



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Αλγόριθμοι δρομολόγησης σε ασύρματα δίκτυα πλέγματος
με χρήση της μαρκοβιανής θεωρίας αποφάσεων**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνα Ι. Χρηστάκου

Επιβλέπων: Παναγιώτης Κωττής

Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2010



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Αλγόριθμοι δρομολόγησης σε ασύρματα δίκτυα πλέγματος
με χρήση της μαρκοβιανής θεωρίας αποφάσεων**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνα Ι. Χρηστάκου

Επιβλέπων: Παναγιώτης Κωττής

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 25^η Ιουνίου 2010

.....
Παναγιώτης Κωττής
Καθηγητής Ε.Μ.Π.

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

.....
Γεώργιος Φικιώρης
Επ.Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2010

.....
Κωνσταντίνα Ι. Χρηστάκου

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π

Copyright © Κωνσταντίνα Ι. Χρηστάκου, 2010.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

ΠΕΡΙΛΗΨΗ

Σκοπός της παρούσας διπλωματικής εργασίας είναι η σχεδίαση ενός αλγορίθμου δρομολόγησης δεδομένων σε ασύρματα δίκτυα πλέγματος που λειτουργούν υπό συνθήκες διαλείψεων με χρήση μαρκοβιανών διαδικασιών λήψης αποφάσεων. Τα δεδομένα στο δίκτυο θεωρείται ότι μεταφέρονται με τη βοήθεια IP πακέτων τα οποία επιθυμούν να φτάσουν από τον κόμβο-πηγή στον κόμβο-προορισμό με την ελάχιστη από-άκρο-σε-άκρο καθυστέρηση και ταυτόχρονα τη μέγιστη αξιοπιστία. Όταν ένα IP πακέτο φθάσει σε έναν κόμβο του δικτύου πρέπει να επιλεγεί σε ποιον από τους γειτονικούς του κόμβους θα δρομολογηθεί. Το βασικό κριτήριο για την απόφαση αυτή είναι η πιθανότητα επιτυχημένης μετάδοσης η οποία επηρεάζεται από την ισχύ εκπομπής των κόμβων. Με χρήση των εξισώσεων Bellman υπολογίζεται η βέλτιστη πολιτική δρομολόγησης για κάθε πακέτο και, ταυτόχρονα, διερευνάται η σχέση μεταξύ καθυστέρησης μετάδοσης και ισχύος εκπομπής.

Στο Κεφάλαιο 1 γίνεται γενική αναφορά στα βασικά χαρακτηριστικά των ασυρμάτων δικτύων πλέγματος και θίγεται το θέμα της δρομολόγησης σε δίκτυα αυτού του τύπου. Στο Κεφάλαιο 2 δίδονται τα βασικά στοιχεία της μαρκοβιανής θεωρίας και παρουσιάζονται οι βασικοί αλγόριθμοι επίλυσης μαρκοβιανών διαδικασιών απόφασης. Στο Κεφάλαιο 3 αναφέρονται ορισμένες εφαρμογές της μαρκοβιανής θεωρίας στην επίλυση προβλημάτων που ανακύπτουν σε δίκτυα επικοινωνιών. Στο Κεφάλαιο 4 παρουσιάζεται η μαρκοβιανή διαδικασία που περιγράφει τον αλγόριθμο δρομολόγησης σε ένα ασύρματο δίκτυο πλέγματος. Τέλος, στο Κεφάλαιο 5 μελετάται η επίδοση του προτεινόμενου αλγορίθμου δρομολόγησης ως προς την ταχύτητα σύγκλισης, την ενεργειακή κατανάλωση, την καθυστέρηση και την κλιμάκωση αυτού με χρήση κατάλληλου μοντέλου προσομοίωσης.

Λέξεις κλειδιά

Ασύρματα δίκτυα πλέγματος, μαρκοβιανές διαδικασίες απόφασης, δρομολόγηση, διαλείψεις λόγω βροχής

ABSTRACT

The subject of this thesis is the design of a routing algorithm for wireless mesh networks operating under fading conditions employing a Markovian Decision Process. Network traffic is assumed to be transferred by IP packets that aim at arriving from the source to the destination node with minimum end-to-end delay and maximum reliability. When an IP packet reaches a specific node, it selects, out of the set of its neighboring nodes, the one that will be forwarded. The basic decision criterion is based on the probability of successful transmission which is related to the transmitted power. Based on the iterative Bellman's equations, the optimal routing policy at a packet level is determined. The energy-delay trade-off is investigated.

The thesis is organized as follows. The first chapter deals with the basic features of wireless mesh networks putting emphasis on the routing algorithms commonly used in such networks. In the next two chapters Markov Decision Theory is thoroughly presented along with relevant applications in communication networks. In Chapter 4 the proposed routing algorithm is formulated as a Markov Decision Process. In Chapter 5, several related factors are studied (convergence rate, energy consumption, scalability) in order to examine the performance of the proposed routing algorithm.

Keywords

Wireless mesh networks, markov decision processes, routing, rain attenuation

Θα ήθελα να εκφράσω τις ειλικρινέστατες ευχαριστίες μου στον επιβλέποντα Καθηγητή κ. Παναγιώτη Κωττή, του οποίου η συμβολή στην ανάπτυξή μου ως μηχανικού τηλεπικοινωνιών ήταν καθοριστική καθόλη τη διάρκεια των σπουδών μου. Θα ήθελα να τον ευχαριστήσω για τις πολύτιμες συμβουλές του που καθόρισαν την εξέλιξη της διπλωματικής. Επιπλέον θα ήθελα να ευχαριστήσω βαθύτατα το Διδάκτορα Ε.Μ.Π. κ. Μάρκο Αναστασόπουλο για το χρόνο που μου διέθεσε, την πολύτιμη βοήθειά του και την άριστη συνεργασία μας κατά την εκπόνηση της διπλωματικής αυτής εργασίας.

Αφιερώνεται στη μητέρα μου,

Σοφία.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1. ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ ΠΛΕΓΜΑΤΟΣ	15
1.1 Εισαγωγή	15
1.2 Βασικά χαρακτηριστικά ασυρμάτων δικτύων πλέγματος	15
1.2.1 Αρχιτεκτονική ασυρμάτων δικτύων πλέγματος	15
1.2.2 Χαρακτηριστικά γνωρίσματα ασυρμάτων δικτύων πλέγματος.....	19
1.2.3 Σχεδιαστικές απαιτήσεις ασυρμάτων δικτύων πλέγματος	19
1.2.4 Εφαρμογές και δυνατότητες ανάπτυξης ασυρμάτων δικτύων πλέγματος.....	22
1.3 Δρομολόγηση σε ασύρματα δίκτυα πλέγματος	23
1.3.1 Σχεδιαστικές απαιτήσεις πρωτοκόλλων δρομολόγησης για εφαρμογή σε ασύρματα δίκτυα πλέγματος	23
1.3.1 Πρωτόκολλα δρομολόγησης σε δίκτυα WMN.....	26
2. ΜΑΡΚΟΒΙΑΝΕΣ ΔΙΑΔΙΚΑΣΙΕΣ ΑΠΟΦΑΣΗΣ	31
2.1 Εισαγωγή	31
2.2 Καθορισμός μοντέλου	32
2.2.1 Ανάλυση παραμέτρων του μαρκοβιανού μοντέλου	33
2.2.2 Κανόνες απόφασης / Πολιτική	37
2.3 Μαρκοβιανές διαδικασίες πεπερασμένου ορίζοντα	38
2.3.1 Κριτήριο συνολικού αναμενόμενου οφέλους.....	38
2.3.2 Εύρεση βέλτιστης πολιτικής.....	40
2.3.3 Εξισώσεις βελτιστοποίησης Bellman (Bellman Equations).....	40
2.3.4 Αλγόριθμος επαγωγής προς τα πίσω (Backward induction)	41
2.4 Μαρκοβιανές διαδικασίες άπειρου ορίζοντα	43
2.4.1 Κριτήρια οφέλους.....	43
2.4.2 Βέλτιστη πολιτική	44
2.4.3 Εξισώσεις Βελτιστοποίησης Bellman	45
2.4.4 Αλγόριθμος αναδρομικής αξίας (Value iteration).....	45
2.4.5 Αλγόριθμος αναδρομικής πολιτικής (Policy iteration)	46

3. ΕΦΑΡΜΟΓΕΣ ΜΑΡΚΟΒΙΑΝΩΝ ΜΟΝΤΕΛΩΝ ΣΕ ΔΙΚΤΥΑ ΕΠΙΚΟΙΝΩΝΙΩΝ	49
3.1 Εισαγωγή	49
3.2 Δυναμική εκχώρηση φάσματος / Γνωστικές ραδιοεπικοινωνίες (Physical/MAC layer)	49
3.2.1 Περιγραφή του προβλήματος ανίχνευσης φάσματος	50
3.2.3 Επίλυση μαρκοβιανής διαδικασίας	52
3.2.4 Συμπεράσματα.....	53
3.3 Δρομολόγηση στο στρώμα δικτύου	53
3.3.1 Διατύπωση και επίλυση μαρκοβιανής διαδικασίας.....	53
3.4 Προγραμματισμός παροχής υπηρεσιών στο στρώμα εφαρμογών	54
3.4.1 Διατύπωση και επίλυση μαρκοβιανής διαδικασίας.....	54
4. ΜΑΘΗΜΑΤΙΚΗ ΑΝΑΛΥΣΗ ΠΡΟΒΛΗΜΑΤΟΣ	57
4.1 Εισαγωγή	57
4.2 Περιγραφή τοπολογίας δικτύου	57
4.2.1 Προσομοίωση τοπολογίας δικτύου	58
4.3 Κατάστρωση μαρκοβιανού μοντέλου	58
4.3.1 Σύνολο καταστάσεων του συστήματος.....	58
4.3.2 Σύνολο δυνατών δράσεων του συστήματος.....	59
4.3.3 Πιθανότητες και πίνακας μετάβασης του συστήματος.....	59
4.3.4 Απώλειες του συστήματος.....	61
4.4 Επίλυση μαρκοβιανού μοντέλου	62
5. ΠΕΡΙΒΑΛΛΟΝ ΠΡΟΣΟΜΟΙΩΣΗΣ ΚΑΙ ΑΡΙΘΜΗΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ	81
5.1 Εισαγωγή	81
5.2 Προσομοίωση της απόσβεσης λόγω βροχής σε δίκτυα επικοινωνιών	81
5.2.1 Διαλείψεις που οφείλονται στην εξασθένηση των Η/Μ κυμάτων λόγω βροχής.....	81
5.2.1.1 Μικροσκοπική περιγραφή του μέσου βροχής.....	82
5.2.1.2 Μακροσκοπική περιγραφή του μέσου βροχής	83
5.2.2 Καθορισμός του μοντέλου απόσβεσης λόγω βροχής.....	84

5.2.2.1 Παραδοχές του μοντέλου απόσβεσης.....	84
5.2.2.2 Εύρεση λογαριθμοκανονικής συνάρτησης αθροιστικής κατανομής (Lognormal Complementary Cumulative Distribution Function- CCDF) για την περιγραφή της απόσβεσης λόγω βροχής σε ένα σημείο του πλέγματος.....	85
5.2.2.3 Χωροχρονική συσχέτιση της απόσβεσης λόγω βροχής	86
5.2.3 Σύνοψη πεδίων απόσβεσης λόγω βροχής	87
5.2.3.1 Πεδίο απόσβεσης λόγω βροχής παραγόμενο από διδιάστατο γκαουσιανό πεδίο	87
5.2.3.2 Δημιουργία χωρικά συσχετισμένων γκαουσιανών πεδίων.....	88
5.2.3.3 Χωρική μεταφορά των βροχοπυρήνων	90
5.2.3.4 Χρονική μετάθεση πεδίου βροχής.....	90
5.2.3.5 Αποτελέσματα προσομοίωσης διαλείψεων λόγω βροχής	91
5.3 Μελέτη δρομολόγησης σε δίκτυα WMN υπό συνθήκες διαλείψεων λόγω βροχής	92
5.3.1 Επίδραση της ισχύος εκπομπής στην από-άκρο-σε-άκρο καθυστέρηση (end- to-end delay).....	93
5.3.2 Επίδραση του μεγέθους του δικτύου στην ταχύτητα σύγκλισης του αλγορίθμου δρομολόγησης	98
5.3.3 Επίδραση του παράγοντα μείωσης του κέρδους λ στην ταχύτητα σύγκλισης του αλγορίθμου	98
5.3.4 Βελτιστοποίηση του αλγορίθμου δρομολόγησης.....	100
ΠΑΡΑΡΤΗΜΑ Α	103
ΠΑΡΑΡΤΗΜΑ Β.....	105
ΠΑΡΑΡΤΗΜΑ Γ	109
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	111

ΚΕΦΑΛΑΙΟ 1

ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ ΠΛΕΓΜΑΤΟΣ

1.1 Εισαγωγή

Στόχος του κεφαλαίου είναι η γνωριμία του αναγνώστη με βασικές έννοιες των ασυρμάτων δικτύων πλέγματος. Αρχικά, παρατίθενται τα βασικά χαρακτηριστικά των δικτύων πλέγματος και, στη συνέχεια, αναλύονται οι δυνατότητες εξέλιξής τους και οι πιθανές εφαρμογές τους. Το κεφάλαιο κλείνει θίγοντας το θέμα της δρομολόγησης σε ασύρματα δίκτυα πλέγματος. Αναφέρονται οι απαιτήσεις που πρέπει να ικανοποιεί ένα πρωτόκολλο δρομολόγησης και παρουσιάζονται τα βασικά είδη δρομολόγησης που εφαρμόζονται σε δίκτυα πλέγματος.

1.2 Βασικά χαρακτηριστικά ασυρμάτων δικτύων πλέγματος

Τα ασύρματα δίκτυα πλέγματος (Wireless Mesh Networks - WMN) δημιουργήθηκαν για να υποστηρίξουν ασύρματη δικτύωση νέας γενιάς. Πρόκειται για δίκτυα που οργανώνονται με δυναμικό τρόπο, παρουσιάζουν συνδεσιμότητα πλέγματος και εμφανίζουν αρκετά κοινά χαρακτηριστικά με τα συμβατικά adhoc δίκτυα. Η ταχεία ανάπτυξή τους οφείλεται στις εξελιγμένες εφαρμογές που μπορούν να υποστηρίξουν σε σχέση με τα κλασικά ασύρματα δίκτυα καθώς επίσης και στα πολλαπλά πλεονεκτήματα που εμφανίζουν, όπως η ευρεία κάλυψη, η ευρωστία, η εύκολη συντήρηση και το χαμηλό κόστος.

1.2.1 Αρχιτεκτονική ασυρμάτων δικτύων πλέγματος

Τα ασύρματα δίκτυα πλέγματος περιλαμβάνουν δύο είδη κόμβων, τους δρομολογητές πλέγματος και τους χρήστες πλέγματος με τα εξής χαρακτηριστικά:

- *Δρομολογητές πλέγματος (mesh routers):*
Οι δρομολογητές πλέγματος διαθέτουν τις ίδιες λειτουργίες γεφύρωσης και χρήσης πυλών με τους δρομολογητές των συμβατικών ασύρματων δικτύων και μάλιστα κατασκευάζονται με βάση τον ίδιο υλικό εξοπλισμό. Εντούτοις, οι δρομολογητές πλέγματος είναι εξοπλισμένοι με πρόσθετες εξελιγμένες δυνατότητες δρομολόγησης και επιτυγχάνουν την ίδια κάλυψη με τους

συμβατικούς δρομολογητές μεταδίδοντας υπό πολύ χαμηλότερη ισχύ εκπομπής, διότι εκμεταλεύονται τη δυνατότητά τους για πολυδιαδρομική δρομολόγηση (multipath routing). Οι πολλαπλές ασύρματες διεπαφές που διαθέτουν προσδίδουν στα δίκτυα πλέγματος υψηλή ευελιξία και επιτρέπουν την εύκολη διασύνδεσή τους τόσο με άλλα είδη δικτύων όσο και με συμβατικούς χρήστες. Οι δρομολογητές πλέγματος, τέλος, παρουσιάζουν μειωμένη κινητικότητα έχοντας ως κύρια λειτουργία τη δόμηση ενός κύριου κορμού-δικτύου στο οποίο να μπορούν να συνδέονται οι χρήστες πλέγματος.

- *Χρήστες πλέγματος(mesh clients):*

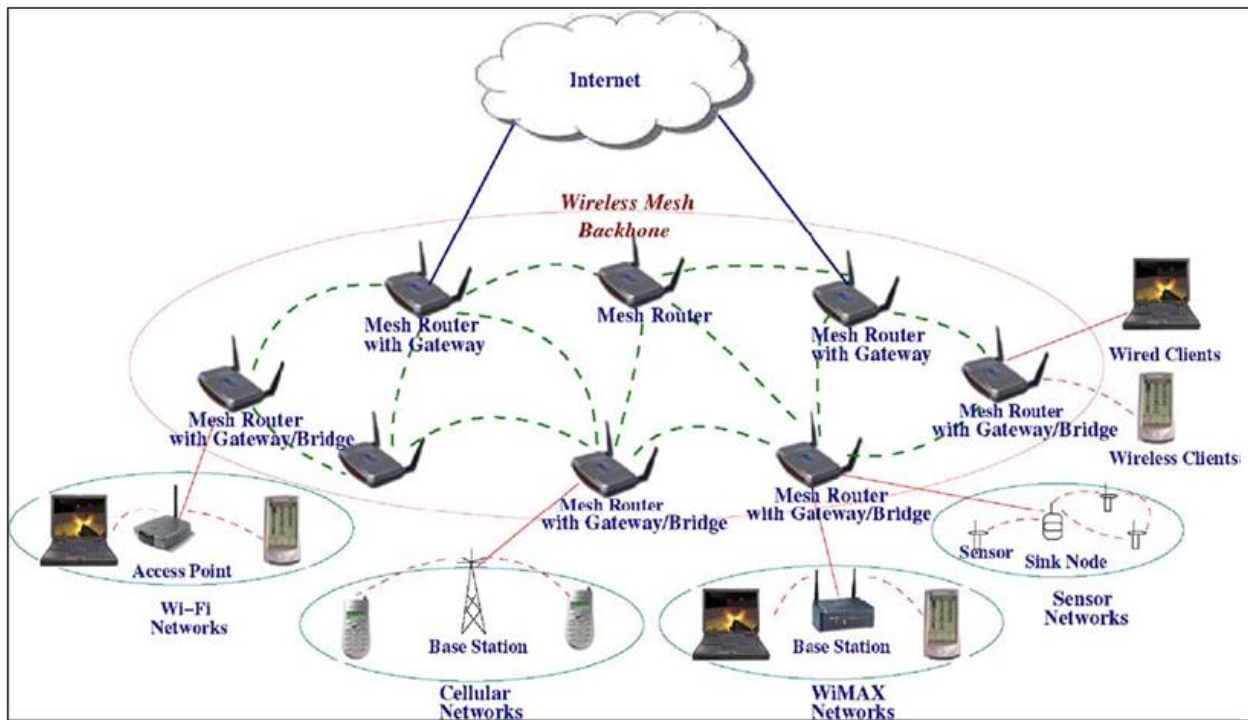
Οι χρήστες πλέγματος διαθέτουν εξελιγμένες λειτουργίες δικτύωσης σε σχέση με τους συμβατικούς χρήστες. Για το λόγο αυτό μπορούν να λειτουργούν και ως δρομολογητές σε ένα δίκτυο WMN. Από την άλλη πλευρά, το λογισμικό και το υλικό τους είναι πολύ απλούστερα σε σχέση με αυτά των δρομολογητών πλέγματος και συνήθως διαθέτουν μόνο μία ασύρματη διεπαφή.

Η αρχιτεκτονική των ασυρμάτων δικτύων πλέγματος μπορεί να διακριθεί σε τρεις κατηγορίες, αντίστοιχα προς τις λειτουργίες που επιτελούν οι δρομολογητές πλέγματος και οι χρήστες πλέγματος στο δίκτυο.

1. Ασύρματα δίκτυα πλέγματος κορμού (Infrastructure/Backbone WMNs)

Η αρχιτεκτονική αυτή είναι ευρέως χρησιμοποιούμενη. Οι δρομολογητές πλέγματος σχηματίζουν τη βασική υποδομή του δικτύου, όπου μπορούν να συνδέονται οι χρήστες. Η υποδομή αυτή, η οποία συνηθέστερα δομείται με βάση την τεχνολογία IEEE 802.11, αποτελεί το βασικό κορμό ο οποίος επιτρέπει τη συνδεσιμότητα μεταξύ δικτύων πλέγματος και ήδη υπάρχοντων ασυρμάτων δικτύων. Η συνδεσιμότητα αυτή επιτυγχάνεται μέσω λειτουργιών γεφύρωσης (στρώμα ζεύξης δεδομένων OSI) και δρομολόγησης (στρώμα δικτύου OSI) οι οποίες επιτελούνται από τους δρομολογητές πλέγματος του δικτύου. Οι τελευταίοι, αφενός, σχηματίζουν μεταξύ τους ζεύξεις οι οποίες οργανώνονται (self-organized) και βελτιστοποιούνται (self-optimized) αυτόματα, και, αφετέρου, είναι υπεύθυνοι για τη σύνδεση στο διαδίκτυο μέσω πυλών. Κατά την αρχιτεκτονική αυτή, οι χρήστες συνδέονται με τους δρομολογητές πλέγματος είτε

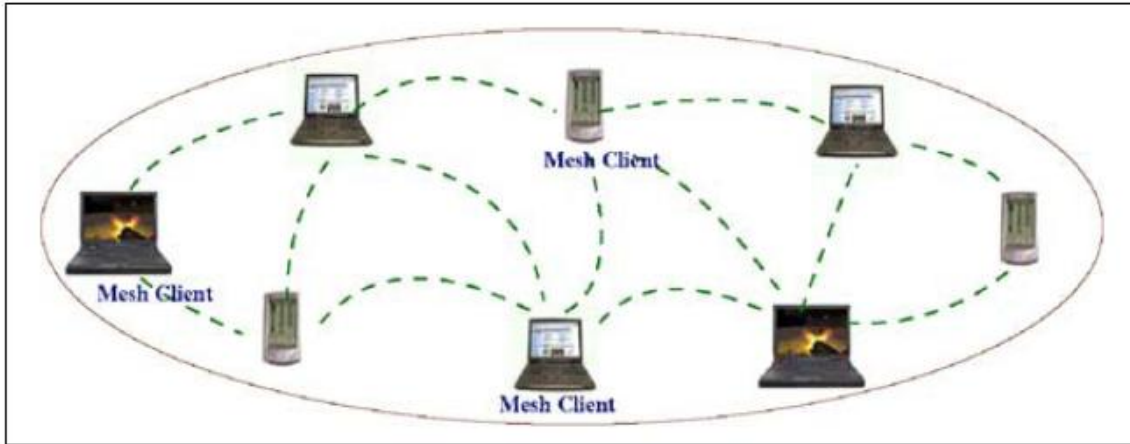
απευθείας, αν διαθέτουν διεπαφές της ίδιας τεχνολογίας με αυτούς, είτε μέσω των σταθμών βάσης.



Σχήμα 1.1: Ασύρματο δίκτυο πλέγματος – αρχιτεκτονική βασικού κορμού

2. Ασύρματα δίκτυα πλέγματος χρηστών (Client WMNs)

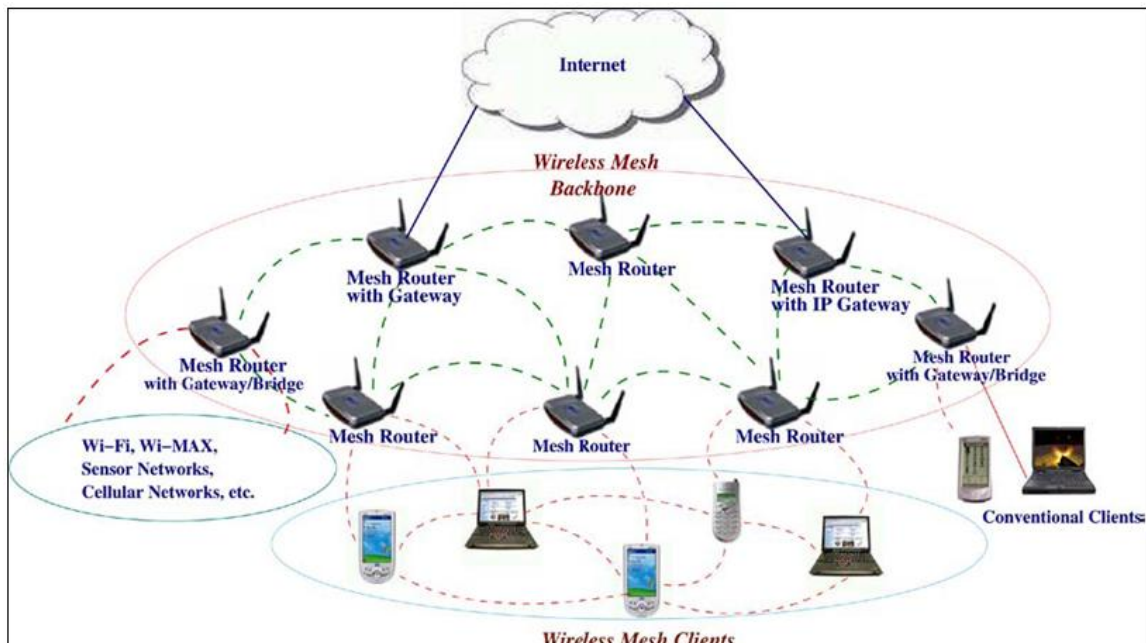
Η αρχιτεκτονική αυτή βασίζεται στη δημιουργία ομότιμης δικτύωσης (peer-to-peer network) μεταξύ των συσκευών των χρηστών. Οι κόμβοι των χρηστών είναι αυτοί που εκτελούν τις λειτουργίες της δρομολόγησης και της συγκρότησης του δικτύου και επιπλέον παρέχουν εφαρμογές τελικού χρήστη. Εφόσον χρησιμοποιείται μόνο ένα είδος ραδιοτεχνολογίας στους κόμβους που τα απαρτίζουν, σε αυτά τα είδη δικτύων δεν απαιτείται η παρουσία δρομολογητών πλέγματος. Η δομή τους θυμίζει πολύ τη δομή των συμβατικών adhoc δικτύων αλλά τα τερματικά οφείλουν να διαθέτουν αυξημένη ευφυΐα καθώς πρέπει να επιτελούν πρόσθετες και πλέον σύνθετες λειτουργίες.



Σχήμα 1. 2: Ασύρματο δίκτυο πλέγματος χρηστών

3. Υβριδικά ασύρματα δίκτυα πλέγματος (Hybrid WMNs)

Η αρχιτεκτονική αυτή συνδυάζει με βέλτιστο τρόπο τις αρχιτεκτονικές βασικού κορμού και πελατών. Οι χρήστες έχουν πρόσβαση στο δίκτυο είτε μέσω των δρομολογητών πλέγματος είτε απευθείας μέσω των άλλων χρηστών. Η ύπαρξη των δρομολογητών πλέγματος επιτρέπει την εύκολη σύνδεση με άλλα είδη δικτύων (διαδίκτυο, WiMax) ενώ οι αυξημένες δυνατότητες των τερματικών των χρηστών προσδίδουν στο δίκτυο μεγαλύτερη συνδεσιμότητα και ποιότητα κάλυψης.



Σχήμα 1.3: Ασύρματο δίκτυο πλέγματος υβριδικής αρχιτεκτονικής

1.2.2 Χαρακτηριστικά γνωρίσματα ασυρμάτων δικτύων πλέγματος

Τα δίκτυα WMN διαθέτουν αρκετά χαρακτηριστικά γνωρίσματα που τους προσδίδουν εξαιρετικές δυνατότητες εφαρμογών σε σχέση με τα κλασικά adhoc δίκτυα. Ειδικότερα, αν θεωρηθεί η περίπτωση του υβριδικού σχήματος αρχιτεκτονικής που συνδυάζει όλα τα πλεονεκτήματα των δικτύων WMN, προκύπτουν τα εξής χαρακτηριστικά:

- Τα δίκτυα WMN υποστηρίζουν κλασική adhoc δικτύωση και έχουν τη δυνατότητα να σχηματίζονται και να οργανώνονται με αυτόματο τρόπο. Επιπλέον, διαθέτουν μηχανισμούς (self-healing) που τους επιτρέπουν να αντιμετωπίζουν αποδοτικά τα προβλήματα που μπορεί να προκύψουν στο δίκτυο, όπως, για παράδειγμα, ενδεχόμενες διακοπές ζεύξεων.
- Η κινητικότητα διαφέρει ανάλογα με το είδος του κόμβου. Οι χρήστες πλέγματος παρουσιάζουν αυξημένη κινητικότητα. Αντίθετα οι δρομολογητές πλέγματος παρουσιάζουν μειωμένη κινητικότητα και εκτελούν πολύπλοκες λειτουργίες με συνέπεια να μην επιβαρύνονται οι χρήστες πλέγματος και οι λοιποί τερματικοί κόμβοι.
- Οι απαιτήσεις ενέργειας διαφέρουν ανάλογα με το είδος του κόμβου. Οι δρομολογητές πλέγματος δεν έχουν ιδιαίτερους περιορισμούς στην κατανάλωση ενέργειας, σε αντίθεση με τους χρήστες πλέγματος οι οποίοι πρέπει να είναι εφοδιασμένοι με πρωτόκολλα αποδοτικής διαχείρισης ενέργειας.
- Οι δρομολογητές πλέγματος υποστηρίζουν τόσο ασύρματη όσο και ενσύρματη πρόσβαση στα δίκτυα WMN. Επιπλέον, τα δίκτυα WMN είναι συμβατά με τα υπάρχοντα ασύρματα δίκτυα και έχουν τη δυνατότητα να συνδέονται με αυτά.
- Τα δίκτυα WMN υποστηρίζουν δρομολόγηση πολλαπλών βημάτων (multihop routing). Συνεπώς, επιτυγχάνεται ευρύτερη γεωγραφική κάλυψη και υψηλότερη τιμή διέλευσης πληροφορίας ενώ μειώνονται οι παρεμβολές.

1.2.3 Σχεδιαστικές απαιτήσεις ασυρμάτων δικτύων πλέγματος

Στη φάση σχεδίασης ενός ασύρματου δικτύου πλέγματος απαιτείται να δοθεί προσοχή στους ακόλουθους παράγοντες, οι οποίοι επηρεάζουν σημαντικά την απόδοση του δικτύου:

1.Τεχνικές ραδιοεπικοινωνίας:

Η πρόοδος στον τομέα των τηλεπικοινωνιών είναι ραγδαία τα τελευταία χρόνια, με αποτέλεσμα να ανακλύπτον διαρκώς νέες ραδιοτεχνολογίες οι οποίες με την εφαρμογή τους στα ασύρματα δίκτυα βελτιστοποιούν την απόδοσή τους. Ήδη χρησιμοποιούνται έξυπνες κεραιές (smart antennas) και συστήματα πολλαπλής εισόδου/πολλαπλής εξόδου (MIMO) για την αύξηση της χωρητικότητας του δικτύου. Τεχνικές, όπως οι γνωστικές ραδιοεπικοινωνίες (cognitive radio) που προσφέρουν ευελιξία καθώς και αυξημένες δυνατότητες για έλεγχο του δικτύου, είναι ακόμη σε αρχικό στάδιο αλλά αναμένεται να αποτελέσουν τη βάση για την περαιτέρω εκμετάλλευση των ασυρμάτων δικτύων. Η ανάπτυξη των προαναφερθεισών τεχνικών ραδιοεπικοινωνίας απαιτεί προσεκτική σχεδίαση των πρωτοκόλλων των ανώτερων στρωμάτων και ιδιαίτερα των πρωτοκόλλων δρομολόγησης (routing protocols) και πρόσβασης στο κοινό μέσο (MAC).

2.Κλιμακωσιμότητα:

Η κλιμακωσιμότητα είναι αναγκαία καθώς, χωρίς την εξασφάλισή της, χειροτερεύει αισθητά η απόδοση των δικτύων WMN, ιδιαίτερα όσο αυξάνεται το μέγεθός τους. Όλα τα πρωτόκολλα, λοιπόν, που εφαρμόζονται στα δίκτυα WMN, από το φυσικό στρώμα μέχρι το στρώμα εφαρμογών, πρέπει να είναι κλιμακώσιμα, ώστε να λειτουργούν αποδοτικά.

3.Συνδεσιμότητα πλέγματος:

Τα περισσότερα πλεονεκτήματα των δικτύων WMN απορρέουν από τη συνδεσιμότητα του πλέγματος και για το λόγο αυτό απαιτείται να είναι αξιόπιστη. Προς τούτο αναπτύσσονται αλγόριθμοι ελέγχου της τοπολογίας και οργάνωσης του δικτύου. Επιπλέον, πρωτόκολλα του στρώματος δρομολόγησης και του φυσικού στρώματος που έχουν γνώση της τοπολογίας του δικτύου βελτιώνουν σε μεγάλο βαθμό την απόδοση του δικτύου.

4.Ευρυζωνικότητα και ποιότητα υπηρεσιών:

Σε αντίθεση με τα συμβατικά adhoc δίκτυα, οι περισσότερες εφαρμογές των δικτύων WMN είναι ευρυζωνικές με διαφορετικές απαιτήσεις ποιότητας υπηρεσιών έκαστη.

Συνεπώς, εκτός από τα συνήθη μέτρα επίδοσης, όπως η από-άκρο-σε-άκρο καθυστέρηση (end-to-end transmission delay) και η δικαιοσύνη (fairness), απαιτούνται πρόσθετα κριτήρια. Ειδικότερα, τα πρωτόκολλα επικοινωνίας πρέπει να λαμβάνουν υπόψη τους μεγέθη όπως η διακύμανση της καθυστέρησης μετάδοσης, το ποσοστό απωλειών πακέτων, η ρυθμαπόδοση ανά κόμβο.

5.Χρηστικότητα:

Η αυξημένη χρηστικότητα θα επιτρέψει την ταχύτερη ανάπτυξη των δικτύων WMN. Για να συμβεί αυτό πρέπει τα διάφορα πρωτόκολλα που θα χρησιμοποιηθούν να σχεδιαστούν με τέτοιο τρόπο ώστε να επιτρέπουν στο δίκτυο να λειτουργεί όσο το δυνατόν περισσότερο αυτόνομα. Η αυτονομία στο δίκτυο απαιτείται κυρίως ως προς τη διαχείριση ισχύος, το δυναμικό έλεγχο της τοπολογίας, την εξασφάλιση της ευρωστίας όταν υπάρχουν διακοπές ζεύξεων και γενικότερα την οργάνωση του δικτύου και την ορθή ρύθμιση των παραμέτρων του. Για την εξασφάλιση της αυτονομίας του δικτύου στους προαναφερθέντες τομείς και την επίβλεψη της απόδοσής του απαιτείται η ανάπτυξη ειδικών μηχανισμών διαχείρισης.

6.Ασφάλεια

Καίτοι υπάρχουν αρκετά σχήματα ασφάλειας που χρησιμοποιούνται σε δίκτυα LAN, κανένα από αυτά δεν είναι ακόμα πλήρως εφαρμόσιμο στα δίκτυα WMN. Τα σχήματα ασφάλειας που χρησιμοποιούνται στα adhoc δίκτυα δε θα μπορούσαν να υιοθετηθούν ως λύση καθώς τα περισσότερα από αυτά δεν μπορούν να υλοποιηθούν πρακτικά. Το πρόβλημα οφείλεται στην κατανεμημένη αρχιτεκτονική των δικτύων WMN, η οποία δεν επιτρέπει την ύπαρξη μιας κεντρικής αρχής που θα διαχειρίζεται τα θέματα ασφάλειας.

7.Συμβατότητα και διαλειτουργικότητα:

Στα δίκτυα WMN πρέπει να υπάρχει πρόσβαση τόσο των χρηστών πλέγματος όσο και συμβατικών χρηστών. Για το λόγο αυτό απαιτείται συμβατότητα του δικτύου με τους τελευταίους. Επιπλέον, απαιτείται η επικοινωνία των δικτύων WMN με άλλα ασύρματα δίκτυα και άρα η ύπαρξη αρκετών ασύρματων διεπαφών στους δρομολογητές πλέγματος και η ανάπτυξη εξελιγμένων λειτουργιών γεφύρωσης.

1.2.4 Εφαρμογές και δυνατότητες ανάπτυξης ασυρμάτων δικτύων πλέγματος

Τα δίκτυα WMN κατηγοριοποιούνται σε ανοικτά και κλειστά δίκτυα. Οποιοσδήποτε χρήστης, με την κατάλληλη ασύρματη διεπαφή, μπορεί να συνδεθεί σε ένα ανοικτό δίκτυο. Σε ένα κλειστό δίκτυο, όμως, υπάρχει κεντρική αρχή που διαχειρίζεται την αποδοχή των χρηστών και επιτρέπει την πρόσβαση τους στο δίκτυο μέσω επαλήθευσης της ταυτότητάς τους (authentication).

Πιθανές εφαρμογές των ανοικτών δικτύων WMN είναι οι εξής:

- *Δικτύωση σε επίπεδο γειτονιάς*
Προσφέρεται εύκολη σύνδεση των χρηστών στο διαδίκτυο με χρήση πυλών και, επιπλέον, δίνεται η δυνατότητα ανάπτυξης εφαρμογών peer-to-peer μεταξύ γειτονικών δικτύων.
- *Συγκρότηση μητροπολιτικών δικτύων (Metropolitan Area Networks -MANs)*
Λόγω της κλιμακωσιμότητας και της ευρείας κάλυψης που μπορούν να παράσχουν τα δίκτυα WMN, είναι η ιδανική λύση για την παροχή ευρυζωνικών υπηρεσιών σε μεγαλύτερες γεωγραφικές ζώνες, όπως ολόκληρες κοινότητες ή ακόμη και πόλεις.
- *Δικτύωση στα μέσα μεταφοράς*
Τα δίκτυα WMN υποστηρίζουν υπηρεσίες όπως η πληροφόρηση των επιβατών και η επικοινωνία μεταξύ των οδηγών. Προς τούτο απαιτείται η ύπαρξη κινητού δικτύου WMN εντός του οχήματος και ταχύτατης σύνδεσης του οχήματος στο διαδίκτυο.

Πιθανές εφαρμογές των κλειστών δικτύων WMN είναι οι εξής:

- *Ευρυζωνική δικτύωση σε επίπεδο οικίας*
Μέχρι τώρα, η ευρυζωνική δικτύωση σε επίπεδο οικίας πραγματοποιείται μέσω της τεχνολογίας IEEE 802.11 WLAN, με σημεία πρόσβασης σε θέσεις που αφήνουν πολλά νεκρά σημεία, δηλαδή σημεία χωρίς κάλυψη. Λύση στα προαναφερθέντα προβλήματα μπορούν να δώσουν τα δίκτυα WMN αντικαθιστώντας τα σημεία πρόσβασης με δρομολογητές πλέγματος. Επιπλέον, με την παρούσα τεχνολογία, η επικοινωνία μεταξύ των κόμβων στο δίκτυο εντός της οικίας δε συμβαίνει απευθείας αλλά πραγματοποιείται με τη μεσολάβηση

ενός διανομέα (hub). Ο διανομέας στέλνει τα μηνύματα που λαμβάνει από έναν κόμβο σε όλους τους κόμβους του δικτύου και συνεπώς προκύπτουν προβλήματα συμφόρησης. Τα δίκτυα WMN δίνουν λύση και σε αυτό το πρόβλημα, επιτρέποντας στους κόμβους να επικοινωνούν απευθείας μεταξύ τους.

- *Δικτύωση σε εταιρικό επίπεδο*

Τα δίκτυα WMN μπορούν να χρησιμοποιηθούν για τη σύνδεση των διαφορετικών ασυρμάτων δικτύων εντός μιας εταιρίας, αποφεύγοντας έτσι το αυξημένο κόστος χρήσης ενσύρματου Ethernet για το σκοπό αυτό.

- *Δίκτυωση για στρατιωτικές εφαρμογές*

Στις στρατιωτικές εφαρμογές απαιτείται η ασφαλής και αξιόπιστη επικοινωνία μεταξύ συγκεκριμένων κόμβων. Τα δίκτυα WMN μπορούν να χρησιμοποιηθούν για τέτοιου είδους εφαρμογές προσφέροντας, μάλιστα, ευρύτερη γεωγραφική κάλυψη.

- *Δικτύωση σε πανεπιστημιακούς χώρους*

Τα σημεία πρόσβασης των δικτύων WMN μπορούν να αξιοποιηθούν σε πανεπιστημιακούς χώρους και να προσφέρουν συνολική δικτυακή κάλυψη με πολύ χαμηλό κόστος.

1.3 Δρομολόγηση σε ασύρματα δίκτυα πλέγματος

Δρομολόγηση (routing) είναι η διαδικασία σύμφωνα με την οποία επιλέγεται μια διαδρομή σε ένα δίκτυο, μέσω της οποίας αποστέλλονται δεδομένα. Η διαδικασία αυτή εκτελείται στο στρώμα δικτύου (OSI Network Layer) με εφαρμογή διάφορων πρωτοκόλλων δρομολόγησης (routing protocols), δηλαδή κανόνων που καθορίζουν τον τρόπο επικοινωνίας και ανταλλαγής μηνυμάτων μεταξύ των δρομολογητών του δικτύου.

1.3.1 Σχεδιαστικές απαιτήσεις πρωτοκόλλων δρομολόγησης για εφαρμογή σε ασύρματα δίκτυα πλέγματος

Ορισμένα πρωτόκολλα δρομολόγησης που εφαρμόζονται σε συμβατικά adhoc δίκτυα είναι εφαρμόσιμα και στα δίκτυα WMN λόγω των πολλών κοινών χαρακτηριστικών που παρουσιάζουν τα δύο είδη δικτύων. Εντούτοις, τα υπάρχοντα πρωτόκολλα παρουσιάζουν μικρή δυνατότητα κλιμάκωσης και δεν λαμβάνουν υπόψη τις διαφορετικές απαιτήσεις

ενέργειας και κινητικότητας των κόμβων. Για το λόγο αυτό είναι αναγκαία η δημιουργία νέων πρωτοκόλλων δρομολόγησης, ειδικά σχεδιασμένων για εφαρμογή σε δίκτυα WMN, ώστε να καλύπτουν τις ακόλουθες απαιτήσεις:

- *Πολλαπλά κριτήρια επίδοσης*

Η διαδρομή δρομολόγησης στα δίκτυα WMN είναι προτιμητέο να επιλέγεται με βάση περισσότερα του ενός κριτήρια επίδοσης. Τα συνήθη κριτήρια επίδοσης είναι:

1. Το πλήθος βημάτων μεταξύ πηγής και προορισμού (minimum hop count), σύμφωνα με το οποίο επιλέγεται η διαδρομή που απαιτεί το ελάχιστο πλήθος βημάτων. Δεν αποδίδει ως μεμονωμένο κριτήριο καθώς δεν λαμβάνει υπόψη την πιθανότητα κακής ποιότητας ή διακοπής μιας ζεύξης μεταξύ κόμβων που συνιστούν την ελάχιστη διαδρομή.

2. Το αναμενόμενο πλήθος αναμεταδόσεων (Expected Transmission Count -ETX), σύμφωνα με το οποίο επιλέγεται η διαδρομή που απαιτεί τις λιγότερες αναμεταδόσεις πακέτων για την επιτυχή εκπομπή τους. Μία εκπομπή θεωρείται επιτυχής, όταν το μεταδιδόμενο σήμα δεν υφίσταται συγκρούσεις ή διαλείψεις και λαμβάνεται, τελικά, από τον κόμβο προορισμό.

3. Ο αναμενόμενος χρόνος εκπομπής (Estimated Transmission Time-ETT), σύμφωνα με το οποίο υπολογίζεται ο μέσος χρόνος που απαιτείται για την επιτυχή εκπομπή ενός πακέτου και επιλέγεται να δρομολογηθεί το πακέτο στη ζεύξη εκείνη που αντιστοιχεί στον ελάχιστο ETT.

4. Η κατανάλωση ενέργειας (Energy consumption), σύμφωνα με το οποίο επιλέγεται η διαδρομή η οποία απαιτεί τη χαμηλότερη κατανάλωση ενέργειας. Το κριτήριο αυτό εφαρμόζεται κυρίως σε κόμβους που έχουν υψηλούς περιορισμούς ενέργειας όπως είναι οι χρήστες πλέγματος που χαρακτηρίζονται από αυξημένη κινητικότητα.

5. Χρόνος μετ' επιστροφής διαδρομής (Round Trip Time -RTT), σύμφωνα με το οποίο επιλέγονται οι ζεύξεις που εμφανίζουν ελάχιστο RTT. Το μέγεθος RTT εκφράζει το χρόνο που απαιτείται για να μεταδοθούν τα δεδομένα από την πηγή στον προορισμό και να σταλεί επιβεβαίωση λήψης τους από τον προορισμό στην πηγή. Το κριτήριο αυτό χρησιμοποιείται συχνά σε συνδυασμό με το κριτήριο του

ελάχιστου αριθμού βημάτων όταν παρουσιάζεται πρόβλημα συμφόρησης στο δίκτυο.

6. Διαθεσιμότητα/ποιότητα ζεύξης (link quality), σύμφωνα με το οποίο επιλέγονται οι ζεύξεις μεταξύ πηγής και προορισμού που είναι διαθέσιμες στο μεγαλύτερο ποσοστό του χρόνου.

- *Ανοχή σε προβλήματα ζεύξεων*

Για την περαιτέρω εκμετάλλευση των δυνατοτήτων των δικτύων WMN τα πρωτόκολλα δρομολόγησης πρέπει να αντιμετωπίζουν αποδοτικά την πιθανή διακοπή ή μη διαθεσιμότητα κάποιας ζεύξης. Η άμεση εύρεση εναλλακτικών διαδρομών σε περίπτωση μη διαθεσιμότητας μιας ζεύξης είναι απαραίτητη για την ευρωστία του δικτύου και την εξασφάλιση υψηλής ποιότητας υπηρεσιών.

- *Εξισορρόπηση φορτίου*

Ένας από τους στόχους των δικτύων WMN είναι η ισορροπημένη κατανομή των πόρων του δικτύου στους κόμβους που το απαρτίζουν. Ο στόχος αυτός πρέπει να λαμβάνεται υπόψη από τα πρωτόκολλα δρομολόγησης και όταν παρατηρείται συμφόρηση σε κάποιο τμήμα του δικτύου, η δρομολόγηση να πραγματοποιείται μέσω εναλλακτικών διαδρομών.

- *Κλιμακωσιμότητα*

Η ιδιότητα αυτή είναι απαραίτητη για πρωτόκολλα δρομολόγησης που εφαρμόζονται σε δίκτυα WMN καθώς το μέγεθος των δικτύων αυτών είναι μεταβλητό και ενδεχομένως μεγάλο. Η επιτυχής δρομολόγηση σε μεγάλα τέτοιου τύπου δίκτυα μπορεί να διαρκέσει επί μεγάλο χρονικό διάστημα αυξάνοντας την από-άκρο-σε-άκρο καθυστέρηση. Επιπλέον, ακόμη και αν προσδιοριστεί η βέλτιστη διαδρομή δρομολόγησης, η κατάσταση των κόμβων που συμμετέχουν σε αυτή μπορεί να μεταβάλλεται δυναμικά.

- *Διαφοροποιημένη αντιμετώπιση των δρομολογητών και των χρηστών πλέγματος*

Όπως έχει ήδη αναφερθεί, οι δρομολογητές πλέγματος, εμφανίζουν μειωμένη κινητικότητα και δεν υπόκεινται σε περιορισμούς ως προς την κατανάλωση ενέργειας. Συνεπώς, για την επικοινωνία τους μπορεί να χρησιμοποιηθεί ένα πρωτόκολλο δρομολόγησης απλούστερο από αυτά που εφαρμόζονται στα ασύρματα adhoc δίκτυα. Αντιθέτως, λόγω της αυξημένης κινητικότητας και των

περιορισμών ενέργειας, οι χρήστες πλέγματος απαιτούν ένα πλέον σύνθετο πρωτόκολλο δρομολόγησης.

- *Διαστρωματική αλληλεπίδραση (cross-layer interaction)*

Τα υπάρχοντα πρωτόκολλα δρομολόγησης δεν λαμβάνουν υπόψη τις διαδικασίες που πραγματοποιούνται στο αμέσως κατώτερο στρώμα MAC (Medium Access Control layer), όπως ορίζει η διαστρωμάτωση κατά OSI. Εντούτοις, η συνεργασία με το στρώμα MAC μπορεί να βελτιώσει σε μεγάλο βαθμό την απόδοση των αλγορίθμων δρομολόγησης που λαμβάνουν χώρα στο στρώμα δικτύου.

1.3.1 Πρωτόκολλα δρομολόγησης σε δίκτυα WMN

Ακολουθεί η περιγραφή ορισμένων από τα γνωστότερα πρωτόκολλα δρομολόγησης που χρησιμοποιούνται σε δίκτυα WMN:

1. Πολυδιαδρομική δρομολόγηση (multipath routing)

Η σχεδίαση ενός αποδοτικού πρωτοκόλλου πολυδιαδρομικής δρομολόγησης για εφαρμογή σε δίκτυα WMN απαιτεί προσεκτική μελέτη του δικτύου και κυρίως της υπηρεσίας την οποία αυτό θα παρέχει στους χρήστες. Οι βασικοί στόχοι όταν εφαρμόζεται πολυδιαδρομική δρομολόγηση είναι η ορθή κατανομή του φορτίου στο δίκτυο και η υψηλή ανθεκτικότητα σε προβλήματα μεταξύ ζεύξεων. Προσδιορίζονται πολλαπλές διαδρομές μεταξύ πηγής και προορισμού και τα δεδομένα αποστέλλονται σε μία από αυτές τις διαδρομές. Αν συμβεί διακοπή κάποιας ζεύξης της διαδρομής αυτής, είτε λόγω κακής ποιότητας είτε λόγω κινητικότητας κάποιου κόμβου, επιλέγεται μια νέα διαδρομή μεταξύ αυτών που προσδιορίστηκαν αρχικά. Κατ' αυτόν τον τρόπο, σε περίπτωση προβλήματος δεν υπολογίζεται εξαρχής νέα διαδρομή δρομολόγησης και παρατηρείται βελτίωση της από-άκρο-σε-άκρο καθυστέρησης και της ρυθμαπόδοσης του δικτύου. Προφανώς η δυνατότητα βελτίωσης σχετίζεται άμεσα με την ύπαρξη περισσότερων της μιας διαδρομών μεταξύ πηγής και προορισμού. Το κύριο μειονέκτημα της πολυδιαδρομικής δρομολόγησης είναι η αυξημένη πολυπλοκότητα που εμφανίζει.

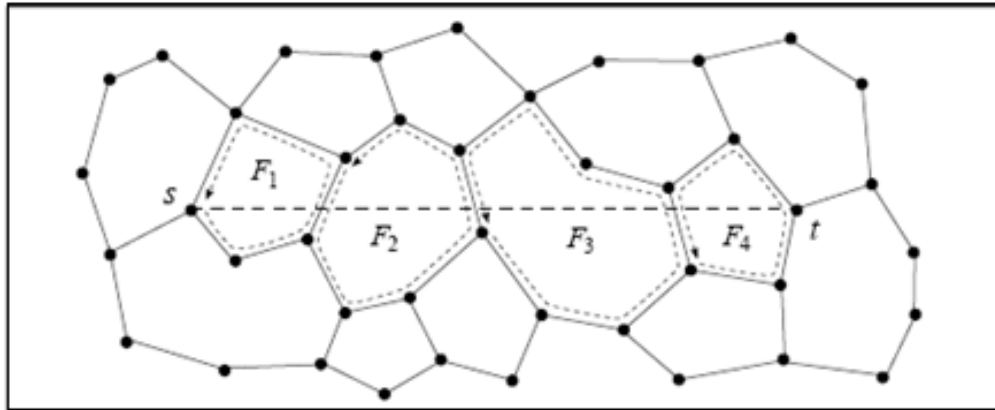
υπόλοιπους κόμβους της ομάδας, οπότε και προκαλείται συμφόρηση. Αν αποφευχθεί αυτό, η ιεραρχική δρομολόγηση αποτελεί μια καλή λύση για δίκτυα WMN καθώς παρέχει και την απαιτούμενη κλιμακωσιμότητα.

3.Γεωγραφική δρομολόγηση (geographic routing)

Σε αντίθεση με τα κλασικά σχήματα δρομολόγησης που υπολογίζουν τη βέλτιστη διαδρομή στηριζόμενα στη γνώση της τοπολογίας του δικτύου, τα σχήματα γεωγραφικής δρομολόγησης προωθούν τα δεδομένα στους κόμβους του δικτύου με βάση μόνο τη γεωγραφική τους θέση. Συνεπώς, οι αλλαγές που συμβαίνουν στην τοπολογία του δικτύου κατά τη διάρκεια της αποστολής των δεδομένων επηρεάζουν πολύ λιγότερο το συγκεκριμένο σχήμα δρομολόγησης. Απαραίτητη προϋπόθεση για την εφαρμογή του πρωτοκόλλου είναι κάθε κόμβος του δικτύου να είναι σε θέση να προσδιορίζει τις γεωγραφικές συντεταγμένες του και η πηγή να γνωρίζει τις συντεταγμένες του κόμβου-προορισμού.

Στα σχήματα γεωγραφικής δρομολόγησης υπάρχει η δυνατότητα τόσο δρομολόγησης σε μοναδική διαδρομή όσο και πολυδιαδρομικής δρομολόγησης. Τα σχήματα εύρεσης μοναδικής διαδρομής βασίζονται κυρίως σε δύο τεχνικές: την άπληστη προώθηση (greedy forwarding) και τη δρομολόγηση που βασίζεται στο γράφο του δικτύου (planar graph routing). Σύμφωνα με την άπληστη προώθηση, τα δεδομένα αποστέλλονται από ένα κόμβο στο γειτονικό του μόνο εφόσον ο δεύτερος βρίσκεται πλησιέστερα στον κόμβο-προορισμό. Η άπληστη προώθηση πραγματοποιείται με βάση τοπικές πληροφορίες για το δίκτυο. Όταν δεν ανευρεθεί γειτονικός κόμβος ο οποίος βρίσκεται πλησιέστερα στον προορισμό, η δρομολόγηση αποτυγχάνει. Τη λύση στο πρόβλημα αυτό παρέχει ο αλγόριθμος planar graph routing, ο οποίος βασίζεται στο ότι οι ζεύξεις στο δίκτυο δημιουργούν ένα γράφο με ακμές που δεν τέμνονται μεταξύ τους (planar graph). Ένας τέτοιος γράφος είναι επίπεδος με κορυφές που σχηματίζουν κλειστές επιφάνειες (faces), όπως στο Σχ.1.5. Θεωρείται η γραμμή που συνδέει την πηγή (s) με τον προορισμό (t) και παρατηρείται ότι τέμνει διαδοχικές επιφάνειες του γράφου του δικτύου. Η δρομολόγηση των δεδομένων πραγματοποιείται ξεκινώντας από την πηγή και διασχίζοντας δεξιόστροφα τα σύνορα διαδοχικών επιφανειών του γράφου (face transversal). Όταν συναντηθεί η ακμή του γράφου που τέμνει την ευθεία πηγής-

προορισμού, κατά τη διάσχιση μιας επιφάνειας, ξεκινά η διάσχιση της γειτονικής επιφάνειας έως ότου προσπελαστεί η επιφάνεια στην οποία ανήκει ο κόμβος προορισμός.



Σχήμα 1.5: Δρομολόγηση βασισμένη στο γράφο του δικτύου

ΚΕΦΑΛΑΙΟ 2

ΜΑΡΚΟΒΙΑΝΕΣ ΔΙΑΔΙΚΑΣΙΕΣ ΑΠΟΦΑΣΗΣ

2.1 Εισαγωγή

Σκοπός του κεφαλαίου αυτού είναι η εξοικείωση του αναγνώστη με βασικές έννοιες της Μαρκοβιανής θεωρίας απόφασης που αποτελεί το βασικό θεωρητικό υπόβαθρο της διπλωματικής εργασίας. Αρχικά, ορίζεται το μαρκοβιανό μοντέλο και αναλύονται οι παράμετροι που το απαρτίζουν. Στη συνέχεια, παρουσιάζονται οι διάφοροι αλγόριθμοι επίλυσης του μοντέλου ανάλογα με τα χαρακτηριστικά του εκάστοτε υπό μελέτη συστήματος.

Στο σημείο αυτό παρατίθενται οι συμβολισμοί των μεγεθών που αναφέρονται στο κεφάλαιο αυτό:

\tilde{T} : σύνολο χρονικών στιγμών, ορίζοντας μαρκοβιανής διαδικασίας

\tilde{S} : σύνολο καταστάσεων συστήματος

\tilde{A} : σύνολο δυνατών δράσεων

r_t : όφελος/ απώλεια του συστήματος τη χρονική στιγμή t

$p_i(\cdot|s, a)$: πιθανότητα μετάβασης σε επόμενη κατάσταση τη χρονική στιγμή t

d_t : ντετερμινιστικός κανόνας απόφασης σύμφωνα με τον οποίο καθορίζεται μια δράση σε μία κατάσταση

$q_{d_t}(\cdot)$: πιθανοτικός κανόνας απόφασης σύμφωνα με τον οποίο καθορίζεται τυχαία μια δράση σε μία κατάσταση

MD: σύνολο ντετερμινιστικών κανόνων απόφασης

MR: σύνολο πιθανοτικών κανόνων απόφασης

π : πολιτική/ ακολουθία κανόνων απόφασης

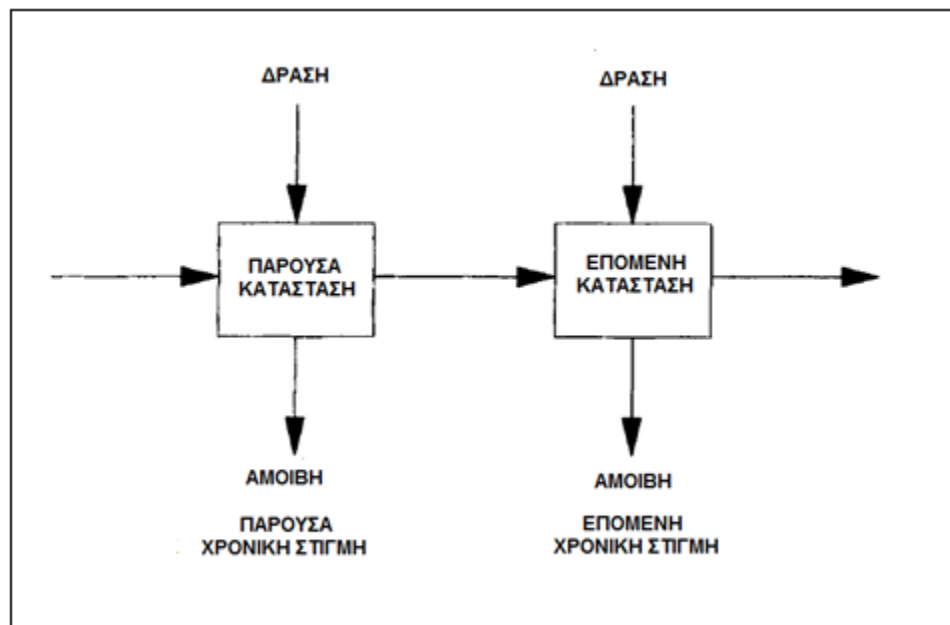
u_t : συνάρτηση οφέλους του συστήματος κατά τη χρονική στιγμή t

u^π : συνάρτηση οφέλους που προκύπτει από την εφαρμογή της πολιτικής π

2.2 Καθορισμός μοντέλου

Οι μαρκοβιανές διαδικασίες απόφασης αποτελούν τμήμα της γενικότερης θεωρίας αποφάσεων και χρησιμοποιούνται ευρέως για τη διατύπωση μοντέλων και τη μελέτη συστημάτων τα οποία εξελίσσονται στο χρόνο με αβεβαιότητα. Ειδικότερα, σύμφωνα με το μαρκοβιανό μοντέλο, σε συγκεκριμένη χρονική στιγμή ένας ελεγκτής επιβλέπει την τρέχουσα κατάσταση ενός υπό μελέτη συστήματος και με βάση αυτήν έχει τη δυνατότητα να επιλέξει μια δράση που θα προκαλέσει, με συγκεκριμένη πιθανότητα, τη μετάβαση σε μία νέα κατάσταση. Ταυτόχρονα, προσδιορίζεται το όφελος (reward) ή η απώλεια (loss) που σχετίζεται με την απόφαση που έλαβε ο ελεγκτής.

Η ουσιώδης διαφορά του μαρκοβιανού μοντέλου απόφασης από τα υπόλοιπα ακολουθιακά μοντέλα απόφασης είναι ότι το όφελος/απώλεια και οι πιθανότητες μετάβασης ικανοποιούν την ιδιότητα Markov, δηλαδή εξαρτώνται μόνο από την τρέχουσα κατάσταση στην οποία βρίσκεται το σύστημα και την τελευταία επιλεγείσα δράση και όχι από όλες τις προηγούμενες καταστάσεις και δράσεις του συστήματος.



Σχήμα 2.1: Εξέλιξη μαρκοβιανής διαδικασίας

Με βάση τα προηγούμενα, ένα μαρκοβιανό μοντέλο απόφασης μπορεί να καθοριστεί από το ακόλουθο πλήθος συνόλων:

$$\{T, S, A, r_i(s, a), p_i(\cdot|s, a)\} \quad (2.1)$$

1. $\{T\}$ το σύνολο από *χρονικές στιγμές/περιόδους* όπου λαμβάνονται οι αποφάσεις
2. $\{S\}$ το σύνολο από *καταστάσεις* στις οποίες μπορεί να βρεθεί το σύστημα
3. $\{A\}$ το σύνολο από *διαθέσιμες δράσεις*
4. $\{r_i(s,a)\}$ το σύνολο από *οφέλη ή απώλειες* οι οποίες εξαρτώνται άμεσα από την επιλεγμένη δράση κάθε φορά καθώς και από την παρούσα κατάσταση στην οποία βρίσκεται το σύστημα
5. $\{p_i(\cdot|s,a)\}$ το σύνολο από *πιθανότητες μετάβασης* μεταξύ καταστάσεων του συστήματος που επίσης εξαρτώνται από την επιλεγμένη δράση και την τρέχουσα κατάσταση του συστήματος

Με χρήση των ανωτέρω οντοτήτων, για να προσδιοριστεί πλήρως μια μαρκοβιανή διαδικασία απόφασης απαιτείται να τεθεί ένας συγκεκριμένος στόχος προς ικανοποίηση, ο οποίος συνήθως είναι η μεγιστοποίηση (ελαχιστοποίηση) κάποιας αθροιστικής συνάρτησης των οφελών (των απωλειών) του συστήματος. Ο στόχος αυτός αποτελεί βασικό κριτήριο για τον προσδιορισμό της κατάλληλης πολιτικής που πρέπει να ακολουθήσει ο ελεγκτής του συστήματος.

2.2.1 Ανάλυση παραμέτρων του μαρκοβιανού μοντέλου

Η κατανόηση των παραμέτρων του μαρκοβιανού μοντέλου είναι επιτακτική καθώς η επιλογή των καταλλήλων παραμέτρων αποτελεί το πρώτο βήμα για την ανάλυση και τη βελτιστοποίηση ενός συστήματος. Προς τούτο, ακολουθεί εκτενής παρουσίαση των συνόλων που απαρτίζουν το μαρκοβιανό μοντέλο:

- *Χρονικές στιγμές/ περίοδοι λήψης των αποφάσεων $\{T\}$*

Το σύνολο αυτό είναι υποσύνολο της ευθείας των μη αρνητικών πραγματικών αριθμών και μπορεί να κατηγοριοποιηθεί ως εξής:

1. Διακριτό ή συνεχές
2. Πεπερασμένο ή άπειρο

Στην περίπτωση όπου υπάρχει διακριτό σύνολο T , η οποία είναι και η συνηθέστερη, ο χρόνος διαιρείται σε περιόδους ή φάσεις και η λήψη των αποφάσεων γίνεται, κατά σύμβαση, στην αρχή κάθε περιόδου. Η μαθηματική περιγραφή του συνόλου είναι $T=\{1,2,3,\dots,N\}$, όπου N πεπερασμένο στην περίπτωση του προβλήματος πεπερασμένου ορίζοντα και $N \rightarrow \infty$ στην περίπτωση του προβλήματος άπειρου ορίζοντα. Όταν N πεπερασμένο γίνεται η παραδοχή ότι δεν λαμβάνεται απόφαση κατά την τελευταία χρονική περίοδο $t=N$.

Στην αντίθετη περίπτωση όπου το σύνολο T είναι συνεχές, οι αποφάσεις μπορούν να ληφθούν με ένα από τους τρεις ακόλουθους τρόπους:

1. Καθόλη τη διάρκεια του T , δηλαδή κατά συνεχή τρόπο
2. Σε χρονικές στιγμές όπου συμβαίνουν γεγονότα ιδιαίτερου ενδιαφέροντος για το υπό μελέτη σύστημα
3. Σε τυχαίες χρονικές στιγμές που επιλέγονται από τον ελεγκτή του συστήματος.

- *Σύνολα καταστάσεων $\{S\}$ και διαθέσιμων δράσεων $\{A\}$:*

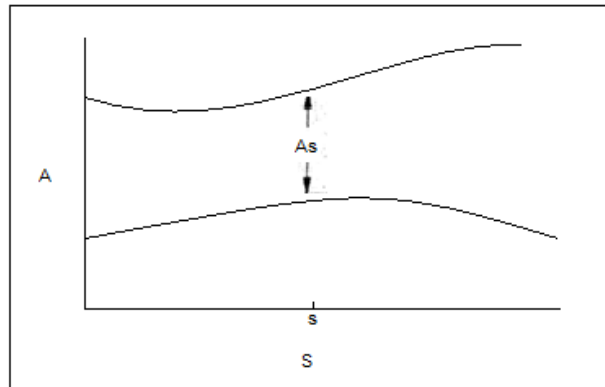
Αρχικά ισχύει η παραδοχή ότι τα στοιχεία των συνόλων $\{S\}$ και $\{A\}$ δεν μεταβάλλονται με το χρόνο. Το σύνολο $\{S\}$ περιλαμβάνει το πλήθος των δυνατών καταστάσεων στις οποίες μπορεί να βρεθεί το σύστημα κατά τη διάρκεια του χρόνου στον οποίο μελετάται. Μια κατάσταση s του συστήματος μπορεί να είναι ντετερμινιστικά γνωστή στον ελεγκτή αλλά, συνηθέστερα, το σύστημα είναι μερικώς παρατηρήσιμο και, κατ' επέκταση, η εκάστοτε κατάστασή του περιγράφεται από μία κατανομή πιθανοτήτων (partially observable Markov decision processes - POMDM).

Όταν το σύστημα βρίσκεται στην κατάσταση $s \in S$, μπορεί να επιλεγεί μία δράση a από το σύνολο των δράσεων A_s που είναι διαθέσιμες ενόσω το σύστημα βρίσκεται στην κατάσταση αυτή. Επομένως, το σύνολο όλων των δυνατών δράσεων $\{A\}$ είναι η ένωση των $\{A_s\}$ για κάθε πιθανή δράση s , $A = \cup_{s \in S} A_s$ [Σχ. 2.2]. Η επιλογή μιας δράσης μπορεί να γίνει είτε με ντετερμινιστικό τρόπο είτε κατά τυχαίο τρόπο. Μετά την επιλογή της το σύστημα μεταβαίνει σε νέα κατάσταση. Για την τυχαία επιλογή μιας δράσης απαιτείται η εκλογή μιας

κατανομής πιθανοτήτων $q(\cdot) \in \mathcal{S}(A_s)$, όπου $\mathcal{S}(A_s)$ είναι το σύνολο των κατανομών πιθανοτήτων στα υποσύνολα του A_s . Το ενδεχόμενο, λοιπόν, να επιλεγεί η δράση a όταν το σύστημα βρίσκεται στην κατάσταση s εμφανίζεται με πιθανότητα $q(a)$.

Τα σύνολα $\{A\}$ και $\{S\}$ μπορεί να είναι:

1. Διακριτά ή συνεχή
2. Πεπερασμένα ή αριθμησίμως άπειρα



Σχήμα 2.2 Σχέση μεταξύ των συνόλων $\{A\}$, $\{A_s\}$ και $\{S\}$

Στην περίπτωση όπου κάποιο από τα $\{S\}$ ή $\{A\}$ εξαρτάται από το χρόνο, η ανωτέρω θεωρία γενικεύεται, ώστε να ισχύει

$$S = \bigcup_{t \in T} S_t \quad \text{και} \quad A_s = \bigcup_{t \in T} A_{s,t}$$

όπου S_t είναι το σύνολο των δυνατών καταστάσεων τη χρονική στιγμή t , ενώ $A_{s,t}$ είναι το σύνολο των επιτρεπόμενων δράσεων για την κατάσταση s κατά τη χρονική στιγμή t .

- Αμοιβές $r_t(s, a)$ και πιθανότητες μετάβασης $p_t(\cdot | s, a)$:

Όταν επιλεγεί μία δράση a κατά τη χρονική περίοδο t και ενόσω το σύστημα βρίσκεται στην κατάσταση s , προκύπτουν άμεσα δύο αποτελέσματα :

1. Λαμβάνεται μία απόδοση $r_t(s, a)$ που θεωρείται όφελος αν είναι θετική και απώλεια για το σύστημα αν είναι αρνητική
2. Ταυτόχρονα, την επόμενη χρονική περίοδο, το σύστημα μεταβαίνει σε νέα κατάσταση που καθορίζεται μέσω της δεσμευμένης πιθανότητας $p_t(\cdot | s, a)$

Ως προς το μέγεθος $r_t(s, a)$, δεν είναι σημαντικός ο τρόπος με τον οποίο λαμβάνεται το όφελος/απώλεια στη διάρκεια του χρόνου, καθώς μπορεί να

λαμβάνεται σε συγκεκριμένες χρονικές στιγμές ή να αθροίζεται κατά τη διάρκεια κάθε περιόδου και να λαμβάνεται στο τέλος της τελευταίας περιόδου. Όμως, σε κάθε περίπτωση, πριν την επιλογή της αντίστοιχης δράσης a απαιτείται η γνώση του ή έστω η γνώση της αναμενόμενης τιμής του,. Επιπλέον, ενώ είναι σημαντικό το όφελος/απώλεια $r_t(s,a)$ να μην εξαρτάται από ενδεχόμενες μελλοντικές δράσεις, είναι συνήθως η εξάρτηση του $r_t(s,a)$ από την κατάσταση j που επιλέγει το υπό μελέτη σύστημα στην επόμενη χρονική περίοδο. Στην περίπτωση αυτή ενδιαφέρει το μέγεθος $r_t(s,a,j)$, το οποίο είναι το όφελος/απώλεια που λαμβάνεται την περίοδο t , όταν η κατάσταση του συστήματος είναι s και επιλέγεται η δράση a που οδηγεί το σύστημα στην κατάσταση j την περίοδο $t+1$. Στην περίπτωση όπου απαιτείται να μην υπάρχει η επίδραση της επόμενης κατάστασης, η αναμενόμενη τιμή του μεγέθους $r_t(s,a)$ υπολογίζεται μέσω της σχέσης

$$r_t(s,a) = \sum_{j \in S} r_t(s, a, j) p_t(j|s, a) \quad (2.2)$$

Στην ανωτέρω εξίσωση, η μη αρνητική συνάρτηση $p_t(j|s,a)$ υποδηλώνει τη δεσμευμένη πιθανότητα το σύστημα να μεταβεί στην κατάσταση $j \in S$ την περίοδο $t+1$ δοθέντος του σύνθετου γεγονότος να έχει επιλεγεί η δράση $a \in A_s$ εφόσον το σύστημα βρίσκεται στην κατάσταση $s \in S$ την προηγούμενη περίοδο t . Η συνάρτηση $p_t(\cdot|s,a)$ ονομάζεται συνάρτηση πιθανότητας μετάβασης για την οποία ισχύει

$$\sum_{j \in S} p_t(j|s, a) = 1$$

Υπάρχει το ενδεχόμενο να συμβούν περισσότερες από μία αλλαγές της κατάστασης του συστήματος στο χρονικό διάστημα μεταξύ των περιόδων t και $t+1$, όμως γίνεται η παραδοχή ότι το μοντέλο που χρησιμοποιείται δεν επηρεάζεται από αυτού του είδους τις μεταβολές. Επιπλέον, υπενθυμίζεται ότι και τα δύο μεγέθη $r_t(s,a)$ και $p_t(\cdot|s,a)$ εξαρτώνται μόνο από την παρούσα κατάσταση του συστήματος και όχι από παλαιότερες. Αυτό ακριβώς είναι που διαφοροποιεί

το μαρκοβιανό μοντέλο αποφάσεων από τα υπόλοιπα ακολουθιακά μοντέλα της θεωρίας αποφάσεων.

2.2.2 Κανόνες απόφασης / Πολιτική

Όταν ένα σύστημα περιγράφεται μέσω των προαναφερθεισών παραμέτρων, έχει τεθεί η βάση ώστε να αποτελεί μαρκοβιανή διαδικασία απόφασης. Συνήθως, αφού καθοριστεί πλήρως το μοντέλο ενός συστήματος, ο στόχος που πρέπει να επιτευχθεί, είναι η βελτιστοποίηση της συμπεριφοράς του είτε μεγιστοποιώντας το όφελος που αυτό μπορεί να αποδώσει είτε ελαχιστοποιώντας τις απώλειες που ανακύπτουν καθώς το σύστημα μεταβαίνει από τη μία κατάσταση στην επόμενη. Κατ' επέκταση, το σημαντικότερο στοιχείο της διαδικασίας είναι η ορθή επιλογή των δράσεων σε κάθε χρονική περίοδο. Οι κανόνες απόφασης περιγράφουν ακριβώς αυτή τη διαδικασία της επιλογής δράσης a για κάθε τρέχουσα κατάσταση s και σε συγκεκριμένη χρονική περίοδο t και μπορούν να διακριθούν σε ντετερμινιστικούς και μη ντετερμινιστικούς.

Οι ντετερμινιστικοί κανόνες απόφασης (Markovian Deterministic - MD) ορίζονται μέσω της συνάρτησης

$$d_t : S \rightarrow A_s \quad (2.3)$$

και καθορίζουν με βεβαιότητα τη δράση η οποία πρέπει να επιλεγεί όταν το σύστημα βρίσκεται στην κατάσταση s την περίοδο t . Από την άλλη πλευρά, οι μη ντετερμινιστικοί κανόνες απόφασης (Markovian Randomized - MR) εμπλέκουν την πιθανότητα να επιλεγεί μία εκ των δυνατών δράσεων στην κατάσταση s και την περίοδο t . Κατ' επέκταση, ο κανόνας d_t καθορίζει μία κατανομή πιθανοτήτων $q_{dt}(\cdot)$ στο σύνολο των δράσεων $\{A_s\}$:

$$d_t : S \rightarrow \mathcal{P}(A_s) \text{ και } q_{dt}(\cdot) \in \mathcal{P}(A_s) \quad (2.4)$$

Αφού καθοριστούν οι κανόνες απόφασης, το επόμενο βήμα είναι η διαμόρφωση της κατάλληλης πολιτικής ή στρατηγικής. Πρακτικά, η πολιτική είναι ένα σύνολο κανόνων απόφασης που πρέπει να εφαρμοστούν στις διαδοχικές χρονικές περιόδους λήψης απόφασης που κατευθύνει τον ελεγκτή του συστήματος να επιλέξει τη σωστή δράση για

το υπό μελέτη σύστημα σε οποιαδήποτε μελλοντική κατάσταση και ανεξάρτητα από τις προηγούμενες καταστάσεις και δράσεις που έχουν επιλεγεί. Ο μαθηματικός ορισμός μιας πολιτικής π είναι

$$\pi = (d_1, d_2, \dots, d_t, \dots, d_{N-1}) \quad (2.5)$$

με $N \leq \infty$ και $d_t \in D_t^{MD}$ ή D_t^{MR} ανάλογα με το είδος των κανόνων απόφασης.

Η ειδική περίπτωση πολιτικής όπου $d_t = d, \forall t \in T$ ονομάζεται στατική, έχει τη μορφή $\pi = (d, d, \dots)$ και συμβολίζεται χάριν συντομίας ως d^∞ . Η σημασία της στατικής πολιτικής είναι μεγάλη, κυρίως για μαρκοβιανές διαδικασίες απόφασης άπειρου ορίζοντα για τις οποίες $N \rightarrow \infty$.

2.3 Μαρκοβιανές διαδικασίες πεπερασμένου ορίζοντα

Στην παρούσα ενότητα θεωρείται ότι το σύνολο T είναι διακριτό, πεπερασμένο και περιλαμβάνει N περιόδους. Επιπλέον, για λόγους απλότητας θεωρείται μόνο το πρόβλημα μεγιστοποίησης του οφέλους που αποδίδει το σύστημα, καθώς ανάλογη είναι η αντιμετώπιση όταν ενδιαφέρει η ελαχιστοποίηση των απωλειών. Αρχικά, τίθενται τα κριτήρια με βάση τα οποία θα προσδιοριστεί η λύση του προβλήματος και ορίζεται η μορφή που πρέπει να έχει η βέλτιστη πολιτική. Στη συνέχεια, καταστρώνονται οι εξισώσεις βελτιστοποίησης (Bellman equations), η λύση των οποίων θα οδηγήσει στην εύρεση της βέλτιστης πολιτικής. Στο τέλος της ενότητας παρουσιάζεται ο αλγόριθμος επαγωγής προς τα πίσω που αποτελεί τη μέθοδο επίλυσης των εξισώσεων Bellman για τις μαρκοβιανές διαδικασίες αυτού του είδους.

2.3.1 Κριτήριο συνολικού αναμενόμενου οφέλους

Με τον όρο $u_N^\pi(s)$ ορίζεται το αναμενόμενο συνολικό όφελος (expected total reward) καθόλη τη χρονική διάρκεια T εφαρμογής της πολιτικής π για την οποία η αρχική κατάσταση του συστήματος, για $t=1$, είναι s . Η αντίστοιχη μαθηματική έκφραση είναι

$$u_N^\pi(s) = E_s^\pi \{ \sum_{t=1}^{N-1} r_t(s_t, d_t(s_t)) + r_N(s_N) \} \quad \text{για } \pi \in \Pi^{MD} \quad (2.6)$$

Έχοντας κάνει την παραδοχή ότι για $t=N$ δεν λαμβάνεται καμία απόφαση από τον ελεγκτή του συστήματος, το όφελος που αποδίδει το σύστημα για $t=N$ δηλώνεται ως $r_N(s_N)$.

Παραλλαγή του ανωτέρω μεγέθους αποτελεί η ακόλουθη έκφραση, η οποία ονομάζεται αναμενόμενη συνολική παρούσα αξία του οφέλους του συστήματος (expected total discounted reward)

$$u_{N,\lambda}^\pi(s) = E_s^\pi \{ \sum_{t=1}^{N-1} \lambda^{t-1} r_t(s_t, d_t(s_t)) + \lambda^{N-1} r_N(s_N) \}, 0 \leq \lambda < 1 \quad (2.7)$$

Ο παράγοντας λ λαμβάνει κατάλληλη τιμή ώστε το αποτέλεσμα της εξίσωσης να αποτελεί την παρούσα αξία του συνολικού οφέλους, δηλαδή την τιμή του για $t=1$. Ειδικότερα θεωρείται ότι το όφελος που αποδίδει το σύστημα σε μελλοντικές περιόδους έχει αξία που διαρκώς μειώνεται σε σχέση με την τιμή του κατά την έναρξη της μαρκοβιανής διαδικασίας $t=1$. Για το λόγο αυτό απαιτείται ο καθορισμός του παράγοντα λ (discount factor) μέσω του οποίου λαμβάνονται υπόψη τα μελλοντικά οφέλη μειωμένα κατά το επιθυμητό ποσοστό.

Με βάση τον ανωτέρω ορισμό, βέλτιστη τιμή του συνολικού αναμενόμενου οφέλους είναι

$$u_N^*(s) = \sup_{\pi \in \Pi^{MD}} u_N^\pi(s), s \in S \quad (2.8)$$

ή εφόσον υπάρχει μέγιστο:

$$u_N^*(s) = \max_{\pi \in \Pi^{MD}} u_N^\pi(s), s \in S$$

Ανάλογες είναι οι εκφράσεις για $\pi \in \Pi^{MR}$.

Στόχος είναι να προσδιοριστεί η κατάλληλη πολιτική που μεγιστοποιεί μία εκ των ανωτέρω συναρτήσεων οφέλους. Προκύπτουν δύο κριτήρια για την επίλυση του προβλήματος, αντίστοιχα προς το ποιο μέγεθος επιδιώκεται να μεγιστοποιηθεί: το κριτήριο συνολικού αναμενόμενου οφέλους και το κριτήριο παρούσας αξίας συνολικού αναμενόμενου οφέλους.

2.3.2 Εύρεση βέλτιστης πολιτικής

Ο προσδιορισμός της βέλτιστης πολιτικής ενός συστήματος μέσω της μαρκοβιανής θεωρίας διατυπώνεται ως εξής

$$\text{εύρεση πολιτικής } \pi^* : u_N^{\pi^*}(s) \geq u_N^\pi(s), s \in S, \forall \pi \in \Pi^{MD} \text{ ή } \Pi^{MR} \quad (2.9)$$

όπου το σύμβολο * υποδηλώνει βέλτιστη τιμή. Το όφελος που επιτυγχάνεται με την εφαρμογή της βέλτιστης πολιτικής ικανοποιεί τη σχέση

$$u_N^{\pi^*}(s) = u_N^*(s), s \in S \quad (2.10)$$

Υπάρχουν συστήματα για τα οποία δεν μπορεί να προσδιοριστεί πολιτική που να ικανοποιεί την ανωτέρω ανισότητα. Στην περίπτωση αυτή αναζητάται μια υποβέλτιστη πολιτική (ε - optimal policy). Το τελευταίο πρόβλημα διατυπώνεται ως ακολούθως

Δεδομένης μιας αυθαίρετης σταθεράς $\varepsilon > 0$, αρκετά μικρής, να προσδιοριστεί η πολιτική π_ε^* ώστε

$$u_N^{\pi_\varepsilon^*}(s) + \varepsilon \geq u_N^\pi(s), s \in S, \forall \pi \in \Pi^{MD} \text{ ή } \Pi^{MR} \quad (2.11)$$

Το αντίστοιχο όφελος ικανοποιεί: $u_N^{\pi_\varepsilon^*}(s) + \varepsilon \geq u_N^*(s), s \in S$

2.3.3 Εξισώσεις βελτιστοποίησης Bellman (Bellman Equations)

Οι εξισώσεις αυτές είναι ιδιαίτερης σημασίας καθώς οι λύσεις τους αντιστοιχούν σε βέλτιστες συναρτήσεις οφέλους και αποτελούν τη βάση για τον προσδιορισμό της βέλτιστης πολιτικής. Η μορφή τους είναι η ακόλουθη:

$$u_t(s_t) = \max_{a \in A_{s_t}} \left\{ r_t(s_t, a) + \sum_{j \in S} p_t(j|s_t, a) u_{t+1}(j) \right\} \quad t < N \quad (2.12)$$

$$u_N(s_N) = r_N(s_N) \quad t = N$$

Το μέγεθος $r_t(s_t, a)$ είναι το άμεσο όφελος που προκύπτει από την εφαρμογή της δράσης a ενόσω το σύστημα βρίσκεται στην κατάσταση s_t . Το μέγεθος $\sum_{j \in S} p_t(j|s_t, a) u_{t+1}(j)$ εκφράζει το αναμενόμενο όφελος από την εφαρμογή της δράσης a ενόσω το σύστημα βρίσκεται στην κατάσταση s_t δεδομένου ότι με την εφαρμογή της δράσης a το σύστημα μπορεί να μεταβεί από την κατάσταση s_t σε κάποια από τις καταστάσεις j με συγκεκριμένη πιθανότητα μετάβασης. Συνεπώς, οι εξισώσεις Bellman εκφράζουν τη μεγιστοποίηση του οφέλους που αποδίδει το σύστημα με την επιλογή μιας δράσης a . Η λύση τους είναι μια ακολουθία συναρτήσεων $u_t : S_t \rightarrow \mathbb{R}$, $t=1, \dots, N$, όπου \mathbb{R} είναι το σύνολο με στοιχεία τα οφέλη που αποδίδει το σύστημα σε κάθε χρονική περίοδο.

Οι εξισώσεις Bellman είναι ουσιώδη εργαλεία της Μαρκοβιανής θεωρίας γιατί διαθέτουν τις εξής ιδιότητες:

1. Οι λύσεις τους για κάθε χρονική περίοδο είναι οι βέλτιστες τιμές οφέλους που αποδίδει το σύστημα την αντίστοιχη περίοδο.
2. Αποτελούν το κριτήριο απόφασης για το αν μια πολιτική είναι βέλτιστη, καθώς εφόσον το αναμενόμενο συνολικό όφελος μιας πολιτικής π ικανοποιεί το σύστημα των εξισώσεων αυτών για $t=1, \dots, N$, η πολιτική αυτή θεωρείται βέλτιστη.
3. Αποτελούν τη βάση μιας αποδοτικής διαδικασίας υπολογισμού βέλτιστης πολιτικής και αποκαλύπτουν τα χαρακτηριστικά που πρέπει να έχει η τελευταία.

2.3.4 Αλγόριθμος επαγωγής προς τα πίσω (Backward induction)

Πρόκειται για αποδοτική μέθοδο επίλυσης μαρκοβιανών προβλημάτων απόφασης διακριτού χρόνου πεπερασμένου ορίζοντα. Ο όρος backward induction ισοδυναμεί με

αυτόν του δυναμικού προγραμματισμού (dynamic programming) αλλά ο τελευταίος αναφέρεται γενικότερα σε ακολουθιακά μοντέλα απόφασης.

Γίνονται οι εξής παραδοχές:

1. Η πολιτική που τελικά προκύπτει είναι ντετερμινιστική μαρκοβιανή
2. Το προτεινόμενο μοντέλο του συστήματος έχει διατυπωθεί κατά τρόπο που επιτυγχάνει μέγιστη τιμή οφέλους οπότε η προκύπτουσα πολιτική είναι βέλτιστη και όχι υποβέλτιστη.

Ως συνέπεια των ανωτέρω προκύπτει η οικογένεια εξισώσεων Bellman που παρουσιάστηκε προηγουμένως που επιλύονται με χρήση του αλγορίθμου επαγωγής προς τα πίσω. Ο τελικός στόχος είναι η εύρεση των κατάλληλων δράσεων που πρέπει να εφαρμοστούν σε κάθε χρονική περίοδο και κατάσταση του συστήματος ώστε το τελικό όφελος να είναι μέγιστο. Ο αλγόριθμος περιλαμβάνει τα ακόλουθα βήματα:

Βήμα 1^ο: Τίθεται $t=N$ και

$$u_N^*(s_N) = r_N(s_N) \quad \forall s_N \in S \quad (2.13)$$

Βήμα 2^ο: Αντικαθίσταται $t-1 \rightarrow t$ και υπολογίζεται το μέγεθος $u_t^*(s_t)$ για κάθε μία από τις καταστάσεις $s_t \in S$ μέσω της εξίσωσης:

$$u_t^*(s_t) = \max_{a \in A_{s_t}} \left\{ r_t(s_t, a) + \sum_{j \in S} p_t(j|s_t, a) u_{t+1}^*(j) \right\} \quad (2.14)$$

επίσης τίθεται:

$$A_{s_t, t}^* = \operatorname{argmax}_{a \in A_{s_t}} \left\{ r_t(s_t, a) + \sum_{j \in S} p_t(j|s_t, a) u_{t+1}^*(j) \right\} \quad (2.15)$$

Βήμα 3^ο: Αν $t=1$ ο αλγόριθμος τερματίζεται, άλλως επαναλαμβάνεται το 2^ο βήμα.

Η μέθοδος αυτή υπολογίζει τη μέγιστη τιμή του οφέλους του συστήματος κάθε χρονική περίοδο, ξεκινώντας από την τελευταία όπου η τιμή του είναι γνωστή. Σε κάθε χρονική

περίοδο υπολογίζεται επίσης και η δράση a η οποία, εφόσον επιλεγεί, αποδίδει τη μέγιστη τιμή της συνάρτησης u_t . Κατ' επέκταση με το πέρας του αλγορίθμου έχει προσδιοριστεί η βέλτιστη πολιτική με κριτήριο το μέγιστο όφελος, καθώς είναι γνωστό το βέλτιστο σύνολο δράσεων $A_{s_t,t}^*$ και επομένως οι βέλτιστοι κανόνες απόφασης $D_t^* \equiv \times_{s \in S} A_{s_t,t}^*$, οπότε βέλτιστη είναι η πολιτική $\pi^* \in \Pi^* \equiv D_1^* \times \dots \times D_{N-1}^*$.

Για κάποια κατάσταση s_t του συστήματος υπάρχει το ενδεχόμενο να προκύψουν δύο ή και περισσότερες δράσεις που αποδίδουν το ίδιο όφελος και άρα το σύνολο $A_{s_t,t}^*$ να περιέχει δύο ή περισσότερα στοιχεία. Στην περίπτωση αυτή δεν προκύπτει μοναδική βέλτιστη πολιτική, αν και η βέλτιστη τιμή οφέλους παραμένει μοναδική.

2.4 Μαρκοβιανές διαδικασίες άπειρου ορίζοντα

Στην ενότητα που ακολουθεί θεωρείται ότι το σύνολο T είναι άπειρο. Επιπλέον, γίνονται οι εξής παραδοχές:

1. Τα οφέλη $r(s,a)$, οι πιθανότητες $p(j|s,a)$ και τα σύνολα των κανόνων απόφασης D^{MD} και D^{MR} δε μεταβάλλονται με το χρόνο
2. Το σύνολο των καταστάσεων $\{S\}$ θεωρείται διακριτό
3. Στόχο αποτελεί η μεγιστοποίηση του οφέλους που αποδίδει το σύστημα
4. Οι μαθηματικές εκφράσεις παρατίθενται μόνο για την περίπτωση των μαρκοβιανών ντετερμινιστικών κανόνων

Αρχικά, αναφέρονται τα διάφορα κριτήρια που πρέπει να ικανοποιούνται από τη βέλτιστη πολιτική και ορίζεται η μορφή της τελευταίας. Ακολουθεί η κατάστρωση των εξισώσεων βελτιστοποίησης και η παρουσίαση δύο ευρέως χρησιμοποιούμενων αλγορίθμων για την επίλυσή τους. Λόγω του άπειρου χρονικού ορίζοντα ανακύπτουν θέματα σύγκλισης των συναρτήσεων οφέλους καθώς και των αλγορίθμων επίλυσης των εξισώσεων βελτιστοποίησης, τα οποία όμως εκφεύγουν από το πλαίσιο της παρούσας εργασίας.

2.4.1 Κριτήρια οφέλους

Τα κριτήρια οφέλους σε προβλήματα άπειρου ορίζοντα ορίζονται όπως και στο πρόβλημα πεπερασμένου ορίζοντα. Όταν εφαρμόζεται η πολιτική $\pi \in \Pi^{\text{MD}}$, το αναμενόμενο συνολικό όφελος ορίζεται μέσω της σχέσης

$$u^\pi(s) = \lim_{N \rightarrow \infty} E_S^\pi \left\{ \sum_{t=1}^N r(s_t, d_t(s_t)) \right\} = \lim_{N \rightarrow \infty} u_N^\pi(s) \quad (2.16)$$

Στην περίπτωση αυτή το μέγιστο όφελος που τελικά αποδίδει το σύστημα υπολογίζεται μέσω της σχέσης

$$u^*(s) = \sup_{\pi \in \Pi^{MD}} u^\pi(s)$$

Η αναμενόμενη συνολική παρούσα αξία του οφέλους ορίζεται μέσω της σχέσης

$$u_\lambda^\pi(s) = \lim_{N \rightarrow \infty} E_S^\pi \left\{ \sum_{t=1}^N \lambda^{t-1} r(s_t, d_t(s_t)) \right\} \text{ με } 0 \leq \lambda < 1 \quad (2.17)$$

Το μέγεθος αυτό χρησιμοποιείται πολύ συχνότερα σε σχέση με το αναμενόμενο συνολικό όφελος. Ο παράγοντας λ έχει ιδιαίτερο ρόλο σε προβλήματα αυτού του είδους καθώς εξασφαλίζει τη σύγκλιση της σειράς και, συνεπώς, την ύπαρξη του ανωτέρω ορίου με μοναδική προϋπόθεση

$$\sup_{s \in S} \sup_{a \in A_s} |r(s, a)| = M < \infty \quad (2.18)$$

Στα μοντέλα όπου εμφανίζεται ο παράγοντας λ η μέγιστη τιμή του οφέλους ορίζεται μέσω της σχέσης

$$u_\lambda^*(s) = \sup_{\pi \in \Pi^{MD}} u_\lambda^\pi(s) \quad (2.19)$$

2.4.2 Βέλτιστη πολιτική

Όταν μελετάται το πρόβλημα μέγιστου αναμενόμενου συνολικού οφέλους, μια πολιτική π^* είναι βέλτιστη όταν ισχύει

$$u^{\pi^*}(s) \geq u^\pi(s) \quad \forall s \in S \text{ και } \forall \pi \in \Pi^{MD} \quad (2.20)$$

Κατ' επέκταση, μια βέλτιστη πολιτική $\pi^* \in \Pi^{MD}$ υπάρχει όταν ικανοποιείται η σχέση

$$u^{\pi^*}(s) = u^*(s) \quad \forall s \in S$$

Όταν μελετάται το πρόβλημα της μέγιστης αναμενόμενης παρούσας αξίας του οφέλους, μια πολιτική π^* είναι βέλτιστη, για δοθέν $\lambda: 0 \leq \lambda < 1$, όταν:

$$u_{\lambda}^{\pi^*}(s) \geq u_{\lambda}^{\pi}(s) \quad \forall s \in S \text{ και } \forall \pi \in \Pi^{MD} \quad (2.21)$$

Και σε αυτή την περίπτωση μια βέλτιστη πολιτική υπάρχει όταν ισχύει:

$$u_{\lambda}^{\pi^*}(s) = u_{\lambda}^*(s) \quad \forall s \in S$$

2.4.3 Εξισώσεις Βελτιστοποίησης Bellman

Οι εξισώσεις βελτιστοποίησης για μαρκοβιανές διαδικασίες άπειρου ορίζοντα, έχουν μοναδική λύση στο σύνολο V , το οποίο είναι το σύνολο των φραγμένων πραγματικών συναρτήσεων οφέλους u με πεδίο ορισμού το σύνολο των καταστάσεων S . Επιπλέον, η βέλτιστη τιμή οφέλους, όπως αυτή ορίστηκε προηγουμένως ικανοποιεί το σύστημα αυτών των εξισώσεων, ενώ η λύση τους οδηγεί στην εύρεση στατικής βέλτιστης πολιτικής. Οι εξισώσεις Bellman λαμβάνουν τη μορφή

$$u(s) = \sup_{\alpha \in A_s} \left\{ r(s, \alpha) + \sum_{j \in S} \lambda p(j|s, \alpha) u(j) \right\} \quad (2.22)$$

2.4.4 Αλγόριθμος αναδρομικής αξίας (Value iteration)

Ο αλγόριθμος αναδρομικής αξίας είναι αρκετά απλός και χρησιμοποιείται ευρέως για την επίλυση του συστήματος των εξισώσεων Bellman. Τα βήματα του αλγορίθμου παρατίθενται χωρίς να γίνεται περαιτέρω διερεύνηση για τις προϋποθέσεις σύγκλισης του αλγορίθμου. Με το πέρας του αλγορίθμου προσδιορίζεται μια στατική υποβέλτιστη (ϵ -optimal) πολιτική $(d_{\epsilon})^{\infty}$ καθώς και μια προσέγγιση του οφέλους που προκύπτει από την εφαρμογή της. Πρακτικά, αν επιλεγθεί αρκετά μικρό ϵ , η προκύπτουσα πολιτική απέχει πολύ λίγο από τη βέλτιστη.

Βήμα 1^ο: επιλέγεται ένα αρχικό όφελος $u^0 \in V$, καθορίζεται η τιμή του $\varepsilon > 0$ και τίθεται $n=0$

Βήμα 2^ο: Για κάθε $s \in S$ υπολογίζεται το μέγεθος $u^{n+1}(s)$ μέσω της εξίσωσης

$$u^{n+1}(s) = \max_{a \in A_s} \left\{ r(s, a) + \sum_{j \in S} \lambda p(j|s, a) u^n(j) \right\} \quad (2.23)$$

Βήμα 3^ο: εφόσον ισχύει $\|u^{n+1} - u^n\| < \varepsilon(1-\lambda)/2\lambda$, γίνεται μετάβαση στο βήμα 4, άλλως αυξάνεται το n κατά 1 και γίνεται μετάβαση στο βήμα 2.

Βήμα 4^ο: Για κάθε $s \in S$ επιλέγεται:

$$d_\varepsilon(s) \in \arg \max_{a \in A_s} \left\{ r(s, a) + \sum_{j \in S} \lambda p(j|s, a) u^{n+1}(j) \right\} \quad (2.24)$$

και ο αλγόριθμος τερματίζεται.

2.4.5 Αλγόριθμος αναδρομικής πολιτικής (Policy iteration)

Ο αλγόριθμος αναδρομικής πολιτικής αντιμετωπίζει καλύτερα τα μαρκοβιανά προβλήματα άπειρου ορίζοντα σε σχέση με το γενικότερο αλγόριθμο αναδρομικής αξίας. Με την εφαρμογή του προκύπτει μια ακολουθία ντετερμινιστικών μαρκοβιανών κανόνων απόφασης $\{d_n\}$ καθώς και μια ακολουθία συναρτήσεων οφέλους $\{u_n\}$, οι οποίες είναι με βεβαιότητα πεπερασμένες αν τα σύνολα $\{A\}$ και $\{S\}$ είναι πεπερασμένα. Η ακολουθία των συναρτήσεων οφέλους $\{u_n\}$ συγκλίνει υπό προϋποθέσεις στη βέλτιστη συνάρτηση οφέλους u_λ^* .

Βήμα 1^ο: τίθεται $n=0$ και επιλέγεται κατά τυχαίο τρόπο ένας κανόνας απόφασης $d_0 \in D$

Βήμα 2^ο: προσδιορίζεται το μέγεθος u^n επιλύοντας την εξίσωση

$$(I - \lambda P_{d_n})u = r_{d_n} \quad (2.25)$$

Βήμα 3^ο: επιλέγεται κατάλληλος κανόνας απόφασης d_{n+1} , ώστε να ικανοποιείται

$$d_{n+1} \in \arg \max_{d \in D} \{r_d + \lambda P_d u^n\} \quad (2.26)$$

επιδιώκεται, εφόσον είναι δυνατό, να ισχύει $d_{n+1} = d_n$

Βήμα 4^ο: Αν $d_{n+1} = d_n$, ο αλγόριθμος τερματίζεται και τίθεται $d^* = d_n$. Άλλωως αυξάνεται ο δείκτης n κατά 1 και γίνεται μετάβαση στο βήμα 2.

ΚΕΦΑΛΑΙΟ 3

ΕΦΑΡΜΟΓΕΣ ΜΑΡΚΟΒΙΑΝΩΝ ΜΟΝΤΕΛΩΝ ΣΕ ΔΙΚΤΥΑ ΕΠΙΚΟΙΝΩΝΙΩΝ

3.1 Εισαγωγή

Η εφαρμογή αναλυτικών μαθηματικών μοντέλων στη μελέτη τηλεπικοινωνιακών προβλημάτων και ειδικότερα στη μελέτη ασυρμάτων δικτύων δεν έχει υπάρξει ιδιαίτερα επιτυχής λόγω της πολυπλοκότητάς και της δυναμικής τοπολογίας των τελευταίων. Το μαρκοβιανό μοντέλο απόφασης δίνει λύσεις σε πολλά νέα προβλήματα και ταυτόχρονα διαχειρίζεται αρκετά παλαιότερα προβλήματα με μεγαλύτερη επιτυχία σε σχέση με άλλες μεθόδους. Ακολουθεί μια περιγραφή τηλεπικοινωνιακών εφαρμογών στις οποίες έχει πρόσφατα χρησιμοποιηθεί επιτυχώς το μαρκοβιανό μοντέλο και έχουν εξαχθεί ενδιαφέροντα συμπεράσματα.

3.2 Δυναμική εκχώρηση φάσματος / Γνωστικές ραδιοεπικοινωνίες (Physical/MAC layer)

Η μαρκοβιανή θεωρία έχει χρησιμοποιηθεί ευρέως σε θέματα διαχείρισης του φάσματος για τη διατύπωση μοντέλων για δευτερεύοντα δίκτυα που αποσκοπούν στη χρήση ανεκμετάλλετων τμημάτων του φάσματος που έχει εκχωρηθεί στατικά σε συγκεκριμένες υπηρεσίες ή σε πρωτεύοντα δίκτυα. Ο στόχος που τίθεται κάθε φορά είναι διαφορετικός, όπως η μεγιστοποίηση της ρυθμιστικής του δευτερεύοντος δικτύου [4] ή η ελαχιστοποίηση του χρόνου αναζήτησης ελεύθερου καναλιού προς μετάδοση [12]. Αντιστοίχως, διατυπώνεται η κατάλληλη πολιτική που ικανοποιεί τον εκάστοτε στόχο. Σε κάθε περίπτωση, πρόκειται για διαστρωματική αντιμετώπιση καθώς το πρόβλημα αφορά, αφενός πρόσβαση σε κοινό μέσο και αφετέρου ανίχνευση φάσματος στο φυσικό στρώμα.

3.2.1 Περιγραφή του προβλήματος ανίχνευσης φάσματος

Όσον αφορά το πρωτεύον δίκτυο, θεωρείται ότι το φάσμα που του έχει αποδοθεί απαρτίζεται από N κανάλια έκαστο με εύρος ζώνης B_n , $n=1,\dots,N$. Η διαθεσιμότητα κάθε καναλιού σε κάθε χρονοσχισμή t εκφράζεται μέσω της μεταβλητής $S_n(t) \in \{0, 1\}$ (κανάλι κατειλημμένο), 1 (κανάλι ελεύθερο)}. Η διαθεσιμότητα του συνολικού φάσματος περιγράφεται ως διακριτή μαρκοβιανή διαδικασία με χώρο καταστάσεων $S \cong \{0, 1\}^N$ και γνωστές πιθανότητες μετάβασης $P(s'|s) \cong P_r\{S(t)=s'|S(t-1)=s\}$ $s, s' \in S$. Υπάρχει η δυνατότητα να θεωρηθεί ότι το πλήθος των καναλιών N είναι μεταβάλλεται με το χρόνο [2] αλλά παραμένει για συγκεκριμένο πλήθος χρονοσχισμών σταθερό επίσης υπάρχουν και εργασίες που θεωρούν ότι η χρησιμοποίηση του φάσματος παραμένει αναλλοίωτη για συγκεκριμένο πλήθος T χρονοσχισμών [3]. Επιπλέον, έχουν μελετηθεί και δίκτυα τα οποία δεν λειτουργούν με δομή χρονοσχισμών [4] επειδή η πρόσβαση των χρηστών σε αυτά γίνεται σε οποιαδήποτε χρονική στιγμή.

Η απόφαση του χρήστη που ανήκει στο δευτερεύον δίκτυο να ανιχνεύσει και, στη συνέχεια, να αποκτήσει πρόσβαση σε κανάλι του πρωτεύοντος δικτύου περιγράφεται, στην πλειονότητα των περιπτώσεων, μέσω μιας μερικώς παρατηρήσιμης μαρκοβιανής διαδικασίας απόφασης (Partially Observable Markov Decision Process - POMDP) [5, Παράρτημα Α]. Η αιτία είναι ότι η στιγμιαία κατάσταση του πρωτεύοντος δικτύου δεν είναι πλήρως γνωστή στο δευτερεύοντα χρήστη, ο οποίος επιλέγοντας μια δράση λαμβάνει εκτός από το αναμενόμενο κέρδος/ζημία απλώς μια παρατήρηση της κατάστασης του δικτύου. Όσο το σύστημα εξελίσσεται, το διάνυσμα των παρατηρήσεων $\Lambda(t)$ (belief vector) παρέχει όλες τις απαραίτητες πληροφορίες για να παρθούν οι αποφάσεις για το πρόβλημα πρόσβασης σε κανάλι του πρωτεύοντος δικτύου. [6] Κάθε στοιχείο του $\Lambda(t)$ αποτελεί τη δεσμευμένη πιθανότητα η κατάσταση του πρωτεύοντος δικτύου στην αρχή της χρονικής σχισμής t να είναι s δεδομένου ότι η εξέλιξη του συστήματος μέχρι εκείνη τη στιγμή είναι $Y(t)$, δηλαδή

$$\Lambda(t) \triangleq \{\Lambda_s(t)\}_{s \in S} \quad \text{όπου} \quad \Lambda_s(t) \triangleq \Pr\{S(t-1)=s|Y(t)\} \quad (3.1)$$

όπου η αλληλουχία καταστάσεων $Y(t)$ εκφάζει την εξέλιξη του συστήματος μέχρι και τη χρονική στιγμή $t-1$. Η αλληλουχία αυτή προκύπτει βάσει τόσο των παρατηρήσεων της

κατάστασης του συστήματος όσο και της πραγματικής κατάστασης του συστήματος που γίνεται γνωστή με την απόφαση πρόσβασης σε αυτό.

3.2.2 Κατάστρωση μαρκοβιανού μοντέλου

Το σύνολο των καταστάσεων $\{S\}$ του υπό εξέταση προβλήματος μπορεί να είναι εξαιρετικά απλό και να περιλαμβάνει μόνο το διάνυσμα της διαθεσιμότητας των καναλιών που παρακολουθούνται από τους δευτερεύοντες χρήστες [7]. Όμως υπάρχουν αναφορές όπου έχουν μελετηθεί διεξοδικότερα ορισμένα χαρακτηριστικά των δευτερευόντων χρηστών, όπως η εναπομένουσα ενέργεια E [8] ή ακόμη και το μέγεθος B του ενταμιευτή [6] όταν δεν μπορούν να υποστηρίξουν μεγάλο αριθμό πακέτων.

Οι δευτερεύοντες χρήστες έχουν την επιλογή να ανιχνεύσουν ένα πλήθος $L \leq N$ καναλιών και να αποφασίσουν αν θα αποκτήσουν πρόσβαση σε οποιαδήποτε από τα κανάλια αυτά ή αν θα παραμείνουν ανενεργοί. Σε ορισμένα προβλήματα υπάρχει περαιτέρω περιορισμός ως προς το πλήθος M των καναλιών στα οποία μπορεί να αποκτήσει πρόσβαση ένας χρήστης, $M < L$. [9] Κατ' επέκταση, το σύνολο των δράσεων $\{A\}$ στην πλέον σύνθετη περίπτωση είναι της μορφής $\{a, \Phi_a\}$, όπου $a \in \{1, \dots, N\}$ το πλήθος των καναλιών που επιλέγονται προς ανίχνευση και Φ_a η απόφαση για πρόσβαση σε ένα κανάλι όταν υπάρχει συγκεκριμένο συμπέρασμα Θ_a για την κατάστασή του καναλιού αυτού [8].

Όταν είναι επιτυχής η πρόσβαση ενός χρήστη σε κανάλι του πρωτεύοντος δικτύου η αμοιβή που αποδίδεται στο δευτερεύον δίκτυο είναι ανάλογη του εύρους ζώνης του καναλιού $R \triangleq B_n$. Εναλλακτικά, θεωρείται ως όφελος το πλήθος των bits που επιτυγχάνει ο δευτερεύων χρήστης να αποστείλει μέσω του καναλιού στο οποίο έχει αποκτήσει πρόσβαση. [9] Σε παραλλαγή του προβλήματος δίνεται η δυνατότητα συνεργασίας μεταξύ των δευτερευόντων χρηστών που επιθυμούν πρόσβαση στο πρωτεύον δίκτυο. Στην περίπτωση αυτή το όφελος θεωρείται μοναδιαίο κάθε φορά που υπάρχει πρόσβαση κάποιου χρήστη σε κανάλι του πρωτεύοντος δικτύου χωρίς να συμβεί σύγκρουση με κάποιον από τους άλλους χρήστες του δευτερεύοντος δικτύου. [10]

Σε κάθε περίπτωση στόχος είναι η εύρεση των κατάλληλων δράσεων ανίχνευσης και πρόσβασης στο μέσο που μεγιστοποιεί το όφελος $\sum_t R(t)$ στο χρονικό ορίζοντα κατά τον

οποίο μελετάται το πρόβλημα. Αυτός είναι πεπερασμένος όταν υπάρχει περιορισμός ενέργειας στο δίκτυο και τίθεται ίσος με τη διάρκεια ζωής της μπαταρίας ενός δευτερεύοντος χρήστη.[8] Η εύρεση της βέλτιστης πολιτικής εκφράζεται μέσω της σχέσης [6]

$$\text{Εύρεση } \pi^* : \pi^* = \operatorname{argmax}_{\pi} E_{\pi}[\sum_{t=1}^T R(t) | \Lambda(1), E(1) = \mathcal{E}_o] \quad (3.2)$$

όπου το άθροισμα μπορεί να επεκταθεί μέχρι $T=\infty$ και η αρχική ενέργεια \mathcal{E}_o μπορεί να παραλειφθεί όταν δεν εξετάζονται συστήματα με περιορισμούς ενέργειας.

3.2.3 Επίλυση μαρκοβιανής διαδικασίας

Το πρόβλημα επιλύεται με χρήση των αναδρομικών μεθόδων δυναμικού προγραμματισμού που έχουν περιγραφεί προηγουμένως. Όταν η συμπεριφορά του συστήματος είναι άγνωστη εφαρμόζονται άλλοι αλγόριθμοι όπως ο αλγόριθμος Q-learning.[11, Παράρτημα Β] Χρησιμοποιείται, όμως, και ο γραμμικός προγραμματισμός, όταν υπάρχει περιορισμός στην πιθανότητα παρεμβολής του δευτερεύοντος χρήστη στο πρωτεύον δίκτυο, με αντίστοιχη διατύπωση του προβλήματος, δηλαδή

$$\max_{\pi} R(\pi), \quad \text{υπό τον περιορισμό } C(\pi) \leq p \quad (3.3)$$

όπου $R(\pi)$ η μέση τιμή του κέρδους όταν ακολουθείται η πολιτική π , $C(\pi)$ η μέση τιμή του κόστους που καταβάλλει ο δευτερεύων χρήστης όταν συμβαίνει σύγκρουση με ένα πρωτεύοντα χρήστη και p η επιτρεπόμενη πιθανότητα παρεμβολής. [4]

Η υπολογιστική πολυπλοκότητα, δηλαδή το πλήθος των πολλαπλασιασμών που απαιτούνται για να υπολογιστούν όλες οι δράσεις ανίχνευσης και πρόσβασης μιας πολιτικής, είναι τεράστια, ιδιαίτερα σε προβλήματα άπειρου χρονικού ορίζοντα. Το γεγονός αυτό έχει οδηγήσει στην υιοθέτηση, αρχικά, μιας άπληστης πολιτικής (greedy approach) για τη λύση του προβλήματος, η οποία βασίζεται στη μεγιστοποίηση του κέρδους του συστήματος σε κάθε χρονική σχισμή [9] ή γενικότερα σε συγκεκριμένο πλήθος w χρονικών σχισμών.[8] Κατ' αυτόν τον τρόπο, υπάρχει αντιστάθμιση μεταξύ της πολυπλοκότητας και της ακρίβειας καθώς η πολιτική αυτή δεν είναι ακριβώς βέλτιστη, αλλά η πολυπλοκότητα της είναι σαφώς μικρότερη από αυτή της βέλτιστης.

3.2.4 Συμπεράσματα

Στις εργασίες που θίγουν το θέμα της δυναμικής εκχώρησης φάσματος προκύπτουν διάφορα συμπεράσματα. Οι πολιτικές ανίχνευσης και πρόσβασης είναι πολιτικές κατωφλίου στην περίπτωση της εργασίας [6]. Στην ίδια εργασία αποδεικνύεται ότι για χαμηλές τιμές της ενέργειας E ο ρυθμός άφιξης δεδομένων λ επηρεάζει σε μεγάλο βαθμό την εύρεση της βέλτιστης πολιτικής. Τέλος, στην εργασία [8] διερευνάται εκτενώς η επίδραση του ποσοστού του χρησιμοποιούμενου φάσματος και οι διαλείψεις του διαύλου στην εύρεση της βέλτιστης πολιτικής.

3.3 Δρομολόγηση στο στρώμα δικτύου

Το μαρκοβιανό μοντέλο έχει υιοθετηθεί σε προβλήματα, όπως η εύρεση νέων τεχνικών δρομολόγησης, που σχετίζονται με το στρώμα δικτύου. Στην περίπτωση αυτή, κάθε κόμβος του δικτύου λαμβάνει αποφάσεις για δρομολόγηση πακέτων βασιζόμενος μόνο σε τοπικές πληροφορίες για το δίκτυο προερχόμενες από τους άμεσους γείτονες του.[13] Το περιβάλλον ενός δικτύου μπορεί να αλλάζει με στοχαστικό τρόπο γεγονός που επιβάλλει η διατύπωση του μοντέλου να βασίζεται σε μερικώς παρατηρήσιμη μαρκοβιανή διαδικασία απόφασης – POMDP και μάλιστα με άγνωστες πιθανότητες μετάβασης του συστήματος. [Παράρτημα Α]

3.3.1 Διατύπωση και επίλυση μαρκοβιανής διαδικασίας

Το σύνολο των καταστάσεων $\{S\}$ είναι το σύνολο των κόμβων που απαρτίζουν το δίκτυο. Το σύνολο των δράσεων $\{A\}$ απαρτίζεται από κάποιες προκαθορισμένες διαδρομές από ένα κόμβο εκκίνησης προς συγκεκριμένο κόμβο-προορισμό. Οι άγνωστες πιθανότητες μετάβασης συμβολίζονται ως $p_{xy}^t(a)$ και εκφράζουν την πιθανότητα επιτυχούς δρομολόγησης μεταξύ των κόμβων x και y , όταν επιλέγεται η διαδρομή $a \in A$ τη χρονική στιγμή t . Λόγω της μερικής γνώσης της κατάστασης του δικτύου, αρκεί κάθε κόμβος να γνωρίζει τη συνάρτηση πυκνότητας πιθανότητας που ακολουθούν οι πιθανότητες μετάβασης $U[P_{xy}^{t+1}(a)|P_{xy}^t(a)]$ (belief state), καθώς, σύμφωνα με αυτή τη συνάρτηση, μπορεί να παραβλέψει τη στοχαστικότητα του συστήματος χρησιμοποιώντας αναμενόμενες τιμές των κρίσιμων μεγεθών (τεχνική CEA- Certainty Equivalence Approximation). Τέλος, το κόστος που προκύπτει κάθε φορά που επιλέγεται μια δράση

αντικατοπτρίζει το πλήθος των βημάτων που απαιτούνται για να διανυθεί η διαδρομή από την πηγή στον προορισμό $C_{xy}(a)$. Η επίλυση του προβλήματος γίνεται με εφαρμογή του αλγορίθμου αναδρομικής αξίας με στόχο τη μείωση του κόστους.

3.3.2 Συμπεράσματα

Ο ελεγκτής του συστήματος δέχεται αιτήσεις δρομολόγησης προς κάποιον προορισμό με ρυθμό άφιξης κ . Εκτός του ρυθμού κ , λαμβάνεται υπόψη μία παράμετρος φ , η οποία χρησιμοποιείται για να δηλώσει την επιτυχή δρομολόγηση σε κάποια διαδρομή. Στόχος είναι η εκλογή των κατάλληλων διαδρομών, ώστε να ικανοποιούνται από πλευράς εκχωρούμενου εύρους ζώνης οι αιτήσεις δρομολόγησης. Επίσης, στόχος είναι η μείωση της συνολικής πιθανότητας αποκλεισμού (Call Blocking Probability) ώστε να διατηρείται υψηλή η ποιότητα υπηρεσίας στο δίκτυο (QoS). Σε δίκτυα με μεγαλύτερη πυκνότητα προκύπτει ότι η λύση του προβλήματος απαιτεί περισσότερα βήματα και άρα τα κόστη είναι μεγαλύτερα. Επιπλέον, προκύπτει ότι, όσο αυξάνονται οι τιμές των φ και κ τόσο αυξάνονται οι επιτυχείς δρομολογήσεις.

3.4 Προγραμματισμός παροχής υπηρεσιών στο στρώμα εφαρμογών

Η μαρκοβιανή θεωρία δεν έχει χρησιμοποιηθεί ακόμη σε μεγάλη έκταση για τη μελέτη προβλημάτων στο στρώμα εφαρμογής. Εντούτοις, ένα από τα ενδιαφέροντα θέματα που έχει διερευνηθεί είναι ο ορθός προγραμματισμός της λειτουργίας ενός πλήθους εξυπηρετητών για την παροχή ασυρμάτων υπηρεσιών δεδομένων με στόχο την εξυπηρέτηση όσο το δυνατόν περισσότερων αιτήσεων χρηστών (service scheduling). [14] Οι παράμετροι του προβλήματος είναι ο ρυθμός άφιξης των αιτήσεων, ο χρόνος απόκρισης των εξυπηρετητών και το κέρδος (ζημία) που αποφέρει η ολοκληρωμένη (αποτυχημένη) παροχή μιας υπηρεσίας.

3.4.1 Διατύπωση και επίλυση μαρκοβιανής διαδικασίας

Γίνεται η παραδοχή ότι η υπό μελέτη μαρκοβιανή διαδικασία εξελίσσεται σε n στάδια. Ένα τυχαίο στοιχείο του συνόλου καταστάσεων σε S είναι της μορφής $\sigma = (q_1, c_1, q_2, c_2, \dots, q_k, c_k)$, όπου q_i είναι το πλήθος των χρηστών της υπηρεσίας κλάσης i που αναμένουν να εξυπηρετηθούν και c_i το πλήθος των χρηστών της υπηρεσίας κλάσης i

που εξυπηρετούνται εκάστοτε, $i=1,2,3,\dots,K$. Για λόγους απλότητας το πλήθος των κλάσεων υπηρεσιών είναι $K=2$, ενώ το πλήθος των εξυπηρετητών $c_{\max}=10$. Θεωρείται ότι λαμβάνουν χώρα τρεις μερικές μεταβάσεις στο σύστημα οι οποίες δεν μπορούν να επηρεαστούν από τον έλεγχο του εξυπηρετητή: η άφιξη μιας αίτησης σε μία από τις κλάσεις υπηρεσίας με μηδενικό κέρδος για το σύστημα, η εξυπηρέτηση ενός χρήστη με θετικό κέρδος και η εγκατάλειψη της αίτησης ενός χρήστη με ζημία για το σύστημα. Για κάθε μετάβαση υπολογίζεται η πιθανότητα η μερική μετάβαση που συμβαίνει να είναι ένα εκ των ανωτέρω τριών ενδεχομένων. Το επόμενο βήμα είναι ο καθορισμός των κανόνων απόφασης $\delta(\sigma, \alpha, n)$ οι οποίοι επιλέγουν με πιθανότητα $p_{\sigma}^{\delta}(\tau|\alpha)$ την επόμενη κατάσταση τ του συστήματος, υπό την προϋπόθεση η παρούσα είναι η σ και λάβει χώρα η μερική μετάβαση α . Με εφαρμογή του κανόνα του Bayes οι πιθανότητες μετάβασης του συστήματος προκύπτουν από τη σχέση

$$p_{\sigma\tau}^{\delta} = \sum_{\alpha} p_{\sigma}^{\delta}(\tau|\alpha)p_{\sigma}(\alpha) \quad (3.4)$$

Το πρόβλημα επιλύεται με κλασικές μεθόδους δυναμικού προγραμματισμού. Η βέλτιστη πολιτική για «εξωτερικές» καταστάσεις του συστήματος, δηλαδή καταστάσεις στις οποίες τουλάχιστον η ουρά αναμονής ενός εξυπηρετητή είναι κενή, είναι πολιτική κατωφλίου. Αντίθετα, η βέλτιστη πολιτική για «εσωτερικές» καταστάσεις, όπου όλες οι ουρές αναμονής είναι μη κενές, ανεξάρτητα από το αν οι εξυπηρετητές είναι απασχολημένοι ή όχι, είναι μια επέκταση του $c\mu$ -κανόνα ($c\mu$ -rule)[15]. Ο $c\mu$ -κανόνας αποτελεί λύση κλειστής μορφής για ένα σύστημα αναμονής, η οποία επιβάλλει την επιλογή ανάμεσα στους χρήστες που αναμένουν να εξυπηρετηθούν, εκείνου που έχει το μεγαλύτερο γινόμενο κέρδους και ρυθμού εξυπηρέτησης. Στην παρούσα μαρκοβιανή διαδικασία επιδιώκεται να μεγιστοποιηθεί, εκτός από το γινόμενο κέρδους και ρυθμού εξυπηρέτησης, το γινόμενο της ποινής αστοχίας για τη μη εξυπηρέτηση ενός χρήστη επί το ρυθμό αστοχίας.

ΚΕΦΑΛΑΙΟ 4

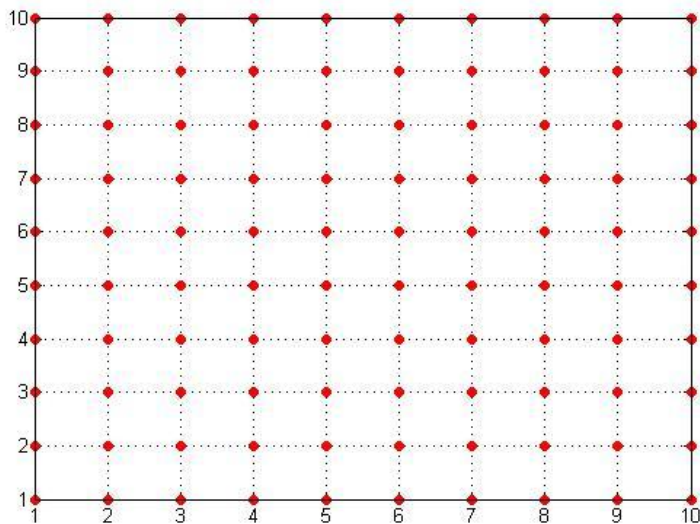
ΜΑΘΗΜΑΤΙΚΗ ΑΝΑΛΥΣΗ ΠΡΟΒΛΗΜΑΤΟΣ

4.1 Εισαγωγή

Η εφαρμογή που θα μελετηθεί με τη βοήθεια της μαρκοβιανής θεωρίας είναι το πρόβλημα δρομολόγησης σε ένα δίκτυο WMN στο οποίο οι ζεύξεις μεταξύ των κόμβων υποφέρουν από διαλείψεις λόγω βροχής.

4.2 Περιγραφή τοπολογίας δικτύου

Το ασύρματο δίκτυο πλέγματος που θα μελετηθεί αποτελείται από N^2 κόμβους ομοιόμορφα κατανεμημένους σε ένα τετραγωνικό πλέγμα διαστάσεων $N \times N$. Κάθε κόμβος του δικτύου μπορεί να εκπέμψει με τέτοια ισχύ ώστε να μπορεί να επικοινωνήσει με κάποιο από τους 8 γειτονικούς του κόμβους. Στο ίδιο πλέγμα θεωρείται ότι λαμβάνει χώρα βροχόπτωση, η οποία διακόπτει ορισμένες από τις ζεύξεις του πλέγματος λόγω διαλείψεων. Οποιοσδήποτε από τους κόμβους μπορεί να θεωρηθεί πηγή ή προορισμός ενώ στόχος είναι η εύρεση της βέλτιστης διαδρομής για τη δρομολόγηση των δεδομένων από την πηγή προς τον προορισμό. Βέλτιστη διαδρομή θεωρείται η διαδρομή με το ελάχιστο πλήθος βημάτων μεταξύ πηγής και προορισμού, η οποία χρησιμοποιεί ζεύξεις που δεν υποφέρουν από έντονες διαλείψεις λόγω βροχής.



Σχήμα 4.1 Τοπολογία δικτύου 100 κόμβων

4.2.1 Προσομοίωση τοπολογίας δικτύου

Η θέση έκαστου των $N \times N$ κόμβων περιγράφεται σε καρτεσιανές συντεταγμένες (x, y) . Το δίκτυο WMN συνολικά μπορεί να περιγραφεί από ένα δισδιάστατο πίνακα M , διαστάσεων $N \times N$, κάθε στοιχείο του οποίου αντιστοιχεί σε κόμβο και μπορεί να λάβει τις τιμές:

- 0, όταν ο κόμβος υφίσταται τέτοιου μεγέθους διαλείψεις λόγω βροχής ώστε να θεωρείται ανενεργός
- 1, όταν ο κόμβος είναι προσπελάσιμος και η ζεύξη δεν υποφέρει από διαλείψεις λόγω βροχής
- 2, όταν πρόκειται για τον κόμβο-πηγή που εκπέμπει δεδομένα προς τον προορισμό
- 3, όταν πρόκειται για τον κόμβο-προορισμό προς τον οποίο αποστέλλονται τα δεδομένα.

4.3 Κατάστρωση Μαρκοβιανού μοντέλου

Σύμφωνα με την ανωτέρω περιγραφή το υπό μελέτη δίκτυο μπορεί να περιγραφεί ως διακριτή μαρκοβιανή διαδικασία απόφασης δύο διαστάσεων με γνωστές τις πιθανότητες μετάβασης και στόχο την εύρεση της ελάχιστης δυνατής διαδρομής για την επιτυχή δρομολόγηση δεδομένων από τον κόμβο-πηγή προς τον κόμβο-προορισμό.

4.3.1 Σύνολο καταστάσεων του συστήματος

Η κατάσταση ενός ενεργού κόμβου του συστήματος, δηλαδή ενός κόμβου που αντιστοιχεί σε μη μηδενικό στοιχείο του πίνακα του δικτύου και έχει δεδομένα προς μετάδοση, δίνεται από το σύνολο:

$$S_j = \{x_j, y_j\}, \quad j \in [1, N^2] \quad (4.1)$$

όπου (x_j, y_j) οι συντεταγμένες της θέσης του κόμβου j .

Το σύνολο, λοιπόν, των δυνατών καταστάσεων στις οποίες μπορεί να βρεθεί το σύστημα κατά την εξέλιξή του, είναι:

$$S = \times_{j \in [1, N^2]} S_j \quad (4.2)$$

και περιλαμβάνει τις θέσεις όλων των ενεργών κόμβων του δικτύου.

Εκτός των ανωτέρω επιτρεπτών καταστάσεων, για την επίλυση του προβλήματος είναι αναγκαίο να θεωρηθεί μία επιπλέον κατάσταση στην οποία είναι δυνατό να βρεθεί το σύστημα, αλλά είναι ανεπιθύμητη. Η κατάσταση αυτή είναι η «κατάσταση ατυχήματος» όπου τα δεδομένα μεταδίδονται λανθασμένα από ένα ενεργό κόμβο σε οποιονδήποτε ανενεργό κόμβο του δικτύου, οπότε και υπάρχει απώλεια πακέτου.

4.3.2 Σύνολο δυνατών δράσεων του συστήματος

Για κάθε δυνατή κατάσταση του συστήματος είναι δυνατό να επιλεγεί μία δράση εκ των 9 δυνατών που περιλαμβάνονται στο σύνολο δράσεων:

$$A = \{(a_x, a_y) : (a_x, a_y) \in \{-1, 0, 1\} \times \{-1, 0, 1\}\} \quad (4.3)$$

Μέσω της δράσης που επιλέγεται κάθε φορά δίνεται η δυνατότητα αλλαγής της κατεύθυνσης μετάδοσης των δεδομένων με διπλό στόχο, αφενός, την αποφυγή μετάδοσης προς ανενεργούς κόμβους και, αφετέρου, την εύρεση της βέλτιστης διαδρομής δρομολόγησης προς τον κόμβο-προορισμό.

Ειδικότερα, η διεύθυνση μετάδοσης των δεδομένων αλλάζει προς την κατεύθυνση του διανύσματος που ορίζεται από το ζεύγος των συντεταγμένων (a_x, a_y) της δράσης που επιλέγεται. Η κατανόηση των συνεπειών της επιλογής μιας εκ των 9 δράσεων είναι σημαντική, καθώς η επίλυση του μαρκοβιανού μοντέλου συνίσταται, πρακτικά, στον ορθό καθορισμό της αλληλουχίας των δράσεων που πρέπει να επιλεγούν ώστε να βελτιστοποιηθεί η δρομολόγηση των δεδομένων προς τον κομβο-προορισμό ξεκινώντας από τον κόμβο-πηγή.

4.3.3 Πιθανότητες και πίνακας μετάβασης του συστήματος

Ως πιθανότητα μετάβασης p_{ij} ορίζεται η δεσμευμένη πιθανότητα η μετάδοση των δεδομένων να γίνει προς τον κόμβο j δεδομένου ότι ο τρέχων κόμβος είναι ο i και έχει

επιλεγεί η δράση (a_x, a_y) . Το μοντέλο καταστρώνεται με γνωστές πιθανότητες μετάβασης p_{ij} , ως εξής:

$$p_{ij} = p(s_j | s_i, (a_x, a_y)) = \begin{cases} p, & \text{η μετάδοση από τον κόμβο } i \text{ να μη πραγματοποιηθεί (αποτυχημένη} \\ & \text{εκπομπή) ανεξάρτητα από το ποιά είναι η βέλτιστη δράση} \\ 1 - p, & \text{η μετάδοση από τον κόμβο } i \text{ να γίνει προς τον κόμβο } j \text{ με εφαρμογή} \\ & \text{της δράσης που έχει προκύψει από τη βέλτιστη πολιτική} \end{cases} \quad (4.4)$$

όπου ο ορισμός αφορά κόμβους i και j που είναι ενεργοί και το μέγεθος $p \in [0,1]$ μπορεί να καθοριστεί ανάλογα με το υπό μελέτη δίκτυο.

Όταν το σύστημα βρεθεί στην τερματική κατάσταση s_j (j : κόμβος-προορισμός) θεωρείται ότι η πιθανότητα παραμονής σε αυτήν είναι μοναδιαία (absorbent state), δηλαδή:

$$p(s_j | s_j, (a_x, a_y))_{j:\text{κόμβος-προορισμός}} = 1 \quad \forall (a_x, a_y) \in A \quad (4.5)$$

Όταν το σύστημα βρεθεί στην «κατάσταση ατύχημα», δηλαδή όταν δρομολογηθούν δεδομένα προς ανενεργό κόμβο του δικτύου, το δρομολογηθέν πακέτο χάνεται, η δρομολόγηση αποτυγχάνει και γίνεται η παραδοχή ότι το σύστημα μεταπίπτει με πιθανότητα 1 στην τερματική κατάσταση:

$$p(s_j | s_i, (a_x, a_y))_{\substack{j:\text{κόμβος-προορισμός} \\ i:\text{ανενεργός κόμβος}}} = 1 \quad \forall (a_x, a_y) \in A \quad (4.6)$$

Εφόσον, ξεκινώντας από τον κόμβο-πηγή, όλες οι μεταδόσεις μεταξύ των κόμβων πραγματοποιηθούν με πιθανότητα $1-p$, η διαδρομή δρομολόγησης είναι η βέλτιστη, δηλαδή αυτή με τα ελάχιστα βήματα μεταξύ πηγής και προορισμού και χωρίς να υπάρχουν αποτυχημένες μεταδόσεις. Σε αντίθετη περίπτωση, στους κόμβους που μεταδίδουν με πιθανότητα p υπάρχει απώλεια των δεδομένων και απαιτείται αναμετάδοση. Είναι φανερό ότι όσο αυξάνεται η τιμή της πιθανότητας p τόσο περισσότερες αναμεταδόσεις πακέτων πραγματοποιούνται.

Με βάση τον ανωτέρω ορισμό των πιθανοτήτων μετάβασης επιβεβαιώνεται ότι η στοχαστική διαδικασία περιγραφής του δικτύου είναι μαρκοβιανή καθώς οι πιθανότητες

μετάβασης εξαρτώνται μόνο από την τρέχουσα κατάσταση και όχι από την προηγούμενη εξέλιξη των καταστάσεων του συστήματος. Επιπλέον, η πιθανότητα p δεν μεταβάλλεται με το χρόνο, οπότε η διαδικασία θεωρείται στατική.

Με βάση τις προαναφερθείσες εξισώσεις, για να επιλυθεί το πρόβλημα της δρομολόγησης απαιτείται να υπολογιστούν τα στοιχεία του πίνακα μετάβασης P . Ο πίνακας μετάβασης της μαρκοβιανής διαδικασίας είναι τρισδιάστατος $P: S \times S \times A$ και κάθε στοιχείο του (i, j, a) αποτελεί ουσιαστικά την πιθανότητα μετάβασης $p(s_j | s_i, (a_x, a_y))$. Μια σημαντική ιδιότητα του πίνακα P είναι ότι για κάθε δράση a , κάθε γραμμή του δισδιάστατου υποπίνακα $P(a): S \times S$ έχει μόνο 2 στοιχεία μη μηδενικά. Η ιδιότητα αυτή οφείλεται στο ότι, όταν επιλεγεί μία δράση ενόσω το σύστημα βρίσκεται σε μία από τις δυνατές καταστάσεις, μπορεί να συμβεί μετάβαση μόνο σε δύο καταστάσεις. Η μία πιθανή κατάσταση είναι αυτή που προκύπτει με εφαρμογή της επιλεγείσας δράσης και η μετάβαση σε αυτήν πραγματοποιείται με πιθανότητα $1-p$. Η άλλη πιθανή μετάβαση είναι η αποτυχημένη εκπομπή του πακέτου και η παραμονή στην τρέχουσα κατάσταση με πιθανότητα p . Το πρώτο μη μηδενικό στοιχείο, λοιπόν, του πίνακα μετάβασης έχει τιμή $1-p$ και είναι η κατάσταση του κόμβου στον οποίο αποστέλλονται τα δεδομένα, όπως θα έπρεπε, με εφαρμογή της βέλτιστης δράσης. Το δεύτερο μη μηδενικό στοιχείο έχει τιμή p και είναι η θέση του ίδιου του κόμβου (διαγώνιο στοιχείο) στον οποίο συμβαίνει αναμετάδοση των δεδομένων. Η ιδιότητα αυτή δεν ισχύει στις δύο τελευταίες γραμμές του πίνακα της «κατάστασης ατυχήματος» και της τερματικής κατάστασης που έχουν μόνο ένα μη μηδενικό στοιχείο με τιμή 1. Αυτό συμβαίνει καθώς, όταν το σύστημα βρεθεί σε μια από τις δύο αυτές καταστάσεις, μεταβαίνει πάντα στην τερματική κατάσταση με πιθανότητα 1.

4.3.4 Απώλειες του συστήματος

Το τελευταίο στοιχείο που απαιτείται για τον πλήρη καθορισμό των παραμέτρων της μαρκοβιανής διαδικασίας είναι ο πίνακας C των απωλειών του συστήματος. Εφόσον το δίκτυο που περιγράφεται μελετάται με σκοπό να βρεθεί η βέλτιστη διαδρομή δρομολόγησης, η απώλεια αποτιμάται σε πλήθος βημάτων.

Ειδικότερα η μετάβαση από μία κατάσταση s_i σε μία κατάσταση s_j , δηλαδή η αποστολή πακέτων από τον κόμβο i στον κόμβο j , με τον κόμβο i να είναι ενεργός και ανεξάρτητα

από το αν ο κόμβος j είναι ενεργός, ανενεργός ή κόμβος-προορισμός θεωρείται ότι αντιστοιχεί σε ένα βήμα. Η μετάβαση από την «κατάσταση ατύχημα», όπως αυτή έχει οριστεί, στην τερματική κατάσταση θεωρείται ότι έχει ποινή ίση με 100 βήματα ενώ όταν το σύστημα καταλήξει στην τερματική κατάσταση, παραμένει σε αυτήν χωρίς να παρουσιάζει απώλειες.

Συνοπτικά, όταν το σύστημα μεταπίπτει από την κατάσταση i στην κατάσταση j , οι απώλειες που προκύπτουν είναι ανεξάρτητες από τη δράση που έχει προκαλέσει την εκάστοτε μετάβαση και δίνονται από τη σχέση:

$$C(s_i, s_j) = \begin{cases} 1, & i: \text{ενεργός κόμβος} \\ 100, & i: \text{ανενεργός κόμβος και} \\ & j: \text{κόμβος προορισμός} \\ 0, & i = j: \text{κόμβος - προορισμός} \end{cases} \quad (4.7)$$

Για την επίλυση του προβλήματος απαιτείται ο πίνακας απωλειών του συστήματος, ο οποίος έχει ίδιες διαστάσεις και μορφή με τον πίνακα μετάβασης P . Συγκεκριμένα, ο πίνακας απωλειών είναι $C: S \times S \times A$ με στοιχεία που προκύπτουν με βάση την ανωτέρω σχέση. Στην περίπτωση όπου το πρόβλημα επιλύεται με μεγιστοποίηση του οφέλους και όχι με ελαχιστοποίηση των απωλειών, απαιτείται ο πίνακας οφέλους R , ο οποίος έχει ακριβώς τα ίδια στοιχεία με τον C αλλά με αντίθετο πρόσημο.

4.4 Επίλυση Μαρκοβιανού μοντέλου

Η μαρκοβιανή στοχαστική διαδικασία θεωρείται άπειρου ορίζοντα καθώς δεν είναι εκ των προτέρων γνωστό πόσος χρόνος απαιτείται για την εύρεση της βέλτιστης διαδρομής δρομολόγησης. Για τη διεκπεραίωση της διαδικασίας χρησιμοποιείται ο αλγόριθμος αναδρομικής πολιτικής (policy iteration) με στόχο τη μεγιστοποίηση του οφέλους. Συγκεκριμένα, ξεκινώντας από μία τυχαία πολιτική η οποία επιλέγεται ώστε να μεγιστοποιεί το άμεσο όφελος του συστήματος σε κάθε κατάσταση, ο αλγόριθμος αυτός, επιδιώκει σε κάθε επανάληψη να υπολογίζει ένα βελτιωμένο διάλυμα πολιτικής, αξιολογώντας το κάθε φορά μέσω του οφέλους που προκύπτει από την εφαρμογή του.

Το πρώτο βήμα του αλγορίθμου είναι ο προσδιορισμός μιας αρχικής άπληστης πολιτικής, δηλαδή μιας αλληλουχίας δράσεων που μεγιστοποιεί το άμεσο όφελος του συστήματος. Προς τούτο θεωρείται μηδενική αρχική συνάρτηση οφέλους του συστήματος, $V_0=0$ και επιλέγεται η πολιτική που μεγιστοποιεί το αναμενόμενο όφελος όπως αυτό εκφράζεται μέσω των εξισώσεων Bellman:

$$V = \max_{a \in A} \{R(s_i, a) + \sum_{s_j \in S} \{\lambda p(s_j | s_i, a) V_0(s_j)\}\} \quad (4.8)$$

$$\pi^* = \arg \max_{a \in A} V |_{V_0=0} = \arg \max_{a \in A} R(s_i, s_j, a) \quad (4.9)$$

Λόγω του άπειρου χρονικού ορίζοντα του μοντέλου, ο παράγοντας $\lambda \in [0,1)$ εισάγεται, για την εξασφάλιση της σύγκλισης του αλγορίθμου και χρησιμοποιείται για τον υπολογισμό της παρούσας αξίας του οφέλους του συστήματος. Στην περίπτωση όπου υπάρχουν δύο ή περισσότερες δράσεις a που μεγιστοποιούν τη συνάρτηση οφέλους V , επιλέγεται αυτή που προκύπτει πρώτη. Κατ' αυτόν τον τρόπο, προκύπτει ένα διάνυσμα μήκους ίσου με το πλήθος των καταστάσεων το οποίο υποδεικνύει σε κάθε κατάσταση ποιά κατάλληλη δράση πρέπει να επιλεγεί.

Το δεύτερο βήμα είναι η αξιολόγηση της πολιτικής που επελέγη, δηλαδή ο υπολογισμός του διανύσματος της συνάρτησης οφέλους που αντιστοιχεί στην πολιτική αυτή. Αυτό επιτυγχάνεται επιλύοντας την εξίσωση

$$V' = (I - \lambda * P')^{-1} * R' \quad (4.10)$$

όπου I ο μοναδιαίος πίνακας διαστάσεων $S \times S$, P' ο πίνακας μετάβασης που αντιστοιχεί στην πολιτική που επελέγη τελευταία και R' ο πίνακας οφέλους που αντιστοιχεί, επίσης, στην πολιτική π^* . Οι πίνακες αυτοί προκύπτουν από τους πίνακες μετάβασης P και οφέλους R του μοντέλου, λαμβάνοντας εκείνα τα στοιχεία που αντιστοιχούν στις στήλες των δράσεων που απαρτίζουν τη βέλτιστη πολιτική και στις γραμμές που υπολογίζονται από τη θέση των δράσεων αυτών στο διάνυσμα της βέλτιστης πολιτικής.

Το τρίτο βήμα είναι η επίλυση των εξισώσεων Bellman με συνάρτηση οφέλους το διάνυσμα V που έχει προκύψει από την αξιολόγηση της πολιτικής στο δεύτερο βήμα

$$V = \max_{a \in A} \{R(s_i, s_j, a) + \sum_{s_j \in S} \{\lambda p(s_j | s_i, a) V'(s_j)\}\} \quad (4.11)$$

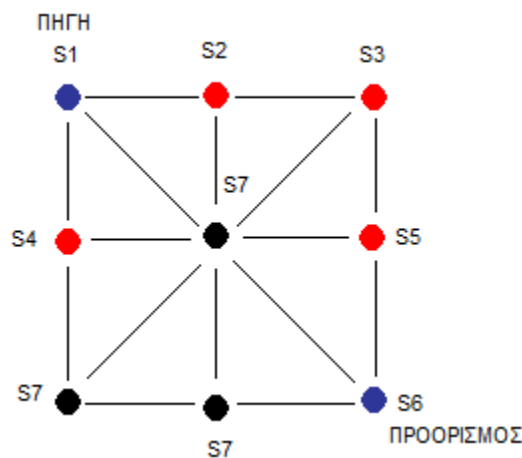
Επομένως, η νέα βέλτιστη πολιτική προκύπτει

$$\pi^* = \arg \max_{a \in A} V = \arg \max_{a \in A} \{R(s_i, s_j, a) + \sum_{s_j \in S} \{\lambda p(s_j | s_i, a) V'(s_j)\}\} \quad (4.12)$$

Στη συνέχεια επαναλαμβάνονται τα βήματα 2 και 3 μέχρι να προσδιοριστεί η βέλτιστη πολιτική. Ο αλγόριθμος συγκλίνει όταν προκύψουν δύο διαδοχικά διανύσματα πολιτικής, τα οποία δεν διαφέρουν σε κανένα στοιχείο, οπότε και τίθεται ως βέλτιστη πολιτική το διάνυσμα αυτό. Λόγω του άπειρου ορίζοντα του μοντέλου τίθεται ένα άνω όριο επαναλήψεων $N=1000$. Υπέρβαση του ορίου αυτού οδηγεί στον τερματισμό του αλγορίθμου και η πολιτική προκύπτει προσεγγιστικά.

Εφαρμογή 4.1

Ας θεωρηθεί το κατωτέρω δίκτυο πλέγματος 9 κόμβων (3x3):



Σχήμα 4.2 Τοπολογία δικτύου 9 κόμβων

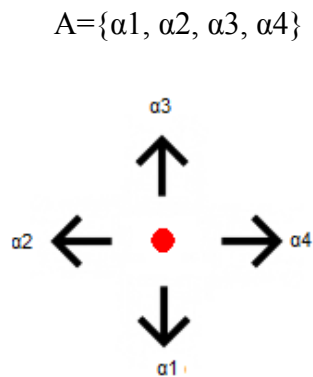
Στόχος είναι η επιτυχής δρομολόγηση δεδομένων από τον κόμβο S1 στον κόμβο S6 μέσω της διαδρομής που περιλαμβάνει το μικρότερο πλήθος βημάτων. Οι κόμβοι με ένδειξη S7 είναι ανενεργοί και συνεπώς δεν μπορούν να χρησιμοποιηθούν για να δρομολογήσουν πακέτα.

Κατάστρωση μαρκοβιανού μοντέλου

Καταστάσεις: Ως καταστάσεις του συστήματος θεωρούνται οι θέσεις των 6 ενεργών κόμβων του δικτύου. Επιπλέον, θεωρείται μια έβδομη κατάσταση ώστε να περιγραφούν όλες οι απαγορευμένες καταστάσεις. Απαγορευμένες καταστάσεις θεωρούνται, σε πρώτη φάση, οι θέσεις των ανενεργών κόμβων προς τις οποίες δεν μπορούν να δρομολογηθούν πακέτα. Ο όρος απαγορευμένη κατάσταση, όμως, χρησιμοποιείται για να συμπεριληφθούν και τα φαινόμενα των άκρων, καθώς οι συνοριακοί κόμβοι του δικτύου δεν επικοινωνούν με τέσσερις γειτονικούς κόμβους. Κατ'έκταση αν σε έναν από τους συνοριακούς κόμβους επιλεγεί η αποστολή των δεδομένων σε κατεύθυνση όπου δεν υπάρχει γειτονικός κόμβος, θεωρείται ότι το σύστημα μεταβαίνει σε απαγορευμένη κατάσταση.

$S = \{S1, S2, S3, S4, S5, S6, S7\}$, όπου S7: κατάσταση ατυχήματος

Δράσεις: Χάριν απλότητας θεωρούνται 4 δυνατές δράσεις δρομολόγησης από κάθε κόμβο, κίνηση προς τα κάτω, προς τα αριστερά, προς τα πάνω και προς τα δεξιά, όπως φαίνεται στο ακόλουθο σχήμα :



Σχήμα 4.3 Δυνατές δράσεις σε κάθε κόμβο του δικτύου

Πιθανότητες: Όταν το σύστημα βρίσκεται σε κάποια κατάσταση, δηλαδή σε κάποιο κόμβο i :

(α) Με πιθανότητα 0.9 το πακέτο μεταδίδεται από τον κόμβο i προς τον κόμβο j όπως επιβάλλει η εκάστοτε δράση

(β) Με πιθανότητα 0.1 υπάρχει απώλεια πακέτου και απαιτείται αναμετάδοση, οπότε το σύστημα παραμένει στην κατάσταση i , δηλαδή το πακέτο παραμένει στον κόμβο i .

Για παράδειγμα, αν στην κατάσταση $S1$ επιλεγεί η δράση $a1$, με πιθανότητα 0.9 το σύστημα μεταβαίνει στην κατάσταση $S4$ ενώ με πιθανότητα 0.1 το σύστημα παραμένει στην κατάσταση $S1$. Αυτό εκφράζεται μέσω των δεσμευμένων πιθανοτήτων

$$p(S_4|S_1, a_1) = 0.9 \quad \text{και} \quad p(S_1|S_1, a_1) = 0.1$$

Επιπλέον, αν τα δεδομένα δρομολογηθούν προς ανενεργό κόμβο ή επιλεγεί κατεύθυνση που οδηγεί εκτός του πλέγματος και το σύστημα βρεθεί στην κατάσταση ατυχήματος $S7$, γίνεται η παραδοχή ότι με πιθανότητα 1 μεταπίπτει στην τερματική κατάσταση $S6$. Τέλος, όταν τα δεδομένα δρομολογηθούν στον κόμβο προορισμό $S6$, με πιθανότητα 1 το σύστημα παραμένει στην τερματική αυτή κατάσταση (absorbent state).

Στη συνέχεια παρουσιάζονται αναλυτικά οι πίνακες μετάβασης για κάθε δράση a_i $i=1,2,3,4$.

$$\text{Δράση } a_1: \quad P(i, j) = P_{ij} = P(S_j|S_i, a_1) = \begin{bmatrix} 0.1 & 0 & 0 & 0.9 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0.1 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0.1 & 0.9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\text{Δράση } a_2: \quad P(i, j) = P_{ij} = P(S_j|S_i, a_2) = \begin{bmatrix} 0.1 & 0 & 0 & 0 & 0 & 0 & 0.9 \\ 0.9 & 0.1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.9 & 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0.1 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\text{Δράση } a_3: \quad P(i, j) = P_{ij} = P(S_j|S_i, a_3) = \begin{bmatrix} 0.1 & 0 & 0 & 0 & 0 & 0 & 0.9 \\ 0 & 0.1 & 0 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0.1 & 0 & 0 & 0 & 0.9 \\ 0.9 & 0 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0.9 & 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\text{Δράση α4: } P(i, j) = P_{ij} = P(S_j | S_i, a_4) = \begin{bmatrix} 0.1 & 0.9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0.9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0.1 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0.1 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Για παράδειγμα, αν στην κατάσταση S2 επιλεγεί η δράση α1, με πιθανότητα 0.9 το σύστημα μεταβαίνει στην κατάσταση S7 ενώ με πιθανότητα 0.1 το σύστημα παραμένει στην κατάσταση S2. Συνεπώς, στον πίνακα μετάβασης της δράσης α1 το στοιχείο $P_{22}=P(2,2)=0.1$ και το στοιχείο $P_{27}=P(2,7)=0.9$. Όταν το σύστημα βρεθεί στην κατάσταση ατυχήματος S7 ή στην τερματική κατάσταση S6, η μετάβασή του γίνεται με βεβαιότητα στην τερματική κατάσταση S6 ανεξάρτητα από την επιλεγείσα δράση. Επομένως, στον πίνακα μετάβασης εκάστης των 4 δράσεων ισχύει

$$P_{76}=P(7,6)=1 \text{ και } P_{66}=P(6,6)=1$$

Κατ'αυτόν τον τρόπο, συμπληρώνεται ο πίνακας μετάβασης για έκαστη των 4 δυνατών δράσεων.

Απώλειες: Όταν μετά την εφαρμογή μιας δράσης το σύστημα μεταπίπτει από μία κατάσταση σε μία άλλη προκύπτει απώλεια η οποία εκφράζεται σε βήματα. Η μετάβαση από ένα ενεργό κόμβο σε οποιονδήποτε άλλο ενεργό ή μη κόμβο θεωρείται ότι κοστίζει ένα βήμα, ενώ αν το σύστημα μεταπέσει σε κατάσταση ατυχήματος υπάρχει απώλεια 100 βημάτων. Τέλος, όταν τα δεδομένα φθάσουν στον προορισμό το σύστημα έχει μηδενική απώλεια. Συνοπτικά, εκφράζοντας τις απώλειες ως αρνητικά οφέλη ισχύει

$$R(S_i, S_j) = \begin{cases} -1, & S_i: \text{ενεργός κόμβος} \\ -100, & S_i: \text{ανενεργός κόμβος και} \\ & S_j: \text{κόμβος προορισμός} \\ 0, & S_i = S_j: \text{κόμβος - προορισμός} \end{cases}$$

Επίλυση μαρκοβιανού μοντέλου

Το πρόβλημα θα λυθεί με τον αλγόριθμο της αναδρομικής πολιτικής όπως αυτός παρουσιάστηκε στο εδ. 2.4.5. Σύμφωνα με τη λογική του αλγορίθμου αυτού, αρχικά επιλέγεται κατά τυχαίο τρόπο μια αλληλουχία δράσεων, δηλαδή για κάθε κατάσταση επιλέγεται κατά τυχαίο τρόπο μια κατεύθυνση δρομολόγησης των δεδομένων. Στη

συνέχεια, οι δράσεις που επιλέχθηκαν αξιολογούνται μέσω του αναμενόμενου οφέλους που αποφέρουν στο δίκτυο και σε κάθε κατάσταση αντιστοιχίζεται μια τιμή οφέλους. Οι τιμές οφέλους χρησιμοποιούνται ώστε σε κάθε κατάσταση να επιλεγεί μια βελτιωμένη δράση που θα αποφέρει αύξηση του οφέλους του συστήματος. Όταν μετά από δύο διαδοχικές επαναλήψεις του αλγορίθμου προκύψει το ίδιο διάνυσμα πολιτικής, δηλαδή η ίδια ακολουθία δράσεων για τις καταστάσεις του συστήματος, το διάνυσμα καθορίζει τη βέλτιστη διαδρομή δρομολόγησης.

Στην περίπτωση του υπό μελέτη δικτύου δεν επιλέγεται μια τυχαία αρχική πολιτική αλλά μία άπληστη πολιτική. Συγκεκριμένα, σε κάθε κατάσταση επιλέγεται η δράση που μεγιστοποιεί το άμεσο αναμενόμενο όφελος για την κατάσταση αυτή. Αυτό προκύπτει με εφαρμογή του κανόνα του Bayes

$$R(S_i, a) = \sum_j p(S_j | S_i, a) R(S_i, S_j)$$

Στη συνέχεια, με βάση την ανωτέρω σχέση, για κάθε κατάσταση υπολογίζεται το άμεσο όφελος που προκύπτει από την εφαρμογή έκαστης των 4 δυνατών δράσεων:

Κατάσταση S1:

$$R(S1, a1) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S1, a2) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S1, a3) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S1, a4) = 0.1 * (-1) + 0.9 * (-1) = -1$$

Κατάσταση S2:

$$R(S2, a1) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S2, a2) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S2, a3) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S2, a4) = 0.1 * (-1) + 0.9 * (-1) = -1$$

Κατάσταση S3:

$$R(S3, a1) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S3, a2) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S3, a3) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S3, a4) = 0.1 * (-1) + 0.9 * (-1) = -1$$

Κατάσταση S4:

$$R(S4, a1) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S4, a2) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S4, a3) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S4, a4) = 0.1 * (-1) + 0.9 * (-1) = -1$$

Κατάσταση S5:

$$R(S5, a1) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S5, a2) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S5, a3) = 0.1 * (-1) + 0.9 * (-1) = -1$$

$$R(S5, a4) = 0.1 * (-1) + 0.9 * (-1) = -1$$

Κατάσταση S6:

$$R(S6, a1) = 1 * 0 = 0$$

$$R(S6, a2) = 1 * 0 = 0$$

$$R(S6, a3) = 1 * 0 = 0$$

$$R(S6, a4) = 1 * 0 = 0$$

Κατάσταση S7:

$$R(S7, a1) = 1 * 100 = 100$$

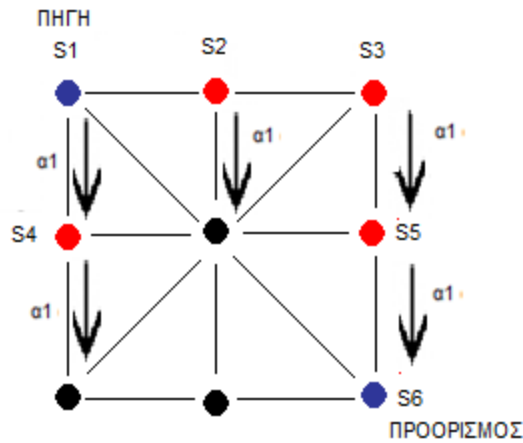
$$R(S7, a2) = 1 * 100 = 100$$

$$R(S7, a3) = 1 * 100 = 100$$

$$R(S7, a4) = 1 * 100 = 100$$

Όπως φαίνεται όλες οι δράσεις έχουν το ίδιο άμεσο όφελος για κάθε κατάσταση. Στην περίπτωση αυτή, επιλέγεται κατά σύμβαση η πρώτη από τις δράσεις που έχουν το ίδιο αποτέλεσμα με κριτήριο το όφελος. Συνεπώς, η αρχική άπληστη πολιτική είναι:

$S1 \rightarrow \alpha1, S2 \rightarrow \alpha1, S3 \rightarrow \alpha1, S4 \rightarrow \alpha1, S5 \rightarrow \alpha1, S6 \rightarrow \alpha1, S7 \rightarrow \alpha1$



Σχήμα 4.4 Άπληστη πολιτική -Επιλεγείσες δράσεις σε κάθε κόμβο

Η πολιτική αυτή δεν είναι βέλτιστη, καθώς όταν η πηγή S1 εκπέμπει θα επιλέξει τη δράση α1 και η εκπομπή θα πραγματοποιηθεί προς τον κόμβο S4. Στον κόμβο αυτό θα επιλεγεί και πάλι η δράση α1 και τα δεδομένα θα σταλούν σε ανενεργό κόμβο. Συνεπώς, με εφαρμογή της πολιτικής αυτής η δρομολόγηση αποτυγχάνει.

1^η επανάληψη:

Αξιολόγηση της πολιτικής

Στο βήμα αυτό αξιολογείται η πολιτική μέσω του οφέλους που της αντιστοιχεί επιλύοντας την εξίσωση:

$$(I - \lambda P_{d_n})u_n = R_{d_n} \leftrightarrow u_n = (I - \lambda P_{d_n})^{-1}R_{d_n}$$

όπου u_n είναι διάνυσμα του οποίου στοιχεία είναι οι συναρτήσεις οφέλους έκαστης των καταστάσεων του συστήματος. Το μέγεθος λ είναι ο παράγοντας μείωσης του οφέλους του συστήματος (discount factor) και λαμβάνεται $\lambda=0.99$. Οι πίνακες P_{d_n} και R_{d_n} είναι οι πίνακες μετάβασης και οφέλους που αντιστοιχούν στην τελευταία επιλεγείσα πολιτική. Στη συγκεκριμένη περίπτωση, ο πίνακας P_{d_n} είναι ο πίνακας μετάβασης της δράσης α1, αφού η δράση α1 έχει επιλεγεί για όλες ανεξαρτήτως τις καταστάσεις. Ο πίνακας R_{d_n} είναι το διάνυσμα που περιλαμβάνει τα άμεσα οφέλη για κάθε κατάσταση από την εφαρμογή της δράσης α1:

$$R_{d_n} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 0 \\ -100 \end{bmatrix}$$

$$\text{Οπότε } u_n = \left(\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} - 0.99 * \begin{bmatrix} 0.1 & 0 & 0 & 0.9 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0.1 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0.1 & 0.9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \right)^{-1} \cdot \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 0 \\ -100 \end{bmatrix} = \begin{bmatrix} -100 \\ -100 \\ -2.207 \\ -100 \\ -1.11 \\ 0 \\ -100 \end{bmatrix}$$

Κατ' αυτόν τον τρόπο, σε κάθε κατάσταση από τις 7 δυνατές του δικτύου αποδίδεται τιμή του οφέλους που αποφέρει η δράση που έχει επιλεγεί.

Βελτίωση της πολιτικής

Έχοντας αξιολογήσει την πολιτική εφαρμόζονται οι εξισώσεις Bellman για κάθε κατάσταση S και για κάθε δράση a και επιλέγεται για κάθε κατάσταση η δράση που μεγιστοποιεί τις εξισώσεις αυτές, δηλαδή

$$d_{n+1}(s) \in \arg \max_{a \in A_s} \{R(s, a) + \sum_{j \in S} \lambda p(j|s, a) u_n(j)\}$$

Κατάσταση S1:

$$V(S1, a1) = -1 + 0.99(0.1 * u(1) + 0.9 * u(4)) = -100$$

$$V(S1, a2) = -1 + 0.99(0.1 * u(1) + 0.9 * u(7)) = -100$$

$$V(S1, a3) = -1 + 0.99(0.1 * u(1) + 0.9 * u(7)) = -100$$

$$V(S1, a4) = -1 + 0.99(0.1 * u(1) + 0.9 * u(2)) = -100$$

Κατάσταση S2:

$$V(S2, a1) = -1 + 0.99(0.1 * u(2) + 0.9 * u(7)) = -100$$

$$V(S2, a2) = -1 + 0.99(0.1 * u(2) + 0.9 * u(1)) = -100$$

$$V(S2, a3) = -1 + 0.99(0.1 * u(2) + 0.9 * u(7)) = -100$$

$$V(S2, a4) = -1 + 0.99(0.1 * u(2) + 0.9 * u(3)) = -12.866$$

Κατάσταση S3:

$$V(S3, a1) = -1 + 0.99(0.1 * u(3) + 0.9 * u(5)) = -2.208$$

$$V(S3, a2) = -1 + 0.99(0.1 * u(3) + 0.9 * u(2)) = -90.318$$

$$V(S3, a3) = -1 + 0.99(0.1 * u(3) + 0.9 * u(7)) = -90.318$$

$$V(S3, a4) = -1 + 0.99(0.1 * u(3) + 0.9 * u(7)) = -90.318$$

Κατάσταση S4:

$$V(S4, a1) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -100$$

$$V(S4, a2) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -100$$

$$V(S4, a3) = -1 + 0.99(0.1 * u(4) + 0.9 * u(1)) = -100$$

$$V(S4, a4) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -100$$

Κατάσταση S5:

$$V(S5, a1) = -1 + 0.99(0.1 * u(5) + 0.9 * u(6)) = -1.11$$

$$V(S5, a2) = -1 + 0.99(0.1 * u(5) + 0.9 * u(7)) = -90.21$$

$$V(S5, a3) = -1 + 0.99(0.1 * u(5) + 0.9 * u(3)) = -3.076$$

$$V(S5, a4) = -1 + 0.99(0.1 * u(5) + 0.9 * u(7)) = -90.21$$

Κατάσταση S6:

$$V(S6, a1) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a2) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a3) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a4) = 0 + 0.99(1 * u(6)) = 0$$

Κατάσταση S7:

$$V(S7, a1) = -100 + 0.99(1 * u(6)) = -100$$

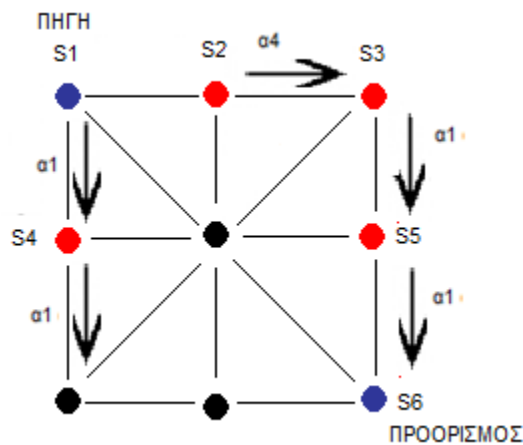
$$V(S7, a2) = -100 + 0.99(1 * u(6)) = -100$$

$$V(S7, a3) = -100 + 0.99(1 * u(6)) = -100$$

$$V(S7, a4) = -100 + 0.99(1 * u(6)) = -100$$

Σύμφωνα με τα παραπάνω αποτελέσματα η νέα πολιτική είναι:

$S1 \rightarrow \alpha1, S2 \rightarrow \alpha4, S3 \rightarrow \alpha1, S4 \rightarrow \alpha1, S5 \rightarrow \alpha1, S6 \rightarrow \alpha1, S7 \rightarrow \alpha1$



Σχήμα 4.5 Επιλεγείσες δράσεις μετά την πρώτη επανάληψη του αλγορίθμου

2^η επανάληψη:

Αξιολόγηση της νέας πολιτικής

Επιλύεται εκ νέου η εξίσωση

$$(I - \lambda P_{a_n})u_n = R_{a_n} \leftrightarrow u_n = (I - \lambda P_{a_n})^{-1}R_{a_n}$$

Εφόσον, όμως, για την κατάσταση S2 έχει επιλεγεί ως βέλτιστη η δράση α4, ο πίνακας P_{a_n} έχει αλλάξει και η δεύτερη γραμμή του είναι η δεύτερη γραμμή του πίνακα μετάβασης της δράσης α4. Αντιθέτως ο πίνακας R_{a_n} παραμένει ίδιος.

$$\text{Οπότε } u_n = \begin{pmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} - 0.99 * \begin{bmatrix} 0.1 & 0 & 0 & 0.9 & 0 & 0 & 0 \\ 0 & 0.1 & 0.9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0.1 & 0.9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{pmatrix}^{-1} \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 0 \\ -100 \end{bmatrix} = \begin{bmatrix} -100 \\ -3.293 \\ -2.207 \\ -100 \\ -1.11 \\ 0 \\ -100 \end{bmatrix}$$

Βελτίωση της πολιτικής

Κατάσταση S1:

$$V(S1, a1) = -1 + 0.99(0.1 * u(1) + 0.9 * u(4)) = -100$$

$$V(S1, a2) = -1 + 0.99(0.1 * u(1) + 0.9 * u(7)) = -100$$

$$V(S1, a3) = -1 + 0.99(0.1 * u(1) + 0.9 * u(7)) = -100$$

$$V(S1, a4) = -1 + 0.99(0.1 * u(1) + 0.9 * u(2)) = -13.384$$

Κατάσταση S2:

$$V(S2, a1) = -1 + 0.99(0.1 * u(2) + 0.9 * u(7)) = -90.426$$

$$V(S2, a2) = -1 + 0.99(0.1 * u(2) + 0.9 * u(1)) = -90.426$$

$$V(S2, a3) = -1 + 0.99(0.1 * u(2) + 0.9 * u(7)) = -90.426$$

$$V(S2, a4) = -1 + 0.99(0.1 * u(2) + 0.9 * u(3)) = -3.292$$

Κατάσταση S3:

$$V(S3, a1) = -1 + 0.99(0.1 * u(3) + 0.9 * u(5)) = -2.208$$

$$V(S3, a2) = -1 + 0.99(0.1 * u(3) + 0.9 * u(2)) = -4.153$$

$$V(S3, a3) = -1 + 0.99(0.1 * u(3) + 0.9 * u(7)) = -90.318$$

$$V(S3, a4) = -1 + 0.99(0.1 * u(3) + 0.9 * u(7)) = -90.318$$

Κατάσταση S4:

$$V(S4, a1) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -100$$

$$V(S4, a2) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -100$$

$$V(S4, a3) = -1 + 0.99(0.1 * u(4) + 0.9 * u(1)) = -100$$

$$V(S4, a4) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -100$$

Κατάσταση S5:

$$V(S5, a1) = -1 + 0.99(0.1 * u(5) + 0.9 * u(6)) = -1.11$$

$$V(S5, a2) = -1 + 0.99(0.1 * u(5) + 0.9 * u(7)) = -90.21$$

$$V(S5, a3) = -1 + 0.99(0.1 * u(5) + 0.9 * u(3)) = -3.076$$

$$V(S5, a4) = -1 + 0.99(0.1 * u(5) + 0.9 * u(7)) = -90.21$$

Κατάσταση S6:

$$V(S6, a1) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a2) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a3) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a4) = 0 + 0.99(1 * u(6)) = 0$$

Κατάσταση S7:

$$V(S7, a1) = -100 + 0.99(1 * u(6)) = -100$$

$$V(S7, a2) = -100 + 0.99(1 * u(6)) = -100$$

$$V(S7, a3) = -100 + 0.99(1 * u(6)) = -100$$

$$V(S7, a4) = -100 + 0.99(1 * u(6)) = -100$$

Επομένως το νέο διάνυσμα πολιτικής είναι: $\alpha_4, \alpha_4, \alpha_1, \alpha_1, \alpha_1, \alpha_1$

$S1 \rightarrow \alpha_4, S2 \rightarrow \alpha_4, S3 \rightarrow \alpha_1, S4 \rightarrow \alpha_1, S5 \rightarrow \alpha_1, S6 \rightarrow \alpha_1, S7 \rightarrow \alpha_1$

3^η επανάληψη :

Αξιολόγηση της νέας πολιτικής

$$(I - \lambda P_{d_n})u_n = R_{d_n} \leftrightarrow u_n = (I - \lambda P_{d_n})^{-1}R_{d_n}$$

$$u_n = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} - 0.99 * \begin{pmatrix} 0.1 & 0.9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0.9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0.1 & 0.9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 0 \\ -100 \end{pmatrix} = \begin{pmatrix} -4.366 \\ -3.293 \\ -2.207 \\ -100 \\ -1.11 \\ 0 \\ -100 \end{pmatrix}$$

Βελτίωση της πολιτικής

Κατάσταση S1:

$$V(S1, a1) = -1 + 0.99(0.1 * u(1) + 0.9 * u(4)) = -90.532$$

$$V(S1, a2) = -1 + 0.99(0.1 * u(1) + 0.9 * u(7)) = -90.532$$

$$V(S1, a3) = -1 + 0.99(0.1 * u(1) + 0.9 * u(7)) = -90.532$$

$$V(S1, a4) = -1 + 0.99(0.1 * u(1) + 0.9 * u(2)) = -4.366$$

Κατάσταση S2:

$$V(S2, a1) = -1 + 0.99(0.1 * u(2) + 0.9 * u(7)) = -90.426$$

$$V(S2, a2) = -1 + 0.99(0.1 * u(2) + 0.9 * u(1)) = -5.216$$

$$V(S2, a3) = -1 + 0.99(0.1 * u(2) + 0.9 * u(7)) = -90.426$$

$$V(S2, a4) = -1 + 0.99(0.1 * u(2) + 0.9 * u(3)) = -3.292$$

Κατάσταση S3:

$$V(S3, a1) = -1 + 0.99(0.1 * u(3) + 0.9 * u(5)) = -2.208$$

$$V(S3, a2) = -1 + 0.99(0.1 * u(3) + 0.9 * u(2)) = -4.153$$

$$V(S3, a3) = -1 + 0.99(0.1 * u(3) + 0.9 * u(7)) = -90.318$$

$$V(S3, a4) = -1 + 0.99(0.1 * u(3) + 0.9 * u(7)) = -90.318$$

Κατάσταση S4:

$$V(S4, a1) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -100$$

$$V(S4, a2) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -100$$

$$V(S4, a3) = -1 + 0.99(0.1 * u(4) + 0.9 * u(1)) = -14.79$$

$$V(S4, a4) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -100$$

Κατάσταση S5:

$$V(S5, a1) = -1 + 0.99(0.1 * u(5) + 0.9 * u(6)) = -1.11$$

$$V(S5, a2) = -1 + 0.99(0.1 * u(5) + 0.9 * u(7)) = -90.21$$

$$V(S5, a3) = -1 + 0.99(0.1 * u(5) + 0.9 * u(3)) = -3.076$$

$$V(S5, a4) = -1 + 0.99(0.1 * u(5) + 0.9 * u(7)) = -90.21$$

Κατάσταση S6:

$$V(S6, a1) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a2) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a3) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a4) = 0 + 0.99(1 * u(6)) = 0$$

Κατάσταση S7:

$$V(S7, a1) = -100 + 0.99(1 * u(6)) = -100$$

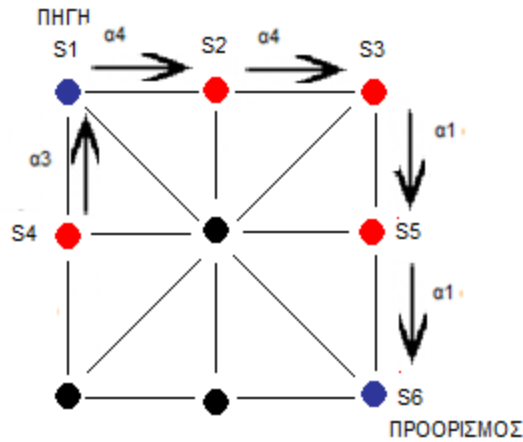
$$V(S7, a2) = -100 + 0.99(1 * u(6)) = -100$$

$$V(S7, a3) = -100 + 0.99(1 * u(6)) = -100$$

$$V(S7, a4) = -100 + 0.99(1 * u(6)) = -100$$

Επομένως το νέο διάνυσμα πολιτικής είναι

S1 → α4, S2 → α4, S3 → α1, S4 → α3, S5 → α1, S6 → α1, S7 → α1



Σχήμα 4.6 Επιλεγείσες δράσεις μετά την τρίτη επανάληψη του αλγορίθμου

Μετά την τρίτη επανάληψη έχει πλέον προσδιοριστεί η βέλτιστη δράση για κάθε κόμβο του δικτύου, οπότε στην επόμενη επανάληψη αναμένεται να μην υπάρξει αλλαγή στο διάνυσμα της πολιτικής.

4^η επανάληψη:

Αξιολόγηση της νέας πολιτικής

$$u_n = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} - 0.99 * \begin{bmatrix} 0.1 & 0.9 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0.9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0.9 & 0 & 0 \\ 0.9 & 0 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.1 & 0.9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 0 \\ -100 \end{bmatrix} = \begin{bmatrix} -4.366 \\ -3.293 \\ -2.207 \\ -5.428 \\ -1.11 \\ 0 \\ -100 \end{bmatrix}$$

Βελτίωση της πολιτικής

Κατάσταση S1:

$$V(S1, a1) = -1 + 0.99(0.1 * u(1) + 0.9 * u(4)) = -6.269$$

$$V(S1, a2) = -1 + 0.99(0.1 * u(1) + 0.9 * u(7)) = -90.532$$

$$V(S1, a3) = -1 + 0.99(0.1 * u(1) + 0.9 * u(7)) = -90.532$$

$$V(S1, a4) = -1 + 0.99(0.1 * u(1) + 0.9 * u(2)) = -4.366$$

Κατάσταση S2:

$$V(S2, a1) = -1 + 0.99(0.1 * u(2) + 0.9 * u(7)) = -90.426$$

$$V(S2, a2) = -1 + 0.99(0.1 * u(2) + 0.9 * u(1)) = -5.216$$

$$V(S2, a3) = -1 + 0.99(0.1 * u(2) + 0.9 * u(7)) = -90.426$$

$$V(S2, a4) = -1 + 0.99(0.1 * u(2) + 0.9 * u(3)) = -3.292$$

Κατάσταση S3:

$$V(S3, a1) = -1 + 0.99(0.1 * u(3) + 0.9 * u(5)) = -2.208$$

$$V(S3, a2) = -1 + 0.99(0.1 * u(3) + 0.9 * u(2)) = -4.153$$

$$V(S3, a3) = -1 + 0.99(0.1 * u(3) + 0.9 * u(7)) = -90.318$$

$$V(S3, a4) = -1 + 0.99(0.1 * u(3) + 0.9 * u(7)) = -90.318$$

Κατάσταση S4:

$$V(S4, a1) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -91.543$$

$$V(S4, a2) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -91.543$$

$$V(S4, a3) = -1 + 0.99(0.1 * u(4) + 0.9 * u(1)) = -5.427$$

$$V(S4, a4) = -1 + 0.99(0.1 * u(4) + 0.9 * u(7)) = -91.543$$

Κατάσταση S5:

$$V(S5, a1) = -1 + 0.99(0.1 * u(5) + 0.9 * u(6)) = -1.11$$

$$V(S5, a2) = -1 + 0.99(0.1 * u(5) + 0.9 * u(7)) = -90.21$$

$$V(S5, a3) = -1 + 0.99(0.1 * u(5) + 0.9 * u(3)) = -3.076$$

$$V(S5, a4) = -1 + 0.99(0.1 * u(5) + 0.9 * u(7)) = -90.21$$

Κατάσταση S6:

$$V(S6, a1) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a2) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a3) = 0 + 0.99(1 * u(6)) = 0$$

$$V(S6, a4) = 0 + 0.99(1 * u(6)) = 0$$

Κατάσταση S7:

$$V(S7, a1) = -100 + 0.99(1 * u(6)) = -100$$

$$V(S7, a2) = -100 + 0.99(1 * u(6)) = -100$$

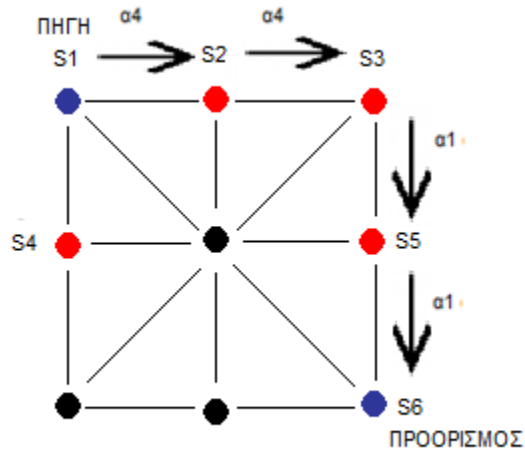
$$V(S7, a3) = -100 + 0.99(1 * u(6)) = -100$$

$$V(S7, a4) = -100 + 0.99(1 * u(6)) = -100$$

Επομένως το νέο διάνυσμα πολιτικής είναι

$S1 \rightarrow a4, S2 \rightarrow a4, S3 \rightarrow a1, S4 \rightarrow a3, S5 \rightarrow a1, S6 \rightarrow a1, S7 \rightarrow a1$

Επειδή στις δύο τελευταίες επαναλήψεις του αλγορίθμου προέκυψε η ίδια ακολουθία δράσεων, θεωρείται ότι ο αλγόριθμος έχει συγκλίνει. Όπως φαίνεται στο ακόλουθο σχήμα η πολιτική που προέκυψε από την επίλυση του αλγορίθμου καθορίζει τη βέλτιστη διαδρομή δρομολόγησης.



Σχήμα 4.7 Βέλτιστη διαδρομή δρομολόγησης

ΚΕΦΑΛΑΙΟ 5

ΠΕΡΙΒΑΛΛΟΝ ΠΡΟΣΟΜΟΙΩΣΗΣ ΚΑΙ ΑΡΙΘΜΗΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1 Εισαγωγή

Στο κεφάλαιο αυτό προσομοιώνονται διάφορες τοπολογίες δικτύων ως μαρκοβιανές διαδικασίες απόφασης και μελετάται σε κάθε μία από αυτές το πρόβλημα της δρομολόγησης υπο συνθήκες διαλείψεων λόγω βροχής. Για το σκοπό αυτό, αρχικά περιγράφεται ένας αλγόριθμος προσομοίωσης του φυσικού διαύλου του προβλήματος, που υποφέρει από διαλείψεις λόγω βροχόπτωσης. Στη συνέχεια, επί του συγκεκριμένου διαύλου θεωρούνται τοπολογίες διαφορετικού πλήθους κόμβων και μελετάται η επίδραση των χαρακτηριστικών τόσο του δικτύου όσο και της βροχής στη δρομολόγηση δεδομένων και στη σύγκλιση του μαρκοβιανού μοντέλου.

5.2 Προσομοίωση της απόσβεσης λόγω βροχής σε δίκτυα επικοινωνιών

Στο σημείο αυτό αφού περιγράφεται θεωρητικά το μέσο διάδοσης που υφίσταται διαλείψεις λόγω βροχοπτώσεων, προτείνεται ένα δυναμικό στοχαστικό μοντέλο για την προσομοίωση της απόσβεσης λόγω βροχής σε μεγάλες γεωγραφικές εκτάσεις. Ειδικότερα, τα χρονομεταβλητά πεδία της απόσβεσης λόγω βροχής υπερτίθενται σε ένα καρτεσιανό δισδιάστατο πλέγμα όπου αναπαριστάται, επίσης, το προς μελέτη δίκτυο. Η προσομοίωση βασίζεται στο μετασχηματισμό Fourier και περιγράφονται τόσο οι χωρικές όσο και οι χρονικές αποκλίσεις της βροχής. Το προτεινόμενο μοντέλο χρησιμοποιεί και επεκτείνει το ήδη υπάρχον μοντέλο της ITU-R και έχει σχεδιαστεί έτσι ώστε να απαιτεί όσο το δυνατόν λιγότερες παραμέτρους εισόδου.

5.2.1 Διαλείψεις που οφείλονται στην εξασθένηση των Η/Μ κυμάτων λόγω βροχής

Η συνεχώς αυξανόμενη ανάγκη για χρησιμοποίηση όλο και υψηλότερων συχνοτήτων στις εφαρμογές των τηλεπικοινωνιών έχει καταστήσει αναπόφευκτη τη μελέτη επίδρασης των ατμοσφαιρικών κατακρημνίσεων και ιδιαίτερα της βροχής στη διάδοση των ραδιοκυμάτων. Τα ρεύματα μετατόπισης που επάγονται στους υδρομετεωρίτες κατά τη

διάδοση ενός ραδιοκύματος προκαλούν δύο σημαντικά φαινόμενα, τη σκέδαση και την απορρόφηση σημαντικού τμήματος της ενέργειας του κύματος. Τα φαινόμενα αυτά επηρεάζουν σημαντικά τις ζεύξεις που λειτουργούν σε συχνότητες άνω των 10GHz. Εκτός των ανωτέρω φαινομένων, σε περιπτώσεις όπου το τηλεπικοινωνιακό δίκτυο αναχρησιμοποιεί τη συχνότητα, η βροχή προκαλεί την αποπόλωση των σημάτων. Συγκεκριμένα, προκαλεί ελάττωση της αναγκαίας απομόνωσης μεταξύ των κάθετα πολωμένων καναλιών με αποτέλεσμα την αμοιβαία παρεμβολή τους και την επακόλουθη μείωση της αξιοπιστίας στη μετάδοση πληροφορίας.

Για τους λόγους αυτούς, οι οποίοι καθιστούν τη βροχόπτωση φαινόμενο που επηρεάζει δυσμενώς τη διάδοση ραδιοκυμάτων σε συχνότητες άνω των 10GHz, στόχος της ενότητας αυτής είναι η προσομοίωση του αντίστοιχου διαύλου διάδοσης κατά τέτοιο τρόπο, ώστε να βρίσκει εφαρμογή σε όσο το δυνατόν περισσότερες κατηγορίες τηλεπικοινωνιακών δικτύων και σε μεγάλο εύρος σεναρίων διάδοσης. Συγκεκριμένα περιγράφεται τόσο η χρονομεταβλητή φύση της απόσβεσης λόγω βροχής όσο και η χωρική συσχέτιση σε μία ζώνη συχνοτήτων.

5.2.1.1 Μικροσκοπική περιγραφή του μέσου βροχής

Η βροχή αποτελεί τυχαία χωροχρονική διαδικασία με έντονη ανομοιογένεια και μελετάται μόνο με στατιστικές μεθόδους καθώς το μέσο της βροχής αποτελείται από μεγάλο πλήθος σταγόνων, των οποίων οι θέσεις, το σχήμα και η κατεύθυνση είναι τυχαίες μεταβλητές. Το μοντέλο που ισχύει για το σχήμα των σταγόνων της βροχής κατά τη διάρκεια μιας βροχόπτωσης δεν είναι σφαιρικό, αλλά πεπλατυσμένο σφαιροειδές που εμφανίζει γωνία κλίσης ως προς τον ορίζοντα. Συγκεκριμένα, για το σχήμα των σταγόνων ακτίνας 170-4000 μm ισχύουν τα εξής:

1. Σταγόνες με ακτίνα $\leq 170\mu\text{m}$ παραμορφώνονται ελάχιστα, οπότε μπορούν να θεωρηθούν σφαιρικές.
2. Σταγόνες με ακτίνες 170-500 μm θεωρείται ότι έχουν σχήμα σφαιροειδούς.
3. Σταγόνες με ακτίνες 500-2000 μm παρουσιάζουν ασύμμετρο σφαιροειδές σχήμα.
4. Σταγόνες με ακτίνα $>2000\mu\text{m}$ αναπτύσσουν μια πλάτυνση προς τα κάτω η οποία αυξάνεται με την αύξηση της ακτίνας.

Η διεύθυνση μιας σταγόνας επηρεάζεται από την κατεύθυνση και την ένταση του ανέμου. Η επίδραση αυτή έχει ως συνέπεια η σταγόνα να εμφανίζει γωνία κλίσης ως προς τον ορίζοντα (canting angle). Η παραμέτρος αυτή έχει μεγάλη σημασία κατά τη μελέτη της αποπόλωσης λόγω βροχής που παρατηρείται σε συστήματα αναχρησιμοποίησης συχνότητας.

Στις συχνότητες ενδιαφέροντος (>10GHz), τα μήκη κύματος είναι της τάξης του cm, οπότε, το μέγεθος των σταγόνων είναι παρόμοιου μεγέθους. Συνεπώς, κάθε σταγόνα δρα ως σκεδαστής για το H/M κύμα. Σε σύγκριση, όμως, με το μήκος κύματος οι σταγόνες της βροχής θεωρείται ότι βρίσκονται αρκετά μακριά μεταξύ τους, οπότε ευσταθεί η υπόθεση της ανεξάρτητης σκέδασης των κυμάτων όταν διέρχονται μέσω βροχής.

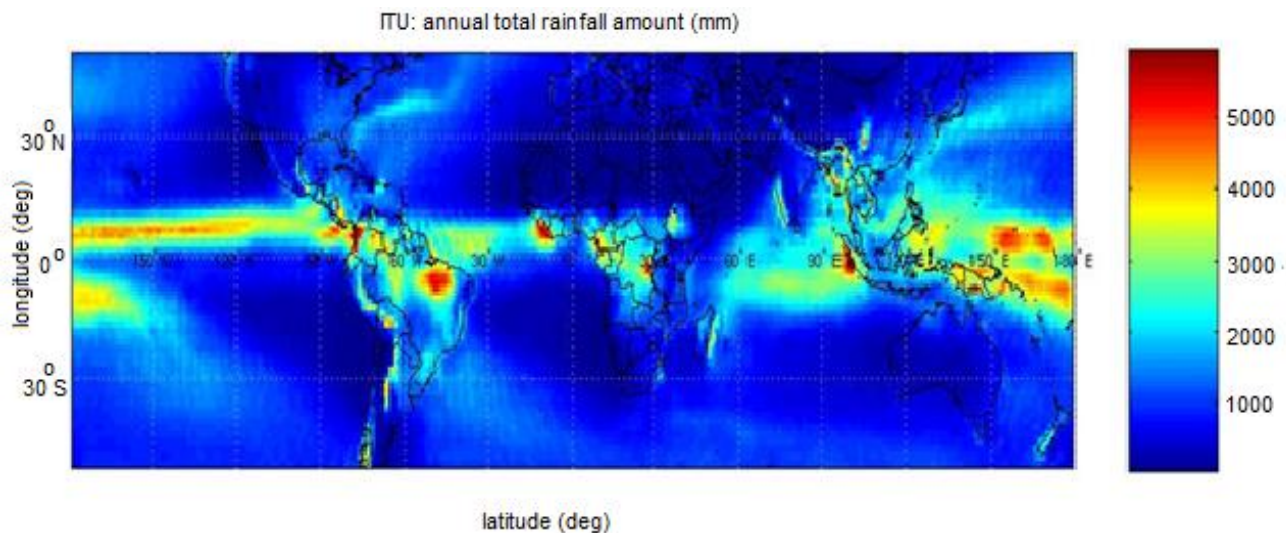
5.2.1.2 Μακροσκοπική περιγραφή του μέσου βροχής

Για τη μαθηματική περιγραφή της απόσβεσης λόγω βροχής ως παράγοντας που επηρεάζει τις διάφορες τηλεπικοινωνιακές εφαρμογές απαιτείται η μελέτη της χωρικής κατανομής της βροχής. Ανάλογα με την εποχή του έτους και τη γεωγραφική περιοχή η βροχή κατηγοριοποιείται σε δύο είδη

- 1.Βροχή στρωσιγενούς τύπου,δηλαδή βροχόπτωση χαμηλής έντασης, η οποία παρουσιάζει χωρική ομοιογένεια και συνεπώς χαμηλή τιμή χωρικής αυτοσυσχέτισης
- 2.Βροχή συνεκτικών βροχοπυρήνων, δηλαδή έντονη βροχόπτωση σε μεμονωμένες περιοχές(βροχοπυρήνες) με αυξημένη χωρική ανομοιογένεια και υψηλή τιμή χωρικής αυτοσυσχέτισης. Εκτός των βροχοπυρήνων επικρατεί στρωσιγενής βροχόπτωση.

Για τη στατιστική μελέτη της απόσβεσης λόγω βροχής απαιτείται η γνώση των στατιστικών ιδιοτήτων της σημειακής έντασης βροχόπτωσης που εμφανίζει χωρική και χρονική εξάρτηση $R=R(x,y,t)$. Η σημειακή ένταση ρυθμός βροχόπτωσης R που μετρείται ως το ύψος της στάθμης του ύδατος που φθάνει στη Γη ανά μονάδα χρόνου και κατά τη διάρκεια βροχόπτωσης αποτελεί τυχαία χωροχρονική διαδικασία. Ο καθορισμός της στατιστικής κατανομής του R είναι απαραίτητος σε κάθε περιοχή όπου πρόκειται να λειτουργήσει τηλεπικοινωνιακό σύστημα με φέρουσα συχνότητα μεγαλύτερη των 10GHz. Προς τούτο και καθώς δεν είναι διαθέσιμα πειραματικά δεδομένα για κάθε περιοχή της Γης, η ITU-R προτείνει τη χρήση δεκαπέντε κλιματικών ζωνών (A,B,C,D,E,F,G,H,J,K,L,M,N,P,Q) που αναφέρονται στα χαρακτηριστικά βροχόπτωσης

των διαφόρων περιοχών της γήινης επιφάνειας. Για έκαστη των ζωνών αυτών είναι γνωστή η αθροιστική κατανομή υπέρβασης του R για διάφορα ποσοστά χρόνου. Η κατανομή διαφέρει αρκετά κατά τόπους και εντός των ζωνών αυτών. Για το λόγο αυτό, η ITU-R με πρόσφατη σύσταση προτείνει τους χάρτες βροχόπτωσης (rainmaps) [ITU-R .837-3], βάσει των οποίων η Γη έχει χωριστεί σε σφαιρικούς τομείς $1.5^{\circ} \times 1.5^{\circ}$ για έκαστο των οποίων υπάρχουν διαθέσιμες παράμετροι περιγραφής της βροχόπτωσης οι οποίες προέκυψαν έπειτα από επεξεργασία από το Ευρωπαϊκό Κέντρο Μέσης Κλίμακας Πρόγνωσης Καιρικών Συνθηκών (ECMWF).



Σχήμα 5.1 Συνολικό ύψος βροχόπτωσης (mm) στην επιφάνεια της Γης σε ετήσια βάση

5.2.2 Καθορισμός του μοντέλου απόσβεσης λόγω βροχής

5.2.2.1 Παραδοχές του μοντέλου απόσβεσης

Για την ορθή κατάστρωση του μοντέλου απαιτούνται οι ακόλουθες παραδοχές:

1. Η απόσβεση λόγω βροχής παρουσιάζει τόσο χωρική όσο και χρονική εξάρτηση και συμβολίζεται ως $A(x,y,t)$. Η παράσταση της απόσβεσης λόγω βροχής πραγματοποιείται στις θέσεις $\{x,y\}$ ενός καρτεσιανού πλέγματος. Σε κάθε σημείο του πλέγματος το μέγεθος της απόσβεσης αναλύεται σε καρτεσιανές συντεταγμένες και δίδεται από το ζεύγος $\{\Delta x, \Delta y\}$.

2. Η κατανομή της απόσβεσης σε κάθε σημείο του πλέγματος θεωρείται λογαριθμοκανονική. Επιπλέον, σε κάθε σημείο θεωρείται ότι βρέχει μόνο για ένα κλάσμα του συνολικού χρόνου σε ετήσια βάση.

3. Η χρονική ενεργειακή φασματική πυκνότητα της απόσβεσης λόγω βροχής παρουσιάζει βαθυπερατή χαρακτηριστική συνάρτηση πρώτης τάξης (1 πόλος).

4. Η χωρική συσχέτιση της απόσβεσης θεωρείται ομογενής και ισοτροπική.

5.2.2.2 Εύρεση λογαριθμοκανονικής συνάρτησης αθροιστικής κατανομής (Lognormal Complementary Cumulative Distribution Function- CCDF) για την περιγραφή της απόσβεσης λόγω βροχής σε ένα σημείο του πλέγματος

Στις περισσότερες κλιματικές ζώνες που έχουν προταθεί από την ITU-R έχει επαληθευτεί με μακροχρόνιες μετρήσεις ότι η λογαριθμοκανονική κατανομή προσεγγίζει με τον καλύτερο τρόπο τη συνάρτηση κατανομής της έντασης βροχόπτωσης. Η αθροιστική κατανομή υπέρβασης της σημειακής έντασης βροχόπτωσης R είναι

$$P(R \geq r) = \frac{1}{2} \operatorname{erfc} \left(\frac{\ln r - \ln R_m}{\sqrt{2} S_r} \right) \quad (5.1)$$

Οι παράμετροι R_m , S_r αποτελούν τις στατιστικές παραμέτρους της λογαριθμοκανονικής κατανομής (μέση τιμή και τυπική απόκλιση της μεταβλητής $\ln R$) και προσδιορίζονται με χρήση διαθέσιμων πειραματικών δεδομένων για την αθροιστική συνάρτηση υπέρβασης της σημειακής έντασης βροχόπτωσης. Η συνάρτηση erfc είναι η γνωστή συμπληρωματική συνάρτηση σφάλματος

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt \quad (5.2)$$

Με αυτή τη λογική λειτουργεί και ο προσομοιωτής που χρησιμοποιείται στην εργασία, καθώς έχει ως στόχο την εκτίμηση της λογαριθμοκανονικής κατανομής που ακολουθεί η απόσβεση λόγω βροχής σε κάθε σημείο του πλέγματος. Προς τούτο, το μοντέλο προσδιορίζει τις στατιστικές παραμέτρους R_m και S_r με παρεμβολή της λογαριθμοκανονικής κατανομής στα δεδομένα της Σύστασης P.837-3 της ITU-R (ITU-R recommendation P.837-3).

5.2.2.3 Χωροχρονική συσχέτιση της απόσβεσης λόγω βροχής

Η από κοινού κατανομή της απόσβεσης σε δύο διαφορετικές διαδρομές εντός του πλέγματος είναι λογαριθμοκανονική με συνάρτηση πυκνότητας πιθανότητας

$$f(A_1, A_2, r) \equiv \Lambda_{A_1 A_2}(M_1, \Sigma_1, r_A, M_2, \Sigma_2) \quad (5.3)$$

όπου τα ζεύγη (M_1, Σ_1) και (M_2, Σ_2) είναι οι στατιστικές παράμετροι R_m και S_r των λογαριθμοκανονικών κατανομών των εντάσεων σημειακής βροχόπτωσης A_1 και A_2 (mm/hr) σε δύο διαφορετικά σημεία ενδιαφέροντος. Η παράμετρος r_A

$$r = \frac{e^{r' \Sigma_1 \Sigma_2 - 1}}{\sqrt{e^{\Sigma_1^2} - 1} \sqrt{e^{\Sigma_2^2} - 1}} \equiv g(r'), \quad r' \in [0, 1] \quad (5.4)$$

Είναι ο παράγοντας συσχέτισης των εντάσεων A_1 και A_2 όπου r' η αλληλοσυσχέτιση των βοηθητικών μεταβλητών

$$u_1 = \frac{(\ln A_1 - M_1)}{\Sigma_1} \quad \text{και} \quad u_2 = \frac{(\ln A_2 - M_2)}{\Sigma_2} \quad (5.5)$$

Αποδεικνύεται ότι αν οι τυχαίες μεταβλητές A_1 και A_2 αντικατασταθούν σύμφωνα με τον παραπάνω μετασχηματισμό από τις τυχαίες μεταβλητές u_1 και u_2 , η από κοινού συνάρτηση πυκνότητας πιθανότητας είναι η κανονική κατανομή με μηδενική μέση τιμή και μοναδιαία τυπική απόκλιση, δηλαδή

$$f(u_1, u_2) = \frac{1}{2\pi\sqrt{1-r'^2}} \exp\left\{-\frac{1}{2(1-r'^2)} (u_1^2 + u_2^2 - 2r'u_1u_2)\right\} \quad (5.6)$$

Τότε

$$r' = \exp\left(-\frac{\rho}{L} - \beta_2 |t|\right) \quad (5.7)$$

όπου

$$\rho = \sqrt{(|x_1 - x_2|)^2 + (|y_1 - y_2|)^2} \quad (5.8)$$

η απόσταση μεταξύ των δύο σημείων του πλέγματος και L χαρακτηριστική παράμετρος που σχετίζεται με το μέγεθος των βροχοπυρήνων.

Υποθέτοντας ότι οι βροχοπυρήνες μετακινούνται ομοιόμορφα με μέση ταχύτητα \bar{V} , εφαρμόζεται το θεώρημα Taylor και η αλληλοσυσχέτιση r' μπορεί να εκφραστεί μέσω της σχέσης

$$r' = \exp\left(-\frac{|\bar{V}|}{L}|t| - \beta_2|t|\right) = \exp(-\beta|t|) \approx \exp\left(-\frac{\beta\rho}{|\bar{V}|}\right) \quad (5.9)$$

όπου β είναι μια σύνθετη σταθερά που εξαρτάται από το ρυθμό της γέννησης και καταστροφής των βροχοπυρήνων, την ταχύτητα και την κατεύθυνση προς την οποία κινούνται και τη χαρακτηριστική διάστασή τους.

5.2.3 Σύνθεση πεδίων απόσβεσης λόγω βροχής

5.2.3.1 Πεδίο απόσβεσης λόγω βροχής παραγόμενο από διδιάστατο γκαουσιανό πεδίο

Η αρχική προσέγγιση συνίσταται στη δημιουργία ενός γκαουσιανού πεδίου $g(x_1, x_2)$, $0 \leq x_1, x_2 \leq N-1$ επί ενός καρτεσιανού πλέγματος $N \times N$ για την περιγραφή της συνάρτησης χωρικής αυτοσυσχέτισης. Αν το πεδίο $g(x_1, x_2)$ είναι πραγματικό και παρουσιάζει μηδενική μέση τιμή και μοναδιαία τυπική απόκλιση, εύκολα μετασχηματίζεται σε πεδίο απόσβεσης λόγω βροχής μέσω του μη γραμμικού μετασχηματισμού

$$A(x_1, x_2) = \exp(\sigma(x_1, x_2) \boxtimes g(x_1, x_2) + m(x_1, x_2)) \quad (5.10)$$

όπου το σύμβολο \boxtimes υποδηλώνει πολλαπλασιασμό πινάκων στοιχείο προς στοιχείο και τα μεγέθη $\{\sigma(x_1, x_2), m(x_1, x_2)\}$ είναι οι στατιστικές παράμετροι της λογαριθμοκανονικής κατανομής οι οποίες υπολογίζονται όπως αναφέρθηκε στην ενότητα 2.2. Η ανωτέρω εξίσωση επιτρέπει τον υπολογισμό της απόσβεσης σε κάθε σημείο από τα N^2 του πλέγματος.

5.2.3.2 Δημιουργία χωρικά συσχετισμένων γκαουσιανών πεδίων

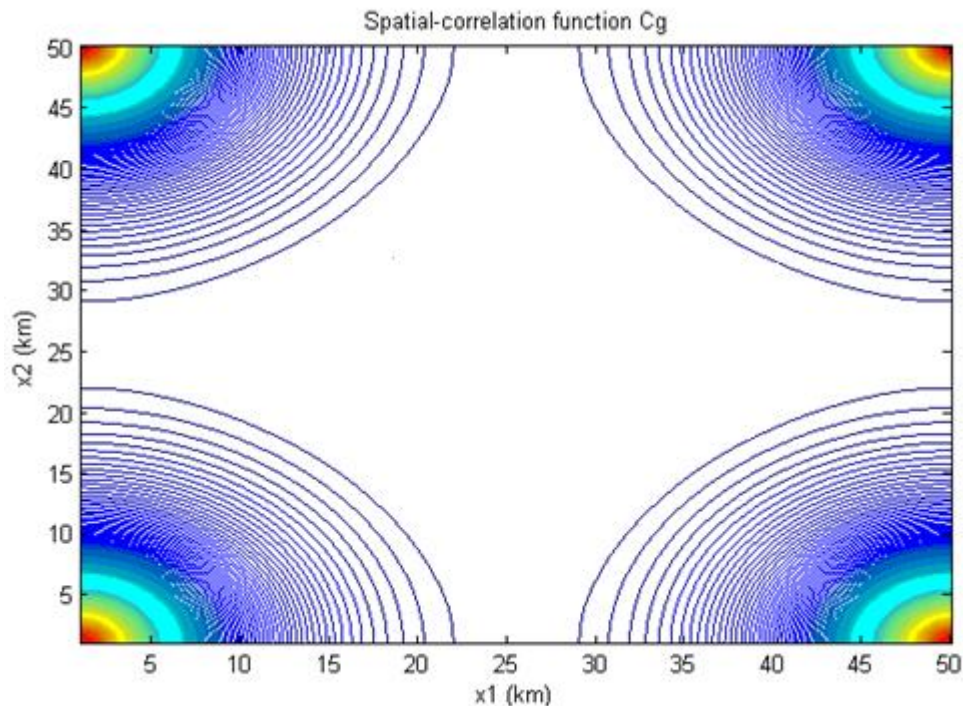
Το κρίσιμο σημείο για την έναρξη της προσομοίωσης είναι η σύνθεση του γκαουσιανού πεδίου $g(x_1, x_2)$. Προς τούτο θεωρείται αρχικά η συνάρτηση χωρικής συσχέτισης $C_g(x_1, x_2)$ των βοηθητικών μεταβλητών

$$u_1(x_1, x_2) = \frac{(\ln A(x_1, x_2) - m(x_1, x_2))}{\sigma(x_1, x_2)} \quad \text{και} \quad u_2(y_1, y_2) = \frac{(\ln A(y_1, y_2) - m(y_1, y_2))}{\sigma(y_1, y_2)} \quad (5.11)$$

Σε συμφωνία με την ανάλυση της προηγούμενης ενότητας, η συνάρτηση $C_g(x_1, x_2)$ έχει τη μορφή

$$C_g(x_1, x_2) = \exp \left[-\frac{\Delta}{L} \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \right] \quad (5.12)$$

όπου Δ (km) είναι η απόσταση μεταξύ 2 σημείων του πλέγματος και L (km) είναι η χαρακτηριστική απόσταση των βροχοπυρήνων. Επιπλέον, η συνάρτηση $C_g(x_1, x_2)$ απαιτείται να εμφανίζει συμμετρία και ως προς τα 4 τεταρτημόρια, όπως στο σχήμα που ακολουθεί



Σχήμα 5.2 Συνάρτηση χωρικής συσχέτισης με συμμετρία και στα 4 τεταρτημόρια

Εφόσον έχει υπολογιστεί η συνάρτηση χωρικής συσχέτισης, το γκαουσιανό πεδίο $g(x_1, x_2)$ παράγεται εφαρμόζοντας τα ακόλουθα βήματα:

1. Δημιουργείται δισδιάστατος μιγαδικός ασυσχέτιστος θόρυβος, $e(k_1, k_2)$, με πραγματική και μιγαδική μηδενική μέση τιμή και μοναδιαίες τυπικές αποκλίσεις, τέτοιος ώστε $e(k_1, k_2) = e^{*(n_1 N - k_1, n_2 N - k_2)}$, $n_1, n_2 \in \mathbb{N}$, όπου το * συμβολίζει το συζυγή μιγαδικό. Η μορφή αυτή του θορύβου επιλέγεται με στόχο ο αντίστροφος μετασχηματισμός Fourier του θορύβου να είναι καθαρά πραγματικό μέγεθος, $\mathcal{F}_2^{-1}\{e(k_1, k_2)\} \in R$.

2. Δημιουργείται πίνακας H με στόχο να αποδοθούν κατάλληλες τιμές ενέργειας στο θόρυβο $e(k_1, k_2)$. Τα στοιχεία του πίνακα είναι:

$$h(k_1, k_2) = \begin{cases} 1, & (k_1, k_2) = (0, 0), \left(0, \frac{N}{2}\right), \left(\frac{N}{2}, 0\right), (N/2, N/2) \\ \frac{1}{2}, & \text{σε οποιαδήποτε άλλη περίπτωση} \end{cases}$$

3. Υπολογίζεται ο δισδιάστατος μετασχηματισμός Fourier της συνάρτησης $C_g(x_1, x_2)$

$$C_g^f(k_1, k_2) = \mathcal{F}_2\{C_g(x_1, x_2)\} \quad (5.13)$$

4. Δημιουργείται μιγαδικός πίνακας A με στοιχεία

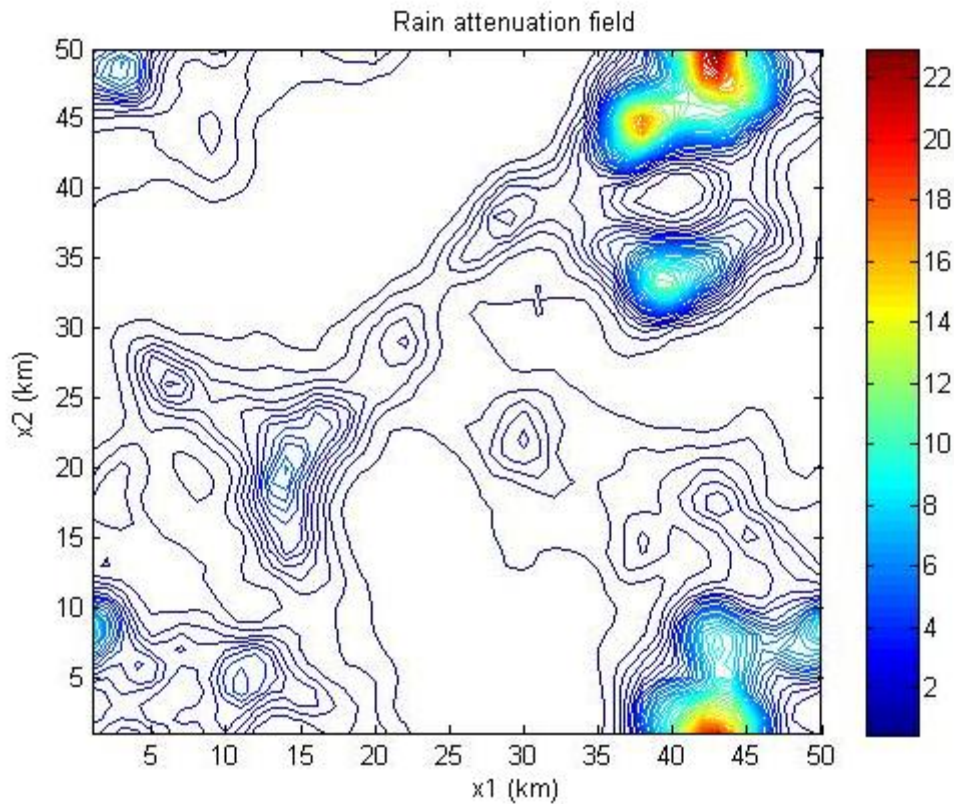
$$a(k_1, k_2) = h(k_1, k_2) \boxtimes C_g^f(k_1, k_2) \boxtimes e(k_1, k_2) \quad (5.14)$$

όπου το σύμβολο \boxtimes δηλώνει πολλαπλασιασμό πινάκων στοιχείο προς στοιχείο.

5. Το χωρικά συσχετισμένο γκαουσιανό πεδίο προκύπτει με αντίστροφο μετασχηματισμό Fourier δηλαδή

$$g(x_1, x_2) = \mathcal{F}_2^{-1}\{a(k_1, k_2)\} \quad (5.15)$$

Στο Σχ.5.3 παρουσιάζεται το πεδίο απόσβεσης λόγω βροχής που δημιουργείται με συνάρτηση χωρικής συσχέτισης αυτή του Σχ.5.2.



Σχήμα 5.3 Πεδίο απόσβεσης λόγω βροχής σε έκταση 50km

5.2.3.3 Χωρική μεταφορά των βροχοπυρήνων

Ως τώρα έχει υποθεθεί ότι το παραγόμενο πεδίο που περιγράφει την απόσβεση λόγω βροχής είναι στατικό. Απαιτείται, λοιπόν, η επέκταση του μοντέλου έτσι ώστε να ληφθεί υπόψη η περίπτωση της μεταφοράς του πεδίου βροχής προς την κατεύθυνση που ορίζει η μέση ταχύτητα μεταφοράς $\bar{\mathbf{V}} = [\bar{V}_x, \bar{V}_y]$ [m/s]. Η επέκταση αυτή του μοντέλου μπορεί εύκολα να υλοποιηθεί στο πεδίο της συχνότητας μέσω της ιδιότητας της χρονικής ολίσθησης του μετασχηματισμού Fourier σύμφωνα με τη σχέση

$$g(x_1 - t\bar{V}_x, x_2 - t\bar{V}_y) = \mathcal{F}_2^{-1}\{a(k_1, k_2) \square \exp[-i \times t \times [k_1, k_2][\bar{V}_x, \bar{V}_y]^T]\} \quad (5.16)$$

5.2.3.4 Χρονική μετάθεση πεδίου βροχής

Τα πεδία της βροχής είναι τυχαία και χρονομεταβλητά. Όπως ήδη παρουσιάστηκε, η απόσβεση που προκαλεί η βροχή σε ένα σημείο του πλέγματος έχει φασματική συμπεριφορά

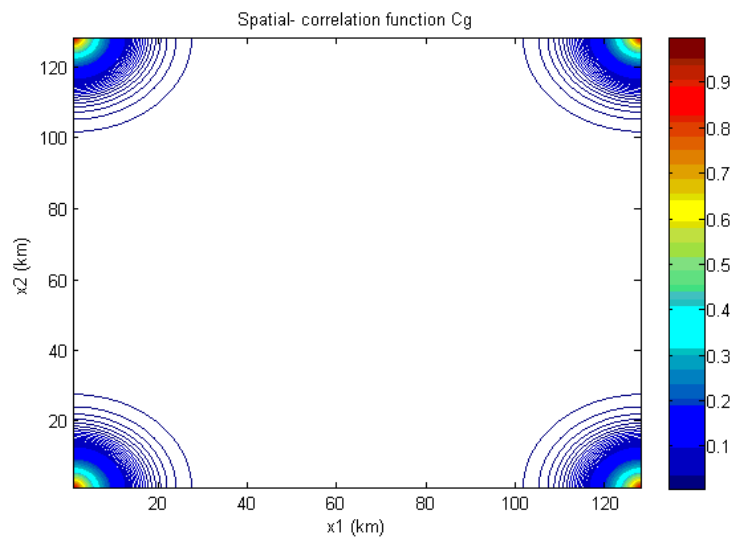
που ασυμπτωτικά τείνει σε βαθυπερατή χαρακτηριστική πρώτης τάξης. Αυτό προϋποθέτει ότι οι βοηθητικές μεταβλητές $u_1(x_1, x_2)$ και $u_2(y_1, y_2)$ πρέπει επίσης να παρουσιάζουν βαθυπερατή συμπεριφορά πρώτης τάξης. Η συνθήκη αυτή μπορεί να υλοποιηθεί θεωρώντας ότι η συνάρτηση $g(x_1, x_2)$ είναι μια αυτοαναδρομική χρονική διαδικασία πρώτης τάξης (auto-regressive process of order one ,AR(1)), δηλαδή μια αλυσίδα Markov. Συγκεκριμένα η διαδικασία AR(1) περιγράφεται μέσω της διαφορικής εξίσωσης

$$\alpha(k_1, k_2, t) = \exp(-\beta_t \Delta t) \alpha(k_1, k_2, t - \Delta t) + \sqrt{1 - \exp(-\beta_t \Delta t)^2} \times z(k_1, k_2, t) \quad (5.17)$$

όπου $z(k_1, k_2, t)$ είναι ο πίνακας του μιγαδικού ασυσχέτιστου θορύβου, Δt το χρονικό βήμα που χρησιμοποιείται κατά την προσομοίωση και β_t είναι χρονική σταθερά κανονικοποίησης.

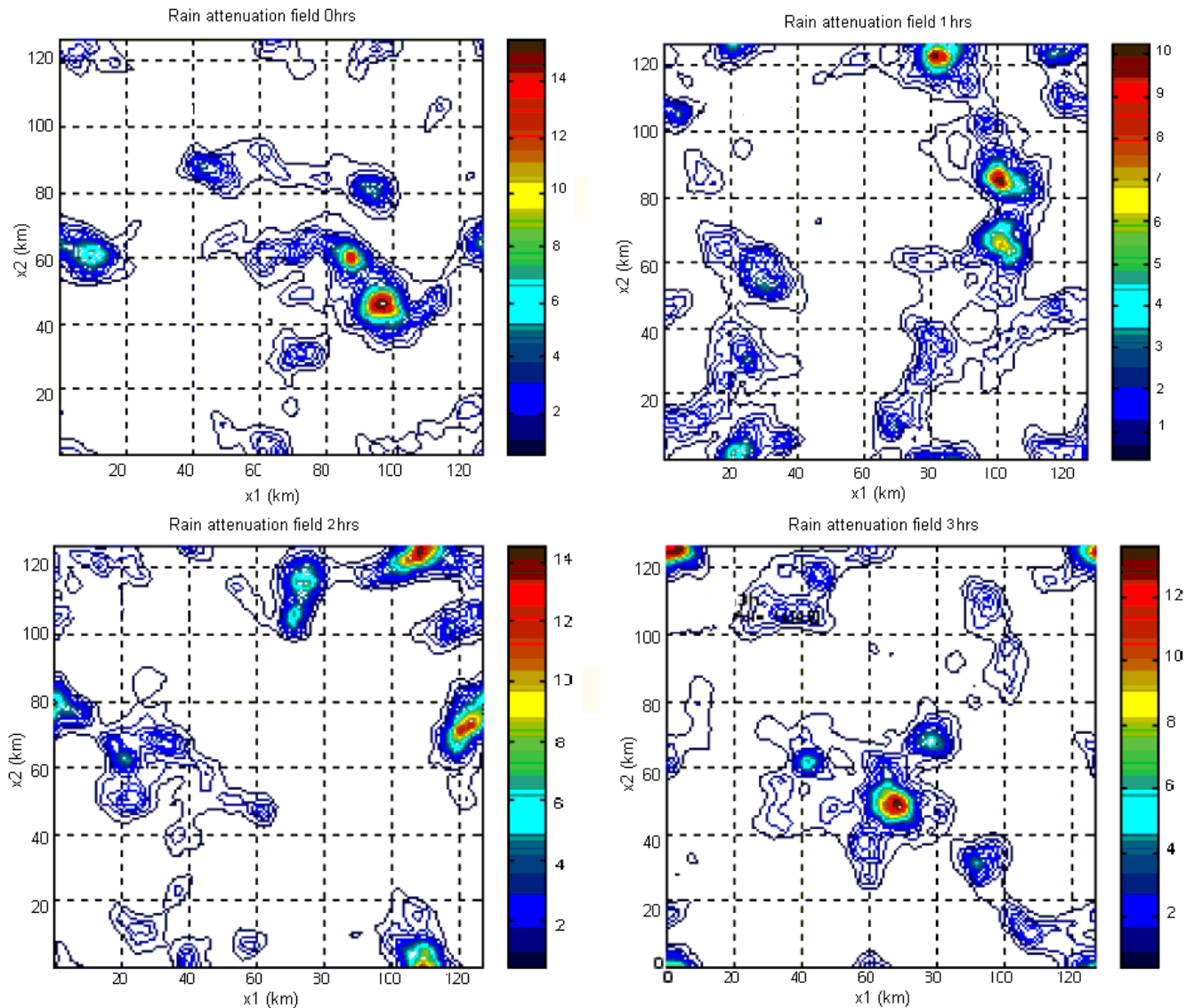
5.2.3.5 Αποτελέσματα προσομοίωσης διαλείψεων λόγω βροχής

Θεωρείται πλέγμα 128^2 σημείων, όπου κάθε σημείο θεωρείται ότι έχει τις ίδιες παραμέτρους λογαριθμοκανονικής κατανομής, $(m, \sigma) = (-1.72, 1.5246)$. Η χαρακτηριστική διάμετρος των βροχοπυρήνων τίθεται $L = 5 \text{ km}$, η ταχύτητα με την οποία κινούνται $\bar{V} = [\bar{V}_x, \bar{V}_y] = [20 \ 0] \text{ km/hr}$ και χρονική σταθερά ορίζεται $\beta_t = 0.5 \text{ hr}^{-1}$. Αρχικά, δημιουργείται η συνάρτηση χωρικής συσχέτισης για το πλέγμα αυτό με βάση την οποία θα παραχθούν τα πεδία απόσβεσης λόγω βροχής



Σχήμα 5. 4 Συνάρτηση χωρικής συσχέτισης για πλέγμα 128^2 σημείων

Η προσομοίωση εκτελείται για ένα συνολικό διάστημα 8h με χρονικό βήμα ίσο με 0.1h και ακολούθως παρουσιάζονται τα γραφήματα που περιγράφουν την απόσβεση που προκαλείται λόγω της βροχής στο πλέγμα ανά διαστήματα 1h.



Σχήμα 5.5 Προσομοίωση 8h πεδίου απόσβεσης λόγω βροχής – παρουσίαση αποτελεσμάτων ανά διαστήματα 1h

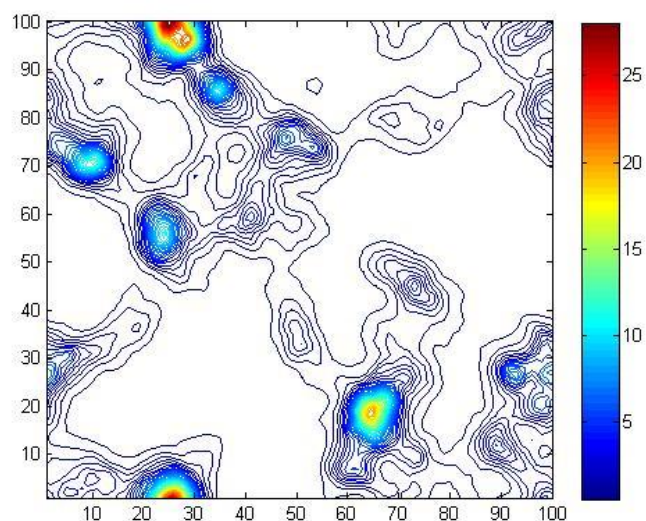
5.3 Μελέτη δρομολόγησης σε δίκτυα WMN υπό συνθήκες διαλείψεων λόγω βροχής

Στην ενότητα αυτή θα διερευνηθεί η επίδραση διαφόρων παραμέτρων της τοπολογίας του δικτύου και του μαρκοβιανού μοντέλου στη σύγκλιση του αλγορίθμου

δρομολόγησης. Αρχικά μελετάται το πώς επηρεάζει η μέση ισχύς εκπομπής των κόμβων του δικτύου τη συνολική καθυστέρηση (end-to-end delay) στη δρομολόγηση των δεδομένων. Στη συνέχεια, διερευνάται η κλιμακωσιμότητα του αλγορίθμου, μέσω της εφαρμογής του σε διαφορετικού μεγέθους τοπολογίες και εξέτασης της ταχύτητας σύγκλισης. Έπειτα αναλύεται, η επίδραση του παράγοντα μείωσης του κέρδους λ (discount factor) στην ταχύτητα σύγκλισης του αλγορίθμου. Τέλος, για συγκεκριμένη τοπολογία δικτύου, προτείνεται η επιλογή συγκεκριμένων χαρακτηριστικών των κόμβων με στόχο τη βελτιστοποίηση του αλγορίθμου δρομολόγησης.

5.3.1 Επίδραση της ισχύος εκπομπής στην από-άκρο-σε-άκρο καθυστέρηση (end- to-end delay)

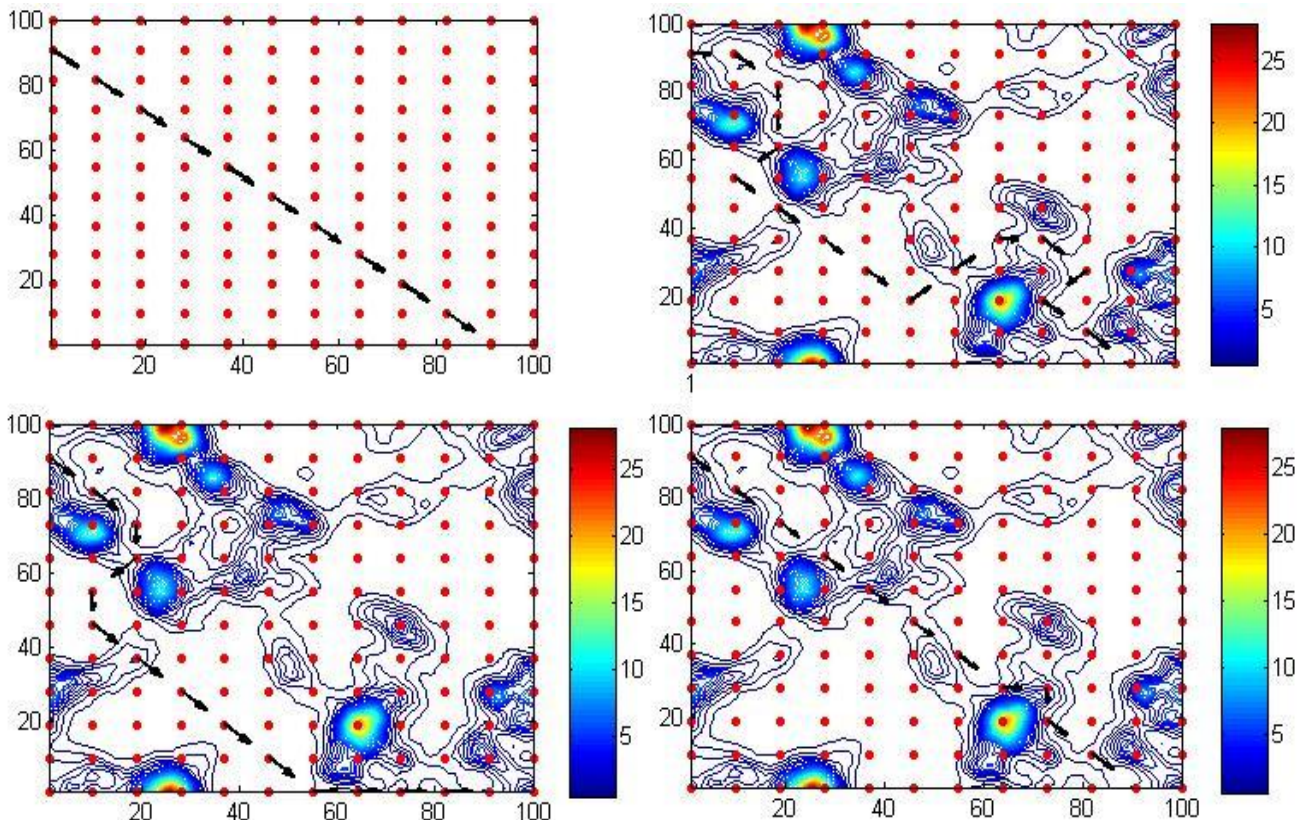
Θεωρείται ένα συγκεκριμένο σενάριο βροχής σε ένα πλέγμα 100x100, όπου κάθε σημείο έχει παραμέτρους λογαριθμοκανονικής κατανομής, $(m,\sigma)=(-1.72,1.5246)$. (Σχ. 5.1) Το σήμα που μεταδίδεται μέσω των βροχοπυρήνων μπορεί να υποστεί απόσβεση μέχρι και 27dB. Στο ίδιο πλέγμα θεωρούνται διαδοχικά τρεις τοπολογίες δικτύων, που περιλαμβάνουν 144, 1156 και 10000 κόμβους, αντίστοιχα, ομοιόμορφα κατανεμημένους στο πλέγμα. Καθορίζεται ο κόμβος - πηγή και ο κόμβος- προορισμός και μελετάται η επίδραση της μέσης ισχύος εκπομπής των κόμβων στη βέλτιστη διαδρομή δρομολόγησης, όπως αυτή προκύπτει από την επίλυση του μακροβιανού μοντέλου.



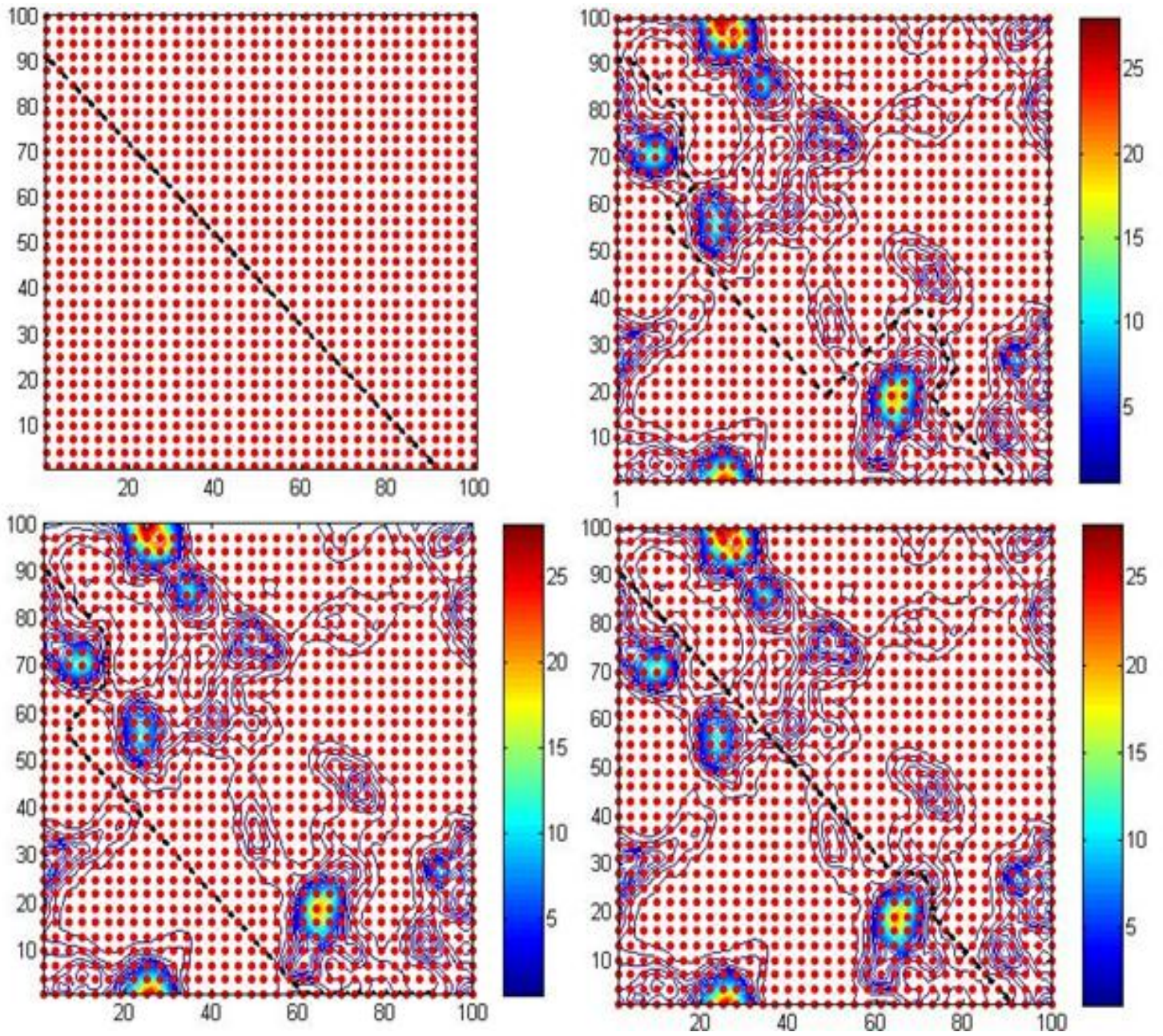
Σχήμα 5. 6 Πεδίο απόσβεσης λόγω βροχής σε πλέγμα 100² σημείων

Όταν δεν υπάρχει βροχόπτωση, η βέλτιστη διαδρομή είναι η ευθεία που συνδέει την πηγή με τον προορισμό και οι κόμβοι εκπέμπουν με ισχύ P_{cs} (ισχύς εκπομπής υπό συνθήκες καθαρού ουρανού). Αντίθετα, όταν υπάρχει βροχόπτωση, η βέλτιστη διαδρομή καθορίζεται σε μεγάλο βαθμό από τη διαθέσιμη ισχύ εκπομπής των κόμβων του πλέγματος. Όταν η μέση ισχύς εκπομπής των κόμβων είναι μικρή, η δρομολόγηση αποτυγχάνει ακόμη και σε ζεύξεις που δεν υφίστανται έντονες διαλείψεις λόγω βροχής. Για το λόγο αυτό αναζητούνται εναλλακτικές διαδρομές δρομολόγησης. Αν αυξηθεί η μέση ισχύς εκπομπής των κόμβων, ορισμένες από τις ανενεργές ζεύξεις καθίστανται διαθέσιμες για δρομολόγηση με αποτέλεσμα να μεταβάλλεται η βέλτιστη διαδρομή δρομολόγησης. Όσο περισσότερο αυξάνεται η ισχύς εκπομπής τόσο η βέλτιστη διαδρομή δρομολόγησης πλησιάζει τη βέλτιστη διαδρομή όταν δεν υπάρχει βροχόπτωση.

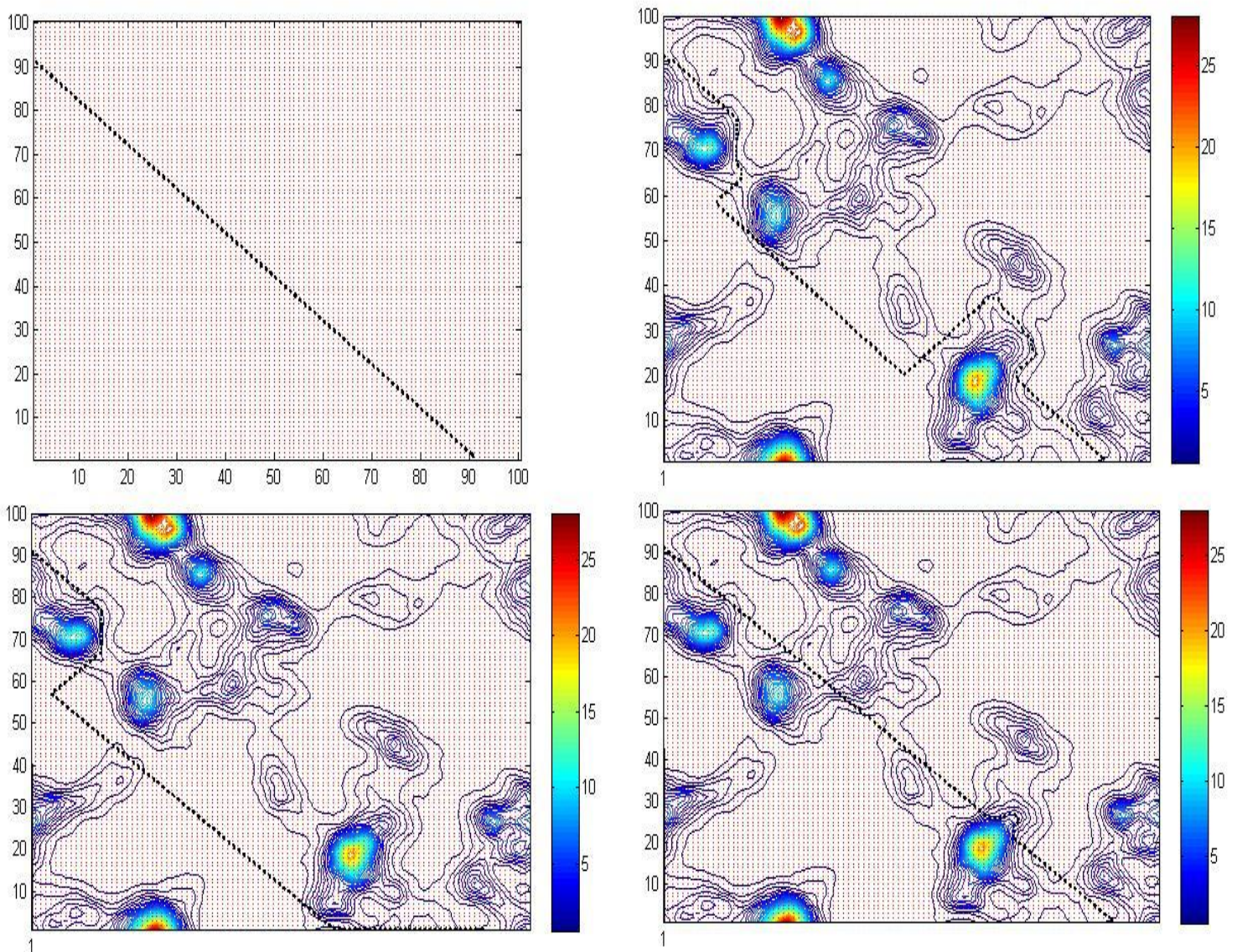
Τα προηγούμενα πιστοποιούνται από τα Σχ. 5.7 ως 5.9. Η πρώτη ομάδα σχημάτων αφορά την τοπολογία των 144 κόμβων, η δεύτερη την τοπολογία των 1156 κόμβων και η τελευταία το πυκνότερο δίκτυο των 10000 κόμβων. Σε κάθε περίπτωση, φαίνεται η βέλτιστη διαδρομή όταν δεν υπάρχουν διαλείψεις λόγω βροχής και πώς αυτή μεταβάλλεται ανάλογα με την αύξηση της ισχύος εκπομπής των κόμβων του δικτύου.



Σχήμα 5. 7 Τοπολογία 144 κόμβων, βέλτιστη διαδρομή για διαφορετικές στάθμες ισχύος εκπομπής



Σχήμα 5. 8 Τοπολογία 1156 κόμβων, βέλτιστη διαδρομή για διαφορετικές στάθμες ισχύος εκπομπής



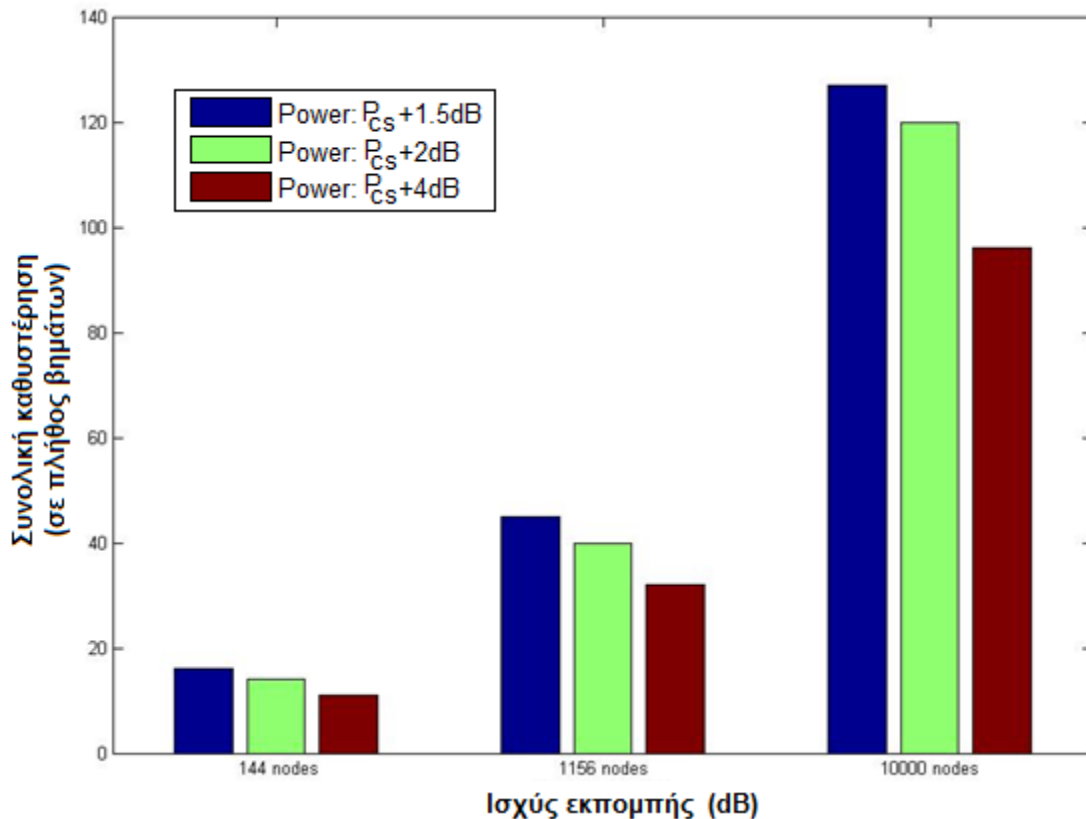
Σχήμα 5. 9 Τοπολογία 10000 κόμβων, βέλτιστη διαδρομή για διαφορετικές στάθμες ισχύος εκπομπής

Φαίνεται ότι όσο μικρότερη είναι η ισχύς εκπομπής των κόμβων του δικτύου τόσο περισσότερα είναι τα ενδιάμεσα βήματα (hops) από τα οποία διέρχονται τα πακέτα. Η αύξηση αυτή των βημάτων έχει ως αποτέλεσμα την αύξηση της από-άκρο-σε-άκρο καθυστέρησης (end-to-end delay) για τη μετάδοση των δεδομένων. Στο Σχ.5.10

απεικονίζεται η επίδραση της ισχύος εκπομπής στο πλήθος των ενδιάμεσων κόμβων, από τους οποίους διέρχονται τα πακέτα προκειμένου να δρομολογηθούν από τον κόμβο-πηγή στον κόμβο-προορισμό για διαφορετικά μεγέθη τοπολογιών. Παρατηρείται ότι η αύξηση του μεγέθους του δικτύου αυξάνει το πλήθος των ενδιάμεσων βημάτων, ενώ η αύξηση της ισχύος εκπομπής προκαλεί τη μείωσή τους.

Ισχύς εκπομπής (dB)	Από-άκρο-σε άκρο καθυστέρηση (σε πλήθος βημάτων)		
	144 κόμβοι	1156 κόμβοι	10000 κόμβοι
$P_{cs}+1.5$	16	45	127
$P_{cs}+2$	14	40	120
$P_{cs}+4$	11	32	96

Πίνακας 5. 1 Επίδραση της ισχύος εκπομπής στη συνολική καθυστέρηση μετάδοσης



Σχήμα 5. 10 Επίδραση της ισχύος εκπομπής στη συνολική καθυστέρηση μετάδοσης

5.3.2 Επίδραση του μεγέθους του δικτύου στην ταχύτητα σύγκλισης του αλγορίθμου δρομολόγησης

Στο σημείο αυτό εξετάζεται η κλιμακωσιμότητα (scalability) του αλγορίθμου. Όσο μεγαλύτερη πυκνότητα παρουσιάζει το δίκτυο απαιτείται περισσότερος χρόνος για να συγκλίνει ο αλγόριθμος. Για συγκεκριμένη στάθμη ισχύος εκπομπής των κόμβων υπολογίζεται ο χρόνος σύγκλισης του αλγορίθμου αναδρομικής πολιτικής για κάθε μία εκ των τριών τοπολογιών (144, 1156 και 10000 κόμβοι). (πίνακας 5.2)

Πλήθος κόμβων	144	1156	10000
Χρόνος σύγκλισης (sec)	0.328	2.11	37.098

Πίνακας 5. 2 Χρόνος σύγκλισης συναρτήσει μεγέθους δικτύου

Τα δεδομένα του ανωτέρω πίνακα υποδεικνύουν ότι η ταχύτητα σύγκλισης είναι πολυωνυμική συνάρτηση δευτέρου βαθμού του μεγέθους του δικτύου με εξίσωση:

$$t = p_1 \cdot x^2 + p_2 \cdot x + p_3$$

όπου t είναι ο χρόνος σύγκλισης, x το πλήθος των κόμβων του δικτύου και οι συντελεστές του πολυωνύμου είναι $p_1 = 2.2172 \cdot 10^{-7}$, $p_2 = 0.0014726$, $p_3 = 0.11134$. Συνεπώς η πολυπλοκότητα του αλγορίθμου είναι $O(n^2)$, όταν εξετάζεται η κλιμακωσιμότητα.

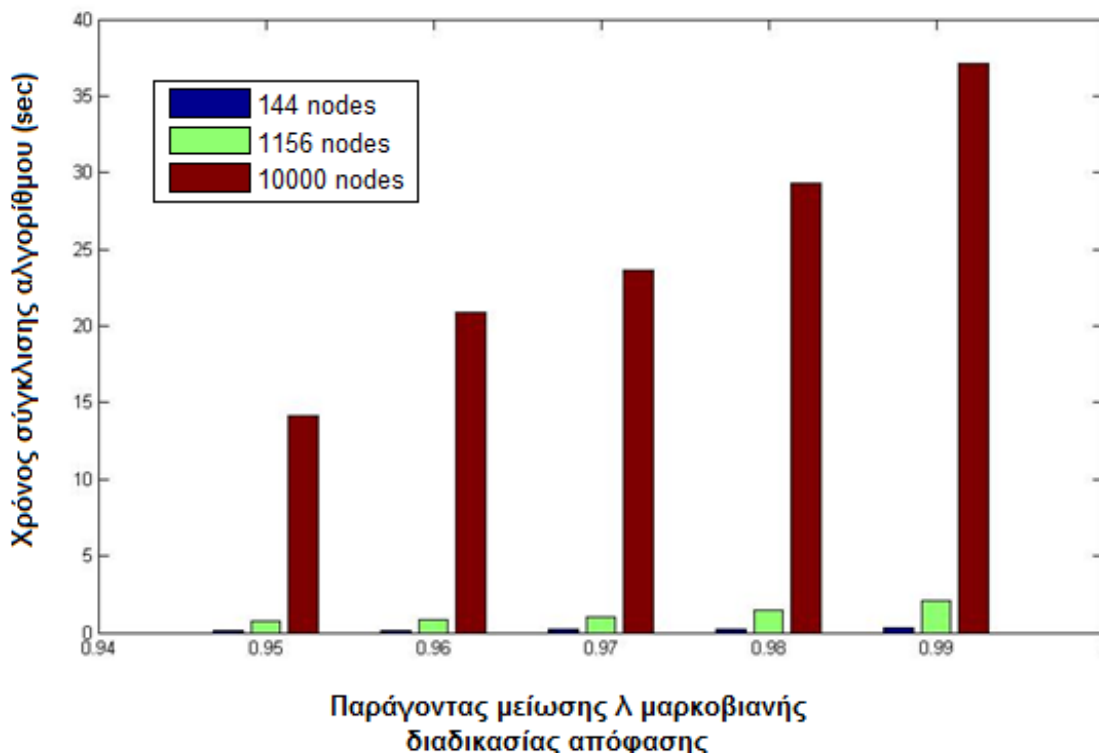
5.3.3 Επίδραση του παράγοντα μείωσης του κέρδους λ στην ταχύτητα σύγκλισης του αλγορίθμου

Ο παράγοντας μείωσης του οφέλους λ (discount factor) εμφανίζεται στις εξισώσεις Bellman της μαρκοβιανής διαδικασίας απόφασης. Η τιμή του κυμαίνεται στο εύρος $[0,1)$ ώστε να εξασφαλίζεται η σύγκλιση του αλγορίθμου. Η τιμή του παράγοντα λ υποδεικνύει το βαθμό στον οποίο λαμβάνονται υπόψη τα παρελθόντα οφέλη του συστήματος για την επιλογή της παρούσας δράσης. Συνεπώς, για μικρότερες τιμές του παράγοντα λ ο αλγόριθμος συγκλίνει ταχύτερα. Τα παραπάνω απεικονίζονται στο

Σχ.5.11, όπου φαίνεται η επίδραση της παραμέτρου λ με τιμές στην περιοχή 0.91÷0.99 για διαφορετικά μεγέθη τοπολογιών.

Παράγοντας μείωσης λ	Χρόνος σύγκλισης(sec)		
	144 κόμβοι	1156 κόμβοι	10000 κόμβοι
0.95	0.125	0.780	14.103
0.96	0.156	0.874	20.904
0.97	0.218	1.061	23.634
0.98	0.250	1.435	29.313
0.99	0.328	2.110	37.098

Πίνακας 5. 3 Μεταβολή ταχύτητας σύγκλισης του αλγορίθμου λόγω της μεταβολής του παράγοντα μείωσης λ



Σχήμα 5. 11 Μεταβολή ταχύτητας σύγκλισης του αλγορίθμου λόγω της μεταβολής του

5.3.4 Βελτιστοποίηση του αλγορίθμου δρομολόγησης

Από τα προηγούμενα προκύπτει ότι, κατά τη δρομολόγηση δεδομένων υπό συνθήκες διαλείψεων λόγω βροχής, υπάρχει σχέση ανταλλαγής (trade-off) μεταξύ της ενέργειας που καταναλώνουν οι κόμβοι (energy consumption) και της συνολικής καθυστέρησης μετάδοσης των δεδομένων. Συγκεκριμένα, εφόσον οι κόμβοι δεν υπόκεινται σε περιορισμούς στην κατανάλωση ενέργειας, μπορούν για τη μετάδοση των δεδομένων να επιλέξουν ζεύξεις που υφίστανται εντονότερες διαλείψεις, ώστε να γίνει δρομολόγηση από την πηγή στον προορισμό με λιγότερα βήματα. Αντιθέτως, όταν οι κόμβοι υπόκεινται σε περιορισμούς ενέργειας (energy constrained nodes), αναγκαστικά επιλέγουν για τη μετάδοση των δεδομένων ζεύξεις που δεν υποφέρουν από διαλείψεις με αποτέλεσμα την αύξηση των ενδιάμεσων βημάτων και, συνεπώς, της συνολικής καθυστέρησης μετάδοσης.

Για τη διερεύνηση της σχέσης μεταξύ καθυστέρησης και καταναλωμένης ενέργειας με στόχο τη βελτιστοποίηση του αλγορίθμου, προτείνεται, η ακόλουθη συνάρτηση κόστους (cost function)

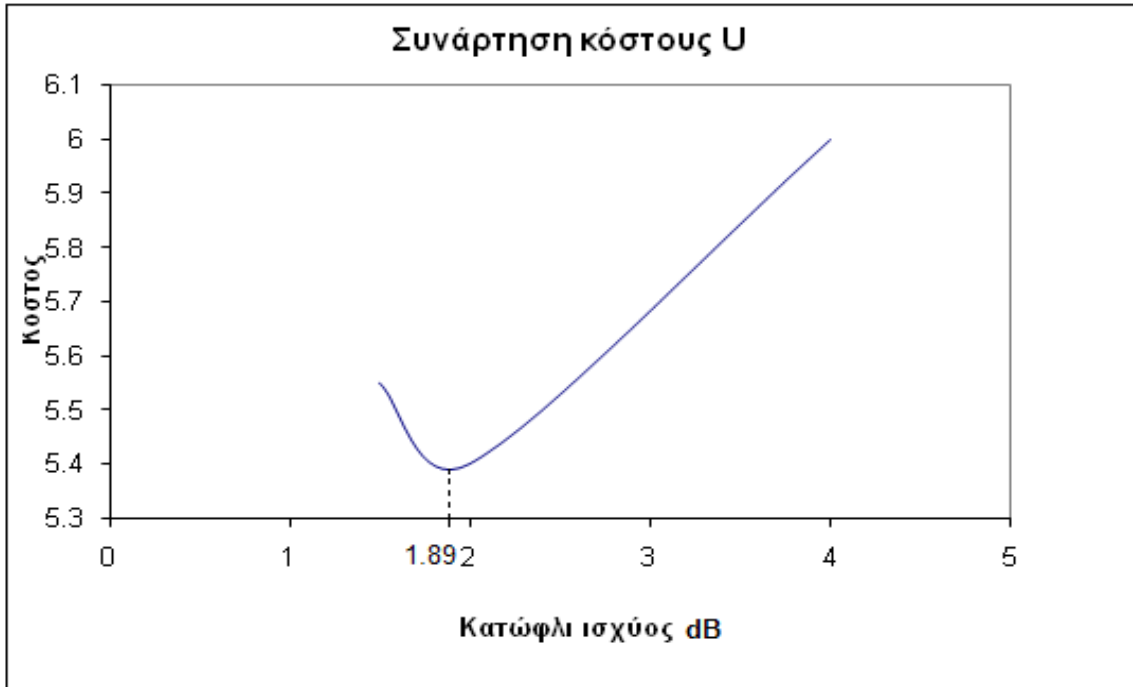
$$U = a \cdot D + b \cdot P$$

όπου D είναι η από άκρο-σε άκρο καθυστέρηση εκφρασμένη σε πλήθος βημάτων, P είναι η καταναλωμένη ενέργεια και a , b σταθεροί συντελεστές βαρύτητας.

Στόχος είναι ο προσδιορισμός του κατωφλίου ισχύος εκπομπής για το οποίο η συνάρτηση κόστους ελαχιστοποιείται. Η συνάρτηση κόστους U για ένα δίκτυο 1156 κόμβων με συντελεστές βαρύτητας $a=0.7$ και $b=0.1$ για διάφορες τιμές κατωφλίου ισχύος υπολογίζεται με βάση τον πίνακα 5.4 και απεικονίζεται στο Σχ.5.12

	Από-άκρο-σε-άκρο καθυστέρηση (σε πλήθος βημάτων)
Ισχύς εκπομπής (dB)	1156 κόμβοι
$P_{cs}+1.5$	45
$P_{cs}+2$	40
$P_{cs}+4$	32

Πίνακας 5. 4 Επίδραση της ισχύος εκπομπής στη συνολική καθυστέρηση μετάδοσης σε δίκτυο 1156 κόμβων



Σχήμα 5. 12 Συνάρτηση κόστους με παραμέτρους την κατανάλωση ενέργειας και τη συνολική καθυστέρηση σε δίκτυο 1156 κόμβων

Από το Σχ.5.12 προκύπτει ότι η τιμή ενέργειας που ελαχιστοποιεί τη συνάρτηση κόστους U είναι 1.89dB για τη δρομολόγηση των δεδομένων στο δίκτυο. Για τιμές ισχύος μικρότερες από αυτό το κατώφλι η συνάρτηση κόστους αυξάνει εξαιτίας της σημαντικής καθυστέρησης μετάδοσης. Αντιθέτως, για τιμές ισχύος μεγαλύτερες από 1.89dB η συνάρτηση κόστους λαμβάνει μεγάλες τιμές εξαιτίας της σημαντικής μείωσης της διάρκειας ζωής του δικτύου.

ΠΑΡΑΡΤΗΜΑ Α

Μερικώς παρατηρήσιμες μαρκοβιανές διαδικασίες απόφασης- Partially Observable Markov Decision Process (POMDP)

Οι μερικώς παρατηρήσιμες μαρκοβιανές διαδικασίες απόφασης είναι απαραίτητες για τη μελέτη προβλημάτων στα οποία το σύνολο των πιθανών καταστάσεων του υπό μελέτη συστήματος δεν είναι πλήρως γνωστό. Αυτό συμβαίνει στις περισσότερες περιπτώσεις προβλημάτων που αφορούν τηλεπικοινωνιακές εφαρμογές, είτε λόγω του θορύβου και των παρεμβολών, είτε λόγω της συχνής μεταβολής του περιβάλλοντος του συστήματος. Κατ' επέκταση, η γνώση των διαδικασιών POMDP είναι απαραίτητη.

Οι παράμετροι μέσω των οποίων περιγράφεται μια μερικώς παρατηρήσιμη μαρκοβιανή διαδικασία δεν διαφέρουν πολύ από αυτές μιας μαρκοβιανής διαδικασίας απόφασης (MDP). Για τον καθορισμό του μοντέλου απαιτούνται ένα σύνολο καταστάσεων $\{S\}$, ένα σύνολο δράσεων $\{A\}$, ένα σύνολο άμεσων οφελών $\{r_t(s,a)\}$ και ένα σύνολο από πιθανότητες μετάβασης του συστήματος από τη μία κατάσταση στην άλλη $\{p_t(\cdot|s,a)\}$. Εκτός, όμως, από τα σύνολα αυτά απαιτείται και ένα σύνολο παρατηρήσεων (observation) Ω της κατάστασης του συστήματος καθώς αυτή δεν είναι πλήρως γνωστή. Στην περίπτωση όπου το σύνολο Ω είναι στοχαστικό, είναι αναγκαία η γνώση της κατανομής πιθανοτήτων $O(\alpha,s,l)$ που ακολουθούν τα στοιχεία του Ω , δηλαδή η γνώση των πιθανοτήτων $p_t(l|\alpha,s)$ που δηλώνουν την πιθανότητα η εικόνα – παρατήρηση του συστήματος να είναι l δεδομένης της δράσης a και της πραγματικής κατάστασης s .

Για να επιλυθεί ένα πρόβλημα POMDP δεν απαιτείται να είναι γνωστή όλη η ιστορία του συστήματος, καθώς αυτό θα καθιστούσε τη διαδικασία μη Μαρκοβιανή. Αντιθέτως, είναι αρκετό να είναι γνωστή η αρχική πιθανότητα να είναι η κατάσταση του συστήματος s (initial belief state)

$$b_0(s)=P(S=s) \quad (A.1)$$

Μετά την εκλογή μιας δράσης και την παρατήρηση της κατάστασης του συστήματος απαιτείται η ανανέωση αυτής της πιθανότητας μέσω των πιθανοτήτων μετάβασης και των πιθανοτήτων των παρατηρήσεων χρησιμοποιώντας τον νόμο του Bayes

$$b'(s') = P(s'|l, a, b) = \frac{P(l|s', \alpha) \sum_{s \in S} P(s'|s, a) b(s)}{\sum_{s' \in S} P(l|s', \alpha) \sum_{s \in S} P(s'|s, a) b(s)} \quad (A.2)$$

ΠΑΡΑΡΤΗΜΑ Β

Αλγόριθμος Q-Learning

Η μέθοδος Q-Learning ανήκει σε μία ευρύτερη κατηγορία μεθόδων, γνωστή ως Reinforcement Learning (RL). Το βασικό χαρακτηριστικό της μεθόδου είναι ότι επιτρέπει στον ελεγκτή ενός συστήματος να δρα βέλτιστα, ανακαλύπτοντας τις συνέπειες των δράσεων που επιλέγει χωρίς να απαιτείται να είναι πλήρως γνωστό το μοντέλο του συστήματος (model-free RL). Ο αλγόριθμος συγκλίνει ακόμη και όταν το μέγεθος του συνόλου των δυνατών καταστάσεων είναι τεράστιο. Η κεντρική ιδέα είναι ο ελεγκτής ενός συστήματος να ανακαλύπτει τη βέλτιστη πολιτική επιλέγοντας σε κάθε κατάσταση συγκεκριμένη δράση και αξιολογώντας, μέσω μιας συνάρτησης οφέλους, τα ζεύγη καταστάσεων-δράσεων κάθε φορά που υπάρχει μετάβαση σε μια νέα κατάσταση.



Σχήμα Β. 1

Γίνεται η παραδοχή ότι το υπό μελέτη σύστημα είναι διακριτού και πεπερασμένου χρόνου με διακριτά σύνολα καταστάσεων $\{S\}$ και δράσεων $\{A\}$. Μέσω της επιλογής των κατάλληλων δράσεων στόχος του ελεγκτή είναι η μεγιστοποίηση του συνολικού αναμενόμενου οφέλους του συστήματος

$$V^\pi(x) = R(s, \pi(s)) + \gamma \sum_s P_{ss'}(\pi(s)) V^\pi(s') \quad (B.1)$$

όπου $V^\pi(x)$ η συνάρτηση οφέλους (value function), x η παρούσα κατάσταση του συστήματος, y η κατάσταση στην οποία γίνεται η μετάβαση, $\pi(x)$ η δράση που επιλέγεται

μέσω της πολιτικής π , P_{xy} η πιθανότητα μετάβασης και R_x το άμεσο όφελος. Η θεωρία του δυναμικού προγραμματισμού αποδεικνύει ότι υπάρχει τουλάχιστον μία βέλτιστη πολιτική που αποδίδει βέλτιστη τιμή οφέλους $V^*(s) = V^{\pi^*}(s) = \max_a V^\pi(s)$, υπό την προϋπόθεση, όμως, να είναι γνωστά τα μεγέθη P_{xy} και R_x .

Ο αλγόριθμος Q-Learning χρησιμοποιείται στην περίπτωση όπου η συμπεριφορά του συστήματος είναι άγνωστη, δηλαδή οι πιθανότητες μετάβασης και το άμεσο όφελος δεν είναι εκ των προτέρων γνωστές. Για συγκεκριμένη πολιτική π , ορίζεται η αξία Q ενός ζεύγους κατάστασης-δράσης μέσω της σχέσης

$$Q^\pi(s, a) = R(s, a) + \gamma \sum_{s' \in S} P_{ss'}(a) V^\pi(s') \quad (B.2)$$

ενώ η βέλτιστη τιμή του δίδεται από τη σχέση

$$Q^*(s, a) = Q^{\pi^*}(s, a) = R(s, a) + \gamma \sum_{s' \in S} P_{ss'}(a) V^{\pi^*}(s') \quad (B.3)$$

Ουσιαστικά, το μέγεθος $Q^\pi(s, a)$ είναι η παρούσα συνολική αναμενόμενη αξία του οφέλους του συστήματος αν επιλεγεί η δράση a στην κατάσταση s και ακολουθηθεί η πολιτική π εφεξής.

Είναι προφανές ότι ισχύει

$$V^* = \max_{a \in A} Q^*(s, a) \quad (B.4)$$

και αυτό επιτρέπει να εκφραστεί το μέγεθος $Q^*(s, a)$ μέσω της αναδρομικής σχέσης

$$Q^*(s, a) = Q^{\pi^*}(s, a) = R(s, a) + \gamma \sum_{s' \in S} P_{ss'}(a) \max_{a' \in A} Q^*(s', a') \quad (B.5)$$

Όταν μία δράση a^* επιτυγχάνει το μέγιστο για την αξία Q, η βέλτιστη πολιτική καταστρώνεται ως $\pi^*(s) = a^*$. Συνεπώς, εφόσον ο ελεγκτής του συστήματος μπορεί να μάθει τις διάφορες τιμές του μεγέθους Q, είναι σε θέση να ανακαλύψει και τη βέλτιστη πολιτική. Αν και μπορεί να προκύψουν περισσότερες της μίας βέλτιστες πολιτικές, οι τιμές $Q^*(s, a)$ είναι μοναδικές.

Ο αλγόριθμος Q-Learning βασίζεται στο ότι οι αρχικές τιμές $Q_0(x,a)$ είναι γνωστές εκ των προτέρων, η γνώση της συμπεριφοράς του συστήματος αποκτάται σε N στάδια καθώς γίνονται γνωστές οι τιμές Q_n και επιπλέον υπάρχει ένας πίνακας στον οποίο αποθηκεύονται οι τιμές Q_n για κάθε ζεύγος (s,a) . Στο n -οστό στάδιο ο αλγόριθμος έχει ως εξής:

ο ελεγκτής παρατηρεί στο στάδιο n την παρούσα κατάσταση του συστήματος s και επιλέγει να εφαρμόσει μια δράση a . Μετά τη μετάβαση στη νέα κατάσταση s' λαμβάνεται το άμεσο όφελος r και προσαρμόζονται οι τιμές Q_n σύμφωνα με το ακόλουθο σχήμα:

$$Q_{n+1}(s, a) = \begin{cases} Q_n(s, a) + \alpha \Delta Q_n(s, a) & \text{εάν } s = s_n \wedge a = a_n \\ Q_n(s, a) & \text{σε κάθε άλλη περίπτωση} \end{cases} \quad (\text{B.6})$$

όπου $\Delta Q_n(s, a) = r_n + \gamma \max_{a' \in A} Q_n(s', a') - Q_n(s, a)$ και α παράμετρος που υποδεικνύει το βαθμό στον οποίο ο ελεγκτής μαθαίνει τη συμπεριφορά του συστήματος (learning rate).

Στο [11] αποδεικνύεται ότι ο αλγόριθμος Q-Learning συγκλίνει, δηλαδή αποδεικνύεται το ακόλουθο θεώρημα:

Θεώρημα:

Δεδομένου ότι τα οφέλη που αποδίδει το σύστημα είναι φραγμένα $|r_n| \leq R$, η παράμετρος γνώσης του συστήματος $\alpha \in [0,1)$ και επιπλέον

$$\sum_{i=1}^{\infty} \alpha_{n^i(s,a)} = \infty, \sum_{i=1}^{\infty} [\alpha_{n^i(s,a)}]^2 < \infty \quad \forall s, a \quad (\text{B.7})$$

αποδεικνύεται ότι με πιθανότητα 1 $Q_n(s,a) \rightarrow Q^*(s,a)$ όταν $n \rightarrow \infty \quad \forall s, a$. Ο δείκτης $n^i(s,a)$ δηλώνει την i -οστή φορά όπου η δράση a επιλέγεται ενόσω το σύστημα βρίσκεται στην κατάσταση s .

ΠΑΡΑΡΤΗΜΑ Γ

ΟΡΟΛΟΓΙΑ

Διεθνής όρος	Ελληνικός όρος	Περιγραφή
Authentication	Ταυτοποίηση	Η διαδικασία ελέγχου της ταυτότητας ενός χρήστη πριν την πρόσβασή του στο δίκτυο.
Backward induction	Προς τα πίσω επαγωγή	Αλγόριθμος επίλυσης μαρκοβιανών διαδικασιών απόφασης πεπερασμένου ορίζοντα.
Call blocking probability	Πιθανότητα αποκλεισμού κλήσης	Η πιθανότητα μια κλήση να μη γίνει δεκτή από το σύστημα.
Clustering	Ομαδοποίηση	Η διαδικασία ομαδοποίησης των κόμβων ενός δικτύου.
Control overhead	Επιβάρυνση ελέγχου	Η επιβάρυνση που προκαλείται από τα μηνύματα που ανταλλάσσονται μεταξύ των κόμβων και δεν αφορούν την πληροφορία του χρήστη.
End-to-end delay	Από-άκρο-σε-άκρο καθυστέρηση	Η καθυστέρηση που εισάγεται κατά τη μετάδοση των δεδομένων από την πηγή προς τον προορισμό.

Expected transmission count	Αναμενόμενο πλήθος αναμεταδόσεων	Μέγεθος που εκφράζει, κατά μέσο όρο, πόσες αναμεταδόσεις έλαβαν χώρα πριν την επιτυχή εκπομπή ενός πακέτου.
Planar graph	Επίπεδος γράφος	Γράφος με ακμές που δεν τέμνονται μεταξύ τους.
Policy iteration	Αναδρομή πολιτικής	Αλγόριθμος επίλυσης μαρκοβιανών διαδικασιών απόφασης άπειρου ορίζοντα.
Round trip time	Χρόνος μετ' επιστροφής διαδρομής	Ο χρόνος που απαιτείται για να μεταδοθούν τα δεδομένα από την πηγή στον προορισμό και να σταλεί επιβεβαίωση λήψης τους από τον προορισμό στην πηγή.
Service scheduling	Προγραμματισμός υπηρεσιών	Ο προγραμματισμός συγκεκριμένου αριθμού εξυπηρετητών για την παροχή υπηρεσιών σε συγκεκριμένο πλήθος χρηστών.
Throughput	Ρυθμός διέλευσης	Μέσος ρυθμός επιτυχημένης παράδοσης πακέτων σε ένα δίαυλο.
Value iteration	Αναδρομή αξίας	Αλγόριθμος επίλυσης μαρκοβιανών διαδικασιών απόφασης άπειρου ορίζοντα.

BIBΛΙΟΓΡΑΦΙΑ

- [1] Martin L. Puterman, “Markov Decision Processes – Discrete Stochastic Dynamic Programming”, New York: Wiley, 1994
- [2] Oliver Yu, Emir Saric, Anfei Li, “Dynamic Control of Open Spectrum Management”, proceedings of the IEEE WCNC 2007, pp.127-132, 2007
- [3] Qing Zhao, Lang Tong, Ananthram Swami, “Decentralized Cognitive Mac for Dynamic Spectrum Access”, in proc. IEEE Symp. New Frontiers Dynamic Spectrum Access Networks Nov.2005, pp.224-232, 2005
- [4] Qianchuan Zhao, Stefan Geirhofer, Lang Tong, Brian M. Sadler, “Optimal Dynamic Spectrum Access via Periodic Channel Sensing”, in proc. of the IEEE WCNC 2007, pp.33-37,2007
- [5] Guy Shani, “Introduction to Partially Observable Markov Decision Processes (POMDPs)”, Machine Learning and Applied Statistics, Microsoft Research, 2009
- [6] Yunxia Chen, Qing Zhao, Ananthram Swami, “Bursty Traffic in Energy-Constrained Opportunistic Spectrum Access”, in proc. of IEEE GLOBECOM 2007, pp.4641-4646, 2007
- [7] Jayakrishnan Unnikrishnan, Venugopal V. Veeravalli, “Dynamic Spectrum Access Policies for Cognitive Radio”
- [8] Yunxia Chen, Qing Zhao, Ananthram Swami, “Distributed Cognitive MAC for Energy-Constrained Opportunistic Spectrum Access”
- [9] Yunxia Chen, Qing Zhao, Ananthram Swami, Lang Tong, “Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework”, IEEE journal on selected areas in communication, VOL.25, NO. 3, April 2007
- [10] Hua Liu, Bhaskar Krishnamachari, Qing Zhao, “Cooperation and Learning in Multiuser Opportunistic Spectrum Access”, in proc. of the ICC 2008 workshop, IEEE Communications Society, 2008
- [11] Fei Yu, Vincent W.S. Wong, Victor C.M. Leung, “Efficient QoS Provisioning for Adaptive Multimedia in Mobile Communication Networks by Reinforcement Learning”, proceedings of the 1st International Conference on Broadband Networks 2004, IEEE Computer Society,2004

- [12]Yin Sixing, Li Shufang, “Optimization for Time Consumption on Channel Searching in Cognitive Radio System with Markov Decision Processes”, IEEE Communications Society 2008
- [13]Jeongsoo Han, “Network – Adaptive QoS Routing Using Local Information”, Y.-T. Kim and M. Takano (Eds.): APNOMS 2006, LNCS 4238, pp. 190–199, 2006.
- [14]William A. Massey, K. G. Ramakrishnan, M. Aravamudan, G. Pai, “Scheduling Algorithms for Downlink Services in Wireless Networks: A Markov Decision Process Approach”, IEEE 2004
- [15]D.P. Bertsekas, “Dynamic Programming and Optimal Control”, Volume 2, Athena Scientific, 1995
- [16]Dr Boris Gremont, “Simulation of Rainfield Attenuation for Satellite Communication Networks”
- [17]Ian F. Akyildiz, Xudong Wang, Kiyon, “A Survey on Wireless Mesh Networks”, IEEE Radio Communications, September 2005
- [18]Hassnaa Moustafa, Usman Javaid, Tinku Mohamed Rasheed, Sidi-Mohammed Senouci, Djamel-Eddine Meddour “A Panorama on Wireless Mesh Networks: Architectures, Applications and Technical Challenges”, WiMeshNets 2006 workshop collocated with the International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks (QShine 2006), August, 2006
- [19] Young-Jin Kim, Ramesh Govindan, Brad Karp, Scott Shenker “On the Pitfalls of Geographic Face Routing”, *DIALMPOMC’05*, September 2, 2005, Cologne, Germany
- [20] Stefan Ruhrup, “Theory and Practice of Geographic Routing”, Department of Computer Science University of Freiburg, Germany, February 2009