχτούν επιπλέον πληροφορία από τον συσχε-

<sup>8</sup>Η οποία ισούται με τη χωρητικότητα του narrow link όπως ορίστηκε προηγουμένως.

τισμό των καθυστερήσεων [16], καθώς επίσης και το γεγονός ότι μειώνονται τα σφάλματα των μετρήσεων.

### Self-Loading Periodic Streams (SLoPS)

Η τεχνική αυτή έχει ως στόχο τον υπολογισμό του end-to-end διαθέσιμου εύρους ζώνης [14]. Βασίζεται στην περιοδική αποστολή  $K$  ίδιου μεγέθους πακέτων από την πηγή στον προορισμό, με δεδομένο ρυθμό  $R$  και στην παρατήρηση των μεταβολών των καθυστερήσεων των πακέτων αυτών. Αυξανόμενες καθυστερήσεις υποδηλώνουν την ύπαρξη καθυστερήσεων αναμονής, οι οποίες κάνουν την εμφάνισή τους όταν  $R > A$ , ενώ σταθερές καθυστερήσεις υποδηλώνουν  $R < A$ <sup>9</sup>. Χρησιμοποιώντας τεχνικές αντίστοιχες της δυαδικής αναζήτησης, ο αποστολέας προσπαθεί να εντοπίσει το σημείο στο οποίο αλλάζει η μονοτονία της συνάρτησης που αντιστοιχεί στις μεταβολές των καθυστερήσεων, διότι στο σημείο εκείνο θα ισχύει ότι  $R = A$ .

Αντίστοιχη τεχνική, αλλά με αποστολή packet trains διαφορετικού μοτίβου αποτελεί η τεχνική που εφαρμόζει το εργαλείο *pathchirp* [16], το οποίο αποστέλλει  $m$  το πλήθος chirp probe trains και στη συνέχεια, πραγματοποιεί στατιστική ανάλυση των μετρήσεων. Το chirp probe train αποτελείται από  $N$  πακέτα, η απόσταση μεταξύ των οποίων αυξάνεται εκθετικά. Κατ'αυτόν τον τρόπο, όχι μόνο επιτυγχάνεται γρήγορος υπολογισμός του διαθέσιμου εύρους ζώνης, με χρήση μόλις λίγων διερευνητικών πακέτων, αλλά λαμβάνεται και χρήσιμη πληροφορία από τον συσχετισμό των καθυστερήσεων. Η εκτίμηση του διαθέσιμου εύρους ζώνης γίνεται μέσω του queuing delay signature<sup>10</sup>.

### Trains of Packet Pairs (TPP)

---

<sup>9</sup>A: Available bandwidth.

<sup>10</sup>Διάγραμμα των καθυστερήσεων αναμονής συναρτήσεως του ρυθμού αποστολής πακέτων. Το σημείο έπειτα από το οποίο η συνάρτηση γίνεται αύξουσα (και διατηρεί τη μονοτονία της) αντιπροσωπεύει ρυθμό αποστολής πακέτων ίσο με το διαθέσιμο εύρος ζώνης. Σε περίπτωση που η μονοτονία δε διατηρείται, η αυξητική τάση της συνάρτησης απλώς υποδηλώνει την ύπαρξη ριπών (bursts) cross-traffic.

Η τεχνική αυτή έχει ως στόχο τον υπολογισμό του διαθέσιμου εύρους ζώνης ενός μονοπατιού [14]. Βασίζεται στην αποστολή πολλαπλών ζευγών πακέτων με αυξανόμενους ρυθμούς από την πηγή στον προορισμό και στη σύγκριση μεταξύ του dispersion των ζευγών αυτών στην πηγή ( $R_0$ ) και στον προορισμό ( $R_m$ ). Εάν  $R_0 = R_m$ , τότε  $R_0 < A$ , ενώ αν  $R_0 > R_m$ , τότε  $R_0 > A$ . Επομένως, το μέγιστο  $R_0$  για το οποίο ισχύει ότι  $R_0 = R_m$  θα αποτελεί πολύ καλή προσέγγιση του διαθέσιμου εύρους ζώνης ( $A$ ).

Η τεχνική *TPP* μοιάζει αρκετά με την τεχνική *SLoPS*, διαφέροντας, κυρίως, στη στατιστική επεξεργασία των μετρήσεων και στον τρόπο αύξησης του  $R_0$  (στην *TPP* αυξάνεται γραμμικά, ενώ στην *SLoPS* αυξάνεται με εφαρμογή δυαδικής αναζήτησης). Επίσης, η *TPP* μπορεί να υπολογίσει και τη χωρητικότητα του tight link (η οποία, όμως, θα διαφέρει από τη χωρητικότητα του μονοπατιού, εάν το tight link είναι διαφορετικό του narrow link), από την ακόλουθη σχέση:

$$R_m = \frac{R_0}{R_0 + R_c} C, \quad (2.35)$$

όπου  $R_c = C - A$  ο μέσος ρυθμός του cross-traffic.

### Recursive Packet Train (RPT)

Η τεχνική αυτή έχει ως στόχο τον εντοπισμό του tight link ενός μονοπατιού [17]. Βασίζεται στην αποστολή 60 διαδοχικών UDP load packets, τα οποία χρησιμεύουν στην απόκτηση πληροφορίας σχετικά με το διαθέσιμο εύρος ζώνης, καθώς και 60 UDP measurement packets, εκ των οποίων τα μισά προηγούνται και τα υπόλοιπα μισά έπονται των load packets. Τα measurement packets έχουν το TTL πεδίο τους ορισμένο με τρόπο τέτοιο ώστε σε κάθε βήμα του μονοπατιού να λήγει τόσο το measurement πακέτο που βρίσκεται στην κορυφή του packet train, όσο και αυτό που βρίσκεται στην ουρά αυτού. Η λήξη των πακέτων αυτών θα προκαλέσει την αποστολή δύο ICMP TTL-expired πακέτων στην πηγή, το dispersion των οποίων μπορεί να χρησιμοποιηθεί ως εκτιμητής του μήκους του packet train στο συγκεκριμένο βήμα/ακμή. Οι ακμές στις οποίες το μήκος του packet train αυξάνεται υποδηλώνουν διαθέσιμο εύρος ζώνης μικρότερο του ρυθμού αποστολής και ονομάζονται choke points. Η downstream ακμή ενός choke point ονομάζεται choke link και η τελευταία choke link του



μονοπατιού αποτελεί την tight link.

### Minimum RTT Packet Trains

Η τεχνική αυτή έχει ως στόχο τον υπολογισμό του χρόνου εξυπηρέτησης του σημείου συμφόρησης (bottleneck service time) και κατ'επέκταση, της διαθέσιμης χωρητικότητας ενός μονοπατιού [18]. Στην προκειμένη εργασία, ως διαθέσιμη χωρητικότητα,  $\alpha_j(\Delta)$ , ορίζεται το πλήθος των δεδομένων που μπορούν να εισαχθούν σε ένα μονοπάτι του δικτύου, τη στιγμή που ένα διερευνητικό πακέτο  $j$  εισάγεται στο μονοπάτι αυτό, ώστε η καθυστέρηση διάβασης (transit delay) να μην ξεπεράσει το μέγιστο επιτρεπτό όριο,  $\Delta$ .

Αρχικά, ορίζεται ο χρόνος εικονικής αναμονής (Virtual Waiting Time)  $u^i(t)$ , σε έναν δρομολογητή  $R_i$ , ως ο χρόνος που πρέπει να περιμένει ένα πακέτο που έφθασε στον  $R_i$  τη χρονική στιγμή  $t$ , έως ότου εξυπηρετηθεί από αυτόν. Ο χρόνος εικονικής αναμονής ενός πακέτου  $j$  σε έναν δρομολογητή  $i$  δίνεται από την ακόλουθη αναδρομική σχέση:

$$u_j^i = \begin{cases} 0, & \text{αν } j \leq 1 \\ \max(u_{j-1}^i + s_j^i - (a_j^i - a_{j-1}^i), 0), & \text{διαφορετικά,} \end{cases} \quad (2.36)$$

όπου  $a_j^i$  η χρονική στιγμή άφιξης του πακέτου  $j$  στον δρομολογητή  $i$  και  $s_j^i$  ο χρόνος εξυπηρέτησης που απαιτείται ώστε να εξυπηρετηθεί ένα πακέτο  $j$  σε έναν δρομολογητή  $i$ .

Στη συνέχεια, ορίζεται μία νέα ποσότητα  $C_j$ , η οποία αντιπροσωπεύει τις καθυστερήσεις που υπόκειται το πακέτο  $j$  και οι οποίες δεν οφείλονται στους χρόνους εξυπηρέτησής του από τους διάφορους δρομολογητές του μονοπατιού και δίνεται από τον ακόλουθο τύπο:

$$C_j = d_j^m - \max(d_j^0 + \tau, d_{j-1}^m + s^{max}), \quad (2.37)$$

όπου  $\tau$  η ελάχιστη καθυστέρηση διάβασης,  $d_j^i$  η χρονική στιγμή της αναχώρησης ενός πακέτου  $j$  από έναν δρομολογητή  $i$  και  $s^{max}$  ο μέγιστος χρόνος εξυπηρέτησης μεταξύ όλων των χρόνων εξυπηρέτησης όλων των δρομολογητών του μονοπατιού.

Έπειτα, ορίζεται ο end-to-end χρόνος εικονικής αναμονής,  $V_j$ , ως ο συνολικός χρόνος που ένα πακέτο  $j$  χρειάστηκε να περιμένει έως ότου εξυπηρετηθεί από όλους τους δρομολογητές

που συνάντησε κατά μήκος του μονοπατιού:

$$V_j = \sum_i u_j^i, \quad (2.38)$$

Συνεπώς, με χρήση των εξισώσεων (2.36), (2.37) και (2.38) προκύπτει η ακόλουθη αναδρομική σχέση για τον χρόνο  $V_j$ :

$$V_j = C_j + \begin{cases} 0, & \text{αν } j \leq 1 \\ \max(V_{j-1} + s^{\max} - (a_j^i - a_{j-1}^i), 0), & \text{διαφορετικά.} \end{cases} \quad (2.39)$$

Τελικά, η διαθέσιμη χωρητικότητα,  $\alpha_j(\Delta)$ , μπορεί να υπολογιστεί μέσω του ακόλουθου τύπου:

$$\alpha_j(\Delta) = \max(\Delta - V_j, 0) \quad (2.40)$$

### TCP-based TOPP & RPT

Η τεχνική αυτή χρησιμοποιείται για την εκτίμηση του διαθέσιμου εύρους ζώνης ενός μονοπατιού και αποτελεί παραλλαγή των προαναφερθέντων μεθόδων TOPP και RPT [19]. Η διαφορά στην *TCP-based TOPP* σε σχέση με την TOPP είναι ότι πραγματοποιεί δυαδική αναζήτηση, αντί της γραμμικής, για την εύρεση του διαθέσιμου εύρους ζώνης.

Η παραλλαγή στην *TCP-based RPT* σε σχέση με την RPT είναι ότι στη θέση των UDP load packets χρησιμοποιούνται TCP RST packets και στη θέση των UDP measurement packets χρησιμοποιούνται  $N$  TCP SYN packets που προηγούνται των load packets και αποστέλλονται σε κλειστές πόρτες άρτιας απόστασης από κάποια πόρτα-βάση (BASE-PORT+2, BASE-PORT+3,...,BASE-PORT+(2N)), καθώς και  $N$  TCP SYN packets που έπονται των load packets και αποστέλλονται σε κλειστές πόρτες περιττής απόστασης από την πόρτα-βάση (BASE-PORT+3, BASE-PORT+5,...,BASE-PORT+(2N+1)). Κατ'αυτόν τον τρόπο, ο αποστολέας λαμβάνει ζεύγη TCP ACK+SYN απαντήσεων, τα οποία τού επιτρέπουν να υπολογίσει το end-to-end dispersion ολόκληρου του packet train. Με χρήση της παραλλαγμένης αυτής τεχνικής γίνεται μία εκτίμηση της end-to-end χωρητικότητας, η οποία στη συνέχεια χρησιμεύει ως ανώτατο όριο στη δυαδική αναζήτηση που

πραγματοποιείται από την *TCP based TOPP* τεχνική, με σκοπό την εύρεση του διαθέσιμου εύρους ζώνης.

Τα πλεονεκτήματα του παραπάνω συνδυασμού παραλλαγών σχετίζονται με τη χαμηλή προτεραιότητα που οι σύγχρονοι δρομολογητές δίνουν στην παραγωγή/προώθηση πακέτων ICMP (γεγονός που οδηγεί σε υποτίμηση της χωρητικότητας από τεχνικές που στηρίζονται στη λήψη τέτοιων πακέτων), με την αποφυγή του φόρτου που δημιουργούν στο δίκτυο τα ICMP replies που προκύπτουν κατά τη λήξη των πακέτων UDP load, καθώς και με τη μείωση του χρόνου που απαιτείται για την ακριβή εκτίμηση του διαθέσιμου εύρους ζώνης, χάρη στη δυαδική αναζήτηση (αντί της γραμμικής).

### Probe Gap Model (PGM)

Η τεχνική αυτή είναι συγγενής της TPP που παρουσιάστηκε στην υποενότητα 2.5.1, με την εξής προσθήκη: κατασκευάζει ένα gap model το οποίο συλλαμβάνει τη σχέση μεταξύ της ανταγωνιστικής κίνησης (cross-traffic) και του input και output gap του ζεύγους των πακέτων που αποστέλλονται. Για τον σκοπό αυτό χρησιμοποιεί τη μέθοδο *IGI* (Initial Gap Increase) για την ταυτοποίηση του μέσου input probing gap για το οποίο το μέσο output probing gap θα ισούται με το μέσο input probing gap. Στο σημείο αυτό -το επονομαζόμενο turning point- τα σφάλματα μέτρησης ελαχιστοποιούνται και η εκτίμηση του μεγέθους της ανταγωνιστικής κίνησης (συνεπώς και του διαθέσιμου εύρους ζώνης) θα είναι ακριβέστερη. Με αυτόν τον τρόπο, αποδεικνύεται η μεγάλη σημασία που κατέχει το αρχικό gap μεταξύ των πακέτων του ζεύγους που αποστέλλεται, στην ακρίβεια της τελικής εκτίμησης [20].

Στο σημείο αυτό πρέπει να διευκρινιστεί ότι όλες οι παραπάνω τεχνικές αποτελούν τεχνικές ενεργούς διερεύνησης (active probing), δηλαδή τεχνικές που εισάγουν στο δίκτυο διερευνητικά πακέτα με σκοπό να λάβουν πληροφορίες και δεν καταγράφουν απλώς τις υπάρχουσες κινήσεις, όπως θα συνέβαινε στην περίπτωση της παθητικής διερεύνησης (passive probing). Παρότι η ενεργή διερεύνηση εισάγει επιπλέον φόρτο στο δίκτυο, παρέχει καλύτερη και πιο στοχευμένη πληροφορία σε σχέση με την παθητική. Η απόφαση για την καταλλη-

λόγηση της ενεργούς διερεύνησης έγκειται στη δυνατότητα (ή όχι) περιορισμού του επιπλέον φόρτου που αυτή εισάγει στο δίκτυο, ώστε να μην προξενούνται αισθητά προβλήματα, τα οποία θα ξεπερνούν τα οφέλη που αυτή παρέχει.

### 2.5.2 Εργαλεία για την εκτίμηση των μετρικών ενός δικτύου

#### Εργαλεία για τον υπολογισμό της χωρητικότητας ανά ακμή

Στη συνέχεια παρουσιάζονται τα εργαλεία για τον υπολογισμό της χωρητικότητας ανά hop, όπως παρουσιάζονται στα [14], [21], καθώς και μία σύντομη περιγραφή αυτών.

- *Pathcar*: Πραγματοποιεί εκτιμήσεις διαφόρων χαρακτηριστικών, όπως η χωρητικότητα, οι καθυστερήσεις αναμονής και ο ρυθμός απωλειών, σε κάθε βήμα ενός μονοπατιού από την πηγή στον προορισμό. Κάνει χρήση της τεχνικής VPS για τον υπολογισμό των προαναφερθέντων προσθετικών μετρικών. Απαιτεί πρόσβαση μόνο στον αποστολέα [22].
- *Clink*: Αποτελεί επανυλοποίηση του *pathcar*, με τη διαφορά ότι χρησιμοποιεί την τεχνική even-odd για την παραγωγή των διαστημάτων εκτίμησης της χωρητικότητας και ότι όταν συναντά κάποια αστάθεια δρομολόγησης (routing instability<sup>11</sup>), συλλέγει δεδομένα από όλα τα μονοπάτια που συναντά, έως ότου κάποιο εξ'αυτών δώσει αρκετά δεδομένα για την κατασκευή μιας ακριβούς εκτίμησης. Απαιτεί πρόσβαση μόνο στον αποστολέα [22].
- *Pchar*: Παρόμοιο με το *pathcar*, αποτελεί ακόμη ένα εργαλείο που υλοποιεί την τεχνική VPS με σκοπό τον χαρακτηρισμό της χωρητικότητας, του latency και των απωλειών των ακμών κατά μήκος ενός μονοπατιού. Χρησιμοποιεί τη βιβλιοθήκη *Libpcap*, για την απόκτηση χρονοσφραγίδων επιπέδου πυρήνα (kernel-level timestamps) και τρία είδη αλγορίθμων γραμμικής παλινδρόμησης, για τον υπολογισμό της κλίσης

<sup>11</sup> Αναφέρεται σε γρήγορες αλλαγές στην προσβασιμότητα του δικτύου και στις πληροφορίες που σχετίζονται με την τοπολογία του. Έχει ως αποτέλεσμα συχνές ενημερώσεις των πινάκων δρομολόγησης, οι οποίες μεταφέρονται σε όλους τους δρομολογητές του πυρήνα του Διαδικτύου [23]

$\beta_i$  της ευθείας  $T_i(L)$ , η οποία μέσω της εξίσωσης (2.31) επιτρέπει τον υπολογισμό της χωρητικότητας,  $C_i$ . Απαιτεί πρόσβαση μόνο στον αποστολέα [22].

### Εργαλεία για τον υπολογισμό της χωρητικότητας μονοπατιού

Στη συνέχεια παρουσιάζονται ορισμένα εργαλεία για τον υπολογισμό της end-to-end χωρητικότητας, καθώς και μία σύντομη περιγραφή αυτών.

- *Bprobe*: Βασίζεται στην τεχνική Packet Pair Dispersion. Για αύξηση της ακρίβειας εκμεταλλεύεται χαρακτηριστικά που προσφέρει το λειτουργικό σύστημα IRIX της SGI (χρονοσφραγίδες υψηλής ευκρίνειας, απόδοση προτεραιότητας στη διαδικασία που παράγει τα διερευνητικά πακέτα - προς αποφυγήν του context switching) και χρησιμοποιεί φίλτράρισμα για την απόρριψη των μετρήσεων που έχουν επηρεαστεί από το cross-traffic, καθώς επίσης και πακέτα μεταβλητού μεγέθους. Επιπλέον, απαιτεί πρόσβαση μόνο στον αποστολέα/πηγή, διότι αποστέλλει πακέτα ICMP echo, στα οποία ο παραλήπτης απαντά με ICMP echo replies [14], [21].
- *Nettimer*: Είναι κυρίως γνωστό ως ένα εργαλείο για τον υπολογισμό της χωρητικότητας ενός μονοπατιού, με χρήση της τεχνικής Packet Pair Dispersion. Χρησιμοποιεί την τεχνική *Kernel density estimation*<sup>12</sup>, για τη στατιστική επεξεργασία των μετρήσεων [14]. Απαιτεί πρόσβαση μόνο στον αποστολέα [22].
- *Pathrate*: Βασίζεται στην τεχνική PPTD για τον υπολογισμό της χωρητικότητας ενός μονοπατιού. Αναλύοντας την κατανομή των μετρήσεων αποκαλύπτει όλες τις τοπικές επικρατούσες τιμές, μία εκ των οποίων αντιστοιχεί στη χωρητικότητα. Έπειτα, υπολογίζει τον ADR, τον οποίο και χρησιμοποιεί ως κάτω όριο της χωρητικότητας.

<sup>12</sup>Αποτελεί στατιστική μέθοδο, διαφορετική των histogram-based μεθόδων, η οποία ουσιαστικά συσχετίζει την πληροφορία όλων των μετρήσεων, σταθμίζοντάς τες βάσει της απόστασής τους από το σημείο στο οποίο γίνεται η εκτίμηση, ταυτοποιώντας, έτσι, την κυρίαρχη επικρατούσα τιμή της κατανομής. Η μορφή του εκτιμητή είναι:  $\hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^N K(\frac{x-x_i}{h})$ , όπου  $K$  ο kernel,  $N$  το πλήθος των μετρήσεων και  $h$  το επονομαζόμενο bandwidth (κάτι αντίστοιχο του bin width στα ιστογράμματα).

Τελικά, εκτιμά πως η χωρητικότητα του μονοπατιού θα είναι η ισχυρότερη τοπική επικρατούσα τιμή που είναι μεγαλύτερη του ADR. Απαιτεί πρόσβαση και εγκατάσταση λογισμικού τόσο στον αποστολέα, όσο και στον παραλήπτη [14], [22].

- *Sprobe*: Βασίζεται στην τεχνική Packet Pair Dispersion και υπολογίζει τη χωρητικότητα ενός μονοπατιού αποστέλλοντας πακέτα TCP SYN, τα οποία αναγκάζουν τον παραλήπτη να απαντά, αποστέλλοντας ο ίδιος πακέτα TCP RST στην πηγή, η οποία, με τη σειρά της, υπολογίζει το dispersion των πακέτων που δέχεται. Επομένως, απαιτείται πρόσβαση μόνο στον αποστολέα/πηγή [14].

### Εργαλεία για τον υπολογισμό του διαθέσιμου εύρους ζώνης μονοπατιού

Στη συνέχεια παρουσιάζονται ορισμένα εργαλεία για τον υπολογισμό του end-to-end διαθέσιμου εύρους ζώνης ενός μονοπατιού, καθώς και μία σύντομη περιγραφή αυτών.

- *Pathload*: Βασίζεται στην τεχνική SLoPS και υπολογίζει ένα εύρος εκτιμήσεων, το κέντρο του οποίου αποτελεί το μέσο διαθέσιμο εύρος ζώνης κατά το χρονικό διάστημα των μετρήσεων, ενώ το ίδιο το εύρος αντιστοιχεί στη διακύμανση αυτού κατά το ίδιο χρονικό διάστημα. Απαιτεί πρόσβαση τόσο στην πηγή/αποστολέα, όσο και στον προορισμό/παραλήπτη [14].
- *PathChirp*: Βασίζεται σε παραλλαγή της τεχνικής SLoPS με αποστολή chirp probe trains όπως αναφέρθηκε στην υποενότητα 2.5.1. Απαιτεί πρόσβαση και εγκατάσταση λογισμικού τόσο στον αποστολέα, όσο και στον παραλήπτη [24].
- *IGI*: Βασίζεται στην τεχνική PGM και καθορίζει το βέλτιστο αρχικό gap μεταξύ του ζεύγους των πακέτων που αποστέλλονται από την πηγή προς τον προορισμό ενός μονοπατιού, για τον υπολογισμό του διαθέσιμου εύρους ζώνης του μονοπατιού αυτού [20]. Απαιτεί πρόσβαση και εγκατάσταση λογισμικού τόσο στον αποστολέα, όσο και στον παραλήπτη [25].
- *Spruce*: Βασίζεται στην τεχνική PGM για τον υπολογισμό του διαθέσιμου εύρους ζώνης ενός μονοπατιού. Θεωρώντας γνωστή τη χωρητικότητα C του μονοπατιού και

ορίζοντας το time gap εισόδου του ζεύγους πακέτων,  $\Delta_{in}$ , μετρά στον αποστολέα το time gap εξόδου του ζεύγους αυτού,  $\Delta_{out}$ , και μέσω του τύπου  $\frac{\Delta_{out}-\Delta_{in}}{\Delta_{in}} \times C$  κάνει μία εκτίμηση του διαθέσιμου εύρους ζώνης του μονοπατιού. Για αύξηση της ακρίβειας αποστέλλει μια σειρά από ζεύγη πακέτων και ορίζει το inter-pair time gap βάσει μιας εκθετικά κατανομημένης συνάρτησης, γεγονός που οδηγεί σε δειγματοληψία Poisson (μείωση του φόρτου που εισάγεται στο δίκτυο - non-intrusive) [26].

Στις εργασίες [14], [22], [27]–[30] πραγματοποιείται συγκριτική ανάλυση της επίδοσης, της εγκυρότητας και της παρεμβατικότητας ορισμένων εκ των εργαλείων που αναφέρθηκαν σε αυτό κεφάλαιο.





## Κεφάλαιο 3

---

### Ουδετερότητα Δικτύου

---

#### 3.1 Γενικά

Η αύξηση του πλήθους των τελικών χρηστών του Διαδικτύου, επομένως και του μεγέθους του, εγείρει προβληματισμούς σχετικά με την ποιότητα των παρεχόμενων υπηρεσιών (*QoS*), τις πηγές εσόδων των διαφόρων φορέων που σχετίζονται με τις υπηρεσίες του, καθώς και τους σχετικούς κανονισμούς για τον έλεγχο των παραπάνω.

Ο όρος ‘ποιότητα υπηρεσιών’ σχετίζεται με την κατάσταση των ακόλουθων παραμέτρων κατά την παροχή κάποιας υπηρεσίας από ένα επικοινωνιακό δίκτυο προς τους συνδρομητές του:

1. Ταχύτητα του δικτύου πρόσβασης (*access – network*)
2. Συμφόρηση στο δίκτυο κορμού (*backbone – network*)
3. Καθυστερήσεις στη μετάδοση (*latency*)
4. Διακύμανση στην καθυστέρηση (*jitter*)
5. Απώλεια πληροφορίας κατά τη μετάδοση.

Η αύξηση του μεγέθους του διαδικτύου οδηγεί σε έναν τεράστιο και συνεχώς αυξανόμενο όγκο δεδομένων προς μετάδοση. Για το λόγο αυτό απαιτείται ο σχεδιασμός/υιοθέτηση 'έξυπνων' μεθόδων διαχείρισης των δεδομένων αυτών, με σκοπό την αποδοτική χρήση της υπάρχουσας χωρητικότητας για την ικανοποίηση των τελικών χρηστών, τη διαχείριση επειγουσών ροών δεδομένων, αλλά και την εξασφάλιση/αύξηση του κέρδους των εμπλεκόμενων φορέων.

Τα παραπάνω οδηγούν στην ανάγκη για ιεράρχηση των διαφόρων κινήσεων δεδομένων, με τρόπο διαφανή και προς όφελος όλων των εμπλεκόμενων. Παρόλα αυτά, έχουν υπάρξει περιπτώσεις διάκρισης των κινήσεων δεδομένων, με τρόπο αδιαφανή και αμφιλεγόμενο. Το 2009, στη Γαλλία ψηφίστηκε νομοθεσία κατά της μη εξουσιοδοτημένης ανάκτησης υλικού που υπόκειται σε πνευματικά δικαιώματα, η οποία οδήγησε μέχρι και σε βεβαιωμένη αποσύνδεση από το Διαδίκτυο χρηστών που πραγματοποίησαν διαμοιρασμό αρχείων. Στον Αραβικό κόσμο, οι Κυβερνήσεις απάντησαν στη διαδικτυακή αγκιτάσια ορισμένων πολιτών μπλοκάροντάς τους, όχι μόνο από το Διαδίκτυο, αλλά και από τις υπηρεσίες κινητής τηλεφωνίας [31].

Επιπλέον, η εκρηκτική αύξηση του μεγέθους του Διαδικτύου το καθιστά πολύ σημαντικό παράγοντα της παγκόσμιας οικονομίας. Συγκεκριμένα, ο τομέας Πληροφορικής και Επικοινωνιών παρήγαγε έσοδα της τάξης των 2000 δις ευρώ παγκοσμίως (7% του παγκόσμιου ΑΕΠ) το 2011 (αναμένοντας να φτάσει στο 20% έως το 2021) και της τάξης των 660 δις ευρώ στην Ευρώπη (5% του Ευρωπαϊκού ΑΕΠ) [31].

Ο τεράστιος οικονομικός ρόλος του διαδικτύου δημιουργεί συμφέροντα, συχνά αντικρουόμενα, τα οποία απαιτούν σχετική νομοθεσία, ώστε να μην πραγματοποιούνται ανεξέλεγκτες ενέργειες καταπάτησης των ανθρωπίνων δικαιωμάτων, όπως αυτές που παρουσιάστηκαν προηγουμένως. Έτσι φτάνουμε στο πρόβλημα της *Ουδετερότητας του Δικτύου*, η οποία ορίζεται ως εξής:

Ένα ουδέτερο δίκτυο αντιμετωπίζει κάθε περιεχόμενο, κάθε ιστότοπο και κάθε εφαρμογή ισότιμα. Επίσης, επιτρέπει σε κάθε συμβατή συσκευή να συνδεθεί σε αυτό και αντιμετωπίζει κάθε κίνηση που διέρχεται από αυτό με τον ίδιο τρόπο [32].

Φυσικά υπάρχουν και εναλλακτικοί ορισμοί της Ουδετερότητας Δικτύου, όπως αυτός που δόθηκε από τους *Hahn* και *Wallsten* και ο οποίος παρουσιάζεται στο [33] :

Η Ουδετερότητα Δικτύου συνήθως σημαίνει ότι οι ευρυζωνικοί πάροχοι υπηρεσιών χρεώνουν τους καταναλωτές μόνο μία φορά για την πρόσβασή τους στο Διαδίκτυο, δεν ευνοούν κάποιον *CP* έναντι κάποιου άλλου και δεν χρεώνουν τους *CPs* για την πληροφορία που στέλνουν μέσω των ευρυζωνικών γραμμών τους στους τελικούς χρήστες.

## 3.2 Παρουσίαση των διαφορετικών προσεγγίσεων του προβλήματος της Ουδετερότητας Δικτύου

### 3.2.1 Κατηγοριοποίηση των δικτυακών κινήσεων ανάλογα με τις απαιτήσεις τους σε εύρος ζώνης

Η αύξηση της χωρητικότητας για την κάλυψη των αυξανόμενων αναγκών εγείρει το ερώτημα του ποια θα πρέπει να είναι η πηγή της χρηματοδότησης προκειμένου κάτι τέτοιο να συμβεί. Η απάντηση στο ερώτημα αυτό είναι αρκετά δύσκολη, διότι οι χρήστες δεν είναι έτοιμοι να πληρώσουν περισσότερα για ένα Διαδίκτυο ευρέως χρησιμοποιούμενο και κομμάτι, της καθημερινότητάς τους και διότι δεν υπάρχει, ακόμη, κάποια ξεκάθαρη στρατηγική σχετικά με τα έσοδα των παρόχων περιεχομένου από διαφημίσεις (π.χ. εάν θα πρέπει να τα εκμεταλλευτούν από κοινού με τους παρόχους υπηρεσιών - ορισμός *side payments*).

Στο [32] παρουσιάζεται ένα μοντέλο για την κατηγοριοποίηση των δικτυακών κινήσεων με τρόπο που να αντανακλά τα διαφορετικά τους χαρακτηριστικά (βαθμός επιρροής, αντοχή, απαιτήσεις) αναφορικά με την υφιστάμενη καθυστέρηση, τη διακύμανση των καθυστερήσεων και την απώλεια πακέτων.

Αρχικά, ορίζεται η *Ποιότητα Εμπειρίας* (*Quality of Experience - QoE*), σε αντιδιαστο-

λή με την Ποιότητα υπηρεσιών (QoS) που ορίστηκε προηγουμένως, η οποία αποτελεί ένα υποκειμενικό μέτρο της ποιότητας των υπηρεσιών, όπως αυτή γίνεται αντιληπτή μέσω της εμπειρίας των χρηστών.

Έπειτα, παρουσιάζεται ο διαχωρισμός των χρηστών σε 3 κατηγορίες: *standard*, *gaming*, *business*. Κάθε κατηγορία μπορεί, ακολούθως, να διαχωριστεί σε κλάσεις. Οι προτεινόμενες κλάσεις φαίνονται στον πίνακα (3.1) και η επιλογή τους αναλύεται στη συνέχεια.

Πίνακας 3.1: Προτεινόμενες κλάσεις χρηστών

| Class                 | Included Services   | Downlink (kbit/s) |
|-----------------------|---|-------------------|
| Standard Class Type 1 | Skype (Voice calling) + Video streaming (Youtube H.264/AAC, 640x360) + Web browsing         | 600               |
| Standard Class Type 2 | Skype (Video calling) + Video streaming (Youtube H.264/AAC, 640x360) + Web browsing         | 698               |
| Gaming Class Type 1   | Teamspeak + Online gaming RPG (WoW) + Web browsing  | 145.9             |
| Gaming Class Type 2   | Teamspeak + Online gaming FPS (UT2k3) + Web browsing  | 196               |
| Business Class Type 1 | H.323 Video-conference + Video streaming (Youtube H.264/AAC, 640x360) + Web browsing        | 954               |
| Business Class Type 2 | H.323 Video-conference (HD) + Video streaming (Youtube H.264/AAC, 1920x1080) + Web browsing | 6990              |

Προκειμένου να δικαιολογηθεί ο παραπάνω διαχωρισμός πρέπει να αναλυθεί η φύση

των διαφόρων υπηρεσιών.

Αρχικά, το *Skype* παρέχει υπηρεσίες αποστολής πακέτων, στιγμιαίων μηνυμάτων, *video-conferencing* και *VoIP*, οι οποίες έχουν διαφορετικές απαιτήσεις αναφορικά με το εύρος ζώνης, ξεκινώντας από τη βασική φωνητική κλήση και φτάνοντας έως τις υψηλής ανάλυσης βίντεο-συνεδριάσεις πολλαπλών συμμετεχόντων.

Έπειτα, το *YouTube* αποτελεί τον διασημότερο ιστότοπο κοινοποίησης βίντεο, επιτρέποντας το ανέβασμα, την κοινοποίηση και την παρακολούθηση αυτών. Και πάλι οι απαιτήσεις σε εύρος ζώνης διαφέρουν ανάλογα με την περίπτωση. Η *default* κωδικοποίηση για το βίντεο είναι η *H.264/MPEG-4 AVC* και για τον ήχο η *AAC*, ενώ οι υποστηριζόμενες αναλύσεις κυμαίνονται από  $640 \times 360$  έως  $1920 \times 1080$  και  $4096 \times 3072$ .

Η πλοήγηση στον ιστό βασίζεται στο μοντέλο πελάτη - εξυπηρετητή. Το ένα τρίτο των αντικειμένων των ιστοσελίδων έχει μέγεθος  $5KB$ , ενώ η δεύτερη πιο κοινή κατανομή μεγεθών ανήκει στο εύρος των  $30-50KB$ . Επομένως τα πιο κοινά αντικείμενα που συναντώνται σε ιστοσελίδες είναι σχετικά μικρά σε μέγεθος. Η έκδοση 1.1 του *HTTP* πρωτοκόλλου χρησιμοποιεί μία *TCP* συνεδρία για την επικοινωνία μεταξύ χρήστη και εξυπηρετητή. Επομένως, αυτό που καθορίζει την απαίτηση σε εύρος ζώνης της διαδικασίας πλοήγησης στον ιστό, δεν είναι ούτε ο χρόνος φόρτωσης της ιστοσελίδας ούτε το μέτρο της ταχύτητας του αντικειμένου του ιστού, αλλά η ίδια η φύση της ασφαλούς και αξιόπιστης *TCP* επικοινωνίας. Προσομοιώσεις έχουν δείξει πως η διεκπεραιωτική ικανότητα των αντικειμένων του ιστού βρίσκεται στην περιοχή των  $20-30kbit/s$ .

Το *Teamspeak* είναι μία *VoIP* εφαρμογή απευθυνόμενη, κυρίως, στη διαδικτυακή *gaming* κοινότητα. Βασίζεται στην κωδικοποίηση *Speex25.9kbit/s* για βέλτιστη καθαρότητα ήχου και επιτρέπει τη συμμετοχή και ταυτόχρονη ομιλία πολλών χρηστών που ανήκουν στο ίδιο κανάλι. Η εφαρμογή αυτή προσφέρει αξιόλογο *QoE*, με μικρή κατανάλωση εύρους ζώνης.

Τα Διαδικτυακά παιχνίδια διακρίνονται στα δύο ακόλουθα είδη: *MMORPG* (*Massively Multiplayer Online Role Playing Game*) και *FPS* (*First-Person Shooter*). Τα παιχνίδια που ανήκουν στην πρώτη κατηγορία διαρκούν αρκετό χρόνο (έως και μερικές ώρες), επομένως έχουν μεγάλες απαιτήσεις σε εύρος ζώνης, σε αντίθεση με αυτά που ανήκουν στη δεύτερη κατηγορία, τα οποία διαρκούν λιγότερο ( $\sim 15'$ ). Επίσης, τα *FPS* απαιτούν την

άμεση αντίδραση των παικτών, επομένως έχουν μικρή ανοχή σε καθυστερήσεις, σε αντίθεση με τα *MMORPG* τα οποία εξελίσσονται σχετικά αργά. Ένας ακόμη παράγοντας διαφοροποίησης των απαιτήσεων των δύο προαναφερθέντων τύπων διαδικτυακών παιχνιδιών είναι τα χρησιμοποιούμενα από αυτά πρωτόκολλα. Κατά κύριο λόγο, τα *MMORPG* χρησιμοποιούν *TCP* πρωτόκολλα, με μεγαλύτερη ανοχή στις καθυστερήσεις σε σχέση με τα *FPS*, που βασίζονται σε *UDP* πρωτόκολλα.

Η εφαρμογή *H.323 Video-conference* απευθύνεται στις απαιτήσεις των επιχειρήσεων για βίντεο - συνεδριάσεις καλής ποιότητας και υψηλής αξιοπιστίας. Η προσφερόμενη ποιότητα μπορεί να είναι είτε κανονική, καταναλώνοντας εύρος ζώνης  $384\text{ kbit/s}$ , είτε υψηλή, καταναλώνοντας  $1920\text{ kbit/s}$ .

Ο στόχος του παραπάνω διαχωρισμού είναι να χρησιμοποιηθεί από τους χειριστές του δικτύου, προκειμένου να μπορέσουν να επιλέξουν τις κλάσεις χρηστών και το πλήθος των χρηστών κάθε κλάσης που το δίκτυό τους μπορεί να υποστηρίξει, ώστε να παρέχουν στους χρήστες τους ικανοποιητικό *QoS*.

### 3.2.2 Μοντελοποίηση των ουδέτερων και μη δικτύων με χρήση Θεωρίας Παιγνίων

Η μοντελοποίηση με χρήση της Θεωρίας Παιγνίων αποσκοπεί στην εύρεση της χρυσής τομής όσον αφορά στο διαμοιρασμό των κερδών από τους τελικούς χρήστες μεταξύ των παρόχων υπηρεσιών και περιεχομένου, αλλά και στην εύρεση στρατηγικών σχετικά με τις χρεώσεις που πρέπει να επιβάλλονται στους τελικούς χρήστες.

Υπάρχουν δύο είδη Παιγνίων: τα συνεργατικά και τα μη-συνεργατικά. Η λύση του Παιγνίου θα πρέπει να είναι τέτοια ώστε όλοι οι συμμετέχοντες (Πάροχοι Περιεχομένου, Πάροχοι Υπηρεσιών, Τελικοί Χρήστες και πιθανόν ο Τομέας της Διαφήμισης) να είναι ικανοποιημένοι. Δηλαδή, οι χρήστες να διαθέτουν πρόσβαση στις υπηρεσίες που επιθυμούν και οι *ISPs* και *CPs* να έχουν τα σωστά κίνητρα και ευκαιρίες για τη διαρκή πραγματοποίηση επενδύσεων και καινοτομιών.

Ένα σχέδιο λύσεων στα μη συνεργατικά παίγνια αποτελεί η Ισορροπία του *Nash*, η οποία

σχετίζεται με την υπόθεση ότι ο κάθε παίκτης γνωρίζει τις στρατηγικές ισορροπίας που θα ακολουθήσουν οι υπόλοιποι παίκτες και κανένας παίκτης δεν έχει να κερδίσει τίποτα αν αλλάξει μόνο τη δική του στρατηγική. Αν κάθε παίκτης έχει επιλέξει μία στρατηγική και κανένας παίκτης δεν έχει όφελος να αλλάξει τη στρατηγική του όσο οι άλλοι παίκτες κρατάνε τη δική τους στρατηγική σταθερή, τότε το σύνολο των στρατηγικών και των αποτελεσμάτων τους αποτελεί Ισορροπία Nash [34].

Στο [31] παρουσιάζονται τα μοντέλα/παίγνια διαφόρων εργασιών και τα συμπεράσματά τους σχετικά με τις σχέσεις που πρέπει να διέπουν τους συμμετέχοντες των παιγνίων.

Σε κάποια μη συνεργατικά παίγνια με έναν πάροχο περιεχομένου (*CP*), έναν πάροχο υπηρεσιών (*ISP*) και ένα πλήθος τελικών χρηστών, χωρίς διαφοροποίηση του *QoS* προέκυψε πως ο *CP* πρέπει να παρέχει παράλληλες πληρωμές στον *ISP*, λόγω των εσόδων που ο πρώτος λαμβάνει από τη διαφήμιση. Οι τιμές αυτές, όμως, είτε θα πρέπει να ορίζονται από κάποια ρυθμιστική αρχή, είτε η πληρωμή στον *ISP* θα πρέπει να γίνεται τόσο από τον *CP* όσο και από τους τελικούς χρήστες, εναλλάξ.

Λαμβάνοντας υπόψιν στην προηγούμενη εργασία το *QoS* που παρέχεται στους τελικούς χρήστες από τον *ISP*, προκύπτει το συμπέρασμα ότι παράλληλες πληρωμές στον *ISP* από τον *CP* μπορεί να οδηγήσουν σε μειωμένο *QoS*. Επίσης, μπορεί να οδηγήσουν σε υποτίμηση, από μεριάς των χρηστών, του *CP* σε σχέση με τον *ISP*, κάτι που μπορεί να επιφέρει την ανάγκη για *reverse payments*, δηλαδή πληρωμές από τον *ISP* στον *CP*, ώστε να μειωθεί η τιμή της υπηρεσίας του *CP*, κάνοντάς τον και πάλι επιθυμητό στους τελικούς χρήστες.

Στη συνέχεια, από κάποια μη συνεργατικά παίγνια με έναν *ISP* και δύο *CPs* και με διαφοροποίηση του *QoS*, συμπεραίνεται ότι τα μη-ουδέτερα δίκτυα συνεπάγονται έλλειψη επενδυτικού κινήτρου τόσο στον *ISP* (και κατ'επέκταση μείωση του παρεχόμενου από αυτόν *QoS*), όσο και στους *CPs*, λόγω του φόβου απαλλοτρίωσης των κερδών τους από τον *ISP*.

Σε ένα άλλο παίγνιο με δύο διασυνδεδεμένους *ISPs* που ανταγωνίζονται τόσο για τους *CPs*, όσο και για τους τελικούς χρήστες, προκύπτει ένας διαφορετικός ορισμός της ουδετερότητας του δικτύου. Σύμφωνα με αυτόν, σε ένα ουδέτερο δίκτυο ο *CP* πρέπει να πληρώνει μόνο τον *ISP* με τον οποίο είναι άμεσα συνδεδεμένος, ενώ σε ένα μη ουδέτερο δίκτυο πρέπει να πληρώνει και τον *ISP* μέσω του οποίου γίνεται προσιτός στους τελικούς

του χρήστες και με τον οποίο δεν είναι άμεσα συνδεδεμένος.

Αναλογικά με την ιδέα της Ισορροπίας του Nash στα μη συνεργατικά παίγνια, υπάρχει η ιδέα του πυρήνα (*core*) στα συνεργατικά παίγνια, όπου καμία απόκλιση από τη μεγάλη συμμαχία  $\mathcal{N}$  (όπου  $\mathcal{N}$  το σύνολο των *ISP* του δικτύου) δεν είναι επικερδής. Η τιμή *Shapley* ορίζεται ως:

$$\phi_i(\mathcal{N}, v) = \frac{1}{\mathcal{N}!} \sum_{\pi \in \Pi} \Delta_i(v, S(\pi, i)), \forall i \in \mathcal{N} \quad (3.1)$$

όπου  $v$  η *worth function*, η οποία υποδεικνύει το κέρδος που παράγεται από ένα υποσύνολο  $\mathcal{S}$  των *ISP*,  $\Pi$  είναι το σύνολο όλων των  $\mathcal{N}!$  διατάξεων του  $\mathcal{N}$  και  $S(\pi, i)$  είναι το σύνολο των παικτών που προηγούνται του  $i$  στη διάταξη  $\pi$ . Αποδεικνύεται ότι η τιμή *Shapley* βρίσκεται στον πυρήνα των κυρτών παιγνίων (*convex games*) [35]. Ως κυρτά παίγνια ορίζονται αυτά των οποίων η χαρακτηριστική συνάρτηση (η οποία περιγράφει την ανταμοιβή κάθε μονάδας [36]) είναι *supermodular*, δηλαδή χαρακτηρίζεται από στρατηγικές συμπληρωματικότητας. Τουτέστιν, όταν ένας παίκτης επιλέγει μία κίνηση μεγαλύτερης τιμής, για να εξασφαλίσει την αύξηση της χρησιμότητάς του, τότε και οι υπόλοιποι παίκτες έχουν κίνητρο να κάνουν το ίδιο για να αναβαθμίσουν τη χρησιμότητά τους [37]. Με άλλα λόγια, τα κίνητρα για την συμμετοχή στη Μεγάλη Συμμαχία αυξάνονται, καθώς αυξάνεται το μέγεθος της συμμαχίας [38]. Η τιμή *Shapley* ουσιαστικά υπολογίζει τη συνεισφορά κάθε παρόχου στο δίκτυο, προκειμένου να κατανειμίσει δίκαια τα κέρδη μεταξύ αυτών.

Πιο συγκεκριμένα, σύμφωνα με το [33], η τιμή *Shapley* αναλαμβάνει να υπολογίσει τη δίκαιη κατανομή των εσόδων των *CPs* που προέρχονται από τις διαφημίσεις και ενεργοποιούνται κατά την πρόσβαση κάποιου συνδρομητή στο δίκτυο, το παρεχόμενο από τον *ISP*, ή/και των εσόδων των *ISPs* που προέρχονται από τις συνδρομές των χρηστών και ενεργοποιούνται από τη ζήτηση του περιεχομένου, του παρεχόμενου από τον *CP*.

Στην εργασία [33], το πρόβλημα της κατανομής των κερδών μοντελοποιείται ως ένα παίγνιο με τρεις παίκτες: τους τελικούς χρήστες, τον πάροχο που επιτρέπει στους τελικούς χρήστες την πρόσβαση στο δίκτυο, επονομαζόμενο και ως *ISP1* και τον πάροχο περιεχομένου μαζί με τον πάροχο υπηρεσιών αυτού, επονομαζόμενος και ως *CP2*. Γίνεται η υπόθεση ότι η ζήτηση-απόκριση των χρηστών ακολουθεί το απλό γραμμικό μοντέλο που δίνεται εν συνε-



χεία:  $D = D_0 - D(p_1 + p_2)$ , όπου  $p_1$  και  $p_2$  οι επιβαλόμενες στους χρήστες τιμές από τον ISP1 και από τον CP2, αντίστοιχα,  $d$  ο συντελεστής απόκρισης στις μεταβολές των τιμών και  $D_0$  η ζήτηση υπό σταθερή χρέωση ( $p_1 = 0 = p_2$ ).

Δεδομένου ότι η ζήτηση παίρνει μη αρνητικές τιμές (δηλαδή  $p_1 + p_2 \leq \frac{D_0}{d} =: p_{max}$ ) και τα έσοδα του  $i$ -οστού παρόχου δίνονται από τη σχέση  $U_i = Dp_i$ , προκύπτουν τα ακόλουθα συμπεράσματα:

- Το σενάριο ενός συνεργατικού παιχνίδι, με συμμαχία μεταξύ των ISP1 και CP2 οφείλει και τους δύο παίκτες, αυξάνοντας τα έσοδά τους σε σύγκριση με το μη συνεργατικό παίγνιο ( $U_{i\_non\_coop}^* = \frac{D_0^2}{9d}$ , ενώ  $U_{i\_coop}^* = \frac{D_0^2}{8d}$ ).
- Το σενάριο ενός μη συνεργατικού παιχνίδι με *side – payments* ( $p_s$ ) από τον πάροχο  $i$  στον πάροχο  $j$  οδηγεί είτε σε ίδια έσοδα όπως και στην περίπτωση του μη συνεργατικού παιχνίδι χωρίς *side – payments*, όταν  $|p_s| \leq \frac{1}{3}p_{max}$ , είτε σε  $U_j^* \geq U_i^*$ , όταν  $p_s \geq \frac{1}{3}p_{max}$  και ο πάροχος  $j$  θέσει το  $p_j = 0$  (ώστε να αυξήσει τη ζήτηση), εξασφαλίζοντας μεγαλύτερα έσοδα μόνο από τα *side – payments*, είτε σε  $U_i^* \geq U_j^*$  εάν  $p_s < -\frac{1}{3}p_{max}$  και  $p_i = 0$ .
- Στην περίπτωση της ύπαρξης επιπλέον διαφημιστικών εσόδων ( $Dp_a$ , όπου  $p_a$  μία σταθερή παράμετρος) για τον CP2, όπως και πριν, όλοι οι παίκτες επωφελούνται της συμμαχίας των ISP1 και CP2, διότι οι τιμές μειώνονται, η ζήτηση αυξάνεται και ως αποτέλεσμα, τα έσοδα και των δύο παρόχων, επίσης, αυξάνονται ( $U_{i\_non\_coop}^* = \frac{D_0+dp_a^2}{9d}$ , ενώ  $U_{i\_coop}^* = \frac{D_0+dp_a^2}{8d}$ ).
- Το σενάριο της ισορροπίας Stackelberg (δηλαδή ενός μη συμμετρικού παιχνίδι στο οποίο ο ένας παίκτης είναι αρχηγός και οι υπόλοιποι είναι ακόλουθοι) έχει ως αποτέλεσμα ο παίκτης - αρχηγός να έχει μεν τα διπλάσια έσοδα από τους παίκτες - ακολούθους, όχι όμως και περισσότερα από το συνεργατικό σενάριο της προηγούμενης περίπτωσης ( $U_{leader}^* = \frac{D_0+dp_a^2}{8d}$ ,  $U_{follower}^* = \frac{D_0+dp_a^2}{16d}$ ).
- Εξετάζεται η περίπτωση όπου ο ISP1 προσφέρει δύο κατηγορίες υπηρεσιών πρόσβασης στο δίκτυο, μία υψηλής προτεραιότητας, με τιμή  $p_h$ , και μία χαμηλής, με τιμή  $p_l$ .

Στην περίπτωση της συνεργασίας μεταξύ του ISP1 και του CP2, προκύπτει πως η βέλτιστη λύση έγκειται στην ύπαρξη usage-based τιμών μόνο από τον CP2 (δηλαδή  $p_h = p_l = 0$ ), με αποτέλεσμα τα έσοδα του ISP1 να προκύπτουν μόνο από flat-rate τιμές. Τα έσοδα και των δύο παρόχων θα είναι  $U_{total} = \frac{D_0^2}{4d}$ , όταν  $p_2 = \frac{p_{max}}{2}$ . Στην περίπτωση της μη συνεργασίας, και πάλι οι usage-based τιμές για τις υπηρεσίες παροχής πρόσβασης στο δίκτυο πρέπει να είναι 0, όμως τώρα ο CP2 θα έχει ξεκάθαρο πλεονέκτημα εσόδων σε σύγκριση με τον ISP1.

### 3.2.3 Ανίχνευση παραβιάσεων της ουδετερότητας σε δίκτυα και εντοπισμός των μη-ουδέτερων ακμών βάσει εξωτερικών παρατηρήσεων

Στο [39] αποδεικνύονται οι συνθήκες υπό τις οποίες μπορούν να ανιχνευθούν παραβιάσεις της ουδετερότητας σε δίκτυα και με χρήση εξωτερικών παρατηρήσεων να εντοπιστούν οι ακμές οι οποίες την παραβιάζουν. Το αν ένα δίκτυο είναι ή όχι ουδέτερο, σχετίζεται με τη συνέπεια ή την ασυνέπεια, αντιστοίχως, εξωτερικών παρατηρήσεων προερχόμενων από διαφορετικά σημεία παρακολούθησης (vantage points).

Στην παρούσα εργασία, η παραβίαση της ουδετερότητας ορίζεται ως η κατάσταση στην οποία κινήσεις προερχόμενες από διαφορετικά μονοπάτια του δικτύου υφίστανται διαφορετική επίδοση ενώ διασχίζουν την ίδια ακμή του δικτύου. Η διαφορετική αυτή αντιμετώπιση από το δίκτυο βασίζεται στον τύπο της κίνησης (που αναγνωρίζεται μέσω των περιεχομένων των επικεφαλίδων Επιπέδου 2 ή 3 ή/και του χρήσιμου φορτίου -payload- μέσω της τεχνικής deep packet inspection) και εκδηλώνεται με διαφορετικούς τρόπους (throttling, shaping, blocking ή/και deprioritization της κίνησης δεδομένων).

Προκειμένου να προσεγγιστεί το πρόβλημα του συναγωγής της ουδετερότητας του δικτύου (network neutrality inference) πραγματοποιούνται οι ακόλουθες υποθέσεις: ο γράφος του δικτύου που διασυνδέει τα διάφορα σημεία παρακολούθησης θεωρείται γνωστός και η κατάσταση κάποιας ακμής του δικτύου θεωρείται ανεξάρτητη της κατάστασης οποιασδήποτε άλλης ακμής του δικτύου.

Το πρόβλημα μοντελοποιείται ως ένας γράφος  $\mathcal{G}(\mathcal{V}, \mathcal{L}, \mathcal{P})$ , όπου  $\mathcal{V}, \mathcal{L}$  το σύνολο των κόμβων και των ακμών του γράφου του δικτύου, αντίστοιχα, και  $\mathcal{P}$  το σύνολο όλων των μονοπατιών του δικτύου που βρίσκονται την παρούσα στιγμή σε χρήση. Επίσης, ορίζονται τα ακόλουθα:  $\lambda$ , μία αλληλουχία διαδοχικών ακμών ( $l \in L$ ) που δε σχηματίζει βρόγχους,  $p \in \mathcal{P}$  κάποιο μονοπάτι σε χρήση το δεδομένο χρονικό διάστημα,  $\pi$  κάποιο σύνολο μονοπατιών (pathset), το οποίο ανήκει στο δυναμοσύνολο του  $\mathcal{P}$ ,  $\mathcal{P}^*$ . Επιπλέον, ορίζονται τα εξής:  $Paths(l)$ , το σύνολο όλων των μονοπατιών που διασχίζουν την ακμή  $l$ ,  $Paths(\lambda)$ , το σύνολο όλων των μονοπατιών που διασχίζουν όλες τις ακμές στο σύνολο  $\lambda$ ,  $Links(p)$ , το σύνολο όλων των ακμών που διασχίζονται από το  $p$  και  $Links(\pi)$ , το σύνολο όλων των ακμών που διασχίζονται από ένα, τουλάχιστον, μονοπάτι του συνόλου μονοπατιών  $\pi$ . Εν συνεχεία, ορίζεται η κλάση επίδοσης (performance class) ως το σύνολο των μονοπατιών που δέχονται ισότιμη μεταχείριση από το δίκτυο,  $c_n, n = 1, \dots, |C|$ , η  $n$ -οστή κλάση θεωρώντας τυχαία διάταξη των  $|C|$  το πλήθος κλάσεων του δικτύου και  $x_k(n) | n = 1, \dots, |C|$ , οι επιδόσεις των ακμών  $l_k$ . Επίσης, ορίζεται η κλάση υψίστης προτεραιότητας (top-priority) για μία μη-ουδέτερη ακμή, ως η κλάση για την οποία η ακμή έχει την υψηλότερη επίδοση. Εδώ πρέπει να τονιστεί η υπόθεση ότι εάν, για οποιοδήποτε χρονικό διάστημα, μία μη-ουδέτερη ακμή παρουσιάζει υπολογίσιμη απώλεια πακέτων στην top-priority κλάση της, τότε θα εισάγει υπολογίσιμες απώλειες και σε οποιαδήποτε άλλη κλάση. Τέλος, ορίζεται το σύνολο των ουδέτερων ακμών του δικτύου,  $L_n$ , το σύνολο των μη-ουδέτερων ακμών,  $L_{\bar{n}}$ , καθώς και οι ακόλουθες μετρικές επίδοσης:  $\hat{x}(n) \equiv \log(\mathbb{P}(\lambda, c_n)) | n = 1, \dots, |C|$ , όπου  $\mathbb{P}(\lambda, c_n)$  η πιθανότητα το  $\lambda$  να μην είναι συμφορημένο (congestion-free) αναφορικά με την κλάση  $c_n$  κατά τη διάρκεια οποιουδήποτε χρονικού διαστήματος και  $y \equiv \log(\mathbb{P}(\pi))$ , όπου  $\mathbb{P}(\pi)$  η πιθανότητα να μην είναι συμφορημένο κανένα μονοπάτι του συνόλου  $\pi$  κατά τη διάρκεια οποιουδήποτε χρονικού διαστήματος.

Δεδομένων όλων των παραπάνω μπορεί να οριστεί ξεκάθαρα το πρόβλημα του neutrality inference ως ένα πρόβλημα το οποίο με είσοδο ένα δίκτυο  $G$  και τις μετρικές επίδοσης,  $y$ , κάθε pathset  $\pi \in \mathcal{P}^*$ , ζητά να υπολογίσει το σύνολο  $L_{\bar{n}}$ .

Ακολούθως, ορίζεται το Σύστημα (3.3), αντίστοιχο του (2.10), με τη διαφορά τώρα, ότι

ο  $\mathbf{A}$  αντιστοιχεί στον πίνακα γενικευμένης δρομολόγησης που ορίζεται ως εξής:

$$A_{ik} = \begin{cases} 1, & \text{if at least one path in pathset } \pi_i \text{ traverses link } l_k \\ 0, & \text{otherwise.} \end{cases} \quad (3.2)$$

$$\vec{y} = A(\Pi) \cdot \vec{x} \quad (3.3)$$

**Λήμμα 3.2.1.** Έστω δίκτυο με  $P$  μονοπάτια. Εάν υπάρχει σύνολο από pathsets  $\Pi \subseteq P^*$ , ώστε το Σύστημα (3.3) να μην έχει λύση, τότε το δίκτυο είναι non-neutral (μη-ουδέτερο).

**Ορισμός 3.2.2.** Έστω non-neutral δίκτυο με  $P$  μονοπάτια. Θα λέμε ότι, η παραβίαση της ουδετερότητας του δικτύου είναι παρατηρήσιμη, όταν υπάρχει ένα σύνολο από pathsets,  $\Pi \subseteq P^*$ , τέτοιο ώστε το Σύστημα (3.3) να μην έχει λύση.

Στη συνέχεια, ορίζεται το *Ισοδύναμο Ουδέτερο Δίκτυο* ενός μη-ουδέτερου δικτύου, το οποίο κατασκευάζεται με αντιστοίχιση των μη-ουδέτερων ακμών του αρχικού δικτύου σε τόσες εικονικές, ουδέτερες ακμές, όσες και οι κλάσεις επίδοσης που διαθέτει η αρχική μη-ουδέτερη ακμή (οι μετρικές επίδοσης θα προκύπτουν ως εξής:  $x_i(1), x_i(2) - x_i(1), \dots, x_i(n) - x_i(n-1)$ , όπου  $x_i(1)$  η επίδοση της top priority κλάσης της  $l_i$  και  $x_i(n)$  η επίδοση της κλάσης με τη μικρότερη προτεραιότητα της  $l_i$ ). Επομένως, αν το αρχικό δίκτυο έχει  $|L_n|$  ουδέτερες ακμές και  $|L_{\bar{n}}|$  μη-ουδέτερες ακμές, το ισοδύναμό του ουδέτερο δίκτυο θα έχει  $|L_n| + |L_{\bar{n}}| \cdot |C|$  ουδέτερες ακμές.

**Θεώρημα 3.2.3.** Έστω μη-ουδέτερο δίκτυο με  $L$  ακμές και το Ουδέτερο Ισοδύναμό του με  $L^+$  ακμές. Η παραβίαση της ουδετερότητας του δικτύου είναι παρατηρήσιμη αν υπάρχει τουλάχιστον μία εικονική ακμή,  $l^+(n) \in L^+$ , διαχωρίσιμη από οποιαδήποτε άλλη ακμή στο  $L$ .

Έως τώρα εξετάστηκαν οι συνθήκες υπό τις οποίες μπορεί να ανιχνευθεί η παραβίαση της ουδετερότητας ενός δικτύου. Στη συνέχεια, εφαρμόζοντας το 3.2.1 για ένα κομμάτι του δικτύου (network slice,  $G_\lambda$ ), επιλεγμένο έτσι ώστε κάθε παρατηρήσιμη παραβίαση της

ουδετερότητας σε αυτό το slice να μπορεί να αποδοθεί μόνο στο  $\lambda$ , μπορεί κανείς να αποφανθεί για την ουδετερότητα της αλληλουχίας ακμών,  $\lambda$ . Για τον σκοπό αυτό, σχηματίζεται το ακόλουθο εξειδικευμένο σύστημα (ως παραλλαγή του (3.3) που σχηματίστηκε προηγουμένως):

$$\vec{y} = A_\lambda(\Pi_\lambda) \cdot \vec{x} \quad (3.4)$$

Για την κατασκευή του (3.4), σχηματίζεται ένα ειδικό σύνολο από pathsets,  $\Pi_\lambda$ , το οποίο θα περιέχει όλα τα ζεύγη μονοπατιών  $p_i, p_j \in P$ , για τα οποία ισχύει ότι  $Links(p_i) \cap Links(p_j) = \lambda$ , καθώς και τα επιμέρους μονοπάτια που συνιστούν τα ζεύγη αυτά. Έπειτα, κατασκευάζεται το slice  $G_\lambda$ , ώστε να αποτελεί ένα λογικό δέντρο δύο επιπέδων: η λογική ακμή που αντιστοιχεί στην αλληλουχία  $\lambda$  αποτελεί το πρώτο επίπεδο και οι λογικές ακμές που αντιστοιχούν στις επιμέρους αλληλουχίες ακμών του δικτύου που δεν εμπεριέχουν τις ακμές  $l_i \in \lambda$ , αποτελούν το δεύτερο επίπεδο.

Κατ'αυτόν τον τρόπο μπορεί να οριστεί η αναγνωρισιμότητα των μη-ουδέτερων ακμών του δικτύου.

**Λήμμα 3.2.4.** Έστω μια αλληλουχία ακμών  $\lambda$  και το σύνολο από pathsets  $\Pi_\lambda$ . Αν το Σύστημα (3.4) δεν έχει λύση, τότε η αλληλουχία  $\lambda$  δεν είναι ουδέτερη (non-neutral).

**Ορισμός 3.2.5.** Έστω μη-ουδέτερη αλληλουχία ακμών  $\lambda$  και το σύνολο από pathsets  $\Pi_\lambda$ . Θα λέμε ότι η αλληλουχία  $\lambda$  είναι αναγνωρίσιμη, όταν το Σύστημα (3.4) δεν έχει λύση.

Τέλος, βάσει όλων όσων προαναφέρθηκαν, κατασκευάζεται και παρουσιάζεται ένας αλγόριθμος για την ταυτοποίηση των μη-ουδέτερων αλληλουχιών ακμών ενός δικτύου.

### 3.2.4 Διαφάνεια στη θέση της ουδετερότητας

Στο [40] παρουσιάζεται μία νέα προσέγγιση του προβλήματος της ουδετερότητας δικτύου, σύμφωνα με την οποία οι τεχνικές προσπάθειες ορισμού της ουδετερότητας ενός δικτύου

είναι καταδικασμένες και τη θέση τους πρέπει να πάρει η προσπάθεια για την επίτευξη της διαφάνειας (transparency) του δικτύου.

Στη συγκεκριμένη εργασία, υποστηρίζεται ότι δεν είναι εφικτός ένας γενικός ορισμός της ουδετερότητας ενός δικτύου και παρουσιάζονται περιπτώσεις, όπου η απάντηση στο αν έχει ή όχι παραβιαστεί η ουδετερότητα είναι διφορούμενη. Για παράδειγμα, ένα δίκτυο είναι ουδέτερο όταν προσφέρει συνδέσεις δίχως συμφόρηση ή συνδέσεις ίσης χωρητικότητας προς όλες τις κινήσεις δεδομένων; Επιπλέον, η προτεραιότητα κινήσεων που σχετίζονται με ιατρικές εφαρμογές έναντι κινήσεων που δεν είναι latency-sensitive (δηλαδή η αύξηση του latency δε τις ζημιώνει αισθητά) αποτελεί παραβίαση της ουδετερότητας;

Στη συνέχεια της εργασίας περιγράφεται ο τρόπος μέσω του οποίου μπορεί να επιτευχθεί η Διαφάνεια του δικτύου. Ο τρόπος αυτός, σχετίζεται με την εξαγωγή/διαμοιρασμό πληροφορίας που αφορά στην κίνηση ενός πεπερασμένου συνόλου πακέτων, τα οποία διαθέτουν ένα κοινό υποσύνολο επικεφαλίδων (aggregate packets), εντός ενός Τομέα Διοίκησης (Administrative Domain - AD) X, από τον AD, σε οποιαδήποτε οντότητα Y βρίσκεται εντός του μονοπατιού που διασχίζεται από το σύνολο των πακέτων αυτών και τα παρατηρεί. Ο σκοπός αυτής της διαδικασίας διαμοιρασμού πληροφορίας (σχετικά με τυχόν απώλειες/καθυστερήσεις που υφίστανται τα πακέτα του συνόλου εντός του X) είναι η εκτίμηση της εμπειρίας των πακέτων εντός του X, η οποία επιτρέπει τον έλεγχο της τήρησης των διαφόρων SLAs μεταξύ των ADs, των διαφόρων παρόχων υπηρεσιών, αλλά και των υπηρεσιών που προσφέρονται από τους παρόχους στον χρήστη.

Παρ'όλα αυτά, η εφαρμογή της παραπάνω διαδικασίας στην πράξη συνοδεύεται από ορισμένα προβλήματα. Πρώτα από όλα, υπάρχουν περιορισμοί σχετικοί με το κόστος (της υλοποίησης, των υπολογισμών, της πιθανής εγκατάστασης νέου εξοπλισμού κλπ.), καθώς και ζητήματα σχετικά με την κλιμακωσιμότητα και ευελιξία της παραπάνω διαδικασίας. Επιπροσθέτως, προκύπτουν θέματα εμπιστοσύνης απέναντι στις οντότητες που εξάγουν την πληροφορία (είναι η πληροφορία αυτή αυθεντική ή αντιμετωπίζονται ευνοϊκά τα πακέτα που δειγματοληπτούνται με σκοπό να χρησιμοποιηθούν στις μετρήσεις;). Τέλος, τίθενται ζητήματα ιδιωτικότητας και ασφάλειας του δικτύου των οντοτήτων που αποκαλύπτουν πληροφορίες για τα πακέτα που τα διασχίζουν.

Τέλος, παρουσιάζονται προτάσεις για την αντιμετώπιση των παραπάνω προβλημάτων, αφήνοντας, όμως, ακόμη ανοιχτά προβλήματα, σχετικά με τον τρόπο που πρέπει να γίνεται η συλλογή/εξαγωγή των πληροφοριών, ώστε να μην εκτίθενται ευαίσθητες πληροφορίες που μπορεί να διακινδυνεύσουν την ασφάλεια των οντοτήτων που τις αποκαλύπτουν, αλλά και σχετικά με το πώς μπορεί να αποτραπεί ο διαμοιρασμός πληροφοριών που δεν ανταποκρίνονται στην πραγματικότητα (dishonest ADs), με χρήση υποδομής της οποίας οι απαιτήσεις να μην ξεπερνούν τα ωφέλη.





## Κεφάλαιο 4

---

### Παρουσίαση του μοντέλου και του πειραματικού σχήματος

---

#### 4.1 Περιγραφή και μοντελοποίηση του προβλήματος

Το πρόβλημα που αποτέλεσε το αντικείμενο της παρούσας διπλωματικής υπάγεται στο πεδίο της *Τομογραφίας Δικτύου* αποσκοπεί στη συναγωγή τόσο της τοπολογίας του δικτύου, όσο και των χαρακτηριστικών κάθε ακμής του δικτύου αυτού, με χρήση μόνο ακραίων μετρήσεων.

Το πρόβλημα προσεγγίστηκε μέσω της υλοποίησης του *συσσωρευτικού αλγορίθμου UP-GMA* (ο οποίος θα παρουσιαστεί αναλυτικά στο Κεφάλαιο 5) και της κατασκευής προσθετικών μετρικών, για την περιγραφή των απωλειών (losses) και των καθυστερήσεων (delays) των ακμών του δικτύου, βάσει των εξισώσεων που παρουσιάστηκαν στην Υποενότητα 2.3.2.

Για τη συλλογή των ακραίων μετρήσεων, που αποτελούν την είσοδο του αλγορίθμου UPGMA, κατασκευάστηκαν τοπολογίες στο jFed (περισσότερες πληροφορίες για το setup θα δοθούν στην ενότητα 4.2). Οι τοπολογίες που κατασκευάστηκαν έχουν τα ακόλουθα χαρακτηριστικά:

1. Διαθέτουν δενδρική, ιεραρχική, δομή.

2. Αποτελούν λογικές τοπολογίες, δηλαδή τοπολογίες όπου διαδοχικοί κόμβοι δίχως διακλαδώσεις συμπίπτουν σε έναν. Μία λογική ακμή μπορεί να αντιστοιχεί σε μία ή περισσότερες φυσικές ακμές.
3. Αποτελούν δυαδικά δένδρα, όπου κάθε κόμβος που δεν είναι φύλλο έχει δύο ακριβώς παιδιά<sup>1</sup> (διαφορετικά η τοπολογία δεν θα ήταν λογική), εξαιρουμένου του αρχικού κόμβου/πηγής (source), που βρίσκεται στην κορυφή του δέντρου και διαθέτει ένα μόνο παιδί.
4. Αποτελούν υπερμετρικά δένδρα, δηλαδή δένδρα όπου οι αποστάσεις από την πηγή προς οποιονδήποτε τελικό κόμβο (φύλλο/end-host) είναι ίσες.

#### 4.1.1 Προσθετικοί και υπερμετρικοί χώροι

Οι αλγόριθμοι ομαδοποίησης (clustering algorithms), στους οποίους εμπίπτει και ο συσσωρευτικός UPGMA αλγόριθμος που υλοποιήθηκε στα πλαίσια της παρούσας διπλωματικής, βασίζονται στην τεχνική της ομαδοποίησης σε clusters, με τρόπο ώστε τα δεδομένα που βρίσκονται εντός του ιδίου cluster να έχουν όσο το δυνατόν περισσότερα κοινά χαρακτηριστικά [41]. Στην περίπτωση της ανακατασκευής δενδρικών τοπολογιών, τα χαρακτηριστικά αυτά αντιστοιχούν στις ανά-ζεύγη αποστάσεις των φύλλων του δένδρου, με τις οποίες κατασκευάζεται πίνακας αποστάσεων, μεγέθους  $i \times i$ , όπου  $i$  το πλήθος των φύλλων του δένδρου. Προκειμένου οι αποστάσεις αυτές να μπορούν να χρησιμοποιηθούν για τη σωστή ανακατασκευή της τοπολογίας, θα πρέπει είτε να είναι προσθετικές (additive), είτε υπερμετρικές (ultrametric). Στην πρώτη περίπτωση χρησιμοποιείται ο αλγόριθμος NeighborJoining, ενώ στη δεύτερη ο αλγόριθμος UPGMA [42].

Στους προσθετικούς πίνακες, όλες οι τετράδες σημείων,  $\{i, j, m, n\}$ , ικανοποιούν τη *Συνθήκη Τεσσάρων Σημείων* (4PC), η οποία γενικεύει την τριγωνική ανισότητα (για  $m = n$ )

<sup>1</sup> Αυτό διαφέρει από τον επίσημο ορισμό του δυαδικού δένδρου, σύμφωνα με τον οποίο κάθε κόμβος μπορεί να έχει το πολύ δύο παιδιά.

και παρουσιάζεται ακολούθως [43]:

$$D(i, j) + D(m, n) \leq \max\{D(i, m) + D(j, n), D(j, m) + D(i, n)\} \quad (4.1)$$

**Ορισμός 4.1.1.** Ένας πίνακας που ικανοποιεί την 4PC ονομάζεται προσθετικός.

**Θεώρημα 4.1.2.** Ένας πίνακας αποστάσεων μπορεί να αναπαρασταθεί από δένδρο ανν είναι προσθετικός.

Στους υπερμετρικούς πίνακες, όλες οι τριάδες σημείων,  $\{x, y, z\}$ , ικανοποιούν μία συνθήκη ισχυρότερη της 4PC [44], [45], η οποία φαίνεται ακολούθως:

$$d(x, y) \leq \max\{d(x, z), d(y, z)\} \quad (4.2)$$

**Λήμμα 4.1.3.** Αν ένας πίνακας είναι υπερμετρικός, τότε είναι και προσθετικός. Το αντίστροφο δεν ισχύει.

Οι υπερμετρικοί πίνακες αποστάσεων επιτρέπουν την ανακατασκευή δένδρων με ρίζα (rooted trees), όπου όλες οι αποστάσεις από τη ρίζα προς τα φύλλα είναι ίσες. Οι προσθετικοί πίνακες αποστάσεων επιτρέπουν την ανακατασκευή δένδρων δίχως ρίζα (unrooted trees), δηλαδή συνδεδεμένων, μη-κατευθυνόμενων γράφων, δίχως βρόγχους<sup>2</sup> [46].

#### 4.1.2 Κατασκευή του πίνακα αποστάσεων βάσει των απωλειών και των καθυστερήσεων των ακμών

Για να είναι εφικτή η αντιστοίχιση των αποστάσεων στις απώλειες και τις καθυστερήσεις των ακμών πρέπει τα χαρακτηριστικά αυτά να υποστούν κατάλληλους μετασχηματισμούς, ώστε να μετατραπούν σε δενδρικές μετρικές. Οι μετασχηματισμοί αυτοί, όπως παρουσιάστηκαν και στην Υποενότητα 2.3.2, είναι οι ακόλουθοι:

<sup>2</sup>Στα δυαδικά δένδρα δίχως ρίζα (binary unrooted trees) όλοι οι εσωτερικοί κόμβοι έχουν βαθμό ακριβώς ίσο με 3, ενώ στα δυαδικά δένδρα με ρίζα όλοι οι εσωτερικοί κόμβοι έχουν ακριβώς δύο παιδιά.

1. Η δενδρική μετρική που σχετίζεται με τις απώλειες των ακμών υπολογίζεται ως ακολούθως:

Ο ρυθμός επιτυχίας στον κόμβο  $i$  υπολογίζεται ως ο μέσος όρος του πλήθους των πακέτων που έφτασαν στον τελικό κόμβο  $i$ , στο χρονικό διάστημα  $t = 1, \dots, n$ :

$$\bar{L}_i = \frac{1}{n} \sum_{t=1}^n L_i^{(t)} \quad (4.3)$$

Ο από κοινού ρυθμός επιτυχίας των τελικών κόμβων  $\{i, j\}$  υπολογίζεται ως ο μέσος όρος του πλήθους των πακέτων που έφτασαν και στον κόμβο  $i$  και στον κόμβο  $j$ , στο χρονικό διάστημα  $t = 1, \dots, n$ :

$$\bar{L}_{ij} = \frac{1}{n} \sum_{t=1}^n L_i^{(t)} L_j^{(t)} \quad (4.4)$$

Οι εξισώσεις (4.3) και (4.4) αποτελούν τη βάση για την κατασκευή του πίνακα αποστάσεων βάσει των απωλειών των ακμών,  $D^L$ , όπου  $D_{ij}^L = \log\left(\frac{\bar{L}_i \bar{L}_j}{\bar{L}_{ij}^2}\right)$ , ενώ τα μήκη των μεμονωμένων ακμών δίνονται από τον τύπο:  $d_i = \log(\bar{L}_i)$ .

Εφόσον το ζητούμενο χαρακτηριστικό των ακμών είναι οι απώλειες, μετά τον υπολογισμό των υπερμετρικών  $d_i$  από τον αλγόριθμο, είναι απαραίτητος ο αντίστροφος μετασχηματισμός:

$$\mathbf{losses\_i} = \mathbf{1} - e^{-d_i} \quad (4.5)$$

2. Η δενδρική μετρική που σχετίζεται με τις καθυστερήσεις των ακμών υπολογίζεται ως ακολούθως:

Η διακύμανση της καθυστέρησης (η οποία ισούται με το  $jitter^2$ ),  $T_i$ , από την πηγή έως τον τελικό κόμβο  $i$ , στο χρονικό διάστημα  $t = 1, \dots, n$ , υπολογίζεται ως εξής:

$$\hat{var}(T_i) = \frac{1}{n-1} \sum_{t=1}^n (T_i^{(t)} - \bar{T}_i)^2 \quad (4.6)$$

Η συνδιακύμανση της καθυστέρησης από την πηγή προς τους τελικούς κόμβους  $i, j$ , υπολογίζεται μέσω των καθυστερήσεων των πακέτων που έφτασαν και στον κόμβο  $i$  και στον κόμβο  $j$ , στο χρονικό διάστημα  $t = 1, \dots, n$ :

$$\hat{cov}(T_i, T_j) = \frac{1}{n-1} \sum_{t=1}^n (T_i^{(t)} - \bar{T}_i)(T_j^{(t)} - \bar{T}_j) \quad (4.7)$$

Οι εξισώσεις (4.6) και (4.7) αποτελούν τη βάση για την κατασκευή του πίνακα αποστάσεων βάσει των καθυστερήσεων των ακμών,  $D^T$ , όπου  $D_{ij}^T = \hat{v}ar(T_i) + \hat{v}ar(T_j) - 2 \cdot \hat{c}ov(T_i, T_j)$ , ενώ τα μήκη των μεμονωμένων ακμών δίνονται από τη διακύμανση της καθυστέρησης:  $d_i = \hat{v}ar(T_i)$ .

Εφόσον το ζητούμενο χαρακτηριστικό των ακμών είναι το jitter, στο τέλος του αλγορίθμου και πριν τυπωθούν τα αποτελέσματα, χρειάζεται να γίνει ο ακόλουθος μετασχηματισμός:

$$\mathbf{jitter.i} = \sqrt{\mathbf{d}_i} \quad (4.8)$$

Δεδομένου ότι οι απώλειες και οι καθυστερήσεις των ακμών αποτελούν δενδρικές (προσθετικές) μετρικές, με κατάλληλη επιλογή των τιμών αυτών, κατά την κατασκευή της τοπολογίας (ώστε να ικανοποιείται η υπερμετρική ιδιότητα 4.2), είναι εφικτό να κατασκευαστούν υπερμετρικοί χώροι και συνεπώς, υπερμετρικοί πίνακες αποστάσεων, οι οποίοι, χορηγούμενοι στον αλγόριθμο UPGMA, θα επιτρέψουν τη σωστή ανακατασκευή της τοπολογίας από αυτόν.

## 4.2 Κατασκευή τοπολογιών και συλλογή μετρήσεων μέσω του jFed

Οι τοπολογίες που αποτέλεσαν τη βάση της πειραματικής διαδικασίας της παρούσας διπλωματικής, δημιουργήθηκαν στο jFed, ένα σύνολο εργαλείων συμβατών με αρχιτεκτονικές SFA (Slice-based Federation Architecture), υλοποιημένο σε Java, με σκοπό την υποστήριξη ενός συστήματος δοκιμών (testbed), που ανήκει στο Fed4FIRE πρότζεκτ και προσφέρει τη δυνατότητα διεξαγωγής πειραμάτων. Τα εργαλεία αυτά διατίθενται μαζί με *Διεπαφή Γραφικών των Χρηστών* (GUI) και *Διεπαφή Γραμμής Εντολών* (CLI).

Κατά τη δημιουργία της τοπολογίας και πριν την έναρξη του πειράματος, γίνεται επιλογή των μηχανημάτων και του λειτουργικού συστήματος αυτών και ορίζονται οι απώλειες και οι καθυστερήσεις των συνδέσεων. Στη συνέχεια, το πείραμα εκκινείται, δεσμεύονται οι κα-

τάλληλοι πόροι και μόλις οι πόροι είναι έτοιμοι για χρήση, ξεκινά το configuration αυτών. Χρησιμοποιείται η σουίτα λογισμικού δρομολόγησης (Routing Software Suite), *Quagga*, η οποία περιέχει υλοποιήσεις διαφόρων πρωτοκόλλων δρομολόγησης. Στην παρούσα εργασία χρησιμοποιήθηκε το πρωτόκολλο *OSPF* (Open Shortest Path First), για unicast εκπομές και το πρωτόκολλο *pimd* (Protocol Independent Multicast Deamon), για multicast εκπομές, ενώ έγινε χρήση και του IP routing manager *Zebra*, που παρέχει δυνατότητες για ενημέρωση των πινάκων δρομολόγησης του πυρήνα (kernel routing table updates), αναζητήσεις διεπαφών (interface lookups) και αναδιανομή των διαδρομών δρομολόγησης (redistribution of routes) μεταξύ των διαφορετικών πρωτοκόλλων. Επιπλέον, γίνεται χρήση του, ενσωματωμένου στο *Quagga*, shell *vttysh*, μέσω του οποίου γίνεται η αντιστοίχιση των διεπαφών σε διευθύνσεις IP, καθώς και ο ορισμός παθητικών διεπαφών (passive interfaces<sup>3</sup>). Τέλος, έγινε χρήση του utility *Traffic Control (tc)* του Linux, το οποίο προσφέρει τη δυνατότητα διαμόρφωσης του χρονοδρομολογητή των πακέτων του πυρήνα (kernel packet scheduler) και εν προκειμένω, χρησιμοποιήθηκε για τη διαμόρφωση του jitter, δηλαδή της διακύμανσης των καθυστερήσεων, των ακμών.

Για την παραγωγή και multicast αποστολή κινήσεων δεδομένων χρησιμοποιείται το εργαλείο *iperf-ssm*, το οποίο τρέχει τόσο στην πηγή, όπου δημιουργούνται τα UDP πακέτα, τα οποία αποστέλλονται στους τελικούς κόμβους, οι οποίοι συνθέτουν ένα multicast group, όσο και στους τελικούς κόμβους, οι οποίοι λαμβάνουν τα πακέτα αυτά. Για την καταγραφή της κίνησης, χρησιμοποιείται το εργαλείο *tcp-dump*, το οποίο, επίσης, τρέχει τόσο στην πηγή όσο και στους τελικούς κόμβους.

Εν συνεχεία, μέσω των scripts, που παρουσιάζονται στο Παράρτημα Β, γίνεται επεξεργασία των αρχείων καταγραφής (log files) που προέκυψαν από το εργαλείο *tcp-dump* και κατασκευή των πινάκων αποστάσεων, όπως περιγράφηκε στην Υποενότητα 4.1.2.

Τέλος, οι πίνακες αποστάσεων που κατασκευάζονται χορηγούνται ως είσοδοι στον αλγόριθμο UPGMA, για την εξαγωγή της τοπολογίας και των επιμέρους μετρικών των ακμών.

---

<sup>3</sup>Οι παθητικές διεπαφές είναι αυτές στις οποίες δεν παραδίδονται πληροφορίες δρομολόγησης (πηγή και τελικοί κόμβοι).

## Κεφάλαιο 5

---

### Παρουσίαση του αλγορίθμου και των αποτελεσμάτων

---

#### 5.1 Unweighted Pair Group Method with Arithmetic mean - UPGMA

Ο αλγόριθμος UPGMA αποτελεί την απλούστερη εκ των μεθόδων που βασίζονται σε πίνακες αποστάσεων. Αποδίδεται στους Sokal και Michener (1958) και πρωταρχικός του στόχος ήταν η ανακατασκευή φυλογενετικών δένδρων (numerical taxonomy). Βασίζεται στην υπόθεση του μοριακού ρολογιού<sup>1</sup> (molecular clock hypothesis) και έχει συχνή εφαρμογή στο πεδίο της Βιολογίας και της Οικολογίας [48].

Στο πρώτο βήμα ο αλγόριθμος εντοπίζει, στον πίνακα αποστάσεων, τους δύο τελικούς κόμβους,  $i$ ,  $j$ , με τη μικρότερη απόσταση και τους συνενώνει. Ο πίνακας αποστάσεων ανανεώνεται ως εξής: αντικαθιστώνται οι γραμμές και οι στήλες του που αντιστοιχούν στους κόμβους  $i$ ,  $j$ , με μία γραμμή και στήλη που αντιστοιχεί στην ομάδα  $\{i, j\}$ . Έπειτα, κατασκευάζεται ο κόμβος  $p$  με τον οποίο τώρα ενώνονται οι  $i$ ,  $j$  (ο τελευταίος κοινός τους πρόγονος

---

<sup>1</sup>Σύμφωνα με την υπόθεση του μοριακού ρολογιού το DNA και οι αλληλουχίες πρωτεϊνών εξελίσσονται σε έναν σχετικά σταθερό ρυθμό με την πάροδο του χρόνου και μεταξύ διαφορετικών οργανισμών. Ως αποτέλεσμα αυτού, η γενετική διαφορά μεταξύ δύο ειδών είναι ανάλογη του χρόνου που παρήλθε από την τελευταία φορά που τα δύο αυτά είδη μοιράζονταν κοινό πρόγονο [47].

- ο πατέρας των  $\{i, j\}$ ). Οι αποστάσεις ανανεώνονται ως εξής:

$$D_{ip} = D_{jp} = \frac{1}{2}D_{ij} \quad (5.1)$$

$$D_{px} = \frac{1}{2}(D_{ix} + D_{jx}), \quad (5.2)$$

όπου  $x$  τελικός κόμβος διάφορος των  $i, j$ . Κάθε φορά που δύο κόμβοι (ή ομάδες κόμβων) ενώνονται και δημιουργείται μία νέα ομάδα, ανανεώνεται ο αριθμός  $n_i$  που αντιπροσωπεύει το πλήθος των στοιχείων της ομάδας  $i$  (και ο οποίος είναι αρχικά ίσος με 1). Οπότε η σχέση (5.2) σε κάθε επόμενο βήμα παίρνει την ακόλουθη μορφή:

$$D_{px} = \frac{1}{n_i + n_j}(n_i D_{ix} + n_j D_{jx}) \quad (5.3)$$

Όλες οι υπόλοιπες αποστάσεις μπορούν να υπολογιστούν μέσω των (5.1) και (5.3). Ο αλγόριθμος τερματίζει όταν ο πίνακας αποστάσεων καταλήξει να έχει ένα και μοναδικό στοιχείο (διάσταση  $1 \times 1$ ).

---

#### Αλγόριθμος 1 UPGMA pseudocode

---

$N = \text{total\_number\_of\_clusters}$

$\text{distances} = \text{empty}$

initially for all clusters:  $|\text{cluster}| = 1$

**while**  $N > 1$  **do**

$\text{find\_clusters}\{i, j\} : D_{ij} = \min(D)$

$p : \text{merge}(\{i\}, \{j\})$  and  $\text{delete}(\{i\}, \{j\})$

$\text{add\_to\_distances}(p, D_{ij}/2)$

$\text{add\_new\_edges } |(p, i)| = D_{ij}/2 - \text{distances}_i$  and  $|(p, j)| = D_{ij}/2 - \text{distances}_j$

**for**  $k = 0$  **to**  $\text{dim}[1]$  **do**

$D_{ik} = D_{ki} = \frac{D_{ik} \cdot |i| + D_{jk} \cdot |j|}{|i| + |j|}$

**end**

$\text{delete\_from\_Distance\_Matrix}(j^{\text{th}}\text{row}, j^{\text{th}}\text{column})$

**end**

---

Η υλοποίηση του αλγορίθμου UPGMA σε Python παρουσιάζεται στο Παράρτημα Β.



## 5.2 Αποτελέσματα και σχολιασμός

Για τη συλλογή των μετρήσεων πραγματοποιήθηκαν δεκατρείς (13) τοπολογίες στο jFed, επτά (7) εξ'αυτών απαρτίζονται από ακμές με καθυστερήσεις (και jitter) και οι υπόλοιπες έξι (6) από ακμές με απώλειες. Οι τοπολογίες κατασκευάστηκαν έτσι ώστε να διαθέτουν τα χαρακτηριστικά που παρουσιάστηκαν στην Ενότητα 4.1.

Οι ακραίες μετρήσεις που συλλέχθηκαν από τα πειράματα που πραγματοποιήθηκαν στο jFed χορηγήθηκαν στον αλγόριθμο Upsilon, ο οποίος επέτρεψε την πλήρη ανακατασκευή των τοπολογιών, με 100% επιτυχία (0% σφάλματα), καθώς και τη συναγωγή των χαρακτηριστικών (καθυστερήσεις ή απώλειες) όλων των ακμών κάθε τοπολογίας, με μεγάλη ακρίβεια (τα σφάλματα για κάθε τοπολογία παρουσιάζονται στη συνέχεια).

Όλες οι τοπολογίες που κατασκευάστηκαν στο jFed, καθώς και αυτές που προέκυψαν από τον Upsilon, παρουσιάζονται στο Παράρτημα Α. Στον Πίνακα 5.1 παρουσιάζονται τα στοιχεία των διαφόρων τοπολογιών, καθώς και τα σφάλματα στην εκτίμηση των χαρακτηριστικών των ακμών ( $d_i$ ) μίας τοπολογίας  $t$ , βάσει του ακόλουθου τύπου:

$$e_t = \frac{1}{n} \sum_{i=1}^n (\hat{d}_i - d_i)^2, \quad (5.4)$$

όπου  $\hat{d}_i$  η εκτίμηση που δίνει ο αλγόριθμος για το jitter ή τις απώλειες της ακμής  $i$ ,  $d_i$  η πραγματική τιμή του εκάστοτε χαρακτηριστικού της ακμής  $i$  και  $n$  το πλήθος των ακμών της τοπολογίας  $t$ .

Το μέσο σφάλμα στην εκτίμηση των καθυστερήσεων και για τις επτά τοπολογίες (με καθυστερήσεις στις ακμές) που κατασκευάστηκαν, είναι  $0.6068ms$ , ενώ το μέσο σφάλμα στην εκτίμηση των απωλειών και για τις έξι τοπολογίες (με απώλειες στις ακμές) που κατασκευάστηκαν, είναι  $0.0133\%$ .

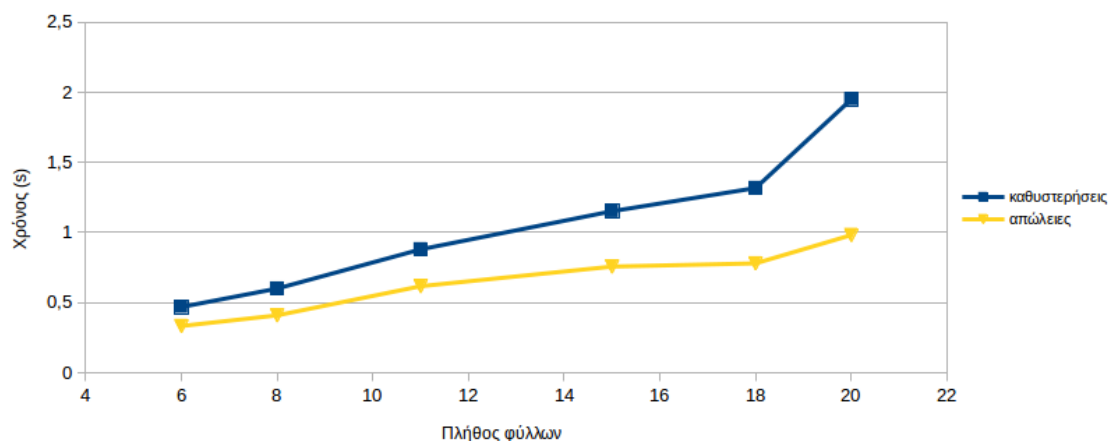
Είναι προφανές ότι υπάρχει (έως και πέντε τάξεις μεγέθους) μεγαλύτερη ακρίβεια στην

Πίνακας 5.1: Χαρακτηριστικά των τοπολογιών που κατασκευάστηκαν και σφάλματα στις εκτιμήσεις του Upsilon.

| Τοπολογία | Πλήθος κόμβων | Πλήθος φύλλων | Σφάλμα         |
|-----------|---------------|---------------|----------------|
| 1         | 10            | 5             | 0.3375ms       |
| 2         | 13            | 7             | 0.2400ms       |
| 3         | 19            | 10            | 0.3117ms       |
| 4         | 27            | 14            | 1.5004ms       |
| 5         | 34            | 17            | 0.6044ms       |
| 6         | 39            | 20            | 0.7408ms       |
| 7         | 47            | 24            | 0.5126ms       |
| 8         | 9             | 5             | $1.2228e - 05$ |
| 9         | 15            | 8             | $0.0004e - 01$ |
| 10        | 19            | 10            | $0.0005e - 02$ |
| 11        | 27            | 14            | $2.1326e - 05$ |
| 12        | 34            | 17            | 0.0007         |
| 13        | 37            | 19            | $1.0068e - 05$ |

εκτίμηση των απωλειών των ακμών, σε σύγκριση με την εκτίμηση των καθυστερήσεων. Το γεγονός αυτό οφείλεται στον εγγενή τρόπο λειτουργίας του Traffic Control των Linux (το οποίο χρησιμοποιείται για τον ορισμό του jitter των ακμών), το οποίο σε μία ακμή με καθυστέρηση  $D$  και ονομαστική τιμή για το jitter  $d$ , εισάγει καθυστέρηση με τιμή προερχόμενη από Ομοιόμορφη Κατανομή στο διάστημα  $[D - d, D + d]$ , με αποτέλεσμα η τιμή του πραγματικού jitter να διαφέρει από την ονομαστική τιμή αυτού και το δίκτυο που προκύπτει να αποτελεί μία κατά προσέγγιση υπερμετρική δενδρική δομή.

Αντίθετα, ο ορισμός των απωλειών των ακμών γίνεται με απόλυτη ακρίβεια (πριν την έναρξη του πειράματος, κατά την κατασκευή της τοπολογίας στο jFed), με αποτέλεσμα το δίκτυο που προκύπτει να πληρεί τα απαιτούμενα χαρακτηριστικά (βλέπε 4.1) και ο Upsilon να δίνει ακριβέστερα αποτελέσματα.



Εικόνα 5.1: Χρόνοι εκτέλεσης των scripts για την κατασκευή των Πινάκων Αποστάσεων, βάσει των καθυστερήσεων και των απωλειών των ακμών.

Επιπροσθέτως, με τη βοήθεια του διαγράμματος 5.1, παρατηρούμε ότι ο χρόνος εκτέλεσης του script για τον υπολογισμό του Πίνακα Αποστάσεων βάσει των απωλειών των ακμών είναι μικρότερος από αυτόν του script για τον υπολογισμό του Πίνακα Αποστάσεων βάσει των καθυστερήσεων των ακμών, για τοπολογίες με ίδιο αριθμό τελικών κόμβων. Αυτό οφείλεται στη μεγαλύτερη υπολογιστική πολυπλοκότητα του δεύτερου script, όπου πέραν της εύρεσης της τομής των διαφόρων συνόλων (δηλαδή των πακέτων που έφθασαν και στους δύο κόμβους ενός ζεύγους τελικών κόμβων) πραγματοποιείται και ο υπολογισμός των συνδιακυμάνσεων των καθυστερήσεων που υπέστησαν τα πακέτα έως τους κόμβους αυτούς.

Τέλος, ο χρόνος εκτέλεσης του UPGMA αλγορίθμου, για τους διάφορους Πίνακες Αποστάσεων, είναι σχετικά σταθερός, με μέση τιμή  $\bar{x} = 0.7265sec$  και τυπική απόκλιση  $sd = \sqrt{(x - \bar{x})^2} = 0.0567sec$ .

Οι διάφοροι κώδικες εκτελέστηκαν σε λειτουργικό σύστημα UBUNTU 16.04 xenial, με CPU Intel Core i7-4510U στα 3.1GHz και μνήμη RAM 8GB.



## Κεφάλαιο 6

---

### Επίλογος

---

Στην παρούσα διπλωματική μελετήθηκαν διάφορες τεχνικές Τομογραφίας Δικτύου, καθώς και ορισμένες προσεγγίσεις του προβλήματος της Ουδετερότητας Δικτύου. Επιπλέον, παρουσιάστηκαν διάφορες τεχνικές και εργαλεία για την εκτίμηση μετρικών που περιγράφουν την κατάσταση των ακμών ενός δικτύου (π.χ. διαθέσιμο εύρος ζώνης, χωρητικότητα).

Εν συνεχεία, υλοποιήθηκε μία τομογραφική τεχνική που παρέχει τη δυνατότητα ανακατασκευής υπερμετρικών τοπολογιών, καθώς και συναγωγής των χαρακτηριστικών των ακμών που τις απαρτίζουν, με γνώση μόνο ακραίων μετρήσεων. Η τεχνική αυτή περιλαμβάνει την υλοποίηση του συσσωρευτικού αλγορίθμου *Urgma* και την αντιστοίχιση των μηκών των ακμών στα χαρακτηριστικά τους (απώλειες και καθυστερήσεις), μέσω της κατασκευής προσθετικών μετρικών.

Συμπερασματικά, προκύπτει ότι ο αλγόριθμος που υλοποιήθηκε επιτυγχάνει την πλήρη ανακατασκευή και των δεκατριών (13) τοπολογιών που κατασκευάστηκαν στο σύστημα δοκιμών του jFed, καθώς επίσης προσφέρει πολύ ικανοποιητικές εκτιμήσεις των εσωτερικών χαρακτηριστικών των τοπολογιών που ανακατασκευάζει.

Πιο συγκεκριμένα, υπολογίζει τις απώλειες των ακμών με σφάλμα 0.0133% (κατά μέσο όρο για όλες τις ακμές και των έξι (6) τοπολογιών όπου τα μήκη των ακμών αντιστοιχούν στις απώλειές τους) και το jitter των ακμών με σφάλμα 0.6068ms (κατά μέσο όρο για όλες

τις ακμές και των επτά (7) τοπολογιών όπου τα μήκη των ακμών αντιστοιχούν στο jitter τους).

Τεχνικές όπως αυτή που υλοποιήθηκε στην παρούσα διπλωματική, ανοίγουν το δρόμο για την ανάπτυξη του πεδίου της Τομογραφίας Δικτύου, ενός πεδίου που παρουσιάζει τεράστια άνθηση τις τελευταίες δεκαετίες, χάρη στην πολλαπλή συνεισφορά του σε ένα Διαδίκτυο καλπάζοντος μεγέθους και σημασίας.

Λαμβάνοντας υπόψιν ότι η ύπαρξη κεντρικών σημείων παρακολούθησης στο Διαδίκτυο είναι αδύνατη και ότι η γνώση της εσωτερικής δομής και των χαρακτηριστικών των επιμέρους δικτύων που το απαρτίζουν θα απαιτούσε τη συνεργασία των εσωτερικών κόμβων, παράγοντας επιπλέον φόρτο, επιβαρύνοντας οικονομικά τους διαχειριστές τους και διακινδυνεύοντας την ασφάλεια και την ιδιωτικότητα τους, η Τομογραφία Δικτύου προσφέρει εναλλακτικές λύσεις για την επίτευξη των παραπάνω, με χρήση μόνον ακραίων μετρήσεων και χωρίς την ανάγκη εγκατάστασης επιπλέον υλικού ή λογισμικού.

Ειδικότερα, προσφέρει τη δυνατότητα δυναμικής δρομολόγησης μέσω της εύρεσης βέλτιστων μονοπατιών δρομολόγησης διαμέσου δικτύων, των οποίων η εσωτερική δομή δεν είναι γνωστή, γεγονός που επιτρέπει στους χρήστες να απολαμβάνουν καλύτερη ποιότητα υπηρεσιών (μικρότερο latency, λιγότερες απώλειες). Επιπροσθέτως, επιτρέπει τη διασφάλιση των διαφόρων συμφωνιών/κανονισμών που διέπουν τις σχέσεις όλων των εμπλεκόμενων μελών του Διαδικτύου, καθώς επίσης, προσφέρει καλύτερη επίβλεψη των δικτύων που απαρτίζουν το Διαδίκτυο, διευκολύνοντας την ανίχνευση και την αντιμετώπιση πιθανών ανωμαλιών ή/και κακόβουλων κινήσεων. Τέλος, η συναγωγή των εσωτερικών χαρακτηριστικών ενός δικτύου επιτρέπει τη μελέτη της συμπεριφοράς του δικτύου αυτού απέναντι σε διαφορετικά είδη κινήσεων δεδομένων, γεγονός που μπορεί να αποκαλύψει παραβιάσεις της Ουδετερότητας, προσφέροντας, έτσι, ένα ακόμη εργαλείο για την εξασφάλιση ενός δίκαιου Διαδικτύου, με σαφώς καθορισμένους στόχους και κανονισμούς.

## Κεφάλαιο 7

---

### Ανοιχτά προβλήματα και συναφής βιβλιογραφία

---

Στην υποενότητα 2.3.2 παρουσιάστηκαν τρία χαρακτηριστικά των ακμών, οι απώλειες, οι καθυστερήσεις (για την ακρίβεια η διακύμανση των καθυστερήσεων) και η χρησιμοποίηση, ως τα χαρακτηριστικά που μπορούν, μετά από κατάλληλο μετασχηματισμό, να αντιστοιχιστούν στις αποστάσεις των ακμών (εφόσον διαθέτουν προσθετικές ιδιότητες). Το ερώτημα που τίθεται και αποτελεί ένα από τα σημαντικότερα ανοιχτά ζητήματα της Τομογραφίας Δικτύου, γύρω από το οποίο πραγματοποιούνται πολλές ερευνητικές εργασίες τα τελευταία χρόνια, είναι το κατά πόσο το διαθέσιμο εύρος ζώνης μπορεί να μετασχηματιστεί κατάλληλα ώστε να αποτελέσει δενδρική μετρική.

Στο [49] υποστηρίζεται ότι, υπό ορισμένες συνθήκες, το διαθέσιμο εύρος ζώνης σχηματίζει υπερμετρικούς χώρους, γεγονός που επιτρέπει την κατασκευή μοντέλου πρόβλεψης, μέσω της εισαγωγής σημείων αναφοράς *landmarks*, για την πραγματοποίηση των μετρήσεων.

Στο [50] προτείνεται μέθοδος για την πρόβλεψη της κατανομής που ακολουθεί το διαθέσιμο εύρος ζώνης, καθώς και τα υπόλοιπα χαρακτηριστικά των ακμών, μέσω της κατασκευής συναρτήσεων μέγιστης πιθανοφάνειας, βάσει της ιδιότητας του διαθέσιμου εύρους ζώνης να καθορίζεται από τη μικρότερη τιμή που αυτό παίρνει κατά μήκος ενός μονοπατιού.

Στο [51] προτείνεται μία διαφορετική προσέγγιση του προβλήματος. Ορίζεται το *Proba-*

*bilistic Available Bandwidth (PAB)* ως ο μέγιστος ρυθμός με τον οποίο μπορεί να εισάγεται ροή κίνησης σε κάποιο μονοπάτι του δικτύου ώστε να επιτυγχάνεται, με δεδομένη πιθανότητα, ρυθμός εξόδου σχεδόν ίσος με το ρυθμό εισόδου. Το πλαίσιο που προτείνεται για την εκτίμηση του *PAB* βασίζεται σε packet trains, Bayesian Inference, γράφους παραγόντων και ενεργή δειγματοληψία.

Το [52] βασίζεται στο [53] (Sequoia system), για να προτείνει μία ακριβέστερη μέθοδο κατασκευής δένδρου πρόβλεψης (prediction tree) για την ενσωμάτωση πληροφορίας σχετικής με το εύρος ζώνης. Ο αλγόριθμος που προτείνεται για την κατασκευή του δένδρου πρόβλεψης βασίζεται στην εισαγωγή κόμβων ώστε να μεγιστοποιείται το *Gromov product* (ή σε παραλλαγή του αλγορίθμου, να ελαχιστοποιείται το σφάλμα πρόβλεψης), ακολουθώντας τη δομή ενός anchor tree, για αύξηση της επίδοσης και της κλιμακωσιμότητας, και πραγματοποιώντας τον ακόλουθο γραμμικό μετασχηματισμό στο διαθέσιμο εύρος ζώνης:  $BW_T(u, v) = C - d_T(u, v)$ , όπου  $C$  μεγάλη θετική σταθερά και  $d_T(u, v)$  η απόσταση μεταξύ των  $u, v$  στο δένδρο πρόβλεψης, όπως υπολογίστηκε έπειτα από μεγιστοποίηση του Gromov product.

Τέλος, το [54] βασίζεται στο last-mile model και σε ευριστικές μεθόδους για την πρόβλεψη του διαθέσιμου εύρους ζώνης. Σύμφωνα με το last-mile model, το διαθέσιμο εύρος ζώνης στα άκρα του δικτύου (network edge), δηλαδή στις ακμές που καταλήγουν στους τελικούς κόμβους (last mile ακμές), είναι αυτό που καθορίζει το διαθέσιμο εύρος ζώνης ολόκληρου του μονοπατιού. Διότι, σύμφωνα με το [55], στο 60% των μονοπατιών που συνδέουν τους τελικούς χρήστες μέσω Δικτύων Ευρείας Ζώνης (WAN), το μικρότερο (διαθέσιμο) εύρος ζώνης (bottleneck bandwidth / tight link) συναντάται στο πρώτο ή στο δεύτερο βήμα από τον τελικό κόμβο.

Ένα επιπλέον ζήτημα, που μπορεί να αποτελέσει άξονα μελλοντικής εργασίας, είναι η πραγματοποίηση των μετρήσεων μέσω unicast αποστολής πακέτων (αντί για multicast), δεδομένου ότι τα προβλήματα κλιμακωσιμότητας που εισάγει η multicast αποστολή έχουν καταστήσει την unicast τον κυρίαρχο τρόπο αποστολής πακέτων, τόσο σε Τοπικά Δίκτυα, όσο και στο Διαδίκτυο [56].

Φυσικά, υπάρχουν πολλά ακόμη ανοιχτά ζητήματα στο πεδίο της *Τομογραφίας και της Ου-*



---

δετερότητας Δικτύου, τα οποία σχετίζονται με τη βελτιστοποίηση της επίδοσης, την επίτευξη της κλιμακωσιμότητας και την αποκεντρωμένη λειτουργία των διαφόρων τεχνικών, ώστε αυτές να μπορούν να εφαρμοστούν σε ένα διαρκώς κλιμακούμενο Διαδίκτυο, σε πραγματικό χρόνο, χωρίς να διακινδυνεύει η ασφάλεια και η ιδιωτικότητα των διαφόρων εμπλεκομένων (πάροχοι, τελικοί χρήστες) και χωρίς να επιφέρεται ασύμφορα μεγάλη αύξηση του κόστους (υλοποίησης/παροχής υπηρεσιών/λήψης υπηρεσιών).



---

## Βιβλιογραφία

---

- [1] <https://repository.kallipos.gr/bitstream/11419/6002/1/08.pdf>.
- [2] [https://mediasrv.uaa.gr/eclass/modules/document/file.php/BIOTECH133/6\\_phylogenetic\\_trees\\_2019.pptx](https://mediasrv.uaa.gr/eclass/modules/document/file.php/BIOTECH133/6_phylogenetic_trees_2019.pptx).
- [3] D. Ghita, K. Argyraki, and P. Thiran, “Toward Accurate and Practical Network Tomography”, *SIGOPS Oper. Syst. Rev.*, vol. 47, no. 1, pp. 22–26, Jan. 2013, ISSN: 0163-5980. DOI: 10.1145/2433140.2433146. [Online]. Available: <http://doi.acm.org/10.1145/2433140.2433146>.
- [4] —, “Network Tomography on Correlated Links”, in *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, ser. IMC ’10, Melbourne, Australia: ACM, 2010, pp. 225–238, ISBN: 978-1-4503-0483-2. DOI: 10.1145/1879141.1879170. [Online]. Available: <http://doi.acm.org/10.1145/1879141.1879170>.
- [5] D. Ghita, C. Karakus, K. Argyraki, and P. Thiran, “Shifting Network Tomography Toward a Practical Goal”, in *Proceedings of the Seventh Conference on Emerging Networking Experiments and Technologies*, ser. CoNEXT ’11, Tokyo, Japan: ACM, 2011, 24:1–24:12, ISBN: 978-1-4503-1041-3. DOI: 10.1145/2079296.2079320. [Online]. Available: <http://doi.acm.org/10.1145/2079296.2079320>.
- [6] M. Coates, R. Castro, R. Nowak, M. Gadhiok, R. King, and Y. Tsang, “Maximum Likelihood Network Topology Identification from Edge-based Unicast Measurements”, *SIGMETRICS Perform. Eval. Rev.*, vol. 30, no. 1, pp. 11–20, Jun.

- 2002, ISSN: 0163-5999. DOI: 10.1145/511399.511337. [Online]. Available: <http://doi.acm.org/10.1145/511399.511337>.
- [7] R. Castro, M. Coates, G. Liang, R. Nowak, and B. Yu, “Network Tomography: Recent Developments”, *Statist. Sci.*, vol. 19, no. 3, pp. 499–517, Aug. 2004. DOI: 10.1214/088342304000000422. [Online]. Available: <https://doi.org/10.1214/088342304000000422>.
- [8] B. Eriksson, G. Dasarathy, P. Barford, and R. Nowak, “Efficient Network Tomography for Internet Topology Discovery”, *IEEE/ACM Transactions on Networking*, vol. 20, no. 3, pp. 931–943, Jun. 2012, ISSN: 1063-6692. DOI: 10.1109/TNET.2011.2175747.
- [9] J. Ni, H. Xie, S. Tatikonda, and Y. R. Yang, “Efficient and Dynamic Routing Topology Inference From End-to-End Measurements”, *IEEE/ACM Transactions on Networking*, vol. 18, no. 1, pp. 123–135, Feb. 2010, ISSN: 1063-6692. DOI: 10.1109/TNET.2009.2022538.
- [10] J. Ni and S. Tatikonda, “A Markov Random Field Approach to Multicast-Based Network Inference Problems”, Aug. 2006, pp. 2769–2773. DOI: 10.1109/ISIT.2006.261566.
- [11] Y. Chen, D. Bindel, H. Song, and R. H. Katz, “An Algebraic Approach to Practical and Scalable Overlay Network Monitoring”, *SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 4, pp. 55–66, Aug. 2004, ISSN: 0146-4833. DOI: 10.1145/1030194.1015475. [Online]. Available: <http://doi.acm.org/10.1145/1030194.1015475>.
- [12] V. Karyotis, E. Stai, and S. Papavassiliou, *Evolutionary Dynamics of Complex Communications Networks*. Boca Raton, FL, USA: CRC Press, Inc., 2013, ISBN: 1466518405, 9781466518407.
- [13] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks”, *Reviews of Modern Physics*, vol. 74, no. 1, pp. 47–97, Jan. 2002, ISSN: 1539-0756. DOI: 10.1103/revmodphys.74.47. [Online]. Available: <http://dx.doi.org/10.1103/RevModPhys.74.47>.

- [14] C. Dovrolis, R. Prasad, M. Murray, and k. claffy k., “Bandwidth estimation: metrics, measurement techniques, and tools”, *IEEE Network*, vol. 17, no. 6, pp. 27–35, Apr. 2003.
- [15] <https://www.comparitech.com/net-admin/throughput-vs-bandwidth/>.
- [16] V. Ribeiro, R. Riedi, J. Navrátil, and L. Cottrell, “pathChirp: Efficient Available Bandwidth Estimation for Network Paths”, *Proceedings of Passive and Active Measurement Workshop*, Apr. 2003. DOI: 10.2172/813038.
- [17] Ningning Hu, Li Li, Z. M. Mao, P. Steenkiste, and Jia Wang, “A measurement study of Internet bottlenecks”, in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 3, Mar. 2005, 1689–1700 vol. 3. DOI: 10.1109/INFCOM.2005.1498450.
- [18] S. Banerjee and A. K. Agrawala, “Estimating available capacity of a network connection”, in *Proceedings IEEE International Conference on Networks 2000 (ICON 2000). Networking Trends and Challenges in the New Millennium*, Sep. 2000, pp. 131–138. DOI: 10.1109/ICON.2000.875780.
- [19] S. Chakravarty, A. Stavrou, and A. D Keromytis, “LinkWidth: A Method to Measure Link Capacity and Available Bandwidth using Single-End Probes”, Jan. 2008.
- [20] N. Hu and P. Steenkiste, “Estimating Available Bandwidth Using Packet Pair Probing”, Apr. 2003.
- [21] <https://www.caida.org/tools/taxonomy/perftaxonomy.xml>.
- [22] F. Abut and M. Leischner, “An Experimental Evaluation of Tools for Estimating Bandwidth-Related Metrics”, *International Journal of Computer Network and Information Security*, vol. 10, pp. 1–11, Jul. 2018. DOI: 10.5815/ijcnis.2018.07.01.
- [23] [http://staff.ii.pw.edu.pl/n2x/papers/Router\\_Instability.pdf](http://staff.ii.pw.edu.pl/n2x/papers/Router_Instability.pdf).
- [24] <http://www.spin.rice.edu/Software/pathChirp/>.
- [25] <https://www.cs.cmu.edu/~hnn/igi/>.

- [26] J. Strauss, D. Katabi, F. Kaashoek, and F. Kaashoek, “A Measurement Study of Available Bandwidth Estimation Tools”, in *Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement*, ser. IMC '03, Miami Beach, FL, USA: ACM, 2003, pp. 39–44, ISBN: 1-58113-773-7. DOI: 10.1145/948205.948211. [Online]. Available: <http://doi.acm.org/10.1145/948205.948211>.
- [27] G. Urvoy-Keller, T. En-Najjary, and A. Sorniotti, “Operational comparison of available bandwidth estimation tools”, *Computer Communication Review*, vol. 38, pp. 39–42, Jan. 2008. DOI: 10.1145/1341431.1341438.
- [28] A. Shriram, M. Murray, Y. Hyun, N. Brownlee, A. Broido, M. Fomenkov, and k. claffy k., “Comparison of Public End-to-End Bandwidth Estimation Tools on High-Speed Links”, in *Passive and Active Network Measurement Workshop (PAM)*, vol. 3431, Boston, MA: PAM 2005, Mar. 2005, pp. 306–320.
- [29] A. Shriram and J. Kaur, “Empirical Evaluation of Techniques for Measuring Available Bandwidth”, in *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*, May 2007, pp. 2162–2170. DOI: 10.1109/INFCOM.2007.250.
- [30] F. Montesino-Pouzols, “Comparative Analysis of Active Bandwidth Estimation Tools”, in *Passive and Active Network Measurement*, C. Barakat and I. Pratt, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 175–184, ISBN: 978-3-540-24668-8.
- [31] E. Altman, J. Rojas-Mora, S. Wong Ramírez, M. Hanawal, and Y. Xu, “Net Neutrality and Quality of Service”, vol. abs/1105.0283, Apr. 2011. DOI: 10.1007/978-3-642-30373-9\_10.
- [32] Z. Jukic, M. Weber, V. Svedek, M. Vukovic, D. Katusic, and G. Jezic, “Technical aspects of network neutrality”, in *Proceedings of the 11th International Conference on Telecommunications*, Jun. 2011, pp. 405–410.

- [33] E. Altman, P. Bernhard, S. Caron, G. Kesidis, J. Rojas-Mora, and S. Wong, “A model of network neutrality with usage-based prices”, *Telecommunication Systems*, vol. 52, no. 2, pp. 601–609, Feb. 2013, ISSN: 1572-9451. DOI: 10.1007/s11235-011-9504-6. [Online]. Available: <https://doi.org/10.1007/s11235-011-9504-6>.
- [34] [https://en.wikipedia.org/wiki/Nash\\_equilibrium](https://en.wikipedia.org/wiki/Nash_equilibrium).
- [35] R. Ma, D. M. Chiu, J. C. Lui, V. Misra, and D. Rubenstein, “Interconnecting eyeballs to content: a shapley value perspective on isp peering and settlement.”, Jan. 2008, pp. 61–66. DOI: 10.1145/1403027.1403041.
- [36] [http://dione.lib.unipi.gr/xmlui/bitstream/handle/unipi/8885/Kallidonis\\_Ioannis.pdf?sequence=1&isAllowed=y](http://dione.lib.unipi.gr/xmlui/bitstream/handle/unipi/8885/Kallidonis_Ioannis.pdf?sequence=1&isAllowed=y).
- [37] [http://dione.lib.unipi.gr/xmlui/bitstream/handle/unipi/9481/Kliafas\\_Konstantinos.pdf?sequence=1&isAllowed=y](http://dione.lib.unipi.gr/xmlui/bitstream/handle/unipi/9481/Kliafas_Konstantinos.pdf?sequence=1&isAllowed=y).
- [38] [https://en.wikipedia.org/wiki/Cooperative\\_game\\_theory](https://en.wikipedia.org/wiki/Cooperative_game_theory).
- [39] Z. Zhang, O. Mara, and K. Argyraki, “Network Neutrality Inference”, *SIGCOMM Comput. Commun. Rev.*, vol. 44, no. 4, pp. 63–74, Aug. 2014, ISSN: 0146-4833. DOI: 10.1145/2740070.2626308. [Online]. Available: <http://doi.acm.org/10.1145/2740070.2626308>.
- [40] C. Pappas, K. Argyraki, S. Bechtold, and A. Perrig, “Transparency Instead of Neutrality”, in *Proceedings of the 14th ACM Workshop on Hot Topics in Networks*, ser. HotNets-XIV, Philadelphia, PA, USA: ACM, 2015, 22:1–22:7, ISBN: 978-1-4503-4047-2. DOI: 10.1145/2834050.2834082. [Online]. Available: <http://doi.acm.org/10.1145/2834050.2834082>.
- [41] [http://repfiles.kallipos.gr/html\\_books/72/kef8.html](http://repfiles.kallipos.gr/html_books/72/kef8.html).
- [42] <http://www.cs.cmu.edu/~durand/03-711/2011/Lectures/Trees11-4.pdf>.
- [43] <http://people.cs.uchicago.edu/~ridg/digbio08/talkaddree.pdf>.

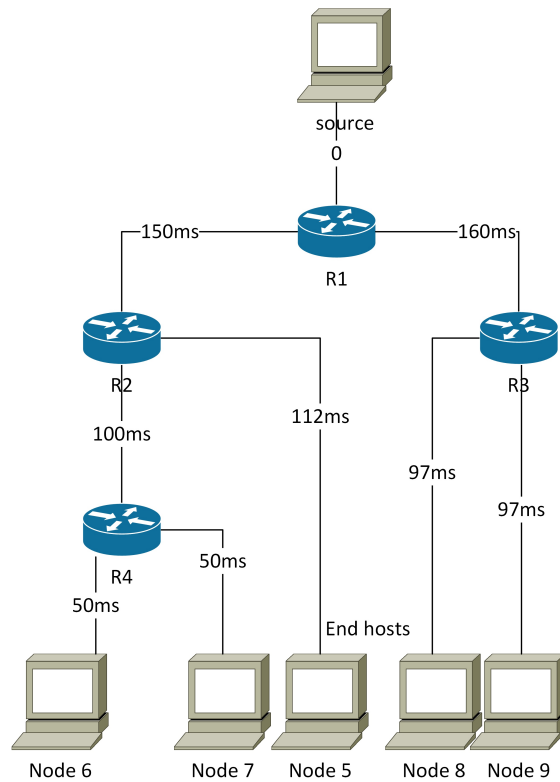
- [44] P. Buneman, “A Note on the Metric Properties of Trees”, English, *Journal of Combinatorial Theory, Series B*, vol. 17, no. 1, pp. 48–50, 1974, ISSN: 0095-8956. DOI: 10.1016/0095-8956(74)90047-1.
- [45] H. Fushing, H. Wang, K. VanderWaal, B. McCowan, and P. Koehl, “Multi-Scale Clustering by Building a Robust and Self Correcting Ultrametric Topology on Data Points”, *PloS one*, vol. 8, e56259, Feb. 2013. DOI: 10.1371/journal.pone.0056259.
- [46] [https://en.wikipedia.org/wiki/Unrooted\\_binary\\_tree](https://en.wikipedia.org/wiki/Unrooted_binary_tree).
- [47] <https://www.nature.com/scitable/topicpage/the-molecular-clock-and-estimating-species-divergence-41971/>.
- [48] <https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/upgma>.
- [49] C. Xing, M. Chen, and L. Yang, “Predicting Available Bandwidth of Internet Path with Ultra Metric Space-Based Approaches”, in *GLOBECOM 2009 - 2009 IEEE Global Telecommunications Conference*, Nov. 2009, pp. 1–6. DOI: 10.1109/GLOCOM.2009.5426036.
- [50] P. A. Chou, C. Rodbro, and U. Dogan, “Prediction of Bandwidth and Additive Metrics for Large Scale Network Tomography”, Tech. Rep. MSR-TR-2016-57, May 2016. [Online]. Available: <https://www.microsoft.com/en-us/research/publication/prediction-bandwidth-additive-metrics-large-scale-network-tomography/>.
- [51] F. Thouin, M. Coates, and M. Rabbat, “Large scale probabilistic available bandwidth estimation”, *Computer Networks*, vol. 55, pp. 2065–2078, Jun. 2011. DOI: 10.1016/j.comnet.2011.02.011.
- [52] S. Song, P. Keleher, B. Bhattacharjee, and A. Sussman, “Decentralized, accurate, and low-cost network bandwidth prediction”, in *2011 Proceedings IEEE INFOCOM*, Apr. 2011, pp. 6–10. DOI: 10.1109/INFCOM.2011.5935251.



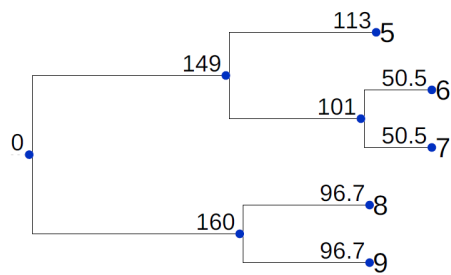
- [53] V. Ramasubramanian, D. Malkhi, F. Kuhn, M. Balakrishnan, A. Gupta, and A. Akella, “On the Treeness of Internet Latency and Bandwidth”, *SIGMETRICS Perform. Eval. Rev.*, vol. 37, no. 1, pp. 61–72, Jun. 2009, ISSN: 0163-5999. DOI: 10.1145/2492101.1555357. [Online]. Available: <http://doi.acm.org/10.1145/2492101.1555357>.
- [54] O. Beaumont, L. Eyraud-Dubois, and Y. Won, “Using the Last-Mile Model as a Distributed Scheme for Available Bandwidth Prediction”, Sep. 2011, pp. 103–116. DOI: 10.1007/978-3-642-23400-2\_11.
- [55] Ningning Hu, Li Li, Z. M. Mao, P. Steenkiste, and Jia Wang, “A measurement study of Internet bottlenecks”, in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 3, Mar. 2005, 1689–1700 vol. 3. DOI: 10.1109/INFCOM.2005.1498450.
- [56] <https://erg.abdn.ac.uk/users/gorry/course/intro-pages/uni-b-mcast.html>.
- [57] <https://ermongroup.github.io/cs228-notes/representation/undirected/>.
- [58] V. Karyotis, “A Markov Random Field Framework for Modeling Malware Propagation in Complex Communications Networks”, *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 4, pp. 551–564, Jul. 2019. DOI: 10.1109/TDSC.2017.2703622.
- [59] W. Xi and J. Baras, “Gibbs Sampler Based Control of Autonomous Vehicle Swarms in the Presence of Sensor Errors”, Jan. 2007, pp. 5084–5090. DOI: 10.1109/CDC.2006.376697.

## Παράρτημα

## Α Τοπολογίες

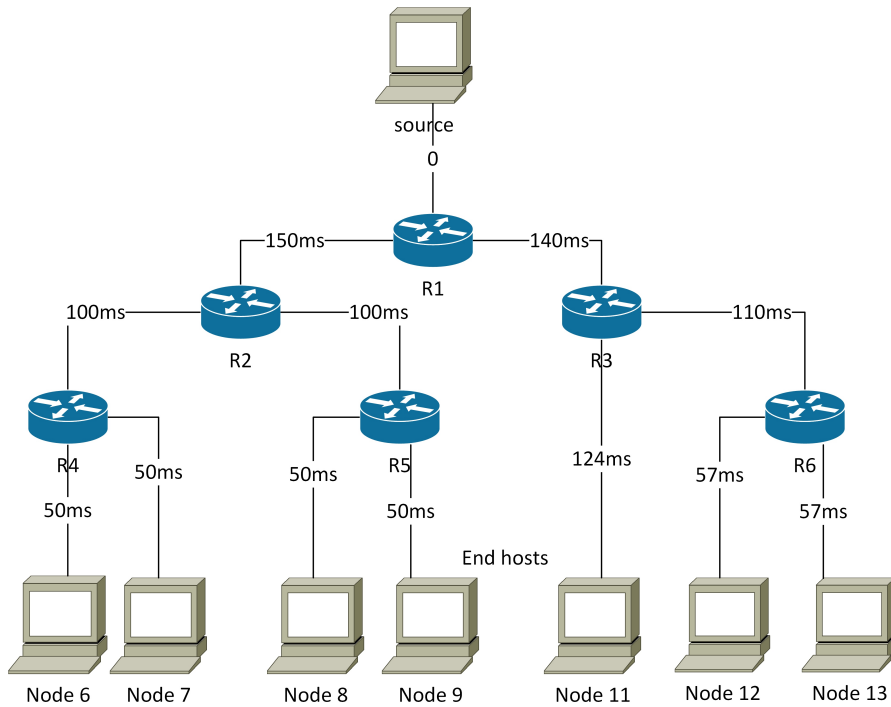


(α') jFed

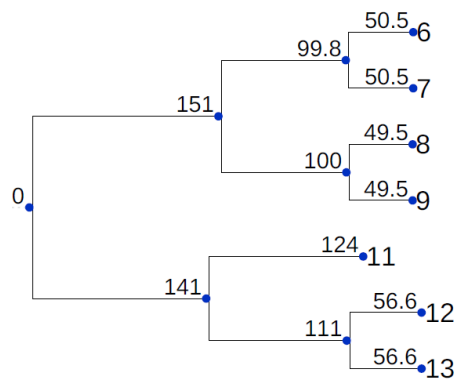


(β') Upsilon

Εικόνα 1: Τοπολογία 1 - ακμές με καθυστερήσεις (επιστροφή στον 5.1).

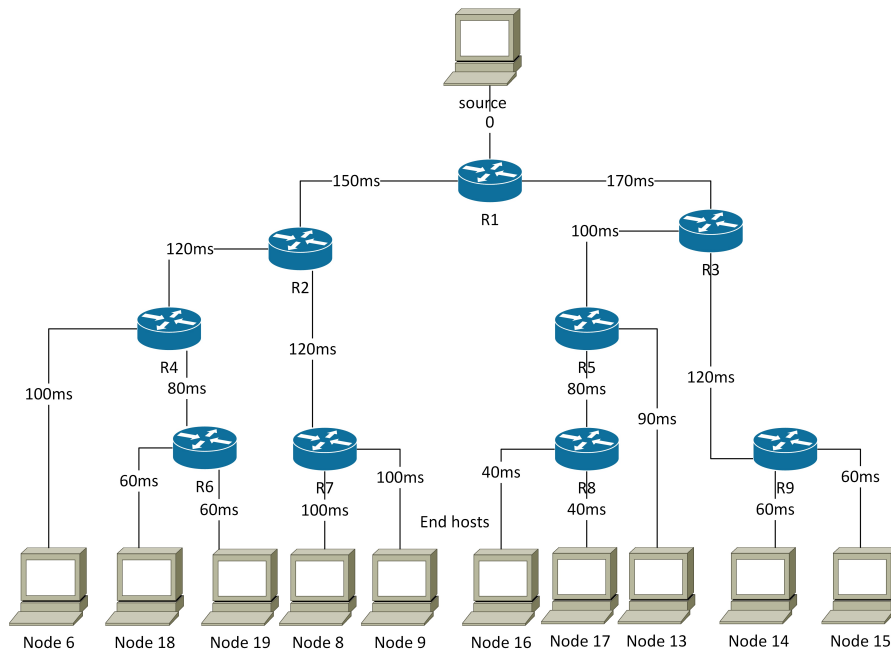


( $\alpha'$ ) jFed

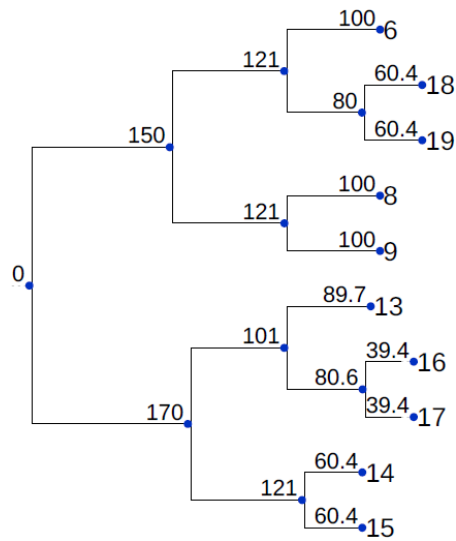


( $\beta'$ ) Upsilon

Εικόνα 2: Τοπολογία 2 - ακμές με καθυστερήσεις (επιστροφή στον 5.1).

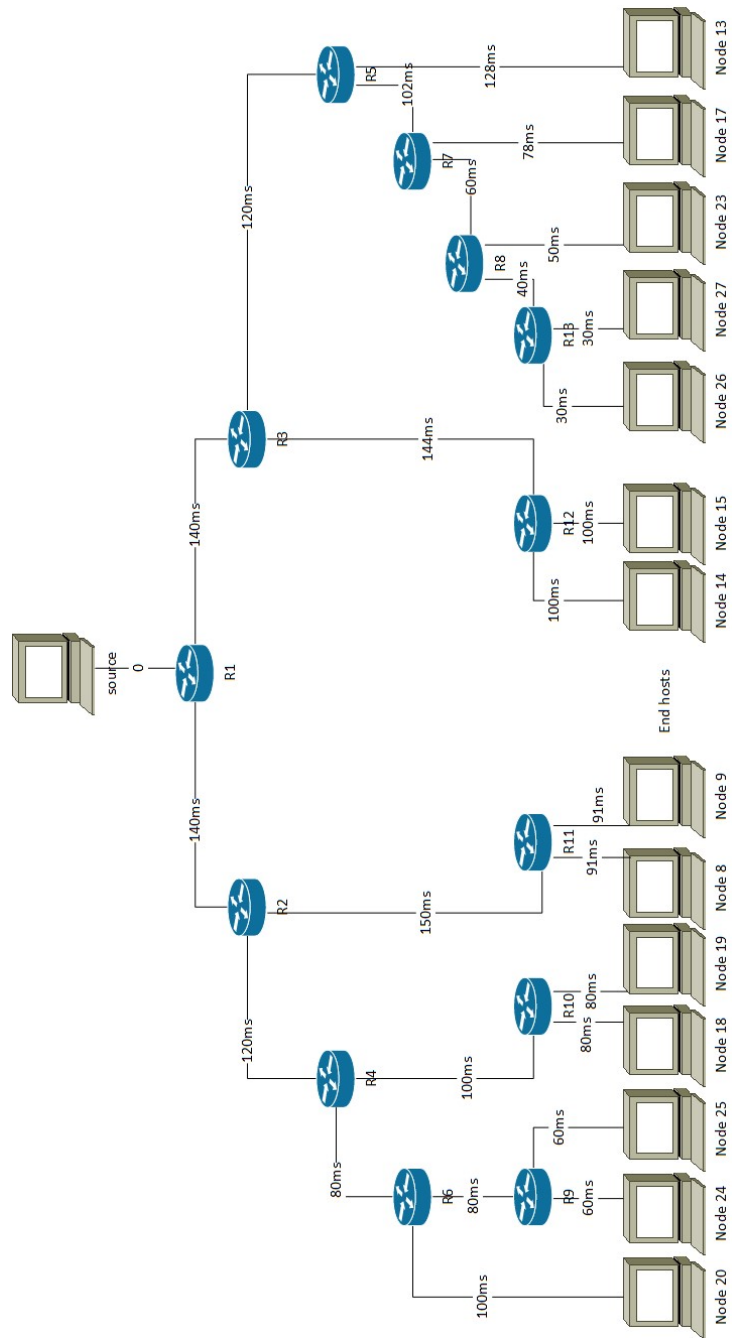


( $\alpha'$ ) jFed

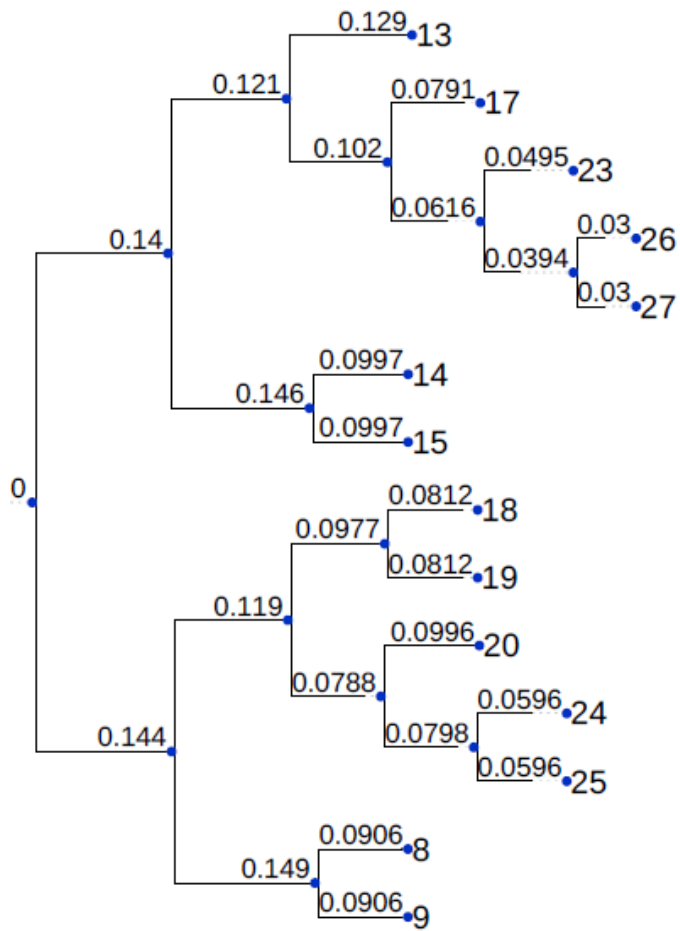


( $\beta'$ ) Upsilon

Εικόνα 3: Τοπολογία 3 - ακμές με καθυστερήσεις (επιστροφή στον 5.1).

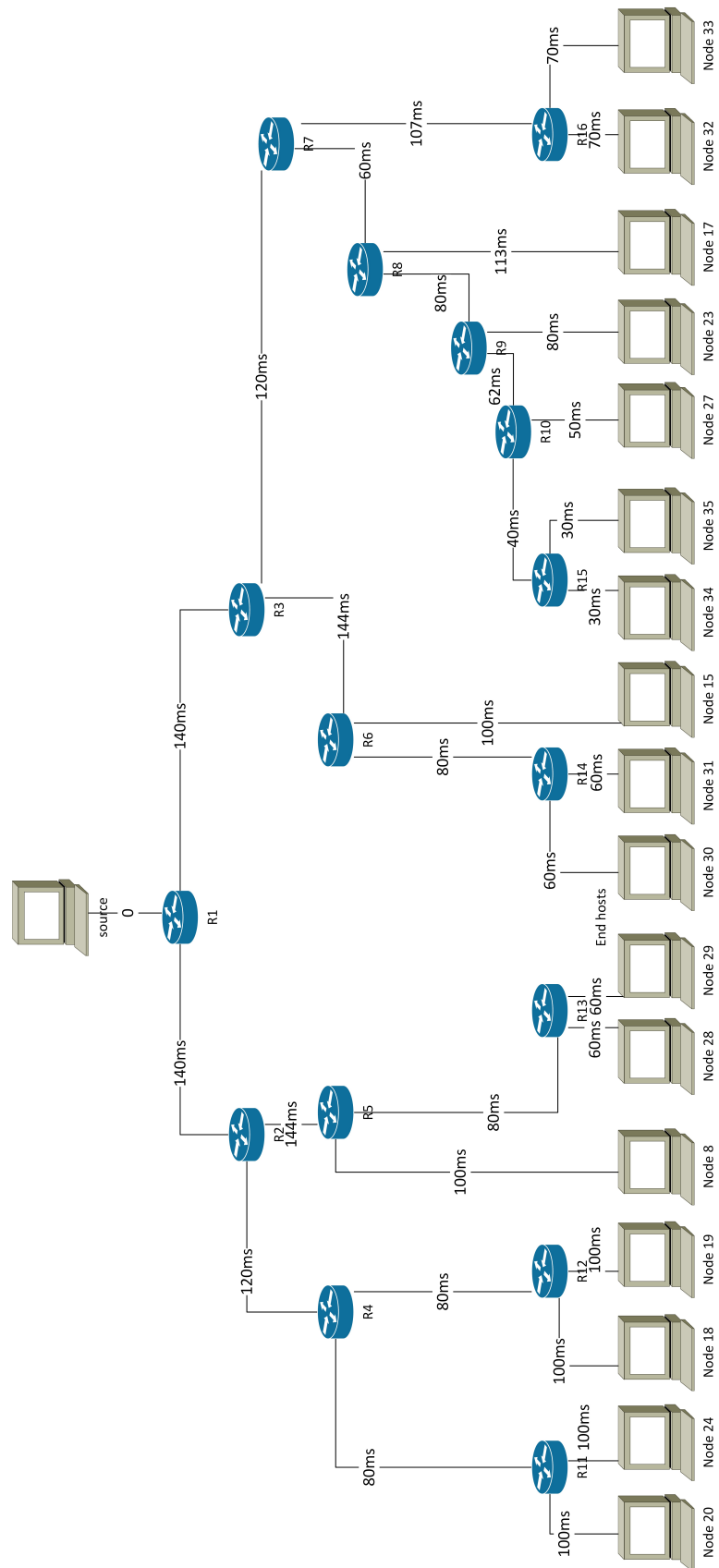


(α) jFed

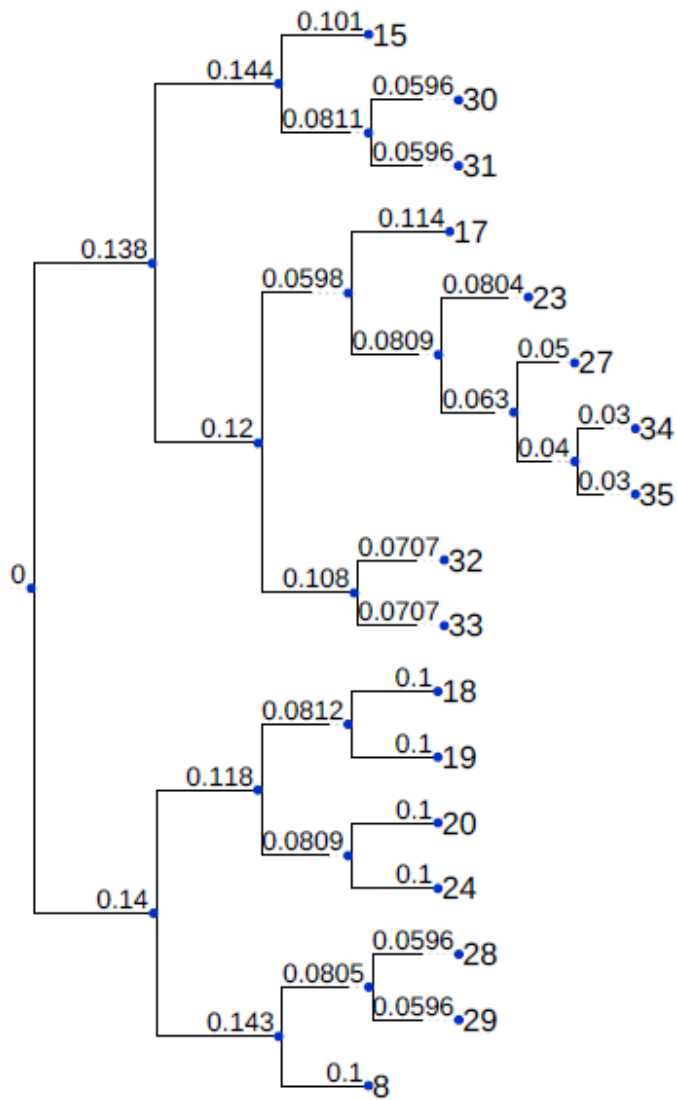


(β') Υψγμα

Εικόνα 4: Τοπολογία 4 - ακμές με καθυστερήσεις (επιστροφή στον 5.1).



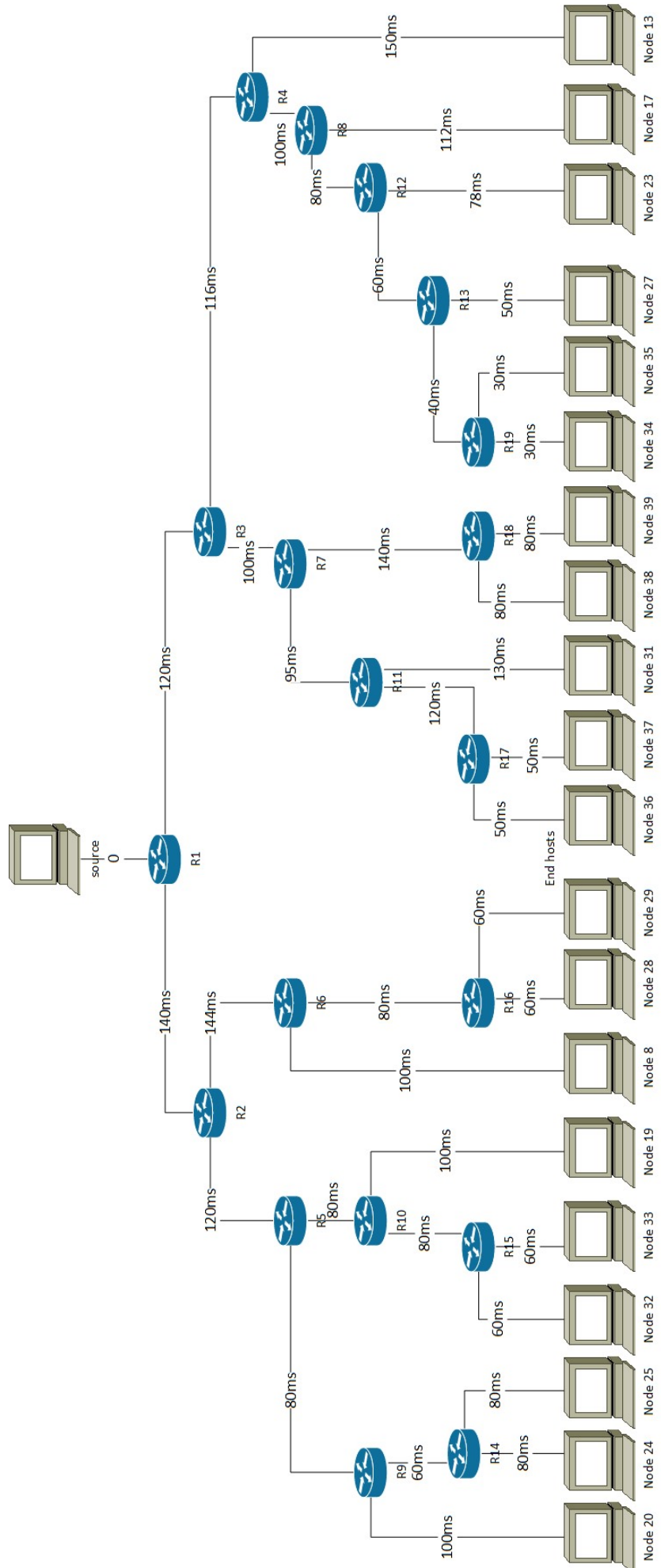
( $\alpha$ ) jFed

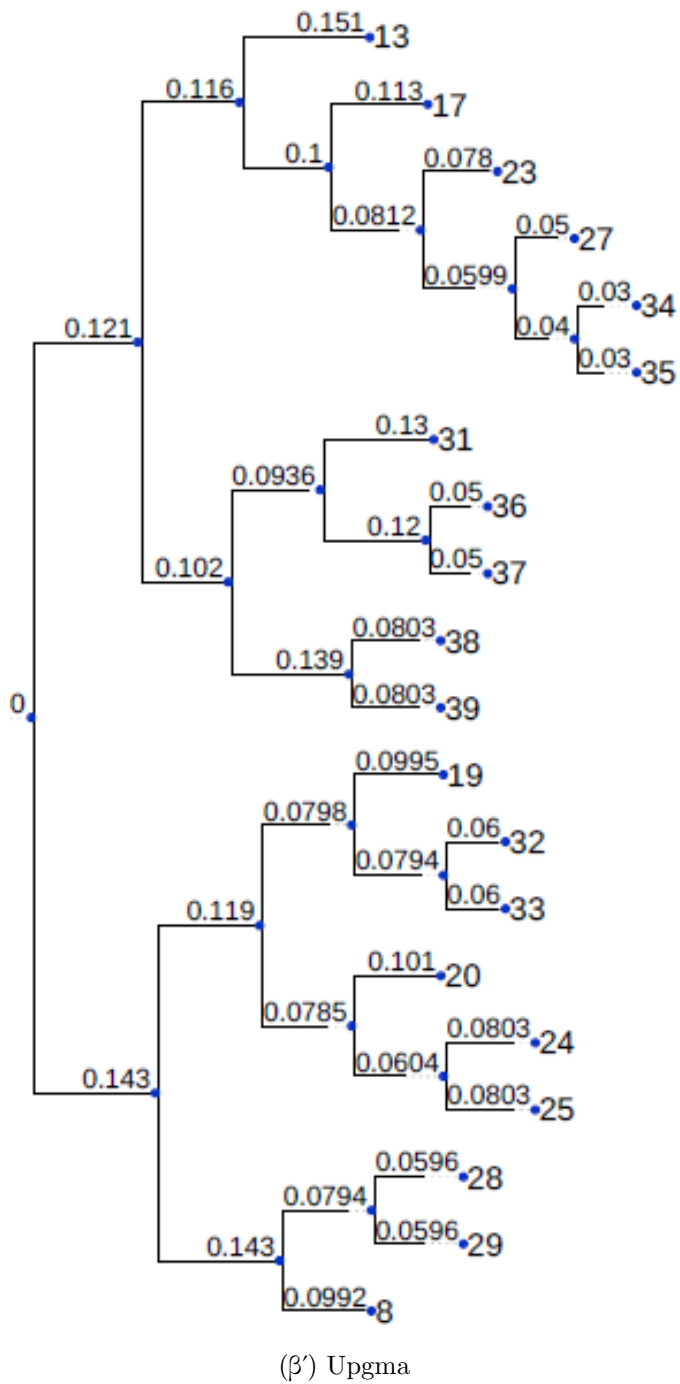


(β') Upsilon

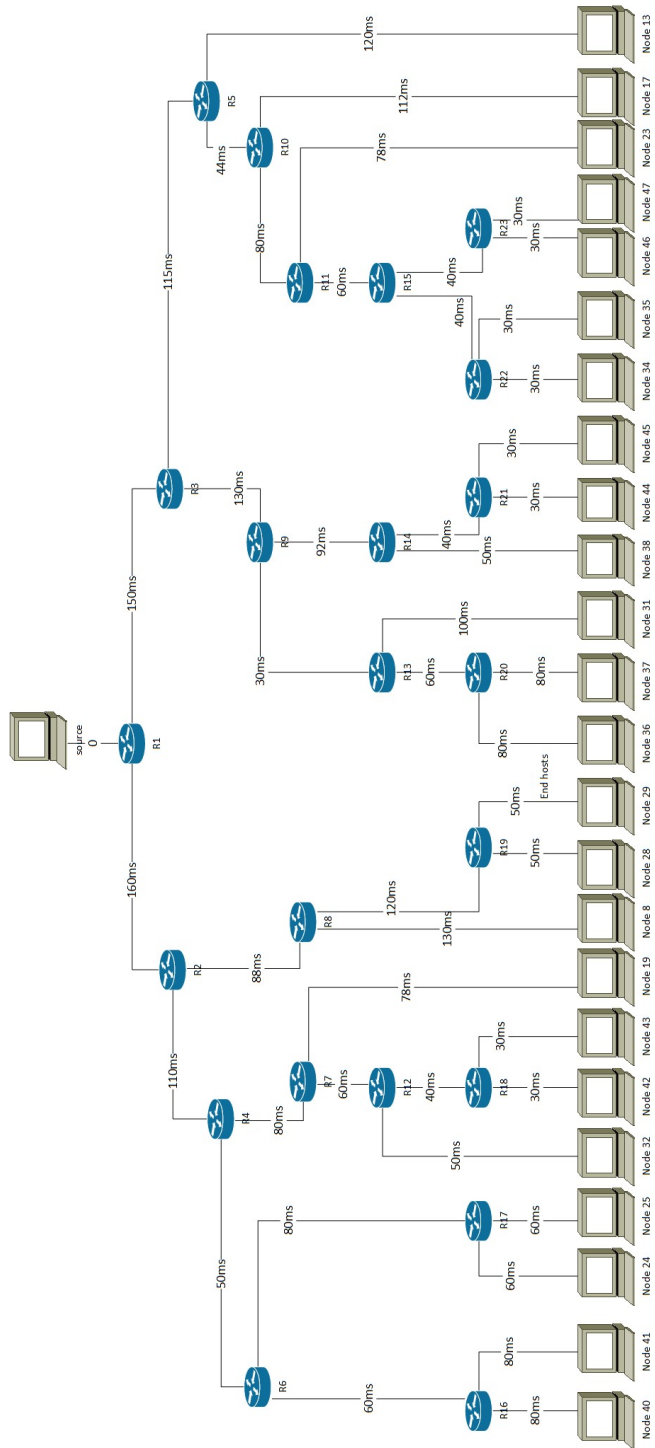
Εικόνα 5: Τοπολογία 5 - ακμές με καθυστερήσεις (επιστροφή στον 5.1).



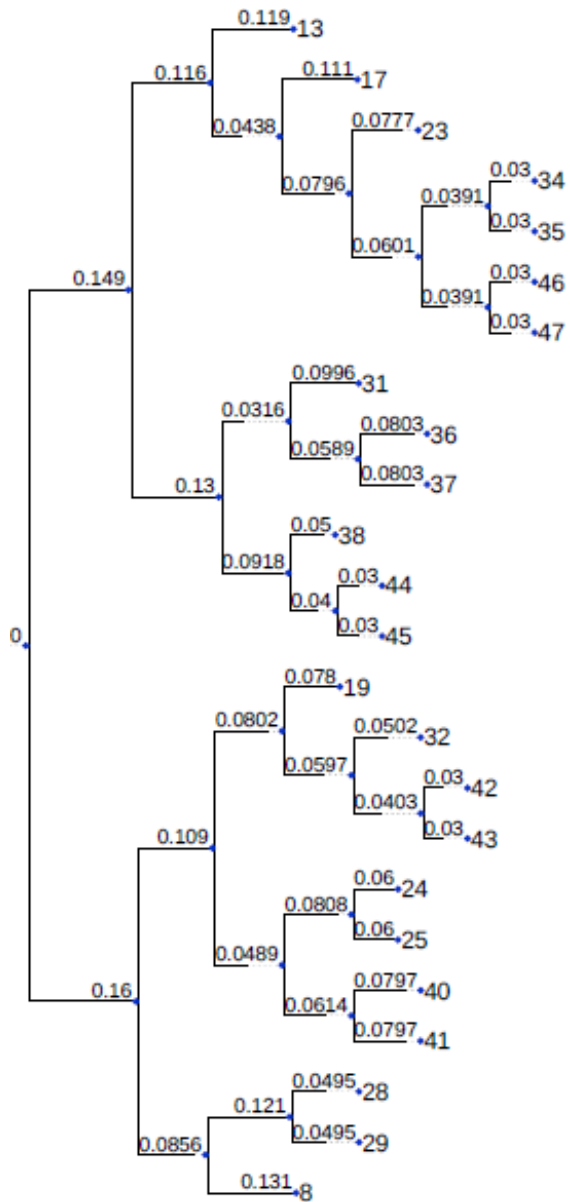




Εικόνα 6: Τοπολογία 6 - ακμές με καθυστερήσεις (επιστροφή στον 5.1).

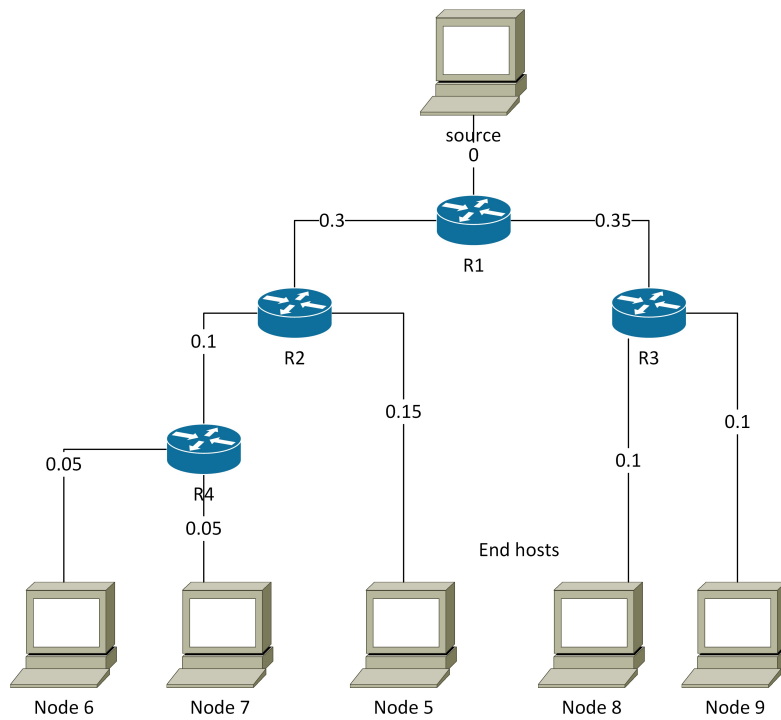


(α) jFed

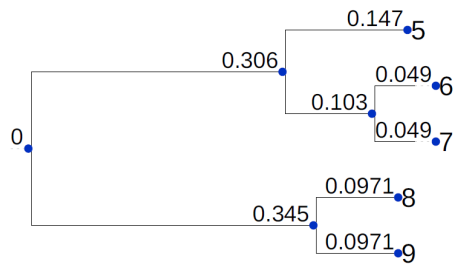


(β') Upsilon

Εικόνα 7: Τοπολογία 7 - ακμές με καθυστερήσεις (επιστροφή στον 5.1).

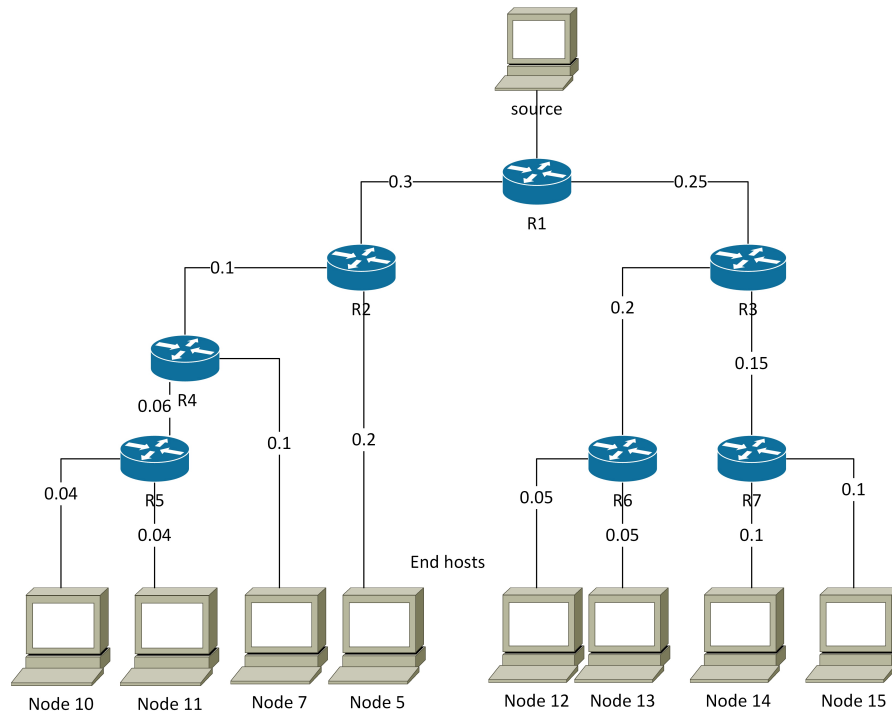


( $\alpha'$ ) jFed

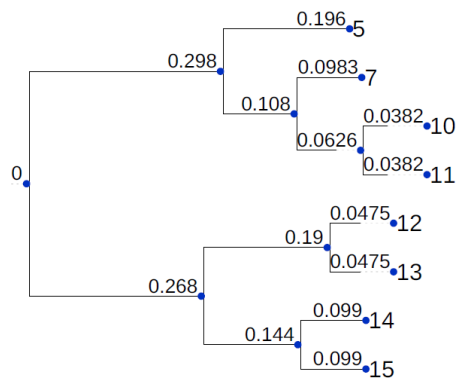


( $\beta'$ ) Upsilon

Εικόνα 8: Τοπολογία 1 - ακμές με απώλειες (επιστροφή στον 5.1).

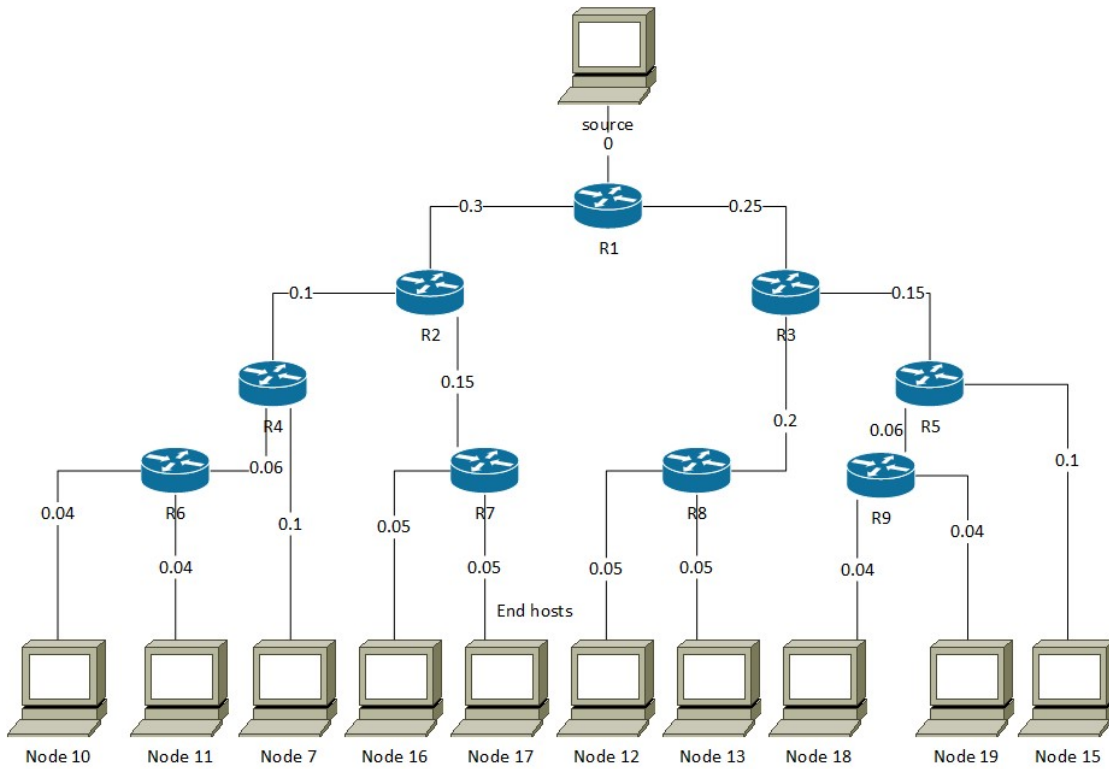


( $\alpha$ ) jFed

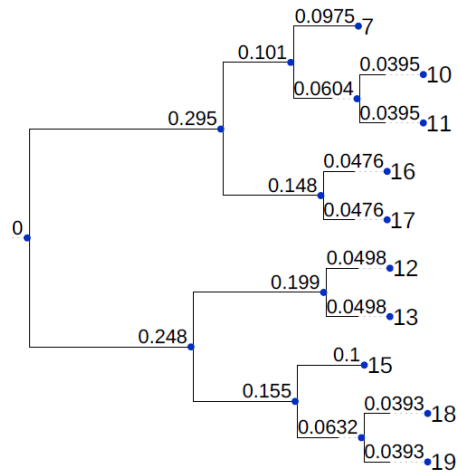


( $\beta$ ) Upsilon

Εικόνα 9: Τοπολογία 2 - ακμές με απώλειες (επιστροφή στον 5.1).

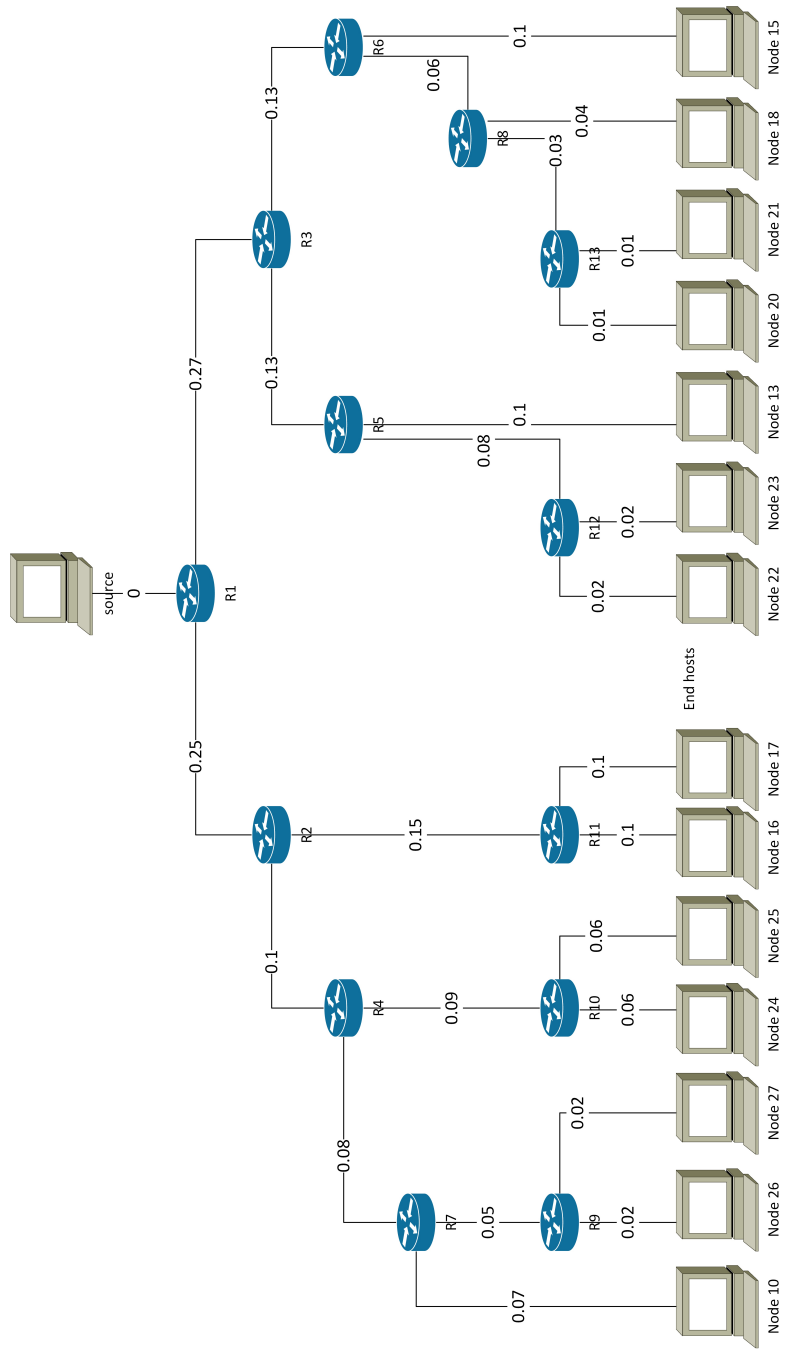


( $\alpha'$ ) jFed



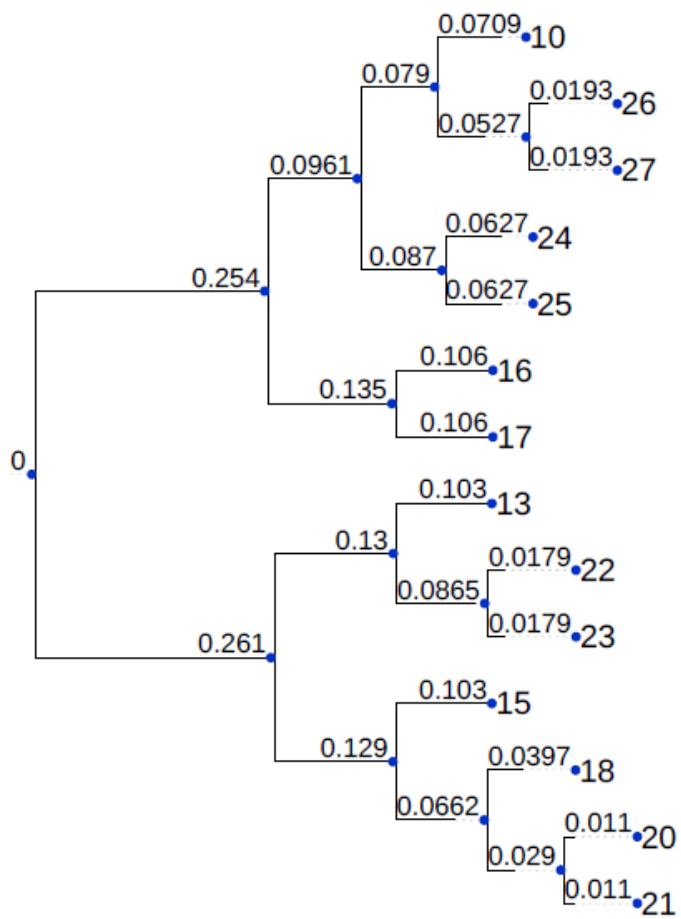
( $\beta'$ )  $\Upsilon$ rgma

Εικόνα 10: Τοπολογία 3 - ακμές με απώλειες (επιστροφή στον 5.1).



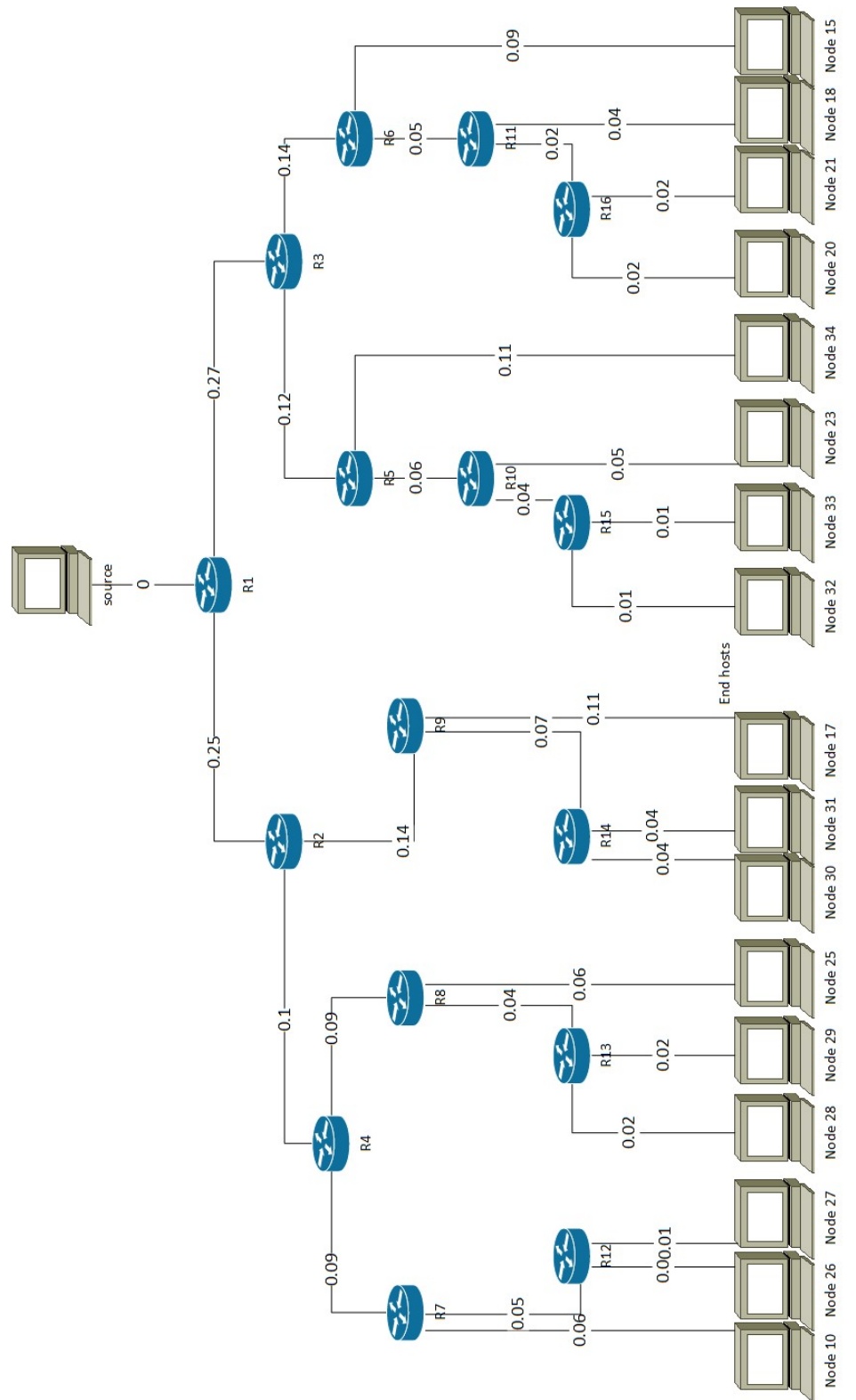
(α) jFed



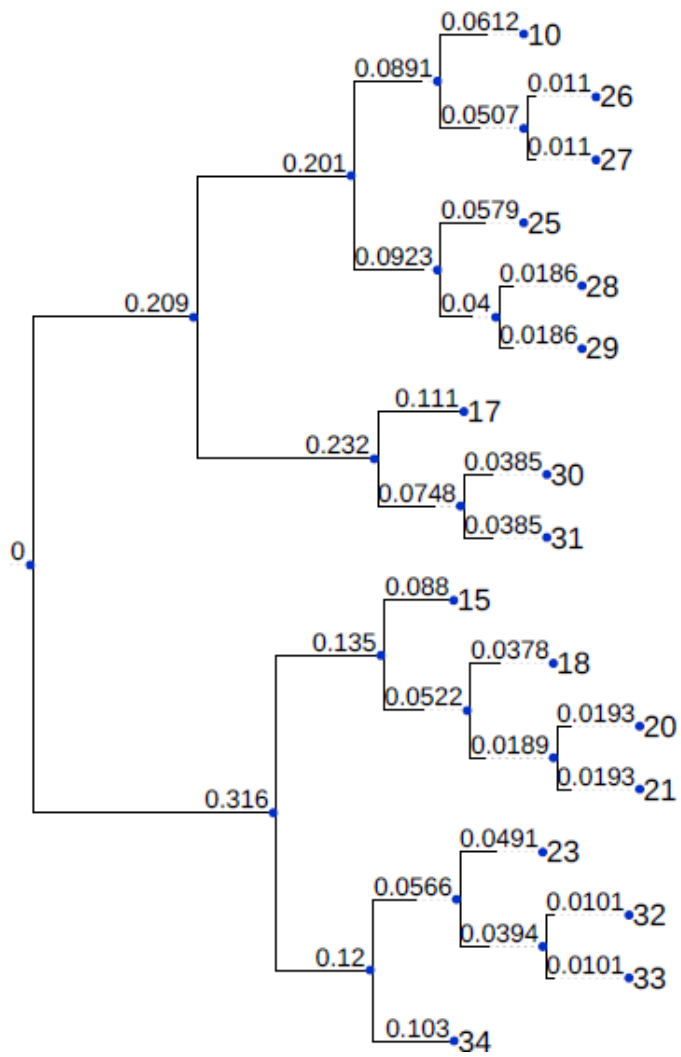


(β') Υργμα

Εικόνα 11: Τοπολογία 4 - ακμές με απώλειες (επιστροφή στον 5.1).

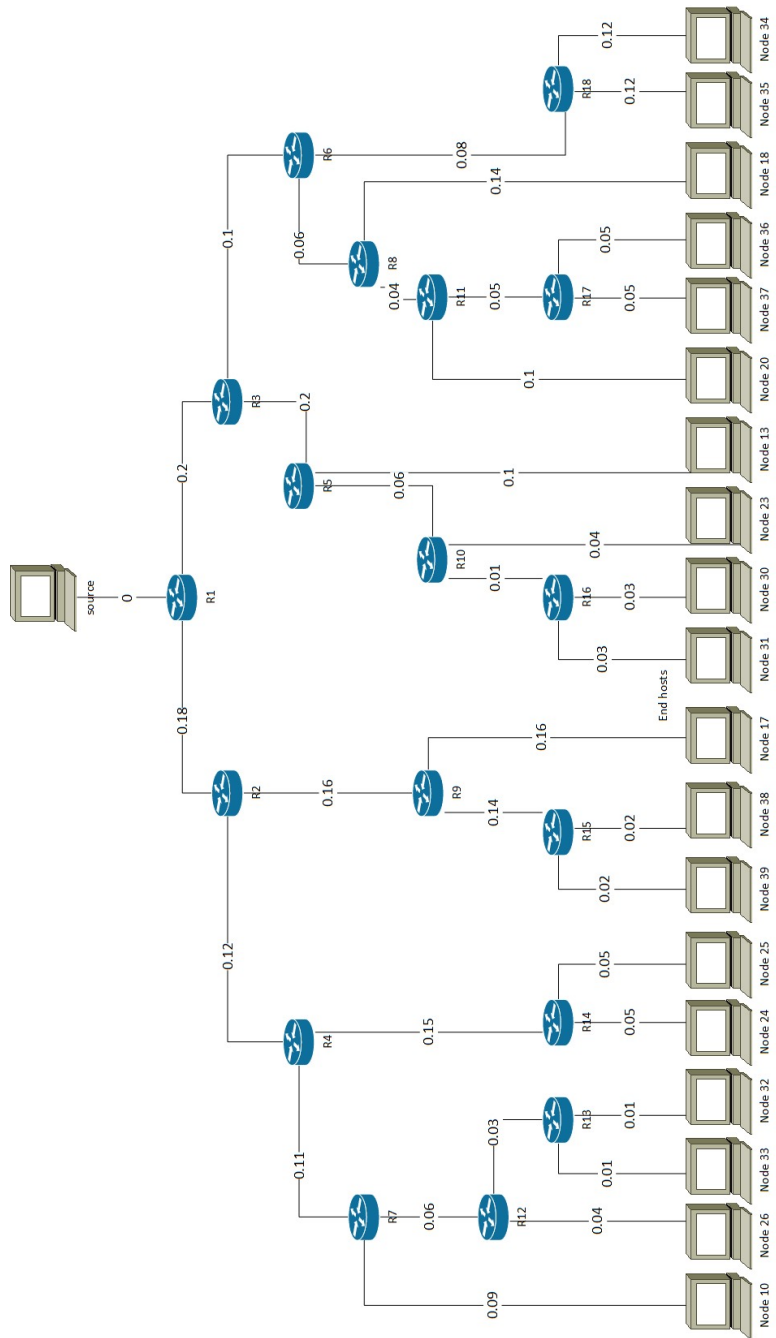


(α) jFed

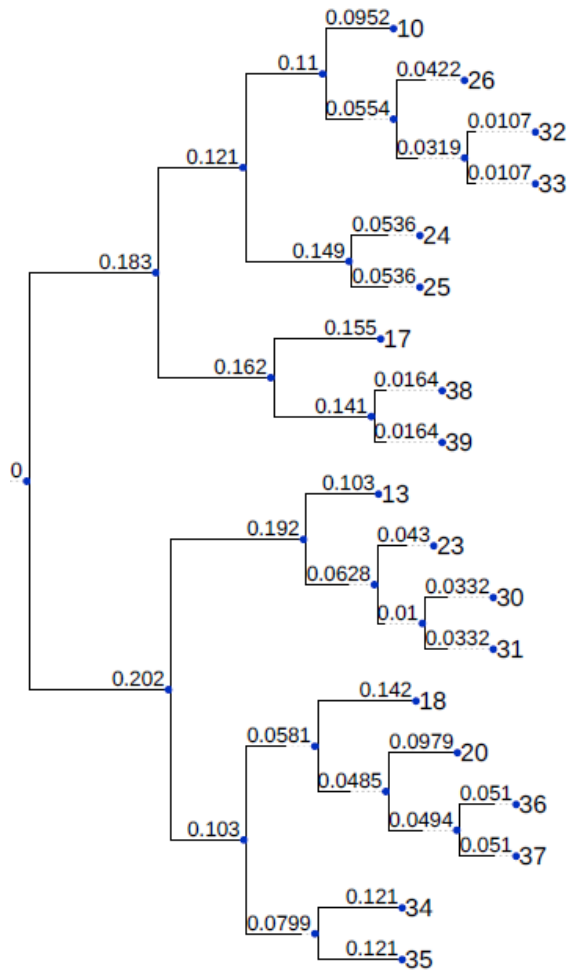


(β') Upsilon

Εικόνα 12: Τοπολογία 5 - ακμές με απώλειες (επιστροφή στον 5.1).



(α') jFed



(β') Upsilon

Εικόνα 13: Τοπολογία 6 - ακμές με απώλειες (επιστροφή στον 5.1).

## B Πηγαίος κώδικας

Υπολογισμός πίνακα αποστάσεων βάσει των καθυστερήσεων των ακμών

```
1 import sys
2 import os
3 import glob
4 from os import walk
5 import math
6 import re
7 import copy
8 import numpy as np
9 from itertools import combinations
10 import random
11 from dict import dict
12
13
14 # compute the covariance of all elements of the lists lst1 & lst2
15 def covariance(lst1, ave1, lst2, ave2, le):
16     return (sum([(x1 - ave1)*(x2 - ave2) for (x1, x2) in zip(lst1, lst2)
17 ])))/(le-1)
18
19 #number of non internal nodes as input1
20 num_of_nds = int(sys.argv[1])
21 #path to the folder that contains the log files as input2
22 path = sys.argv[2]
23 os.chdir(path)
24
25 lin = 0
26 # find all possible combinations of the terminal nodes in pairs
27 comb = list(combinations(range(num_of_nds-1), 2))
28 # Distance matrix
29 Dist = [[0 for x in range(num_of_nds-1)] for y in range(num_of_nds-1)]
30 t = {}
31 times = [{} for i in range(num_of_nds)]
32 delays = [{} for l in range(num_of_nds-1)]
```

```
32 nums = []
33 ave = []
34 cov = []
35 dij = []
36 keys = []
37 inter_keys = []
38 L1 = []
39 L2 = []
40 L = []
41 x = 0
42 # open only files with 10k measurements named "nodeX_10k.txt", where X
    the id of the node
43 for filename in sorted(glob.glob(os.path.join(path, '*k.txt'))):
44     num = filename.split('node', 1)[1]
45     num = num.split('_', 1)[0]
46     if num != '0':
47         nums.append(str(num))
48     lin = 0
49     with open(filename, 'r') as f:
50         for line in f:
51 # make a dictionary where key is the ID of the packet & value is it's
    delay
52         if lin == 0 or lin % 2 == 0:
53             m = re.search(r"(?<=id )(\d+)", line)
54             t[int(m.group(1))] = float(line[6:15])
55             lin += 1
56         times[x] = copy.copy(t)
57         t.clear()
58         x += 1
59
60 for x in range(0, num_of_nds-1):
61     for y in times[x+1]:
62 # for the packets that have reached terminal node x, compute their delay
    from the source
63         if times[0].get(y) is not None:
64             dif = float(times[x + 1][y] - times[0][y])
```

```

65         if dif >= 0:
66             delays[x][y] = dif
67         else:
68             delays[x][y] = dif + 60
69
70 # compute average delay from the source to terminal node x
71 for x in range(num_of_nds-1):
72     L = list(delays[x].values())
73     ave.append(np.average(L))
74     L.clear()
75
76 # for all pairs of terminal nodes, find the packets that have reached
77   both nodes & compute their covariance
77 for x in range(len(comb)):
78     inter_keys = list(set(delays[comb[x][0]].keys()).intersection(set
79   (delays[comb[x][1]].keys())))
79     l1l = len(inter_keys)
80     for p in range(l1l):
81         L1.append(delays[comb[x][0]][inter_keys[p]])
82         L2.append(delays[comb[x][1]][inter_keys[p]])
83     M = np.array([L1, L2])
84     cov.append(np.cov(M, ddof=1))
85     L1.clear()
86     L2.clear()
87     M = []
88     inter_keys.clear()
89 # compute the distance between two terminal nodes using (4.7)
90 for x in range(len(comb)):
91     dij.append(cov[x][0][0]+cov[x][1][1]-2*cov[x][0][1])
92
93 # construct Distance matrix
94 for x in range(num_of_nds-1):
95     Dist[x][x] = 0
96
97 for x in range(0, len(comb)):
98     Dist[comb[x][0]][comb[x][1]] = float("{0:.4f}".format(dij[x]))

```



```

99     Dist[comb[x][1]][comb[x][0]] = float("{0:.4f}".format(dij[x]))
100
101 with open("output.txt", 'w') as f:
102     for x in range(num_of_nds-1):
103         if x < num_of_nds-2:
104             f.write(nums[x] + ',')
105         else:
106             f.write(nums[x] + '\n')
107     f.writelines(','.join(str(j) for j in i) + '\n' for i in Dist)

```

Υπολογισμός πίνακα αποστάσεων βάσει των απωλειών των ακμών

```

1 import sys
2 import os
3 import math
4 import glob
5 import re
6 import copy
7 from itertools import combinations
8 import numpy as np
9
10
11 # number of non internal nodes as input1
12 num_of_nds = int(sys.argv[1])
13 # path to the folder that contains the log files as input2
14 path = sys.argv[2]
15 os.chdir(path)
16
17
18 nums = []
19 Dist = [[0 for x in range(num_of_nds-1)] for y in range(num_of_nds-1)]
20 ids = set()
21 id_sets = [set() for i in range(num_of_nds)]
22 # this is the number of common packets that any two destination hosts
    receive (not including the source)
23 inter = []

```

```

24 packets = []
25 di = []
26 dij = []
27 Xi = []
28 Xij = []
29 products = []
30 x = 0
31 for filename in sorted(glob.glob(os.path.join(path, '*10k.txt'))):
32     num = filename.split('node', 1)[1]
33     num = num.split('_', 1)[0]
34     if num != '0':
35         nums.append(str(num))
36         with open(filename, 'r') as f:
37             for line in f:
38                 m = re.search(r"(?<=id )(\d+)", line)
39                 if m:
40                     ids.add(m.group(1))
41 # list of sets containing the ids of the packets received at any end host
42     # (or sent from the source)
43     id_sets[x] = ids.copy()
44     ids.clear()
45     x += 1
46 # find all possible combinations of the terminal nodes in pairs
47 comb = list(combinations(range(num_of_nds - 1), 2))
48 # find which packets have reached both destinations of each pair
49 for l in list(combinations(range(1, num_of_nds), 2)):
50     inter.append(len(id_sets[l[0]].intersection(id_sets[l[1]])))
51
52 # compute number of packets that have reached each end host (or that were
53 # sent from the source - packets[0])
54 for x in range(num_of_nds):
55     packets.append(len(id_sets[x]))
56 # compute average success rate for each end host (4.3)
57 for m in range(num_of_nds-1):
58     Xi.append(packets[m+1]/packets[0])

```

```

58 # compute average joint success rate for each pair of end hosts (4.4)
59 for m in range(len(comb)):
60     Xij.append(inter[m]/packets[0])
61
62 for l in comb:
63     products.append(Xi[l[0]]*Xi[l[1]])
64
65 # length metric for each network edge
66 # for m in range(num_of_nds-1):
67 #     di.append(math.log(1/Xi[m]))
68
69 #distance metric between any two end hosts
70 for m in range(len(comb)):
71     dij.append(math.log((products[m])/Xij[m]**2))
72
73 #construction of Distance matrix
74 for x in range(num_of_nds-1):
75     Dist[x][x] = 0
76
77 for x in range(0, len(comb)):
78     Dist[comb[x][0]][comb[x][1]] = float("{0:.4f}".format(dij[x]))
79     Dist[comb[x][1]][comb[x][0]] = float("{0:.4f}".format(dij[x]))
80
81 with open("output1.txt", 'w') as f:
82     for x in range(num_of_nds-1):
83         if x < num_of_nds-2:
84             f.write(nums[x] + ',')
85         else:
86             f.write(nums[x] + '\n')
87     f.writelines(','.join(str(j) for j in i) + '\n' for i in Dist)

```

### Αλγόριθμος *Urgma*

```

1 import sys
2 import numpy as np
3 import _tkinter

```

```
4 import matplotlib.pyplot as plt
5 import networkx as nx
6 import math
7 import pydot
8 import graphviz
9 from hierarchy import hierarchy_pos
10 from ete3 import Tree, TreeStyle, TreeNode
11
12
13 # the Distance matrix is needs to be symmetric
14 def check_symmetric(a, rtol=1e-05, atol=1e-08):
15     return np.allclose(a, a.T, rtol=rtol, atol=atol)
16
17
18 filename_1 = str(sys.argv[1])
19 file1 = open(filename_1, 'r')
20 # if delo=1 then we need to compute losses, else if delo=0 we need to
    compute delays
21 delo = int(sys.argv[2])
22 # names of end hosts
23 nodes = list(np.loadtxt(filename_1, delimiter=',', max_rows=1, dtype=str)
    )
24 # the Distance matrix
25 dist_matrix = np.loadtxt(filename_1, skiprows=1, delimiter=',', dtype=
    float)
26 leaves = list(nodes)
27 dtree = {}
28 # insert leaves into the tree
29 for i in leaves:
30     dtree.update({i: i})
31 edges = []
32
33 if not(check_symmetric(dist_matrix)):
34     sys.exit("The matrix you gave is not symmetric.")
35
36 dim = dist_matrix.shape
```

```
37
38 G = nx.Graph()
39 t = ''
40 distances = []
41 weights = {}
42 num_of_elem = []
43 merge = ()
44 # keep the number of elements of each merged set (initially is 1)
45 for x in range(0, dim[0]):
46     num_of_elem.append(1)
47 Min = dist_matrix[0, 1]
48 L = []
49 height = 0
50 length = 0
51 while dim[0] > 1:
52     Min = dist_matrix[0, 1]
53     #locate minimum distance
54
55     for r in range(0, dim[0]):
56         for c in range(0, dim[1]):
57             if 0 < dist_matrix[r, c] < Min:
58                 Min = dist_matrix[r, c]
59                 L.clear()
60                 L.append((r, c))
61             elif dist_matrix[r, c] == Min:
62                 L.append((r, c))
63 # break tie arbitrarily considering the pair constructed from the first
64 # element of the list
65 # the matrix is considered to be symmetric
66 # D is the distance from the selected (to be merged) end hosts to their
67 # father
68 D = Min / 2
69 merge = (L[0][0], L[0][1])
70 dif1 = 0
71 dif2 = 0
```

```

71 # distances is used in order to compute inner edges' lengths (i.e.
    lengths of edges connecting nodes that were created by merging other
    nodes)
72 for x in range(0, len(distances)):
73     if distances[x][0] == nodes[merge[0]]:
74         dif1 = distances[x][1]
75 for x in range(0, len(distances)):
76     if distances[x][0] == nodes[merge[1]]:
77         dif2 = distances[x][1]
78
79 if delo == 1:
80     # Compute weights as losses (4.5)
81     weight1 = float("{0:.4f}".format(1-math.exp(-(D - dif1))))
82     weight2 = float("{0:.4f}".format(1-math.exp(-(D - dif2))))
83 else:
84     # Compute weights as jitter (4.8)
85     weight1 = float("{0:.4f}".format(1000*math.sqrt(D - dif1)))
86     weight2 = float("{0:.4f}".format(1000*math.sqrt(D - dif2)))
87
88
89 dtree.update({(nodes[merge[0]]+nodes[merge[1]]): str('(')+str(dtree[
nodes[merge[0]]])+':'+str(weight1)+','+str(dtree[nodes[merge[1]]])+':'+
str(weight2)+')'})
90 L.clear()
91
92 nodes[merge[0]] = nodes[merge[0]] + nodes[merge[1]]
93 del nodes[merge[1]]
94
95 # in the list distances (in tuples), keep the distance between the new
    node (that was created from merging the selected nodes) and the leaves
    it ends up to.
96 distances.append((nodes[merge[0]], D))
97
98 # compute all new distances based on (5.3)
99 for c in range(0, dim[1]):
100     if c != merge[0] and c != merge[1]:

```

---

```
101         new = (dist_matrix[merge[0], c]*num_of_elem[merge[0]] +
dist_matrix[merge[1], c]*num_of_elem[merge[1]])/(num_of_elem[merge
[0]]+num_of_elem[merge[1]])
102         dist_matrix[merge[0], c] = new
103         dist_matrix[c, merge[0]] = new
104     dist_matrix = np.delete(dist_matrix, merge[1], 0)
105     dist_matrix = np.delete(dist_matrix, merge[1], 1)
106     num_of_elem[merge[0]] = num_of_elem[merge[0]] + num_of_elem[merge[1]]
107     del num_of_elem[merge[1]]
108     dim = dist_matrix.shape
109
110 # print the dendrogram
111 t = Tree(str(dtree[nodes[merge[0]]]+'s'+';'), format=1)
112 ts = TreeStyle()
113 ts.show_leaf_name = True
114 ts.show_branch_length = True
115 t.show(tree_style=ts)
```

## C Θεμελιώδη στοιχεία των Markov Random Fields

Τα Τυχαία Πεδία Markov (Markov Random Fields - MRF), αποτελούν ένα τρόπο αναπαράστασης και οπτικοποίησης κατανομών πιθανότητας, με τη χρήση μη κατευθυνόμενων γράφων, με σκοπό την αναπαράσταση υποθέσεων ανεξαρτησίας των ακμών, που τα Μπεϋζιανά δίκτυα (Bayesian Networks), τα οποία κάνουν χρήση κατευθυνόμενων γράφων, αδυνατούν να αναπαραστήσουν [57]. Στη συνέχεια, θα παρουσιαστούν τα θεμελιώδη στοιχεία των MRF όπως αυτά αναφέρονται στις εργασίες [58], [59].

Έστω πεπερασμένο σύνολο  $\mathcal{S}$ ,  $|\mathcal{S}| = \mathcal{N}$ , του οποίου τα στοιχεία  $s \in \mathcal{S}$  αναφέρονται ως sites και αντιστοιχούν στους κόμβους ενός δικτύου. Ορίζεται ο φασικός χώρος  $\Lambda$  ως το σύνολο των πιθανών καταστάσεων κάθε στοιχείου  $s \in \mathcal{S}$ . Μία συλλογή  $\mathcal{X} = \{\mathcal{X}_s, \forall s \in \mathcal{S}\}$  τυχαίων μεταβλητών με τιμές στον φασικό χώρο  $\Lambda$ , καλείται Τυχαίο Πεδίο (Random Field - RF) στο  $\mathcal{S}$ , με φάσεις στο  $\Lambda$ . Μία διαμόρφωση (configuration)  $\mathbf{x} = \{x_s, \forall s \in \mathcal{S}, x_s \in \Lambda\}$ , αντιστοιχεί σε μία εκ των πιθανών καταστάσεων του συστήματος (δικτύου) και ο χώρος γινομένου (product space)  $\Lambda^{\mathcal{N}}$ ,  $x \in \Lambda^{\mathcal{N}}$  αποτελεί τον χώρο διαμόρφωσης (configuration space) του συστήματος.

Εν συνεχεία, ορίζεται ένα γειτονικό σύστημα (neighborhood system) στο  $\mathcal{S}$  ως η οικογένεια  $\mathcal{N} = \{\mathcal{N}_s\}_{s \in \mathcal{S}}$  των υποσυνόλων  $\mathcal{N}_s \subset \mathcal{S}$ , για τα οποία ισχύει ότι  $\forall s, r \in \mathcal{S}$ ,  $s \notin \mathcal{N}_s$  και  $r \in \mathcal{N}_s$  αν και μόνο αν  $s \in \mathcal{N}_r$ . Το  $\mathcal{N}_s$  αποτελεί τη γειτονιά του site  $s$ . Το RF  $\mathcal{X}$  καλείται Markov Random Field αναφορικά με το  $\mathcal{N}$ , αν  $\forall s \in \mathcal{S}$  ισχύει το ακόλουθο:

$$\mathbb{P}(\mathcal{X}_s = x_s | \mathcal{X}_r = x_r, r \neq s) = \mathbb{P}(\mathcal{X}_s = x_s | \mathcal{X}_r = x_r, r \in \mathcal{N}_s) \quad (1)$$

Τελικά, η κατάσταση στην οποία βρίσκεται ένας κόμβος ενός MRF δικτύου εξαρτάται μόνο από τις καταστάσεις των γειτονικών του κόμβων, εάν οι τυχαίες μεταβλητές που τις περιγράφουν είναι παρατηρήσιμες. Διαφορετικά, προκύπτουν επιπλέον εξαρτήσεις, οι οποίες, όμως, εκφράζονται ως από κοινού πιθανότητες (και όχι ως υπό συνθήκη πιθανότητες, όπως στην περίπτωση των Μπεϋζιανών Δικτύων), απαιτώντας λιγότερη πρότερη γνώση (prior knowledge) σε σχέση με τα Μπεϋζιανά Δίκτυα [57].



