

ΑΠΟΔΟΤΙΚΗ ΑΝΑΚΤΗΣΗ ΔΕΔΟΜΕΝΩΝ ΣΕ  
ΚΑΤΑΝΕΜΗΜΕΝΑ ΣΥΣΤΗΜΑΤΑ  
ΑΣΥΡΜΑΤΩΝ ΔΙΚΤΥΩΝ ΑΙΣΘΗΤΗΡΩΝ

Νικόλαος Χρ. Δημόκας

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ  
ΕΓΚΡΙΘΕΙΣΑ ΑΠΟ ΤΟ ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ  
ΤΟΥ ΑΡΙΣΤΟΤΕΛΕΙΟΥ ΠΑΝΕΠΙΣΤΗΜΙΟΥ ΘΕΣΣΑΛΟΝΙΚΗΣ

Φεβρουάριος 2010



η διατριβή αφιερώνεται

στους γονείς μου  
Χρήστο και Αποστολίνα



# Σύνοψη και Κυριότερα Επιτεύγματα της Διατριβής

Η παρούσα διατριβή εξετάζει το ζήτημα της ανάκτησης πληροφορίας σε ασύρματα περιβάλλοντα και ειδικότερα σε Ασύρματα Δίκτυα Αισθητήρων, με κίνητρο την ελάττωση της καθυστέρησης πρόσβασης των αισθητήρων στα προσφερόμενα δεδομένα και την ταυτόχρονη μείωση στην κατανάλωση ενέργειας των αισθητήρων με στόχο τη βελτίωση της επίδοσης και την επιμήκυνση της διάρκειας λειτουργίας του δικτύου.

Συγκεκριμένα, εστιάζει το ενδιαφέρον της στη γενικότερη τεχνική του συνεργατικού caching και στην ταυτόχρονη συνέπεια της cache, η οποία πρέπει να διασφαλίζεται. Παρουσιάζει ένα καινοτόμο αλγόριθμο ομαδοποίησης των αισθητήρων, ώστε να μειωθεί το πλήθος των εκπεμπόμενων μηνυμάτων και συνεπώς να επιμηκυνθεί η χρονική διάρκεια λειτουργίας του δικτύου. Αντιμετωπίζοντας επίσης τη ραγδαία εξάπλωση των ασύρματων δικτύων που χρησιμοποιούν τη μέθοδο της εναέριας εκπομπής δεδομένων, η διατριβή διερευνά τη δημιουργία εναέριων καταλόγων για την καλύτερη απόδοση του συστήματος.

Οι κυριότερες συνεισφορές της διατριβής συνοψίζονται στα εξής:

- Αποτελεί μία από τις πρώτες ερευνητικές προσπάθειες που πραγματεύονται το ζήτημα του συνεργατικού caching και της συνέπειας των δεδομένων σε Ασύρματα Πολυμεσικά Δίκτυα Αισθητήρων. Η παρούσα διατριβή παρουσιάζει μια νέα πολιτική συνεργατικού caching και μια νέα τεχνική διατήρησης της συνέπειας της cache. Η σχεδιάσή τους περιλαμβάνει την ανάλυση της τοπολογίας του δικτύου ώστε να επιλεγούν οι κόμβοι αισθητήρες που θα αναλάβουν το ρόλο του διαμεσολαβητή για να συντονίσουν τις αποφάσεις caching, ώστε να μειωθεί το κόστος επικοινωνίας, η κατανάλωση της ενέργειας, η καθυστέρηση προσπέλασης των δεδομένων και να αξιοποιηθεί βέλτιστα ο αποθηκευτικός χώρος της cache κάθε συνεργαζόμενου αισθητήρα. Η πολιτική επιλογής των κόμβων ήταν αποτέλεσμα έμπνευσης από την ερευνητική περιοχή των Κοινωνικών Δικτύων. Τα πρωτόκολλα συνεργατικού caching και συνέπειας της cache αποδείχθηκαν καλύτερα από τις ανταγωνιστικές μεθόδους που προέρχονται από την περιοχή των ασύρματων ad hoc δικτύων.
- Η εισαγωγή μιας νέας μετρικής για τον υπολογισμό της σπουδαιότητας ενός

## VI

κόμβου αισθητήρα και η εφαρμογή της για τη δημιουργία πρωτοκόλλου συνεργατικού caching σε Ασύρματα Δίκτυα Αισθητήρων. Η νέα μετρική συγκρίθηκε με αντίστοιχες από την ερευνητική περιοχή των Κοινωνικών Δικτύων και αποδείχθηκε ότι συμβάλλει στη βελτίωση της απόδοσης των δικτύων.

- Αποτελεί από τις πρώτες ερευνητικές προσπάθειες που πραγματοποιούνται τη δημιουργία εναέριων καταλόγων για μερικώς διατεταγμένα δεδομένα σε Ασύρματα Δίκτυα. Ο νέος κατάλογος αποδείχθηκε ότι αποτελεί μια γενίκευση δύο προηγούμενων υψηλής απόδοσης δομών καταλόγων. Ο προτεινόμενος κατάλογος είναι αποδοτικός ως προς τη μείωση της καταναλισκόμενης ενέργειας χωρίς να επιβαρύνει τη καθυστέρηση πρόσβασης.

## Σύντομο Βιογραφικό

Ο Νικόλαος Χρ. Δημόκας γεννήθηκε στα Γιαννιτσά, του νομού Πέλλης στις 26 Απριλίου 1978. Είναι μόνιμος κάτοικος Γιαννιτσών. Έλαβε το Πτυχίο Επιστήμης Υπολογιστών το Δεκέμβριο του 2001 και το Μεταπτυχιακό Δίπλωμα Ειδίκευσης τον Ιούλιο του 2004 από το Τμήμα Επιστήμης Υπολογιστών της Σχολής Θετικών Επιστημών του Πανεπιστημίου Κρήτης. Κατά το χρονικό διάστημα από τον Ιούλιο του 2001 έως τον Ιανουάριο του 2002 και από το Μάιο του 2004 μέχρι και το Δεκέμβριο του 2005, εργάστηκε ως μηχανικός έρευνας και ανάπτυξης στο Ινστιτούτο Πληροφορικής του Ιδρύματος Τεχνολογίας και Έρευνας στο Ηράκλειο Κρήτης. Έλαβε το Διδακτορικό Δίπλωμα το Φεβρουάριο του 2010 από το Τμήμα Πληροφορικής της Σχολής Θετικών Επιστημών του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης.





# Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές, ειλικρινείς ευχαριστίες μου στον κ. Ιωάννη Μανωλόπουλο, Καθηγητή του Τμήματος Πληροφορικής του Α.Π.Θ., και κύριο επιβλέποντα της διατριβής μου για την επιστημονική του διορατικότητα που με οδήγησε στη συγκεκριμένη ερευνητική κατεύθυνση. Επίσης, τον ευχαριστώ καθώς μου έδωσε τη δυνατότητα να συμμετάσχω σε διάφορα ερευνητικά και αναπτυξιακά έργα, γεγονός που μου προσέφερε πολύτιμες εμπειρίες και στήριξε ουσιαστικά τα οικονομικά μου. Θα ήθελα επίσης να ευχαριστήσω τον κ. Ιωάννη Βλαχάβα, Καθηγητή του Τμήματος Πληροφορικής του Α.Π.Θ., και την κ. Αθηνά Βακάλη, Αναπληρώτρια Καθηγήτρια του Τμήματος Πληροφορικής του Α.Π.Θ., οι οποίοι διετέλεσαν μέλη της τριμελούς επιτροπής επίβλεψης της διατριβής μου, καθώς και την κ. Ελένη Καρατζά, Καθηγήτρια του Τμήματος Πληροφορικής του Α.Π.Θ., τον κ. Γεώργιο Παπαδημητρίου, Αναπληρωτή Καθηγητή του Τμήματος Πληροφορικής του Α.Π.Θ., τον κ. Πέτρο Νικοπολιτίδη, Επίκουρο Καθηγητή του Τμήματος Πληροφορικής του Α.Π.Θ., και την κ. Νιόβη Παυλίδου, Καθηγήτρια του Τμήματος Ηλεκτρολόγων Μηχανικών και Μηχανικών Ηλεκτρονικών Υπολογιστών του Α.Π.Θ., οι οποίοι διετέλεσαν μέλη της εξεταστικής επιτροπής της διατριβής μου.

Θέλω να ευχαριστήσω τον καλό φίλο και συν-συγγραφέα Δημήτριο Κατσαρό, Λέκτορα στο Τμήμα Μηχανικών Η/Υ, Τηλεπικοινωνιών και Δικτύων του Πανεπιστημίου Θεσσαλίας, για τις συμβουλές του κατά τη διάρκεια του διδακτορικού καθώς και για την πολύτιμη συνεργασία. Επίσης θα ήθελα να ευχαριστήσω το συν-συγγραφέα μου σε κάποιες ερευνητικές εργασίες, καθηγητή του Τμήματος Μηχανικών Η/Υ, Τηλεπικοινωνιών και Δικτύων του Πανεπιστημίου Θεσσαλίας κ. Λέανδρο Τασσιούλα για την πολύτιμη συνεργασία του.

Θα ήθελα επίσης να ευχαριστήσω όλα τα μέλη του Εργαστηρίου Τεχνολογίας και Επεξεργασίας Δεδομένων του Τμήματος Πληροφορικής Α.Π.Θ. για την άριστη συνεργασία που διατηρήσαμε. Ξεχωριστές ευχαριστίες θέλω να απευθύνω στους καλούς φίλους που απέκτησα εκεί, τους Γιάννη Καρύδη, Μαρία Κοντάκη, Απόστολο Παπαδόπουλο, Δημήτρη Ραφαηλίδη και Αλέξη Παπαδημητρίου.

Πάνω απ' όλα όμως θέλω να ευχαριστήσω τους γονείς μου, Χρήστο και Αποστολίνα και τον αδελφό μου Γιώργο. Μου συμπαραστάθηκαν σε κάθε βήμα μου, σε κάθε προσπάθειά μου, τόσο ηθικά όσο και υλικά. Πραγματικά ένα μεγάλο της διατριβής ανήκει σε αυτούς. Θα ήθελα τέλος να ευχαριστήσω όλους όσους με αγαπούν.



# ΠΕΡΙΕΧΟΜΕΝΑ

<b>1</b>	<b>ΕΙΣΑΓΩΓΗ</b>	<b>1</b>
1.1	ΑΠΟΔΟΤΙΚΗ ΔΙΑΧΥΣΗ ΠΛΗΡΟΦΟΡΙΑΣ ΣΤΑ ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ ΑΙΣΘΗΤΗΡΩΝ . . . . .	1
1.2	ΣΥΝΕΙΣΦΟΡΕΣ ΤΗΣ ΔΙΑΤΡΙΒΗΣ . . . . .	3
1.3	ΒΙΒΛΙΟΓΡΑΦΙΑ . . . . .	5
<b>2</b>	<b>ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΓΙΑ ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ ΑΙΣΘΗΤΗΡΩΝ</b>	<b>7</b>
2.1	ΕΙΣΑΓΩΓΗ . . . . .	7
2.2	ΠΡΩΤΟΚΟΛΛΑ CACHING . . . . .	15
2.2.1	Πολιτικές αντικατάστασης . . . . .	19
2.3	ΤΕΧΝΙΚΕΣ ΔΙΑΤΗΡΗΣΗΣ ΣΥΝΕΠΕΙΑΣ ΤΗΣ CACHE . . . . .	22
2.3.1	Υπόβαθρο . . . . .	23
2.4	ΠΡΩΤΟΚΟΛΛΑ ΟΜΑΔΟΠΟΙΗΣΗΣ . . . . .	27
2.5	ΚΑΤΑΛΟΓΟΙ ΣΕ ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ . . . . .	31
2.6	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	33
<b>3</b>	<b>ΣΥΝΕΡΓΑΤΙΚΟ CACHING ΣΕ ΑΣΥΡΜΑΤΑ ΠΟΛΥΜΕΣΙΚΑ ΔΙΚΤΥΑ ΑΙΣΘΗΤΗΡΩΝ</b>	<b>35</b>
3.1	ΕΙΣΑΓΩΓΗ . . . . .	35
3.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	39
3.2.1	Υπόβαθρο . . . . .	39
3.2.2	Γιατί χρειαζόμαστε ένα νέο πρωτόκολλο caching; . . . . .	44
3.3	ΠΡΩΤΟΚΟΛΛΟ NiCoCa ΓΙΑ ΣΥΝΕΡΓΑΤΙΚΟ CACHING . . . . .	45
3.3.1	Μέτρηση σπουδαιότητας ενός κόμβου αισθητήρα . . . . .	46
3.3.2	Προκαταρκτικά για το πρωτόκολλο NiCoCa . . . . .	48
3.3.3	Συνιστώσα ανεύρεσης cache . . . . .	50
3.3.4	Συνιστώσα αντικατάστασης της cache . . . . .	54
3.4	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	55
3.4.1	Μοντέλο προσομοίωσης . . . . .	56
3.4.2	Μετρικές απόδοσης . . . . .	58
3.4.3	Πειραματικά αποτελέσματα . . . . .	58
3.5	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	67

<b>4</b>	<b>ΣΥΝΕΠΕΙΑ ΤΗΣ CACHE ΣΕ ΑΣΥΡΜΑΤΑ ΠΟΛΥΜΕΣΙΚΑ ΔΙΚΤΥΑ ΑΙΣΘΗΤΗΡΩΝ</b>	<b>69</b>
4.1	ΕΙΣΑΓΩΓΗ . . . . .	69
4.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	72
4.2.1	Κίνητρο και Συνεισφορές . . . . .	75
4.3	ΜΟΝΤΕΛΟ ΔΙΚΤΥΟΥ . . . . .	76
4.4	ΠΡΩΤΟΚΟΛΛΟ ΣΥΝΕΠΕΙΑΣ NICC . . . . .	77
4.4.1	Προσέγγιση push-pull . . . . .	80
4.4.2	Απαιτήσεις συνέπειας . . . . .	84
4.4.3	Πολιτική αντικατάστασης της cache . . . . .	85
4.5	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	86
4.5.1	Μοντέλο προσομοίωσης . . . . .	86
4.5.2	Μετρικές απόδοσης . . . . .	88
4.5.3	Αποτίμηση . . . . .	89
4.6	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	106
<b>5</b>	<b>ΑΠΟΔΟΤΙΚΟ ΣΥΝΕΡΓΑΤΙΚΟ CACHING ΣΕ ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ ΑΙΣΘΗΤΗΡΩΝ</b>	<b>109</b>
5.1	ΕΙΣΑΓΩΓΗ . . . . .	109
5.1.1	Κίνητρο και συνεισφορές . . . . .	111
5.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	112
5.3	ΜΟΝΤΕΛΟ ΔΙΚΤΥΟΥ . . . . .	112
5.4	ΝΕΑ ΣΧΗΜΑΤΑ ΣΥΝΕΡΓΑΤΙΚΟΥ CACHING . . . . .	113
5.4.1	Power community index . . . . .	113
5.4.2	Εκτεταμένος power community index . . . . .	116
5.4.3	Πρωτόκολλα extPCICC και PCICC . . . . .	118
5.5	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	124
5.5.1	Μοντέλο προσομοίωσης . . . . .	124
5.5.2	Μετρικές απόδοσης . . . . .	126
5.5.3	Αποτίμηση . . . . .	126
5.6	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	133
<b>6</b>	<b>ΟΜΑΔΟΠΟΙΗΣΗ ΣΕ ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ ΑΙΣΘΗΤΗΡΩΝ</b>	<b>135</b>
6.1	ΕΙΣΑΓΩΓΗ . . . . .	135
6.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	138
6.3	ΜΟΝΤΕΛΟ ΔΙΚΤΥΟΥ . . . . .	142
6.4	ΤΟ ΝΕΟ ΠΡΩΤΟΚΟΛΛΟ ΟΜΑΔΟΠΟΙΗΣΗΣ . . . . .	142
6.4.1	Το πρωτόκολλο ομαδοποίησης . . . . .	143
6.5	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	147
6.5.1	Αποτίμηση . . . . .	149
6.6	ΕΦΑΡΜΟΓΕΣ ΟΜΑΔΟΠΟΙΗΣΗΣ . . . . .	155
6.7	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	159

<b>7</b>	<b>ΚΑΤΑΛΟΓΟΠΟΙΗΣΗ ΣΕ ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ</b>	<b>161</b>
7.1	ΕΙΣΑΓΩΓΗ . . . . .	161
7.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	165
7.2.1	Προκαταρκτικά . . . . .	165
7.2.2	Υπόβαθρο . . . . .	167
7.2.3	Κατάλογοι για μερικώς διατεταγμένα δεδομένα . . . . .	169
7.2.4	Γενικεύοντας τους προηγούμενους καταλόγους . . . . .	169
7.3	PARTIAL ORDERING BROADCAST INDEX . . . . .	170
7.3.1	Κατεύθυνση αναζήτησης στα υποδένδρα κάθε ομάδας . . . . .	171
7.4	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	172
7.4.1	Επίδραση του αριθμού των αντικειμένων . . . . .	174
7.4.2	Επίδραση του αριθμού των ομάδων . . . . .	175
7.4.3	Επίδραση του μεγέθους της ομάδας . . . . .	177
7.4.4	Επίδραση της δημοτικότητας της ομάδας . . . . .	177
7.5	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	179
<b>8</b>	<b>ΕΠΙΛΟΓΟΣ</b>	<b>181</b>
8.1	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	181
8.2	ΔΡΟΜΟΙ ΜΕΛΛΟΝΤΙΚΗΣ ΕΡΕΥΝΑΣ . . . . .	184
<b>A</b>	<b>ΛΙΣΤΑ ΕΡΕΥΝΗΤΙΚΩΝ ΕΡΓΑΣΙΩΝ</b>	<b>203</b>



# ΛΙΣΤΑ ΣΧΗΜΑΤΩΝ

2.1	Παράδειγμα Ασύρματου Δικτύου Αισθητήρων. . . . .	8
2.2	Εφαρμογή ΑΔΑ σε πολεμικό πεδίο. . . . .	10
2.3	Τρία σχεδιαστικά ζητήματα για τη διατήρηση της συνέπειας της cache. . . . .	24
3.1	Υπολογισμός δείκτη $NI$ για δύο παραδείγματα γράφων. . . . .	48
3.2	Αίτηση από τον κόμβο $SN_i$ προωθείται στον caching κόμβο $SN_g$ . . . . .	51
3.3	Αλγόριθμος αναζήτησης πρωτοκόλλου NiCoCa. . . . .	53
3.4	Πολιτική αντικατάστασης πρωτοκόλλου NiCoCa. . . . .	54
3.5	Πολιτική αποδοχής πρωτοκόλλου NiCoCa. . . . .	55
3.6	Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB, $\theta=0.0$ και $0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) με 100 αισθητήρες. . . . .	59
3.7	Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB, $\theta=0.0$ και $0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 100 αισθητήρες. . . . .	59
3.8	Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB, $\theta=0.0$ και $0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) με 100 αισθητήρες. . . . .	60
3.9	Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB, $\theta=0.0$ και $0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 100 αισθητήρες. . . . .	60
3.10	Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB, $\theta=0.0$ και $0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) με 500 αισθητήρες. . . . .	61
3.11	Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB, $\theta=0.0$ και $0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 500 αισθητήρες. . . . .	61
3.12	Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB, $\theta=0.0$ και $0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) με 500 αισθητήρες. . . . .	62
3.13	Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB, $\theta=0.0$ και $0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 500 αισθητήρες. . . . .	62
3.14	Επίδραση μεγέθους της cache στον αριθμό των μηνυμάτων (δεδομένα μεγέθους MB, $\theta=0.0$ και $0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 500 αισθητήρες. . . . .	62

3.15	Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB, $\theta=0.8$ ) σε ένα αραιό ( $d=7$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 1000 αισθητήρες. . . . .	63
3.16	Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB, $\theta=0.8$ ) σε ένα αραιό ( $d=7$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 1000 αισθητήρες. . . . .	63
3.17	Επίδραση μεγέθους της cache στο πλήθος των μηνυμάτων (δεδομένα μεγέθους MB, $\theta=0.8$ ) σε ένα αραιό ( $d=7$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 1000 αισθητήρες. . . . .	64
3.18	Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους KB, $\theta=0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) για πλήθος αισθητήρων 100, 500 και 1000. . . . .	65
3.19	Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους KB, $\theta=0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) για πλήθος αισθητήρων 100, 500 και 1000. . . . .	65
3.20	Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους KB, $\theta=0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) για πλήθος αισθητήρων 100, 500 και 1000. . . . .	66
3.21	Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους KB, $\theta=0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) για πλήθος αισθητήρων 100, 500 και 1000. . . . .	66
4.1	Παράδειγμα πρωτοκόλλου NICC. . . . .	79
4.2	Αλγόριθμος push και πολιτική αντικατάστασης πρωτοκόλλου NICC. . . . .	81
4.3	Αλγόριθμος pull πρωτοκόλλου NICC. . . . .	82
4.4	Παράδειγμα λειτουργίας πρωτοκόλλου NICC όταν απαιτείται ισχυρή συνέπεια ή $\Delta$ συνέπεια και η τιμή του $\Delta$ είναι 0. . . . .	84
4.5	Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες. . . . .	90
4.6	Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες. . . . .	90
4.7	Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες. . . . .	91
4.8	Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες. . . . .	91
4.9	Επίδραση μεγέθους της cache στον αριθμό των μηνυμάτων (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες. . . . .	91
4.10	Επίδραση μεγέθους της cache στον αριθμό των μηνυμάτων (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες. . . . .	93
4.11	Καθυστέρηση πρόσβασης σε δεδομένα μεγέθους MB σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 100 αισθητήρες. . . . .	93



4.12 Καθυστερήση πρόσβασης σε δεδομένα μεγέθους MB σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 500 αισθητήρες. . . . .	93
4.13 Δικτυακή κίνηση σε δεδομένα μεγέθους MB σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 100 αισθητήρες. . . . .	94
4.14 Δικτυακή κίνηση σε δεδομένα μεγέθους MB σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 500 αισθητήρες. . . . .	94
4.15 % διαφορά στην κατανάλωση ενέργειας (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες, όταν εκτελούνται το NICC και το RPCC. . . . .	94
4.16 % διαφορά στην κατανάλωση ενέργειας (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100, αισθητήρες όταν εκτελούνται το NICC και το RPCC. . . . .	96
4.17 Επίδραση μεγέθους της cache (δεδομένα μεγέθους KB, $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΠΔΑ ( $d=4$ ) με 100 αισθητήρες. . . . .	97
4.18 Επίδραση μεγέθους της cache (δεδομένα μεγέθους KB, $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 100 αισθητήρες. . . . .	97
4.19 Επίδραση μεγέθους της cache (δεδομένα μεγέθους KB, $\theta=0.8$ ) στα: (Επάνω) hits, (Μέση) καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΠΔΑ ( $d=4$ ) με 500 αισθητήρες. . . . .	98
4.20 Επίδραση μεγέθους της cache (δεδομένα μεγέθους KB, $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΔΑ ( $d=10$ ) με 500 αισθητήρες. . . . .	98
4.21 Καθυστερήση πρόσβασης (δεδομένα μεγέθους KB) σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 100 αισθητήρες. . . . .	99
4.22 Καθυστερήση πρόσβασης (δεδομένα μεγέθους KB) σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 500 αισθητήρες. . . . .	99
4.23 Δικτυακή κίνηση (δεδομένα μεγέθους KB) σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 100 αισθητήρες. . . . .	100
4.24 Δικτυακή κίνηση (δεδομένα μεγέθους KB) σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 500 αισθητήρες. . . . .	100
4.25 % διαφορά στην κατανάλωση ενέργειας (δεδομένα μεγέθους KB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες όταν εκτελούνται τα NICC και RPCC. . . . .	100
4.26 % διαφορά στην κατανάλωση ενέργειας (δεδομένα μεγέθους KB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες, όταν εκτελούνται το NICC και RPCC. . . . .	101

4.27	Επίδραση τιμής <i>TTL</i> στα: (Επάνω) αριθμό μηνυμάτων και καθυστέρηση πρόσβασης, (Κάτω) hits σε ένα ΑΠΔΑ με 500 αισθητήρες.	102
4.28	Επίδραση της αντικατάστασης της cache στην καθυστέρηση πρόσβασης σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες.	102
4.29	Επίδραση της αντικατάστασης της cache στην καθυστέρηση πρόσβασης σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες.	103
4.30	Επίδραση της αντικατάστασης της cache στα hits σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες.	104
4.31	Επίδραση της αντικατάστασης της cache στα hits σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες.	104
4.32	Μέσος όρος αντιγράφων ανά αντικείμενο για διαφορετικές τιμές βαθμού σε ένα ΑΠΔΑ με 100 και 500 αισθητήρες.	105
4.33	Μέσος όρος αντιγράφων ανά διαμεσολαβητή για διαφορετικές τιμές βαθμού σε ένα ΑΠΔΑ με 100 και 500 αισθητήρες.	105
4.34	Μέσος όρος διαμεσολαβητών για διαφορετικές τιμές βαθμού σε ένα ΑΠΔΑ με 100 και 500 αισθητήρες.	106
5.1	Οι δείκτες <i>NI</i> για ένα μικρό γράφο.	114
5.2	Υπολογισμός του <i>PCI</i> για ένα γράφο.	115
5.3	Ένα παράδειγμα μη ορθής συμπεριφοράς του <i>PCI</i> .	117
5.4	Υπολογισμός του δείκτη <i>extPCI</i> για ένα γράφο.	118
5.5	Το πρωτόκολλο <i>extPCICC</i> .	120
5.6	Μια αίτηση από τον κόμβο <i>Q</i> προωθείται στον caching κόμβο <i>G</i> .	121
5.7	Επίδραση του μεγέθους της cache ( $\theta=0.0$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΔΑ ( $d=4$ ) με 100 αισθητήρες.	127
5.8	Επίδραση του μεγέθους της cache ( $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΔΑ ( $d=4$ ) με 100 αισθητήρες.	127
5.9	Επίδραση του μεγέθους της cache ( $\theta=0.0$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΔΑ ( $d=10$ ) με 100 αισθητήρες.	128
5.10	Επίδραση του μεγέθους της cache ( $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΔΑ ( $d=10$ ) με 100 αισθητήρες.	128
5.11	Επίδραση του μεγέθους της cache ( $\theta=0.0$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΔΑ ( $d=4$ ) με 500 αισθητήρες.	130
5.12	Επίδραση του μεγέθους της cache ( $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΔΑ ( $d=4$ ) με 500 αισθητήρες.	131
5.13	Επίδραση του μεγέθους της cache ( $\theta=0.0$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΔΑ ( $d=10$ ) με 500 αισθητήρες.	131

5.14	Επίδραση του μεγέθους της cache ( $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΔΑ ( $d=10$ ) με 500 αισθητήρες. . . . .	132
6.1	Επίδραση του αριθμού των κόμβων στο σχηματισμό των ομάδων για: (Επάνω) βαθμός ίσος με 4 και 7, (Κάτω) βαθμός ίσος με 10. . . . .	150
6.2	Επίδραση του αριθμού των κόμβων στη μετάδοση για: (Επάνω) βαθμός ίσος με 4 και 7, (Κάτω) βαθμός ίσος με 10. . . . .	151
6.3	Επίδραση του μέσου βαθμού κόμβου στο πλήθος μεταδόσεων για ένα δίκτυο: (α) 100 αισθητήρων, (β) 300 αισθητήρων, (γ) 500 αισθητήρων, (δ) 1000 αισθητήρων. . . . .	153
6.4	Εναπομείνασα ενέργεια για βαθμό ίσο με 10 για ένα δίκτυο: (α) 100 αισθητήρων, (β) 300 αισθητήρων, (γ) 500 αισθητήρων, (δ) 1000 αισθητήρων. . . . .	154
6.5	Εναπομείνασα ενέργεια για βαθμό ίσο με 7 για ένα δίκτυο: (α) 100 αισθητήρων, (β) 300 αισθητήρων, (γ) 500 αισθητήρων, (δ) 1000 αισθητήρων. . . . .	155
6.6	Εναπομείνασα ενέργεια για βαθμό ίσο με 4 για ένα δίκτυο: (α) 100 αισθητήρων, (β) 300 αισθητήρων, (γ) 500 αισθητήρων, (δ) 1000 αισθητήρων. . . . .	156
6.7	Η απόδοση του πρωτοκόλλου GESG σε δικτυακές εφαρμογές, (Επάνω) πρώτος θάνατος κόμβου, (Κάτω) τελευταίος θάνατος κόμβου. . . . .	158
6.8	Ο συνολικός αριθμός από δεδομένα που έλαβε ο σταθμός βάσης για ένα δίκτυο: (Επάνω) 300 και 500 κόμβων, (Κάτω) 1000 κόμβων. . . . .	159
7.1	Οφέλη από την αξιοποίηση της κυρτότητας στην προσπέλαση και του μη ισοζυγισμένου δένδρου. . . . .	167
7.2	Επίδραση του αριθμού των αντικειμένων στην απόδοση του δένδρου. . . . .	175
7.3	Επίδραση του αριθμού των ομάδων στην απόδοση του δένδρου. . . . .	176
7.4	Επίδραση του μεγέθους της ομάδας στην απόδοση του δένδρου. . . . .	178
7.5	Επίδραση της δημοτικότητας της ομάδας. . . . .	178



# ΚΕΦΑΛΑΙΟ 1

## Εισαγωγή

### Περιεχόμενα

---

1.1	ΑΠΟΔΟΤΙΚΗ ΔΙΑΧΥΣΗ ΠΛΗΡΟΦΟΡΙΑΣ ΣΤΑ ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ ΑΙΣΘΗΤΗΡΩΝ . . . . .	1
1.2	ΣΥΝΕΙΣΦΟΡΕΣ ΤΗΣ ΔΙΑΤΡΙΒΗΣ . . . . .	3
1.3	ΒΙΒΛΙΟΓΡΑΦΙΑ . . . . .	5

---

### 1.1 Αποδοτική Διάχυση Πληροφορίας στα Ασύρματα Δίκτυα Αισθητήρων

Τα τελευταία χρόνια, τα Ασύρματα Δίκτυα Αισθητήρων (Wireless Sensor Networks) προκάλεσαν μεγάλο ενδιαφέρον τόσο στην ερευνητική κοινότητα όσο και στη βιομηχανία και έδωσαν τεράστια ώθηση στην ανάπτυξη ενός μεγάλου αριθμού εφαρμογών. Η ανάπτυξη των Ασύρματων Δικτύων Αισθητήρων (ΑΔΑ) περιλαμβάνουν δύο θεμελιώδεις στόχους. Ο πρώτος σχετίζεται με τη μείωση της κατανάλωσης ενέργειας των αισθητήρων, ενώ ο δεύτερος με τη μείωση της καθυστέρησης πρόσβασης (latency) των δεδομένων. Για να αντιμετωπισθεί το πρόβλημα αυτό, θα μπορούσαμε να αυξήσουμε το εύρος ζώνης (bandwidth) του υποκείμενου δικτύου, δηλαδή του ΑΔΑ. Αυτή η λύση όμως δεν είναι πάντα επιθυμητή ή εφικτή, εξαιτίας είτε των περιορισμών που τίθενται στα ΑΔΑ είτε της σημαντικής οικονομικής επιβάρυνσης ή της ανεπάρκειας της τεχνολογίας.

Ο κύριος λόγος για την παρατηρούμενη καθυστέρηση οφείλεται στους περιορισμούς που τίθενται στα ΑΔΑ. Οι αισθητήρες είναι συσκευές με περιορισμένη ισχύ μπαταρίας, υπολογιστικές δυνατότητες και μνήμη. Η ποιότητα του καναλιού σε ένα ΑΔΑ είναι πολύ χαμηλή, ενώ το εύρος ζώνης πολύ περιορισμένο. Παράλληλα, η επικοινωνία μεταξύ δύο αισθητήρων ενός ΑΔΑ, οι οποίοι δεν βρίσκονται ο ένας στην ακτίνα μετάδοσης του άλλου, επιτυγχάνεται μέσω των ενδιάμεσων κόμβων που αναμεταδίδουν τα μηνύματα έτσι ώστε να δημιουργηθεί ένα μονοπάτι επικοινωνίας μεταξύ των δύο κόμβων (multi-hop<sup>1</sup> επικοινωνία). Συνεπώς το multi-hop μοντέλο επικοινωνίας σε συνδυασμό με την πυκνή ανάπτυξη των αισθητήρων και το περιορισμένο εύρος ζώνης έχει ως επακόλουθο την αύξηση της μέσης καθυστέρησης πρόσβασης στα προσφερόμενα δεδομένα.

Στην παρούσα διατριβή εξετάζονται μέθοδοι διάχυσης της πληροφορίας με ταυτόχρονη μείωση της κατανάλωσης ενέργειας στα ΑΔΑ. Στόχος είναι η βελτίωση της παρατηρούμενης ταχύτητας πρόσβασης των αισθητήρων στην προσφερόμενη πληροφορία και η επιμήκυνση του χρόνου ζωής του δικτύου. Αξιοποιείται η τεχνική του caching<sup>2</sup>, δηλαδή η αποθήκευση της πληροφορίας σε ενδιάμεσους “αντιπροσώπους” των κόμβων πηγών, η τεχνική της συνέπειας της cache<sup>3</sup>, η τεχνική της ομαδοποίησης (clustering) και η τεχνική των καταλόγων ή ευρετηρίων (indexes). Η έννοια του αντιπροσώπου όπως χρησιμοποιείται στην περίπτωση αυτή αναφέρεται στην cache ενός κόμβου αισθητήρα του ΑΔΑ. Η έννοια του κόμβου πηγής αντιπροσωπεύει τον αισθητήρα που παράγει κάποια δεδομένα και αποτελεί την πηγή αυτών των δεδομένων.

Αντικείμενο της παρούσης μελέτης είναι η ελάττωση της καθυστέρησης πρόσβασης στα δεδομένα και η επιμήκυνση του χρόνου ζωής του δικτύου. Το πρώτο ζήτημα που ερευνάται είναι η αναγνώριση και ο συντονισμός, με δυναμικό και κατανεμημένο τρόπο, των αισθητήρων που θα αναλάβουν να πραγματοποιήσουν το caching των δεδομένων, καθώς και η αναγνώριση των αντικειμένων που θα πρέπει να στεγάζονται στις caches. Το δεύτερο ζήτημα σχετίζεται με τη διατήρηση της συνέπειας των δεδομένων που είναι τοποθετημένα στις caches, ενώ το τρίτο αναφέρεται στην αποδοτική δημιουργία ομάδων αισθητήρων ώστε να μειωθεί το πλήθος των εκπεμπόμενων μηνυμάτων και συνεπώς να επιμηκυνθεί η χρονική διάρκεια λειτουργίας του δικτύου. Τέλος, εξετάζουμε τη δημιουργία εναέριων

<sup>1</sup> Η λέξη hop έχει αποδοθεί στα ελληνικά με την έκφραση σκέλος (εναέριος) διαδρομής, αλλά προτιμούμε τη χρήση της καθιερωμένης αγγλικής έκφρασης.

<sup>2</sup> Η λέξη caching έχει αποδοθεί στα ελληνικά με την έκφραση εναποθήκευση, αλλά προτιμούμε τη χρήση της καθιερωμένης αγγλικής έκφρασης.

<sup>3</sup> Η λέξη cache έχει αποδοθεί στα ελληνικά με την έκφραση κρυφή μνήμη, αλλά προτιμούμε τη χρήση της καθιερωμένης αγγλικής έκφρασης.

καταλόγων για την καλύτερη απόδοση του συστήματος. Στην επόμενη ενότητα περιγράφονται σαφέστερα οι συνεισφορές της παρούσας διατριβής.

## 1.2 Συνεισφορές της Διατριβής

Στο Κεφάλαιο 2 παρουσιάζονται διάφορες εισαγωγικές έννοιες κατάλληλες για την καλύτερη κατανόηση των ζητημάτων που θα εξετασθούν στα επόμενα κεφάλαια της παρούσας διατριβής.

Στο Κεφάλαιο 3 εξετάζεται το πρόβλημα του συνεργατικού caching στην ερευνητική περιοχή των Ασύρματων Πολυμεσικών Δικτύων Αισθητήρων (ΑΠΔΑ). Οι καινοτόμες εφαρμογές των ΑΠΔΑ απαιτούν την επανεξέταση του σχεδιαστικού μοντέλου υπολογισμού-επικοινωνίας που ακολουθήθηκε στα παραδοσιακά ΑΔΑ. Το μοντέλο αυτό επικεντρώθηκε κυρίως στη μείωση της κατανάλωσης ενέργειας, αποσκοπώντας στην επιμήκυνση του χρόνου ζωής των δικτύων αισθητήρων. Ωστόσο, οι εφαρμογές που υλοποιούνται από τα ΑΠΔΑ έχουν ένα δεύτερο στόχο εξίσου σημαντικό με τη μείωση της κατανάλωσης ενέργειας. Ο στόχος αυτός είναι η παράδοση Ποιότητας Υπηρεσίας (Quality of Service – QoS) σε επίπεδο εφαρμογής και η αντιστοίχιση αυτής της απαίτησης στις μετρικές του δικτυακού επιπέδου (network layer), όπως είναι η καθυστέρηση πρόσβασης.

Ως λύση στο προηγούμενο ζήτημα επινοήθηκε το πρωτόκολλο συνεργατικού caching, το οποίο ονομάζεται “NiCoCa”. Το NiCoCa σχεδιάστηκε έτσι ώστε να ανιχνεύει τους κόμβους αισθητήρες που βρίσκονται στα “κεντρικότερα” σημεία των γειτονιών του δικτύου και να τους αναθέτει το ρόλο του διαμεσολαβητή για να συντονίσουν τις αποφάσεις caching, ώστε να μειωθεί το κόστος επικοινωνίας, η κατανάλωση της ενέργειας, η καθυστέρηση πρόσβασης των δεδομένων και να αξιοποιηθεί βέλτιστα ο αποθηκευτικός χώρος της cache κάθε συνεργαζόμενου αισθητήρα. Προκύπτει ότι το νέο πρωτόκολλο είναι καλύτερο από το ανταγωνιστικό του ως προς τα global hits<sup>4</sup>, τα remote hits και την καθυστέρηση πρόσβασης.

Στο Κεφάλαιο 4 εξετάζεται το πρόβλημα της συνέπειας της cache (cache consistency). Η τεχνική του caching προκαλεί μια πρόκληση: πώς θα διατηρήσουμε τα περιεχόμενα της cache *συνεπή ή φρέσκα* (fresh) σε σχέση με τα αντίστοιχα πρωτότυπά που βρίσκονται στην πηγή. Οι στρατηγικές που βασίζονται στο push, υποφέρουν από περιττές μεταδόσεις ενημερώσεων και προκαλούν “καταιγίδες” (burst) στη λειτουργία του κόμβου πηγή, ενώ οι στρατηγικές που βασίζονται στο

<sup>4</sup> Η λέξη hit έχει αποδοθεί στα ελληνικά με την έκφραση *κτύπημα*, αλλά προτιμούμε τη χρήση της καθιερωμένης αγγλικής έκφρασης.

pull, υποφέρουν από μία πλημμυρίδα μηνυμάτων με αποτέλεσμα μεγάλη επιβάρυνση στην επικοινωνία και στην κατανάλωση ενέργειας. Λαμβάνοντας υπόψη τους περιορισμένους πόρους των αισθητήρων, του εύρους μετάδοσης αλλά κυρίως της ανάγκης για ενημερωμένα και έγκυρα δεδομένα σχεδιάσαμε μια push-pull πολιτική.

Ως λύση στα προηγούμενα ζητήματα επινοήθηκε μία πολιτική συνέπειας της cache με όνομα “NICC”. Δείχνεται ότι η πολιτική αυτή είναι καλύτερη από την ανταγωνιστική της και ότι μπορεί να προσφέρει υψηλές επιδόσεις ακόμα και για μικρή cache. Παράλληλα προτείνεται μία νέα πολιτική αντικατάστασης της cache, η οποία υπερτερεί της ανταγωνιστικής της πολιτικής και συμβάλλει στη βελτίωση της απόδοσης του πρωτοκόλλου NICC.

Στο Κεφάλαιο 5 μελετούμε το ζήτημα του συνεργατικού caching σε AΔΑ. Κίνητρό μας υπήρξε η βελτίωση της επιλογής των κόμβων που θα πραγματοποιήσουν το συντονισμό και τη συνεργασία μεταξύ των αισθητήρων και η ανάπτυξη μιας νέας αποδοτικής πολιτικής ελέγχου εισδοχής στη cache, η οποία θα λαμβάνει υπόψη τα δεδομένα που έχουν τοποθετηθεί στις caches των γειτονικών αισθητήρων.

Ως λύση στο προηγούμενο πρόβλημα επινοήθηκαν νέα πρωτόκολλα, τα επονομαζόμενα “PCICC” και “extPCICC”, που βασίστηκαν σε δύο νέες μετρικές για την επιλογή των σημαντικών κόμβων αισθητήρων. Η πρώτη μετρική *PCI* αξιοποιεί την πληροφορία που σχετίζεται με την 1-hop γειτονιά ενός κόμβου, ενώ η δεύτερη μετρική *extPCI* αποτελεί βελτίωση της πρώτης και ο υπολογισμός της βασίζεται στη γνώση των συνδέσεων μεταξύ των κόμβων της 2-hop γειτονιάς. Δείχνεται ότι οι αλγόριθμοι είναι καλύτεροι από τον ανταγωνιστή τους.

Στο Κεφάλαιο 6 εξετάζεται το ζήτημα της ομαδοποίησης των κόμβων. Στα πλαίσια της ομαδοποίησης διακρίνουμε δύο μεγάλες οικογένειες τεχνικών. Αυτές που στηρίζονται στη κατασκευή ενός κυρίαρχου συνόλου (dominating set) και αυτές που βασίζονται στη θεώρηση μόνο της ενέργειας. Η πρώτη οικογένεια τεχνικών υποφέρει από το γεγονός ότι μόνο ένα υποσύνολο των κόμβων του δικτύου είναι υπεύθυνο για την αναμετάδοση των μηνυμάτων με αποτέλεσμα να προκαλείται ταχεία κατανάλωση της ενέργειας σε αυτούς τους κόμβους. Η δεύτερη οικογένεια τεχνικών χρησιμοποιεί την εναπομείνασα ενέργεια κάθε κόμβου για τον προσδιορισμό του κόμβου που θα ηγείται της ομάδας, αγνοώντας τα δικτυακά τοπολογικά χαρακτηριστικά.

Ως λύση στα προηγούμενα ζητήματα περιγράφεται ένας νέος ενεργειακά αποδοτικός κατανεμημένος αλγόριθμος ομαδοποίησης, ο οποίος ονομάζεται “GESC”. Ο νέος αλγόριθμος αξιοποιεί την τοπική δομή του δικτύου και την εναπομείνασα



ενέργεια των γειτονικών κόμβων για να ορίσει ένα νέο τρόπο δυναμικού υπολογισμού των κόμβων κεφαλών ομάδων (cluster heads). Δείχνεται ότι ο αλγόριθμος είναι καλύτερος από τους ανταγωνιστές του, τόσο στη μείωση της καταναλισκόμενης ενέργειας, όσο και στον αριθμό των μηνυμάτων που αναμεταδίδονται στο δίκτυο.

Στο Κεφάλαιο 7 εξετάζεται το ζήτημα της δημιουργία καταλόγων για πληροφορίες που εκπέμπονται σε ΑΔΑ υπό το πρίσμα των μερικώς διατεταγμένων δεδομένων. Έχει αποδειχθεί ότι η διαστρωμάτωση των καταλόγων με τα δεδομένα μπορεί να αντιμετωπίσει τη διασπάθιση της ενέργειας εξαιτίας της σειριακής αναζήτησης των δεδομένων στις μεταδόσεις εκπομπής χωρίς να επιβαρύνεται η καθυστέρηση πρόσβασης. Ωστόσο όλες οι προταθείσες τεχνικές υποθέτουν μια συνολική διάταξη δεδομένων, χωρίς κάποια να θεωρεί τη γενική περίπτωση της μερικής διάταξης. Επιπρόσθετα, είναι ισορροπημένες δομές κάτι που δεν ταιριάζει στη γραμμική δομή (μία διάσταση) του ασύρματου μέσου και δεν λαμβάνουν υπόψη την κυρτότητα (skewness) στο πρότυπο προσπέλασης. Τέλος, όλες οι δομές καταλόγων απαιτούν ισομεγέθη αντικείμενα.

Ως λύση στα προηγούμενα προβλήματα επινοήθηκε ένας νέος μη ισορροπημένος κατάλογος κατάλληλος για κυρτά πρότυπα προσπέλασης. Ο προτεινόμενος κατάλογος αποτελεί μια γενίκευση δύο προηγούμενων αντίστοιχων μεθόδων υψηλής απόδοσης. Ο νέος κατάλογος είναι αποδοτικός ως προς τη μείωση της καταναλισκόμενης ενέργειας χωρίς να επιβαρύνει την καθυστέρηση πρόσβασης.

### 1.3 Βιβλιογραφία

Το Κεφάλαιο 2 περιέχει βασικές έννοιες και ζητήματα των ΑΔΑ και αποτελεί μία βιβλιογραφική ανασκόπηση των ζητημάτων που πραγματεύεται η διατριβή. Το Κεφάλαιο 3 βασίζεται στο υλικό της εργασίας [43, 44]. Το Κεφάλαιο 4 βασίζεται στο υλικό της εργασίας [47]. Το Κεφάλαιο 5 βασίζεται στο υλικό της εργασίας [45]. Το Κεφάλαιο 6 βασίζεται στο υλικό της εργασίας [42, 46]. Το Κεφάλαιο 7 βασίζεται στο υλικό της εργασίας [80]. Ο πλήρης κατάλογος των ερευνητικών εργασιών βρίσκεται στο *Παράρτημα Α* της παρούσης διατριβής.



## ΚΕΦΑΛΑΙΟ 2

# Βασικές Έννοιες για Ασύρματα Δίκτυα Αισθητήρων

### Περιεχόμενα

---

2.1	ΕΙΣΑΓΩΓΗ . . . . .	7
2.2	ΠΡΩΤΟΚΟΛΛΑ CACHING . . . . .	15
2.3	ΤΕΧΝΙΚΕΣ ΔΙΑΤΗΡΗΣΗΣ ΣΥΝΕΠΕΙΑΣ ΤΗΣ CACHE . . . . .	22
2.4	ΠΡΩΤΟΚΟΛΛΑ ΟΜΑΔΟΠΟΙΗΣΗΣ . . . . .	27
2.5	ΚΑΤΑΛΟΓΟΙ ΣΕ ΑΣΥΡΜΑΤΑ ΔΙΚΤΥΑ . . . . .	31
2.6	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	33

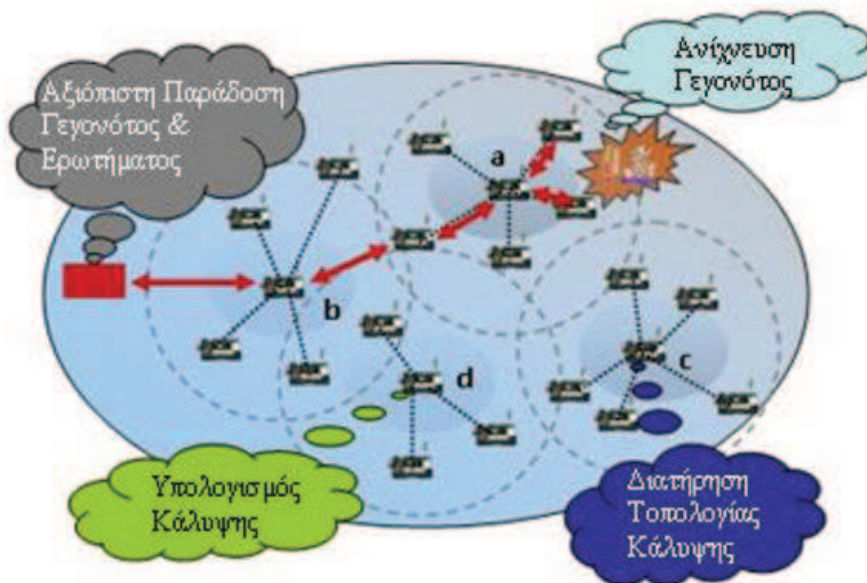
---

### 2.1 Εισαγωγή

Την τελευταία δεκαετία οι εξελίξεις στη σμίκρυνση των ηλεκτρονικών κυκλωμάτων, στις ασύρματες επικοινωνίες και στη δημιουργία κυκλωμάτων χαμηλής κατανάλωσης ενέργειας έχουν προσφέρει τη δυνατότητα ανάπτυξης χαμηλού κόστους, χαμηλής κατανάλωσης ενέργειας, πολυ-λειτουργικών αισθητήρων που είναι μικροί σε μέγεθος και μπορούν να επικοινωνούν σε μικρές αποστάσεις. Αυτοί οι μικροσκοπικοί αισθητήρες που είναι ικανοί να επεξεργάζονται δεδομένα, να επικοινωνούν και να αλληλεπιδρούν με το περιβάλλον, έδωσαν πνοή στην ιδέα

των δικτύων αισθητήρων, τα οποία βασίζονται στη συνεργατική προσπάθεια ενός μεγάλου αριθμού αισθητήρων.

Ένα δίκτυο αισθητήρων (Σχήμα 2.1) συντίθεται από ένα μεγάλο αριθμό κόμβων αισθητήρων που αναπτύσσονται με πυκνό τρόπο είτε μέσα στην περιοχή στόχο, είτε πολύ κοντά σε αυτή. Η θέση των κόμβων δεν χρειάζεται να είναι προκαθορισμένη αλλά επιτρέπεται η ανάπτυξή τους σε μη προσβάσιμες περιοχές. Επιπρόσθετα, αυτό σημαίνει ότι τα πρωτόκολλα και οι αλγόριθμοι που απευθύνονται σε ΑΔΑ πρέπει να διέπονται από δυνατότητες αυτο-οργάνωσης. Ένα άλλο μοναδικό στοιχείο των δικτύων αισθητήρων είναι η συνεργατική προσπάθεια των κόμβων τους. Οι κόμβοι αισθητήρες είναι εφοδιασμένοι με ένα μικροεπεξεργαστή. Επομένως, αντί να στέλνουν τα λαμβανόμενα δεδομένα απευθείας στους υπεύθυνους κόμβους για την επεξεργασία και τη συνάθροιση των δεδομένων, οι κόμβοι αισθητήρες χρησιμοποιούν τις δυνατότητές τους στην επεξεργασία, έτσι ώστε να πραγματοποιούν τοπικά απλούς υπολογισμούς και να μεταδίδουν μόνο τα απαιτούμενα και μερικώς επεξεργασμένα δεδομένα.



Σχήμα 2.1. Παράδειγμα Ασύρματου Δικτύου Αισθητήρων.

Τα δίκτυα αισθητήρων μπορούν να αποτελούνται από διαφορετικούς τύπους αισθητήρων όπως για παράδειγμα είναι οι σεισμικοί, οι θερμικοί, οι οπτικοί, οι υπέρυθροι, οι ακουστικοί, οι αισθητήρες ραντάρ κ.α. και μπορούν να παρακολου-

θούν μία μεγάλη ποικιλία περιβαλλοντικών συνθηκών που περιλαμβάνουν τα εξής [50]:

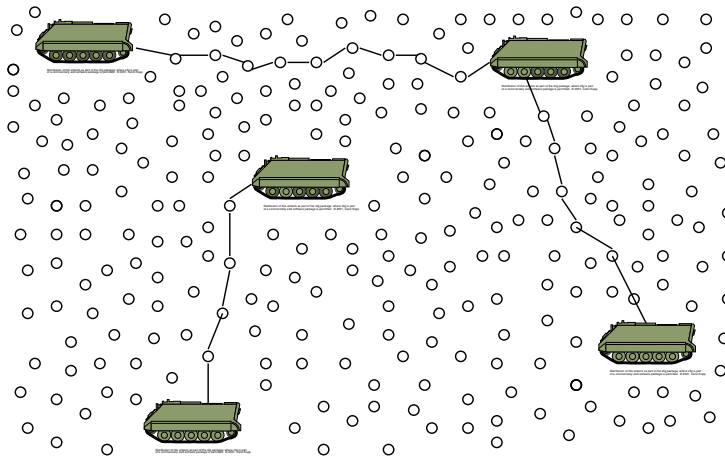
- θερμοκρασία,
- υγρασία,
- κίνηση οχημάτων,
- πίεση,
- επίπεδα θορύβου,
- ακτινοβολία,
- την παρουσία ή απουσία συγκεκριμένων αντικειμένων,
- σύνθεση εδάφους,
- επίπεδα πίεσης σε προσαρτημένα αντικείμενα και
- τα τρέχοντα χαρακτηριστικά όπως είναι η ταχύτητα, η κατεύθυνση και το μέγεθος ενός αντικειμένου.

Τα ΑΔΑ [5, 78] προκάλεσαν το ενδιαφέρον της ερευνητικής κοινότητας και της βιομηχανίας και έδωσαν τεράστια ώθηση στην ανάπτυξη ενός πλήθους εφαρμογών, όπως είναι η ειδοποίηση στην περίπτωση καταστροφών, ο έλεγχος του περιβάλλοντος, η παρακολούθηση της βιομηχανικής παραγωγής, ο έλεγχος αποθεμάτων και εμπορευμάτων, η γεωργία ακρίβειας, η απομακρυσμένη παρακολούθηση ιατρικών δεδομένων, η βοήθεια ηλικιωμένων, η παρακολούθηση στρατευμάτων σε πολεμικά πεδία , κ.α. Για παράδειγμα, ας εξετάσουμε δύο εφαρμογές δικτύων αισθητήρων:

- *Παρακολούθηση φιλίων δυνάμεων, εξοπλισμού και πυρομαχικών (Σχήμα 2.2).* Οι διοικητές των στρατιωτικών δυνάμεων στο πεδίο της μάχης μπορούν διαρκώς να παρακολουθούν και να εποπτεύουν την κατάσταση των φιλίων δυνάμεων, την κατάσταση και τη διαθεσιμότητα του εξοπλισμού και των πυρομαχικών με τη χρήση ενός δικτύου αισθητήρων. Κάθε ομάδα στρατιωτών, όχημα, εξοπλισμός και ζωτικής σημασίας πολεμοφόδια μπορούν να φέρουν μικρούς αισθητήρες που αναφέρουν την κατάστασή τους. Οι αναφορές, που αποστέλλονται, συλλέγονται από τους sink<sup>1</sup> κόμβους και

---

<sup>1</sup> Η λέξη sink έχει αποδοθεί στα ελληνικά με την έκφραση δεξαμενή, αλλά προτιμούμε τη χρήση της καθιερωμένης αγγλικής έκφρασης.



Σχήμα 2.2. Εφαρμογή ΑΔΑ σε πολεμικό πεδίο.

στέλνονται στους διοικητές των στρατευμάτων. Τα δεδομένα μπορούν επίσης να προωθηθούν στα υψηλότερα επίπεδα της στρατιωτικής ιεραρχίας, ενώ συναθροίζονται με δεδομένα άλλων μονάδων σε κάθε επίπεδο.

- *Ανίχνευση πυρκαγιάς σε δασική περιοχή.* Οι κόμβοι αισθητήρες μπορούν να αναπτυχθούν με τυχαίο και ταυτοχρόνως πυκνό τρόπο σε ένα δάσος και συνεπώς μπορούν να αναμεταδίδουν την ακριβή θέση και κατεύθυνση της πυρκαγιάς στους τελικούς χρήστες πριν η πυρκαγιά λάβει ανεξέλεγκτες διαστάσεις. Εκατομύρια αισθητήρες μπορούν να αναπτυχθούν και να ενοποιηθούν χρησιμοποιώντας ραδιοφωνικές συχνότητες και οπτικά συστήματα. Οι κόμβοι αισθητήρες μπορούν επίσης να εφοδιασθούν με αποδοτικές μεθόδους ενίσχυσης της ισχύος τους [12], όπως είναι τα ηλιακά κύτταρα, επειδή με αυτό τον τρόπο μπορούν να λειτουργούν για μήνες ακόμα και χρόνια. Οι κόμβοι αισθητήρες συνεργάζονται μεταξύ τους για να πραγματοποιούν κατανεμημένη παρακολούθηση και να ξεπερνούν τα φυσικά εμπόδια, όπως είναι τα δένδρα, οι πέτρες, κ.α.

Η πραγματοποίηση των προηγούμενων αλλά και άλλων εφαρμογών σε δίκτυα αισθητήρων απαιτεί τεχνικές των ασύρματων ad hoc δικτύων. Αν και έχουν προταθεί πολλά πρωτόκολλα και εφαρμογές για τα παραδοσιακά ασύρματα ad hoc δίκτυα, εντούτοις αυτά δεν μπορούν να ικανοποιήσουν πλήρως τα μοναδικά χαρακτηριστικά και τις απαιτήσεις των εφαρμογών σε δίκτυα αισθητήρων. Για να διαφωτίσουμε πλήρως αυτό το σημείο παραθέτουμε συνοπτικά τις διαφορές μεταξύ των δικτύων αισθητήρων και των ad hoc δικτύων [103]:

- Το πλήθος των κόμβων σε ένα δίκτυο αισθητήρων μπορεί να είναι πολύ μεγαλύτερος από το πλήθος των κόμβων ενός ad hoc δικτύου.
- Οι κόμβοι αισθητήρες αναπτύσσονται σε μια περιοχή με πυκνό τρόπο.
- Οι κόμβοι αισθητήρες είναι επιρρεπείς στις αποτυχίες.
- Η τοπολογία ενός δικτύου αισθητήρων στις περισσότερες περιπτώσεις είναι στατική.
- Οι κόμβοι αισθητήρες χρησιμοποιούν κυρίως τη μετάδοση με εκπομπή (broadcasting), ενώ τα περισσότερα ad hoc δίκτυα βασίζονται σε επικοινωνίες σημείο-προς-σημείο.
- Οι κόμβοι αισθητήρες εμφανίζουν σημαντικούς περιορισμούς ως προς την ενέργεια της μπαταρίας, την υπολογιστική ισχύ και τη μνήμη.
- Οι κόμβοι αισθητήρες μπορεί να μην έχουν γενικό αναγνωριστικό λόγω του πλήθους τους.

Η υποστήριξη μιας τόσο μεγάλης ποικιλίας εφαρμογών θα ήταν αδύνατη για μια και μοναδική πραγματοποίηση ενός ΑΔΑ. Ωστόσο, εμφανίζονται αδιαφιλονίκητα γνώρισματα σχετικά με τα χαρακτηριστικά και τους απαιτούμενους μηχανισμούς αυτών των συστημάτων και συνεπώς η πραγματοποίηση αυτών των χαρακτηριστικών είναι η σημαντικότερη πρόκληση που αντιμετωπίζεται σε αυτή την κατηγορία δικτύων.

Το πρώτο και σημαντικότερο χαρακτηριστικό είναι η περιορισμένη ισχύς της μπαταρίας και η απαίτηση για χαμηλή κατανάλωση ενέργειας. Οι κόμβοι αισθητήρες βασίζονται συνήθως σε μια μπαταρία περιορισμένης διάρκειας ζωής και η αντικατάστασή της δεν είναι δυνατή είτε εξαιτίας φυσικών περιορισμών (οι αισθητήρες βρίσκονται σε ωκεανούς ή σε εχθρικό περιβάλλον), είτε επειδή δεν είναι συμφέρον για τον ιδιοκτήτη του ΑΔΑ. Αν και η τεχνολογία των μπαταριών έχει βελτιωθεί με την πάροδο των χρόνων, η πρόοδος αυτή είναι αργή και η μπαταρία παραμένει ένα προβληματικό σημείο (bottleneck) για τα ΑΔΑ. Στις περισσότερες περιπτώσεις, τα κυκλώματα που είναι υπεύθυνα για την επικοινωνία και οι κεραίες είναι τα σημαντικότερα συστατικά ενός αισθητήρα, τα οποία καταναλώνουν το μεγαλύτερο ποσοστό της ενέργειας. Η επέκταση του χρόνου λειτουργίας των αισθητήρων μπορεί να επιτευχθεί με την εξοικονόμηση ενέργειας. Ταυτόχρονα με το χαρακτηριστικό της περιορισμένης ισχύς της μπαταρίας εμφανίζεται

και η ασύμμετρη κατανάλωση ενέργειας καθώς ο πομποδέκτης καταναλώνει περισσότερη ενέργεια κατά τη μετάδοση (uplink) από ό,τι κατά τη λήψη (downlink).

Ένα τρίτο χαρακτηριστικό που πρέπει να ληφθεί υπόψη κατά τη σχεδίαση πρωτοκόλλων για ΑΔΑ είναι η κλιμακωσιμότητα. Η αρχιτεκτονική και τα πρωτόκολλα για ΑΔΑ πρέπει να είναι ικανά να κλιμακώνονται (δηλαδή, να αξιοποιούν) οποιοδήποτε πλήθος αισθητήρων. Στις περισσότερες περιπτώσεις τα δίκτυα αισθητήρων περιλαμβάνουν πολύ μεγαλύτερο πλήθος κόμβων από ότι τα κινητά ad hoc δίκτυα (MANETs).

Οι εφαρμογές των δικτύων αισθητήρων δεν προϋποθέτουν γενικά το χαρακτηριστικό της κίνησης. Επομένως, το δίκτυο είναι απελευθερωμένο από το φόρτο εργασίας της αναγνώρισης νέων επικοινωνιακών μονοπατιών και συνακόλουθα υποφέρει λιγότερο από τα μηνύματα ελέγχου. Έτσι, ένα ΑΔΑ μπορεί να χρησιμοποιήσει απλούστερες μεθόδους δρομολόγησης από ότι τα MANET. Ένα ακόμα χαρακτηριστικό που πρέπει να αξιοποιηθεί κατά τη σχεδίαση πρωτοκόλλων είναι η διασύνδεση με βάση τα δεδομένα. Ο στόχος ενός συμβατικού δικτύου επικοινωνίας είναι η μεταφορά bits από μία μηχανή σε μία άλλη, αλλά ο πραγματικός σκοπός ενός δικτύου αισθητήρων είναι η παροχή πληροφοριών και απαντήσεων, όχι απλά αριθμών [72]. Εξίσου σημαντική όμως είναι και η απαίτηση που προκύπτει από τη μεγάλη ποικιλία στην πυκνότητα ανάπτυξης των κόμβων αισθητήρων στην περιοχή παρακολούθησης. Η παράταξη των κόμβων αισθητήρων μπορεί να μην είναι κανονική και είναι δυνατόν να ποικίλλει σημαντικά, ανάλογα με την εφαρμογή, στη διάσταση του χώρου και του χρόνου. Συνεπώς, σε κάποια σημεία του δικτύου μπορεί να παρατηρείται πολύ μεγάλη συγκέντρωση αισθητήρων, ενώ κάποια άλλα να είναι πολύ αραιά.

Μια επιπρόσθετη σημαντική πρόκληση είναι το χαμηλής ποιότητας κανάλι. Η ποιότητα του καναλιού σε ένα ΑΔΑ είναι πολύ χειρότερη από αυτή των ενσύρματων δικτύων λόγω του θορύβου, των παρεμβολών, του fading κ.α. Η εκπομπή ενός μηνύματος μπορεί να διακοπεί εξαιτίας κάποιου κτηρίου, δένδρου, εμποδίου, θορύβου ή άλλων παραγόντων.

Η επικοινωνία μεταξύ δύο κόμβων αισθητήρων ενός ΑΔΑ, οι οποίοι δεν βρίσκονται ο ένας στην ακτίνα μετάδοσης του άλλου, επιτυγχάνεται μέσω των ενδιάμεσων κόμβων που αναμεταδίδουν τα μηνύματα έτσι ώστε να δημιουργηθεί ένα μονοπάτι επικοινωνίας μεταξύ των δύο κόμβων. Επειδή ο αριθμός των κόμβων αισθητήρων είναι πολύ μεγάλος και σχηματίζουν μία πυκνή τοπολογία, οι γειτονικοί κόμβοι μπορεί να βρίσκονται πολύ κοντά μεταξύ τους. Έτσι είναι επιθυμητό η multi-hop επικοινωνία σε δίκτυα αισθητήρων να καταναλώνει λιγότερη ισχύ από ότι η παραδοσιακή single-hop επικοινωνία. Η multi-hop επικοινωνία μπορεί



επίσης αποδοτικά να ανταπεξέλθει κάποια επακόλουθα της διάδοσης του σήματος, τα οποία παρατηρούνται σε ασύρματες επικοινωνίες μεγάλων αποστάσεων. Παρ' όλα αυτά, ο multi-hop τρόπος επικοινωνίας των ΑΔΑ συνεπάγεται ότι η χωρητικότητα κάθε ασύρματου συνδέσμου εξαρτάται από το επίπεδο παρεμβολής μεταξύ των κόμβων.

Η σημαντικότερη απαίτηση, από όσες περιγράφηκαν προηγουμένως, την οποία πρέπει να πληροί η σχεδίαση ενός πρωτοκόλλου για δίκτυα αισθητήρων είναι η ελαχιστοποίηση της κατανάλωσης ενέργειας και συνεπώς η επιμήκυνση του χρόνου ζωής του δικτύου. Τα τελευταία χρόνια όμως ανακύπτει η ανάγκη για υποστήριξη μικρής καθυστέρησης πρόσβασης των δεδομένων. Πολλές εφαρμογές που εκτελούνται σε ΑΔΑ απαιτούν τη βελτιστοποίηση των επικοινωνιών μεταξύ των αισθητήρων ώστε να εξυπηρετούν την αίτηση για δεδομένα με μικρή καθυστέρηση πρόσβασης των δεδομένων και με ελάχιστη κατανάλωση ενέργειας. Πιο συγκεκριμένα, οι πρόσφατες τεχνολογικές εξελίξεις με την ανάπτυξη φθηνών καμερών CMOS και μικροφώνων είχαν ως αποτέλεσμα τη δημιουργία των ΑΠΔΑ [3]. Το κύριο χαρακτηριστικό που εμφανίζουν οι εφαρμογές που υλοποιούνται στα ΑΠΔΑ είναι ότι εκτός από την απαιτούμενη χαμηλή κατανάλωση ενέργειας, οφείλουν να παραδώσουν τα δεδομένα με μικρή χρονική καθυστέρηση.

Μια τεχνική που έχει προταθεί για τη μείωση της καθυστέρησης πρόσβασης των δεδομένων και η οποία συμβάλλει και στη μείωση της κατανάλωσης ενέργειας είναι η τεχνική του συνεργατικού caching, δηλαδή της προσωρινής αποθήκευσης δεδομένων εγγύτερα στον κόμβο πελάτη. Το caching σώζει υπολογιστικούς πόρους, συγκεκριμένα εύρος ζώνης δικτύου, αφού λιγότερα πακέτα ταξιδεύουν στο δίκτυο, και χρόνο αφού επιτυγχάνονται μικρότεροι χρόνοι απόκρισης. Παράλληλα, επειδή ελαττώνονται τα μηνύματα που ταξιδεύουν στο δίκτυο, μειώνεται και η κατανάλωση ενέργειας στους κόμβους αισθητήρες. Η τεχνική του caching έχει ευρέως εξερευνηθεί στα πεδία των Λειτουργικών Συστημάτων και των Βάσεων Δεδομένων και συνεχίζει να αποτελεί μια ελκυστική περιοχή έρευνας [95]. Παρομοίως, η τεχνική του caching έχει διερευνηθεί εκτενώς στα πλαίσια του Παγκόσμιου Ιστού [51, 82, 101], στα πλαίσια των ασύρματων κυψελοειδών δικτύων [81, 131] και σε MANET [113, 127, 149]. Το caching μπορεί να υλοποιηθεί σε διάφορα σημεία κατά μήκος της διαδρομής της ροής των δεδομένων από τον κόμβο πηγή μέχρι τον τελικό κόμβο παραλήπτη. Το σημείο που επιλέγεται είναι πολύ σημαντικό για τα ΑΔΑ επειδή καθορίζει το πλήθος των μηνυμάτων που διαδίδονται στο δίκτυο και την καθυστέρηση με την οποία θα εξυπηρετηθεί η αίτηση του κόμβου αισθητήρα. Εξίσου σημαντικό κομμάτι της κάθε cache είναι ο αλγόριθμος που καθορίζει ποια αντικείμενα θα στεγασθούν στην cache [32]. Αυτός ο

μηχανισμός αποκαλείται *πολιτική αντικατάστασης* (replacement policy).

Η τεχνική του caching έχει ως παρενέργεια το πρόβλημα της διατήρησης των περιεχομένων της cache ώστε να είναι *συνεπής* (consistent ή coherent) ή *φρέσκα* (fresh) σε σχέση με τα αντίστοιχα πρωτότυπα που βρίσκονται στην πηγή. Η συνέπεια της cache έχει ιδιαίτερη σημασία για τα δίκτυα αισθητήρων, ιδιαίτερα σε εφαρμογές όπως για παράδειγμα σε ΑΔΑ που αναπτύσσονται σε πολεμικά πεδία, όπου οι μετρούμενες ποσότητες πρέπει να είναι ακριβείς και έγκυρες αν και τροποποιούνται συχνά. Συνεπώς τα cached αντίγραφα πρέπει αναλόγως να *ακυρώνονται* (invalidated) ή να *ενημερώνονται* (updated).

Η απόδοση ενός ΑΔΑ μπορεί να ωφεληθεί σημαντικά από την τεχνική της ομαδοποίησης, η οποία μελετήθηκε εκτενώς στα πλαίσια των MANET [156] και συντελεί στην οργάνωση των κόμβων αισθητήρων του δικτύου σε ομάδες, ώστε να επιτευχθεί η αποδοτική συνάθροιση της πληροφορίας και να περιορισθεί η μετάδοση των δεδομένων κυρίως μέσα στην κάθε ομάδα. Συνεπώς, η ομαδοποίηση συμβάλλει στη μείωση του αριθμού των μηνυμάτων που διαδίδονται στο δίκτυο και επομένως στην κατανάλωση της ενέργειας των αισθητήρων. Παράλληλα, μειώνεται ο ανταγωνισμός για την απόκτηση του ασύρματου καναλιού και η καθυστέρηση πρόσβασης των δεδομένων.

Στο υπόλοιπο του κεφαλαίου, θα παρουσιάσουμε πληροφορίες που σχετίζονται με τα ζητήματα του συνεργατικού caching, της συνέπειας της cache, της τεχνικής ομαδοποίησης καθώς και των εναέριων καταλόγων στο περιβάλλον των ΑΔΑ. Δεν θα παρουσιάσουμε μια εξαντλητική αναφορά όλων των σχετικών ζητημάτων, καθώς κάτι τέτοιο μπορεί να βρεθεί στις εργασίες [3, 4, 50].

Ο στόχος του κεφαλαίου είναι διπλός. Αρχικά παρουσιάσαμε και διασαφήνισαμε τις ιδιαιτερότητες του περιβάλλοντος των ΑΔΑ, εξαιτίας των οποίων χρειαζόμαστε νέες λύσεις για το συνεργατικό caching, τις πολιτικές αντικατάστασης, τη συνέπεια της cache, τη ομαδοποίηση και τους εναέριους καταλόγους. Στη συνέχεια θα δείξουμε πώς αυτές οι ιδιαιτερότητες κατέστησαν τις παλιές λύσεις (σχεδιασμένες για Λειτουργικά Συστήματα, Βάσεις Δεδομένων, ασύρματα κυψελοειδή δίκτυα και MANET) μη ικανοποιητικές για τα ΑΔΑ, και πώς έδωσαν ώθηση στη δημιουργία νέων. Δεύτερος στόχος του κεφαλαίου είναι να παρουσιάσει συνοπτικά μια κατηγοριοποίηση των τεχνικών που έχουν προταθεί μέχρι στιγμής και να περιγράψει τους σημαντικότερους αλγόριθμους κάθε κατηγορίας.

## 2.2 Πρωτόκολλα Caching

Η τεχνική του caching μελετήθηκε σε μεγάλο βάθος στα πεδία των Λειτουργικών Συστημάτων, των Βάσεων Δεδομένων και στα πλαίσια του Παγκόσμιου Ιστού τόσο για συνεργατικές [28, 51, 111] όσο και για μη συνεργατικές [82, 101] αρχιτεκτονικές. Αν και το πρόβλημα του caching μελετήθηκε εκτενώς στους ανωτέρω ερευνητικούς τομείς, τα ιδιαίτερα χαρακτηριστικά που εμφανίζονται στα ΑΔΑ και γενικότερα στα ασύρματα δίκτυα είναι μοναδικά όπως περιγράφηκαν στην αρχή του κεφαλαίου. Στη συνέχεια, αρχικά θα αναλύσουμε τα σημεία που πρέπει να στοχεύει η σχεδίαση ενός πρωτοκόλλου caching για ΑΔΑ. Επίσης, θα αναφέρουμε τις κυριότερες εργασίες που πραγματοποιήθηκαν στα πλαίσια του συνεργατικού caching, τόσο στα MANET όσο και στα ΑΔΑ. Τέλος θα περιγράψουμε το ζήτημα της πολιτικής αντικατάστασης, το οποίο σχετίζεται άμεσα με το caching.

Όπως τονίστηκε στην αρχή του κεφαλαίου, οι κυριότερες ιδιότητες που εμφανίζονται στα ΑΔΑ είναι η περιορισμένη ισχύς της μπαταρίας, η ασύμμετρη κατανάλωση ενέργειας, η ανεπάρκεια εύρους ζώνης, το χαμηλής ποιότητας κανάλι και η μη ύπαρξη κίνησης. Παράλληλα όμως στα πλαίσια της σχεδίασης πρωτοκόλλων caching εμφανίζεται ένα ακόμα σημαντικό χαρακτηριστικό που είναι η περιορισμένη cache. Το μέγεθος της cache των αισθητήρων είναι μικρό συγκρινόμενο με τις ενσύρματες συσκευές. Συνεπώς, η επιλογή ενός αποδοτικού αλγορίθμου προσπέλασης της cache και μιας αποδοτικής πολιτικής αντικατάστασης είναι ζωτικής σημασίας. Εξίσου σημαντική ιδιότητα που πρέπει να ληφθεί υπόψη κατά τη σχεδίαση των πρωτοκόλλων είναι η κατανεμημένη προσπέλαση και ενημέρωση των δεδομένων. Ένας αισθητήρας δεν γνωρίζει εάν τα δεδομένα έχουν ενημερωθεί ή έχουν προσπελασθεί από άλλους αισθητήρες εκτός και αν ενημερωθεί από το δίκτυο.

Εξαιτίας των προηγούμενων προκλήσεων, οι στόχοι στη σχεδίαση των πολιτικών caching δεν είναι ίδιοι όπως αυτοί του caching σε ενσύρματα δίκτυα (για παράδειγμα στον Παγκόσμιο Ιστό). Οι στόχοι που τίθενται είναι οι εξής:

- **Αύξηση του cache hit:** Υπάρχουν τρεις διαφορετικοί τύποι hits: local, remote και global. Ένα local hit συμβαίνει όταν το ζητούμενο αντικείμενο βρίσκεται στην cache του κόμβου που υποβάλλει το ερώτημα και είναι έγκυρο, ενώ ένα remote hit συμβαίνει όταν το ζητούμενο αντικείμενο βρίσκεται στην cache ενός άλλου κόμβου. Τέλος, ένα global hit συμβαίνει όταν το ζητούμενο αντικείμενο αναχτάται από τον κόμβο πηγή. Προφανώς, μια καλή πολιτική caching και κυρίως συνεργατικού caching πρέπει να

εμφανίζει υψηλό αριθμό local/remote hits και χαμηλό αριθμό global hits.

- **Μείωση του κόστους επικοινωνίας:** Το κόστος επικοινωνίας είναι ένα πολύ σημαντικό ζήτημα, επειδή το εύρος ζώνης είναι περιορισμένο. Συνεπώς μια αποτελεσματική και αποδοτική πολιτική caching θα πρέπει να συμβάλλει στη μείωση του πλήθους των διαδιδόμενων μηνυμάτων.
- **Μείωση της κατανάλωσης ενέργειας:** Η μείωση της κατανάλωσης της ενέργειας είναι ένα εξίσου σημαντικό ζήτημα που σχετίζεται με τη μείωση του κόστους επικοινωνίας. Όσο λιγότερα μηνύματα διαχέονται στο δίκτυο, τόσο λιγότερη ενέργεια καταναλώνεται από τους κόμβους αισθητήρες.
- **Μείωση της καθυστέρησης πρόσβασης των δεδομένων:** Το caching συνδέεται άμεσα με τη μείωση στη καθυστέρηση πρόσβασης, καθώς η ανάγκη των δεδομένων γίνεται ταχύτερα.

Θεωρώντας ότι κάθε κόμβος αισθητήρας έχει ένα τοπικό αποθηκευτικό χώρο, μπορούμε να διεκπεραιώσουμε τις προαναφερθείσες προκλήσεις των δικτύων αισθητήρων με την αξιοποίηση της τεχνικής του συνεργατικού caching. Στο συνεργατικό caching, πολλοί κόμβοι αισθητήρες διαμοιράζουν και συντονίζουν τα δεδομένα της cache χωρίς να χρειάζεται πάντα να επισκέπτονται τους κόμβους πηγές των δεδομένων. Η συνεργασία μεταξύ των κόμβων αισθητήρων είναι επιβεβλημένη εξαιτίας των περιορισμών χώρου της cache, στην ανεπάρκεια του εύρους ζώνης και στην ισχύ της μπαταρίας. Η διάρκεια ζωής της μπαταρίας μπορεί να επεκταθεί αν μειώσουμε τον όγκο της επικοινωνίας που πραγματοποιείται μεταξύ των αισθητήρων. Η τεχνική του caching των χρήσιμων δεδομένων για κάθε κόμβο αισθητήρα είτε στον τοπικό αποθηκευτικό χώρο είτε στους γειτονικούς κόμβους μπορεί να επιμηκύνει τη διάρκεια ζωής του δικτύου, να μειώσει το υψηλό κόστος επικοινωνίας και να μειώσει το φόρτο εργασίας των πηγών των δεδομένων. Επιπρόσθετα, το caching μπορεί να είναι πολύ αποδοτικό στη μείωση των μεταδόσεων μηνυμάτων σε μεγάλη έκταση του δικτύου. Επομένως μπορεί να μειώσει την παρεμβολή και ανταπεξέλθει τις συνθήκες μεταβλητού καναλιού. Το συνεργατικό caching μπορεί επίσης να επιταχύνει την επεξεργασία των δεδομένων, επειδή η επεξεργασία και η παράδοση των δεδομένων δεν είναι ανεξάρτητες οντότητες. Συνεπώς, το συνεργατικό caching έχει ως αποτέλεσμα την ελάττωση της καθυστέρησης πρόσβασης, της κατανάλωσης ενέργειας και της απώλειας πακέτων.

Αν και η αναγκαιότητα για αποδοτικές πολιτικές caching στα ΑΔΑ έχει αναδειχθεί σε διάφορες μελέτες [40, 98] στην πρόσφατη βιβλιογραφία, παρ' όλα αυτά

δεν έχει προταθεί κάποια προηγμένη πολιτική. Ωστόσο, στα πλαίσια του συγγενούς ερευνητικού πεδίου των MANET έχουν προταθεί αρκετά πρωτόκολλα caching. Όμως, το κυριότερο κίνητρο για την ανάπτυξη των πρωτοκόλλων στα MANET είναι η κινητικότητα των κόμβων, και συνεπώς όλες οι ερευνητικές συνεισφορές προσπαθούν είτε να διαμορφώσουν είτε να αξιοποιήσουν την κίνηση.

Το πρωτόκολλο GroCoca [34] προσπαθεί να ομαδοποιήσει τους κόμβους με βάση το πρότυπο αίτησης των δεδομένων και το πρότυπο της κίνησής τους. Τα πρωτόκολλα ECOR [117], Zone Cooperative [24] και Cluster Cooperative [23] προσπαθούν να ομαδοποιήσουν τους κόμβους με βάση τη γεωγραφική εγγύτητα ή χρησιμοποιώντας αλγόριθμους ομαδοποίησης οι οποίοι αναπτύχθηκαν για MANET [12]. Αυτό συμβαίνει, επειδή η επικοινωνία με τους κόμβους που βρίσκονται στη γειτονιά είναι χαμηλού κόστους τόσο ως προς την κατανάλωση ενέργειας όσο και ως προς τον αριθμό των ανταλλασσόμενων μηνυμάτων, ενώ επιπλέον η ανάκτηση των δεδομένων από γειτονικούς κόμβους συνεπάγεται πολύ μικρότερη καθυστέρηση από ότι η ανάκτηση από τους κόμβους πηγές. Στο πρωτόκολλο Zone Cooperative (ZC) οι κινητοί κόμβοι του δικτύου που ανήκουν στη γειτονιά (ζώνη) ενός συγκεκριμένου κόμβου σχηματίζουν ένα σύστημα συνεργατικής cache για τον κόμβο. Σε περίπτωση αποτυχίας εύρεσης του ζητούμενου δεδομένου στην τοπική cache, ο κόμβος αρχικά κάνει αναζήτηση στους γειτονικούς του κόμβους που σχηματίζουν τη ζώνη. Αν η αίτηση για το δεδομένο δεν ικανοποιηθεί από αυτούς, τότε η αίτηση προωθείται στον επόμενο κόμβο που βρίσκεται στο μονοπάτι προς την πηγή των δεδομένων. Στο πρωτόκολλο Cluster Cooperative (CC) το κινητό τυχαίο δίκτυο διαιρείται σε ισομεγέθεις ομάδες με κριτήριο τη γεωγραφική τοποθεσία. Σε κάθε ομάδα, το πρωτόκολλο επιλέγει έναν κόμβο ως κόμβο κατάστασης της cache. Ο επιλεγμένος κόμβος αναλαμβάνει τη διατήρηση της πληροφορίας που σχετίζεται με την cache κάθε κόμβου που βρίσκεται στην ομάδα. Έτσι κάθε φορά που ένας κόμβος αναζητά κάποια δεδομένα, απευθύνεται πρώτα στον κόμβο που συντονίζει την ομάδα και αν δεν μπορεί να εξυπηρετηθεί, τότε στέλνει την αίτηση στον κόμβο πηγή. Παρόμοιο με τα προηγούμενα δύο πρωτόκολλα είναι και το πρωτόκολλο ECOR, όπου κάθε κινητός κόμβος σχηματίζει μια ζώνη συνεργασίας με τους γείτονες. Στα πλαίσια λειτουργίας του ECOR όταν μια αίτηση για κάποιο δεδομένο καταφθάνει σε ένα κόμβο, αυτός αναζητά ένα αντίγραφο του δεδομένου στη ζώνη συνεργασίας του πριν προωθήσει την αίτηση στον κόμβο πηγή. Το πρωτόκολλο όμως παρουσιάζει κάποια μειονεκτήματα, όπως είναι ο μεγάλος αριθμός ανταλλασσόμενων μηνυμάτων και συνακόλουθα η κατανάλωση ενέργειας ιδιαίτερα σε περιπτώσεις όπου η πυκνότητα των κόμβων είναι μεγάλη και η ακτίνα της ζώνης συνεργασίας είναι επίσης

μεγάλη.

Τα μόνα πρωτόκολλα που παρεκκλίνουν από τις προαναφερθείσες προσεγγίσεις και προσπαθούν να αξιοποιήσουν την *τοπικότητα* (locality) τόσο των δεδομένων όσο και των κόμβων με ένα ομοιογενή τρόπο, περιγράφονται στη μελέτη [149] και είναι τα εξής: CachePath, CacheData, και HybridCache. Στο πρωτόκολλο CacheData, οι ενδιάμεσοι κόμβοι (οι κόμβοι μεταξύ του αιτούντος κόμβου και του κόμβου πηγή) αποθηκεύουν τα δεδομένα, ώστε να καταφέρουν να εξυπηρετήσουν μελλοντικές αιτήσεις για τα ίδια δεδομένα χωρίς να χρειασθεί να προωθήσουν τα μηνύματα προς τους κόμβους πηγές. Στο πρωτόκολλο CachePath, οι κινητοί κόμβοι αποθηκεύουν τη διαδρομή και τη χρησιμοποιούν για να ανακατευθύνουν μελλοντικές αιτήσεις προς τους γειτονικούς κόμβους που έχουν τα δεδομένα αντί να στείλουν τις αιτήσεις προς τις απομακρυσμένες πηγές. Ένα υψηλής απόδοσης αμάλγαμα των δύο πρωτοκόλλων είναι το πρωτόκολλο HybridCache που μπορεί να βελτιώσει την απόδοση λαμβάνοντας υπόψη τα πλεονεκτήματα των δύο πρωτοκόλλων, ενώ ταυτόχρονα αποφεύγει τις αδυναμίες τους.

Στα Κεφάλαια 3 και 5 παρουσιάζουμε δύο νέες τεχνικές συνεργατικού caching για ΑΔΑ και ΑΠΔΑ, οι οποίες προσπαθούν να μειώσουν την καθυστέρηση πρόσβασης, την κατανάλωση ενέργειας και το πλήθος μηνυμάτων που ταξιδεύουν στο δίκτυο. Ένα σημαντικό τμήμα των τεχνικών είναι η ανακάλυψη των σημαντικών κόμβων που θα αναλάβουν να συντονίσουν τις caches των γειτονικών κόμβων ώστε να μειωθεί η καθυστέρηση πρόσβασης των δεδομένων. Οι κόμβοι θα πρέπει να βρίσκονται σε κεντρικά σημεία του δικτύου ώστε να ανακαλύπτονται εύκολα και να μπορούν να εξυπηρετούν μεγάλο τμήμα του. Οι σημαντικοί κόμβοι αναλαμβάνουν το ρόλο του διαμεσολαβητή μεταξύ των κόμβων πηγών και των cache κόμβων. Επιπλέον, οι σημαντικοί κόμβοι αποθηκεύουν μεταδεδομένα που σχετίζονται με τα δεδομένα που έχουν αποθηκευμένα στην cache οι γειτονικοί κόμβοι, ώστε να εξυπηρετούν τα ερωτήματα για δεδομένα που τίθενται είτε από τους γειτονικούς κόμβους είτε από άλλους κόμβους μέσω των οποίων τα ερωτήματα τυγχάνει να διέρχονται. Ένα δεύτερο σημαντικό στοιχείο των τεχνικών είναι ότι αξιοποιούν την εναπομείνασα ενέργεια των κόμβων ώστε να διαμοιράζουν και συντονίζουν τα cached δεδομένα των κόμβων με ενεργειακά αποδοτικό τρόπο.

Επιπρόσθετα προτείνουμε μια νέα πολιτική εισδοχής των δεδομένων στην cache. Σε ένα συνεργατικό περιβάλλον η πολιτική εισδοχής στην cache πρέπει να λαμβάνει υπόψη την απόσταση από όπου προέρχεται το δεδομένο, ώστε να εκμεταλλεύεται αποδοτικά τις caches των γειτονικών κόμβων. Όσο πιο απομακρυσμένος είναι ο κόμβος που αποστέλλει το δεδομένο, τόσο πιθανότερο είναι

να τοποθετηθεί στην cache. Αυτό οφείλεται στο γεγονός ότι τα δεδομένα που ανακτώνται από γειτονικούς κόμβους έχουν μικρή καθυστέρηση πρόσβασης και μικρή επιβάρυνση στην επικοινωνία. Τέλος, χρησιμοποιούμε μια πολιτική αντικατάστασης της cache, η οποία ενεργοποιείται κάθε φορά που η cache είναι πλήρης και υπάρχει ανάγκη για τοποθέτηση ενός νέου αντικειμένου. Λαμβάνοντας υπόψη την πολιτική εισδοχής στην cache, η οποία δεν επιτρέπει να γίνονται cached δεδομένα που προέρχονται από γειτονικούς κόμβους, και για να βελτιωθεί η απόδοση του συστήματος η πολιτική αντικατάστασης ενεργοποιείται χωριστά από κάθε κόμβο αλλά συντονίζεται από τους γειτονικούς σημαντικούς κόμβους. Με τον τρόπο αυτό κάθε απομακρυσμένο αντικείμενο μπορεί να τοποθετηθεί σε ένα γειτονικό κόμβο εφόσον υπάρχει ελεύθερος χώρος και το αντικείμενο δεν έχει γίνει cached από κάποιο άλλο κόμβο. Από τα προηγούμενα αντιλαμβανόμαστε ότι πραγματοποιείται μια εικονική (virtual) ενοποίηση των cache των γειτονικών κόμβων.

Στη συνέχεια θα παρουσιασθούν οι σημαντικότερες πολιτικές αντικατάστασης, οι οποίες έχουν προταθεί στη βιβλιογραφία.

### 2.2.1 Πολιτικές αντικατάστασης

Κάθε πολιτική caching θα πρέπει να περιλαμβάνει έναν αποδοτικό αλγόριθμο αντικατάστασης [2, 7]. Η πολιτική αντικατάστασης είναι ζωτικής σημασίας για την ομαλή και βέλτιστη απόδοση της cache, καθώς καθορίζει ποια δεδομένα πρέπει να απομακρυνθούν προκειμένου να σχηματισθεί ο απαιτούμενος ελεύθερος χώρος στην cache ώστε να αποθηκευθούν τα νέα δεδομένα που έχουν ληφθεί.

Ο *LRU* [90, 115, 135, 138] απομακρύνει το αντικείμενο που προσπελάστηκε λιγότερο πρόσφατα. Αυτός είναι ο δημοφιλέστερος αλγόριθμος που χρησιμοποιείται σήμερα και εκμεταλλεύεται τη χρονική τοπικότητα των αναφορών. Είναι εύκολος στην υλοποίησή του, επειδή χρειάζεται να διατηρεί μόνο ένα σωρό (heap) με τα IDs των αντικειμένων της cache. Η χωρική του πολυπλοκότητά είναι  $O(n)$  (όπου  $n$  το πλήθος των αντικειμένων της cache), ενώ η χρονική είναι σταθερή  $O(1)$ . Αυτός είναι ο κύριος λόγος που χρησιμοποιείται σε πολλά εμπορικά προϊόντα. Μια γενίκευση του *LRU* είναι ο *LRU-K* [99], ο οποίος αντικαθιστά το αντικείμενο του οποίου η  $k$ -οστή αναφορά είναι μακρύτερα πίσω στο χρόνο. Ο *FIFO* αντικαθιστά το αντικείμενο που εισήχθη πρώτο στην cache, ενώ ο *LFU* [7, 138] εκείνο με το μικρότερο αριθμό αναφορών. Μια παραλλαγή του είναι ο *LFU-Aging* [110], που εξετάζει τον αριθμό αναφορών (frequency count) ενός αντικειμένου, αλλά και την ηλικία του (age) στην cache. Ο *LRU* και ο *LFU*

βελτιστοποιούν τη μετρική Byte Hit Ratio. Μια πολιτική κατάλληλη για τον Παγκόσμιο Ιστό είναι η *Size* [138], που αντικαθιστά το μεγαλύτερο αντικείμενο. Η *Size* στοχεύει στη βελτίωση της μετρικής Cache Hit Ratio, αφού ευνοεί τα μικρά αντικείμενα.

Ο *LRU-MIN* [138] είναι μια παραλλαγή του *LRU* που προσπαθεί να ελαττώσει τον αριθμό των αντικαθιστώμενων αντικειμένων. Έστω  $s_i$  το μέγεθος του αντικειμένου που πρέπει να στεγασθεί στην cache, αλλά δεν υπάρχει αρκετός διαθέσιμος χώρος. Εάν στην cache υπάρχουν αντικείμενα μεγέθους τουλάχιστον  $s_i$ , τότε ο *LRU-MIN* απομακρύνει το λιγότερο πρόσφατα προσπελασμένο αντικείμενο (least recently used). Εάν δεν υπάρχουν τέτοια αντικείμενα, τότε με βάση τον *LRU* αρχίζει να απομακρύνει τα αντικείμενα μεγέθους τουλάχιστον  $s_i/2$ , κατόπιν τα αντικείμενα μεγέθους  $s_i/4$ , κοκ, μέχρι να αδειάσει ο απαραίτητος χώρος.

Η πολιτική  $\text{LOG}_2(\text{size})$  [138] είναι μια άλλη πολιτική που βασίζεται σε κλειδί, η οποία χρησιμοποιεί το  $\log_2(\text{size})$  ως πρωτεύον κλειδί και το χρόνο της τελευταίας αναφοράς ως δευτερεύον κλειδί. Ο *Hyper-G* [138] είναι ένας ακόμη αλγόριθμος που χρησιμοποιεί τη συχνότητα αναφοράς ως πρωτεύον κλειδί, το πρόσφατο (recency) της αναφοράς ως δευτερεύον και το μέγεθος του αντικειμένου ως τριτεύον κλειδί.

Στην ουσία οι παραδοσιακές πολιτικές αντικατάστασης και οι πολιτικές που βασίζονται σε κλειδιά αποτυγχάνουν να ενοποιήσουν όλους τους παράγοντες σε μια τιμή που θα χαρακτηρίζει τη χρησιμότητα ενός αντικειμένου στην cache. Αυτό ακριβώς προσπαθούν να επιτύχουν οι πολιτικές που βασίζονται στη λειτουργία (function-based). Αναθέτουν μια *τιμή χρησιμότητας* (utility value) σε κάθε αντικείμενο, η οποία είναι συνάρτηση διαφόρων παραγόντων, όπως το πρόσφατο της αναφοράς, το μέγεθος, το κόστος ανάκτησης, κ.τ.λ.

Ο *GreedyDual-Size* [18] (*GD-Size*) είναι ένας αλγόριθμος βασισμένος στον *GreedyDual* [150], ο οποίος συνδυάζει το πρόσφατο της αναφοράς με το κόστος ανάκτησης  $c_i$  και το μέγεθος ενός αντικειμένου  $s_i$ . Η τιμή χρησιμότητας ενός αντικειμένου που εισέρχεται στην cache είναι:

$$UV_{GD-Size} = \frac{c_i}{s_i}$$

Όταν απαιτείται χώρος στην cache, τότε απομακρύνεται το αντικείμενο με τη μικρότερη  $UV_{GD-Size}$ . Μετά από κάθε αντικατάσταση, η  $UV_{GD-Size}$  όλων των αντικειμένων ελαττώνεται κατά ένα ποσό ίσο με τη  $UV_{GD-Size}$  του αντικειμένου που αντικαταστάθηκε. Μετά από κάθε αναφορά αντικειμένου η  $UV_{GD-Size}$  του παίρνει τιμή ίση με  $\frac{c_i}{s_i}$ .



Ο Size-Adjusted LRU [2] ενσωματώνει το κόστος ανάκτησης και το μέγεθος ενός αντικειμένου στον παραδοσιακό *LRU*, αλλά και ζητήματα συνέπειας της cache. Συνήθως, στα αντικείμενα ανατίθεται μια τιμή, ο *Χρόνος Ζωής* (Time-To-Live – TTL) από την πηγή τους ή η τιμή *Εκπνοής* (Expires), οι οποίες προσδιορίζουν το χρονικό διάστημα εγκυρότητας του συγκεκριμένου αντικείμενου. Μπορούμε να εκμεταλλευθούμε αυτές τις τιμές, για να αποφασίσουμε για τη φρεσκάδα του αντικειμένου και να ενσωματώσουμε αυτές τις παραμέτρους στην πολιτική αντικατάστασης.

Ο αλγόριθμος Least Normalized Cost Replacement for the Web with Updates [121] εκτός των άλλων λαμβάνει υπόψη και το κόστος για να διαπιστωθεί η εγκυρότητα του αντικειμένου. Στόχος του είναι να ελαχιστοποιήσει το χρόνο απόκρισης, και για το λόγο αυτό βελτιστοποιεί τη μετρική Delay Savings Ratio και όχι τη μετρική Hit Ratio. Ο αλγόριθμος λαμβάνει υπόψη το μέσο αριθμό αναφορών στο αντικείμενο, τη μέση καθυστέρηση πρόσβασης, το μέσο αριθμό ελέγχων εγκυρότητας και τη μέση καθυστέρηση για τον έλεγχο της εγκυρότητάς του.

Οι πολιτικές αντικατάστασης που αποδίδουν καλύτερα στο συνεργατικά caching των ΑΔΑ όσο και των MANET είναι οι εκείνες που βασίζονται σε λειτουργίες. Αυτό οφείλεται στο γεγονός ότι η χρησιμότητα ενός αντικειμένου μπορεί να υπολογισθεί καλύτερα με βάση το μέγεθος, το κόστος ανάκτησης, το ρυθμό προσπέλασης, το ρυθμό ενημέρωσης κ.α. Ένα σημαντικός περιορισμός που τίθεται στα ΑΔΑ είναι η περιορισμένη ισχύς της μπαταρίας και συνεπώς η πολιτική αντικατάστασης πρέπει να στοχεύει και στη μείωση της κατανάλωσης της ενέργειας. Έτσι η ανάκτηση ενός αντικειμένου από έναν απομακρυσμένο κόμβο έχει ως αποτέλεσμα τη διασπάθιση μεγαλύτερης ποσότητας ενέργειας από ότι από ένα γειτονικό. Επομένως, η ανάκτηση ενός αντικειμένου από έναν απομακρυσμένο κόμβο συνεπάγεται μεγαλύτερο πλήθος διαδιδόμενων μηνυμάτων. Επιπλέον, όσο λιγότερο ενημερώνεται ένα αντικείμενο, τόσο πιθανότερο είναι να παραμείνει στην cache καθώς συνεπάγεται μικρότερη επιβάρυνση στην επικοινωνία. Παράλληλα ο ρυθμός προσπέλασης καθορίζει τη δημοφιλία ενός αντικειμένου, ενώ προτιμάται το μέγεθος του αντικειμένου να είναι μικρό καθώς έτσι περισσότερα αντικείμενα μπορούν να φιλοξενηθούν στην cache.

### 2.3 Τεχνικές Διατήρησης Συνέπειας της Cache

Η συνέπεια ορίζεται ως ο βαθμός της διαφοράς μεταξύ του αυθεντικού δεδομένου που είναι αποθηκευμένο στην πηγή των δεδομένων και του cached αντιγράφου του [25, 75]. Μπορεί επίσης να ορισθεί ανάλογα με το πόσο παλιό είναι το δεδομένο, πόσες ενημερώσεις έχουν συμβεί στο δεδομένο ή ακόμα με το πόση πληροφορία έχει μεταβληθεί. Μια στρατηγική για τη διατήρηση της συνέπειας των δεδομένων που βρίσκονται στην cache σε ΑΔΑ μπορεί να επικεντρωθεί σε τρία σχεδιαστικά ζητήματα: το *επίπεδο συνέπειας* (consistency level), τον *έλεγχο συνέπειας* (consistency control) και τη *διατήρηση της κατάστασης της cache* (cache status maintenance).

Υπάρχουν τρία επίπεδα συνέπειας της cache: *ισχυρή (SC)*, *δέλτα (DC)* και *ασθενής (WC)*. Η ισχυρή συνέπεια ονομάζεται επίσης και συνέπεια της τελευταίας τιμής [17], όπου κάθε κόμβος πελάτης θα χρησιμοποιήσει τα cached δεδομένα μόνο εάν αυτά δεν έχουν ενημερωθεί ποτέ από τον κόμβο πηγή από τη στιγμή που έγιναν cached στον πελάτη. Το αντίθετο της ισχυρής συνέπειας είναι η ασθενής [25]. Σε ορισμένες εφαρμογές, ένας πελάτης δεν έχει πρόβλημα να χρησιμοποιήσει κάποια παλιά δεδομένα και συνεπώς μπορεί να χρησιμοποιηθεί ένα σχήμα αδύναμης συνέπειας. Ένα παράδειγμα από τον Παγκόσμιο Ιστό είναι όταν οι άνθρωποι διαβάζουν on-line νέα. Οι περισσότεροι άνθρωποι δεν ενδιαφέρονται εάν η ιστοσελίδα ενημερώνεται άμεσα. Αντίθετα σε κάποιες εφαρμογές απαιτείται η χρήση ισχυρής συνέπειας, όπως για παράδειγμα είναι οι εφαρμογές διακίνησης μετοχών.

Θεωρούμε ότι το  $V_i^t$  συμβολίζει την έκδοση ενός αντικειμένου  $D_i$  που βρίσκεται σε ένα κόμβο πηγή, ενώ το  $T_i^t$  συμβολίζει τη *χρονοσφραγίδα* (timestamp) του  $D_i$  στον κόμβο αισθητήρα  $SN_j$  τη χρονική στιγμή  $t$ . Ο αριθμός έκδοσης του αντικειμένου ισούται με 0 όταν το αντικείμενο δημιουργείται από τον κόμβο πηγή και αυξάνεται κατά ένα σε κάθε διαδοχική ενημέρωση. Εάν η έκδοση του  $D_i$  στον cache κόμβο είναι πάντα επίκαιρη και σύμφωνη με αυτή του κόμβου πηγή, τότε όταν φθάσει στον κόμβο  $SN_j$  ένα ερώτημα για το αντικείμενο  $D_i$  ικανοποιείται η ισχυρή συνέπεια (SC):  $\forall t, \forall j, T_i^t = V_i^t$ . Εάν κάθε ανάγνωση του αντικειμένου  $D_i$  δεν είναι ποτέ παλαιότερη από ένα χρονικό διάστημα  $\Delta$  σε σχέση με το αυθεντικό αντικείμενο που βρίσκεται στον κόμβο πηγή, τότε ικανοποιείται η  $\Delta$  συνέπεια (DC):  $\forall t, \forall j$ , υπάρχει  $\tau$ , όπου  $0 \leq \tau \leq \Delta$ , ώστε ο χρόνος παλαιότητας  $T_i^t = V_i^{t-\tau}$ . Εάν μια ανάγνωση του  $D_i$  δεν ανταποκρίνεται απαραίτητα στην πιο ενημερωμένη έκδοση του αυθεντικού αντικειμένου αλλά κάποια ορθή παλαιότερη έκδοση, τότε ικανοποιείται η ασθενής συνέπεια (WC):  $\forall t, \forall j$ , υπάρχει  $\tau$ , ώστε ο

χρόνος παλαιότητας  $T_i^t = V_i^{t-\tau}$ .

Ο έλεγχος συνέπειας μπορεί να αρχικοποιηθεί τόσο από τον κόμβο πηγή (source-initiated), όσο και από τους cache κόμβους (cache-initiated). Στην πρώτη προσέγγιση, τα μηνύματα μη εγκυρότητας εκπέμπονται σε όλους τους cache κόμβους έτσι ώστε να γνωστοποιηθεί η κατάσταση ενημέρωσης των αντικειμένων. Στη δεύτερη προσέγγιση, οι cache κόμβοι θα στείλουν ερωτήματα στους κόμβους πηγές (pull) προκειμένου να διαπιστώσουν εάν τα αντικείμενα είναι έγκυρα. Η πρώτη προσέγγιση έχει ως συνέπεια το φόρτο εργασίας για τη διατήρηση της συνέπειας των αντικειμένων να τον επωμίζεται ο κόμβος πηγή. Εναλλακτικά, θα μπορούσε να θεωρηθεί μία push (PS) λειτουργία, καθώς ο κόμβος πηγή ωθεί τα μηνύματα των ενημερώσεων στους cache κόμβους. Αντίθετα, η δεύτερη προσέγγιση είναι μία μορφή pull (PL) λειτουργίας, που απαιτεί από κάθε cache κόμβο να “τραβά” τα μηνύματα ενημερώσεων από τον κόμβο πηγή.

Ο κόμβος πηγή έχει τη δυνατότητα είτε να καταγράφει την κατάσταση των δικών του cache κόμβων είτε όχι. Στην *stateful* (SF) προσέγγιση, η πηγή καταγράφει ποιοι κόμβοι έχουν αποθηκεύσει ποιά αντικείμενα. Επομένως, όταν ο κόμβος πηγή αποθηκεύει δεδομένα σχετικά με το ιστορικό των προσπελάσεων και ενημερώσεων, τότε ονομάζεται *stateful* [135]. Όταν ο κόμβος πηγή επιθυμεί να επικοινωνήσει με τους cache κόμβους, τότε μπορεί να στείλει είτε unicast είτε multicast μηνύματα αντί να πραγματοποιήσει μετάδοση με εκπομπή. Αυτή η προσέγγιση είναι περισσότερο αποδοτική και μειώνει τη χρήση του εύρους ζώνης. Μειονέκτημα όμως αποτελεί το γεγονός ότι ο κόμβος πηγή πρέπει να αποθηκεύει την κατάσταση των caches για όλους τους cache κόμβους. Συνεπώς, η προσέγγιση δεν είναι κλιμακώσιμη σε μεγάλο αριθμό κόμβων. Η *stateless* (SL) προσέγγιση δεν απαιτεί από τον κόμβο πηγή να γνωρίζει την κατάσταση των cache κόμβων. Αντιθέτως, ο κόμβος πηγή αποθηκεύει το ιστορικό ενημερώσεων και σε περιοδικά διαστήματα “πλημμυρίζει” το δίκτυο με ακυρωτικά μηνύματα που περιλαμβάνουν πληροφορία σχετικά με τις ενημερώσεις των αντικειμένων. Η *stateless* προσέγγιση επιβαρύνει σημαντικά την επικοινωνία καθώς αυξάνεται δραματικά το πλήθος των διαχεόμενων μηνυμάτων.

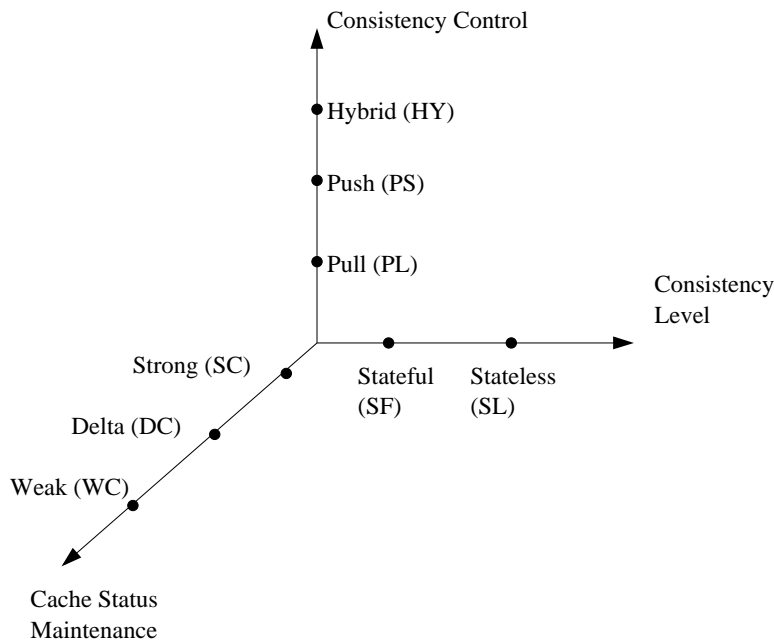
Στη συνέχεια, θα παρουσιάσουμε τις κυριότερες τεχνικές διατήρησης της συνέπειας της cache.

### 2.3.1 Υπόβαθρο

Οι τεχνικές διατήρησης της cache μπορούν να χωρισθούν ανάλογα με την ακολουθούμενη στρατηγική τους ως προς τα τρία σχεδιαστικά ζητήματα: επίπεδο

συνέπειας, έλεγχος συνέπειας και διατήρηση κατάστασης της cache [22]. Γενικότερα θα μπορούσαμε να τις κατηγοριοποιήσουμε σε push-based και pull-based.

Στο Σχήμα 2.3 παρουσιάζεται ένας τρισδιάστατος χώρος που παριστά κάθε σχεδιαστικό ζήτημα ως μία διάσταση. Μπορεί να δημιουργηθεί μια μεγάλη ποικιλία σχημάτων διατήρησης της συνέπειας της cache συνδυάζοντας τις ιδιότητες κάθε διάστασης. Αυτό επιτυγχάνεται επειδή οι τρεις διαστάσεις είναι ανεξάρτητες μεταξύ τους. Μία CS-CC-CL συμβολοσειρά αντιστοιχεί σε μία στρατηγική, όπου το CL παριστά το επίπεδο συνέπειας (SC, DC και WC), το CC τον έλεγχο συνέπειας (PS, PL και HY) και το CS τη διατήρηση της κατάστασης της cache (SF, SL και HY).



Σχήμα 2.3. Τρία σχεδιαστικά ζητήματα για τη διατήρηση της συνέπειας της cache.

Η απλούστερη pull-based προσέγγιση στα παραδοσιακά κατανομημένα συστήματα είναι η Poll-each-read [19, 148] και μπορεί να χαρακτηριστεί ως SL-PL-SC. Η τεχνική περιλαμβάνει την αποστολή ενός μηνύματος ελέγχου εγκυρότητας στον αντίστοιχο κόμβο πηγή κάθε φορά που το ζητά κάποιος πελάτης. Εάν το αντικείμενο είναι φρέσκο, τότε στέλνεται κατευθείαν στον πελάτη, αλλιώς ανακτάται από τον κόμβο πηγή και κατόπιν αποστέλλεται. Το κύριο πλεονέκτημα της τεχνικής είναι ότι η cache ποτέ δεν επιστρέφει στους πελάτες της μη φρέσκα δεδομένα, όμως έχει αρκετά μειονεκτήματα. Επιβάλλει μεγάλο φόρτο στους κόμβους πη-

γές εξαιτίας των πολλών μηνυμάτων ελέγχου εγκυρότητας, δημιουργεί μεγάλη δικτυακή κίνηση εξαιτίας των μηνυμάτων αυτών, αλλά και καθυστέρηση πρόσβασης για τους πελάτες, στην περίπτωση που τα αντικείμενα δεν έχουν τροποποιηθεί.

Η τεχνική διατήρησης της συνέπειας Poll-with-time-out-period [69] είναι ένα είδος SL-PL-DC στρατηγικής που θεωρεί ότι τα cached δεδομένα παραμένουν έγκαιρα για μία χρονική περίοδο τουλάχιστον  $t$ , αφού ο cache κόμβος έχει επιβεβαιώσει τα δεδομένα. Όταν  $t=0$ , τότε η στρατηγική είναι ισοδύναμη με την Poll-each-read. Όταν η χρονική περίοδος αυξάνεται, τότε αυξάνεται και η πιθανότητα οι cache κόμβοι να παρέχουν μη φρέσκα δεδομένα στους χρήστες, αλλά τα δεδομένα θα είναι το πολύ  $t$  χρονικό διάστημα παλαιά. Αντίθετα, επιλέγοντας μικρή τιμή για αυτό το διάστημα, μπορούμε να ελαττώσουμε το ποσοστό των μη φρέσκων αντικειμένων που προσφέρονται στους πελάτες, αλλά αυξάνουμε τον αριθμό των μηνυμάτων ελέγχου εγκυρότητας που ταξιδεύουν στο δίκτυο.

Στην άλλη άκρη του φάσματος των επιλογών, βρίσκεται η push-based τεχνική, όπου οι κόμβοι πηγές θα πρέπει να ενημερώνουν τις caches για όποιες αλλαγές συμβαίνουν στα δεδομένα. Αυτή η τεχνική λέγεται Invalidation ή Callback [100] και προφανώς απαιτεί από τους κόμβους πηγές να καταγράφουν ποιές caches στεγάζουν ποιά δεδομένα. Πριν ένας κόμβος πηγή μεταβάλλει τα αποθηκευμένα του δεδομένα θα πρέπει να ειδοποιήσει όλους τους cache κόμβους που έχουν αντίγραφά τους. Όταν συμβεί ένα cache hit, τότε ο cache κόμβος χρησιμοποιεί απευθείας το δικό του αντίγραφο και δεν χρειάζεται να επικοινωνήσει με τον κόμβο πηγή για να διαπιστώσει την εγκυρότητά του. Συνεπώς, η τεχνική χρησιμοποιεί τη stateful στρατηγική, την push στρατηγική για τον έλεγχο συνέπειας και εγγυάται ισχυρή συνέπεια (SC). Στα μειονεκτήματα της τεχνικής είναι ότι μπορεί να αυξηθούν πάρα πολύ οι πληροφορίες που πρέπει να συντηρούν οι κόμβοι πηγές, ειδικά όταν υπάρχουν πολλές caches και πολλά αντίγραφα των δεδομένων της πηγής στις caches αυτές. Επιπλέον, προκαλεί “καταιγίδες” (bursts) στη λειτουργία του κόμβου πηγή, όταν πρόκειται να διανείμει τα μηνύματα ακύρωσης.

Το Invalidation Report (IR) είναι ένα push-based σχήμα που υιοθετεί τη stateful στρατηγική και εγγυάται ισχυρή συνέπεια. Υπάρχουν δύο διαφορετικά είδη τέτοιων σχημάτων: το ασύγχρονο IR και το σύγχρονο IR [10]. Στο ασύγχρονο IR ο κόμβος πηγή εκπέμπει εκθέσεις ακυρότητας (invalidation reports) στους cache κόμβους μόλις συμβεί μια ενημέρωση. Μια έκθεση ακυρότητας περιέχει πληροφορίες ενημέρωσης για το αντικείμενο που μεταβλήθηκε και πληροφορίες ενημέρωσης για τα άλλα αντικείμενα που ενημερώθηκαν κατά τη διάρκεια της περιόδου πριν την τρέχουσα ενημέρωση. Η διαφορά μεταξύ του IR και του CB

είναι ότι το CB δεν περιλαμβάνει πληροφορίες ενημέρωσης για τα άλλα δεδομένα που ενημερώθηκαν. Το σύγχρονο IR εκπέμπει εκθέσεις ακυρότητας περιοδικά. Κάθε έκθεση ακυρότητας περιλαμβάνει πληροφορίες ενημέρωσης των δεδομένων που μεταβλήθηκαν στις προηγούμενες  $w$  περιόδους, όπου το  $w$  ονομάζεται παράθυρο εκπομπής έκθεσης ακυρότητας (invalidation report broadcasting window), ενώ κάθε περίοδος συμβολίζεται με  $L$ . Επομένως, κάθε IR περιέχει πληροφορίες για τις ενημερώσεις των δεδομένων το προηγούμενο  $w \times L$  χρονικό διάστημα.

Αμάλγαμα των τεχνικών pull-based και push-based είναι η τεχνική Leases [59]. Σύμφωνα με την τεχνική αυτή κάθε κόμβος πηγή πληροφορεί τον cache κόμβο για όποιες αλλαγές συμβαίνουν στα δεδομένα, αλλά μόνο για ένα σχετικά σύντομο και καλά καθορισμένο χρονικό διάστημα, ενώ μετά την παρέλευση αυτού του διαστήματος, οι cache κόμβοι είναι υπεύθυνοι για να εξακριβώσουν την εγκυρότητα των δεδομένων. Αυτή η μέθοδος προσπαθεί να συνδυάσει τα πλεονεκτήματα των δυο προαναφερθεισών τεχνικών κάνοντας προσεκτική ρύθμιση του διαστήματος για το οποίο ο κόμβος πηγή ενημερώνει τις caches για όποιες αλλαγές συμβαίνουν στα δεδομένα.

Συμπερασματικά οι push-based στρατηγικές είναι περισσότερο κατάλληλες για σταθερά δίκτυα. Αυτό συμβαίνει επειδή μπορούν να παρέχουν καλές εγγυήσεις συνέπειας στους χρήστες που είναι διαρκώς on-line και σε απόσταση που μπορεί να τους φθάσει ο κόμβος πηγή. Οι στρατηγικές αυτές παρουσιάζουν μικρή καθυστέρηση πρόσβασης των δεδομένων. Παρ' όλα αυτά οι push-based στρατηγικές υποφέρουν από περιττές μεταδόσεις ενημερώσεων και προκαλούν καταιγίδες στη λειτουργία του κόμβου πηγή, όταν πρόκειται να διανείμει τα μηνύματα ακύρωσης.

Αντίθετα, οι pull-based στρατηγικές είναι κατάλληλες για δυναμικά δίκτυα αλλά η πλημμυρίδα μηνυμάτων συνεπάγεται μεγάλη επιβάρυνση στην επικοινωνία. Παράλληλα, ο μεγάλος όγκος διαχόμενων μηνυμάτων συνεπάγεται τη διασπάθιση της ενέργειας των κόμβων με αποτέλεσμα τη μείωση του χρόνου ζωής του δικτύου. Το πλεονέκτημα της stateful στρατηγικής είναι ότι ο κόμβος πηγή αποφεύγει τη μη απαραίτητη πλημμυρίδα μηνυμάτων. Μειονέκτημα όμως αποτελεί το γεγονός ότι οι κόμβοι πηγές θα πρέπει να αποθηκεύουν μεγάλο όγκο πληροφορίας σχετικά με τα δεδομένα που έχουν οι cache κόμβοι. Η stateless προσέγγιση είναι εύκολη στην υλοποίηση και διαχείριση αλλά συνεπάγεται τη μετάδοση πολλών μη απαραίτητων μηνυμάτων με αποτέλεσμα τη μείωση της ισχύος της μπαταρίας των κόμβων.

Επομένως, μια υβριδική πολιτική μεταξύ των δύο θα είναι η ιδανική λύση. Οι μελέτες [21, 118] προτείνουν μια τέτοια υβριδική πολιτική (push-pull-based) για MANET. Το σχήμα Proximity Regions for Caching in Cooperative MP2P

(PReCinCt) διαιρεί την τοπολογία του δικτύου σε γεωγραφικές περιοχές [118]. Κάθε γεωγραφική περιοχή είναι υπεύθυνη για ένα σύνολο κλειδιών που απεικονίζουν δεδομένα, ενώ κάθε κλειδί αντιστοιχίζεται σε μια περιοχή με βάση μια γεωγραφική συνάρτηση κατακερματισμού. Το PReCinCt χρησιμοποιεί ένα υβριδικό μηχανισμό push-pull για να διατηρήσει τη συνέπεια των δεδομένων μεταξύ των αντιγράφων που βρίσκονται μέσα στο δίκτυο. Παρ' όλα αυτά η χρησιμοποιούμενη τεχνική της πλημμυρίδας έχει ως αποτέλεσμα τη μεγάλη αύξηση των μηνυμάτων που ανταλλάσσονται στο δίκτυο και την ταχεία κατανάλωση της ενέργειας, ιδιαίτερα όταν τα όρια κάθε περιοχής είναι μεγάλα και το δίκτυο πυκνό. Το επιπλέον κόστος επικοινωνίας δημιουργείται επίσης όταν η τοποθεσία κάθε περιοχής θα πρέπει γνωστοποιηθεί σε ολόκληρο το δίκτυο. Παρόμοια, το σχήμα Relay Peer-based Caching Consistency (RPCC) αξιοποιεί τους ενδιάμεσους κόμβους που ονομάζονται κόμβοι αναμεταδότες (relay nodes) για να υλοποιήσει την υβριδική push-pull πολιτική [21]. Με τη βοήθεια των κόμβων αναμεταδοτών που βρίσκονται μεταξύ του κόμβου πηγή και των cache κόμβων, μπορούν να εφαρμοσθούν η push-based και η pull-based στρατηγική, γεγονός που έχει ως αποτέλεσμα τη μείωση του κόστους επικοινωνίας. Ο κόμβος πηγή ωθεί τα δεδομένα στους κόμβους αναμεταδότες και οι cache κόμβοι ανακτούν τα δεδομένα από τους κόμβους αναμεταδότες. Αυτές οι δύο λειτουργίες μπορούν να γίνουν ασύγχρονα και ταυτόχρονα.

Ωστόσο, οι προαναφερθείσες πολιτικές δεν λαμβάνουν υπόψη το κριτήριο της ελαχιστοποίησης της καθυστέρησης πρόσβασης των δεδομένων, η οποία είναι ζωτικής σημασίας για τα ΑΠΔΑ. Επιπλέον, και οι δύο πολιτικές κατά τη σχεδίαση της στρατηγικής συνέπειας της cache λαμβάνουν υπόψη την κίνηση που εκτελούν οι κόμβοι. Στα πλαίσια όμως των ΑΠΔΑ θεωρούμε ότι οι κόμβοι είναι στατικοί. Ο στόχος μας είναι η ανάπτυξη ενός push-pull based πρωτοκόλλου διατήρησης της συνέπειας της cache ώστε να επωφεληθούμε από τα θετικά τόσο της push-based στρατηγικής όσο και της pull-based στρατηγικής. Το πρωτόκολλο που προτείνουμε στο Κεφάλαιο 4 στοχεύει στη μείωση της επιβάρυνσης της επικοινωνίας, της καθυστέρησης πρόσβασης των δεδομένων και της κατανάλωσης ενέργειας.

## 2.4 Πρωτόκολλα Ομαδοποίησης

Η τεχνική της ομαδοποίησης έχει διερευνηθεί ευρέως στα πλαίσια των MANET [6, 9, 11, 12, 27, 30, 57, 89, 106, 108, 112, 124, 140, 152]. Κατά την ομαδοποίηση οι κόμβοι του δικτύου διαιρούνται σε διαφορετικές ομάδες σύμφωνα με

κάποιους κανόνες. Ένας κόμβος που ανήκει σε μια ομάδα μπορεί να βρίσκεται σε διαφορετική κατάσταση ή λειτουργία, όπως *κεφαλή ομάδας* (cluster head – CH), *πύλη ομάδας* (clustergateway) ή *μέλος ομάδας* (cluster member). Η διαδικασία ομαδοποίησης περιλαμβάνει την εκλογή ενός κόμβου ως κεφαλής σε κάθε ομάδα, ώστε να συντονίσει τους κόμβους. Η κεφαλή ομάδας είναι υπεύθυνη για να λαμβάνει τις μετρημένες τιμές από τους κόμβους της ομάδας, να τις συναθροίζει και να στέλνει την τελική πληροφορία μέσω άλλων κεφαλών ομάδας στους sink κόμβους.

Μία δομή ομάδας, ως ένας αποδοτικός τρόπος ελέγχου της τοπολογίας παρέχει πολλά οφέλη [108]. Αρχικά, μια δομή ομάδας διευκολύνει την επαναχρησιμοποίηση των πόρων για να αυξηθεί η απόδοση του συστήματος [89]. Δύο ομάδες μπορούν να χρησιμοποιήσουν την ίδια συχνότητα ή code set, εάν οι ομάδες δεν είναι γειτονικές. Επιπρόσθετα, μια ομάδα μπορεί να οργανώσει και να συντονίσει καλύτερα τις μεταδόσεις γεγονότων με τη βοήθεια ενός ειδικού κόμβου, όπως είναι η κεφαλή της ομάδας. Αυτή η τεχνική μπορεί να εξοικονομήσει αρκετούς πόρους που χρησιμοποιούνται για τις αναμεταδόσεις μηνυμάτων και επομένως να μειωθεί ο αριθμός των συγκρούσεων στις μεταδόσεις [76, 141]. Ένα ακόμα όφελος σχετίζεται με τη δρομολόγηση. Οι κεφαλές και οι πύλες ομάδων μπορούν να σχηματίσουν ένα εικονικό επικαλυπτόμενο δίκτυο κορμού και έτσι η δημιουργία και η εξάπλωση της πληροφορίας δρομολόγησης μπορεί να περιοριστεί μόνο σε αυτό το σύνολο κόμβων. Επομένως, η τεχνική της ομαδοποίησης μπορεί να συντηρήσει το εύρος ζώνης επειδή περιορίζει το πλαίσιο των αλληλεπιδράσεων μεταξύ των ομάδων στους κόμβους κεφαλές και μειώνει τις ανταλλαγές μηνυμάτων μεταξύ των κόμβων αισθητήρων [153]. Οι κόμβοι αισθητήρες θα νοιάζονται μόνο για τη σύνδεσή τους με τους κόμβους κεφαλές και δεν θα επηρεάζονται από τις αλλαγές στο επίπεδο διασύνδεσης των κεφαλών των ομάδων [68]. Επιπρόσθετα, στο επίπεδο του κόμβου κεφαλή μπορούν να υλοποιηθούν στρατηγικές διαχείρισης έτσι ώστε να βελτιωθεί περαιτέρω η λειτουργία του δικτύου και να επιμηκυνθεί ο χρόνος ζωής της μπαταρίας κάθε αισθητήρα αλλά και ολόκληρου του δικτύου [153]. Ο κόμβος κεφαλή μπορεί να προγραμματίζει τις δραστηριότητες μέσα στην ομάδα ώστε οι κόμβοι να εναλλάσσονται μεταξύ της κατάστασης “ύπνου” (κατάσταση χαμηλής κατανάλωσης ενέργειας) και της ενεργούς κατάστασης. Με τον τρόπο αυτό μπορεί να επιτευχθεί σημαντική μείωση στην κατανάλωση της ενέργειας. Ένα άλλο πλεονέκτημα που απορρέει από την τεχνική της ομαδοποίησης και την εκλογή της κεφαλής της ομάδας είναι ότι κάθε κόμβος κεφαλή μπορεί να συναθροίζει τα δεδομένα που συλλέγει από τους αισθητήρες της ομάδας και συνεπώς να μειώνει τον αριθμό των μηνυμάτων που αναμεταδίδονται



στο δίκτυο [38]. Τέλος η δομή της ομάδας δημιουργεί την εντύπωση ενός πολύ μικρότερου και σταθερότερου δικτύου.

Οι περισσότερες εργασίες στα πλαίσια των MANET στόχευαν στην επιλογή των καταλληλότερων κόμβων ως κεφαλές ομάδων και αυτό επετεύχθη με τη χρήση της έννοιας των κυρίαρχων συνόλων (dominating sets – DS). Πιο συγκεκριμένα, κυρίαρχο σύνολο είναι ένα υποσύνολο των κόμβων του δικτύου τέτοιο ώστε κάθε κόμβος του δικτυακού γράφου είτε να ανήκει στο κυρίαρχο σύνολο, είτε να είναι γείτονας ενός κόμβου που ανήκει στο σύνολο αυτό. Μια έρευνα των μεθόδων αυτού του τύπου μπορεί να βρεθεί στη μελέτη [12]. Το κύριο μειονέκτημα των αλγορίθμων είναι το γεγονός ότι οι κόμβοι που ανήκουν στο κυρίαρχο σύνολο είναι υπεύθυνοι για τη διευθέτηση όλης της επικοινωνίας και επομένως μπορούν να διασπαθίσουν την ενέργεια τους πολύ σύντομα. Για να αντιμετωπισθεί το πρόβλημα του υπερβολικού αριθμού εκπομπών, κάποιες μέθοδοι αξιοποιούν τα αναγνωριστικά των κόμβων όπως για παράδειγμα οι μελέτες [6, 11, 27, 124, 140]. Παρ' όλα αυτά, οι προσεγγίσεις αυτές παρουσιάζουν το μειονέκτημα της μη εύρεσης όλων των πλεονασματικών εκπομπών. Αυτό οφείλεται στο γεγονός ότι το εμποδίζει η ρύθμιση με βάση το αναγνωριστικό ή το βάρος του κόμβου. Επιπρόσθετα, ορισμένες μέθοδοι, όπως αυτές που περιγράφονται στις μελέτες [9, 57, 89, 106, 124] δεν αξιοποιούν πλήρως την υπάρχουσα πληροφορία. Για παράδειγμα, η χρήση του βαθμού ενός κόμβου ως προβάδισμα όταν αποφασίζεται εάν θα επιλεγεί ως κεφαλή ομάδας μπορεί να μην έχει ως αποτέλεσμα την καλύτερη τοπική απόφαση. Τελικά, κάποιες μέθοδοι δημιουργούν μεγάλο αριθμό ομάδων [140], ή απαιτούν υπερβολικό κόστος επικοινωνίας [6]. Στη συνέχεια, θα εξετάσουμε συνοπτικά κάποιες από τις μεθόδους που αναπτύχθηκαν στα πλαίσια των MANET.

Οι Baker και Ephremides [9] πρότειναν έναν αλγόριθμο που ονόμασαν “το υψηλότερο αναγνωριστικό κορυφής” (highest vertex ID). Στον αλγόριθμο αυτό, κάθε κορυφή σαρώνει το σύνολο των γειτόνων και επιλέγει ως κεφαλή ομάδας το γείτονα με το υψηλότερο αναγνωριστικό. Ένας παρόμοιος αλγόριθμος είναι ο αλγόριθμος του χαμηλότερου αναγνωριστικού που πρότειναν οι Gerla και Tsai [57], όπου κάθε κορυφή με το χαμηλότερο αναγνωριστικό στη γειτονιά επιλέγεται ως κεφαλή ομάδας. Οι Gerla and Tsai ανέπτυξαν ακόμα ένα αλγόριθμο που ονόμασαν “αλγόριθμο του υψηλότερου βαθμού” σύμφωνα με τον οποίο η κορυφή με το μεγαλύτερο βαθμό επιλέγεται ως κεφαλή ομάδας.

Οι Das και Bharghavan [35, 36] παρείχαν κατανεμημένες υλοποιήσεις των αλγορίθμων των Ghua και Khuller [61]. Οι Wu και Li [139, 140] βελτίωσαν τον κατανεμημένο αλγόριθμο των Das και Bharghavan με ένα τοπικό κατανεμημένο αλγόριθμο για συνδεδεμένα κυρίαρχα σύνολα, όπου κάθε κόμβος χρειάζεται να

γνωρίζει μόνο την 2-hop γειτονιά του. Αρχικά κάθε κόμβος σημειώνει τον εαυτό του ως F δείχνοντας έτσι ότι δεν έχει ακόμα καλυφθεί. Στην πρώτη φάση, ένας κόμβος σημειώνει τον εαυτό του ως T εάν υπάρχουν δύο μη άμεσα συνδεδεμένοι γείτονες. Στη δεύτερη φάση ένας κόμβος που έχει σημειώσει τον εαυτό του ως T μετατρέπει την ένδειξη σε F εάν ισχύει κάποιος από τους εξής κανόνες. Σύμφωνα με το πρώτο κανόνα, ένας κόμβος σημειώνει τον εαυτό του ως F όταν το κλειστό σύνολο γειτόνων του (περιλαμβάνει όλους τους άμεσους γείτονες και τον εαυτό του) συμπεριλαμβάνεται εξολοκλήρου στο σύνολο γειτόνων ενός κυρίαρχου γειτονικού κόμβου και έχει μικρότερο αναγνωριστικό (ID) από το γειτονικό κόμβο. Σύμφωνα με το δεύτερο κανόνα, ένας κόμβος διαγράφει τον εαυτό του από το CDS όταν το ανοιχτό σύνολο γειτόνων του (περιλαμβάνει όλους τους άμεσους γείτονές του) συμπεριλαμβάνεται εξολοκλήρου στα σύνολα γειτόνων δύο κυρίαρχων συνδεδεμένων γειτονικών κόμβων και έχει το μικρότερο αναγνωριστικό. Ο Stojmenovic [124] πρότεινε έναν αλγόριθμο για τη βελτίωση της απόδοσης του πρωτοκόλλου των Wu και Li, σύμφωνα με τον οποίο οι κόμβοι ταξινομούνται ως εξής. Ένας κόμβος ονομάζεται ενδιάμεσος εάν υπάρχουν δύο γειτονικοί κόμβοι που δεν συνδέονται άμεσα. Δια-πύλη (intergateway) ονομάζεται ένας κόμβος που δεν διαγράφεται από τους κυρίαρχους κόμβους ύστερα από την εφαρμογή του Κανόνα 1 του πρωτοκόλλου των Wu και Li, ενώ πύλη (gateway) ονομάζεται ένας κόμβος που δεν διαγράφεται ύστερα από την εφαρμογή και του Κανόνα 2.

Όμως στα ΑΔΑ υπάρχει ένας αυστηρός περιορισμός σχετικά με την κατανάλωση της ενέργειας καθώς οι μπαταρίες δεν μπορούν να επαναφορτισθούν. Επομένως, η σχεδίαση αλγορίθμων ομαδοποίησης πρέπει να λαμβάνει υπόψη την ενέργεια των κόμβων και να παρέχει μηχανισμούς για τη διευθέτηση του προβλήματος της κατανάλωσης ενέργειας λόγω των επαναλαμβανόμενων μεταδόσεων από τους ίδιους κόμβους. Η κατανάλωση ενέργειας είναι ένας παράγοντας που είχε μελετηθεί στα MANET, όπως για παράδειγμα στο πρωτόκολλο SPAN [29] αλλά εξετάζεται εκτενώς στα ΑΔΑ, όπως για παράδειγμα στο πρωτόκολλο LEACH [66], στο πρωτόκολλο HEAD [151], στο πρωτόκολλο EECS [145] κ.α. (δες τις μελέτες [1, 152] για ολοκληρωμένες έρευνες). Η οικογένεια αυτή πρωτοκόλλων προτείνει τρόπους εναλλαγής του ρόλου της κεφαλής ομάδας μεταξύ των κόμβων της ομάδας. Επιπλέον, τα προτεινόμενα πρωτόκολλα χρησιμοποιούν την εναπομείνουσα ενέργεια κάθε κόμβου για να προσδιορίσουν την κεφαλή ομάδας. Παρ' όλα αυτά, αυτή η οικογένεια μεθόδων αγνοεί τα τοπολογικά χαρακτηριστικά του δικτύου.

Στη μελέτη [66] προτείνεται ένας κατανεμημένος αλγόριθμος ομαδοποίησης, όπου οι κόμβοι αισθητήρες εκλέγουν τους εαυτούς τους ως κεφαλές με κάποια

πιθανότητα και δημοσιοποιούν στη συνέχεια την απόφασή τους. Οι απομένοντες κόμβοι εισέρχονται σε μια ομάδα με κριτήριο το κόστος επικοινωνίας με την κεφαλή της ομάδας. Έπειτα ο ρόλος της κεφαλής ομάδας εναλλάσσεται περιοδικά μεταξύ των κόμβων, έτσι ώστε να εξισορροπηθεί η κατανάλωση της ενέργειας. Στη μελέτη [151], οι συγγραφείς προτείνουν έναν κατανεμημένο αλγόριθμο ομαδοποίησης, ο οποίος προσπαθεί να συγκεράσει το ενεργειακό και το επικοινωνιακό κόστος. Σε αυτόν τον αλγόριθμο γίνεται προσπάθεια για επιμήκυνση του χρόνου ζωής του δικτύου μέσω της κατανεμημένης κατανάλωσης της ενέργειας. Επιπρόσθετα ο αλγόριθμος τερματίζει τη διαδικασία ομαδοποίησης σε ένα σταθερό αριθμό επαναλήψεων. Ο αλγόριθμος HEED υποθέτει ότι όλοι οι κόμβοι είναι εξίσου σημαντικοί και ότι η κατανάλωση της ενέργειας δεν είναι ομοιόμορφη μεταξύ των κόμβων. Ο αλγόριθμος δεν κάνει υποθέσεις σχετικά με την πυκνότητα των κόμβων ή τις δυνατότητές τους.

Στο Κεφάλαιο 6 θα παρουσιάσουμε ένα δυναμικό ενεργειακά αποδοτικό κατανεμημένο πρωτόκολλο ομαδοποίησης για ΑΔΑ, το οποίο ονομάζεται *GESC*. Το πρωτόκολλο *GESC* λαμβάνει υπόψη τόσο την εναπομείνασα ενέργεια όσο και τα τοπολογικά στοιχεία του δικτύου σε αντίθεση με τα προηγούμενα πρωτόκολλα που ελάμβαναν υπόψη μόνο τον ένα από τους δύο παράγοντες. Το προτεινόμενο πρωτόκολλο αξιοποιεί την τοπική δομή του δικτύου και την εναπομείνασα ενέργεια των γειτονικών κόμβων για να ορίσει ένα νέο τρόπο δυναμικού υπολογισμού των κεφαλών ομάδων.

## 2.5 Κατάλογοι σε Ασύρματα Δίκτυα

Στα ασύρματα δίκτυα υπεισέρχονται δύο αντιμαχόμενοι στόχοι. Ο πρώτος στόχος επικεντρώνεται στη μείωση της κατανάλωσης ενέργειας των κινητών πελατών, ενώ ο δεύτερος αποσκοπεί στη μείωση της καθυστέρησης πρόσβασης των δεδομένων. Οι κόμβοι των ΑΔΑ είναι φτωχοί σε πόρους με περιορισμένες δυνατότητες και βασίζονται στη μετάδοση με εκπομπή για να ανακτήσουν τις απαιτούμενες πληροφορίες. Για να αντιμετωπισθεί το μειονέκτημα της μεγάλης κατανάλωσης ενέργειας εξαιτίας της ακολουθιακής αναζήτησης των δεδομένων σε ασύρματα δίκτυα που βασίζονται στη μετάδοση με εκπομπή χωρίς παράλληλα να επιβαρύνεται η καθυστέρηση πρόσβασης των δεδομένων, έχει προταθεί η παρεμβολή ενός καταλόγου.

Ας θεωρήσουμε το ακόλουθο σενάριο που είναι εμπνευσμένο από τα σύγχρονα αυτοματοποιημένα πεδία μάχης, όπου αναπτύσσεται ένα *ad hoc* δίκτυο. Ένας

ισχυρός ενεργειακά κόμβος έχει σχεδιασθεί να παρέχει πληροφορίες σχετικά με κινήσεις τακτικής στα φίλια στρατεύματα. Ο κόμβος αυτός εκπέμπει αναφορές σε μικρές κινητές μονάδες περιορισμένης ενέργειας και τις οποίες φέρουν για παράδειγμα είτε απλοί (ή ομάδες από) στρατιώτες, είτε ανώνυμα οχήματα εξερεύνησης. Προφανώς, όλοι η εκπεμπόμενη πληροφορία δεν είναι δυνατόν να ενδιαφέρει όλες τις μονάδες που λαμβάνουν τις εκπομπές. Επίσης υπάρχει η περίπτωση όπου το “περιγραφικό κλειδί” των τμημάτων της πληροφορίας δεν επιβάλλει κάποια συνολική διάταξη μεταξύ των τμημάτων της πληροφορίας.

Στο προηγούμενο σενάριο οι καταναλωτές της πληροφορίας χρειάζεται να ανακτήσουν τα δεδομένα όσο το δυνατόν γρηγορότερα, με μικρή καθυστέρηση πρόσβασης. Επιπλέον, οι καταναλωτές είναι κόμβοι με περιορισμένη ενέργεια. Επομένως, πρέπει να αποφεύγουν τη διαρκή παρακολούθηση του ασύρματου καναλιού επειδή η παρακολούθηση του καναλιού επιβάλλει στον κόμβο την κατανάλωση ενέργειας. Επιπλέον γίνεται αντιληπτό ότι δεν υπάρχει συνολική αλλά μόνο μερική διάταξη των δεδομένων. Τέλος, το πρότυπο προσπέλασης στην εκπεμπόμενη πληροφορία εμφανίζει κυρτότητα. Αυτό σημαίνει ότι κάποια δεδομένα παρουσιάζουν μεγαλύτερο ενδιαφέρον και ζήτηση από κάποια άλλα.

Είναι προφανές ότι όταν μεταδίδονται δεδομένα στο ασύρματο κανάλι, ο κόμβος δεν έχει εναλλακτική από το να παρακολουθεί συνεχώς το ασύρματο μέσο μέχρι να λάβει τα δεδομένα. Η διαρκής παρακολούθηση όμως του καναλιού σημαίνει μεγάλη κατανάλωση ενέργειας. Για να επιτρέψουμε στους κόμβους να βρίσκονται σε μια κατάσταση ύπνου και να μη σπαταλούν ενέργεια, πρέπει με κάποιο τρόπο να ενσωματώσουμε καταλόγους μέσα στα δεδομένα εκπομπής ή πριν από αυτά. Χρησιμοποιώντας τον *εναέριο κατάλογο* (air indexing) [74], ο οποίος περιλαμβάνει πληροφορία σχετικά με το χρόνο άφιξης των δεδομένων, οι κόμβοι δεν χρειάζεται να παρακολουθούν διαρκώς το κανάλι. Αντίθετα, μπορούν να “ξυπνούν” σε συγκεκριμένες χρονικές στιγμές όταν τα ζητούμενα δεδομένα πρόκειται να μεταδοθούν.

Μια προσαρμογή της ιδέας των  $B^+$ -δένδρων σε ασύρματα περιβάλλοντα περιγράφηκε πρώτη φορά στη μελέτη [74], όπου τα φύλλα του  $B^+$ -δένδρου αποθηκεύουν το χρόνο άφιξης κάθε δεδομένου στο ασύρματο κανάλι και όχι διευθύνσεις στο δίσκο. Παρομοίως, μια προσαρμογή της παραδοσιακής τεχνικής κατακερματισμού σε ασύρματα περιβάλλοντα παρουσιάστηκε νωρίτερα στη μελέτη [73], ενώ αργότερα μια γενίκευση περιγράφηκε στη μελέτη [142]. Υβριδικές τεχνικές μεταξύ των δύο προσεγγίσεων έχουν περιγραφεί στις μελέτες [116, 144], ενώ εφαρμογές των δένδρων υπογραφών (signature trees) αναφέρονται στη μελέτη [85]. Οι μελέτες αυτές δεν προτείνουν νέα σχήματα αλλά διαφορετικές μεθόδους κατα-

νομής για τους κόμβους του δενδρικού καταλόγου. Σε όλες αυτές τις εργασίες έχει υποτεθεί πρώτον ότι υπάρχει μια συνολική διάταξη μεταξύ των δεδομένων που μεταδίδονται και δεύτερον ότι το πρότυπο προσπέλασης είναι ομοιόμορφο, γεγονός που σημαίνει ότι όλα τα δεδομένα προσπελούνται ισοπίθانا.

Παρεκκλίνοντας από την ομοιόμορφη πιθανότητα προσπέλασης, συναντούμε ποικίλες εργασίες που θεωρούν την επίδραση της κυρτότητας στην προσπέλαση έτσι ώστε να σχεδιάσουν αποδοτικά σχήματα καταλόγων. Στη μελέτη [119] προτάθηκε μια νέα εκδοχή του Αλφαβητικού Δένδρου  $k$  τάξης ( $k$ -ary alphabetic tree) [71], ενώ οι μελέτες [143, 146] προσαρμόζουν τον κατάλογο της μελέτης [74] για να διαχειριστεί τη μη ομοιομορφία στη προσπέλαση. Τέλος, όλες οι εργασίες [77, 92, 119, 143, 146] θεωρούν ότι υπάρχει μια συνολική διάταξη μεταξύ των μεταδιδόμενων δεδομένων. Όμως υπάρχει μόνο μία εργασία που παρεκκλίνει και από τις δύο υποθέσεις [31]. Παρ' όλα αυτά, γίνεται η υπόθεση της μη ύπαρξης διάταξης για τα δεδομένα και συνεπώς ο προτεινόμενος κατάλογος δεν αποτελεί δομή αναζήτησης.

Στο Κεφάλαιο 7 θα παρουσιάσουμε ένα νέο μη ισοζυγισμένο δενδρικό κατάλογο, που ονομάζουμε *Partial Ordering Broadcast Index* (POBI). Ο νέος κατάλογος μπορεί να διαπλέκεται ή να προηγείται των δεδομένων που εκπέμπονται έτσι ώστε να ελαχιστοποιηθεί η σπατάλη ενέργειας και ο καθυστέρηση πρόσβασης των δεδομένων. Η προτεινόμενη δομή έχει υλοποιηθεί για περιπτώσεις ρεαλιστικής κύρτωσης στο πρότυπο προσπέλασης των δεδομένων και όχι για μη ρεαλιστικά ομοιόμορφα πρότυπα, ενώ επιπλέον μπορεί να διαχειριστεί μερικώς διατεταγμένα δεδομένα κάτι που δεν είχε παρουσιασθεί σε καμία προηγούμενη έρευνα.

## 2.6 Συμπεράσματα

Στο κεφάλαιο αυτό παρουσιάσαμε και διασαφήνισαμε αρχικά βασικές έννοιες του περιβάλλοντος των AΔA, εξαιτίας των οποίων χρειαζόμαστε νέες λύσεις στο ζήτημα του συνεργατικού caching, τις πολιτικές αντικατάστασης, τη συνέπεια της cache, την ομαδοποίηση και τους εναέριους καταλόγους. Στη συνέχεια παρουσιάσαμε τις κυριότερες τεχνικές συνεργατικού caching που αναφέρονται στη βιβλιογραφία. Το caching των χρήσιμων δεδομένων για κάθε αισθητήρα είτε στον τοπικό αποθηκευτικό χώρο είτε στους γειτονικούς αισθητήρες μπορεί να επιμηκύνει το χρόνο ζωής ολόκληρου του δικτύου. Επιπρόσθετα, το caching μπορεί να είναι πολύ αποτελεσματικό στη μείωση των μεταδόσεων σε μεγάλη έκταση του δικτύου και επομένως στη μείωση των παρεμβολών και στην επιτυχή αντιμετώπιση

των συνθηκών μεταβλητού καναλιού. Οι τεχνικές όμως που έχουν προταθεί είτε δεν λαμβάνουν υπόψη τα ιδιαίτερα χαρακτηριστικά των ΑΔΑ, είτε δεν αξιοποιούν ταυτόχρονα και τα δεδομένα και την τοπικότητα ενός κόμβου. Στο πλαίσιο της έρευνας στις πολιτικές αντικατάστασης αναδείξαμε ότι η χρησιμότητα ενός αντικειμένου μπορεί να υπολογισθεί αποδοτικότερα με βάση το μέγεθος, το κόστος ανάκτησης, το ρυθμό προσπέλασης και το ρυθμό ενημέρωσης.

Στο πλαίσιο του ζητήματος της συνέπειας της cache μελετήσαμε τις τεχνικές που αναπτύχθηκαν στο πεδίο των κινητών τυχαίων δικτύων, των λειτουργικών συστημάτων και τον Παγκόσμιο Ιστό. Οι πολιτικές που υλοποιήθηκαν δεν λαμβάνουν υπόψη το κριτήριο της ταυτόχρονης μείωσης της καθυστέρησης προσπέλασης των δεδομένων και της κατανάλωσης ενέργειας των αισθητήρων, το οποίο είναι ζωτικής σημασίας για τα ΑΠΔΑ. Παράλληλα αναδείξαμε τη σπουδαιότητα της υβριδικής push/pull τεχνικής που πρέπει να διέπει τις πολιτικές συνέπειας της cache που αναπτύσσονται τόσο σε ΑΔΑ όσο και σε ΑΠΔΑ. Στη συνέχεια εξετάσαμε το ζήτημα της ομαδοποίησης σε ΑΔΑ, η οποία επιτρέπει τη συνάθροιση της πληροφορίας, περιορίζει τη μετάδοση των δεδομένων κυρίως μέσα στην ομάδα, μειώνει το συναγωνισμό για την απόκτηση του ασύρματου καναλιού και ελαττώνει την κατανάλωση ενέργειας. Η μελέτη των τεχνικών ομαδοποίησης που αναφέρονται στη βιβλιογραφία αναδεικνύει την ανάγκη για την ανάπτυξη μιας νέας πολιτικής η οποία θα λαμβάνει υπόψη τόσο την εναπομένουσα ενέργεια όσο και τα τοπολογικά χαρακτηριστικά του δικτύου σε αντίθεση με τις πολιτικές που αναπτύχθηκαν οι οποίες αξιοποιούσαν μόνο τον ένα από τους δύο παράγοντες. Τέλος στην Παράγραφο 2.5 μελετήσαμε το ζήτημα κατασκευής εναέριων ευρετηρίων σε Ασύρματα Δίκτυα. Από τη βιβλιογραφική έρευνα που διεξήχθη προκύπτει η ανάγκη για τη δημιουργία ενός νέου μη ισοζυγισμένου καταλόγου για μερικώς διατεταγμένα δεδομένα έτσι ώστε να μειωθεί η διασπάθιση ενέργειας και η καθυστέρηση προσπέλασης των δεδομένων.

## ΚΕΦΑΛΑΙΟ 3

# Συνεργατικό Caching σε Ασύρματα Πολυμεσικά Δίκτυα Αισθητήρων\*

### Περιεχόμενα

---

3.1	ΕΙΣΑΓΩΓΗ . . . . .	35
3.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	39
3.3	ΠΡΩΤΟΚΟΛΛΟ NiCoCa ΓΙΑ ΣΥΝΕΡΓΑΤΙΚΟ CACHING	45
3.4	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	55
3.5	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	67

---

### 3.1 Εισαγωγή

Τα ΑΔΑ έχουν εμφανισθεί στο προσκήνιο κατά τη διάρκεια των τελευταίων ετών λόγω των επιτευγμάτων στη σχεδίαση κυκλωμάτων χαμηλής κατανάλωσης ενέργειας και στην ανάπτυξη κατάλληλου λογισμικού, τα οποία επέτρεψαν τη δημιουργία μικροσκοπικών συσκευών που είναι ικανές να υπολογίζουν, να ελέγχουν και να επικοινωνούν μεταξύ τους [5, 78]. Ένα ΑΔΑ αποτελείται από

---

\*Το κεφάλαιο αντλεί τα περιεχόμενά του από τις εργασίες [43, 44].

ασύρματα διασυνδεδεμένες συσκευές που μπορούν να αλληλεπιδρούν με το περιβάλλον ελέγχοντας και ανιχνεύοντας φυσικές παραμέτρους. Τα ΑΔΑ προσέλκυσαν το αυξανόμενο ενδιαφέρον τόσο της επιστημονικής κοινότητας όσο και από της βιομηχανίας. Αυτό το ενδιαφέρον μπορεί να αποδοθεί στις πολλές νέες εντυπωσιακές εφαρμογές που δημιουργήθηκαν ως αποτέλεσμα της ανάπτυξης των μεγάλης κλίμακας ΑΔΑ. Ενδεικτικά αναφέρονται εφαρμογές όπως: ο έλεγχος του περιβάλλοντος, η βοήθεια σε περίπτωση θεομηνίας, η παρακολούθηση της ορθής λειτουργίας των μηχανών, η γεωργία ακρίβειας, το έξυπνο κτίριο, εφαρμογές υγείας και τηλεματικής κ.α. Τα χαρακτηριστικά που εμφανίζονται στα ΑΔΑ και οι ιδιαίτερες απαιτήσεις που πρέπει να ικανοποιούνται από τις εφαρμογές παρουσιάστηκαν με λεπτομέρεια στο Κεφάλαιο 2.

Η παραγωγή φθηνών CMOS καμερών και μικροφώνων, τα οποία μπορούν να αποκτήσουν πλούσιο πολυμεσικό περιεχόμενο από το περιβάλλον, δημιούργησε ένα νέο κύμα στην εξέλιξη των ΑΔΑ. Για παράδειγμα, το Cyclops είναι ένα ελαφρύ αποδοτικό λογισμικό [107], το οποίο μπορεί να προσαρμοσθεί σε αισθητήρες MICA2<sup>1</sup> ή MICAz. Έτσι, μια νέα τάξη ΑΔΑ ήρθε στη ζωή, τα ΑΠΔΑ [3]. Τα ΑΠΔΑ δεν ενισχύουν απλώς τις υπάρχουσες εφαρμογές των ΑΔΑ αλλά δημιουργούν και νέες όπως είναι: (α) τα ασύρματα πολυμεσικά δίκτυα παρακολούθησης τα οποία συντίθενται από μικροσκοπικές κάμερες [84] και μπορούν να επικοινωνούν, να επεξεργάζονται και να αποθηκεύουν δεδομένα σχετικά με εγκλήματα και τρομοκρατικές επιθέσεις, (β) η αποφυγή κυκλοφοριακής κίνησης και συστήματα ελέγχου τα οποία παρακολουθούν την κίνηση των αυτοκινήτων και προσφέρουν ταξιδιωτικές συμβουλές για να εμποδισθεί η συμφόρηση των αυτοκινήτων, (γ) ο έλεγχος της βιομηχανικής παραγωγικής διαδικασίας ο οποίος θα γίνει πραγματικότητα με τα ΑΠΔΑ που θα προσφέρουν χρονικά κρίσιμη (time critical) πληροφορία σχετιζόμενη με την πίεση, τη θερμοκρασία, κ.α.

Οι καινοτόμες εφαρμογές των ΑΠΔΑ προκάλεσαν το ενδιαφέρον της επιστημονικής κοινότητας επειδή απαιτείται η επανεξέταση του σχεδιαστικού μοντέλου υπολογισμού-επικοινωνίας, το οποίο ακολουθήθηκε στα παραδοσιακά ΑΔΑ [3]. Το μοντέλο αυτό επικεντρώθηκε κυρίως στη μείωση της κατανάλωσης ενέργειας αποσκοπώντας στην επιμήκυνση του χρόνου ζωής των δικτύων αισθητήρων. Ωστόσο, οι εφαρμογές που υλοποιούνται από τα ΑΠΔΑ έχουν ένα δεύτερο στόχο εξίσου σημαντικό με τη μείωση της κατανάλωσης ενέργειας. Σε επίπεδο εφαρμογής, στόχος είναι η παράδοση Ποιότητας Υπηρεσίας (Quality of Service – QoS) και η αντιστοίχιση αυτής της απαίτησης στις μετρικές του επιπέδου δικτύου (network

<sup>1</sup><http://www.xbow.com>



layer), όπως είναι η καθυστέρηση πρόσβασης. Αυτός ο στόχος είχε (σχεδόν) αγνοηθεί στις προηγούμενες ερευνητικές προσπάθειες στα παραδοσιακά ΑΔΑ.

Ο στόχος για Ποιότητα Υπηρεσίας στο διαδίκτυο σε ότι αφορά στην παράδοση πολυμεσικού περιεχομένου έχει ακολουθηθεί σε αρχιτεκτονικές όπως είναι η Diffserv και η Intserv, αλλά αυτά τα πρωτόκολλα και οι τεχνικές δεν αντιμετώπιζαν τους αυστηρούς περιορισμούς και το αφιλόξενο περιβάλλον των ΑΔΑ. Πιο συγκεκριμένα, τα ΑΠΔΑ χαρακτηρίζονται κυρίως από:

- **Περιορισμό πόρων:** οι αισθητήρες είναι συσκευές με πολύ περιορισμένη ενέργεια, υπολογιστική ισχύ και μνήμη.
- **Μεταβλητή χωρητικότητα καναλιού:** η multi-hop φύση των ΑΠΔΑ που λειτουργούν με βάση την τεχνική αποθήκευση-και-προώθηση (store-and-forward) εξαιτίας της απουσίας των σταθμών βάσης, συνεπάγεται ότι η χωρητικότητα κάθε ασύρματου συνδέσμου εξαρτάται από το επίπεδο παρεμβολής μεταξύ των κόμβων, το οποίο επιβαρύνεται από τις λειτουργίες εκπομπής.
- **Επεξεργασία πολυμεσικών δεδομένων μέσω διασύνδεσης:** σε πολλές εφαρμογές των ΑΠΔΑ, ένας μοναδικός αισθητήρας δεν είναι ικανός να ικανοποιήσει μία αίτηση, αλλά πρέπει να συνεργασθούν πολύ διαφορετικοί αισθητήρες για να την απαντήσουν. Για παράδειγμα, ένας αισθητήρας εφοδιασμένος με μια κάμερα που παρακολουθεί μια κινούμενη ομάδα ανθρώπων δεν μπορεί να μετρήσει τον ακριβή αριθμό και να καθορίσει την κατεύθυνσή τους. Σε αυτή την περίπτωση είναι απαραίτητη η συνεργασία των γειτονικών αισθητήρων για να καλυφθεί ολόκληρη η ομάδα των ανθρώπων. Συνεπώς, οι αισθητήρες απαιτείται να αποθηκεύουν πολυμεσικά δεδομένα, π.χ., εικόνα, βίντεο, που χρειάζονται από τις εφαρμογές, και επίσης να ανακτούν πολυμεσικά δεδομένα από απομακρυσμένους αισθητήρες με μικρή καθυστέρηση πρόσβασης.

Λαμβάνοντας υπόψη τους προηγούμενους περιορισμούς/απαιτήσεις, η επίτευξη του στόχου για Ποιότητα Υπηρεσιών σε εφαρμογές σε ΑΠΔΑ έγινε μια πολύ απαιτητική υπόθεση. Η λύση του προβλήματος θα μπορούσε να επιτευχθεί με διάφορους τρόπους, όπως για παράδειγμα η *προσαρμοσμένη στο κανάλι ροή* [58] (channel-adaptive streaming) ή η *συνδυασμένη κωδικοποίηση πηγής-καναλιού* (joint source-channel coding) [48]. Ωστόσο, οι τεχνικές αυτές δεν μπορούν να παράσχουν λύσεις στα τρία προαναφερθέντα θέματα. Σε αυτό το κεφάλαιο ερευνάται διεξοδικά η λύση του συνεργατικού caching πολυμεσικού περιεχομένου σε

αισθητήρες, έτσι ώστε να μπορέσουν να ικανοποιηθούν και τα τρία χαρακτηριστικά. Στα πλαίσια του συνεργατικού caching, πολλοί αισθητήρες διαμοιράζονται τα αποθηκευμένα δεδομένα και συνεργάζονται ώστε να μειωθεί το κόστος επικοινωνίας και να αξιοποιηθεί βέλτιστα ο αποθηκευτικός χώρος της cache κάθε συνεργαζόμενου αισθητήρα. Η θεώρηση μας είναι ότι κάθε αισθητήρας έχει μια τοπική αποθηκευτική χωρητικότητα, όπως για παράδειγμα είναι η μνήμη flash. Αν και υπάρχουν μνήμες flash με αποθηκευτική χωρητικότητα πολλών GB, π.χ. οι NAND flash [94] και οι τεχνολογικές εξελίξεις δείχνουν ότι γίνονται διαρκώς φθηνότερες, μεγαλύτερες και περισσότερο κοινές, εντούτοις δεν θεωρούμε ότι οι αισθητήρες έχουν τεράστιες αποθηκευτικές δυνατότητες, έτσι ώστε να καλύψουμε ένα μεγαλύτερο φάσμα εφαρμογών και αρχιτεκτονικών.

Η διάρκεια ζωής της μπαταρίας μπορεί να επεκταθεί αν καταφέρουμε να μειώσουμε την ποσότητα της επικοινωνίας. Το caching των χρήσιμων δεδομένων για κάθε αισθητήρα είτε στον τοπικό αποθηκευτικό χώρο είτε στους γειτονικούς αισθητήρες μπορεί να επιμηκύνει το χρόνο ζωής ολόκληρου του δικτύου. Επιπρόσθετα, το caching μπορεί να είναι πολύ αποτελεσματικό στη μείωση των μεταδόσεων σε μεγάλη έκταση του δικτύου και επομένως στη μείωση των παρεμβολών και στην επιτυχή αντιμετώπιση των συνθηκών μεταβλητού καναλιού. Τελικά, το caching μπορεί να επιταχύνει την επεξεργασία των πολυμεσικών δεδομένων, επειδή η επεξεργασία και η παράδοση των πολυμεσικών περιεχομένων δεν είναι ανεξάρτητες και η αλληλεπίδρασή τους έχει σημαντικό αντίκτυπο στα επίπεδα της παρεχόμενης Ποιότητας Υπηρεσιών [3].

Σε αυτό το κεφάλαιο ερευνούμε την τεχνική του caching στα πλαίσια των ΑΠΔΑ. Η ανάγκη για αποτελεσματικές και αποδοτικές τεχνικές caching σε δίκτυα αισθητήρων έχει αναδειχθεί αρκετές φορές στη βιβλιογραφία [40, 98], αλλά δεν έχει προταθεί κάποια εκλεπτυσμένη πολιτική παρότι υπάρχουν αρκετά πρωτόκολλα caching σε άλλα πεδία (δες τις σχετικές εργασίες στην Παράγραφο 3.2). Επιπρόσθετα παρουσιάζουμε ένα νέο υψηλής απόδοσης πρωτόκολλο συνεργατικού caching, το οποίο ονομάζεται *Node Importance-based Cooperative Caching* (NiCoCa), και το συγκρίνουμε με την καλύτερη τεχνική που έχει προταθεί στα πλαίσια των MANETs. Η πειραματική αποτίμηση δείχνει ότι η προτεινόμενη μέθοδος υπερισχύει της ανταγωνίστριας.

## 3.2 Σχετικές Εργασίες

### 3.2.1 Υπόβαθρο

Η τεχνική του caching έχει ευρέως διερευνηθεί στα πεδία των Λειτουργικών Συστημάτων και των Βάσεων Δεδομένων αλλά ακόμα αποτελεί μια ελκυστική περιοχή έρευνας [95]. Παρομοίως, η τεχνική του caching έχει εκτενώς διερευνηθεί στα πλαίσια του Παγκόσμιου Ιστού, τόσο για συνεργατικές [51] όσο και για μη συνεργατικές [82, 101] αρχιτεκτονικές. Οι Wessels και Claffy [137] εισήγαγαν το πρωτόκολλο Internet Cache Protocol (ICP) για να παράσχουν επικοινωνία μεταξύ caching proxies με την ανταλλαγή μηνυμάτων. Τα πρωτόκολλα Cache Digests [111] και Summary Cache [51] επιτρέπουν σε proxies να ανταλλάσουν πληροφορίες σχετικά με τα αποθηκευμένα περιεχόμενά τους. Στο [28] μελετάται μια αρχιτεκτονική συνεργατικού ιεραρχικού caching για τον Παγκόσμιο Ιστό. Παρ' όλα αυτά, το πρόβλημα που διαπραγματεύονται τα ασύρματα δίκτυα είναι διαφορετικό από αυτό των ενσύρματων δικτύων. Αυτές οι αρχιτεκτονικές και πρωτόκολλα θεωρούν συνήθως μια σταθερή μη μεταβαλλόμενη τοπολογία δικτύου και απαιτούν ισχυρές υπολογιστικές και επικοινωνιακές δυνατότητες. Λόγω των περιορισμένων πόρων (π.χ. εύρος ζώνης, ισχύς μπαταρίας και υπολογιστική ισχύς) των κόμβων αισθητήρων, αυτές οι τεχνικές δεν μπορούν να προσαρμοσθούν καλά σε ΑΔΑ.

Στα πλαίσια των ασύρματων κυψελοειδών δικτύων έχει προταθεί ένας σημαντικός αριθμός από πολιτικές caching [81, 131]. Αυτές οι πολιτικές όμως προϋποθέτουν την ύπαρξη πολύ ισχυρότερων κόμβων συγκριτικά με τους κόμβους αισθητήρες. Επιπλέον, η επικοινωνία μεταξύ των κόμβων και των σταθμών βάσης θεωρείται ότι είναι 1-hop και οι σταθμοί βάσης παρέχουν τα απαραίτητα δεδομένα.

Ένας σημαντικός αριθμός σχημάτων αντιγράφων δεδομένων (data replication schemes) [63, 64] και σχημάτων caching [113, 127, 149] έχει προταθεί για να διευκολυνθεί η πρόσβαση στα δεδομένα σε MANETs. Τα σχήματα των αντιγράφων δεδομένων μελετούν το ζήτημα της κατανομής αντιγράφων ώστε να βελτιωθεί η πρόσβαση στα δεδομένα. Οι τεχνικές αυτές απαιτούν τις περισσότερες φορές μια εκ των προτέρων γνώση της τοπολογίας του δικτύου.

Τα σχήματα caching όμως δεν διευκολύνουν την πρόσβαση στα δεδομένα με βάση τη γνώση της κατανομής των δεδομένων. Στην τεχνική SimpleCache [149] τα ζητούμενα δεδομένα γίνονται πάντα cached μόνο στον αιτούντα κόμβο. Ο κόμβος χρησιμοποιεί το αντίγραφο που έχει στην cache για να μπορέσει να εξυ-

πηρετήθει μεταγενέστερες αιτήσεις από άλλους κόμβους για το ίδιο δεδομένο. Σε περίπτωση cache miss ο κόμβος που ζητά το δεδομένο θα πρέπει να το αντλήσει από την πηγή των δεδομένων. Η αύξηση όμως της απόστασης μεταξύ του κόμβου που αιτείται το δεδομένο και του caching κόμβου έχει ως αποτέλεσμα την αύξηση του χρόνου απόκρισης.

Οι περισσότερο σχετικές ερευνητικές εργασίες προς το προτεινόμενο πρωτόκολλο είναι τα πρωτόκολλα συνεργατικού caching, τα οποία έχουν αναπτυχθεί στην περιοχή των MANET. Το κυριότερο κίνητρο για την ανάπτυξη των πρωτοκόλλων είναι η κινητικότητα των κόμβων, και συνεπώς όλες οι ερευνητικές συνεισφορές προσπαθούν είτε να την διαμορφώσουν είτε να την αξιοποιήσουν. Πρόσφατα προτάθηκε στη βιβλιογραφία μια πολιτική συνεργατικού caching, η οποία ονομάζεται Cooperative Caching (CoCa) [33]. Η πολιτική CoCa διευκολύνει τους κινητούς κόμβους να διαμοιράσουν τα δεδομένα που έχουν στην cache, έτσι ώστε να μειωθεί ο αριθμός των αιτήσεων που απευθύνεται στον εξυπηρετητή και ο αριθμός των άστοχων προσβάσεων (access misses) σε δεδομένα. Οι εισηγητές της πολιτικής την επέκτειναν με ένα σχήμα συνεργατικού caching βασισμένο σε ομάδες, το οποίο ονομάστηκε GroCoCa [34]. Σύμφωνα με το GroCoCa η απόφαση εάν ένα δεδομένο θα πρέπει να τοποθετηθεί στην cache εξαρτάται από δύο παράγοντες. Ο πρώτος είναι η κινητικότητα του κόμβου και ο δεύτερος είναι η συνάφεια των δεδομένων που ζητούνται από τους κόμβους. Ο κινητός σταθμός υποστήριξης εκτελεί έναν αυξητικό αλγόριθμο ομαδοποίησης για να μπορέσει να ομαδοποιήσει τους κινητούς κόμβους του δικτύου με βάση το πρότυπο της κίνησής τους σε στενά συνδεδεμένες ομάδες. Στο πρωτόκολλο GroCoCa η ομοιότητα στις αιτήσεις για δεδομένα μπορεί να υπολογισθεί με μια μετρική ομοιότητας βασισμένη στη συχνότητα. Το πρωτόκολλο GroCoCa βελτιώνει την απόδοση του συστήματος καθώς μειώνει την κατανάλωση ενέργειας.

Σε αντίθεση με το πρωτόκολλο NiCoCa οι προαναφερθείσες ερευνητικές εργασίες επικεντρώνονται σε single-hop MANET, και όχι σε multi-hop ΑΠΔΑ. Οι Papadopoulou και Schulzrinne [102] πρότειναν μια αρχιτεκτονική που ονομάζεται 7DS. Οι συγγραφείς ανέπτυξαν ένα ζεύγος πρωτοκόλλων για να διευκολυνθεί ο διαμοιρασμός και η διάχυση της πληροφορίας μεταξύ των χρηστών. Η αρχιτεκτονική λειτουργεί σε δύο καταστάσεις. Η πρώτη είναι η κατάσταση prefetch, η οποία βασίζεται στη μελλοντική πληροφορία που θα χρειασθούν οι χρήστες, ενώ η δεύτερη ονομάζεται κατάσταση κατά απαίτηση (on-demand) και περιλαμβάνει την αναζήτηση δεδομένων όταν τίθεται μια αίτηση από ένα χρήστη στην 1-hop γειτονιά. Βασισμένη στη συνεργατική συμπεριφορά, η αρχιτεκτονική χρησιμοποιεί το μοντέλο ομότιμων (peer-to-peer) δικτύων και το μοντέλο πελάτη-εξυπηρετητή.

Αντίθετα με τη προτιεγόμενη προσέγγιση, η αρχιτεκτονική 7DS στοχεύει επίσης τα single-hop ασύρματα περιβάλλοντα και συνεπώς δεν διερευνώνται κατάλληλα η διαχείριση της cache, συμπεριλαμβανομένου του ελέγχου αποδοχής ενός αντικειμένου στην cache καθώς και η πολιτική αντικατάστασης. Οι Sailhan και Issarny [114] πρότειναν μια στρατηγική συνεργατικής διαχείρισης της cache μεταξύ των κινούμενων τερματικών που αλληλεπιδρούν μέσω ενός τυχαίου δικτύου. Το ζήτημα που απασχολεί τους συγγραφείς είναι η διαμόρφωση ενός τυχαίου δικτύου κινούμενων τερματικών που συνεργάζονται για την ανταλλαγή ιστοσελίδων. Η στρατηγική αυτή στόχευε στη μείωση της καθυστέρησης πρόσβασης στον Παγκόσμιο Ιστό καθώς και στη βελτιστοποίηση της καταναλισκόμενης ενέργειας. Η υλοποίησή της γίνεται ένα επίπεδο επάνω από το πρωτόκολλο Zone Routing Protocol (ZRP).

Ορισμένα πρωτόκολλα προσπαθούν να σχηματίσουν ομάδες κόμβων με βάση είτε τη γεωγραφική γειτνίαση ή χρησιμοποιώντας αλγορίθμους ομαδοποίησης που είναι ευρέως γνωστοί από το ερευνητικό πεδίο των MANET. Τα πρωτόκολλα αυτά είναι το Zone Cooperative [24], το Cluster Cooperative [23] και το ECOR [117]. Στο πρωτόκολλο Zone Cooperative (ZC) οι κινητοί κόμβοι που ανήκουν στη γειτονιά (ζώνη) ενός συγκεκριμένου κόμβου σχηματίζουν ένα σύστημα συνεργατικής cache για τον κόμβο. Αυτό συμβαίνει, επειδή η επικοινωνία με τους κόμβους του δικτύου που βρίσκονται στη γειτονιά είναι χαμηλού κόστους, τόσο με βάση την κατανάλωση ενέργειας όσο και με βάση τον αριθμό των ανταλλασσόμενων μηνυμάτων. Κάθε κόμβος του δικτύου περιλαμβάνει μια cache όπου αποθηκεύει τα συχνά προσβάσιμα δεδομένα. Τα δεδομένα της cache εξυπηρετούν όχι μόνο τις αναζητήσεις του ίδιου του κόμβου που τα έχει αποθηκευμένα, αλλά και τις διερχόμενες αιτήσεις που τίθενται από άλλους κόμβους. Σε περίπτωση αποτυχίας εύρεσης του ζητούμενου δεδομένου στην τοπική cache, ο κόμβος αρχικά κάνει αναζήτηση στους γειτονικούς κόμβους που σχηματίζουν τη ζώνη. Αν η αίτηση για το δεδομένο δεν ικανοποιηθεί από αυτούς, τότε η αίτηση προωθείται στον επόμενο κόμβο που βρίσκεται στο μονοπάτι προς την πηγή δεδομένων. Το πρωτόκολλο περιλαμβάνει επίσης μια πολιτική αντικατάστασης της cache, η οποία βασίζεται στον υπολογισμό μιας τιμής για κάθε αποθηκευμένο αντικείμενο. Η τιμή υπολογίζεται με βάση το μέγεθος του αντικειμένου, τη δημοτικότητά του, την απόσταση και το χρόνο ζωής (Time-To-Live – TTL) του αντικειμένου. Οι πειραματικές μετρήσεις ανέδειξαν βελτιώσεις στο cache hit ratio και στη μέση καθυστέρηση πρόσβασης σε σύγκριση με άλλες στρατηγικές caching.

Το πρωτόκολλο Cluster Cooperative (CC) αποτελεί ένα σχήμα caching σε MANET. Στόχος του είναι η μείωση του υψηλού κόστους αναζήτησης cache και

η παροχή καλύτερης απόδοσης. Οι συγγραφείς διαίρεσαν ολόκληρο το κινητό τυχαίο δίκτυο σε ίσου μεγέθους ομάδες/συμπλέγματα με κριτήριο τη γεωγραφική τοποθεσία. Σε κάθε σύμπλεγμα, το πρωτόκολλο επιλέγει έναν υπερ-κόμβο ως κόμβο κατάστασης της cache. Ο υπερ-κόμβος αναλαμβάνει να διατηρεί την πληροφορία που σχετίζεται με τη cache κάθε κόμβου που βρίσκεται μέσα στο σύμπλεγμα. Ως υπερ-κόμβος ορίζεται ο πρώτος εισερχόμενος κόμβος σε ένα σύμπλεγμα. Η κατάσταση της cache του συμπλέγματος για ένα κόμβο είναι η λίστα από τα cached αντικείμενα μαζί με το πεδίο TTL του κάθε αντικειμένου. Η πολιτική αντικατάστασης της cache είναι παρόμοια με αυτή του πρωτοκόλλου ZC. Ωστόσο το πρωτόκολλο ZC παρουσιάζει μειονεκτήματα, όπως το υψηλό κόστος επικοινωνίας και η κατανάλωση ενέργειας. Αυτό συμβαίνει επειδή κάθε κόμβος επάνω στο μονοπάτι προς τον κόμβο πηγή θα πρέπει να εκπέμπει το μήνυμα στους γειτονικούς κόμβους για να διαπιστώσει αν υπάρχει το ζητούμενο αντικείμενο σε κάποια cache. Σε αντίθεση με το πρωτόκολλο ZC, το πρωτόκολλο CC παρουσιάζει μικρότερο κόστος ως προς τα μηνύματα που ανταλλάσσονται μεταξύ των κόμβων, γιατί οι ενδιάμεσοι κόμβοι (οι κόμβοι μεταξύ του κόμβου που υποβάλλει την αίτηση και του κόμβου που θα εξυπηρετήσει την αίτηση) εκπέμπουν το μήνυμα της αίτησης μόνο στον υπερ-κόμβο.

Στην ίδια κατηγορία με τα δύο προηγούμενα πρωτόκολλα (ZC, CC) συγκαταλέγεται και το πρωτόκολλο ECOR, όπου κάθε κινητός κόμβος σχηματίζει μια ζώνη συνεργασίας (cooperation zone) με τους γείτονες κόμβους. Αυτό επιτυγχάνεται με την ανταλλαγή μηνυμάτων και έχει ως αποτέλεσμα το διαμοιρασμό των δεδομένων που βρίσκονται στις caches. Με αυτό το τρόπο είναι εφικτή η μείωση του εύρους ζώνης και η μείωση της διασπάθισης ενέργειας για κάθε ανάκτηση αντικειμένου. Στα πλαίσια λειτουργίας του ECOR όταν μια αίτηση για κάποιο δεδομένο καταφθάνει σε ένα κόμβο, αυτός αναζητά ένα αντίγραφο του δεδομένου στη ζώνη συνεργασίας του πριν προωθήσει την αίτηση στον κόμβο πηγή. Οι δημιουργοί του πρωτοκόλλου ανέπτυξαν ένα αναλυτικό μοντέλο για να καθορίσουν τη βέλτιστη ακτίνα της ζώνης συνεργασίας. Το μοντέλο βασίστηκε στην τοποθεσία του κόμβου, στη δημοτικότητα του αντικειμένου και στη δικτυακή πυκνότητα κόμβων. Σύμφωνα με το πρωτόκολλο κάθε κόμβος εκπέμπει κάθε διαφοροποίηση που έχει συμβεί στα αντικείμενα που βρίσκονται στην cache στους κόμβους που ανήκουν στη ζώνη συνεργασίας. Κάθε κόμβος διατηρεί ένα πίνακα με πληροφορίες για τα αντικείμενα που βρίσκονται στις caches των γειτονικών κόμβων. Το πρωτόκολλο όμως παρουσιάζει κάποια μειονεκτήματα, όπως είναι ο μεγάλος αριθμός ανταλλασσόμενων μηνυμάτων και επομένως η κατανάλωση ενέργειας, ιδιαίτερα σε περιπτώσεις μεγάλης πυκνότητας κόμβων και μεγάλης ακτίνας ζώνης συνεργα-

σίας. Τέλος, μειονέκτημα αποτελεί ακόμα το υψηλό κόστος επικοινωνίας, όταν η ακτίνα της ζώνης είναι μεγάλη και κάποιος κόμβος λανθασμένα κατευθύνει μια αίτηση σε ένα γειτονικό κόμβο.

Τα μόνα πρωτόκολλα που προσπάθησαν να αξιοποιήσουν ταυτόχρονα και τα δεδομένα και την τοπικότητα ενός κόμβου (η θέση του κόμβου σε σχέση με του υπόλοιπους γειτονικούς κόμβους) περιγράφονται στη μελέτη [149] και είναι τα εξής: CachePath, CacheData, και HybridCache. Στο πρωτόκολλο CacheData, οι ενδιαμέσοι κόμβοι μπορούν να αποθηκεύσουν δεδομένα για να εξυπηρετήσουν μελλοντικές αιτήσεις, αντί να φέρουν τα δεδομένα από τον κόμβο πηγή. Ένας ενδιαμέσος κόμβος αποθηκεύει τοπικά κάποιο διερχόμενο δεδομένο όταν αντιλαμβάνεται ότι είναι δημοφιλές (ζητείται από πολλούς διαφορετικούς κόμβους), ενώ δεν το αποθηκεύει όταν όλες οι αιτήσεις προέρχονται από τον ίδιο κόμβο. Αυτός ο κανόνας σχεδιάστηκε για να μειώσει την απαίτηση σε χώρο της cache. Στο πρωτόκολλο CachePath, ένας κινητός κόμβος μπορεί να αποθηκεύσει την πληροφορία σχετικά με το μονοπάτι σε κάποιο κοντινό κόμβο που έχει αποστείλει μια αίτηση για ένα δεδομένο, καθώς προωθεί το δεδομένο. Η πληροφορία για το μονοπάτι μπορεί να χρησιμοποιηθεί σε μελλοντικές αιτήσεις έτσι ώστε να τις κατευθύνει προς τον κόμβο που είχε ζητήσει και λάβει το δεδομένο παλαιότερα. Η τεχνική αυτή έχει ως αποτέλεσμα την ελάττωση του χρησιμοποιούμενου εύρους ζώνης και τη μείωση της καθυστέρησης στην αναζήτηση της πληροφορίας, αφού τα δεδομένα μπορούν να αποκτηθούν μέσω συντομότερης απόστασης.

Οι συγγραφείς προτείνουν ακόμα κάποιες τεχνικές βελτιστοποίησης της απόδοσης των πρωτοκόλλων. Ένας ενδιαμέσος κόμβος αποθηκεύει μόνο τον τελικό κόμβο (κόμβο προορισμό) επειδή το μονοπάτι από τον τρέχοντα κόμβο προς τον κόμβο προορισμό προκύπτει από το πρωτόκολλο δρομολόγησης. Μια άλλη τεχνική που προτείνεται για τη βελτίωση της απόδοσης είναι κάθε ενδιαμέσος κόμβος να αποθηκεύει τον κόμβο προορισμό μόνο όταν αυτός βρίσκεται εγγύτερα σε σχέση με τον κόμβο πηγή. Όμως σύμφωνα με τους συγγραφείς, το πρωτόκολλο CachePath εμφανίζει ένα σημαντικό μειονέκτημα που μπορεί να οδηγήσει σε μείωση της αξιοπιστίας του πρωτοκόλλου και σε αύξηση του πλήθους των ανταλλασσόμενων μηνυμάτων. Η τοποθέτηση στην cache ενός μονοπατιού μπορεί να μην είναι αξιόπιστη, επειδή είτε το αντικείμενο έχει απαρχαιωθεί είτε δεν μπορεί να βρεθεί ο κόμβος που έχει το μονοπάτι. Το υβριδικό πρωτόκολλο HybridCache συνδυάζει τα πρωτόκολλα CacheData and CachePath, ενώ παράλληλα προσπαθεί να αποφύγει τις αδυναμίες τους. Όταν ένας κόμβος προωθεί ένα αντικείμενο στο πρωτόκολλο HybridCache, τότε με βάση κάποια κριτήρια τοποθετεί στην cache είτε το αντικείμενο είτε το μονοπάτι. Τα κριτήρια περιλαμβάνουν το

μέγεθος, το χρονικό όριο ζωής του αντικειμένου και τον αριθμό από hops που ένα αποθηκευμένο μονοπάτι μπορεί να εξοικονομήσει, ο οποίος δηλώνεται ως  $H_{save}$ . Η τιμή  $H_{save}$  δηλώνει τη διαφορά μεταξύ της απόστασης προς τον κόμβο πηγή του αντικειμένου και της απόστασης προς τον caching κόμβο. Η τιμή αυτή πρέπει να είναι μεγαλύτερη από μια τιμή κατώφλι. Παρ' όλα αυτά, σύμφωνα με τις προηγούμενες μεθόδους η πληροφορία που βρίσκεται στην cache ενός κόμβου δεν μπορεί να διαμοιρασθεί εάν ο κόμβος δεν βρίσκεται επάνω στο μονοπάτι που συνδέει το κόμβο που θέτει την αίτηση με τον κόμβο πηγή. Επιπλέον, οι τιμές κατώφλια που χρησιμοποιούνται από τις ευριστικές μεθόδους πρέπει να τεθούν προσεκτικά ώστε να επιτευχθεί καλή απόδοση του συστήματος.

Τέλος, στην περιοχή του caching σε ΑΔΑ υπάρχουν ερευνητικές εργασίες που στοχεύουν στην τοποθέτηση των caches και επομένως δεν τις εξετάζουμε [105, 128]. Επίσης μερικώς σχετικά προς το προτεινόμενο πρωτόκολλο NiCoCa είναι και τα πρωτόκολλα δρομολόγησης, τα οποία έχουν προταθεί στα πλαίσια των MANET [155].

### 3.2.2 Γιατί χρειαζόμαστε ένα νέο πρωτόκολλο caching;

Τα πρωτόκολλα που έχουν προταθεί για συνεργατικό caching σε MANET παρουσιάζουν διάφορους περιορισμούς. Η κατηγορία των πρωτοκόλλων που πραγματοποιούν αρχικά ομαδοποίηση των κόμβων και μετά αξιοποιούν την ομαδοποίηση (δηλαδή, τους κόμβους κεφαλές κάθε παραγόμενης ομάδας) για να συντονίσουν τις αποφάσεις caching, κληρονομούν τις αδυναμίες που προέρχονται από την επιλογή των κεφαλών αυτών. Για παράδειγμα, στις μελέτες [23, 24], θεωρείται ότι οι κόμβοι μιας ομάδας έχουν την ίδια εμβέλεια επικοινωνίας. Δηλαδή, όλοι οι κόμβοι μιας ομάδας έχουν 1-hop απόσταση μεταξύ τους. Επιπρόσθετα, οι κόμβοι δεν αποθηκεύουν την πληροφορία που προέρχεται από κάποιον 1-hop γείτονά τους. Έτσι, οι κεφαλές που δεν βρίσκονται σε κάποιο σημαντικό τμήμα των μονοπατιών του δικτύου δεν μπορούν να εξυπηρετήσουν αποδοτικά τα μέλη της ομάδας τους, καθώς δεν έχουν γρήγορη πρόσβαση στα ζητούμενα δεδομένα. Η ζώνη συνεργασίας που σχηματίζεται στη μελέτη [117] με την επιλογή της καταλληλότερης ακτίνας εκπομπής, συνεπάγεται ένα μεγάλο κόστος επικοινωνίας, επειδή κάθε κόμβος που βρίσκεται μέσα στην ακτίνα πρέπει να στέλνει/λαμβάνει κάθε αλλαγή που συμβαίνει στην cache σε όλους τους άλλους κόμβους που ανήκουν μέσα στην ακτίνα εκπομπής. Εν τέλει, το πρωτόκολλο HybridCache είναι στενά συνδεδεμένο με το χρησιμοποιούμενο πρωτόκολλο δρομολόγησης και συνεπώς εάν ένας κόμβος δεν βρίσκεται επάνω στη διαδρομή που έχει επιλέξει το



πρωτόκολλο δρομολόγησης δεν μπορεί να τοποθετήσει στην cache το δεδομένο ή το μονοπάτι και συνακόλουθα δεν μπορεί να εξυπηρετήσει την αίτηση ακόμα και εάν έχει το ζητούμενο δεδομένο.

Παρακινούμενοι από τις αδυναμίες και τα ελαττώματα των προηγούμενων πρωτοκόλλων συνεργατικού caching και λαμβάνοντας υπόψη τα μοναδικά χαρακτηριστικά και τους περιορισμούς που τίθενται στα ΑΠΔΑ που συντίθενται επί το πλείστον από στατικούς και όχι κινητούς κόμβους, προτείνουμε ένα νέο πρωτόκολλο συνεργατικού caching. Το προτεινόμενη πρωτόκολλο NiCoCa βασίζεται στην ιδέα της αξιοποίησης της δικτυακής τοπολογίας, έτσι ώστε να ανακαλύψει τους κόμβους που είναι σημαντικότεροι από κάποιους άλλους. Η σημαντικότητα σχετίζεται με την τοποθεσία ενός κόμβου μέσα στο δίκτυο και την εναπομένουσα ενέργειά του. Ενσωματώνοντας και τους δύο παράγοντες στη σχεδίαση της πολιτικής caching εξασφαλίζουμε μεγαλύτερη διάρκεια ζωής για το δίκτυο και μικρή καθυστέρηση πρόσβασης στην ανάκτηση πολυμεσιικών δεδομένων. Περιληπτικά, τα κύρια σημεία συνεισφοράς του πρωτοκόλλου NiCoCa είναι τα εξής:

- Ο ορισμός μιας μετρικής για τον υπολογισμό της σημαντικότητας ενός κόμβου αισθητήρα μέσα στην τοπολογία του δικτύου, η οποία θα συνεπάγεται μικρή καθυστέρηση πρόσβασης στην ανάκτηση των δεδομένων.
- Η περιγραφή του πρωτοκόλλου συνεργατικού caching, η οποία λαμβάνει υπόψη την εναπομένουσα ενέργεια των κόμβων αισθητήρων.
- Η ανάπτυξη των αλγορίθμων για την αναζήτηση και εύρεση των ζητούμενων πολυμεσιικών δεδομένων, και η πολιτική αντικατάστασης της cache.
- Η αξιολόγηση της απόδοσης του πρωτοκόλλου και η σύγκρισή του με την καλύτερη τεχνική που έχει προταθεί στα πλαίσια των MANET. Η πειραματική αποτίμηση και σύγκριση έχει γίνει με τη χρήση του λογισμικού προσομοίωσης J-Sim [123].

### 3.3 Πρωτόκολλο NiCoCa για Συνεργατικό Caching

Ένα από τα σημαντικότερα σημεία του πρωτοκόλλου NiCoCa είναι η εκτίμηση της σπουδαιότητας των αισθητήρων με βάση την τοπολογία του δικτύου και τη συνεργασία μεταξύ των κόμβων, η οποία επιτυγχάνεται μέσω των αισθητήρων αυτών. Η ιδέα βασίζεται στο ότι εάν ανακαλυφθούν οι κόμβοι που συμμετέχουν σε ένα σημαντικό ποσοστό των (συντομότερων) μονοπατιών που συνδέουν άλλους

κόμβους, τότε αυτοί οι κόμβοι ονομάζονται *σημαντικοί κόμβοι*. Επομένως, οι σημαντικοί κόμβοι μπορεί να επιλεγούν ως συντονιστές για τις αποφάσεις caching, π.χ., ως διαμεσολαβητές (mediators) για να παρέχουν πληροφορία σχετικά με την πρόσβαση των ζητούμενων δεδομένων ή ακόμα και ως σημεία caching. Οι κόμβοι απαρτίζουν το κύριο σημείο του πρωτοκόλλου συνεργατικού caching. Η συνεργασία μεταξύ γειτονικών κόμβων και ο έλεγχος κατανάλωσης ενέργειας πραγματοποιείται από τους κόμβους διαμεσολαβητές που περιέχουν δείκτες σχετικά με τα δεδομένα που έχουν τοποθετηθεί στις caches, την εναπομένουσα ενέργεια και τον ελεύθερο χώρο σε κάθε cache των γειτονικών κόμβων. Συνεπώς, κάθε φορά που μια αίτηση για δεδομένα καταφθάνει σε ένα διαμεσολαβητή, αυτός αποφασίζει ποιός γείτονας με τη μεγαλύτερη εναπομένουσα ενέργεια μπορεί να την απαντήσει. Οι διαμεσολαβητές λαμβάνουν μέρος επίσης στη φάση αντικατάστασης της cache, η οποία συμβαίνει σε ένα γειτονικό κόμβο. Η συνεργασία μεταξύ του κόμβου που αντικαθιστά κάποια δεδομένα της cache και των διαμεσολαβητών λαμβάνει χώρα για να μην απομακρυνθούν οριστικά τα δεδομένα από τη γειτονική περιοχή. Λαμβάνοντας υπόψη τον ελεύθερο χώρο στην cache και τα δεδομένα που βρίσκονται στην cache κάθε γειτονικού κόμβου, οι διαμεσολαβητές αποφασίζουν που πρέπει να αποθηκευθούν τα αντικαθιστώμενα δεδομένα. Εάν τα αντικαθιστώμενα δεδομένα υπάρχουν κάπου αλλού στη γειτονιά, τότε απλά διαγράφονται.

### 3.3.1 Μέτρηση σπουδαιότητας ενός κόμβου αισθητήρα

Ένα ΑΠΔΑ συνοψίζεται ως ένας γράφος  $G(V, E)$ , όπου  $V$  το σύνολο των κόμβων και  $E$  το σύνολο των ασύρματων συνδέσεων μεταξύ των κόμβων. Για  $u, v \in E$  μια ακμή  $e = (u, v)$  υπάρχει εάν και μόνο εάν ο κόμβος  $u$  βρίσκεται μέσα στο εύρος μετάδοσης του κόμβου  $v$ , και το αντίστροφο. Συνεπώς, όλοι οι σύνδεσμοι του γράφου είναι διπλής κατεύθυνσης, π.χ., εάν ο κόμβος  $u$  βρίσκεται μέσα στο εύρος μετάδοσης του κόμβου  $v$ , τότε και ο κόμβος  $v$  βρίσκεται μέσα στο εύρος μετάδοσης του κόμβου  $u$ . Το σύνολο των γειτονικών κόμβων ενός κόμβου  $v$  αναπαρίσταται ως  $N_1(v)$ , δηλαδή  $N_1(v) = \{u : (v, u) \in E\}$ . Το σύνολο των 2-hop κόμβων ενός κόμβου  $v$ , π.χ., οι κόμβοι που είναι γείτονες των γειτόνων του κόμβου  $v$  εκτός από τους γείτονες του κόμβου  $v$  αναπαρίσταται ως  $N_2(v)$ , όπου  $N_2(v) = \{w : (u, w) \in E, \text{ όπου } w \neq v \text{ και } w \notin N_1 \text{ και } (v, u) \in E\}$ . Τέλος το συνδυασμένο σύνολο των 1-hop και 2-hop γειτόνων ενός κόμβου  $v$  δηλώνεται ως  $N_{12}(v)$ . Τέλος, θεωρείται ότι το δίκτυο βρίσκεται σε συνδεδεμένη κατάσταση.

**ΟΡΙΣΜΟΣ 3.1** (Η όψη του τοπικού δικτύου για ένα κόμβο  $v$ ).

Η όψη του τοπικού δικτύου ενός γράφου  $G(V, E)$  σε σχέση με τον κόμβο  $v \in V$

είναι ο τεχνητός υπογράφος του  $G$  που σχετίζεται με το σύνολο των κόμβων που ανήκουν στο  $N_{12}(v)$  και δηλώνεται ως  $LN_v$ .

Ένα μονοπάτι από τον κόμβο  $u \in V$  στον  $w \in V$  αποτελεί μία ακολουθία κορυφών και ακμών, ξεκινώντας από το  $u$  και τελειώνοντας στο  $w$ . Το μήκος ενός μονοπατιού είναι το πλήθος των ενδιάμεσων ακμών. Συμβολίζουμε με  $d_G(u, w)$  την απόσταση μεταξύ των  $u$  και  $w$ , δηλαδή το ελάχιστο μήκος ενός μονοπατιού που συνδέει το  $u$  και το  $w$  στο  $G$ , όπου  $\forall u, v, w \in V$  εξ ορισμού ισχύει  $d_G(v, v) = 0$  και  $d_G(u, w) = d_G(w, u)$ . Πρέπει να σημειωθεί ότι η απόσταση δεν σχετίζεται με το κόστος των συνδέσμων του δικτύου (π.χ. καθυστέρηση πρόσβασης), αλλά είναι μια αφαιρετική μετρική που μετρά το πλήθος των hops.

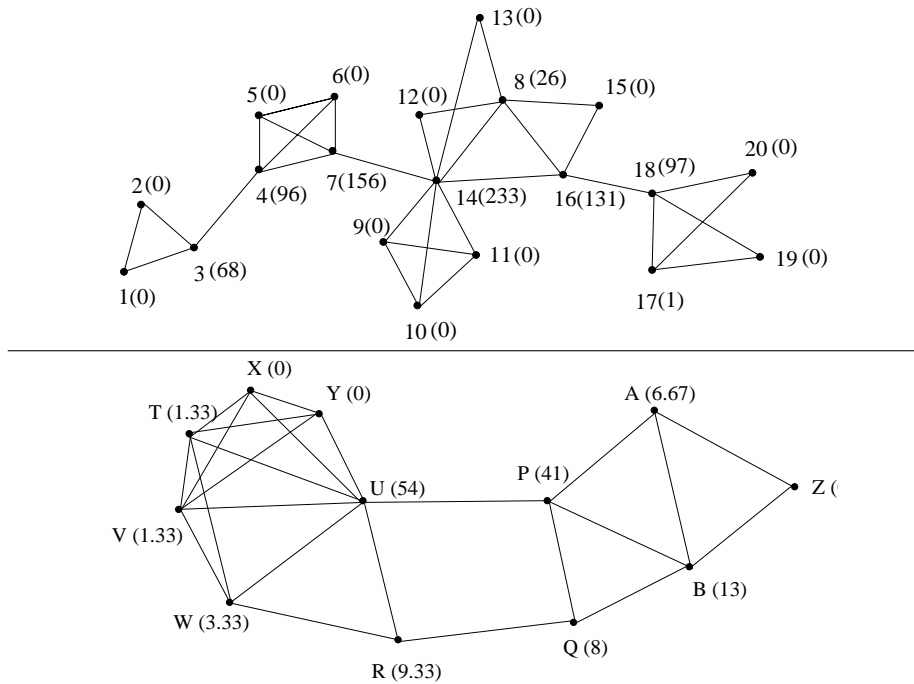
Θεωρούμε ότι ο συμβολισμός  $\sigma_{uw}$  δηλώνει τον αριθμό των συντομότερων μονοπατιών από το  $u \in V$  στο  $w \in V$  (εξ ορισμού ισχύει  $\sigma_{uu} = 0$  και  $\sigma_{uw} = \sigma_{wu}$ ). Επιπλέον, θεωρούμε ότι το  $\sigma_{uw}(v)$  δηλώνει το πλήθος των συντομότερων μονοπατιών από το  $u$  στο  $w$ , τα οποία διέρχονται από κάποια κορυφή  $v \in V$ . Συνεπώς, ορίζουμε το δείκτη σπουδαιότητας κόμβου (node importance index),  $NI(v)$ , για έναν κόμβο  $v$  ως:

$$NI(v) = \sum_{u \neq v \neq w \in V} \frac{\sigma_{uw}(v)}{\sigma_{uw}}. \quad (3.1)$$

Μεγάλες τιμές για το δείκτη  $NI$  ενός κόμβου  $v$  υποδηλώνουν ότι ο κόμβος  $v$  μπορεί να επικοινωνήσει με άλλους κόμβους μέσω σχετικά σύντομων μονοπατιών, ή ότι ο κόμβος  $v$  βρίσκεται σε ένα σχετικά μεγάλο πλήθος από συντομότερα μονοπάτια που συνδέουν άλλους κόμβους του δικτύου. Η εικονογράφηση της μετρικής παρουσιάζεται στο Σχήμα 3.1. Οι αριθμοί στις παρενθέσεις δηλώνουν το δείκτη  $NI$  του αντίστοιχου κόμβου θεωρώντας ολόκληρη την τοπολογία του ΑΠΔΑ.

Προφανώς, όταν μετρούμε το δείκτη  $NI$  για κάθε κόμβο θεωρώντας ολόκληρο του δίκτυο αποκτούμε μια πολύ κατατοπιστική εικόνα των κόμβων που βρίσκονται σε ένα μεγάλο πλήθος συντομότερων μονοπατιών που συνδέουν άλλους κόμβους του δικτύου. Είναι πολύ εύκολο να επιβεβαιωθεί ότι ακόμα και αν υπολογίζουμε τους δείκτες  $NI$  των κόμβων λαμβάνοντας υπόψη μόνο τη  $k$ -hop γειτονιά (για  $k=2$  ή  $3$ ), η εικόνα που θα αποκτήσουμε για τη σχετική σπουδαιότητα των κόμβων παραμένει πολύ ακριβής. Για τιμή  $k=1$  ο δείκτης  $NI$  ενός κόμβου αισθητήρα είναι ισοδύναμος με το βαθμό του κόμβου.

Ο ψευδοκώδικας του αλγορίθμου ComputeNI για τον υπολογισμό του δείκτη  $NI$  ενός κόμβου μπορεί να βρεθεί στη μελέτη [15]. Ο αλγόριθμος είναι ικανός να διαχειρίζεται πολλά συντομότερα μονοπάτια μεταξύ σε δύο κόμβων.



Σχήμα 3.1. Υπολογισμός δείκτη  $NI$  για δύο παραδείγματα γράφων.

**ΘΕΩΡΗΜΑ 3.2** ([15]). *Η πολυπλοκότητα του αλγορίθμου  $ComputeNI$  είναι  $O(n \times m)$  για ένα γράφο με  $n$  κορυφές και  $m$  ακμές.*

Με τη πρώτη ματιά ο υπολογισμός του  $NI$  φαίνεται ακριβός, δηλαδή  $O(m \times n^2)$  πράξεις για την 2-hop γειτονιά που αποτελείται από  $n$  κόμβους και  $m$  συνδέσεις. Ευτυχώς, μπορούμε να έχουμε καλύτερη απόδοση κάνοντας κάποιες έξυπνες παρατηρήσεις.

### 3.3.2 Προκαταρκτικά για το πρωτόκολλο NiCoCa

Υιοθετώντας το μοντέλο που παρουσιάζεται στη μελέτη [149], υποθέτουμε ότι η πηγή των πολυμεσικών δεδομένων είναι το κέντρο δεδομένων (data center). Η υπόθεση δεν είναι καθόλου περιοριστική και απλά εγγυάται ότι κάθε αίτηση, εάν δεν εξυπηρετείται από άλλους κόμβους και εάν δεν έχει λήξει, θα εξυπηρετηθεί τελικά από το κέντρο δεδομένων.

Αρχικά υποθέτουμε ότι κάθε κόμβος έχει γνώση της 2-hop γειτονιάς του. Αυτή η πληροφορία αποκτάται μέσω της περιοδικής ανταλλαγής μηνυμάτων beacon με τους γειτονικούς κόμβους. Υποθέτουμε επίσης ότι μπορούμε να εκχωρή-

σοιμε χρονικά διαστήματα (time slots) στους κόμβους έτσι ώστε να μην συμβαίνει παρεμβολές, δηλαδή δύο κόμβοι να εκπέμπουν στο ίδιο time slot. Χρησιμοποιώντας την προηγούμενη εκχώρηση, οι κόμβοι μπορούν να μεταδίδουν μηνύματα beacon για να ανακαλύψουν τους γείτονές τους και παράλληλα να αποκτήσουν κανάλι για τη μετάδοση των αιτήσεών τους για δεδομένα. Η εκχώρηση καθορίζεται χρησιμοποιώντας τον αλγόριθμο D2-coloring [56]. Αν και μπορούν να χρησιμοποιηθούν παραδοσιακότερες μέθοδοι για την επιδιαιτησία του καναλιού, αυτή η προσέγγιση συνεισφέρει στη μείωση της καθυστέρησης πρόσβασης καθώς αποφεύγονται οι παρεμβολές. Στην συνέχεια κάθε κόμβος υπολογίζει το δείκτη  $NI$  των 1-hop γειτόνων του και χρησιμοποιεί αυτή την πληροφορία για να μπορέσει να χαρακτηρίσει κάποιους από τους γειτονικούς του κόμβους ως *διαμεσολαβητές*. Ένας κόμβος επιλέγει ως διαμεσολαβητές το ελάχιστο σύνολο των γειτόνων με τις μεγαλύτερες τιμές δείκτη  $NI$ , οι οποίοι καλύπτουν την 2-hop γειτονιά του. Κάθε κόμβος είναι υπεύθυνος να ειδοποιεί τους γείτονές του ποιους από αυτούς έχει επιλέξει ως διαμεσολαβητές. Συνεπώς ένας κόμβος μπορεί να βρίσκεται σε δύο καταστάσεις, να είναι είτε διαμεσολαβητής είτε κοινός (ordinary). Πρέπει να τονίσουμε εδώ ότι ένας κόμβος δεν είναι ανάγκη να γνωρίζει τη σπουδαιότητα των γειτόνων του, αλλά χρειάζεται να γνωρίζει τη σπουδαιότητά τους στη γειτονιά του. Φυσικά, εάν η έννοια της γειτονιάς επεκταθεί και καλύπτει ολόκληρο το δίκτυο (π.χ. κάθε κόμβος γνωρίζει την  $k$ -hop γειτονιά του, όπου  $k$  ισούται με τη διάμετρο του δικτύου), τότε είναι προφανές ότι ο κόμβος θα έχει γνώση για τους δείκτες  $NI$  όλων των κόμβων του δικτύου.

Η αποστολή των αιτήσεων για δεδομένα πραγματοποιείται μέσω ενός συνηθισμένου πρωτοκόλλου δρομολόγησης για δίκτυα αισθητήρων, όπως είναι το Ad-hoc On-demand Distance Vector (AODV) [104]. Ένας κόμβος τοποθετεί πάντα στην cache ένα δεδομένο που έχει ζητήσει, ενώ γνωρίζει επίσης την εναπομείνουσα ενέργεια στην μπαταρία του και τον ελεύθερο χώρο που έχει στην cache. Κάθε κόμβος αποθηκεύει τα εξής μεταδεδομένα που σχετίζονται με ένα πολυμεσικό αντικείμενο που βρίσκεται στην cache:

- Το αναγνωριστικό του αντικειμένου (dataID) και το ίδιο το αντικείμενο,
- Το μέγεθος του αντικειμένου ( $s_i$ ),
- Το χρονικό όριο ζωής (TTL) του αντικειμένου,
- για κάθε αντικείμενο τις χρονικές στιγμές των  $K$  πιο πρόσφατων προσβάσεων. Συνήθως,  $K=2$  ή  $3$ .

- κάθε αντικείμενο χαρακτηρίζεται είτε ως  $O$  (δηλαδή: own, ίδιο) είτε ως  $H$  (δηλαδή: hosted, φιλοξενούμενο). Εάν ένα  $H$  αντικείμενο ζητηθεί από τον ίδιο τον κόμβο που το έχει στην cache, τότε η κατάστασή του μεταβάλλεται σε  $O$ .

Όταν ένας κόμβος αποκτά το πολυμεσικό δεδομένο που είχε ζητήσει, τότε το τοποθετεί στην cache και εκπέμπει ένα μικρό πακέτο με το αναγνωριστικό του δεδομένου, το χρονικό όριο ζωής του δεδομένου, την εναπομείνασα ενέργειά του και τον ελεύθερο χώρο στην cache. Οι κόμβοι διαμεσολαβητές που είναι 1-hop γείτονες του κόμβου αποθηκεύουν την πληροφορία που μεταδόθηκε. Είναι χρήσιμο να επισημάνουμε στο σημείο αυτό ότι το σύνολο των διαμεσολαβητών περιλαμβάνει τους κόμβους που έχει επιλέξει ο συγκεκριμένος κόμβος, καθώς επίσης και κάθε άλλο διαμεσολαβητή που έχει επιλεγεί από γειτονικούς κόμβους. Εν κατακλείδι, κάθε διαμεσολαβητής αποθηκεύει την εναπομείνασα ενέργεια και τον ελεύθερο χώρο στη cache για κάθε ένα από τους 1-hop γείτονες του, και για κάθε ένα αναγνωριστικό αντικειμένου που έχει ακούσει από τη διαδικασία εκπομπής, το χρονικό όριο ζωής και τους κόμβους που έχουν αποθηκεύσει το δεδομένο. Στο υπόλοιπο του κεφαλαίου χρησιμοποιούμε ένα σύνολο μεταβλητών και συμβόλων που παρουσιάζονται στον Πίνακα 3.1.

Μεταβλητή	Περιγραφή
$d_i$	Αντικείμενο $i$ ( $1 \leq i \leq 1000$ )
$\text{request}(d_i)$	Αίτηση για το αντικείμενο $i$
$N_i$	Κόμβος $i$
FS	Ελεύθερος χώρος στην cache
RE	Εναπομείνασα ενέργεια
PCT	Πίνακας Εγγύτητας
ipacket	Πακέτο με το αναγνωριστικό του $d_i$ , FS και RE

Πίνακας 3.1. Λίστα συμβόλων πρωτοκόλλου NiCoCa.

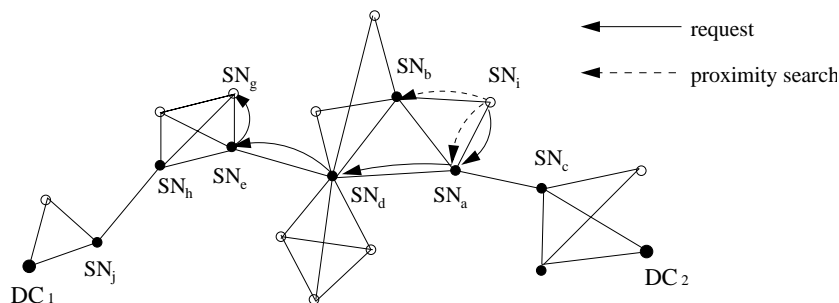
### 3.3.3 Συνιστώσα ανεύρεσης cache

Όταν ένας κόμβος εκδίδει μια αίτηση για ένα πολυμεσικό αντικείμενο, τότε αρχικά το αναζητά στην τοπική cache. Εάν το αντικείμενο βρεθεί εκεί (local cache hit), τότε ενημερώνεται ο χρόνος της πιο πρόσφατης πρόσβασης του αντικειμένου. Διαφορετικά (local cache miss), η αίτηση εκπέμπεται και λαμβάνεται από τους

διαμεσολαβητές. Εάν δεν μπορεί να εξυπηρετήσει την αίτηση κάποιος από αυτούς (proximity cache miss), τότε η αίτηση προωθείται προς το κέντρο δεδομένων.

Στην περίπτωση που ένας μη 1-hop διαμεσολαβητής λάβει την αίτηση, τότε ερευνά τη δική του cache. Εάν συμπεραίνει ότι η αίτηση μπορεί να ικανοποιηθεί από ένα γειτονικό κόμβο (remote cache hit), τότε σταματά την προώθηση της αίτησης προς το κέντρο δεδομένων, και προωθεί την αίτηση στο γειτονικό κόμβο. Υπάρχει όμως η περίπτωση περισσότεροι του ενός γειτονικοί κόμβοι να μπορούν να εξυπηρετήσουν την αίτηση. Τότε επιλέγεται ο κόμβος με τη μεγαλύτερη εναπομείνουσα ενέργεια, ώστε να επιτύχουμε ισορροπία στην κατανάλωση ενέργειας μεταξύ των γειτονικών κόμβων. Εάν η αίτηση δεν μπορεί να ικανοποιηθεί από το διαμεσολαβητή, τότε αυτός δεν την προωθεί επαναληπτικά στους δικούς του διαμεσολαβητές. Η τεχνική αυτή έχει επιλεγεί γιατί είναι πολύ πιθανόν οι διαμεσολαβητές να έχουν επίσης επιλεγεί και από το πρωτόκολλο δρομολόγησης (AODV) και επομένως μπορεί να αποφευχθεί ένας μεγάλος αριθμός μηνυμάτων. Συνεπώς, κατά τη διάρκεια της διαδικασίας προώθησης μιας αίτησης προς το κέντρο δεδομένων η αναζήτηση πραγματοποιείται μόνο σε εκείνους τους κόμβους που συμπεριλαμβάνονται στο μονοπάτι για το κέντρο δεδομένων.

Λαμβάνοντας υπόψη την ιδέα αυτή, περιγράφουμε το πρωτόκολλο ανεύρεσης cache και τη συνεργασία μεταξύ των κόμβων, έτσι ώστε να καθορισθεί το μονοπάτι προς το κόμβο με το ζητούμενο αντικείμενο ή προς το κέντρο δεδομένων. Για παράδειγμα, ας υποθέσουμε ότι ο κόμβος  $SN_i$  στο Σχήμα 3.2 δημοσιοποιεί μια αίτηση για ένα αντικείμενο  $x$  που βρίσκεται αποθηκευμένο στο κέντρο δεδομένων  $DC_1$  και το οποίο είναι επίσης τοποθετημένο στις caches των κόμβων  $SN_g$  και  $SN_h$ . Οι γραμμοσκιασμένοι κόμβοι θεωρούμε ότι είναι οι διαμεσολαβητές. Αρχικά ο κόμβος κάνει αναζήτηση στη δική του cache. Εάν διαπιστώσει ότι το



Σχήμα 3.2. Αίτηση από τον κόμβο  $SN_i$  προωθείται στον caching κόμβο  $SN_g$ .

αντικείμενο δεν είναι διαθέσιμο στην τοπική cache, τότε στέλνει μια αίτηση αναζήτησης εγγύτητας (proximity search request) στους γειτονικούς διαμεσολαβητές  $SN_a$  και  $SN_b$ . Με τη λήψη της αίτησης αναζήτησης, κάθε διαμεσολαβητής πραγματοποιεί αναζήτηση στον πίνακα εγγύτητας (proximity cache table). Εάν το αντικείμενο βρεθεί, κάθε διαμεσολαβητής απαντά με ένα πακέτο δείκτη, το οποίο περιέχει το αναγνωριστικό του αντικειμένου και την εναπομείνουσα ενέργεια του κόμβου με τη μεγαλύτερη ισχύ μπαταρίας και το αντικείμενο αποθηκευμένο. Ο κόμβος  $SN_i$  επιλέγει τον κόμβο που έχει καταναλώσει τη μικρότερη ενέργεια και του στέλνει ένα πακέτο αίτησης. Ο caching κόμβος ανταποκρίνεται με ένα πακέτο απάντησης, το οποίο περιλαμβάνει το ζητούμενο αντικείμενο.

Όπως φαίνεται στο Σχήμα 3.2, εάν το αντικείμενο δεν βρίσκεται αποθηκευμένο σε κάποιο γειτονικό κόμβο, τότε ο κόμβος  $SN_i$  αποστέλλει το πακέτο αίτησης προς το κέντρο δεδομένων. Όταν ο κόμβος  $SN_x$  (όπου  $x \in \{d, e\}$ ) λάβει ένα πακέτο αίτησης, πραγματοποιεί αναζήτηση στην τοπική cache και στον πίνακα εγγύτητας. Αν το αντικείμενο δεν βρεθεί, τότε ο κόμβος  $SN_x$  προωθεί την αίτηση μέσω του μονοπατιού στο  $DC_1$ . Όταν ο κόμβος  $SN_e$  ανακαλύψει ότι το αντικείμενο βρίσκεται στις caches κάποιων γειτονικών κόμβων, τότε επιλέγει τον κόμβο που έχει ξοδέψει τη μικρότερη ποσότητα ενέργειας και ανακατευθύνει το πακέτο αίτησης προς τον caching κόμβο. Ο caching κόμβος στέλνει ένα πακέτο απάντησης, όπου περιλαμβάνεται το αντικείμενο  $x$  κατά μήκος του μονοπατιού δρομολόγησης μέχρι αυτό να φθάσει στον αρχικό κόμβο. Όταν ο κόμβος λάβει το αντικείμενο, ενημερώνει τους 1-hop διαμεσολαβητές σχετικά με το νέο αντικείμενο που έχει εισέλθει στην cache. Αυτό πραγματοποιείται με την αποστολή ενός πακέτου που περιλαμβάνει το αναγνωριστικό του αντικειμένου. Τέλος, αν δεν υπάρχει αρκετός αποθηκευτικός χώρος, τότε καλείται το πρωτόκολλο αντικατάστασης της cache για να καθορισθούν τα δεδομένα που θα απομακρυνθούν από την cache. Ο ψευδοκώδικας για τον αλγόριθμο αναζήτησης παρουσιάζεται στο Σχήμα 3.3.

Για κάθε αίτηση που δημοσιοποιείται, προκύπτει μία από τις εξής τέσσερις περιπτώσεις:

1. Local hit (LH): Το ζητούμενο αντικείμενο βρίσκεται στην cache του κόμβου που δημοσιοποιεί την αίτηση. Εάν το αντικείμενο είναι έγκυρο (το χρονικό όριο ζωής δεν έχει παρέλθει), τότε δεν εκτελείται το πρωτόκολλο NiCoCa.
2. Proximity hit (PH): Το ζητούμενο αντικείμενο είναι τοποθετημένο στην cache ενός κόμβου που βρίσκεται στη 2-hop γειτονιά του κόμβου που το χρειάζεται. Σε αυτή την περίπτωση οι διαμεσολαβητές επιστρέφουν στον



**(A) Αλγόριθμος αναζήτησης cache**

```

if ( $d_i$  is in local cache of requester node) then
    send ipacket to CHs;
    return;
if (requester node is CH and  $d_i$ 's id in PCT) then
    select caching node with largest RE;
    send request( $d_i$ ) to caching node;
else
    requester node sends request( $d_i$ ) to CHs;
    when CHs answers or time elapsed
    if (caching nodes found) then
        select caching node with largest RE;
        send request( $d_i$ ) to caching node;
    else
        send request( $d_i$ ) to data center;
    when  $N_i$  receives request( $d_i$ )
    if ( $N_i$  has a valid copy) then
        send  $d_i$  to requester node;
    else if ( $N_i$  is CH and  $d_i$ 's id in PCT) then
        select caching node with largest RE;
        redirect request( $d_i$ ) to caching node;
    else
        forward request( $d_i$ ) to caching node;

```

Σχήμα 3.3. Αλγόριθμος αναζήτησης πρωτοκόλλου NiCoCa.

αιτούντα κόμβο την τοποθεσία του κόμβου που έχει αποθηκευμένο το αντικείμενο.

3. Remote hit (RH): Το ζητούμενο αντικείμενο βρίσκεται στην cache ενός κόμβου που έχει τουλάχιστον ένα διαμεσολαβητή επί του μονοπατιού που συνδέει τον αιτούντα κόμβο με το κέντρο δεδομένων.
4. Global hit (GH): Το ζητούμενο αντικείμενο αποκτάται από το κέντρο δεδομένων.

### 3.3.4 Συνιστώσα αντικατάστασης της cache

Αν και η χωρητικότητα της cache των αισθητήρων μπορεί να είναι της τάξης των GB (π.χ., NAND flash), είναι αναγκαία η ανάπτυξη μιας αποτελεσματικής και αποδοτικής πολιτικής αντικατάστασης για την επιτυχή αντιμετώπιση του υπερβολικού μέγεθους των πολυμεσικών δεδομένων που δημιουργούνται σε ένα ΑΠΔΑ. Το πρωτόκολλο NiCoCa συνίσταται στην ακόλουθη πολιτική που αποτελείται από τέσσερα βήματα:

- ΒΗΜΑ 1.: Σε περίπτωση ανάγκης, ένας κόμβος εκκαθαρίζει πρώτα από την cache τα αντικείμενα που έχει αποθηκεύσει για λογαριασμό άλλων κόμβων και στην συνέχεια απομακρύνει άλλα αντικείμενα, αν απαιτείται. Κάθε cached αντικείμενο χαρακτηρίζεται είτε ως O είτε ως H. Σε περίπτωση ενός local hit, η κατάσταση ενός H αντικειμένου μεταβάλλεται σε O. Εάν ο διαθέσιμος χώρος εξακολουθεί να είναι μικρότερος από αυτόν που απαιτείται, τότε εκτελείται το Βήμα 2.
- ΒΗΜΑ 2.: Υπολόγισε την ακόλουθη συνάρτηση για κάθε cached αντικείμενο  $i$ :  $cost(i) = \frac{s_i}{TTL_i} * \frac{now - tk - th\_access}{K}$ . Το υποψήφιο αντικείμενο προς

#### (B) Πολιτική αντικατάστασης

**while** (current node has not enough FS )

Select a valid  $d_i$  with largest

value and store it temporary;

Send to CHs  $d_i$ 's id;

Remove the valid  $d_i$ ;

when a CH gets  $d_i$ 's id

**if** (CH gets  $d_i$ 's id and  $d_i$ 's id not in PCT) **then**

select caching node with largest RE and FS;

send answer to requester node;

when current node get answers from CHs

**foreach** (temporary stored  $d_i$ )

**if** (there is no other caching node) **then**

Select caching node with least RE and largest FS;

Send  $d_i$  to new caching node;

Remove temporary stored  $d_i$ ;

Σχήμα 3.4. Πολιτική αντικατάστασης πρωτοκόλλου NiCoCa.

αντικατάσταση είναι αυτό με το μεγαλύτερο κόστος.

- **ΒΗΜΑ 3.:** Πληροφόρησε τους διαμεσολαβητές σχετικά με τα υποψήφια προς αντικατάσταση αντικείμενα. Εάν βρίσκονται αποθηκευμένα σε κάποιον άλλο γειτονικό κόμβο, τότε πληροφορία επιστρέφεται στον κόμβο που εκτελεί τη διαδικασία αντικατάστασης και το αντικείμενο διαγράφεται από την cache. Εάν το αντικείμενο δεν βρίσκεται κάπου αλλού αποθηκευμένο, τότε απομακρύνεται από την cache του αρχικού κόμβου και προωθείται στον κόμβο με τη μεγαλύτερη εναπομείνουσα ενέργεια. Σε κάθε περίπτωση, οι διαμεσολαβητές ενημερώνουν τα αποθηκευμένα μεταδεδομένα που διατηρούν σχετικά με τη νέα κατάσταση.
- **ΒΗΜΑ 4.:** Ο κόμβος που δέχεται στην cache το δεδομένο που απομακρύνθηκε, πληροφορεί τους διαμεσολαβητές με τη συνηθισμένη διαδικασία εκπομπής και το αντικείμενο χαρακτηρίζεται ως H.

Ο ψευδοκώδικας για την πολιτική αντικατάστασης και την πολιτική αποδοχής στην cache παρουσιάζονται στο Σχήματα 3.4 και 3.5 αντίστοιχα.

#### (C) Πολιτική αποδοχής στην cache

when the packet with  $d_i$  obtained from current node

**if** (current node is packet's destination) **then**

**if** (there is enough FS) **then**

cache  $d_i$ ;

send ipacket to CHs;

**else**

call Replacement Policy ;

when CH gets an ipacket

**if** (CH get ipacket) **then**

store  $d_i$ 's id, RE and FS in PCT;

Σχήμα 3.5. Πολιτική αποδοχής πρωτοκόλλου NiCoCa.

### 3.4 Πειραματική Αποτίμηση

Η απόδοση του πρωτοκόλλου NiCoCa αποτιμήθηκε μέσω πειραματικών προσομοιώσεων. Στην παρούσα παράγραφο παρουσιάζεται ένας μεγάλος αριθμός πειρα-

μάτων με διάφορες παραμέτρους και μια σύγκριση μεταξύ του πρωτοκόλλου NiCoCa και της επικρατούσας τεχνικής HybridCache στα πλαίσια των MANET [149].

### 3.4.1 Μοντέλο προσομοίωσης

Αρχικά αναπτύχθηκε ένα μοντέλο προσομοίωσης με βάση το δικτυακό προσομοιωτή J-Sim [123]. Το πρωτόκολλο AODV [104] χρησιμοποιήθηκε στις πειραματικές προσομοιώσεις για να δρομολογεί την κίνηση των δεδομένων στο ΑΔΑ. Επίσης, επιλέχθηκε το πρωτόκολλο IEEE 802.11 ως πρωτόκολλο πρόσβασης του μέσου MAC και το μοντέλο ελεύθερου χώρου (free space model) ως μοντέλο ασύρματης διασποράς. Το εύρος ζώνης του ασύρματου δικτύου ήταν 2 Mbps.

Τα πρωτόκολλο δοκιμάστηκε σε μια μεγάλη ποικιλία από τοπολογίες δικτύων, έτσι ώστε να επιτευχθεί η προσομοίωση των δικτύων αισθητήρων με διαφορετικού βαθμού πυκνότητα κόμβων, από 4 μέχρι 10. Τα πειράματα διεξήχθησαν επίσης με διαφορετικό πλήθος κόμβων, το οποίο ποικίλλει μεταξύ 100 και 1000. Επιπρόσθετα, πειραματικά αποτιμήσαμε την απόδοση των πρωτοκόλλων χρησιμοποιώντας δύο διαφορετικά σύνολα δεδομένων. Κάθε δεδομένο έχει μέγεθος που ποικίλει ομοιόμορφα μεταξύ 1KB και 10KB για το πρώτο σύνολο, και από 1MB μέχρι 5MB για το δεύτερο.

Η δικτυακή τοπολογία αποτελείται από πολλά τετράγωνα τμήματα πλέγματος, όπου τοποθετούνται ένας ή περισσότεροι κόμβοι. Ο αριθμός των τετράγωνων τμημάτων του πλέγματος εξαρτάται από τον αριθμό των κόμβων και την πυκνότητα τοποθέτησης τους. Οι τοπολογίες δημιουργούνται ως εξής: η τοποθεσία κάθε ενός από τους  $n$  κόμβους κατανέμεται ομοιόμορφα μεταξύ των σημείων  $(x=0, y=0)$  και  $(x=500, y=500)$ . Η μέση πυκνότητα  $d$  υπολογίζεται μέσω της ταξινόμησης όλων των  $n \times (n - 1)/2$  ακμών του δικτύου με βάση το μήκος τους σε αύξουσα σειρά. Το μέγεθος κάθε τμήματος του πλέγματος, το οποίο αντιστοιχεί στην τιμή του  $d$ , ισούται με την τετραγωνική ρίζα του 2 πολλαπλασιαζόμενη με το μήκος της ακμής στη θέση  $n \times d/2$  της ταξινομημένης ακολουθίας. Δύο κόμβοι θεωρούνται γείτονες όταν τοποθετούνται μέσα στο ίδιο τμήμα του πλέγματος ή σε γειτονικά. Η περιοχή προσομοίωσης έχει μέγεθος  $500m \times 500m$  και διαιρείται σε ίσου μεγέθους τετραγωνικά τμήματα πλέγματος. Ξεκινώντας από το κάτω αριστερά άκρο, ονομάζουμε τα τμήματα κατά στήλες ως  $1, 2, \dots$

Το μοντέλο ερωτημάτων το οποίο υιοθετεί ένας κόμβος είναι όμοιο με προηγούμενες μελέτες [149]. Κάθε αισθητήρας δημιουργεί ερωτήσεις “read-only”. Εάν ένας κόμβος δεν λάβει απάντηση σε ένα ερώτημα που έχει αποστείλει, τότε περιμένει για ένα χρονικό διάστημα ( $tw$ ) πριν ξαναστείλει ένα νέο ερώτημα. Το πρότυπο

πρόσβασης των κόμβων είναι ανεξάρτητο από την τοποθεσία. Αυτό σημαίνει ότι οι κόμβοι αποφασίζουν ανεξάρτητα για τα δεδομένα που τους ενδιαφέρουν. Κάθε κόμβος δημιουργεί ερωτήματα για δεδομένα ακολουθώντας μια ομοιόμορφη κατανομή. Επιπλέον, ως πρότυπο πρόσβασης χρησιμοποιήθηκε και η Zipfian κατανομή με  $\theta=0.8$ . Σύμφωνα με τη Zipfian κατανομή, ομάδες κόμβων που βρίσκονται σε γειτονικά τμήματα του πλέγματος (25 τμήματα με μέγεθος  $100\text{m} \times 100\text{m}$ ) ακολουθούν το ίδιο πρότυπο πρόσβασης. Εξετάσαμε την απόδοση των πρωτοκόλλων και για τα δύο πρότυπα πρόσβασης. Στην περίπτωση του προτύπου πρόσβασης με βάση τη Zipfian κατανομή, πραγματοποιήσαμε πειράματα με διάφορες τιμές  $\theta$  μεταξύ 0.0 και 1.0.

Παρόμοια με τη μελέτη [149], δύο πηγές δεδομένων τοποθετούνται στις αντίθετες γωνίες της περιοχής προσομοίωσης. Η πηγή δεδομένων 1 τοποθετείται στο σημείο  $(x=0, y=0)$ , ενώ η πηγή δεδομένων 2 τοποθετείται στο σημείο  $(x=500, y=500)$ . Παράλληλα υπάρχουν  $N/2$  δεδομένα σε κάθε πηγή δεδομένων. Τα δεδομένα με άρτιο αναγνωριστικό τοποθετούνται στην πηγή δεδομένων 1, ενώ τα δεδομένα με περιττό στην πηγή 2. Το μέγεθος κάθε δεδομένου κατανέμεται ομοιόμορφα μεταξύ  $s_{min}$  και  $s_{max}$ . Θεωρούμε επίσης ότι τα δεδομένα δεν ενημερώνονται. Τέλος οι πηγές δεδομένων εξυπηρετούν τα ερωτήματα με βάση την πολιτική FCFS (first-come-first-served). Οι παράμετροι της προσομοίωσης παρουσιάζονται στον Πίνακα 3.2.

Παράμετροι	Εξ ορισμού	Εμβέλεια τιμών
Πλήθος αντικειμένων ( $N$ )	1000	
$S_{min}$ (KB)	1	
$S_{max}$ (KB)	10	
$S_{min}$ (MB)	1	
$S_{max}$ (MB)	5	
Πλήθος κόμβων ( $n$ )	500	100-1000
Εύρος ζώνης (Mbps)	2	
Διάστημα αναμονής ( $t_w$ )	10 sec για KB αντικείμενα 100 sec για MB αντικείμενα	
Μέγεθος cache (KB)	800	200-1200
Μέγεθος cache (MB)	125	25-250
Zipfian κύρτωση ( $\theta$ )	0.8	0.0-1.0

Πίνακας 3.2. Παράμετροι προσομοίωσης.

### 3.4.2 Μετρικές απόδοσης

Οι μετρημένες ποσότητες περιλαμβάνουν τον αριθμό των hits (local, remote και global), τη μέση καθυστέρηση πρόσβασης για την απόκτηση των ζητούμενων αντικειμένων και το πλήθος των ανταλασσόμενων μηνυμάτων. Είναι φανερό ότι ένας μικρός αριθμός global hits συνεπάγεται μικρότερη συμφόρηση του δικτύου και συνεπώς λιγότερες συγκρούσεις και απώλειες πακέτων. Επιπλέον, ένας μεγάλος αριθμός remote hits αποδεικνύει την αποτελεσματικότητα της συνεργασίας στη μείωση του αριθμού των global hits. Ένας μεγάλος αριθμός local hits δεν συνεπάγεται μια αποτελεσματική πολιτική συνεργατικού caching, εκτός αν συνδυάζεται με ένα μικρό αριθμό global hits. Αυτό συμβαίνει γιατί το κόστος των global hits περιορίζει σημαντικά τα οφέλη που προκύπτουν από τα local hits.

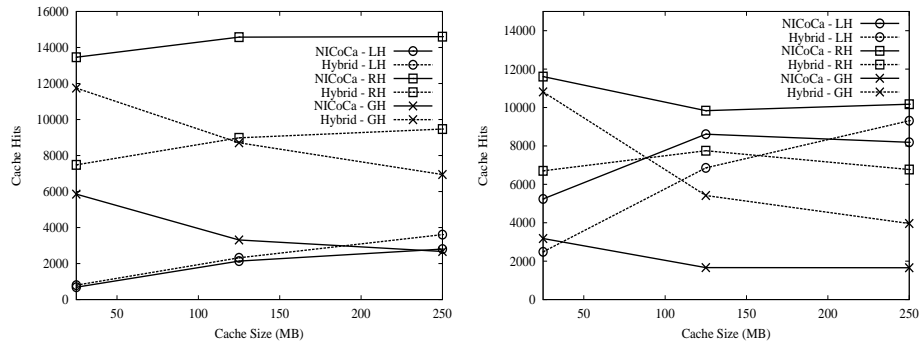
### 3.4.3 Πειραματικά αποτελέσματα

Εκτελέσαμε ένα μεγάλο αριθμό πειραμάτων μεταβάλλοντας το μέγεθος του δικτύου, τον τρόπο προσπέλασης από τους κόμβους και το μέγεθος της cache. Πιο συγκεκριμένα, πραγματοποιήσαμε πειράματα με 100, 500 και 1000 αισθητήρες, για μέγεθος cache ίσο με το 1%, 5% και 10% του συνολικού μεγέθους όλων των πολυμεσικών δεδομένων, για πρότυπα προσπέλασης όπου το  $\theta$  λαμβάνει τιμές από 0.0 (ομοιόμορφο πρότυπο πρόσβασης) μέχρι 1.0 (πολύ ασύμμετρο πρότυπο πρόσβασης), για μέσο όρο πυκνότητας των αισθητήρων ίσο με 4, 7 (πολύ αραιό και αραιό δίκτυο αισθητήρων) και 10 (πυκνό δίκτυο αισθητήρων) και τέλος για δεδομένα με μέγεθος από λίγα KB μέχρι λίγα MB. Για κάθε διαφορετικό σύνολο τιμών των παραμέτρων μετρήσαμε τον αριθμό των hits (local, remote, global), την καθυστέρηση πρόσβασης και το πλήθος των μηνυμάτων. Στην συνέχεια θα παρουσιάσουμε ένα αντιπροσωπευτικό σύνολο αποτελεσμάτων, επειδή υπάρχουν πολλές ανεξάρτητες παράμετροι και τρεις εξαρτημένες μετρικές απόδοσης. Τα πειραματικά αποτελέσματα χωρίζονται σε δύο μεγάλες κατηγορίες. Η πρώτη αποτελείται από πειράματα με δεδομένα μεγέθους λίγων MB, ενώ η δεύτερη αποτελείται από πειράματα με δεδομένα μεγέθους λίγων KB.

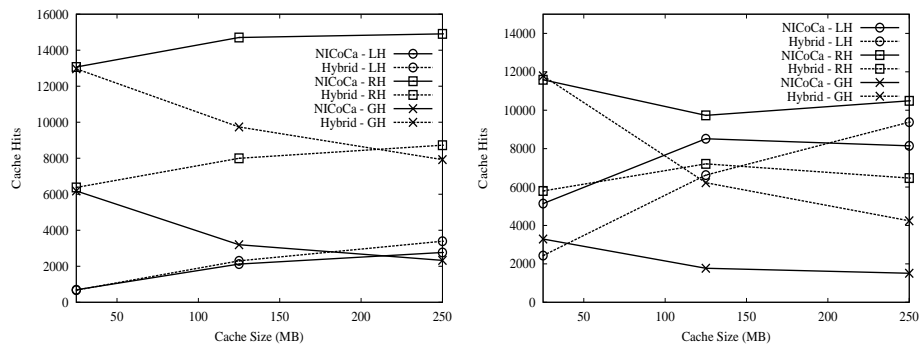
#### 3.4.3.1 Πειράματα με δεδομένα μεγέθους MB

Ο σκοπός των πειραμάτων είναι να εξετάσουμε την απόδοση των πρωτοκόλλων caching στο χειρισμό μεγάλων πολυμεσικών δεδομένων, όπως για παράδειγμα είναι τα αρχεία video, ήχου και εικόνας.

Τα Σχήματα 3.6-3.7 δείχνουν τον αριθμό των hits που επιτυγχάνεται από



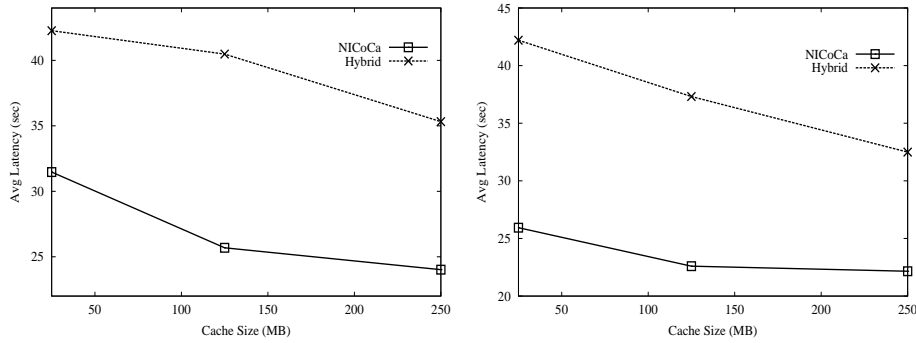
Σχήμα 3.6: Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB,  $\theta=0.0$  και  $0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) με 100 αισθητήρες.



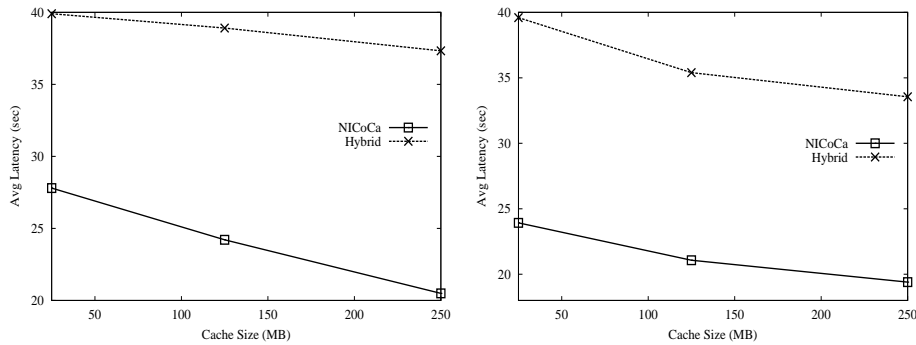
Σχήμα 3.7: Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB,  $\theta=0.0$  και  $0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 100 αισθητήρες.

τα δύο πρωτόκολλα για ένα μικρό δίκτυο (αραιό και πυκνό, αντίστοιχα) στην περίπτωση ομοιόμορφου και ασύμμετρου μοντέλου προσπέλασης. Η πρώτη παρατήρηση είναι ότι ο αριθμός των local hits αυξάνει και για τα δύο πρωτόκολλα όσο το μοντέλο προσπέλασης των δεδομένων γίνεται περισσότερο κυρτό. Το ενδιαφέρον είναι ότι αν και για το ομοιόμορφο μοντέλο προσπέλασης το πρωτόκολλο HybridCache είναι ελαφρώς καλύτερο από το πρωτόκολλο NiCoCa στην περίπτωση των local hits, ωστόσο η κατάσταση αντιστρέφεται όταν τα ερωτήματα αφορούν σε μικρότερο αριθμό δεδομένων. Αυτό δείχνει την καλύτερη απόδοση της πολιτικής αντικατάστασης και αποδοχής του πρωτοκόλλου NiCoCa. Λαμβάνοντας υπόψη τον αριθμό των global hits, το πρωτόκολλο NiCoCa πετυχαίνει το μισό αριθμό από global hits σε σχέση με το πρωτόκολλο HybridCache, ενώ η διαφορά στην απόδοση αυξάνεται σε περισσότερο πυκνά δίκτυα. Ειδικότερα, το πρωτόκολλο NiCoCa διατηρεί σχεδόν αμετάβλητο τον αριθμό των global hits.

Η εξήγηση είναι η σχετική απόδοση των δύο αλγορίθμων, και πιο συγκεκριμένα τα remote cache hits. Για πιο αραιές τοπολογίες το πρωτόκολλο NiCoCa είναι δύο φορές καλύτερο από το πρωτόκολλο HybridCache, ενώ η διαφορά γίνεται εμφανέστερη για πυκνότερα δίκτυα. Επομένως, το προτεινόμενο πρωτόκολλο αποδεικνύεται ότι είναι αποδοτικότερο επειδή προσπαθεί να αξιοποιήσει την τοπολογία του δικτύου. Τα αποτελέσματα αυτά αντανακλούν στο χρόνο προσπέλασης που παράγεται από τα δύο πρωτόκολλα (δες τα Σχήματα 3.8-3.9).



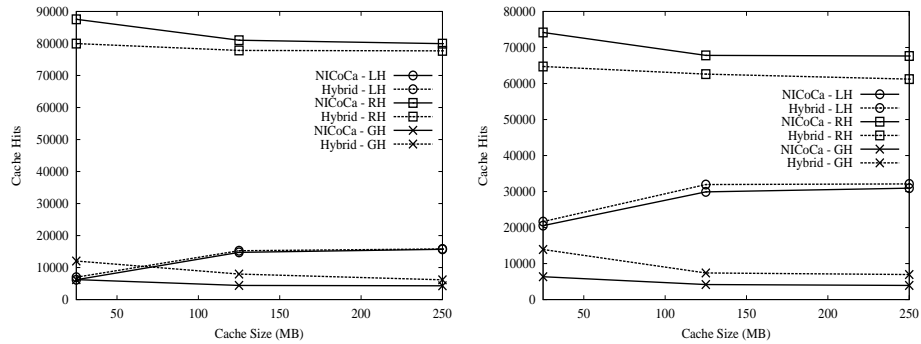
Σχήμα 3.8: Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB,  $\theta=0.0$  και 0.8) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) με 100 αισθητήρες.



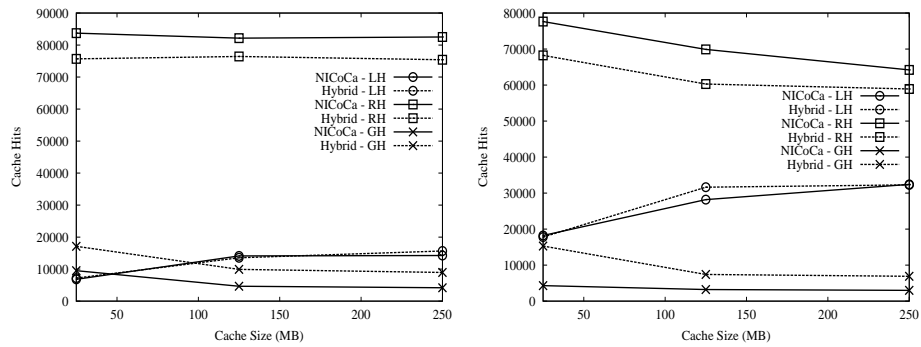
Σχήμα 3.9: Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB,  $\theta=0.0$  και 0.8) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 100 αισθητήρες.

Σε μεγαλύτερα δίκτυα, για παράδειγμα με 500 κόμβους (Σχήματα 3.10-3.11), η υπεροχή του πρωτοκόλλου NiCoCa ως προς τα hits είναι ακόμα εμφανέστερη, αλλά τα αποτελέσματα δεν είναι τόσο εντυπωσιακά, αν περιορίζουμε τη διασπορά πολλών κόμβων μέσα στην ίδια γεωγραφική περιοχή, και δημιουργούμε πολύ περισσότερα αντίγραφα δεδομένων. Η υπεροχή αντανακλάται επίσης και στο χρόνο καθυστέρησης κατά την προσπέλαση των δεδομένων, όπως παρουσιάζεται στα





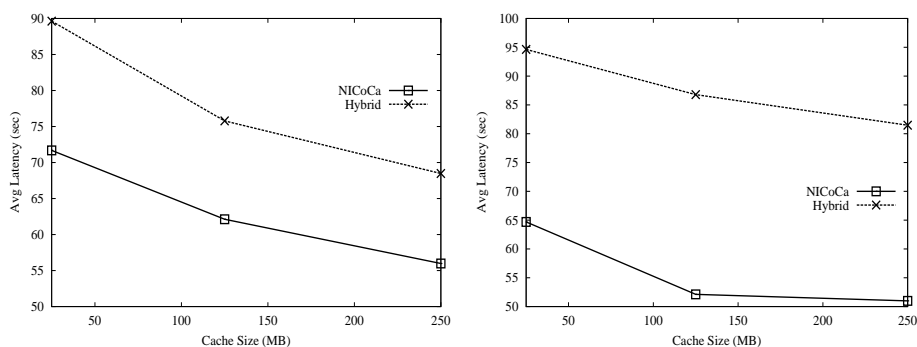
Σχήμα 3.10: Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB,  $\theta=0.0$  και  $0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) με 500 αισθητήρες.



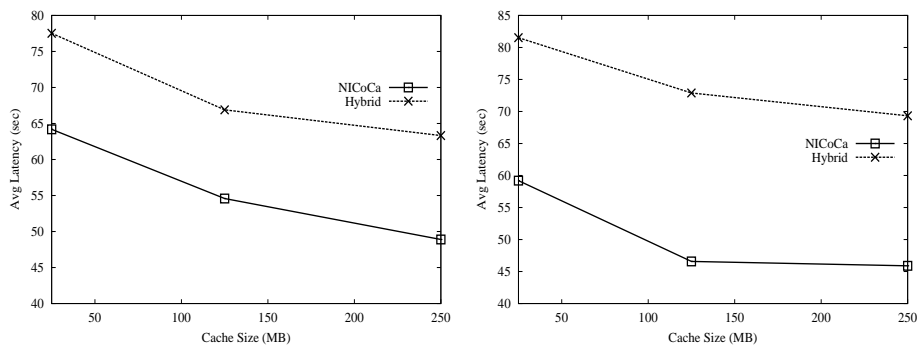
Σχήμα 3.11: Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB,  $\theta=0.0$  και  $0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 500 αισθητήρες.

Σχήματα 3.12-3.13. Επιπλέον παρατηρούμε ότι όπως μετακινούμαστε σε μεγαλύτερα δίκτυα η καθυστέρηση προσπέλασης επίσης αυξάνεται. Αυτό συμβαίνει επειδή τα πυκνότερα δίκτυα (περισσότεροι κόμβοι στην ίδια περιοχή) έχουν αρνητική επίδραση στην αποδοτικότητα της επικοινωνίας καθώς λαμβάνουν χώρα περισσότερες συγκρούσεις, ανταγωνισμός για την απόκτηση του ασύρματου μέσου και διαγραφή πακέτων.

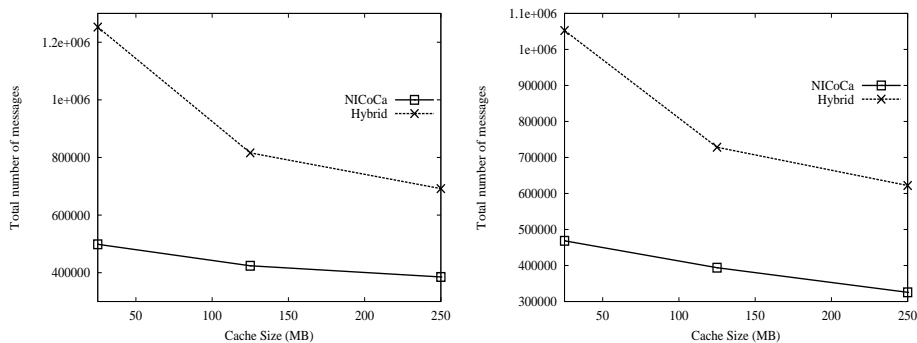
Σε αυτό το σημείο είναι ενδιαφέρον να σημειώσουμε το συνολικό αριθμό των ανταλασσόμενων μηνυμάτων μεταξύ των κόμβων του δικτύου, ο οποίος είναι επίσης μια μετρική που μπορεί να εξηγήσει τη συνολική ενέργεια που καταναλώνεται στο δίκτυο (Σχήμα 3.14). Όσο περισσότερα μηνύματα εκπέμπει και λαμβάνει ένας αισθητήρας, τόσο περισσότερη ενέργεια καταναλώνει. Για πυκνές τοπολογίες δικτύου με ομοιόμορφο και ασύμμετρο πρότυπο προσπέλασης, το πρωτόκολλο NiCoCa στέλνει τα μισά μηνύματα σε σχέση με τα μηνύματα



Σχήμα 3.12: Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB,  $\theta=0.0$  και 0.8) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) με 500 αισθητήρες.



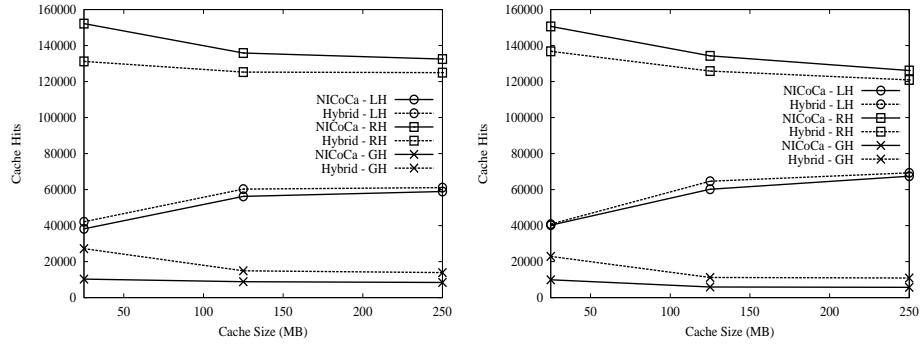
Σχήμα 3.13: Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB,  $\theta=0.0$  και 0.8) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 500 αισθητήρες.



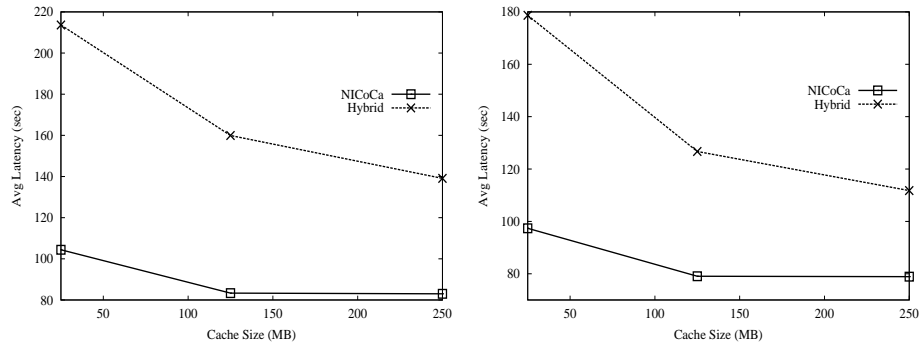
Σχήμα 3.14: Επίδραση μεγέθους της cache στον αριθμό των μηνυμάτων (δεδομένα μεγέθους MB,  $\theta=0.0$  και 0.8) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 500 αισθητήρες.

που διαχέονται στο δίκτυο από το πρωτόκολλο HybridCache. Τα αποτελέσματα βελτιώνονται περισσότερο και το πλεονέκτημα του πρωτοκόλλου NiCoCa γίνεται

μεγαλύτερο όταν το πρότυπο προσπέλασης γίνεται περισσότερο ασύμμετρο. Τέλος, τα ίδια συμπεράσματα επιβεβαιώνονται και για μεγαλύτερα δίκτυα, τα οποία περιλαμβάνουν 1000 αισθητήρες (Σχήματα 3.15-3.17).

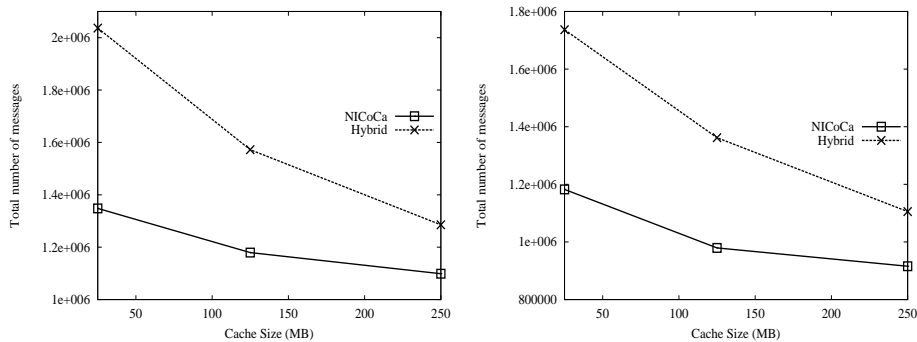


Σχήμα 3.15: Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB,  $\theta=0.8$ ) σε ένα αραιό ( $d=7$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 1000 αισθητήρες.



Σχήμα 3.16: Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB,  $\theta=0.8$ ) σε ένα αραιό ( $d=7$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 1000 αισθητήρες.

Γενικά, για όλες τις τοπολογίες δικτύου το πρωτόκολλο NiCoCa πετυχαίνει περισσότερα remote hits και λιγότερα global hits από το πρωτόκολλο Hybrid-Cache. Η διαφορά στην απόδοση μεγαλώνει υπέρ του πρωτοκόλλου NiCoCa όταν εξετάζουμε πυκνότερα ΑΔΑ. Ειδικότερα στην περίπτωση των πολύ πυκνών τοπολογιών δικτύου, το NiCoCa πετυχαίνει διπλάσιο αριθμό από remote hits συγκρινόμενο με το HybridCache και μόνο τα μισά global hits. Όταν εξετάζουμε τον αριθμό των local hits, τότε παρατηρούμε ότι για αραιά δίκτυα αισθητήρων το πρωτόκολλο HybridCache πετυχαίνει ελαφρώς περισσότερα local hits από το NiCoCa, αλλά αυτή η διαφορά εκλείπει εντελώς σε πυκνότερες τοπολογίες δικτύου. Παράλληλα, το μικρό κέρδος στην απόδοση που έχει το HybridCache



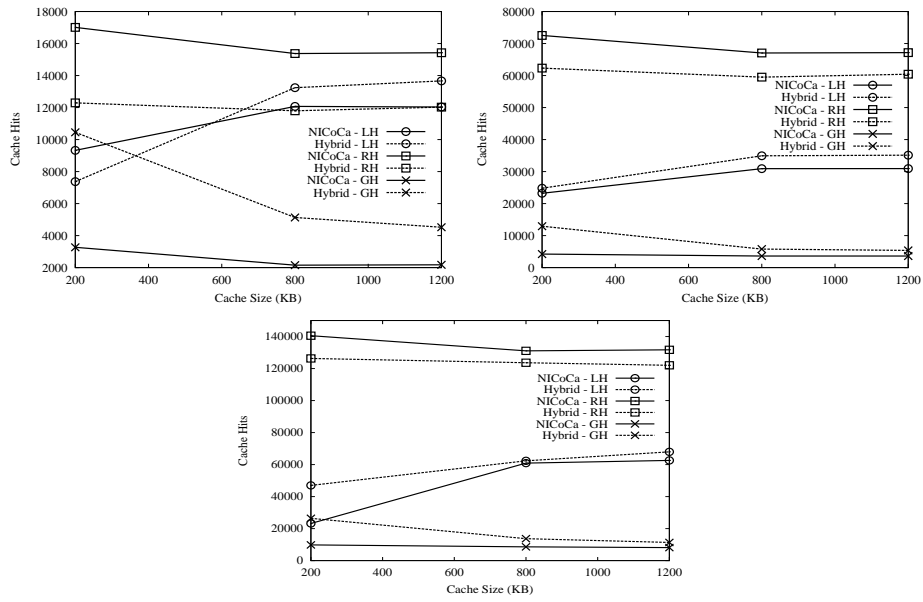
Σχήμα 3.17: Επίδραση μεγέθους της cache στο πλήθος των μηνυμάτων (δεδομένα μεγέθους MB,  $\theta=0.8$ ) σε ένα αραιό ( $d=7$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 1000 αισθητήρες.

δεν αποτελεί πλεονέκτημα, επειδή την ίδια στιγμή εμφανίζει διπλάσιο αριθμό από global hits.

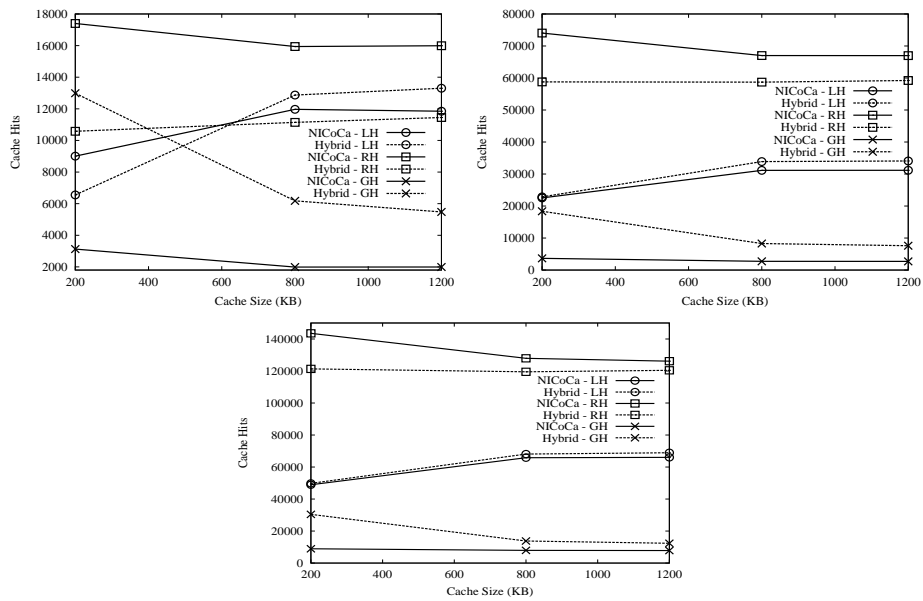
### 3.4.3.2 Πειράματα με δεδομένα μεγέθους KB

Μια σημαντική ερώτηση είναι κατά πόσο αυτά τα αποτελέσματα διατηρούνται όταν τα δεδομένα είναι μικρότερου μεγέθους, π.χ. λίγων KB. Παρότι αναμένουμε ότι τα ΑΠΔΑ θα διαχειρίζονται αρχεία εικόνας και video, είναι πιθανότατο τα ΑΠΔΑ να διαχειρίζονται ταυτόχρονα και πολύ μικρότερα δεδομένα. Για να ερευνησουμε την απόδοση των πρωτοκόλλων συνεργατικού caching στην περίπτωση αυτή, πραγματοποιούμε το ίδιο σύνολο πειραμάτων αλλά για δεδομένα μεγέθους KB και παρουσιάζουμε στη συνέχεια ένα υποσύνολο των αποτελεσμάτων.

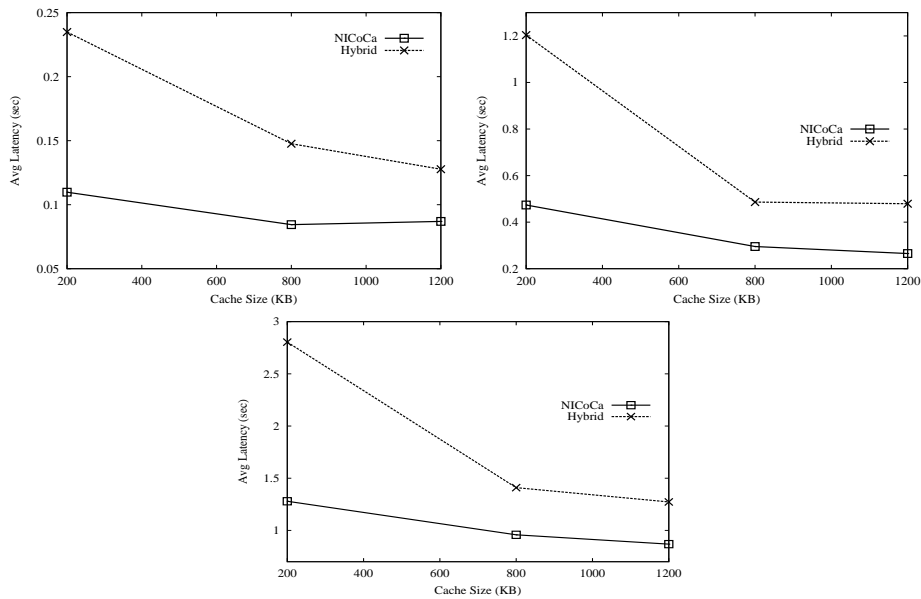
Οι γενικές παρατηρήσεις που καταγράψαμε για τα πειράματα με δεδομένα μεγέθους MB, εξακολουθούν να ισχύουν και σε αυτή την περίπτωση. Το πρωτόκολλο NiCoCa πετυχαίνει σημαντικά μικρότερο αριθμό global hits και μεγαλύτερο αριθμό remote hits από το πρωτόκολλο HybridCache. Δεν αξίζει να σχολιάσουμε κάθε γράφημα (Σχήματα 3.18-3.21), καθώς το NiCoCa υπερισχύει σε όλες τις περιπτώσεις. Το νέο πρωτόκολλο πετυχαίνει 25% περισσότερα remote hits και 50% λιγότερα global hits.



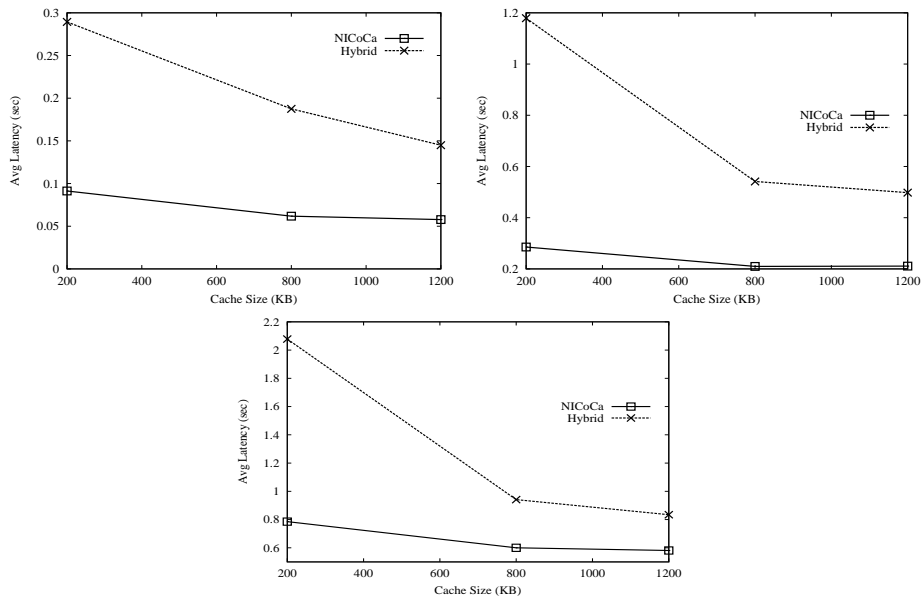
Σχήμα 3.18: Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους KB,  $\theta=0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) για πλήθος αισητηρών 100, 500 και 1000.



Σχήμα 3.19: Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους KB,  $\theta=0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) για πλήθος αισητηρών 100, 500 και 1000.



Σχήμα 3.20: Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους KB,  $\theta=0.8$ ) σε ένα αραιό ΑΠΔΑ ( $d=7$ ) για πλήθος αισθητήρων 100, 500 και 1000.



Σχήμα 3.21: Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους KB,  $\theta=0.8$ ) σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) για πλήθος αισθητήρων 100, 500 και 1000.

### 3.5 Συμπεράσματα

Στο κεφάλαιο αυτό εξετάσαμε το πρόβλημα του συνεργατικού caching στην περιοχή των ΑΠΔΑ. Οι πρόσφατες εξελίξεις στη σμίκρυνση των ηλεκτρονικών κυκλωμάτων και στη δημιουργία κυκλωμάτων χαμηλής κατανάλωσης ενέργειας, σε συνδυασμό με την ανάπτυξη φθηνών καμερών CMOS και μικροφώνων έδωσαν ζωή στην περιοχή αυτή. Τα ΑΠΔΑ θα δώσουν νέα ώθηση και εντυπωσιακές δυνατότητες στα ΑΔΑ με αποτέλεσμα να δημιουργηθούν πολλές νέες εφαρμογές.

Η σχεδίαση πρωτοκόλλων πρέπει να λάβει υπόψη τα μοναδικά χαρακτηριστικά των ΑΠΔΑ και κυρίως να παρέχει Ποιότητα Υπηρεσίας σε επίπεδο εφαρμογής. Το θέμα αυτό έχει αγνοηθεί στα παραδοσιακά ΑΔΑ. Στην προσπάθεια επίτευξης του στόχου, αναπτύξαμε ένα πρωτόκολλο συνεργατικού caching, το οποίο ονομάζεται NiCoCa, και είναι κατάλληλο για ΑΠΔΑ. Το πρωτόκολλο ανιχνεύει τους κόμβους που βρίσκονται στα κεντρικότερα σημεία των γειτονιών του δικτύου και τους αναθέτει το ρόλο του διαμεσολαβητή για να συντονίσουν τις αποφάσεις caching. Το προτεινόμενο πρωτόκολλο αποτιμήθηκε με το δικτυακό προσομοιωτή J-Sim και η απόδοσή του συγκρίθηκε με την καλύτερη μέθοδο συνεργατικού caching από την περιοχή των MANET. Τα παραγόμενα αποτελέσματα αποδεικνύουν την ανωτερότητα του προτεινόμενου πρωτοκόλλου που μπορεί να μειώσει τα global hits κατά ένα μέσο ποσοστό 50% και να αυξήσει τα remote hits εξαιτίας της αποδοτικής συνεργασίας των αισθητήρων κατά ένα μέσο ποσοστό 40%.

Συνοψίζοντας, το NiCoCa είναι ένα πρακτικό, εύρωστο και αποδοτικό πρωτόκολλο συνεργατικού caching για τα ΑΠΔΑ.





## ΚΕΦΑΛΑΙΟ 4

# Συνέπεια της Cache σε Ασύρματα Πολυμεσικά Δίκτυα Αισθητήρων \*

### Περιεχόμενα

---

4.1	ΕΙΣΑΓΩΓΗ . . . . .	69
4.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	72
4.3	ΜΟΝΤΕΛΟ ΔΙΚΤΥΟΥ . . . . .	76
4.4	ΠΡΩΤΟΚΟΛΛΟ ΣΥΝΕΠΕΙΑΣ NICC . . . . .	77
4.5	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	86
4.6	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	106

---

### 4.1 Εισαγωγή

Στο Κεφάλαιο 3 αναφέραμε την απαρχή των ΑΠΔΑ που επιτεύχθηκε χάρις στη δημιουργία κυκλωμάτων χαμηλής κατανάλωσης ενέργειας, σε συνδυασμό με την κατασκευή φθηνών CMOS καμερών και μικροφώνων, τα οποία μπορούν να αποκτήσουν πλούσιο πολυμεσικό περιεχόμενο από το περιβάλλον.

---

\*Το κεφάλαιο αντλεί τα περιεχόμενά του από την εργασία [47].

Θεωρώντας ότι κάθε κόμβος έχει ένα τοπικό αποθηκευτικό χώρο, μπορούμε να διεκπεραιώσουμε τις προκλήσεις που τίθενται στη σχεδίαση και υλοποίηση των ΑΠΔΑ με την αξιοποίηση της τεχνικής του συνεργατικού caching, όπως περιγράψαμε στο Κεφάλαιο 3. Στο συνεργατικό caching, πολλοί κόμβοι διαμοιράζουν και συντονίζουν τα δεδομένα της cache χωρίς να χρειάζεται πάντα να επισκέπτονται τα κέντρα δεδομένων. Η διάρκεια ζωής της μπαταρίας μπορεί να επεκταθεί αν καταφέρουμε να μειώσουμε τον όγκο της πραγματοποιούμενης επικοινωνίας μεταξύ των αισθητήρων. Η τεχνική του caching των χρήσιμων δεδομένων για κάθε κόμβο, είτε στον τοπικό αποθηκευτικό χώρο είτε στους γειτονικούς κόμβους, μπορεί να επιμηκύνει τη διάρκεια ζωής του δικτύου, να μειώσει το υψηλό κόστος επικοινωνίας και να μειώσει το φόρτο εργασίας των πηγών των δεδομένων. Επιπρόσθετα, το caching μπορεί να είναι πολύ αποδοτικό στη μείωση των μεταδόσεων μηνυμάτων σε μεγάλο εύρος του δικτύου. Επομένως μπορεί να μειώσει την παρεμβολή και να ανταπεξέλθει στις συνθήκες μεταβλητού καναλιού. Το συνεργατικό caching μπορεί επίσης να επιταχύνει την επεξεργασία των πολυμεσικών δεδομένων, επειδή η επεξεργασία και η παράδοση των δεδομένων δεν είναι ανεξάρτητες οντότητες. Συνεπώς, το συνεργατικό caching έχει ως αποτέλεσμα τη μείωση της καθυστέρησης πρόσβασης, της κατανάλωσης ενέργειας και της απώλειας πακέτων.

Παρ' όλα αυτά μαζί με το caching εμφανίζεται το πρόβλημα συνέπειας της cache. Πολλές πολιτικές συνέπειας της cache έχουν προταθεί στη βιβλιογραφία των Βάσεων Δεδομένων, των Λειτουργικών Συστημάτων, των κυψελοειδών ασύρματων δικτύων και των MANET (δες τις σχετικές εργασίες στην Παράγραφο 4.2). Καμία όμως από αυτές δεν ανταποκρίνεται στις προκλήσεις των ΑΠΔΑ. Στο κεφάλαιο αυτό, προτείνουμε μια νέα πολιτική συνέπειας της cache σε περιβάλλον συνεργατικού caching, ώστε να ανταποκριθούμε στις προηγούμενες προκλήσεις. Παραδείγματα σεναρίων είναι τα εξής:

- Σε ένα πεδίο μάχης οι αισθητήρες διασκορπίζονται σε μια μεγάλη περιοχή και κάθε κόμβος είναι εξοπλισμένος με μια συσκευή επικοινωνίας και μια μικροκάμερα για τη λήψη φωτογραφιών της τοπικής περιοχής. Οι αισθητήρες ενημερώνουν για τις φωτογραφίες που λαμβάνουν και τις διαμοιράζονται μεταξύ τους ώστε να δημιουργήσουν μια καταλληλότερη όψη της παρακολουθούμενης περιοχής. Αυτό συμβαίνει επειδή κάθε μικροκάμερα μπορεί λάβει μια περιορισμένη όψη της περιοχής, είτε εξαιτίας της θέσης του αισθητήρα, είτε εξαιτίας εμποδίων που υπάρχουν κοντά του. Παρ' όλα αυτά, κάθε αισθητήρας ζητά και λαμβάνει ένα μεγάλο αριθμό φωτογραφιών

που έχουν ληφθεί από άλλους κόμβους μέσω multi-hop επικοινωνίας και κατασκευάζει μια μεγάλη και ακριβή εικόνα της παρακολουθούμενης περιοχής. Κάθε φωτογραφία έχει ένα μοναδικό αναγνωριστικό που αποτελείται από την περιοχή, τον αισθητήρα, και το χρόνο λήψης της φωτογραφίας. Σύμφωνα με αυτά τα στοιχεία, ένας αισθητήρας μπορεί να ζητήσει μια φωτογραφία που τον ενδιαφέρει και να συμπεραίνει σχετικά με την εγκυρότητά της. Επομένως, κάθε στρατιωτική μονάδα μπορεί να επικοινωνήσει με κοντινούς κόμβους για να αποκτήσει μια ακριβή όψη της περιοχής στόχου. Αυτή η λειτουργία είναι κρίσιμη επειδή κάθε στρατιωτική μονάδα μπορεί να παρακολουθεί περισσότερο αποδοτικά την αντίστοιχη περίμετρο.

- Τα ΑΠΔΑ μπορούν επίσης να χρησιμοποιηθούν στην παρακολούθηση του περιβάλλοντος. Ένας μεγάλος αριθμός αισθητήρων μπορεί να αναπτυχθεί σε μια δασική περιοχή για να εμποδισθεί μια περιβαλλοντική καταστροφή, όπως είναι η πυρκαγιά και η απόθεση απορριμμάτων. Αυτό μπορεί να επιτευχθεί μέσω φωτογραφιών που λαμβάνουν οι αισθητήρες. Η διασπορά και ο διαμοιρασμός πολυμεσικών δεδομένων μπορεί να προσφέρει μια ακριβή όψη της παρακολουθούμενης περιοχής. Συνεπώς, οι ανθρώπινες ομάδες που παρακολουθούν και προστατεύουν τη δασική περιοχή μπορούν να επικοινωνήσουν με λίγους αισθητήρες για να αποκτήσουν μια λεπτομερή όψη της δασικής περιοχής ευθύνης τους. Αυτό οφείλεται στο γεγονός ότι κάθε αισθητήρας μπορεί να συλλέξει μέσω multi-hop επικοινωνίας τις εικόνες που έχουν ληφθεί από άλλους κόμβους και να δημιουργήσει μια μεγάλη εικόνα της αντίστοιχης υποπεριοχής.

Σε αυτό το κεφάλαιο προτείνουμε έναν καινοτόμο υψηλής απόδοσης πρωτόκολλο συνέπειας της cache για συνεργατικό caching, το οποίο ονομάζεται *Node Importance-based Cache Consistency* (NICC). Το πρωτόκολλο NICC είναι ένας υβριδικός push/pull αλγόριθμος που χρησιμοποιεί το ελάχιστο κόστος ανταλλαγής μηνυμάτων για να διατηρήσει τη συνέπεια στην cache. Με την εισαγωγή των διαμεσολαβητών, η απόδοση της ακύρωσης του caching μπορεί να βελτιωθεί σημαντικά. Επειδή δεν υπάρχει πρότερη εργασία στο πεδίο της εγκυρότητας της cache για ΑΠΔΑ, συγκρίνουμε το προτεινόμενο πρωτόκολλο με την καλύτερη πολιτική συνέπειας της cache από το ερευνητικό πεδίο των MANET. Η υλοποίηση των πρωτοκόλλων και η πειραματική αποτίμηση γίνεται με τη χρήση του περιβάλλοντος προσομοίωσης J-Sim [123]. Στην Παράγραφο 4.5 θα παρουσιάσουμε την πειραματική αξιολόγηση των δύο μεθόδων, η οποία αναδεικνύει την ανωτερότητα του προτεινόμενου πρωτοκόλλου NICC έναντι του ανταγωνιστή του. Στην

Παράγραφο 4.2 θα παρουσιάσουμε τις σπουδαιότερες σχετικές μελέτες.

## 4.2 Σχετικές Εργασίες

Το θέμα της συνέπειας της cache έχει εκτενώς διερευνηθεί στο πεδίο των Βάσεων Δεδομένων, των Λειτουργικών Συστημάτων, τον Παγκόσμιο Ιστό και στα ασύρματα κυβελοειδή (1-hop) δίκτυα. Η μελέτη [22] περιέχει μία εξαιρετική κατηγοριοποίηση των δημοφιλέστερων τεχνικών που έχουν αναπτυχθεί για τη διατήρηση της εγχυρότητας και συνέπειας της cache στα προηγούμενα περιβάλλοντα. Γενικότερα, υπάρχουν δύο γενικές πολιτικές για τη γνωστοποίηση σχετικά με τις ενημερώσεις που έχουν γίνει στα δεδομένα από τους κόμβους πηγές. Η πρώτη πολιτική περιλαμβάνει τη διάχυση (push) των αλλαγών μέσω ακυρωτικών μηνυμάτων. Σύμφωνα με τη δεύτερη πολιτική, οι κόμβοι που έχουν δεδομένα στην cache θα στείλουν ερωτήματα στους κόμβους πηγές (pull) για τα δεδομένα που τους ενδιαφέρουν.

Στις push-based στρατηγικές ο κόμβος πηγή πρέπει συνεχώς ή περιοδικά να στέλνει τα δεδομένα που έχει σε κάθε κόμβο στην cache. Οι μεταβολές που πραγματοποιούνται σε ένα δεδομένο προωθούνται σε ολόκληρο το δίκτυο μέσω της τεχνικής της πλημμυρίδας (flooding). Κάθε κόμβος που λαμβάνει ένα ακυρωτικό μήνυμα θα πρέπει να ελέγξει την cache και να ακυρώσει αν χρειάζεται τα κατάλληλα δεδομένα. Το πλεονέκτημα της πολιτικής push είναι οι καλές εγγυήσεις συνέπειας που προσφέρει και η απλότητά του. Το σχήμα όμως υποφέρει από περιττές μεταδόσεις ενημερώσεων, επειδή οι προηγούμενοι ενδιαφερόμενοι κόμβοι δεν συνεπάγεται ότι συνεχίζουν να ενδιαφέρονται για τα ίδια δεδομένα. Επίσης αυξάνει το πλήθος των μηνυμάτων ελέγχου εξαιτίας των διαδοχικών εκπομπών των ακυρωτικών μηνυμάτων και συνεπάγεται επεξεργασία των μηνυμάτων σε κάθε κόμβο.

Στις pull-based στρατηγικές, όταν ένας κόμβος χρειάζεται δεδομένα, τότε στέλνει ένα μήνυμα στον εξυπηρετητή και ο εξυπηρετητής απαντά με την αποστολή των δεδομένων. Έτσι, οι cache κόμβοι είναι υπεύθυνοι για τη διατήρηση της συνέπειας των δεδομένων που βρίσκονται στην cache. Το πλεονέκτημα του σχήματος είναι οι εγγυήσεις ισχυρής συνέπειας και είναι κατάλληλο για δυναμικά δίκτυα όπου υπάρχουν συχνές αποσυνδέσεις και συνδέσεις κόμβων. Παρ' όλα αυτά, το σχήμα υποφέρει από μεγάλο αριθμό μηνυμάτων polling, τα οποία διαχέονται στο δίκτυο. Αυτό συμβαίνει γιατί οι cache κόμβοι δεν γνωρίζουν πότε ή πόσο συχνά λαμβάνουν χώρα οι ενημερώσεις από τους κόμβους πηγές.

Επιπλέον, η καθυστέρηση πρόσβασης αυξάνεται καθώς ο χρόνος απόκρισης περιλαμβάνει μια επιπρόσθετη καθυστέρηση, η οποία οφείλεται στα ερωτήματα που πρέπει να θέσουν οι cache κόμβοι προς τους κόμβους πηγές. Εξαιτίας του υψηλού κόστους επικοινωνίας, η κατανάλωση ενέργειας αυξάνεται. Επομένως, μια υβριδική πολιτική μεταξύ των δύο θα είναι η ιδανική λύση. Κάποιες προσπάθειες για ανάπτυξη της υβριδικής πολιτικής έχουν προταθεί στη βιβλιογραφία [65].

Πράγματι, οι μελέτες [21, 118] προτείνουν μια τέτοια υβριδική πολιτική για MANET, και αποτελούν τις πιο σχετιζόμενες ερευνητικές εργασίες στο πρωτόκολλο που προτείνουμε. Το σχήμα Proximity Regions for Caching in Cooperative MP2P (PReCinCt) διαιρεί το δίκτυο σε γεωγραφικές περιοχές, όπου κάθε μία ευθύνεται για ένα σύνολο κλειδιών που απεικονίζουν δεδομένα [118]. Κάθε peer χρησιμοποιεί έναν πίνακα περιοχής όπου αποθηκεύει πληροφορίες σχετικά με τις τοποθεσίες όλων των περιοχών. Επιπλέον, κάθε κλειδί αντιστοιχίζεται σε μια περιοχή βάσει μιας γεωγραφικής συνάρτησης κατακερματισμού. Το πρωτόκολλο χρησιμοποιεί ένα πρωτόκολλο δρομολόγησης, το οποίο βασίζεται στην τοποθεσία για να δρομολογήσει την ανάκτηση των δεδομένων. Το PReCinCt ενσωματώνει ένα σχήμα συνεργατικού caching ώστε να μειώσει το χρησιμοποιούμενο εύρος ζώνης. Το σχήμα συνεργατικού caching αποθηκεύει τα σχετικά δεδομένα σε ένα σύνολο από peers μιας περιοχής. Όταν ένας peer ζητά κάποια δεδομένα, τότε αρχικά αποστέλλει το ερώτημα στην περιοχή όπου βρίσκεται. Εάν τα δεδομένα δεν βρεθούν, το ερώτημα προωθείται προς την περιοχή που έχει τα δεδομένα. Ο πρώτος peer μέσα στην περιοχή προορισμού που λαμβάνει το ερώτημα, το προωθεί σε όλους τους peers της περιοχής για να μπορέσει να ανακαλύψει που βρίσκονται αποθηκευμένα τα δεδομένα. Το σχήμα PReCinCt είναι επίσης εφοδιασμένο με ένα μηχανισμό συνέπειας της cache. Το σχήμα caching κατά τη διάρκεια αντικατάστασης της cache λαμβάνει υπόψη τη δημοτικότητα των δεδομένων, το μέγεθος και την απόσταση της περιοχής προορισμού. Παράλληλα το PReCinCt χρησιμοποιεί ένα υβριδικό μηχανισμό push/pull για να διατηρήσει τη συνέπεια των δεδομένων μεταξύ των αντιγράφων που βρίσκονται στο δίκτυο. Ωστόσο, η χρησιμοποιούμενη τεχνική πλημμυρίδας έχει ως αποτέλεσμα την αύξηση της κίνησης στο δίκτυο και το άδειασμα της μπαταρίας, ιδιαίτερα όταν οι περιοχές είναι μεγάλες και πυκνές σε κόμβους. Το επιπλέον κόστος επικοινωνίας δημιουργείται επίσης όταν θα πρέπει η τοποθεσία κάθε περιοχής να γνωστοποιηθεί σε ολόκληρο το δίκτυο. Ένα ακόμα ελάττωμα του σχήματος PReCinCt είναι ότι κάθε peer είναι γνώστης της ακριβούς γεωγραφικής τοποθεσίας όπου βρίσκεται. Αυτό προϋποθέτει ένα περιορισμό στα ΑΔΑ καθώς κάθε αισθητήρας θα πρέπει να είναι εξοπλισμένος με συσκευή GPS. Τέλος, δεν είναι ξεκάθαρο ποιο είναι

το βέλτιστο πλήθος των περιοχών του δικτύου και το πλήθος των αισθητήρων σε κάθε περιοχή. Συνεπώς, οι περιοχές είναι πολύ δύσκολο να ορισθούν δυναμικά σε ένα πολεμικό πεδίο, όπου οι αισθητήρες διασκορπίζονται σε μια τεράστια περιοχή και δεν είναι εφοδιασμένοι με συσκευή GPS.

Η πολιτική Relay Peer-based Caching Consistency (RPCC) [21] αξιοποιεί τους ενδιαμέσους κόμβους που ονομάζονται αναμεταδότες (relay nodes) για να υλοποιήσει την υβριδική push/pull πολιτική. Με τη βοήθεια των αναμεταδοτών μεταξύ του κόμβου πηγή και των cache κόμβων, μπορούν να εφαρμοσθούν η push-based και η pull-based στρατηγική, γεγονός που έχει ως αποτέλεσμα τη μείωση του κόστους επικοινωνίας και της καθυστέρησης στην ικανοποίηση ενός ερωτήματος. Ο κόμβος πηγή ωθεί τα δεδομένα στους αναμεταδότες και οι cache κόμβοι ανακτούν τα δεδομένα από τους αναμεταδότες. Αυτές οι δύο λειτουργίες μπορούν να γίνουν ασύγχρονα και ταυτόχρονα. Οι peers που μπορούν να λάβουν τα ακυρωτικά μηνύματα που στέλνονται από τους κόμβους πηγές (κάθε κόμβος πηγή στέλνει τα ακυρωτικά μηνύματα σε μια απόσταση μικρότερη από ένα συγκεκριμένο αριθμό από hops) είναι υποψήφιοι να γίνουν αναμεταδότες. Οι συγγραφείς χρησιμοποιούν τρεις παραμέτρους για να ορίσουν το κριτήριο επιλογής των κόμβων που θα έχουν το ρόλο του αναμεταδότη. Η πρώτη παράμετρος είναι ο συντελεστής του ρυθμού προσπέλασης (coefficient of access rate – CAR). Η δεύτερη παράμετρος είναι ο συντελεστής της σταθερότητας ενός peer (coefficient of stability – CS), ενώ η τρίτη παράμετρος είναι ο συντελεστής του επιπέδου ενέργειας ενός κόμβου (coefficient of energy level – CE). Οι συγγραφείς καθορίζουν επίσης ένα κατώφλι για κάθε παράμετρο ( $th_{CAR}$ ,  $th_{CS}$  και  $th_{CE}$ ). Ένας peer θεωρείται υποψήφιος αναμεταδότης εάν ικανοποιεί τη σχέση  $(CAR < th_{CAR}) \wedge (CS > th_{CS}) \wedge (CE > th_{CE})$ . Μόνο αφού λάβει ένα μήνυμα έγκρισης μπορεί ο υποψήφιος αναμεταδότης peer να μεταλλαχθεί σε αναμεταδότη peer. Παρ' όλα αυτά η πολιτική RPCC εμφανίζει κάποια μειονεκτήματα. Η χρήση των κατωφλιών είναι ένα ελάττωμα επειδή είναι αρκετά δύσκολο να ορισθούν τιμές εφαρμόσιμες για όλες τις δικτυακές τοπολογίες, για διαφορετικό πλήθος και για διαφορετικές τιμές πυκνότητας των αισθητήρων. Επιπρόσθετα, ένα επιπλέον επικοινωνιακό κόστος δημιουργείται κατά τη διάρκεια εκπομπής των ακυρωτικών μηνυμάτων από τους κόμβους πηγές. Αυτό οφείλεται στην τεχνική της πλημμυρίδας. Συνεπώς στην περίπτωση μιας πυκνής δικτυακής τοπολογίας, ο αριθμός των μηνυμάτων που διαδίδονται στο δίκτυο είναι σημαντικά αυξημένος.

Ωστόσο, καμία από τις προαναφερθείσες πολιτικές δεν λαμβάνει υπόψη το κριτήριο της ελαχιστοποίησης της χρονικής καθυστέρησης προσπέλασης των δεδομένων, η οποία είναι ζωτικής σημασίας για τα ΑΠΔΑ. Επιπλέον, και οι δύο πο-

λιτικές κατά τη σχεδίαση της στρατηγικής συνέπειας της cache λαμβάνουν ειδική μέριμνα για την κίνηση που εκτελούν οι κόμβοι, κάτι που ελλείπει εντελώς από τα ΑΠΔΑ. Ο στόχος μας είναι παρεμφερής με αυτόν της πολιτικής RPCC, αλλά το πρωτόκολλο NICC είναι περισσότερο αποδοτικό ως προς το κόστος επικοινωνίας, τη διασπάθιση της ενέργειας και τη χρονική καθυστέρηση στην απάντηση ερωτημάτων.

Πολύ στενά σχετιζόμενες με την παρούσα εργασία είναι τα πρωτόκολλα συνεργασίας cache που πραγματοποιήθηκαν στο πεδίο των ΑΔΑ [43] και των MANET [62, 117, 149, 86, 34, 24], αλλά αυτά τα πρωτόκολλα δεν περιλαμβάνουν κάποιο μηχανισμό διατήρησης της εγκυρότητας της cache. Στη μελέτη [97] οι συγγραφείς παρουσιάζουν έναν αλγόριθμο τοποθέτησης αναμεταδοτών σε περιοχές του δικτύου όπου παρατηρείται συμφόρηση. Λιγότερο σχετικές με τη παρούσα εργασία είναι οι πολιτικές τοποθέτησης της cache που αναπτύχθηκαν στα ΑΔΑ [105, 129] και τα MANET [130], και οι αρχιτεκτονικές και τα πρωτόκολλα που αναπτύχθηκαν για βασισμένα στο διαδίκτυο ασύρματα ad hoc και mesh δίκτυα [88, 37].

#### 4.2.1 Κίνητρο και Συνεισφορές

Αντλώντας κίνητρο από τις αδυναμίες των υπαρχόντων πρωτοκόλλων συνέπειας της cache και συνεργατικού caching και λαμβάνοντας υπόψη τα μοναδικά χαρακτηριστικά των ΑΠΔΑ, προτείνουμε μια συνεργατική πολιτική συνέπειας της cache, η οποία ονομάζεται NICC. Το πρωτόκολλο NICC βασίζεται στην ιδέα της αξιοποίησης της τοπολογίας του δικτύου, έτσι ώστε να βρεθούν οι περισσότερο σημαντικοί κόμβοι. Οι σπουδαιότερες συνεισφορές του πρωτοκόλλου συνοψίζονται στα εξής:

- Το πρωτόκολλο συνεργατική συνέπειας της cache διατηρεί την εγκυρότητα των δεδομένων σε ΑΠΔΑ με τη χρήση μιας μετρικής που αναδεικνύει τη σπουδαιότητα ενός αισθητήρα σε μια τοπολογία δικτύου. Η ανακάλυψη των σπουδαίων κόμβων συνεπάγεται τη μείωση του χρόνου καθυστέρησης κατά την ανάκτηση των δεδομένων.
- Περιγραφή του πρωτοκόλλου συνεργατικής συνέπειας της cache.
- Ανάπτυξη του αλγορίθμου για διασπορά των ενημερώσεων στους caching κόμβους και για αντικατάσταση των λιγότερο πολύτιμων δεδομένων.
- Πειραματική αποτίμηση του πρωτοκόλλου και σύγκρισή του με την καλύτερη πολιτική συνέπειας της cache από το ερευνητικό πεδίο των MANET

με τη χρήση του πακέτου προσομοίωσης (J-Sim).

### 4.3 Μοντέλο Δικτύου

Ένα ΑΠΔΑ αποτελείται από μια συλλογή ασύρματα συνδεδεμένων αισθητήρων που αναπτύσσονται σε μια μεγάλη περιοχή. Στα ΑΠΔΑ δεν υπάρχει η έννοια του εξυπηρετητή που διατηρεί τη συνέπεια των δεδομένων. Κάθε κόμβος έχει ένα μοναδικό χαρακτηριστικό  $SN_i$ , όπου  $1 \leq i \leq l$  και  $l$  ο συνολικός αριθμός αισθητήρων. Παρομοίως, κάθε δεδομένο έχει επίσης ένα μοναδικό αναγνωριστικό  $D_i$ , όπου  $1 \leq i \leq m$  και  $m$  ο συνολικός αριθμός των δεδομένων. Κάθε δεδομένο ανατίθεται σε ένα κόμβο. Έτσι, κάθε κόμβος αποτελεί την πηγή δεδομένων ενός μόνο αντικειμένου (για παράδειγμα, κάθε αισθητήρας με τη χρήση μιας μικροκάμερας μπορεί να συλλάβει μια φωτογραφία μιας συγκεκριμένης τοποθεσίας). Συνακόλουθα υποθέτουμε ότι η πηγή δεδομένων του αντικειμένου  $D_i$  είναι ο κόμβος  $SN_i$ . Διάφορα αντίγραφα του ίδιου αντικειμένου μπορούν να είναι διασκορπισμένα στο δίκτυο και να έχουν αποθηκευθεί είτε από διαμεσολαβητές είτε από cache κόμβους. Κάθε κόμβος έχει περιορισμένο χώρο στην cache, έτσι ώστε μόνο ένα μικρό μέρος του  $D$  να μπορεί να τοποθετηθεί στην cache. Εάν ο χώρος της cache είναι πλήρης και ο κόμβος πρέπει να αποθηκεύσει ένα νέο αντικείμενο, τότε πρέπει να επιλέξει έναν αριθμό από αντικείμενα από τη cache για αντικατάσταση. Οι μεταβολές σε ένα αντικείμενο πραγματοποιούνται μόνο από τον κόμβο πηγή που έχει στην κατοχή του το γνήσιο αντίγραφο του αντικειμένου. Συνεπώς, μόνο ο κόμβος πηγή έχει την πιο ενημερωμένη έκδοση του αντικειμένου. Στη συνέχεια συνοψίζονται οι παραδοχές μας:

- Οι αισθητήρες είναι στατικοί. Στην πλειοψηφία των εφαρμογών οι αισθητήρες δεν κινούνται.
- Οι σύνδεσμοι είναι διπλής κατεύθυνσης, έτσι εάν ένας κόμβος  $SN_i$  μπορεί να επικοινωνεί με ένα κόμβο  $SN_j$ , τότε ο κόμβος  $SN_i$  είναι επίσης μέσα στην εμβέλεια μετάδοσης του κόμβου  $SN_j$ .
- Οι υπολογιστικές και επικοινωνιακές δυνατότητες είναι ίδιες για όλους τους κόμβους.
- Ένα αντικείμενο ανατίθεται σε κάθε κόμβο και τα δεδομένα δεν έχουν το ίδιο μέγεθος. Έτσι, κάθε κόμβος αποτελεί μια πηγή δεδομένων.
- Κάθε κόμβος μπορεί να τοποθετήσει στην cache έναν αριθμό δεδομένων, ο οποίος ορίζεται από το μέγεθος της cache.



- Κάθε κόμβος υπολογίζει τους σημαντικούς γείτονές του (διαμεσολαβητές). Επομένως, ένας κόμβος μπορεί να χαρακτηρίζεται ως κοινός ή διαμεσολαβητής.
- Κάθε κόμβος χρησιμοποιεί ένα πρωτόκολλο δρομολόγησης (π.χ., AODV) για να αποστέλλει τα ερωτήματα στους κόμβους πηγής.
- Κάθε διαμεσολαβητής για κάθε δεδομένο/αντικείμενο που βρίσκεται στην cache αποθηκεύει τις εξής πληροφορίες:
  - *dataID*, το αναγνωριστικό του αντικειμένου.
  - το αναγνωριστικό του κόμβου πηγή που διατηρεί το γνήσιο αντίγραφο του αντικειμένου.
  - το αντικείμενο.
  - το μέγεθος του αντικειμένου, το οποίο είναι σημαντικό να το γνωρίζει ο διαμεσολαβητής για τη συνάρτηση αντικατάστασης της cache.
  - *TTR* (time to refresh), καθώς χρειάζεται κάθε διαμεσολαβητής να γνωρίζει τη χρονική στιγμή ανανέωσης ενός αντικειμένου, έτσι ώστε να κρίνει την εγκυρότητα της έκδοσης του αντικειμένου που έχει στην cache.
  - τη μέση απόσταση (σε hops) των κόμβων που έχουν ζητήσει το αντικείμενο.
  - το ρυθμό προσπέλασης του αντικειμένου, ο οποίος αναδεικνύει τη συχνότητα αίτησης ενός αντικειμένου.
  - το ρυθμό ενημέρωσης του αντικειμένου, ο οποίος καθορίζει τη συχνότητα μεταβολής του.
  - την έκδοση του αντικειμένου, η οποία ανατίθεται από τον κόμβο πηγή.

Στο υπόλοιπο του κεφαλαίου χρησιμοποιούμε ένα σύνολο χαρακτήρων και συμβόλων, τα οποία παρατίθεται στο Πίνακα 4.1.

## 4.4 Πρωτόκολλο Συνέπειας NICC

Η εκπλήρωση της απαίτησης για μικρή καθυστέρηση πρόσβασης σε ΑΠΔΑ είχε ως αποτέλεσμα να προσπαθήσουμε να αναγνωρίσουμε τους κόμβους που είναι σημαντικότεροι από άλλους ώστε να συντονίζουν τη διαδικασία ελέγχου της συνέπειας της cache και να προωθούν τα μηνύματα ενημερώσεων. Αυτοί οι κόμβοι

Μεταβλητή	Περιγραφή
<i>dataID</i>	Αναγνωριστικό του αντικειμένου
$d_i$	Αντικείμενο $i$
<i>SN</i>	Κόμβος αισθητήρας
<i>MN</i>	Κόμβος διαμεσολαβητής
<i>CN</i>	Κόμβος cache
<i>RN</i>	Αιτών Κόμβος
<i>DC</i>	Κέντρο δεδομένων
<i>src – req</i>	Αναγνωριστικό αιτούντος κόμβου
<i>src – med</i>	Αναγνωριστικό διαμεσολαβητή
<i>TTL</i>	Χρονικό όριο ζωής ενός ακυρωτικού μηνύματος
<i>TTR</i>	Χρονική στιγμή ανανέωσης αντικειμένου στο διαμεσολαβητή
<i>DUI</i>	Χρονικό διάστημα μεταξύ ενημερώσεων στο κόμβο πηγή
<i>CL</i>	Επίπεδο συνέπειας (αδύναμο, $\Delta$ , ισχυρό)
<i>AR</i>	Ρυθμός πρόσβασης ενός αντικειμένου
<i>UR</i>	Ρυθμός ενημέρωσης ενός αντικειμένου
<i>AvgHops</i>	Μέσος όρος από hops μεταξύ του κόμβου που ζητά το αντικείμενο και του κόμβου που απαντά

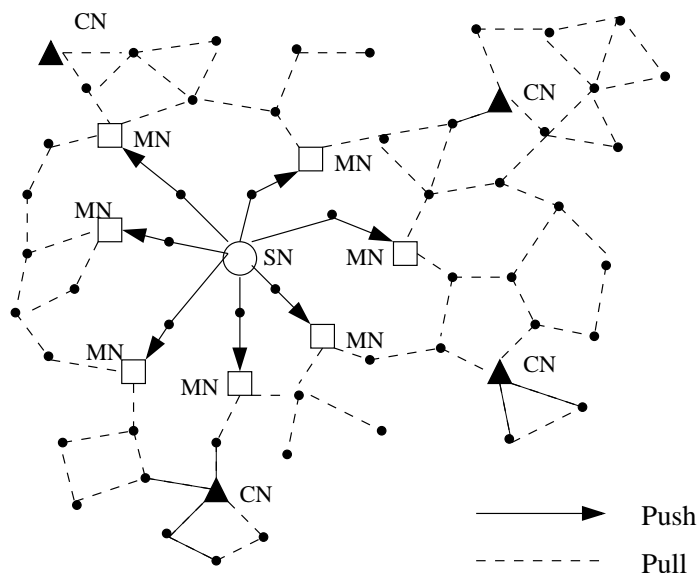
Πίνακας 4.1. Λίστα συμβόλων πρωτοκόλλου NICC.

θα ονομάζονται διαμεσολαβητές και θα παίξουν το ρόλο των αντιπροσώπων των κόμβων πηγών. Επομένως, οι διαμεσολαβητές θα αναλάβουν την επικοινωνία με τους cache κόμβους που έχουν αποθηκευμένα αντίγραφα των δεδομένων.

Οι διαμεσολαβητές αποτελούν το κομβικό σημείο του πρωτοκόλλου συνέπειας. Κάθε αισθητήρας επιλέγει έναν αριθμό διαμεσολαβητών στα κεντρικότερα σημεία της δικτυακής γειτονιάς του αισθητήρα. Η επιλογή των διαμεσολαβητών είναι τοπικού χαρακτήρα με μικρό κόστος επικοινωνίας και πραγματοποιείται στην αρχική φάση της ζωής του δικτύου, επειδή τα δίκτυα αισθητήρων θεωρούνται σχετικά στατικά. Οι διαμεσολαβητές βρίσκονται μεταξύ του κόμβου πηγή και των cache κόμβων και είναι πλησιέστερα στον κόμβο πηγή. Η απόσταση μεταξύ του κόμβου πηγή και των διαμεσολαβητών ορίζεται από την τιμή *TTL*, την οποία αναθέτει ο κόμβος πηγή στα ακυρωτικά μηνύματα. Όταν ένας διαμεσολαβητής λάβει ένα ακυρωτικό μήνυμα και η απόσταση από τον κόμβο πηγή ισούται με την τιμή *TTL*, τότε μπορεί να τοποθετήσει στην cache το αντικείμενο.

Εξαιτίας της εγγύτητας μεταξύ του κόμβου πηγή και των διαμεσολαβητών,

η επικοινωνία είναι push-based επειδή το κόστος επικοινωνίας θα είναι σχετικά χαμηλό. Επομένως, η πηγή και οι διαμεσολαβητές μπορούν να σχηματίσουν ένα επικαλυπτόμενο δίκτυο που θα προσφέρει ένα έγκυρο αντικείμενο στους cache κόμβους. Την ίδια στιγμή η επικοινωνία μεταξύ των διαμεσολαβητών και των cache κόμβων είναι pull-based. Κάθε φορά που φθάνει μια νέα αίτηση για δεδομένα σε ένα cache κόμβο, αυτός αποστέλλει την αίτηση στον κόμβο πηγή για να λάβει έγκυρα δεδομένα. Ένας διαμεσολαβητής όμως μπορεί να απαντήσει στην αίτηση εάν βρίσκεται κατά μήκος της διαδρομής προς την πηγή. Εξαιτίας των λειτουργιών push και pull, οι οποίες μπορούν να γίνουν ασύγχρονα και ταυτόχρονα, τόσο το κόστος επικοινωνίας όσο και η καθυστέρηση πρόσβασης μπορούν να μειωθούν. Το Σχήμα 4.1 απεικονίζει ένα παράδειγμα που παρουσιάζει την προσέγγιση NICC.



Σχήμα 4.1. Παράδειγμα πρωτοκόλλου NICC.

Η σημαντικότητα των κόμβων έχει εξηγηθεί σε βάθος στην Παράγραφο 3.3.1 του Κεφαλαίου 3. Χρησιμοποιώντας τη μέθοδο που περιγράφηκε, μπορούμε να ορίσουμε τη σημαντικότητα των κόμβων στις τοπικές γειτονιές του δικτύου. Οι Παράγραφοι 4.4.1-4.4.3 παρουσιάζουν τα κύρια συστατικά του πρωτοκόλλου NICC.

#### 4.4.1 Προσέγγιση push-pull

Ο κόμβος πηγή ωθεί τα δεδομένα και τα ακυρωτικά μηνύματα σε όλους τους κόμβους που βρίσκονται σε απόσταση  $TTL$ . Κάθε ακυρωτικό μήνυμα περιέχει μια τιμή  $TTL$ , η οποία καθορίζει την απόσταση στην οποία μπορεί να φθάσει το μήνυμα. Η τιμή  $TTL$  αρχικοποιείται από τον κόμβο πηγή και μειώνεται κάθε φορά που το μήνυμα διέρχεται από ένα αισθητήρα. Ένας κόμβος μπορεί να ακούσει ένα ακυρωτικό μήνυμα όταν βρίσκεται σε απόσταση μικρότερη από  $TTL$  hops από τον κόμβο πηγή. Επιπρόσθετα, θεωρούμε ότι οι κόμβοι ανταλλάσσουν μηνύματα “Hello” (beacon μηνύματα), τα οποία περιλαμβάνουν τη λίστα των γειτόνων τους. Επομένως, κάθε κόμβος γνωρίζει την 2-hop γειτονιά του.

Οι διαμεσολαβητές προσδιορίζονται χωριστά από κάθε κόμβο σύμφωνα με την τιμή σπουδαιότητας. Κάθε κόμβος υπολογίζει το δείκτη  $NI$  των 1-hop γειτόνων του. Στην συνέχεια χρησιμοποιεί αυτή την πληροφορία για να χαρακτηρίσει ορισμένους από τους γείτονες ως διαμεσολαβητές. Οι διαμεσολαβητές είναι οι κόμβοι που αποτελούν το ελάχιστο σύνολο των γειτόνων με το μεγαλύτερο δείκτη  $NI$ , οι οποίοι καλύπτουν την 2-hop γειτονιά του κόμβου. Οι διαμεσολαβητές αναλαμβάνουν την επανεκπομπή των ακυρωτικών μηνυμάτων που στέλνουν οι κόμβοι πηγές μέσα σε  $TTL$  απόσταση και χρησιμοποιούν αυτά τα μηνύματα για να διαπιστώσουν εάν τα δεδομένα που έχουν στην cache είναι έγκυρα. Η επανεκπομπή των ακυρωτικών μηνυμάτων συνεισφέρει στη μείωση του κόστους επικοινωνίας, το οποίο συνεπάγεται η push-based στρατηγική. Συνεπώς, ελαχιστοποιείται ο αριθμός των διαδιδόμενων μηνυμάτων όταν γίνεται η μετάδοση των ακυρωτικών μηνυμάτων. Κάθε πηγή στέλνει περιοδικά ακυρωτικά μηνύματα. Η περιοδικότητα καθορίζεται από την παράμετρο *διάστημα ενημέρωσης δεδομένων* (data update interval),  $DUI$ . Τα μηνύματα περιλαμβάνουν το αναγνωριστικό του αντικείμενου, το αντικείμενο (εάν έχει αλλαχθεί διαφορετικά την τιμή null), την τιμή  $TTL$ , και τον αριθμό της έκδοσης. Εάν το αντικείμενο έχει μεταβληθεί κατά τη διάρκεια μιας περιόδου, τότε η πηγή θα εκδώσει ένα μήνυμα ενημέρωσης και θα το στείλει στους διαμεσολαβητές για να τους πληροφορήσει σχετικά με την ενημέρωση. Το μήνυμα ενημέρωσης περιλαμβάνει το αντικείμενο που έχει μεταβληθεί. Για κάθε αντικείμενο, κάθε κόμβος πηγή διατηρεί την εξής πληροφορία:

1. το αναγνωριστικό του αντικείμενου ( $dataID$ ),
2. το αντικείμενο,
3. το μέγεθος,
4. την έκδοση.

Το γεγονός ότι ο κόμβος πηγή καθορίζει μια *TTL* τιμή, δεν συνεπάγεται ότι κάθε ενδιαμέσος κόμβος θα τοποθετήσει στην cache το αντικείμενο. Το αντικείμενο αποθηκεύεται μόνο από τους εκλεγμένους διαμεσολαβητές των οποίων η απόσταση από τον κόμβο πηγή (σε αριθμό hops) ισούται με την τιμή *TTL*.

#### **Ο Αλγόριθμος push**

**if** (*DUI=0*) **then**

*DC* broadcast invalidation message;

renew *DUI*;

**if** (*MN* receives an invalidation message) **then**

**if** (invalidation message is not in local cache) **then**

cache temporary invalidation message;

**if** (*TTL* > 1) **then**

*TTL*=*TTL*-1;

broadcast invalidation message

**else**

**if** (there is enough *FS*) **then**

cache  $d_i$

**else**

call Replacement Policy

**else**

**if** (*TTL*=1) **then**

**if** (local version equals invalidation version) **then**

renew *TTR*

**else**

cache  $d_i$

when *DC* issues an update message

**if** (*MN* receives an update message and *TTL*=1) **then**

update  $d_i$

#### **Πολιτική αντικατάστασης**

**foreach** (cached  $d_i$ )

calculate value( $d_i$ )

Data item with the smaller value is replaced

Επομένως, δημιουργούμε ένα δακτύλιο κόμβων γύρω από τον κόμβο πηγή, όπου ωθούνται τα ακυρωτικά μηνύματα και αποθηκεύονται τα έγκυρα αντίγραφα. Κάθε διαμεσολαβητής μπορεί να τοποθετήσει αντικείμενα από διαφορετικές πηγές. Ο ψευδοκώδικας για τον αλγόριθμο push παρουσιάζεται στο Σχήμα 4.2.

Η επικοινωνία μεταξύ των διαμεσολαβητών και των αντίστοιχων cache κόμβων είναι pull-based. Οι cache κόμβοι βρίσκονται μεταξύ των κόμβων που υπο-

### Ο Αλγόριθμος pull

```

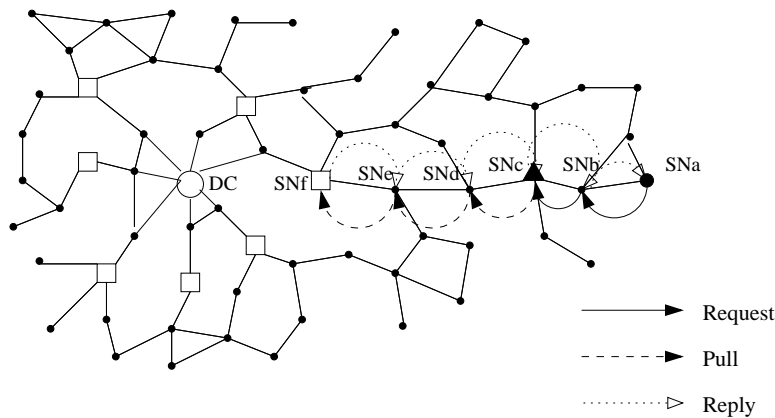
RN requests  $d_i$ 
when CN receives request( $d_i$ )
if ( $d_i$  is in local cache of CN) then
    if (CL=weak) then
        send reply( $d_i$ ) to RN;
    else if (CL=delta and  $D > 0$ ) then
        send reply( $d_i$ ) to RN;
    else
        send request( $d_i$ ) to DC;
when MN receives request ( $d_i$ )
if ( $d_i$  is in local cache of MN) then
    if (TTR > 0) then
        if (MN has a fresher and valid version of  $d_i$ ) then
            send reply( $d_i$ ) to CN;
        if (request's version equals MN's version) then
            send ack( $d_i$ ) to CN;
    else
        forward request( $d_i$ ) to DC
when DC receives request( $d_i$ )
    if (DC has a fresher version of  $d_i$ ) then
        send reply( $d_i$ ) to CN;
    if (request's version equals DC's version) then
        send ack( $d_i$ ) to CN;
when CN receives reply( $d_i$ ) stores new version of  $d_i$  and broadcast reply( $d_i$ ) to RN
when CN receives ack( $d_i$ ) broadcast reply( $d_i$ ) to RN

```

Σχήμα 4.3. Αλγόριθμος pull πρωτοκόλλου NICC.

βάλλουν ερωτήματα για κάποια αντικείμενα και των διαμεσολαβητών. Κάθε φορά που ένας κοινός κόμβος ζητά κάποιο αντικείμενο, το ερώτημά αρχικά κατευθύνεται σε ένα cache κόμβο. Στην περίπτωση που απαιτείται ισχυρή συνέπεια (όχι απλώς αδύναμη ή  $\Delta$  συνέπεια), τότε ο cache κόμβος προωθεί το ερώτημα μέσω της διαδρομής μέχρι να φθάσει σε ένα διαμεσολαβητή, του οποίου η απόσταση από την πηγή ισούται με  $TTL$  hops. Κάθε ερώτημα περιλαμβάνει το αναγνωριστικό του αντικειμένου ( $dataID$ ), την έκδοση και το δικό του αναγνωριστικό ( $src - reqthe$ ). Εάν ο διαμεσολαβητής έχει ένα έγκυρο αντίγραφο και πιο πρόσφατο από την έκδοση που περιλαμβάνεται στο ερώτημα, τότε στέλνει ένα πακέτο απάντησης, το οποίο περιλαμβάνει το αναγνωριστικό του αντικειμένου ( $dataID$ ), το αντικείμενο, την έκδοση και το αναγνωριστικό του ( $src - med$ ). Εάν η έκδοση του ερωτήματος είναι όμοια με αυτή που έχει αποθηκευμένη ο διαμεσολαβητής, τότε αποστέλλεται ένα AckPacket με το αναγνωριστικό του αντικειμένου, την έκδοση και το αναγνωριστικό του διαμεσολαβητή. Στην περίπτωση που δεν υπάρχει έγκυρο αντικείμενο, ο διαμεσολαβητής κρατά το ερώτημα μέχρι να φθάσει το επόμενο ακυρωτικό μήνυμα και στη συνέχεια αποστέλλει την κατάλληλη απάντηση. Ένα αντικείμενο που βρίσκεται στην cache ενός διαμεσολαβητή θεωρείται έγκυρο όταν η τιμή  $TTR$  είναι μεγαλύτερη του μηδενός. Στην περίπτωση όμως που δεν βρεθεί κάποιος διαμεσολαβητής κατά μήκος της διαδρομής προς την πηγή, τότε το ερώτημα φθάνει τελικά στην πηγή που αναλαμβάνει να απαντήσει. Ο κόμβος που ζητά ένα αντικείμενο δεν το αποθηκεύει επειδή δεν είναι σε θέση να λαμβάνει τα ακυρωτικά μηνύματα (θεωρούμε την περίπτωση που ζητείται ισχυρή συνέπεια των δεδομένων – όχι αδύναμη ή  $\Delta$  συνέπεια). Έτσι, κάθε φορά ο κόμβος προσπαθεί να επικοινωνήσει είτε με κάποιο διαμεσολαβητή είτε με τον κόμβο πηγή. Ο ψευδοκώδικας για τον αλγόριθμο push παρουσιάζεται στο Σχήμα 4.3.

Στο σημείο αυτό περιγράφουμε το πρωτόκολλο NICC βασισμένοι στην προηγούμενη ιδέα. Για παράδειγμα, υποθέτουμε ότι ο αισθητήρας  $SN_a$  στο Σχήμα 4.4 εκδίδει μια αίτηση για ένα αντικείμενο  $x$  που βρίσκεται στο κέντρο δεδομένων  $DC$  και έχει προωθηθεί στους διαμεσολαβητές. Η τιμή  $TTL$  θεωρούμε ότι ισούται με 2. Οι κόμβοι με σχήμα τετραγώνου υποθέτουμε ότι είναι οι διαμεσολαβητές του δικτύου αισθητήρων, ενώ οι κόμβοι με σχήμα τριγώνου υποθέτουμε ότι είναι οι cache κόμβοι. Αρχικά ο αισθητήρας  $SN_a$  στέλνει μια αίτηση στο cache κόμβο  $SN_c$ . Λαμβάνοντας την αίτηση ο cache κόμβος ελέγχει το επίπεδο συνέπειας που ζητείται. Εάν απαιτείται  $\Delta$  συνέπεια και η τιμή του  $\Delta$  είναι μηδέν ή απαιτείται ισχυρή συνέπεια, τότε ο cache κόμβος θα προωθήσει την αίτηση προς το κέντρο δεδομένων. Ακολουθώντας τη διαδρομή προς το κόμβο  $DC$  η αίτηση φθάνει στον αισθητήρα  $SN_f$  που είναι διαμεσολαβητής για το ζητούμενο αντικείμενο  $x$ .



Σχήμα 4.4: Παράδειγμα λειτουργίας πρωτοκόλλου NICC όταν απαιτείται ισχυρή συνέπεια ή  $\Delta$  συνέπεια και η τιμή του  $\Delta$  είναι 0.

Εάν ο διαμεσολαβητής έχει ένα έγκυρο και πρόσφατο αντίτυπο του αντικειμένου στην cache, τότε στέλνει ένα μήνυμα απάντησης στον cache κόμβο  $SN_c$  με τη νέα έκδοση του αντικειμένου. Ο cache κόμβος αποθηκεύει τη νέα έκδοση του αντικειμένου  $x$  και προωθεί την απάντηση στον κόμβο  $SN_a$ . Εάν η έκδοση που περιγράφεται στην αίτηση είναι όμοια με αυτή που έχει ο διαμεσολαβητής, τότε αποστέλλεται ένα AckPacket μήνυμα.

#### 4.4.2 Απαιτήσεις συνέπειας

Το πρωτόκολλο NICC παρέχει ερωτήματα με διαφορετικά επίπεδα συνέπειας. Σε αυτή την περίπτωση εισάγουμε την έννοια του cache κόμβου. Οι cache κόμβοι τοποθετούνται μεταξύ των κόμβων που ζητούν τα αντικείμενα και τους διαμεσολαβητές. Ένα ερώτημα προωθείται αρχικά σε ένα cache κόμβο. Εάν απαιτείται μόνο αδύναμη συνέπεια, τότε ο cache κόμβος στέλνει αμέσως πίσω μια απάντηση στον κόμβο που έθεσε το ερώτημα. Εάν απαιτείται  $\Delta$  συνέπεια και η τιμή του  $\Delta$  είναι μεγαλύτερη από το μηδέν, τότε ο cache κόμβος εξυπηρετεί άμεσα το ερώτημα. Όμως, εάν η τιμή του  $\Delta$  είναι ίση με το μηδέν ή απαιτείται ισχυρή συνέπεια, τότε ο cache κόμβος στέλνει το ερώτημα σε ένα διαμεσολαβητή. Ο διαμεσολαβητής μπορεί να συγκρίνει τις εκδόσεις και να απαντήσει στον cache κόμβο. Εάν ο διαμεσολαβητής έχει μία έγκυρη και πρόσφατη έκδοση του αντικειμένου, τότε στέλνει ένα μήνυμα απάντησης στον cache κόμβο. Διαφορετικά, θα περιμένει το επόμενο ακυρωτικό μήνυμα από τον κόμβο πηγή. Συνεπώς, η έννοια του cache κόμβου χρησιμοποιείται στην περίπτωση που το πρωτόκολλο πρέπει να



εξυπηρετήσει αιτήματα για  $\Delta$  ή αδύναμη συνέπεια. Οι cache κόμβοι μπορούν να επιλεγούν με βάση το πρωτόκολλο NiCoCa [43].

#### 4.4.3 Πολιτική αντικατάστασης της cache

Αν και η χωρητικότητα της cache των αισθητήρων μπορεί να είναι της τάξης των GB (π.χ. NAND flash), η ανάπτυξη μιας αποτελεσματικής και αποδοτικής πολιτικής αντικατάστασης της cache είναι ένας πολύ σημαντικός στόχος. Η πολιτική αντικατάστασης της cache χρησιμοποιείται σε κάθε κόμβο, είτε πρόκειται για διαμεσολαβητή είτε για cache (στο υπόλοιπο της παραγράφου εξετάζουμε τους διαμεσολαβητές) και λαμβάνει υπόψη τα εξής χαρακτηριστικά των αντικειμένων:

1. ρυθμός ενημέρωσης,
2. ρυθμός προσπέλασης,
3. μέση τιμή hops των κόμβων που ζητούν το αντικείμενο,
4. μέγεθος του αντικειμένου.

Όταν ένας διαμεσολαβητής λάβει ένα μήνυμα αίτησης, εξάγει το πλήθος hops που διήνυσε το μήνυμα. Σύμφωνα με το πλήθος αυτό, ο διαμεσολαβητής υπολογίζει το μέσο πλήθος hops για κάθε αποθηκευμένο αντικείμενο. Όσο μεγαλύτερη είναι η μέση απόσταση των κόμβων που αιτούνται το αντικείμενο, τόσο περισσότερο επικτό είναι να παραμείνει στην cache. Αυτό συμβαίνει για να πετύχουμε μείωση της καθυστέρηση πρόσβασης, επειδή ο απομακρυσμένος αισθητήρας δεν θα πρέπει να περιμένει μέχρι η αίτησή του να φθάσει στον κόμβο πηγή. Ο ρυθμός προσπέλασης δείχνει τη συχνότητα αίτησης ενός αντικειμένου της cache, ενώ ο ρυθμός ενημέρωσης αναδεικνύει τη συχνότητα μεταβολής ενός αντικειμένου. Όσο υψηλότερος είναι ο ρυθμός προσπέλασης, τόσο πιθανότερο είναι να παραμείνει το αντικείμενο στην cache. Αντίθετα όσο υψηλότερος είναι ο ρυθμός ενημέρωσης ενός αντικειμένου, τόσο λιγότερο πιθανό είναι να παραμείνει στην cache. Τέλος, όσο μεγαλύτερο είναι το μέγεθος ενός αντικειμένου, τόσο πιθανότερο είναι να αντικατασταθεί. Επομένως, κατά τη διάρκεια αντικατάστασης της cache, ο διαμεσολαβητής υπολογίζει μια τιμή για κάθε αντικείμενο και αντικαθίσταται το αντικείμενο με τη μικρότερη τιμή. Η συνάρτηση αποτίμησης, για κάθε αντικείμενο  $i$  που βρίσκεται στην cache, είναι η εξής:

$$Value(d_i) = \frac{AR_i \times AvgHops_i}{UR_i \times Size_i}$$

όπου το  $AR$  αναφέρεται στο ρυθμό προσπέλασης,  $AvgHops$  είναι ο μέσος όρος των hops και το  $UR$  συμβολίζει το ρυθμό ενημέρωσης.

## 4.5 Πειραματική Αποτίμηση

Η απόδοση του πρωτοκόλλου NICC αποτιμήθηκε μέσω πειραματικών προσομοιώσεων. Στο παρόν κεφάλαιο παρουσιάζεται ένας μεγάλος αριθμός πειραμάτων με διάφορες παραμέτρους και μια σύγκριση μεταξύ της απόδοσης του πρωτοκόλλου NICC και της αποτελεσματικότερης τεχνικής που έχει προταθεί στα πλαίσια των MANET, η οποία ονομάζεται RPCC [21]. Επιπρόσθετα τα δύο πρωτόκολλα συγκρίνονται και με ένα πρωτόκολλο που δεν εμπεριέχει τη τεχνική του caching και ονομάζεται NC (No Caching), το οποίο αποδεικνύεται πως έχει τη χειρότερη επίδοση σε όλα τα πειράματα.

### 4.5.1 Μοντέλο προσομοίωσης

Αρχικά αναπτύχθηκε ένα μοντέλο προσομοίωσης με βάση το δικτυακό προσομοιωτή J-Sim [123]. Το πρωτόκολλο AODV [104] χρησιμοποιήθηκε για να δρομολογεί την κίνηση των δεδομένων στο APDA. Συνακόλουθα, το πρωτόκολλο IEEE 802.11 επιλέχθηκε ως πρωτόκολλο πρόσβασης του μέσου MAC και το μοντέλο ελεύθερου χώρου ως το μοντέλο ασύρματης διασποράς. Το εύρος ζώνης του ασύρματου δικτύου ήταν 2 Mbps. Επιπρόσθετα, τα χαρακτηριστικά της ραδιοεπικοινωνίας συνοψίζονται στον Πίνακα 4.2.

Λειτουργία	Δαπανώμενη Ενέργεια
Transmitter/Receiver Electronics	$E_{elec}=50\text{nJ/bit}$
Transmit Amplifier if $d_{toBS} \leq d_0$	$e_{fs}=10\text{pJ/bit/m}^2$
Transmit Amplifier if $d_{toBS} \geq d_0$	$e_{mp}=0.0013\text{pJ/bit/m}^4$
Data Aggregation	$E_{DA}=5\text{nJ/bit/signal}$

Πίνακας 4.2. Χαρακτηριστικά εκπομπής και λήψης στις προσομοιώσεις.

Τα πρωτόκολλα δοκιμάστηκαν σε μια μεγάλη ποικιλία τοπολογιών δικτύων αισθητήρων, με διαφορετικού βαθμού πυκνότητες κόμβων, από 4 μέχρι 10. Τα πειράματα διεξήχθησαν επίσης με διαφορετικό πλήθος κόμβων που ποικίλλει μεταξύ 100 και 500 κόμβων. Επιπρόσθετα, αποτιμήσαμε πειραματικά την απόδοση των πρωτοκόλλων χρησιμοποιώντας δύο διαφορετικά σύνολα δεδομένων. Κάθε

δεδομένο έχει μέγεθος που κατανέμεται ομοιόμορφα μεταξύ 1KB και 10KB για το πρώτο σύνολο, και από 1MB μέχρι 5MB για το δεύτερο.

Η τοπολογία του δικτύου αποτελείται από πολλά τετράγωνα τμήματα πλέγματος όπου τοποθετούνται ένας ή περισσότεροι κόμβοι. Ο αριθμός των τετράγωνων τμημάτων του πλέγματος εξαρτάται από τον αριθμό των κόμβων και την πυκνότητα τοποθέτησής τους. Οι τοπολογίες δημιουργούνται ως εξής: η τοποθεσία κάθε ενός από τους  $n$  κόμβους κατανέμεται ομοιόμορφα ανάμεσα στα σημεία  $(x=0, y=0)$  και  $(x=2000, y=2000)$ . Η μέση πυκνότητα  $d$  υπολογίζεται μέσω της ταξινόμησης όλων των  $n \times (n-1)/2$  ακμών του δικτύου με βάση το μήκος τους σε αύξουσα σειρά. Το μέγεθος κάθε τμήματος του πλέγματος, το οποίο αντιστοιχεί στην τιμή του  $d$  ισούται με την τετραγωνική ρίζα του 2 πολλαπλασιαζόμενη με το μήκος της ακμής στη θέση  $n \times d/2$  της ταξινομημένης ακολουθίας. Δύο αισθητήρες θεωρούνται γείτονες όταν τοποθετούνται μέσα στο ίδιο τμήμα του πλέγματος ή σε γειτονικά. Η περιοχή προσομοίωσης έχει μέγεθος  $2000m \times 2000m$  και διαιρείται σε ίσου μεγέθους τετραγωνικά τμήματα πλέγματος. Ξεκινώντας από το κάτω αριστερά άκρο, ονομάζουμε τα τμήματα κατά στήλες ως  $1, 2, \dots$

Το μοντέλο ερωτημάτων που ακολουθεί ένας αισθητήρας, είναι όμοιο με προηγούμενες μελέτες [21]. Κάθε αισθητήρας δημιουργεί read-only ερωτήματα με ένα εκθετικά κατανεμημένο διάστημα. Επίσης, οι κόμβοι δημιουργούν μια ανεξάρτητη ροή ενημερώσεων για τα δεδομένα τους με ένα εκθετικά κατανεμημένο διάστημα ενημέρωσης. Το πρότυπο προσπέλασης των αισθητήρων είναι ανεξάρτητο από την τοποθεσία, δηλαδή οι αισθητήρες αποφασίζουν ανεξάρτητα για τα δεδομένα που τους ενδιαφέρουν. Κάθε αισθητήρας δημιουργεί ερωτήματα για δεδομένα με βάση μια ομοιόμορφη κατανομή. Επιπλέον, ως πρότυπο πρόσβασης χρησιμοποιήθηκε και η Zipfian κατανομή με  $\theta=0.8$ . Σύμφωνα με τη Zipfian κατανομή, ομάδες κόμβων που βρίσκονται σε γειτονικά τμήματα του πλέγματος (25 τμήματα με μέγεθος  $400m \times 400m$ ) ακολουθούν το ίδιο πρότυπο πρόσβασης. Εξετάσαμε την απόδοση των πρωτοκόλλων και για τα δύο πρότυπα πρόσβασης. Στην περίπτωση του προτύπου πρόσβασης με βάση τη Zipfian κατανομή, πραγματοποιήσαμε πειράματα με διάφορες τιμές  $\theta$  μεταξύ 0.0 και 1.0. Στη συνέχεια θα παρουσιάσουμε μόνο τα αποτελέσματα για  $\theta=0.8$  επειδή είναι το αντιπροσωπευτικότερο πρότυπο πρόσβασης. Ένα δεδομένο ανατίθεται σε κάθε αισθητήρα και επομένως ο αριθμός των δεδομένων είναι 100 στην περίπτωση της τοπολογίας 100 αισθητήρων και 500 στην περίπτωση της τοπολογίας 500 αισθητήρων. Οι παράμετροι του συστήματος παρουσιάζονται στον Πίνακα 4.3.

Παράμετροι	Εξ ορισμού	Φάσμα τιμών
Πλήθος αντικειμένων ( $N$ )	100	100-500
$S_{min}$ (KB)	1	
$S_{max}$ (KB)	10	
$S_{min}$ (MB)	1	
$S_{max}$ (MB)	5	
Zipfian $\theta$	0.8	
Πλήθος ερωτημάτων κόμβου	200	100-200
Εύρος ζώνης (Mbps)	2	
Διάστημα μεταξύ ερωτημάτων	20s για αντικείμενα KB-size 3m για αντικείμενα MB-size	
Διάστημα μεταξύ ενημερώσεων	2m για αντικείμενα KB-size 20m για αντικείμενα MB-size	
Μέγεθος cache (KB)	75	10-100
Μέγεθος cache (MB)	37.5	15-65
$TTL$	3 hops	3-6

Πίνακας 4.3. Παράμετροι προσομοίωσης.

#### 4.5.2 Μετρικές απόδοσης

Οι ποσότητες που μετρούμε στα πειράματα είναι ο αριθμός από hits (local, remote και global), ο μέση καθυστέρηση πρόσβασης, η εναπομείνασα ενέργεια, η επίδραση της τιμής  $TTL$  των ακυρωτικών μηνυμάτων και η επίδραση της πολιτικής αντικατάστασης στην απόδοση του συστήματος. Μια συχνά χρησιμοποιούμενη μετρική είναι ο συνολικός αριθμός των μηνυμάτων που διαδίδονται μέσα στο δίκτυο και τα οποία συμμετέχουν στο προτεινόμενο πρωτόκολλο. Το κόστος των μηνυμάτων περιλαμβάνει όλα τα μηνύματα των ερωτημάτων και απαντήσεων κατά τη διαδικασία αναζήτησης των αντικειμένων και ανάκτησής τους. Το επιπλέον κόστος από τα μηνύματα δρομολόγησης δεν λαμβάνονται υπόψη επειδή ο αριθμός των μηνυμάτων λόγω δρομολόγησης είναι ο ίδιος και για τα δύο πρωτόκολλα. Οι τρεις διαφορετικοί τύποι από hits είναι οι εξής:

1. Local hit (LH): το ζητούμενο αντικείμενο βρίσκεται στην cache του κόμβου που εκδίδει το ερώτημα και είναι έγκυρο.
2. Remote hit (RH): το ζητούμενο αντικείμενο βρίσκεται στην cache ενός κόμβου που έχει τουλάχιστον ένα διαμεσολαβητή επί της διαδρομής που

συνδέει τον κόμβο που ζήτησε το αντικείμενο και τον κόμβο πηγή.

3. Global hit (GH): το ζητούμενο αντικείμενο αποκτάται από τον κόμβο πηγή.

Είναι φανερό ότι ένας μικρός αριθμός από global hits συνεπάγεται μικρότερη συμφόρηση του δικτύου και συνεπώς λιγότερες συγκρούσεις και απώλειες πακέτων. Επιπλέον, μεγάλος αριθμός από remote hits αποδεικνύει την αποτελεσματικότητα της συνεργασίας στη μείωση του αριθμού των global hits. Ένας μεγάλος αριθμός από local hits δεν συνεπάγεται μια αποτελεσματική πολιτική συνεργατικού caching, εκτός και αν συνδυάζεται με ένα μικρό αριθμό από global hits. Αυτό συμβαίνει γιατί το κόστος των global hits περιορίζει σημαντικά τα οφέλη που προκύπτουν από τα local hits. Η ανεξάρτητη παράμετρος σε όλα τα πειράματα είναι το μέγεθος της cache του αισθητήρα. Σε κάθε περίπτωση ο τοπικός αποθηκευτικός χώρος ενός αισθητήρα ποικίλλει από 1% μέχρι 10% του συνολικού μεγέθους όλων των πολυμεσικών δεδομένων.

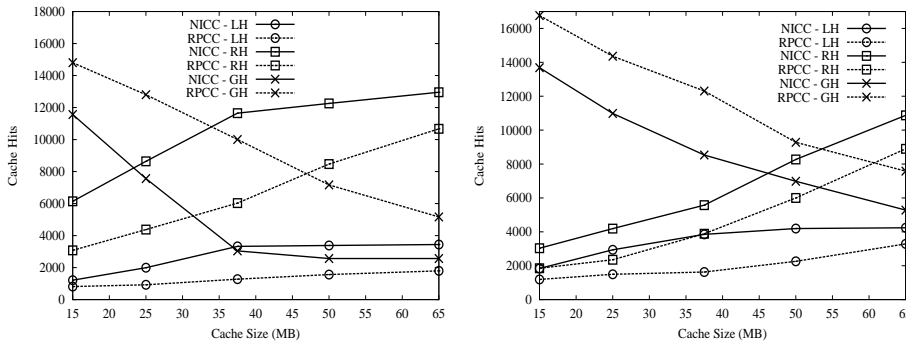
### 4.5.3 Αποτίμηση

Εκτελέσαμε ένα μεγάλο αριθμό πειραμάτων μεταβάλλοντας το μέγεθος του δικτύου (το πλήθος των αισθητήρων του δικτύου), μεταβάλλοντας τον τρόπο προσπέλασης από τους κόμβους και το μέγεθος της cache. Πιο συγκεκριμένα, πραγματοποιήσαμε πειράματα με 100 και 500 αισθητήρες, για μέγεθος cache ίσο με το 1%, 5% και 10% του συνολικού μεγέθους όλων των πολυμεσικών δεδομένων, για πρότυπα προσπέλασης με  $\theta=0.8$ , για μέσο όρο πυκνότητας των αισθητήρων ίσο με 4 (αραιό δίκτυο αισθητήρων) και 10 (πυκνό δίκτυο αισθητήρων) και τέλος για δεδομένα με μέγεθος από λίγα KB μέχρι λίγα MB. Για κάθε διαφορετικό σύνολο τιμών των παραμέτρων μετρήσαμε τον αριθμό των hits (local, remote, global), την καθυστέρηση πρόσβασης και το πλήθος των μηνυμάτων. Τα πειραματικά αποτελέσματα χωρίζονται σε δύο μεγάλες κατηγορίες. Η πρώτη αποτελείται από πειράματα με δεδομένα μεγέθους λίγων MB, ενώ η δεύτερη αποτελείται από πειράματα με δεδομένα μεγέθους λίγων KB. Ακολουθώντας την ίδια μεθοδολογία με το πρωτόκολλο RPPC, θέσαμε την τιμή του  $TTL=3$ . Η επίδραση διαφορετικών τιμών  $TTL$  θα συζητηθεί στην Παράγραφο 4.5.3.3.

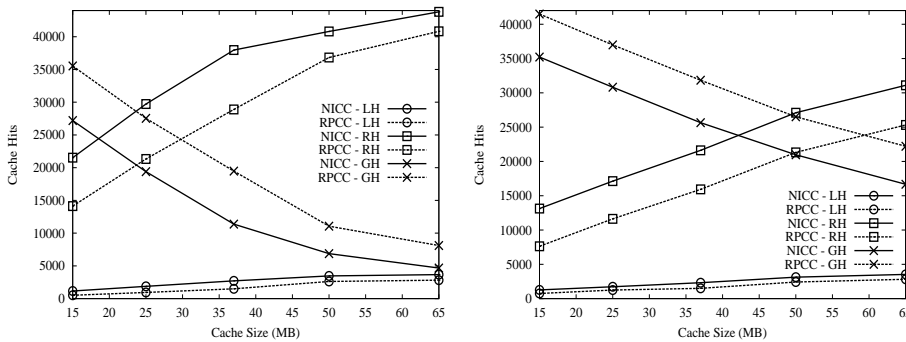
#### 4.5.3.1 Πειράματα με δεδομένα μεγέθους MB

Σκοπός είναι να εξετάσουμε την απόδοση των πρωτοκόλλων συνέπειας στο χειρισμό μεγάλων πολυμεσικών δεδομένων, όπως για παράδειγμα σε αρχεία video, ήχου και εικόνας. Όλα τα σχήματα δείχνουν ότι και οι δύο πολιτικές έχουν

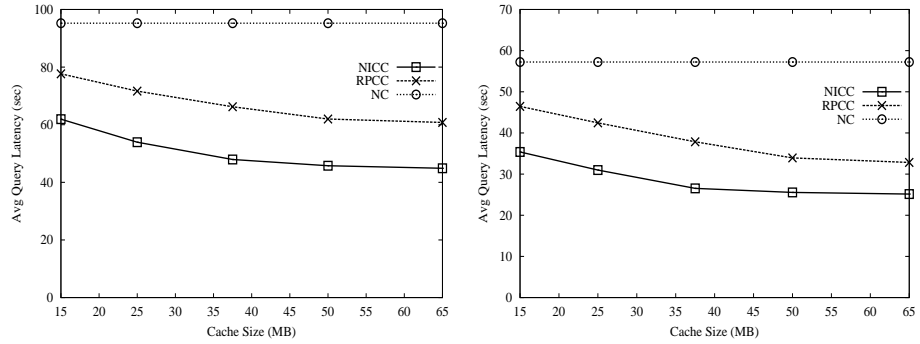
καλύτερη μέση καθυστέρηση πρόσβασης, ποσοστό hit (local, remote και global) και κόστος μηνυμάτων όταν το μέγεθος της cache αυξάνεται από 15MB σε 65MB και αυτό συμβαίνει τόσο για αραιές όσο και για πυκνές τοπολογίες. Αυτό προκύπτει επειδή περισσότερα χρήσιμα δεδομένα μπορούν να βρεθούν στη τοπική cache καθώς το μέγεθός της μεγαλώνει. Τα αποτελέσματα εμφανίζονται στα Σχήματα 4.5-4.6. Η πρώτη παρατήρηση είναι ότι το προτεινόμενο πρωτόκολλο έχει πάντα καλύτερη απόδοση από τον ανταγωνιστή του. Το NICC πετυχαίνει πάντα υψηλότερο αριθμό από remote hits σε σχέση με το πρωτόκολλο RPCC, επειδή οι διαμεσολαβητές είναι τοποθετημένοι σε κεντρικά σημεία του δικτύου. Επιπρόσθετα το NICC πετυχαίνει μικρότερο αριθμό global hits από το RPCC. Συνεπώς, οι κόμβοι πηγές δεν έχουν μεγάλο φόρτο εργασίας και δεν αποτελούν σημείο αποτυχίας του δικτύου. Τέλος, μικρότερα είναι τα οφέλη του προτεινόμενου πρωτοκόλλου έναντι του ανταγωνιστικού σε ότι αφορά τα local hits.



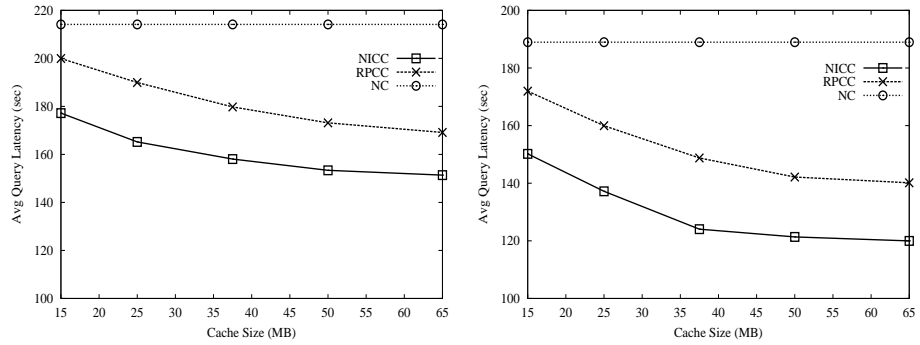
Σχήμα 4.5: Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες.



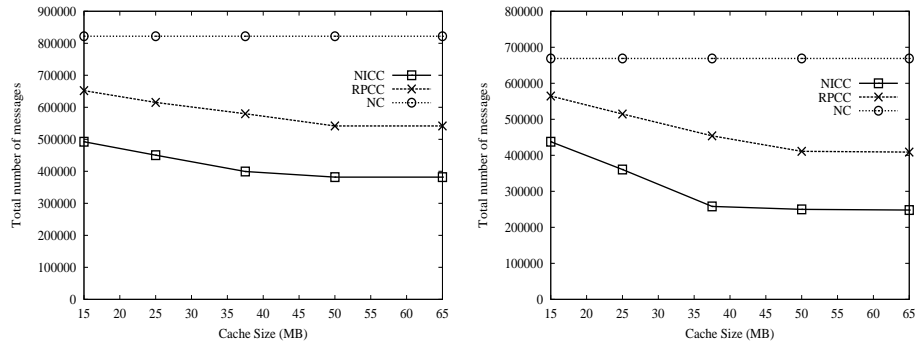
Σχήμα 4.6: Επίδραση μεγέθους της cache στα hits (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες.



Σχήμα 4.7: Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισητήρες.



Σχήμα 4.8: Επίδραση μεγέθους της cache στην καθυστέρηση πρόσβασης (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισητήρες.



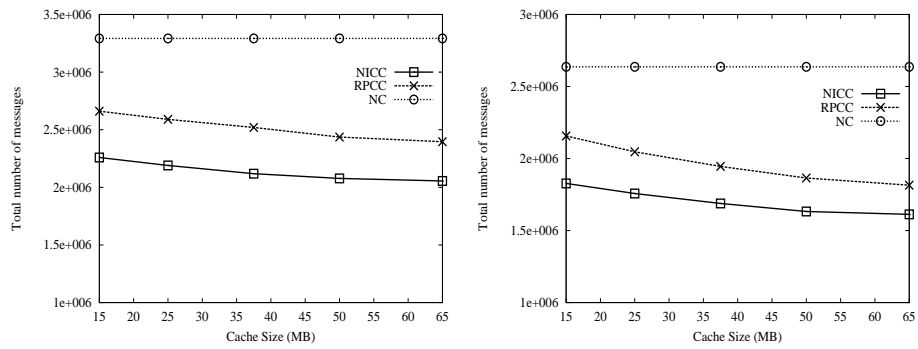
Σχήμα 4.9: Επίδραση μεγέθους της cache στον αριθμό των μηνυμάτων (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισητήρες.

Η δεύτερη κύρια παρατήρηση είναι ότι η απόδοση του NICC και του RPCC πρωτοκόλλου γίνεται καλύτερη για αραιότερα δίκτυα αισθητήρων. Αυτή η παρατήρηση ισχύει και για τις επόμενες μετρικές. Η εξήγηση είναι ότι το RPCC λαμβάνει καλύτερες αποφάσεις ως προς την εκλογή των αναμεταδοτών σε πυκνότερες τοπολογίες. Αυτό γίνεται φανερό όταν εξετάζουμε τη σχετική εναπομείνουσα ενέργεια ανά αισθητήρα (δες τα Σχήματα 4.15 και 4.16). Επομένως, οι υψηλά παραμετροποιησιμες μέθοδοι για την επιλογή των κόμβων με ειδικούς ρόλους (διαμεσολαβητές ή αναμεταδότες), έχουν σημαντική συνεισφορά στην απόδοση των πρωτοκόλλων συνέπειας της cache.

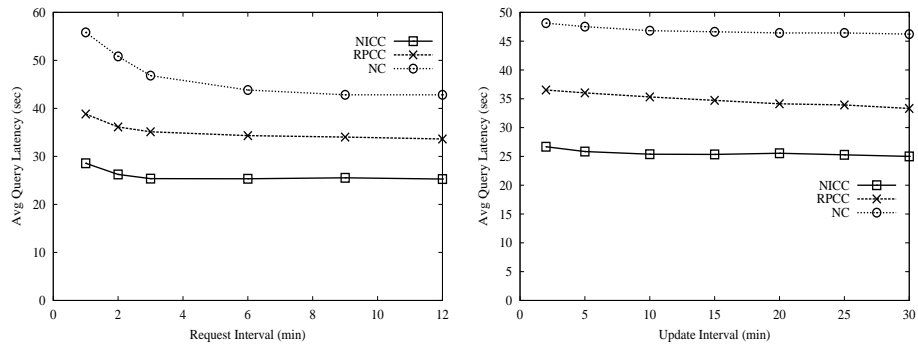
Στην συνέχεια αποτιμήσαμε την απόδοση των αλγορίθμων λαμβάνοντας υπόψη τη μέση καθυστέρηση πρόσβασης, η οποία επιτυγχάνεται για διάφορα μεγέθη cache όταν τα μεταδιδόμενα αντικείμενα έχουν μέγεθος λίγα MB και για τοπολογίες δικτύου, είτε αραιές είτε πυκνές. Τα αποτελέσματα παρουσιάζονται στα Σχήματα 4.7 και 4.8. Η χρονική καθυστέρηση που συμβαίνει στο πρωτόκολλο NICC είναι 20% με 30% μικρότερη από αυτή του πρωτοκόλλου RPCC. Εξαιτίας της πολιτικής αντικατάστασης της cache και των κεντρικών σημείων όπου βρίσκονται οι διαμεσολαβητές, τα ερωτήματα εξυπηρετούνται γρηγορότερα στο NICC από ότι στο RPCC. Μπορούμε επίσης να διαπιστώσουμε ότι η μείωση του χρόνου απόκρισης στο προτεινόμενο πρωτόκολλο για μεγέθη της cache μεταξύ 37.5 και 65MB δεν είναι σημαντικά μεγάλη. Επιπλέον, είναι άξιο αναφοράς ότι το NICC προσεγγίζει πάντα τη βέλτιστη απόδοση όταν το μέγεθος της cache είναι ίσο με 37.5MB. Αυτό αποδεικνύει τη μικρή απαίτηση που έχει το πρωτόκολλο σε χωρητικότητα της cache.

Επιπρόσθετα, αποτιμήσαμε την απόδοση των αλγορίθμων με βάση τον αριθμό των μεταδιδόμενων μηνυμάτων. Τα αποτελέσματα εμφανίζονται στα Σχήματα 4.9 και 4.10. Τα συγκριτικά αποτελέσματα ακολουθούν τα ίδια συμπεράσματα που εξήχθησαν από τα προηγούμενα πειράματα. Υπάρχει μια στενή σύνδεση μεταξύ των μηνυμάτων και της καθυστέρησης πρόσβασης, επειδή όσο μεγαλύτερος είναι ο αριθμός των μηνυμάτων που μεταδίδονται τόσο μεγαλύτερος είναι και ο ανταγωνισμός για την απόκτηση του ασύρματου καναλιού. Επομένως μπορούν να συμβούν πολλές συγκρούσεις που οδηγούν σε αύξηση της μέσης καθυστέρησης πρόσβασης. Ο αριθμός των μηνυμάτων στις αραιές τοπολογίες δικτύου είναι μεγαλύτερος από ότι στις πυκνές. Αυτό οφείλεται στον αριθμό hops που πρέπει να διέλθει ένα ερώτημα μέχρι να φθάσει σε ένα κόμβο με το ζητούμενο αντικείμενο. Στις αραιές τοπολογίες ο αριθμός hops είναι μεγαλύτερος συγκριτικά με τις πυκνές τοπολογίες.

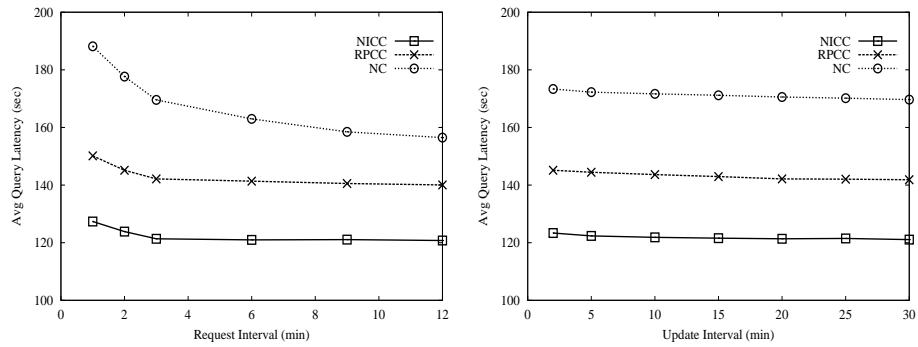




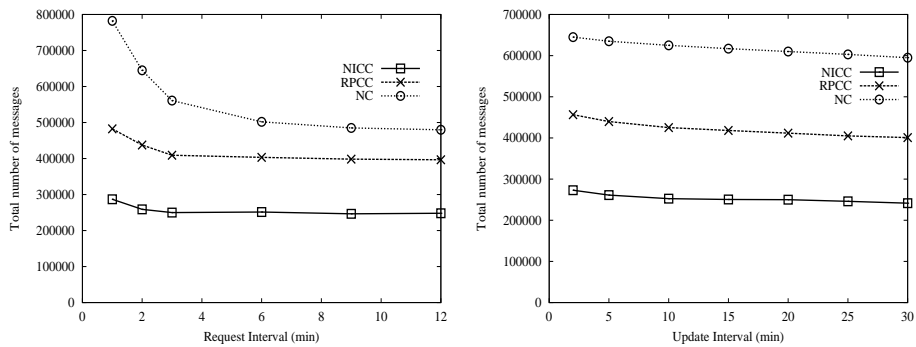
Σχήμα 4.10: Επίδραση μεγέθους της cache στον αριθμό των μηνυμάτων (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες.



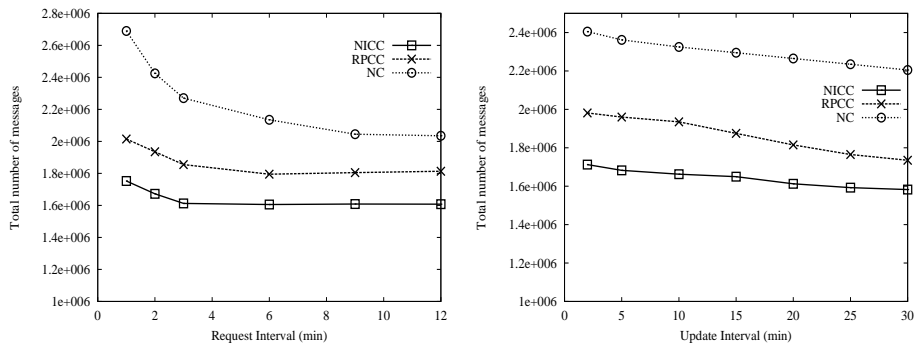
Σχήμα 4.11: Καθυστερήση πρόσβασης σε δεδομένα μεγέθους MB σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 100 αισθητήρες.



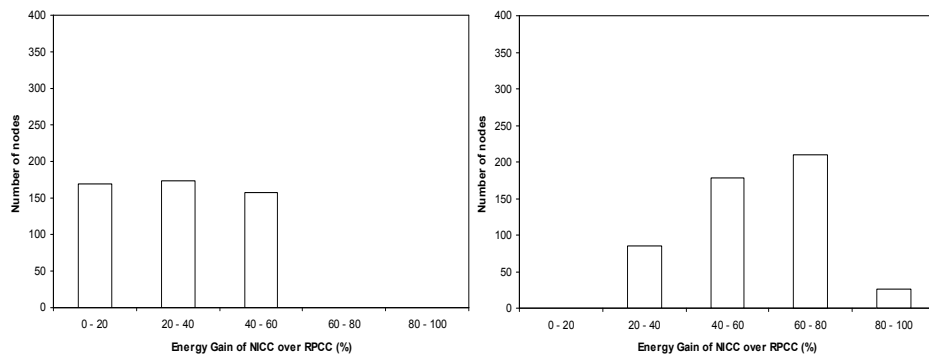
Σχήμα 4.12: Καθυστερήση πρόσβασης σε δεδομένα μεγέθους MB σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 500 αισθητήρες.



Σχήμα 4.13: Δικτυακή κίνηση σε δεδομένα μεγέθους MB σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 100 αισθητήρες.



Σχήμα 4.14: Δικτυακή κίνηση σε δεδομένα μεγέθους MB σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 500 αισθητήρες.



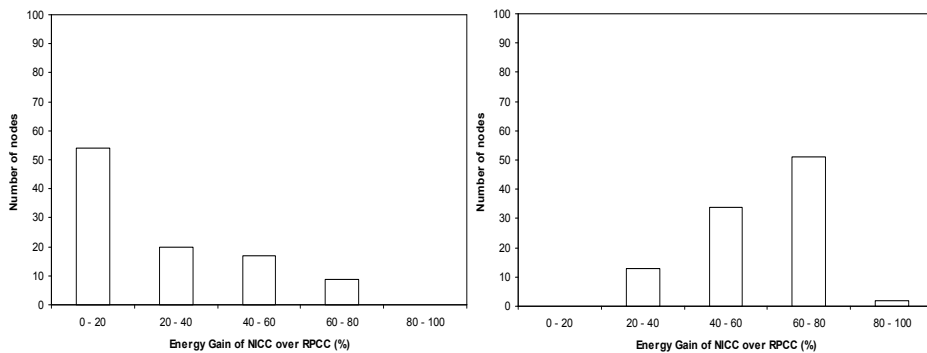
Σχήμα 4.15: % διαφορά στην κατανάλωση ενέργειας (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες, όταν εκτελούνται το NICC και το RPCC.

Τα Σχήματα 4.11-4.12 αποτιμούν την απόδοση με όρους καθυστέρησης πρόσβασης για διάφορες τιμές διαστημάτων ερωτήματος (request interval) και ενημέρωσης (update interval). Τα αριστερά γραφήματα των άνω σχημάτων δείχνουν ότι η καθυστέρηση πρόσβασης είναι λίγο μεγαλύτερη για μικρές τιμές διαστήματος ερωτήματος. Αυτό συμβαίνει επειδή η κίνηση στο δίκτυο αυξάνεται όσο μειώνεται το διάστημα ερωτήματος. Έτσι, η καθυστέρηση συμβαίνει εξαιτίας του ανταγωνισμού για την απόκτηση του ασύρματου καναλιού. Είναι γεγονός ότι και τα δύο σχήματα παρουσιάζουν μια σαφώς καλύτερη απόδοση του πρωτοκόλλου NICC έναντι του πρωτοκόλλου RPCC. Στο NICC τα ερωτήματα εξυπηρετούνται από τους διαμεσολαβητές που βρίσκονται σε κεντρικά σημεία. Συνεπώς η χρονική καθυστέρηση στην εξυπηρέτηση του ερωτήματος είναι μικρότερη. Τα δεξιά γραφήματα των σχημάτων δείχνουν τα αποτελέσματα με τη χρήση διαφορετικών διαστημάτων ενημέρωσης. Όπως παρατηρούμε δεν υπάρχει σημαντική διαφορά των δύο πρωτοκόλλων σε σχέση με τη χρονική καθυστέρηση όταν μεγαλώνει το διάστημα ενημέρωσης. Επιπλέον, τα δύο πρωτόκολλα έχουν μια ελαφρώς χειρότερη απόδοση για μικρά διαστήματα ενημέρωσης, επειδή περισσότερα μηνύματα invalidation διαδίδονται από τους κόμβους πηγές. Εν τέλει, προκύπτει ότι το πρωτόκολλο NICC αποδίδει καλύτερα από το πρωτόκολλο RPCC.

Λιγότερα μηνύματα επικοινωνίας οδηγούν σε μικρότερη δικτυακή κίνηση και κατανάλωση ενέργειας. Επομένως, ένας καλός αλγόριθμος θα πρέπει να στέλνει ένα μικρό αριθμό μηνυμάτων. Τα Σχήματα 4.13-4.14 αποτιμούν την απόδοση των πρωτοκόλλων με όρους δικτυακής κίνησης για διάφορες τιμές διαστήματος ερωτήματος και διαστήματος ενημέρωσης. Όπως παρατηρούμε, το πρωτόκολλο RPCC οδηγεί πάντα σε ένα μεγάλο κόστος επικοινωνίας. Επίσης, παρατηρούμε ότι το διάστημα ερωτήματος και το διάστημα ενημέρωσης δεν επηρεάζουν την απόδοση των δύο πρωτοκόλλων εκτός από την περίπτωση των μικρών τιμών.

Από τα αποτελέσματα για τον αριθμό των hits και των μεταδοθέντων μηνυμάτων, εξάγουμε το συμπέρασμα ότι το προτεινόμενο πρωτόκολλο είναι ενεργειακά αποδοτικότερο έναντι του ανταγωνιστή του. Το συμπέρασμα αυτό μπορεί να αναδειχθεί και από τη συγκριτική διαφορά μεταξύ του ποσού της καταναλωθείσας ενέργειας κάθε κόμβου, αφού εκτελέσει το πρωτόκολλο NICC και το πρωτόκολλο RPCC. Τα αποτελέσματα εμφανίζονται στα Σχήματα 4.15-4.16. Είναι ενδεικτικό ότι δεν υπάρχει κάποιος κόμβος που να εκτελεί το NICC με λιγότερη εναπομείνασα ενέργεια από την αντίστοιχη ενέργεια που έχει αφότου έχει ολοκληρώσει την εκτέλεση του RPCC. Στο NICC κάθε κόμβος επιλέγει τους σημαντικότερους κόμβους της 2-hop γειτονιάς του σύμφωνα με την τοπική όψη του δικτύου. Επομένως, γειτονικοί κόμβοι μπορεί να επιλέξουν διαφορετικούς κόμ-

βους ως διαμεσολαβητές. Η διαδικασία επιλογής των διαμεσολαβητών εξαρτάται από τους κόμβους που ξεκινούν τη διαδικασία εκπομπής ενός μηνύματος. Έτσι, εξαιτίας της κατανευμένης εκλογής των διαμεσολαβητών από κάθε αισθητήρα, το πρωτόκολλο NICC επιτυγχάνει μια ισορροπημένη κατανάλωση της ενέργειας μεταξύ των κόμβων. Συνακόλουθα, το NICC αποφεύγει το μειονέκτημα της κεντρικής διαδικασίας εκλογής των διαμεσολαβητών, η οποία θα οδηγούσε σε ταχεία κατανάλωση της ενέργειας της μπαταρίας των διαμεσολαβητών.

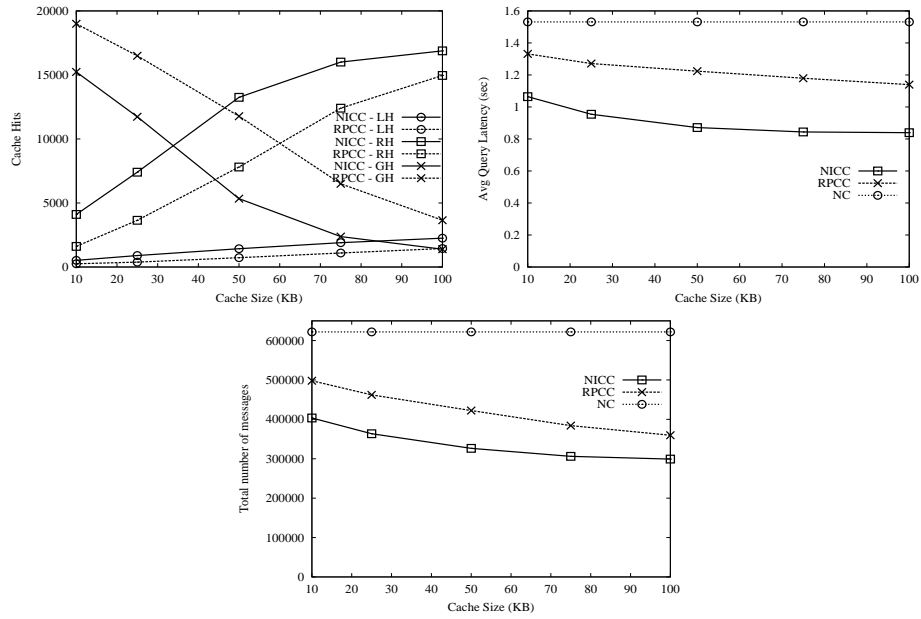


Σχήμα 4.16: % διαφορά στην κατανάλωση ενέργειας (δεδομένα μεγέθους MB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100, αισθητήρες όταν εκτελούνται το NICC και το RPCC.

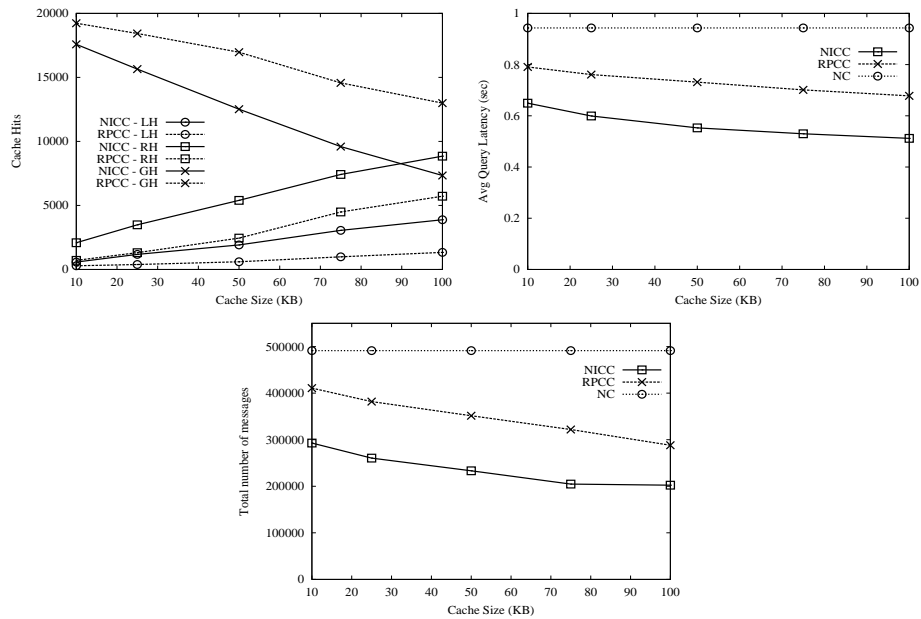
Επιπλέον, τόσο στην περίπτωση των KB-sized αντικειμένων (Σχήματα 4.25-4.26) όσο και στην περίπτωση των MB-sized αντικειμένων, παρατηρούμε ότι τα οφέλη ως προς την ενέργεια είναι μεγαλύτερα για το πρωτόκολλο NICC όταν η τοπολογία του δικτύου είναι πυκνή, επειδή η επιλογή των διαμεσολαβητών είναι περισσότερο ακριβής. Τέλος, όλα τα σχήματα παρουσιάζουν τη μεγάλη κατανάλωση ενέργειας των αισθητήρων που προκαλεί το πρωτόκολλο RPCC. Μια τέτοια κατάσταση είναι ανεπιθύμητη επειδή μπορεί να προκαλέσει διαίρεση του δικτύου.

#### 4.5.3.2 Πειράματα με δεδομένα μεγέθους KB

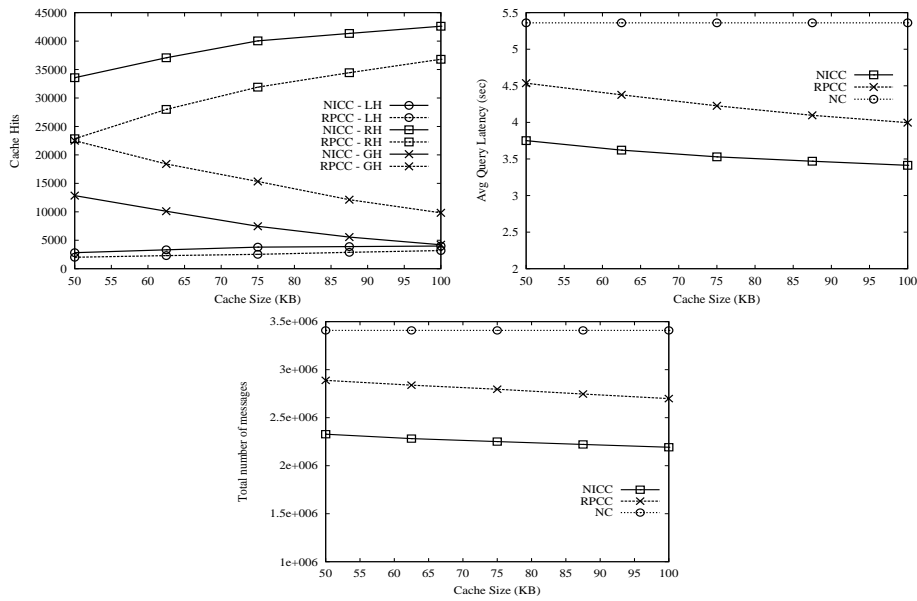
Μια σημαντική ερώτηση είναι εάν αυτά τα συγκριτικά αποτελέσματα διατηρούνται όταν το δίκτυο αισθητήρων διαχειρίζεται μικρότερου μεγέθους δεδομένα, π.χ. με μέγεθος λίγων KB. Αν και αναμένουμε ότι τα ΑΠΔΑ θα διαχειρίζονται αρχεία εικόνας και video, παρ' όλα αυτά είναι πολύ πιθανό τα ΑΠΔΑ να διαχειρίζονται ταυτόχρονα και πολύ μικρότερα δεδομένα. Για να εξερευνήσουμε την απόδοση των πρωτοκόλλων συνέπειας της cache σε αυτή την περίπτωση, πραγματοποιούμε το ίδιο σύνολο πειραμάτων αλλά για δεδομένα μεγέθους KB και παρουσιάζουμε ένα υποσύνολο των αποτελεσμάτων στη συνέχεια.



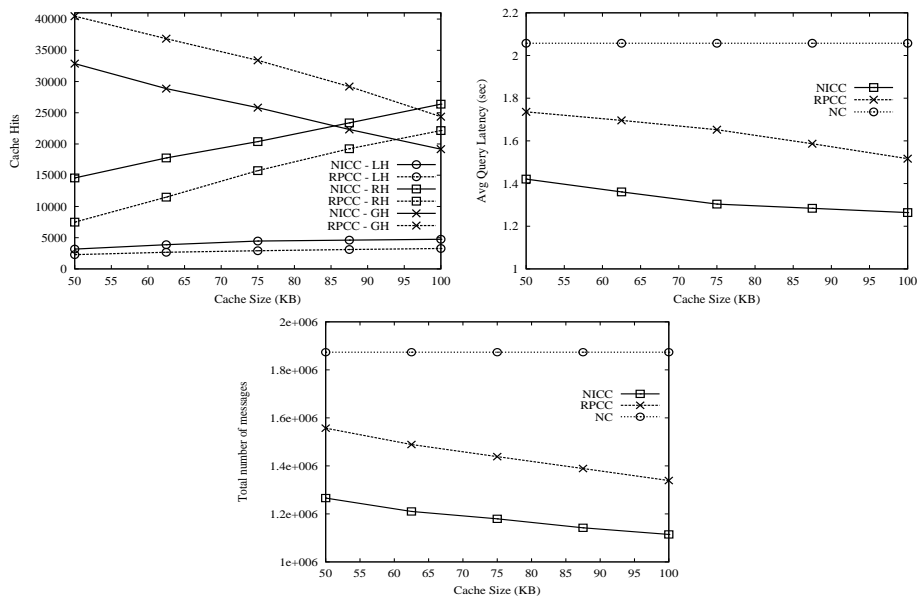
Σχήμα 4.17: Επίδραση μεγέθους της cache (δεδομένα μεγέθους KB,  $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΠΔΑ ( $d=4$ ) με 100 αισθητήρες.



Σχήμα 4.18: Επίδραση μεγέθους της cache (δεδομένα μεγέθους KB,  $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΠΔΑ ( $d=10$ ) με 100 αισθητήρες.

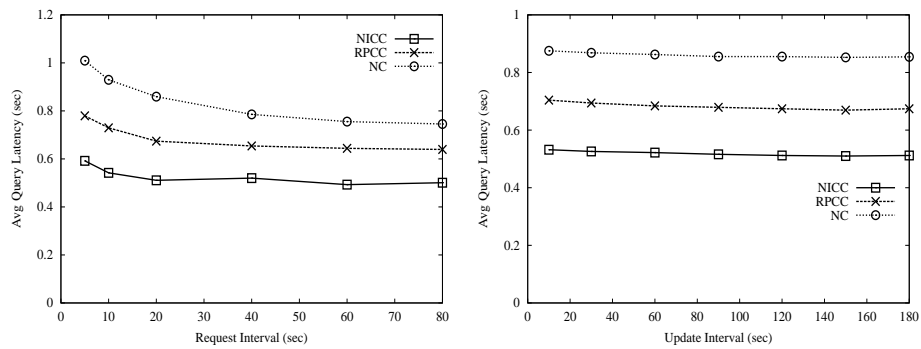


Σχήμα 4.19: Επίδραση μεγέθους της cache (δεδομένα μεγέθους KB,  $\theta=0.8$ ) στα: (Επάνω) hits, (Μέση) καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΠΔΑ ( $d=4$ ) με 500 αισθητήρες.

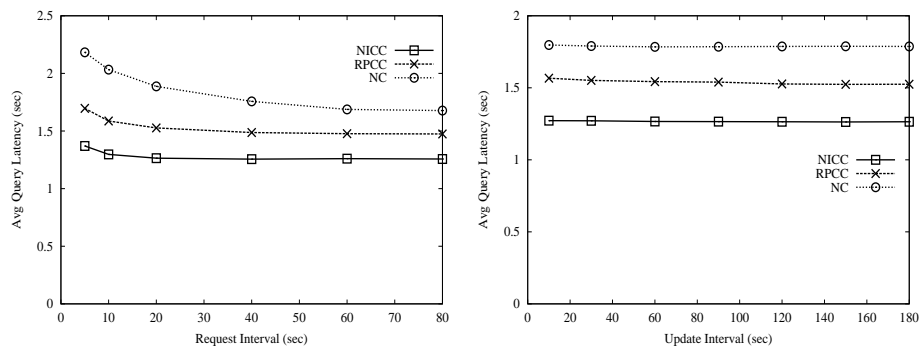


Σχήμα 4.20: Επίδραση μεγέθους της cache (δεδομένα μεγέθους KB,  $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΔΑ ( $d=10$ ) με 500 αισθητήρες.

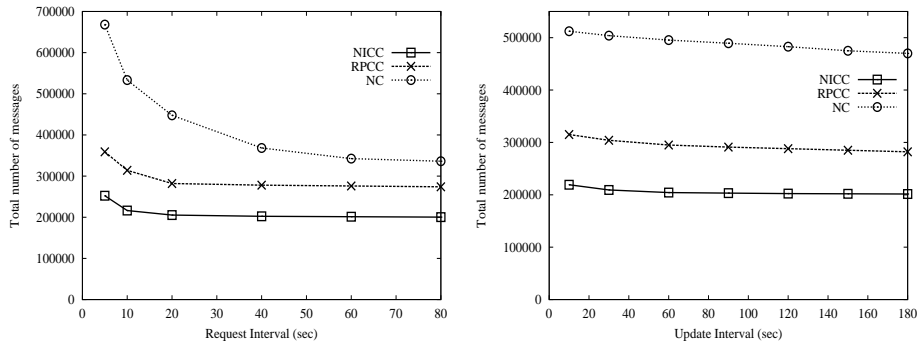
Οι γενικές παρατηρήσεις που καταγράψαμε στην περίπτωση των πειραμάτων με δεδομένα μεγέθους MB, εξακολουθούν να ισχύουν και σε αυτή την περίπτωση. Το πρωτόκολλο NICC επιτυγχάνει σημαντικά μικρότερο αριθμό από global hits και μεγαλύτερο αριθμό από remote hits συγκριτικά με το πρωτόκολλο RPCC. Ο σχολιασμός των πειραμάτων δεν παρατίθεται καθώς όλα τα γραφήματα της απόδοσης (Σχήματα 4.17-4.26), καταδεικνύουν τη σαφή υπεροχή του πρωτοκόλλου NICC. Το προτεινόμενο πρωτόκολλο πετυχαίνει ξανά 25% περισσότερα remote hits και 50% λιγότερα global hits σε σχέση με το πρωτόκολλο RPCC. Επίσης τα αποτελέσματα επίσης που σχετίζονται με την καθυστέρηση πρόσβασης και τον αριθμό των μηνυμάτων, παρουσιάζουν παρόμοια συμπεριφορά με αυτή του προηγούμενου συνόλου πειραμάτων.



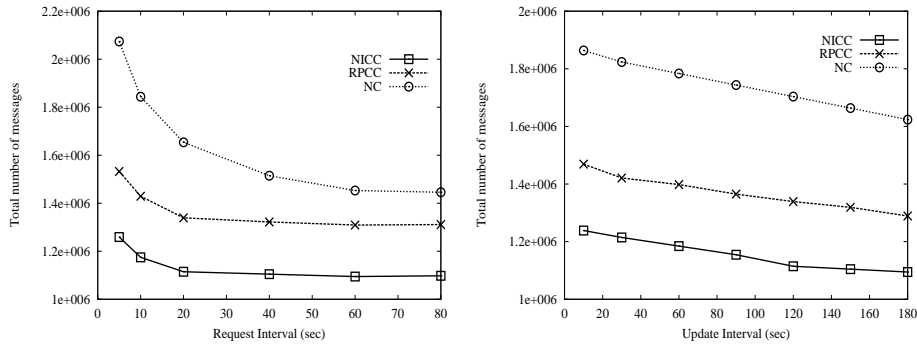
Σχήμα 4.21: Καθυστέρηση πρόσβασης (δεδομένα μεγέθους KB) σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 100 αισθητήρες.



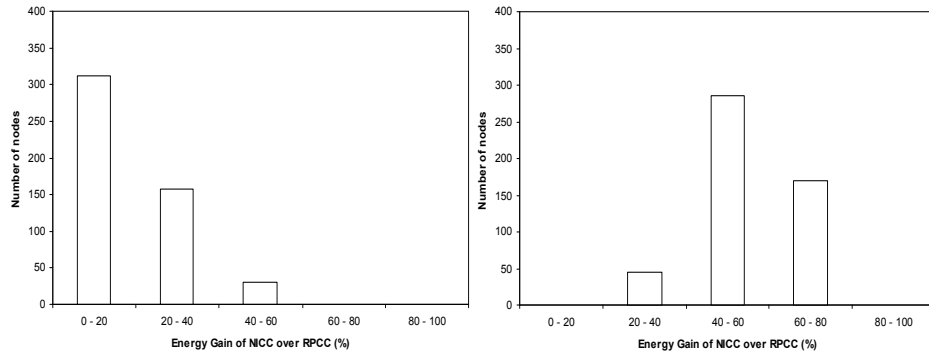
Σχήμα 4.22: Καθυστέρηση πρόσβασης (δεδομένα μεγέθους KB) σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 500 αισθητήρες.



Σχήμα 4.23: Δικτυακή κίνηση (δεδομένα μεγέθους KB) σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 100 αισθητήρες.



Σχήμα 4.24: Δικτυακή κίνηση (δεδομένα μεγέθους KB) σύμφωνα με το διάστημα ερωτήματος και το διάστημα ενημέρωσης σε ένα ΑΠΔΑ με 500 αισθητήρες.

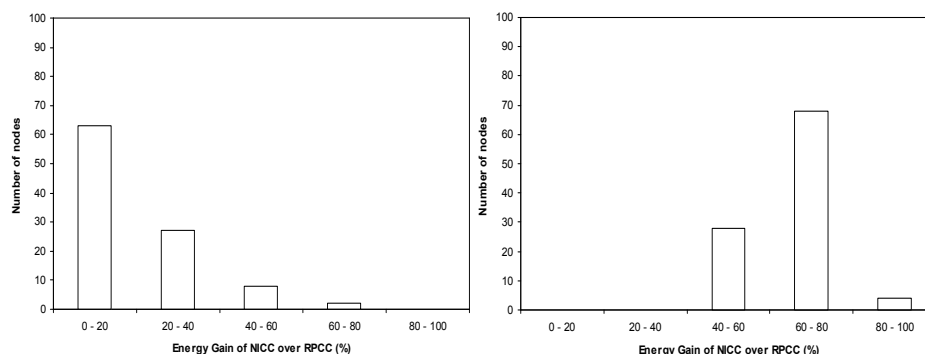


Σχήμα 4.25: % διαφορά στην κατανάλωση ενέργειας (δεδομένα μεγέθους KB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες όταν εκτελούνται τα NCC και RPCC.

#### 4.5.3.3 Επίδραση τιμής $TTL$ στην απόδοση του συστήματος

Κάθε cache κόμβος μπορεί να συνδεθεί αμέσως με ένα διαμεσολαβητή, εάν υπάρχουν πολλοί διαμεσολαβητές στο δίκτυο. Αυτό συνεπάγεται ότι η επικοινωνία





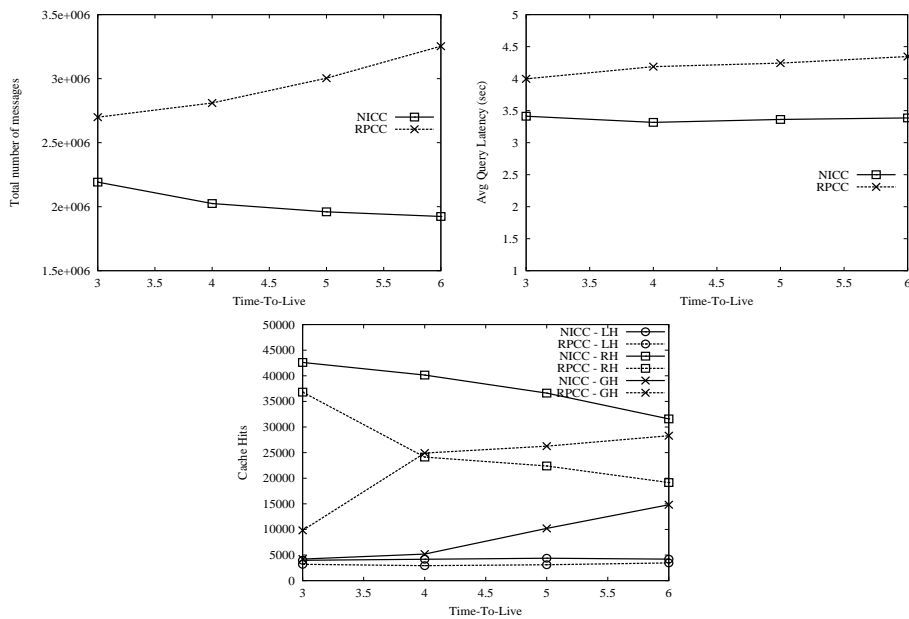
Σχήμα 4.26: % διαφορά στην κατανάλωση ενέργειας (δεδομένα μεγέθους KB) σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες, όταν εκτελούνται το NICC και RPCC.

μεταξύ των κόμβων πηγών και των διαμεσολαβητών θα αυξηθεί αφού περισσότερα μηνύματα θα πρέπει να διαδίδονται στο δίκτυο. Παρ' όλα αυτά, θα μειωθεί το κόστος επικοινωνίας μεταξύ των διαμεσολαβητών και των cache κόμβων. Ο αριθμός των κόμβων από όπου θα πρέπει να διέλθει ένα ερώτημα μέχρι να φθάσει σε ένα διαμεσολαβητή, θα μειωθεί. Αντίθετα, εάν ο αριθμός των διαμεσολαβητών είναι μικρός, τότε θα επιτύχουμε μια μείωση στο κόστος της επικοινωνίας μεταξύ των κόμβων πηγών και των διαμεσολαβητών αλλά θα αυξηθούν τα μηνύματα που θα διαχέονται στο δίκτυο, όταν θα επικοινωνούν οι cache κόμβοι με τους διαμεσολαβητές.

Για να διαπιστώσουμε ποια είναι η καλύτερη απόσταση που πρέπει να έχουν οι διαμεσολαβητές σε σχέση με τις πηγές, διεξάγουμε ένα σύνολο πειραμάτων όπου διαφοροποιούμε την τιμή  $TTL$  από 3 μέχρι 6 hops. Τα πειραματικά αποτελέσματα εμφανίζονται στο Σχήμα 4.27. Το πρωτόκολλο NICC παρουσιάζει μια ελαφρώς καλύτερη απόδοση ως προς στον αριθμό των μηνυμάτων όταν μεγαλώνει η τιμή  $TTL$ . Η καθυστέρηση πρόσβασης παραμένει όμως σταθερή σε όλες τις περιπτώσεις. Αντίθετα, στο πρωτόκολλο RPCC ο αριθμός των μηνυμάτων αυξάνεται όσο γίνεται μεγαλύτερη η τιμή  $TTL$ . Επίσης, η καθυστέρηση πρόσβασης αυξάνεται σύμφωνα με την τιμή  $TTL$ .

#### 4.5.3.4 Πειράματα αντικατάστασης της cache

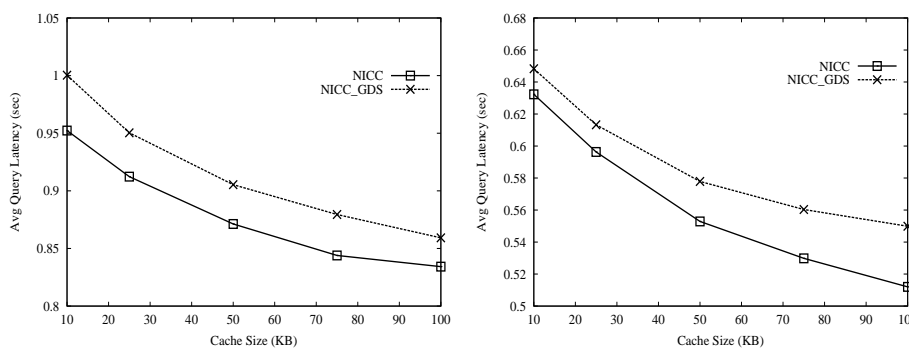
Η πολιτική αντικατάστασης της είναι ένα πολύ σημαντικό τμήμα του πρωτοκόλλου NICC. Είναι απαραίτητος ένας αποδοτικός και αποτελεσματικός αλγόριθμος αντικατάστασης για να διαχειρισθεί το μεγάλο αριθμό δεδομένων που δημιουργούνται σε ένα δίκτυο αισθητήρων. Σε αυτό το σύνολο πειραμάτων η πολιτική αντικατά-



Σχήμα 4.27: Επίδραση τιμής  $TTL$  στα: (Επάνω) αριθμό μηνυμάτων και καθυστέρηση πρόσβασης, (Κάτω) hits σε ένα ΑΠΔΑ με 500 αισθητήρες.

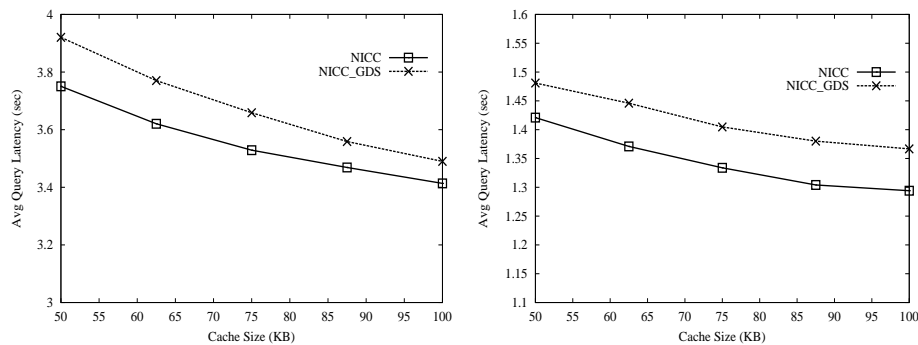
στασης της cache που παρουσιάστηκε στη Παράγραφο 4.4.3 συγκρίνεται με την πολιτική αντικατάστασης GD-Size [18].

Τα Σχήματα 4.28 και 4.29 παρουσιάζουν την επίδραση των πολιτικών αντικατάστασης στην καθυστέρηση πρόσβασης. Η προτεινόμενη πολιτική υπερτερεί έναντι της πολιτικής GD-Size για όλα τα μεγέθη της cache. Ο πρώτη εξήγηση είναι η παράμετρος  $AvgHops$  που συμπεριλαμβάνεται από τη συνάρτηση αντικατά-



Σχήμα 4.28: Επίδραση της αντικατάστασης της cache στην καθυστέρηση πρόσβασης σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες.

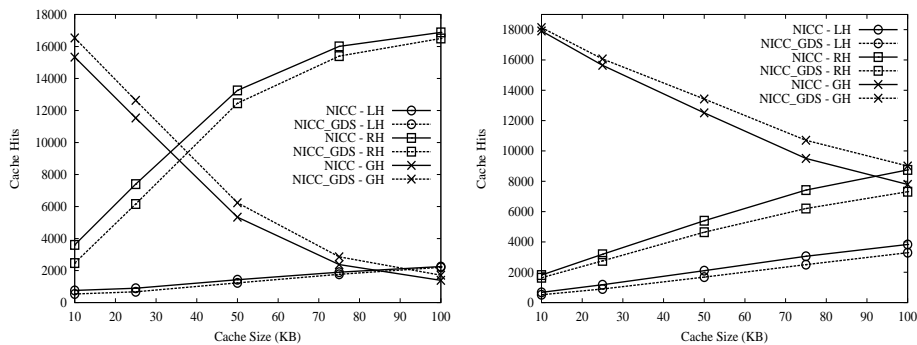
στασης κατά τη διάρκεια υπολογισμού της τιμής χρησιμότητας ενός αντικειμένου. Όσο μεγαλώνει η απόσταση μεταξύ του κόμβου που θέτει την αίτηση και του κόμβου που απαντά, τόσο πιθανότερο είναι να παραμείνει το αντικείμενο στην cache. Συνεπώς, η προτεινόμενη πολιτική προσπαθεί να κρατήσει στην cache τα αντικείμενα που έχουν ζητηθεί από μακρινούς κόμβους. Η αποθήκευση αυτών των δεδομένων μειώνει την καθυστέρηση πρόσβασης, την κατανάλωση ενέργειας και σώζει εύρος ζώνης για επακόλουθες αιτήσεις για τα ίδια αντικείμενα. Ο δεύτερος λόγος είναι ότι η πολιτική GD-Size δεν λαμβάνει υπόψη το ρυθμό ενημέρωσης και τη δημοτικότητα ενός αντικειμένου. Η προτεινόμενη πολιτική μεροληπτεί υπέρ των αντικειμένων που δεν μεταβάλλονται συχνά και έχουν μεγάλο ρυθμό προσπέλασης.



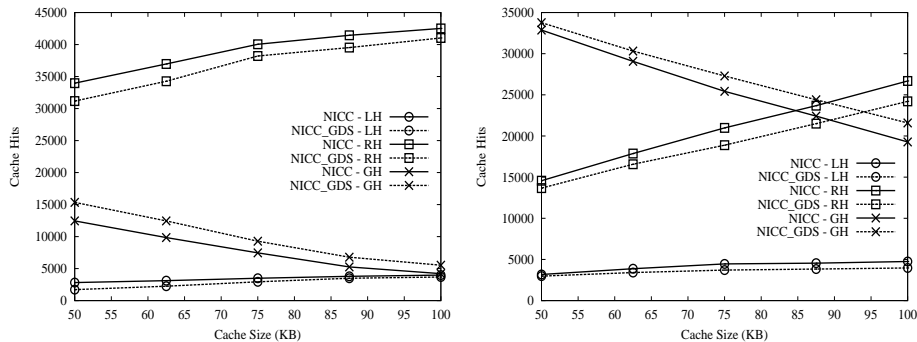
Σχήμα 4.29: Επίδραση της αντικατάστασης της cache στην καθυστέρηση πρόσβασης σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες.

Τα Σχήματα 4.30 και 4.31 αποτιμούν την απόδοση με όρους hits. Παρατηρούμε ότι η προτεινόμενη πολιτική πετυχαίνει καλύτερα αποτελέσματα στα local, remote και global cache hits συγκρινόμενα με τα αποτελέσματα που παράγονται από την πολιτική GD-Size. Αυτό συμβαίνει επειδή η προτεινόμενη πολιτική ευνοεί τις απομακρυσμένες αιτήσεις και τα μικρά δεδομένα με μεγάλη δημοτικότητα. Επιπρόσθετα, η πολιτική GD-Size δεν λαμβάνει υπόψη το ρυθμό ενημέρωσης.

Τέλος, είναι πολύ ενδιαφέρουσα η διαφορά της απόδοσης μεταξύ των αραιών και των πυκνών δικτυακών τοπολογιών (Σχήματα 4.28-4.31). Στα αραιά δίκτυα η διαφορά της απόδοσης υπέρ της προτεινόμενης πολιτικής μειώνεται όσο αυξάνεται το μέγεθος της cache. Αυτό οφείλεται στο γεγονός ότι για μεγάλα μεγέθη χωρητικότητας της cache η ανάγκη αντικατάστασης της cache μειώνεται σημαντικά, επειδή υπάρχει αρκετός χώρος για να αποθηκευθεί μεγάλος όγκος δεδομένων. Επιπρόσθετα, στα αραιά δίκτυα ο αριθμός των γειτονικών κόμβων είναι μικρός και συνεπώς τα δεδομένα που χρειάζεται να αποθηκευθούν είναι λίγα. Παρ' όλα



Σχήμα 4.30: Επίδραση της αντικατάστασης της cache στα hits σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 100 αισθητήρες.



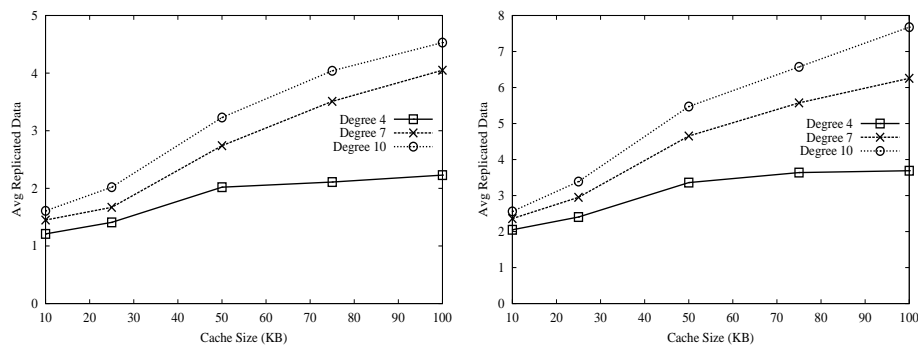
Σχήμα 4.31: Επίδραση της αντικατάστασης της cache στα hits σε ένα αραιό ( $d=4$ ) και ένα πυκνό ( $d=10$ ) ΑΠΔΑ με 500 αισθητήρες.

αυτά στα πυκνά δίκτυα η προτεινόμενη πολιτική αποδίδει καλύτερα από τη GD-Size όταν μεγαλώνει το μέγεθος της cache. Αυτό συμβαίνει εξαιτίας του μεγάλου αριθμού γειτονικών κόμβων. Όταν το μέγεθος της cache είναι πολύ μικρό οι αντικαταστάσεις των δεδομένων είναι πολύ συχνές, ενώ τα υποψήφια για αποθήκευση δεδομένα είναι πολύ περισσότερα από ότι στα αραιά δίκτυα. Συνεπώς, τα local και τα remote cache hits μειώνονται καθώς αυξάνονται τα global hits.

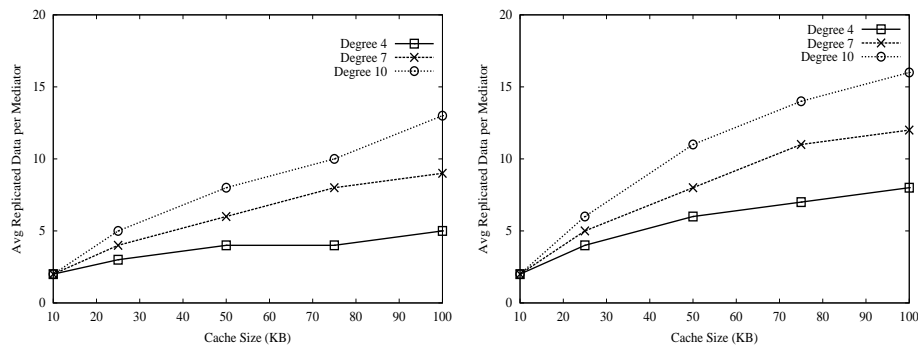
#### 4.5.3.5 Επίδραση της πυκνότητας δικτύου και του μεγέθους της cache στον αριθμό αντιγράφων και διαμεσολαβητών

Σε αυτό το σύνολο πειραμάτων εξετάζουμε την επίδραση του μεγέθους της cache και της πυκνότητας του δικτύου στον αριθμό των διαμεσολαβητών και τον αντιγράφων στο πρωτόκολλο NICC. Το Σχήμα 4.32 παρουσιάζει το μέσο όρο αντιγράφων ενός δεδομένου που έχει τοποθετηθεί στις caches των διαμεσολαβητών.

Όταν ο χώρος της cache είναι μικρός, ο αριθμός των αντιγράφων είναι επίσης μικρός. Ωστόσο, όταν αυξάνεται το μέγεθος της cache, τότε παρατηρούμε ότι ο αριθμός των αντιγράφων επίσης αυξάνεται. Παρ' όλα αυτά, όταν μεταβάλλουμε την πυκνότητα των κόμβων παρατηρούμε διαφορετική απόδοση του συστήματος. Πιο συγκεκριμένα, όταν ο αριθμός των γειτονικών κόμβων είναι μεγάλος, τότε ο αριθμός των γειτονικών διαμεσολαβητών μεγαλώνει. Συνεπώς, ο αριθμός των αντιγράφων που ωθούνται στους διαμεσολαβητές είναι επίσης μεγάλος.



Σχήμα 4.32: Μέσος όρος αντιγράφων ανά αντικείμενο για διαφορετικές τιμές βαθμού σε ένα ΑΠΔΑ με 100 και 500 αισθητήρες.

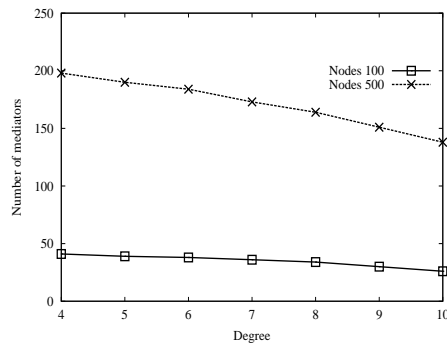


Σχήμα 4.33: Μέσος όρος αντιγράφων ανά διαμεσολαβητή για διαφορετικές τιμές βαθμού σε ένα ΑΠΔΑ με 100 και 500 αισθητήρες.

Το Σχήμα 4.33 απεικονίζει το μέσο όρο αντιγράφων που έχουν αποθηκευθεί σε κάθε διαμεσολαβητή. Είναι φανερό ότι για μεγάλες τιμές πυκνότητας κόμβων και χωρητικότητας της cache, ο αριθμός των αντιγράφων είναι επίσης μεγάλος. Αυτό συμβαίνει επειδή κάθε διαμεσολαβητής έχει περισσότερους γειτονικούς κόμβους και επομένως λαμβάνει περισσότερα αντίγραφα. Έτσι, κάθε διαμεσολαβητής μπορεί να αξιοποιήσει καλύτερα το χώρο της cache τοποθετώντας σε αυτήν

περισσότερα αντικείμενα.

Τελικά, στο Σχήμα 4.34 εξετάζουμε το μέσο όρο των διαμεσολαβητών που δημιουργούνται σε ολόκληρο το δίκτυο. Το εξαγόμενο συμπέρασμα είναι ότι για αραιά δίκτυα ο αριθμός των διαμεσολαβητών είναι αρκετά μεγάλος. Όμως, για πυκνά δίκτυα ο αριθμός των διαμεσολαβητών μειώνεται στο 25-30% του πλήθους των κόμβων του δικτύου. Αυτό οφείλεται στο γεγονός ότι πολλοί κόμβοι ανιχνεύουν τους ίδιους κόμβους ως σημαντικούς.



Σχήμα 4.34: Μέσος όρος διαμεσολαβητών για διαφορετικές τιμές βαθμού σε ένα ΑΠΔΑ με 100 και 500 αισθητήρες.

## 4.6 Συμπεράσματα

Στο παρόν κεφάλαιο διαπραγματευθήκαμε το ζήτημα της συνέπειας της cache αισθητήρων που εξυπηρετούνται από ασύρματα δίκτυα μέσω της αποστολής δεδομένων σε ένα κανάλι εκπομπής.

Καταγράψαμε τα μειονεκτήματα των μέχρι στιγμής γνωστών πολιτικών και θέσαμε όρους για τα χαρακτηριστικά που θα πρέπει να παρουσιάζει μια πολιτική συνέπειας της cache για πολυμεσικά δεδομένα σε ΑΠΔΑ. Περιγράψαμε την πολιτική NICC και τη συγκρίναμε με την ανταγωνιστική της μέθοδο ως προς τον αριθμό των hits (local, remote, global), την καθυστέρηση προσπέλασης, τον αριθμό των διαδιδόμενων μηνυμάτων που και την καταναλισκόμενη ενέργεια από τους αισθητήρες. Πρακτικά, η προτεινόμενη πολιτική αποτιμήθηκε με το προσομοιωτή J-Sim και τα αποτελέσματα αποδεικνύουν την ανωτερότητα της έναντι της ανταγωνιστικής της. Σύμφωνα με την πολιτική NICC κάθε κόμβος ανιχνεύει τους κόμβους που βρίσκονται στα κεντρικότερα σημεία των δικτυακών γειτονιών και αναθέτει σε αυτούς το ρόλο του διαμεσολαβητή για να συντονίζουν τις απο-

φάσεις που σχετίζονται με την εγκυρότητα και τη συνέπεια των δεδομένων που βρίσκονται στις caches.

Στην παρούσα μελέτη ασχοληθήκαμε επίσης με το θέμα της αντικατάστασης της cache και προτείναμε μια νέα πολιτική που λαμβάνει υπόψη το ρυθμό προσέλασης και ενημέρωσης κάθε αντικειμένου, το μέγεθος του αντικειμένου και την απόσταση των κόμβων που ζητούν το αντικείμενο. Η πολιτική αντικατάστασης αποτιμήθηκε πειραματικά και συγκρίθηκε με την πολιτική αντικατάστασης GD-Size. Τα αποτελέσματα αποδεικνύουν μια σαφή υπεροχή της προτεινόμενης πολιτικής.





## ΚΕΦΑΛΑΙΟ 5

# Αποδοτικό Συνεργατικό Caching σε Ασύρματα Δίκτυα Αισθητήρων\*

### Περιεχόμενα

---

5.1	ΕΙΣΑΓΩΓΗ . . . . .	109
5.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	112
5.3	ΜΟΝΤΕΛΟ ΔΙΚΤΥΟΥ . . . . .	112
5.4	ΝΕΑ ΣΧΗΜΑΤΑ ΣΥΝΕΡΓΑΤΙΚΟΥ CACHING	113
5.5	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	124
5.6	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	133

---

### 5.1 Εισαγωγή

Στο Κεφάλαιο 2 αναφερθήκαμε αναλυτικά στις προκλήσεις που αντιμετωπίζονται κατά τη σχεδίαση και υλοποίηση των ΑΔΑ. Ας εξετάσουμε για παράδειγμα ένα δίκτυο αισθητήρων, το οποίο έχει αναπτυχθεί σε ένα πολεμικό πεδίο, όπου οι αισθητήρες έχουν διασκορπισθεί σε μια μεγάλη περιοχή. Κάθε αισθητήρας είναι εξοπλισμένος με μια μικροκάμερα που μπορεί να τραβήξει φωτογραφίες μιας

\*Το κεφάλαιο αντλεί τα περιεχόμενά του από την εργασία [45].

πολύ μικρής περιοχής γύρω από τη θέση του. Οι αισθητήρες ενημερώνουν για τις φωτογραφίες που λαμβάνουν και τις διαμοιράζονται μεταξύ τους ώστε να μπορέσουν να δημιουργήσουν μια περισσότερο ολοκληρωμένη και ακριβή εικόνα της παρακολουθούμενης περιοχής. Ο διαμοιρασμός είναι απαραίτητος επειδή κάθε μικροκάμερα μπορεί να συλλάβει μια περιορισμένη όψη της συνολικής περιοχής είτε εξαιτίας της θέσης του αισθητήρα είτε επειδή υπάρχουν εμπόδια κοντά του. Συνεπώς, κάθε αισθητήρας μπορεί να ζητήσει και να λάβει ένα μεγάλο αριθμό φωτογραφιών, που έχουν παρθεί από κάποιους άλλους αισθητήρες μέσω multi-hop επικοινωνίας. Κατόπιν, κάθε αισθητήρας είναι σε θέση να απαντήσει σε ερωτήματα σχετικά με “υψηλού επιπέδου” γεγονότα, π.χ. παρουσία του εχθρού σε μια περιοχή.

Σε ένα εχθρικό περιβάλλον, είναι φανερό ότι η φόρτιση της μπαταρίας του αισθητήρα δεν είναι ένας εύκολος και εφικτός στόχος. Επίσης, οι τοποθεσίες των αισθητήρων δεν έχουν αποφασισθεί από κάποιον αλγόριθμο τοποθέτησης και συνεπώς η επικοινωνία μεταξύ των αισθητήρων είναι αυστηρά multi-hop.

Η πλειοψηφία των εφαρμογών που εκτελούνται σε ΑΔΑ απαιτούν τη βελτιστοποίηση των επικοινωνιών μεταξύ των αισθητήρων ώστε να εξυπηρετούν τη ζήτηση για δεδομένα με μικρή καθυστέρηση προσπέλασης και με ελάχιστη κατανάλωση ενέργειας. Η διάρκεια ζωής της μπαταρίας μπορεί να επεκταθεί εάν καταφέρουμε να μειώσουμε τον όγκο της επικοινωνίας. Έτσι, η τεχνική του caching των δεδομένων για κάθε αισθητήρα είτε στον τοπικό αποθηκευτικό χώρο είτε στους γειτονικούς αισθητήρες μπορεί να βοηθήσει στην επιμήκυνση της ζωής του δικτύου. Επιπρόσθετα, το caching μπορεί πολύ αποδοτικά να μειώσει την ανάγκη για μεταδόσεις σε μεγάλο εύρος του δικτύου και επομένως να μειώσει τις παρεμβολές και να αντιμετωπίσει τις συνθήκες μεταβλητού καναλιού. Το συνεργατικό caching έχει προταθεί ως μια αποτελεσματική και αποδοτική τεχνική για την επίτευξη των προηγούμενων στόχων [44, 117, 149] (για περισσότερες λεπτομέρειες δες την Παράγραφο 3.2).

Η θεμελιώδης άποψη σε όλα τα προτεινόμενα σχήματα συνεργατικού caching για δίκτυα αισθητήρων είναι η εύρεση των κόμβων που θα υλοποιήσουν το στόχο της συνεργασίας και θα λάβουν τις αποφάσεις caching, π.χ., προς ποιους κόμβους θα προωθηθούν τα αιτήματα για δεδομένα; ποιοι κόμβοι θα αποφασίσουν σχετικά με το ποια δεδομένα θα τοποθετηθούν στις caches ποιών κόμβων;

### 5.1.1 Κίνητρο και συνεισφορές

Οι αρχικές μελέτες στο συνεργατικό caching σε AΔΑ και MANET [86, 149], δεν έλαβαν υπόψη την επιλογή των κόμβων που θα έχουν ειδικούς ρόλους στο πρωτόκολλο συνεργασίας. Η μελέτη [44] κατέδειξε τη σπουδαιότητα της προσεκτικής επιλογής αυτών των κόμβων καθώς θα πρέπει να βρίσκονται σε κεντρικά σημεία της τοπολογίας. Με βάση αυτήν την ιδέα προτάθηκε ένα συνεργατικό σχήμα, όπου οι αισθητήρες ειδικού ρόλου επιλέγονταν με βάση την ικανότητα να επηρεάζουν την επικοινωνία μεταξύ ζευγών άλλων κόμβων. Αυτή η ικανότητα ποσοτικοποιήθηκε με τον υπολογισμό του  $NI$  για κάθε αισθητήρα. Η μετρική αποτελεί μια τοπική έκδοση του πολύ γνωστού δείκτη της *ενδιάμεσης κεντρικότητας* (betweenness centrality), ο οποίος χρησιμοποιείται στα κοινωνικά δίκτυα.

Παρ' όλα αυτά, η μετρική  $NI$  παρουσιάζει τα εξής μειονεκτήματα:

- Ο υπολογισμός της από ένα αισθητήρα απαιτεί λεπτομερή γνώση της συνδεσιμότητας των 1-hop γειτόνων του, π.χ., ο αισθητήρας πρέπει να ανταλλάσει το σύνολο των 1-hop γειτόνων του με κάθε ένα 1-hop γείτονά του. Συνεπώς, μεγαλύτερα πακέτα διαδίδονται στο δίκτυο.
- Αν και είναι ο υπολογισμός του δείκτη  $NI$  είναι αρκετά γρήγορος, εντούτοις δεν αποτελεί μια λειτουργία με σταθερή πολυπλοκότητα  $O(1)$ , το οποίο μπορεί να είναι μειονέκτημα όταν η δικτυακή τοπολογία αλλάζει αρκετά γρήγορα.
- Ο δείκτης  $NI$  μπορεί να παραπλανήσει επειδή επηρεάζεται πολύ από την ύπαρξη απομονωμένων κόμβων στα όρια του δικτύου (αυτό το θέμα θα εξηγηθεί στην Παράγραφο 5.4.1 μέσω ενός παραδείγματος).
- Όταν ο δείκτης  $NI$  υπολογίζεται με βάση μεγαλύτερες γειτονιές και όχι απλά για την 2-hop γειτονιά ενός κόμβου, μπορεί να αλλάξει σημαντικά επειδή εισέρχονται στον υπολογισμό περισσότεροι κόμβοι και συνδέσεις.

Με κίνητρο τις προαναφερθείσες αδυναμίες, στο κεφάλαιο αυτό παρουσιάζουμε δύο νέες μετρικές για να αποτιμήσουμε τη σπουδαιότητα ενός κόμβου, έτσι ώστε να αναλάβει ειδικούς ρόλους στο περιβάλλον συνεργασίας. Βασισμένοι σε αυτές τις μετρικές, περιγράφουμε τα νέα πρωτόκολλα συνεργατικού caching. Πιο συγκεκριμένα, το κεφάλαιο αυτό εισάγει τις εξής συνεισφορές:

- Περιγράφει μια νέα μετρική κεντρικότητας για κόμβους, η οποία ονομάζεται  *$\mu$ -power community index ( $\mu$ -PCI)*. Η νέα μετρική είναι περισσότερο

πληροφοριακή από το βαθμό κόμβου, είναι πιο εύρωστη και υπολογίζεται ταχύτερα από τη μετρική  $NI$ . Επιπρόσθετα, δεν επηρεάζεται από τους απομονωμένους κόμβους.

- Περιγράφει επίσης μια επεκταμένη έκδοση της μετρικής  $\mu\text{-PCI}$ , την  $\text{extPCI}$  που βελτιώνει την επιλογή των αισθητήρων ειδικού ρόλου. Ο υπολογισμός της μετρικής  $\text{extPCI}$  απαιτεί το ίδιο κόστος επικοινωνίας με τη μετρική  $NI$ , αλλά είναι πιο εύρωστη από τις μετρικές  $\mu\text{-PCI}$  και  $NI$ . Τέλος δεν επηρεάζεται επίσης από τους απομονωμένους κόμβους.
- Αναπτύσσει τα νέα πρωτόκολλα συνεργατικού caching για ΑΔΑ, τα οποία ονομάζονται *PCI Cooperative Caching* (PCICC) και *extPCICC*. Τα πρωτόκολλα συγκρίνονται με την καλύτερη γνωστή μέθοδο μέσω πειραματικών προσομοιώσεων.

## 5.2 Σχετικές Εργασίες

Οι σχετικές εργασίες που πραγματοποιήθηκαν στα πλαίσια του caching και του συνεργατικού caching τόσο σε ΑΔΑ όσο και σε ΑΠΔΑ έχουν παρουσιασθεί αρχικά στο Κεφάλαιο 2 και περισσότερο λεπτομερώς στη Παράγραφο 3.2. Στη επόμενη παράγραφο θα περιγράψουμε το μοντέλο του δικτύου.

## 5.3 Μοντέλο Δικτύου

Στο παρόν κεφάλαιο υποθέτουμε ότι ένα ΑΔΑ αποτελείται από  $N$  αισθητήρες. Καμία υπόθεση δεν γίνεται σχετικά με τη διάμετρο και την πυκνότητα του δικτύου. Θεωρούμε επίσης τις εξής γενικές ιδιότητες ενός ΑΔΑ:

- Οι κόμβοι είναι στατικοί, οι σύνδεσμοι επικοινωνίας είναι διπλής κατεύθυνσης, ενώ οι αισθητήρες επικοινωνούν με multi-hop τρόπο.
- Οι υπολογιστικές και επικοινωνιακές δυνατότητες είναι οι ίδιες για όλους τους κόμβους. Επιπλέον, οι αισθητήρες δεν χρειάζεται να είναι εφοδιασμένοι με υλικό τύπου GPS.
- Οι αισθητήρες χρησιμοποιούν ένα πρωτόκολλο δρομολόγησης (π.χ., AODV) για να στείλουν αιτήσεις στους κόμβους πηγές. Κάθε κόμβος έχει χωρητικότητα cache ίση με  $C$  bytes, και υπολογίζει το δείκτη  $PCI$ . Οι γειτονικοί

κόμβοι με το μεγαλύτερες τιμές  $PCI$  ορίζονται ως κοινοτικοί (community) caching κόμβοι. Συνεπώς, ένα κόμβος μπορεί να βρίσκεται σε δύο καταστάσεις: κοινός ή κοινοτικός caching κόμβος (CCN).

- Όπως στις μελέτες [149, 44] και χωρίς βλάβη της γενικότητας, θεωρούμε ότι ο παροχέας των δεδομένων είναι η πηγή δεδομένων. Έτσι, κάθε αίτημα μπορεί να ικανοποιηθεί ακόμα και αν το δεδομένο έχει απομακρυνθεί από τις caches όλων των αισθητήρων. Η ύπαρξη της πηγής των δεδομένων δεν επηρεάζει την απόδοση των πρωτοκόλλων.

## 5.4 Νέα Σχήματα Συνεργατικού Caching

Στο Κεφάλαιο 3.1 δώσαμε έμφαση στην ανακάλυψη των γειτόνων που είναι κεντρικότεροι σε ένα δίκτυο αισθητήρων. Το αντικείμενο αυτό έχει περιγραφεί επίσης στη μελέτη [49]. Η έννοια της κεντρικότητας επιδέχεται πολλές ερμηνείες. Ο δείκτης  $NI$  παρουσιάζει κάποια μειονεκτήματα που περιγράφηκαν στην Παράγραφο 5.1.1. Στο παρόν κεφάλαιο ακολουθούμε μια άλλη διαδρομή για να ορίσουμε ένα νέο δείκτη κεντρικότητας, ο οποίος αποτελεί αμάλγαμα μεταξύ του δείκτη  $NI$  και του βαθμού του κόμβου. Η επόμενη παράγραφος παρέχει τον ορισμό του δείκτη, ενώ η Παράγραφος 5.4.2 περιγράφει μια επεκταμένη εκδοχή του δείκτη κεντρικότητας. Εν τέλει, στην Παράγραφο 5.4.3 περιγράφονται στα συστατικά στοιχεία του πρωτοκόλλου συνεργατικού caching.

### 5.4.1 Power community index

Ένα ΑΔΑ παρίσταται ως ένας γράφος  $G(V, E)$ , όπου  $V$  είναι το σύνολο των κόμβων αισθητήρων  $SN_1, SN_2, \dots$ , και  $E \subseteq V \times V$  είναι το σύνολο των ασύρματων συνδέσεων μεταξύ των κόμβων. Μια ακμή  $e = (u, v)$ , όπου  $u, v \in E$  υπάρχει εάν και μόνο εάν ο κόμβος  $u$  βρίσκεται στο πεδίο μετάδοσης του κόμβου  $v$ . Όλοι οι σύνδεσμοι του γράφου είναι διπλής κατεύθυνσης, δηλαδή, εάν ο κόμβος  $u$  βρίσκεται στο πεδίο μετάδοσης του κόμβου  $v$ , τότε επίσης ο κόμβος  $v$  βρίσκεται στο πεδίο μετάδοσης του κόμβου  $u$ . Βαθμός  $D_i$  ενός κόμβου αισθητήρα  $SN_i$  είναι το πλήθος των άμεσων συνδέσεων του  $SN_i$  με τους υπόλοιπους κόμβους. Το δίκτυο θεωρείται ότι βρίσκεται σε συνδεδεμένη κατάσταση.

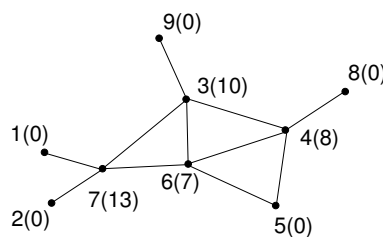
Σε ορισμένες περιπτώσεις, ο βαθμός ενός αισθητήρα είχε χρησιμοποιηθεί ως μετρική για την κεντρικότητα. Παρατηρώντας το Σχήμα 5.1, βλέπουμε ότι οι κόμβοι 3, 4, 7 και 6 είναι εξίσου κεντρικοί, καθώς ο βαθμός τους είναι 4 (οι

αριθμοί στις παρενθέσεις). Επιπρόσθετα, εάν υπολογίσουμε την ενδιάμεση κεντρικότητα για κάθε αισθητήρα του συνολικού γράφου, τότε ο κόμβος 7 είναι ο κεντρικότερος, ακολουθούμενος από τους κόμβους 3 και 4, ενώ στη συνέχεια έρχεται ο κόμβος 6. Αυτό είναι κάπως διαισθητικό επειδή ο κόμβος 6 έχει όλους τους κόμβους του δικτύου στη γειτονιά του (σε μια απόσταση 2-hops). Ξεκινώντας από αυτή την παρατήρηση, προτείνουμε μια νέα μετρική κεντρικότητας, την οποία ονομάζουμε  $\mu$ -Power Community Index ( $\mu$ -PCI) και την ορίζουμε ως εξής:

**ΟΡΙΣΜΟΣ 5.1.** Ο δείκτης  $\mu$ -Power Community Index ενός αισθητήρα  $v$  είναι ίσος με  $k$ , εάν δεν υπάρχουν περισσότεροι από  $\mu \times k$  αισθητήρες στην  $\mu$ -hop γειτονιά του  $v$  με βαθμό ίσο ή μεγαλύτερο από  $k$ , ενώ οι υπόλοιποι αισθητήρες που βρίσκονται μέσα στην περιοχή έχουν βαθμό ίσο ή μικρότερο του  $k$ .

Σε αυτό το σημείο πρέπει να σημειώσουμε ότι καθώς υπολογίζουμε το βαθμό των αισθητήρων που βρίσκονται στην περιοχή που ορίζεται από την  $\mu$ -hop γειτονιά του αισθητήρα  $v$ , είναι πιθανό στον υπολογισμό του βαθμού να συνεισφέρουν συνδέσεις σε αισθητήρες εκτός της περιοχής. Τέτοιες περιπτώσεις δεν εξαιρούνται από τον ανωτέρω ορισμό (και δεν θα πρέπει) επειδή ο ορισμός προσπαθεί να ανακαλύψει τους αισθητήρες εκείνους που μπορούν να ασκήσουν τη μέγιστη επιρροή στους άλλους αισθητήρες.

Είναι φανερό ότι οι κόμβοι με μεγαλύτερο βαθμό είναι πιθανότερο να είναι ισχυροί, επειδή μπορούν απευθείας να επηρεάσουν περισσότερους κόμβους. Παρ' όλα αυτά, η ισχύς των κόμβων εξαρτάται επίσης και από το βαθμό των 1-hop γειτόνων τους. Μεγάλη τιμή για το δείκτη  $\mu$ -PCI ενός κόμβου  $v$  δηλώνει ότι ο αισθητήρας  $v$  μπορεί να επικοινωνήσει με άλλους αισθητήρες μέσω σχετικά σύντομων διαδρομών (όπως συμβαίνει με το δείκτη  $NI$ ) ή ότι ο αισθητήρας  $v$  βρίσκεται σε μια πυκνή περιοχή του δικτύου.

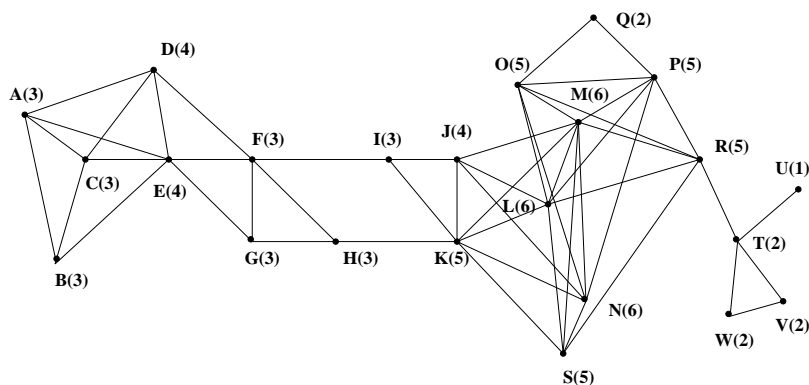


Σχήμα 5.1. Οι δείκτες  $NI$  για ένα μικρό γράφο.

Καθώς ο στόχος μας είναι οι εφαρμογές για ΑΔΑ, το ενδιαφέρον μας εστιάζεται σε μια τοπική έκδοση της μετρικής. Θέτοντας  $\mu=1$  λαμβάνουμε τον απλό δείκτη Power Community Index ( $PCI$ ), ο οποίος ορίζεται ως εξής:

**ΟΡΙΣΜΟΣ 5.2.** Ο δείκτης *Power Community Index*  $PCI(v)$  ενός αισθητήρα  $v$  ισούται με  $k$ , εάν δεν υπάρχουν περισσότεροι από  $k$  1-hop γείτονες του  $v$  με βαθμό ίσο ή μεγαλύτερο του  $k$ , ενώ οι υπόλοιποι 1-hop γείτονες του  $v$  έχουν βαθμό ίσο ή μικρότερο του  $k$ .

Σύμφωνα με τον ανωτέρω ορισμό, παρατηρούμε στο Σχήμα 5.1 ότι  $PCI(7) = PCI(4)=2$ , όπου  $PCI(6)=PCI(3)=3$ . Ο υπολογισμός των δεικτών  $PCI$  για ένα μεγαλύτερο γράφο παρουσιάζεται στο Σχήμα 5.2, όπου κάθε κόμβος χαρακτηρίζεται από ένα ζεύγος  $ID(PCI)$ .



Σχήμα 5.2. Υπολογισμός του  $PCI$  για ένα γράφο.

Είναι φανερό ότι όταν υπολογίζουμε τους  $\mu$ - $PCI$  δείκτες για κάθε αισθητήρα αποκτούμε μια πολύ εμπειριστωμένη εικόνα ως προς τους αισθητήρες που βρίσκονται σε σημαντικά θέσεις του δικτύου. Παρ'όλα αυτά, εξαιτίας των περιορισμών στην ενέργεια και στο εύρος ζώνης που επιβάλλονται από το ΑΔΑ, οι προσπάθειες μας επικεντρώνονται στον καταναμημένο τοπικό υπολογισμό του  $\mu$ - $PCI$ , π.χ.  $PCI$ . Επομένως, θεωρούμε μόνο τον  $PCI$  δείκτη ενός αισθητήρα στην προσπάθειά μας να μεγιστοποιήσουμε τα οφέλη σε ότι αφορά την επικοινωνία και την ενέργεια.

Ο ορισμός του  $PCI$  προσφέρει κάποια σημαντικά πλεονεκτήματα επειδή κάθε αισθητήρας προκειμένου να υπολογίσει το  $PCI$  χρειάζεται να ανταλλάξει με τους 1-hop γείτονές του μόνο τους βαθμούς και όχι τη λεπτομερή συνδεσιμότητα. Σε αυτό το σημείο πρέπει να σημειώσουμε ότι ο  $NI$  δείκτης [44] είναι βασισμένος

στην εύρεση του αριθμού των συντομότερων μονοπατιών που διέρχονται από ένα κόμβο. Ο υπολογισμός του δείκτη  $NI$  υποθέτει ότι κάθε αισθητήρας γνωρίζει την 1-hop γειτονιά κάθε ενός 1-hop γείτονά του. Συνεπώς, ο υπολογισμός του  $NI$  περιλαμβάνει τρία βήματα: (α) αρχικά κάθε κόμβος εκπέμπει το αναγνωριστικό του (ID), (β) κάθε αισθητήρας εκπέμπει τα αναγνωριστικά των 1-hop γειτόνων του, και (γ) κάθε κόμβος υπολογίζει τα συντομότερα μονοπάτια για το τοπικό δίκτυο και αποφασίζει σχετικά με τους σημαντικούς γείτονες. Σε αντίθεση με το δείκτη  $NI$ , ο υπολογισμός του  $PCI$  απαιτεί στο δεύτερο βήμα από κάθε αισθητήρα την εκπομπή μόνο του ακεραίου αριθμού που αντιστοιχεί στο βαθμό. Τέλος, στο τρίτο βήμα δεν χρειάζεται να πραγματοποιηθούν εξεζητημένοι υπολογισμοί. Συνεπώς, ο υπολογισμός του  $PCI$  συνεπάγεται λιγότερο υπολογιστικό και επικοινωνιακό κόστος.

#### 5.4.2 Εκτεταμένος power community index

Είναι φανερό ότι ένα κόμβος είναι ισχυρός και μπορεί να επηρεάσει άλλους κόμβους όταν οι βαθμοί των 1-hop γειτόνων του είναι μεγάλοι. Σε αυτή την περίπτωση, οι μεγάλες τιμές για το δείκτη  $PCI$  ενός κόμβου  $v$  δείχνουν ότι ο κόμβος μπορεί να επικοινωνήσει με άλλους κόμβους μέσω σχετικά σύντομων μονοπατιών ή ότι ο κόμβος  $v$  βρίσκεται σε μια πυκνή περιοχή του δικτύου. Αν και η μετρική  $PCI$  είναι πολύ χρήσιμη με μικρό υπολογιστικό και επικοινωνιακό κόστος, δεν μπορεί πάντα να αναδείξει τη σημαντικότητα ενός κόμβου. Ο δείκτης  $PCI$  ενός κόμβου  $v$  υπολογίζεται βασισμένος στους βαθμούς των αισθητήρων που είναι 1-hop γείτονες. Όμως ο βαθμός ενός αισθητήρα μπορεί να περιλαμβάνει συνδέσεις σε αισθητήρες που είναι επίσης 1-hop γείτονες του  $v$ . Επομένως είναι φανερό ότι ο υπολογισμός του  $PCI$  μπορεί να επηρεασθεί από τις συνδέσεις που εγκαθιδρύονται μεταξύ 1-hop γειτόνων ενός κόμβου  $v$ . Σε αυτή την περίπτωση, ο κόμβος  $v$  δεν επηρεάζει πολλούς διαφορετικούς κόμβους, επειδή η πλειονότητα των αισθητήρων βρίσκονται στο πεδίο επικοινωνίας του  $v$ .

Εξετάζοντας το Σχήμα 5.3 παρατηρούμε ότι ο κόμβος  $A$  έχει την ίδια τιμή  $PCI$  και για τις δύο τοπολογίες, επειδή οι κόμβοι  $B, C, D, E$ , έχουν τον ίδιο βαθμό. Οι κόμβοι  $B, C$  έχουν βαθμό ίσο με 4, ενώ ο κόμβος  $D$  ίσο με 5 και ο κόμβος  $E$  ίσο με 6. Έτσι, η τιμή  $PCI$  του κόμβου  $A$  είναι ίση με 4 και για τις δύο τοπολογίες. Αυτό όμως είναι κάπως μη ακριβές επειδή ο κόμβος  $A$  έχει στη γειτονιά του (σε απόσταση 2-hops μακριά) 7 κόμβους στην αριστερή τοπολογία και 16 κόμβους στη δεξιά τοπολογία. Αυτό το αποτέλεσμα αναδεικνύει ότι ο δείκτης  $PCI$  δεν μπορεί πάντα να συλλάβει τη σπουδαιότητα ενός κόμβου και αυτό



συμβαίνει όταν υπάρχουν κλίκες, δηλαδή πλήρεις υπογράφοι, όπου όλοι οι κόμβοι τους είναι συνδεδεμένοι μεταξύ τους. Η ύπαρξη κλικών σε ένα γράφο υπολογίζεται από τη μετρική του *συντελεστή ομαδοποίησης* (clustering coefficient) [136], η οποία παρέχει τοπική πληροφορία και υπολογίζεται ως εξής:

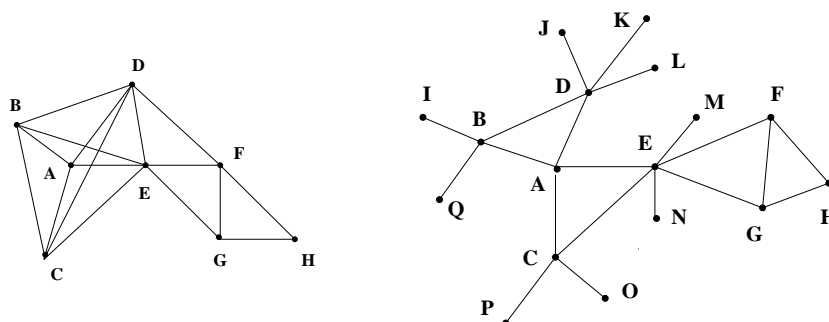
$$C(v) = \frac{2 \times L_v}{d_v \times (d_v - 1)}. \quad (5.1)$$

όπου  $d_v$  είναι ο βαθμός ενός κόμβου  $v$ , ενώ  $L_v$  είναι το πλήθος των συνδέσεων μεταξύ των  $d_v$  γειτόνων του κόμβου  $v$ . Ο συντελεστής ομαδοποίησης του δικτύου είναι ο μέσος όρος όλων των  $C_v$ . Είναι φανερό ότι οι μικρές τιμές του  $C(v)$  φανερώνουν ότι οι γείτονες του κόμβου  $v$  έχουν λίγες συνδέσεις μεταξύ τους.

Ξεκινώντας από την προηγούμενη παρατήρηση, δηλαδή ότι το *PCI* δεν μπορεί πάντα να φανερώνει τη σημαντικότητα ενός κόμβου, και λαμβάνοντας υπόψη το συντελεστή ομαδοποίησης, προτείνουμε μια νέα μετρική κεντρικότητας, την οποία ονομάζουμε *extended Power Community Index* (*extPCI*) και υπολογίζουμε ως εξής:

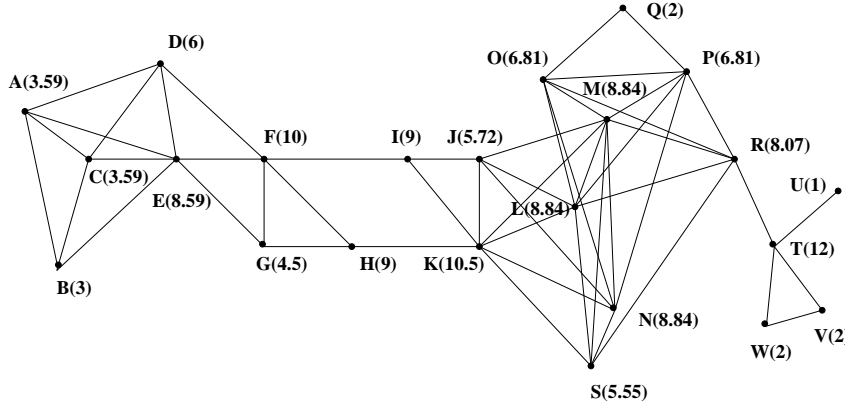
$$extPCI(v) = \frac{PCI(v)}{C(v)} \quad (5.2)$$

Ο δείκτης *extPCI* προσφέρει περισσότερα πλεονεκτήματα επειδή η υλοποίηση του είναι απλή και το υπολογιστικό κόστος χαμηλό. Αν και ο δείκτης *PCI* έχει χαμηλό κόστος επικοινωνίας, η λεπτομερής συνδεσιμότητα που εκπέμπεται από κάθε αισθητήρα στο *extPCI* έχει ως αποτέλεσμα μια ακριβέστερη επιλογή των κόμβων ειδικού ρόλου. Ο υπολογισμός του *extPCI* περιλαμβάνει τρία βήματα (όπως και του *NI*): (α) κάθε κόμβος εκπέμπει το αναγνωριστικό του, (β) κάθε αισθητήρας εκπέμπει τα αναγνωριστικά των 1-hop γειτόνων του, και (γ) κάθε κόμβος  $v$  υπολογίζει το *extPCI*( $v$ ). Το *extPCI* όπως και το *PCI* αποδίδουν καλύτερα από το *NI* επειδή το *NI* επηρεάζεται πολύ από την ύπαρξη των απομο-



Σχήμα 5.3: Ένα παράδειγμα μη ορθής συμπεριφοράς του *PCI*.

νωμένων κόμβων στα όρια του δικτύου. Ο υπολογισμός των δεικτών  $extPCI$  για ένα μεγάλο γράφο φαίνεται στο Σχήμα 5.4, όπου κάθε κόμβος χαρακτηρίζεται από ένα ζεύγος  $ID(extPCI)$ .



Σχήμα 5.4. Υπολογισμός του δείκτη  $extPCI$  για ένα γράφο.

### 5.4.3 Πρωτόκολλα $extPCICC$ και $PCICC$

Σε αυτή την παράγραφο θα παρουσιάσουμε τα προτεινόμενα σχήματα συνεργατικού caching, π.χ., τα μεταδεδομένα που αποθηκεύονται σε κάθε αισθητήρα, τον αλγόριθμο για την επιλογή των αισθητήρων που θα συντονίσουν τις αποφάσεις caching, την προώθηση των αιτημάτων για δεδομένα, την πολιτική αντικατάστασης της cache και τη στρατηγική αποδοχής στη cache. Τα δύο πρωτόκολλα ακολουθούν τις ίδιες αρχές εκτός από το γεγονός ότι το  $PCICC$  υπολογίζει τη μετρική  $PCI$ , ενώ το  $extPCICC$  υπολογίζει τη μετρική  $extPCI$ . Στη συνέχεια θα παρουσιάσουμε τη λειτουργία του πρωτοκόλλου  $extPCICC$  και θα εξηγήσουμε μόνο τις διαφορετικές λειτουργίες του πρωτοκόλλου  $PCICC$ .

Αρχικά, υποθέτουμε ότι κάθε αισθητήρας είναι γνώστης των 1-hop γειτόνων του. Αυτό επιτυγχάνεται με την ανταλλαγή των “HELLO” μηνυμάτων. Στη συνέχεια, κάθε κόμβος εκπέμπει τους 1-hop γείτονές του στη γειτονιά του. Συνεπώς, κάθε κόμβος μπορεί να αποκτήσει τη λίστα των 1-hop γειτόνων του και τους 1-hop γείτονες κάθε γειτονικού κόμβου. Η μόνη διαφορά μεταξύ του πρωτοκόλλου  $extPCICC$  και του πρωτοκόλλου  $PCICC$  είναι ότι κάθε κόμβος που εκτελεί το πρωτόκολλο  $PCICC$  εκπέμπει μόνο το βαθμό του στους γείτονές του. Η πληροφορία που μεταδίδεται στο  $PCICC$  είναι μόνο δύο αριθμοί, δηλαδή το αναγνωριστικό του αισθητήρα και ο βαθμός του. Επομένως, κάθε κόμβος μπορεί

να αποκτήσει την λίστα των 1-hop γειτόνων του και το βαθμό που έχει κάθε γειτονικός κόμβος. Θεωρούμε ότι είμαστε σε θέση να καθορίσουμε τα χρονικά διαστήματα (time slots) εκπομπής των κόμβων έτσι ώστε να μην συμβαίνουν παρεμβολές, π.χ., δύο κόμβοι δεν μεταδίδουν στο ίδιο χρονικό διάστημα. Ένα τέτοιο σχήμα είναι ο αλγόριθμος D2-coloring που περιγράφεται στη μελέτη [56].

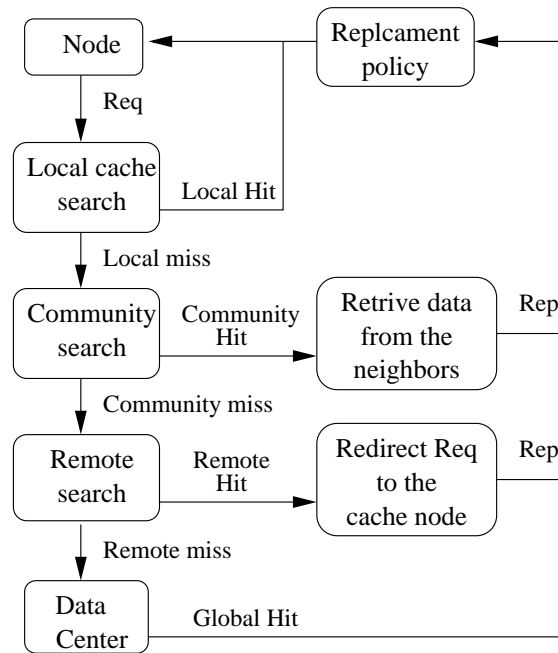
Στη συνέχεια, κάθε αισθητήρας υπολογίζει και εκπέμπει τη μετρική *extPCI* (ενώ σύμφωνα με το πρωτόκολλο PCICC κάθε κόμβος υπολογίζει και μεταδίδει τη μετρική *PCI*). Αυτό συμβαίνει επειδή κάθε κόμβος χρειάζεται τον extended Power Community Index (Power Community Index για το πρωτόκολλο PCICC) των γειτονικών κόμβων για να χαρακτηρίσει κάποιους από τους γειτονικούς κόμβους ως κοινοτικούς caching κόμβους (CCNs). Ένα κόμβος  $v$  καθορίζει τους δικούς του CCNs ως το ελάχιστο σύνολο των 1-hop αισθητήρων με τις μεγαλύτερες τιμές *extPCI* που καλύπτουν τη 2-hop γειτονιά του κόμβου  $v$ .

Θεωρούμε επίσης ότι κάθε αισθητήρας είναι σε γνώση της εναπομείνουσας ενέργειας και του ελεύθερου χώρου της cache. Επιπρόσθετα, κάθε κόμβος αποθηκεύει τα εξής δεδομένα και μεταδεδομένα:

- Το αναγνωριστικό του αντικειμένου και το ίδιο το αντικείμενο.
- Την καθυστέρηση πρόσβασης για να αποκτηθεί ένα αντικείμενο.
- Το μέγεθος  $Size_i$  ενός αντικειμένου  $i$ .
- Το χρονικό διάστημα *TTL* (Time-To-Live) για κάθε αντικείμενο.
- Για κάθε αντικείμενο της cache ο κόμβος αποθηκεύει τις χρονικές στιγμές των  $K$  πιο πρόσφατων προσπελάσεων του αντικειμένου (συνήθως,  $K=2$  ή  $3$ ).
- Κάθε αντικείμενο της cache χαρακτηρίζεται είτε ως *O* (δηλαδή, own) είτε ως *H* (δηλαδή, hosted). Εάν ένα *H*-αντικείμενο ζητηθεί από έναν cache κόμβο, τότε η κατάσταση του αντικειμένου μεταβάλλεται σε *O*.
- Κάθε *O* cached αντικείμενο αναγνωρίζεται επίσης και ως κοινοτικό (προέρχεται από τη 2-hop γειτονιά) και δικτυακό (προέρχεται εκτός της 2-hop γειτονιάς).

Το πρωτόκολλο extPCICC χρησιμοποιεί ένα αλγόριθμο εύρεσης μιας cache (περιγράφεται στη συνέχεια και εξηγείται μέσω παραδείγματος στο Σχήμα 5.6) για να βρει τον κόμβο που έχει το ζητούμενο αντικείμενο. Όταν ένας κόμβος εκδίδει ένα αίτημα για κάποιο αντικείμενο, τότε ερευνά την τοπική του cache. Εάν

το αντικείμενο βρεθεί εκεί (local cache hit), τότε ενημερώνονται οι  $K$  πιο πρόσφατες προσπελάσεις. Διαφορετικά (local cache miss), η αίτηση μεταδίδεται και λαμβάνεται από τους CCNs. Εάν δεν απαντήσει κάποιος από αυτούς (community cache miss), τότε το αίτημα δρομολογείται προς την πηγή των δεδομένων.



Σχήμα 5.5. Το πρωτόκολλο extPCICC.

Όταν ένας μη 1-hop CCN λάβει το αίτημα, ερευνά την τοπική cache. Εάν αντιληφθεί ότι η αίτηση μπορεί να ικανοποιηθεί από ένα γειτονικό κόμβο (remote cache hit), τότε σταματά την προώθηση του αιτήματος προς τον κόμβο πηγή και το προωθεί προς το γειτονικό κόμβο. Εάν περισσότεροι του ενός κόμβοι μπορούν να εξυπηρετήσουν την αίτηση, τότε επιλέγεται ο κόμβος με τη μεγαλύτερη εναπομείνασα ενέργεια. Στην περίπτωση όμως που το αίτημα δεν μπορεί να ικανοποιηθεί από το CCN, τότε προωθείται προς τον κόμβο πηγή. Το Σχήμα 5.5 δείχνει τη συμπεριφορά του πρωτοκόλλου extPCICC για ένα αίτημα κόμβου. Συνοπτικά, για κάθε αίτημα που εκδίδεται από έναν κόμβο μπορεί να λαμβάνει χώρα μια από τις εξής τέσσερις περιπτώσεις:

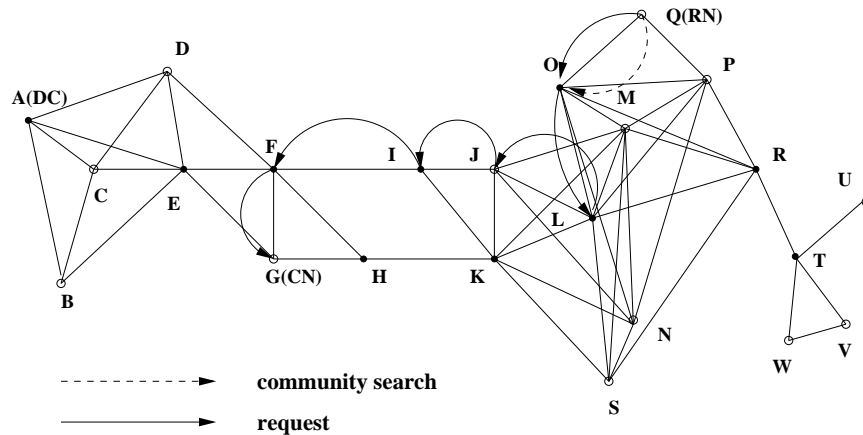
1. Local hit (LH): Το αντικείμενο που ζητείται βρίσκεται στην cache του κόμβου που δημοσιοποιεί το αίτημα. Εάν το αντικείμενο είναι έγκυρο (το χρονικό όριο ζωής δεν έχει παρέλθει), τότε το πρωτόκολλο NiCoCa δεν

εκτελείται.

2. Community hit (CH): Το ζητούμενο αντικείμενο είναι τοποθετημένο στην cache ενός κόμβου που βρίσκεται στη 2-hop γειτονιά του κόμβου που το χρειάζεται. Σε αυτή την περίπτωση οι CCN κόμβοι επιστρέφουν στον αιτούντα κόμβο την τοποθεσία του κόμβου που έχει αποθηκευμένο το αντικείμενο.
3. Remote hit (RH): Το ζητούμενο αντικείμενο βρίσκεται στην cache ενός κόμβου που έχει τουλάχιστον ένα CCN κόμβο που περιλαμβάνεται στο μονοπάτι που συνδέει τον αιτούντα κόμβο με την πηγή δεδομένων.
4. Global hit (GH): Το ζητούμενο αντικείμενο αποκτάται από την πηγή δεδομένων.

Όταν ένας κόμβος λάβει το αντικείμενο που έχει ζητήσει, τότε το τοποθετεί στη cache και εκπέμπει ένα μικρό πακέτο που περιέχει το αναγνωριστικό και το *TTL* του αντικειμένου, την εναπομείνασα ενέργεια και τον ελεύθερο χώρο της cache. Οι κόμβοι CCN που είναι 1-hop γείτονες του κόμβου αποθηκεύουν την πληροφορία. Κάθε CCN αποθηκεύει την εναπομείνασα ενέργεια και τον ελεύθερο χώρο της cache για κάθε ένα 1-hop γείτονά του. Επιπλέον, κάθε CCN αποθηκεύει για κάθε αναγνωριστικό αντικειμένου που συνέλαβε να εκπέμπεται, την τιμή *TTL* και τους κόμβους που έχουν το αντικείμενο.

Το Σχήμα 5.6 απεικονίζει ένα παράδειγμα εύρεσης cache. Στο παράδειγμα αυτό θεωρούμε ότι ο κόμβος *G* περιέχει στη τοπική cache το αντικείμενο  $d_i$  που



Σχήμα 5.6. Μια αίτηση από τον κόμβο *Q* προωθείται στον caching κόμβο *G*.

έχει ζητήσει ο κόμβος  $Q$ . Επιπλέον, οι κόμβοι  $E, F$  και  $H$  είναι οι CCNs του κόμβου  $G$ , ενώ ο κόμβος  $O$  είναι ο CCN του  $Q$ . Αρχικά, ο κόμβος  $Q$  στέλνει μια αίτηση στο δικό του CCN. Ο CCN πραγματοποιεί μια αναζήτηση στη πληροφορία που έχει αποθηκεύσει και στην περίπτωση ενός community cache miss αποστέλλεται ένα μήνυμα αποτυχίας στο κόμβο  $Q$  που πλέον στέλνει την αίτηση στον κόμβο πηγή (κόμβος  $A$ ). Όταν ένας ενδιαμέσος κόμβος λαμβάνει την αίτηση, τότε πραγματοποιεί αναζήτηση στην τοπική cache και στον community cache πίνακα. Εάν το αντικείμενο δεν βρεθεί, τότε η αίτηση προωθείται διαμέσου του μονοπατιού προς το κόμβο πηγή. Τελικά, ο κόμβος  $F$  λαμβάνει την αίτηση και το ζητούμενο αντικείμενο ανακαλύπτεται στο community cache πίνακα. Η αίτηση ανακατευθύνεται στον caching κόμβο  $G$  και ένα μήνυμα απάντησης στέλνεται στο κόμβο  $Q$ .

#### 5.4.3.1 Συνιστώσα αντικατάστασης της cache

Το πρωτόκολλο extPCICC χρησιμοποιεί μια αποτελεσματική και αποδοτική πολιτική αντικατάστασης για να διαχειρισθεί κατάλληλα το χώρο της cache. Στη λειτουργία αντικατάστασης ένα ή περισσότερα αντικείμενα απομακρύνονται από την τοπική cache (για να δημιουργηθεί ο απαραίτητος χώρος) και ένα νέο τοποθετείται. Το πρωτόκολλο extPCICC χρησιμοποιεί την εξής πολιτική:

- A. Αρχικά κάθε κόμβος απομακρύνει τα αντικείμενα που έχει αποθηκεύσει για λογαριασμό άλλου κόμβου. Κάθε αντικείμενο χαρακτηρίζεται είτε ως  $O$  (δηλαδή, own) είτε ως  $H$  (δηλαδή, hosted). Στην περίπτωση ενός local hit η κατάσταση ενός αντικειμένου μεταβάλλεται σε  $O$ . Εάν ο διαθέσιμος χώρος στη cache είναι ακόμα μικρότερος από τον απαιτούμενο, τότε εκτέλεσε το βήμα B.
- B. Κάθε αντικείμενο  $O$  αναγνωρίζεται επίσης ως κοινοτικό (έχει αποκτηθεί από την 2-hop γειτονιά) και δικτυακό (έχει αποκτηθεί από κάποιο κόμβο εκτός της 2-hop γειτονιάς). Ένας κόμβος απομακρύνει έναν αριθμό κοινοτικών αντικειμένων εκτελώντας το βήμα C, μέχρι να σχηματισθεί ο απαιτούμενος χώρος στην cache. Εάν ο διαθέσιμος χώρος είναι ακόμα μικρότερος από τον απαιτούμενο, τότε εκτέλεσε το βήμα C για τα δικτυακά αντικείμενα της cache.
- C. Για κάθε αντικείμενο  $i$  της cache υπολογίζεται η συνάρτηση αντικατάστασης:  $cost(i) = \frac{Lat_i \times Size_i}{TTL_i \times AR_i}$ , οπότε τα αντικείμενα με τις μεγαλύτερες τιμές απομακρύνονται από την cache. Όταν ένας κόμβος λαμβάνει μια απάντηση,

τότε όσο μικρότερη είναι η καθυστέρηση πρόσβασης ενός αντικειμένου, τόσο πιθανότερο είναι να παραμείνει το αντικείμενο στην cache. Ο ρυθμός προσπέλασης (access rate,  $AR$ ) εμφανίζει τη συχνότητα αίτησης ενός αντικειμένου, ενώ η τιμή  $TTL$  καθορίζει την εγυρότητα του αντικειμένου. Ένα αντικείμενο παραμένει στην cache όταν τα  $AR$  και  $TTL$  είναι μεγάλα. Τέλος, όσο μεγαλύτερο είναι το μέγεθος ενός αντικειμένου, τόσο πιθανότερο είναι να απομακρυνθεί από την cache.

- D. Ενημερώνουμε τους κόμβους CCN σχετικά με τα αντικείμενα που είναι υποψήφια προς απομάκρυνση. Εάν το αντικείμενο είναι επίσης τοποθετημένο στην cache ενός άλλου κόμβου της κοινότητας, τότε οι CCNs μεταδίδουν ένα μήνυμα διαγραφής και το αντικείμενο απομακρύνεται από την τοπική cache. Διαφορετικά, κάθε CCN στέλνει ένα μήνυμα που περιέχει τον κόμβο με τη μεγαλύτερη εναπομείνασα ενέργεια και αρκετό χώρο στην cache. Σε αυτή την περίπτωση το αντικείμενο ωθείται στην cache του κόμβου με τη μεγαλύτερη εναπομείνασα ενέργεια. Τελικά, οι κόμβοι CCN ενημερώνουν τα μεταδεδομένα τους.
- E. Ο κόμβος που υποδέχεται στην cache το αντικείμενο που απομακρύνθηκε, ενημερώνει τους κόμβους CCN.

#### 5.4.3.2 Έλεγχος εισδοχής στη cache

Όταν ένας κόμβος λάβει το ζητούμενο αντικείμενο, τότε εκτελείται ένας έλεγχος εισδοχής στη cache για να καθορισθεί αν θα πρέπει το αντικείμενο να τοποθετηθεί στην cache. Το caching ενός αντικειμένου δεν είναι πάντα η ενδεδειγμένη λύση, επειδή μπορεί να οδηγήσει σε μείωση των cache hits [20]. Σε αυτό το κεφάλαιο, η πολιτική ελέγχου εισδοχής στην cache επιτρέπει σε ένα κόμβο να τοποθετεί στην cache κάθε ζητούμενο αντικείμενο με την προϋπόθεση ότι υπάρχει ο απαραίτητος χώρος. Όμως, όταν η τοπική cache είναι πλήρης, τότε η πολιτική προτιμά την τοποθέτηση στην cache των αντικειμένων που δεν βρίσκονται στη 2-hop γειτονιά. Αυτό συμβαίνει επειδή τα αντικείμενα που βρίσκονται στη γειτονιά μπορούν να αποκτηθούν σχετικά γρήγορα, σε αντίθεση με αυτά που προέρχονται από μακρινούς κόμβους.

Αρχικά, κάθε κόμβος μπορεί να τοποθετήσει στην cache κάθε ζητούμενο αντικείμενο συμπεριλαμβανομένων όσων προέρχονται από τη 2-hop γειτονιά. Έτσι, τα cached αντικείμενα διαιρούνται σε κοινοτικά (αυτά που προέρχονται από κόμβους εκτός της 2-hop γειτονιάς και χαρακτηρίζονται με το bit 0) και δικτυακά

(όσα προέρχονται από τη 2-hop γειτονιά και χαρακτηρίζονται με το bit 1). Όταν η τοπική cache είναι πλήρης και φθάνει ένα νέο ζητούμενο αντικείμενο, τότε η πολιτική εισδοχής στην cache ελέγχει αρχικά εάν η απάντηση προέρχεται από τη 2-hop γειτονιά. Σε αυτή την περίπτωση ο κόμβος ελέγχει εάν υπάρχουν κοινοτικά αντικείμενα στην cache και απομακρύνει τόσα ώστε να δημιουργηθεί ο απαιτούμενος ελεύθερος χώρος. Τελικά, το ζητούμενο αντικείμενο τοποθετείται στην cache. Εάν όμως δεν υπάρχουν κοινοτικά αντικείμενα, τότε το νέο αντικείμενο δεν αποθηκεύεται. Τα ζητούμενα αντικείμενα που δεν προέρχονται από τη 2-hop γειτονιά πάντα τοποθετούνται στην cache.

## 5.5 Πειραματική Αποτίμηση

Η απόδοση των πρωτοκόλλων extPCICC και PCICC αποτιμήθηκε μέσω πειραματικών προσομοιώσεων. Διεξήγαμε ένα μεγάλο αριθμό πειραμάτων με διάφορες παραμέτρους και συγκρίναμε την απόδοση του πρωτοκόλλου extPCICC και του πρωτοκόλλου PCICC με την καλύτερη πολιτική συνεργατικού caching για ΑΠΔΑ, η οποία ονομάζεται NICoCa και παρουσιάστηκε στο Κεφάλαιο 3.

### 5.5.1 Μοντέλο προσομοίωσης

Όλα τα πρωτόκολλα υλοποιήθηκαν και αποτιμήθηκαν με το δικτυακό προσομοιωτή J-Sim [123]. Στις προσομοιώσεις χρησιμοποιήθηκε το πρωτόκολλο AODV [104] για τη δρομολόγηση της κίνησης των δεδομένων στο ΑΠΔΑ. Συνακόλουθα, το πρωτόκολλο IEEE 802.11 επιλέχθηκε ως πρωτόκολλο πρόσβασης του μέσου MAC και το πρωτόκολλο free space model ως το μοντέλο ασύρματης διασποράς. Το εύρος ζώνης του ασύρματου δικτύου ήταν 2 Mbps. Συνοπτικά, τα χαρακτηριστικά της ραδιοεπικοινωνίας έχουν συνοψισθεί στον Πίνακα 4.2.

Τα πρωτόκολλα δοκιμάστηκαν σε μια μεγάλη ποικιλία τοπολογιών, ώστε να επιτευχθεί η προσομοίωση δικτύων με κόμβους διαφορετικού βαθμού, από 4 μέχρι 10. Έτσι είμαστε σε θέση να εξετάσουμε τόσο τις αραιές δικτυακές τοπολογίες όσο και τις πυκνές. Τα πειράματα διεξήχθησαν επίσης με διαφορετικό πλήθος κόμβων που ποικίλλει μεταξύ 100 και 500. Επιπρόσθετα, αποτιμήσαμε πειραματικά την απόδοση των πρωτοκόλλων χρησιμοποιώντας σύνολα με δεδομένα μεγέθους που ποικίλει ομοιόμορφα μεταξύ 1KB και 10KB.

Η δικτυακή τοπολογία αποτελείται από πολλά τετράγωνα τμήματα πλέγματος όπου τοποθετούνται ένας ή περισσότεροι κόμβοι. Ο αριθμός των τετράγωνων τμημάτων εξαρτάται από το πλήθος των κόμβων και την πυκνότητα της τοποθέ-



τησής τους. Οι τοπολογίες δημιουργούνται ως εξής: η τοποθεσία κάθε ενός από τους  $n$  κόμβους κατανέμεται ομοιόμορφα μεταξύ των σημείων  $(x=0, y=0)$  και  $(x=500, y=500)$ . Η μέση πυκνότητα  $d$  υπολογίζεται μέσω της ταξινόμησης όλων των  $n \times (n - 1)/2$  ακμών του δικτύου με βάση το μήκος τους σε αύξουσα σειρά. Το μέγεθος κάθε τμήματος του πλέγματος, το οποίο αντιστοιχεί στην τιμή του  $d$  ισούται με την τετραγωνική ρίζα του 2 πολλαπλασιαζόμενη με το μήκος της ακμής στη θέση  $n \times d/2$  της ταξινομημένης ακολουθίας. Δύο κόμβοι θεωρούνται γείτονες όταν τοποθετούνται μέσα στο ίδιο τμήμα του πλέγματος ή σε γειτονικά. Η περιοχή προσομοίωσης έχει μέγεθος  $500\text{m} \times 500\text{m}$  και διαιρείται σε ίσου μεγέθους τετραγωνικά τμήματα πλέγματος. Ξεκινώντας από το κάτω αριστερά άκρο, ονομάζουμε τα τμήματα κατά στήλες ως  $1, 2, \dots$ .

Το μοντέλο ερωτημάτων που ακολουθεί ένας κόμβος είναι όμοιο με προηγούμενες μελέτες [149]. Κάθε αισθητήρας δημιουργεί ερωτήματα read-only. Εάν ένας κόμβος δεν λάβει απάντηση σε ένα ερώτημα που έχει αποστείλει, τότε περιμένει για ένα χρονικό διάστημα ( $t_w$ ) πριν ξαναστείλει ένα νέο. Το πρότυπο πρόσβασης των αισθητήρων είναι ανεξάρτητο από την τοποθεσία, δηλαδή οι κόμβοι αποφασίζουν ανεξάρτητα για τα δεδομένα που τους ενδιαφέρουν. Κάθε κόμβος δημιουργεί ερωτήματα για δεδομένα ακολουθώντας μια ομοιόμορφη κατανομή. Επιπλέον, ως πρότυπο πρόσβασης χρησιμοποιήθηκε και η Zipfian κατανομή με  $\theta=0.8$ . Σύμφωνα με τη Zipfian κατανομή, ομάδες κόμβων που βρίσκονται σε γειτονικά τμήματα του πλέγματος (25 τμήματα με μέγεθος  $100\text{m} \times 100\text{m}$ ) ακολουθούν το ίδιο πρότυπο πρόσβασης. Εξετάσαμε την απόδοση των πρωτοκόλλων και για τα δύο πρότυπα πρόσβασης. Στην περίπτωση του προτύπου πρόσβασης με βάση τη Zipfian κατανομή, πραγματοποιήσαμε πειράματα με τιμές  $\theta$  μεταξύ 0.0 και 1.0.

Παρόμοια με τη μελέτη [149], δύο πηγές δεδομένων τοποθετούνται στις αντίθετες γωνίες της περιοχής προσομοίωσης. Η πηγή δεδομένων 1 τοποθετείται στο σημείο  $(x=0, y=0)$  και η πηγή δεδομένων 2 τοποθετείται στο σημείο  $(x=500, y=500)$ . Παράλληλα υπάρχουν  $N/2$  δεδομένα σε κάθε πηγή δεδομένων. Τα δεδομένα με άρτιο αναγνωριστικό τοποθετούνται στη πηγή δεδομένων 1, ενώ τα δεδομένα με περιττό στην πηγή 2. Το μέγεθος κάθε δεδομένου ποικίλει ομοιόμορφα μεταξύ  $s_{min}$  και  $s_{max}$ . Θεωρούμε επίσης ότι τα δεδομένα δεν ενημερώνονται. Τέλος οι πηγές δεδομένων εξυπηρετούν τα ερωτήματα με βάση την πολιτική FCFS (first-come-first-served). Οι παράμετροι της προσομοίωσης παρουσιάζονται στον Πίνακα 5.1.

Παράμετροι	Εξ ορισμού	Εμβέλεια τιμών
Πλήθος αντικειμένων ( $N$ )	1000	
$S_{min}$ (KB)	1	
$S_{max}$ (KB)	10	
Αριθμός κόμβων ( $n$ )	500	100-500
Εύρος ζώνης (Mbps)	2	
Διάστημα αναμονής ( $t_w$ )	10 sec για KB αντικείμενα	
Μέγεθος cache (KB)	800	200-1200
Zipfian κύρτωση ( $\theta$ )	0.8	0.0-1.0

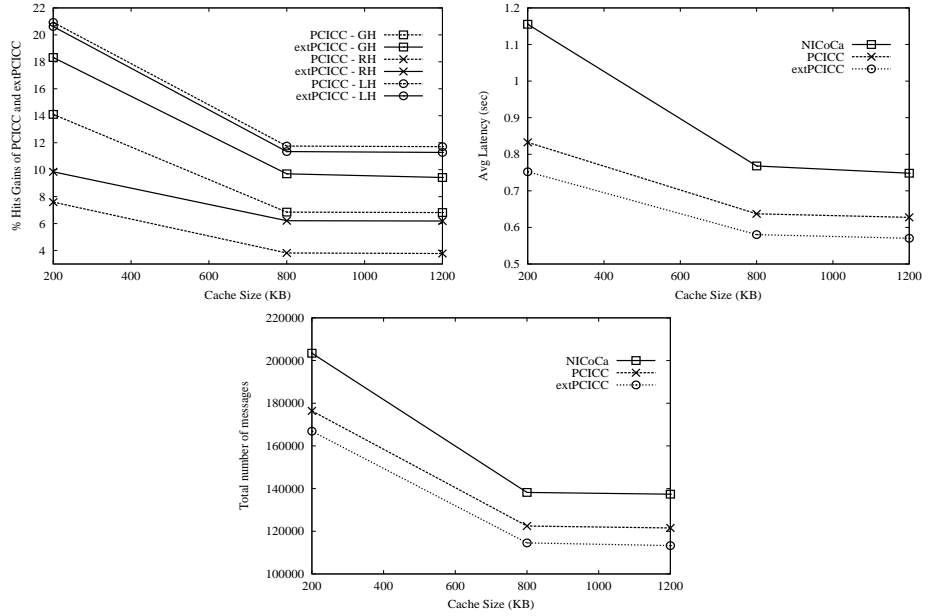
Πίνακας 5.1. Παράμετροι προσομοίωσης.

### 5.5.2 Μετρικές απόδοσης

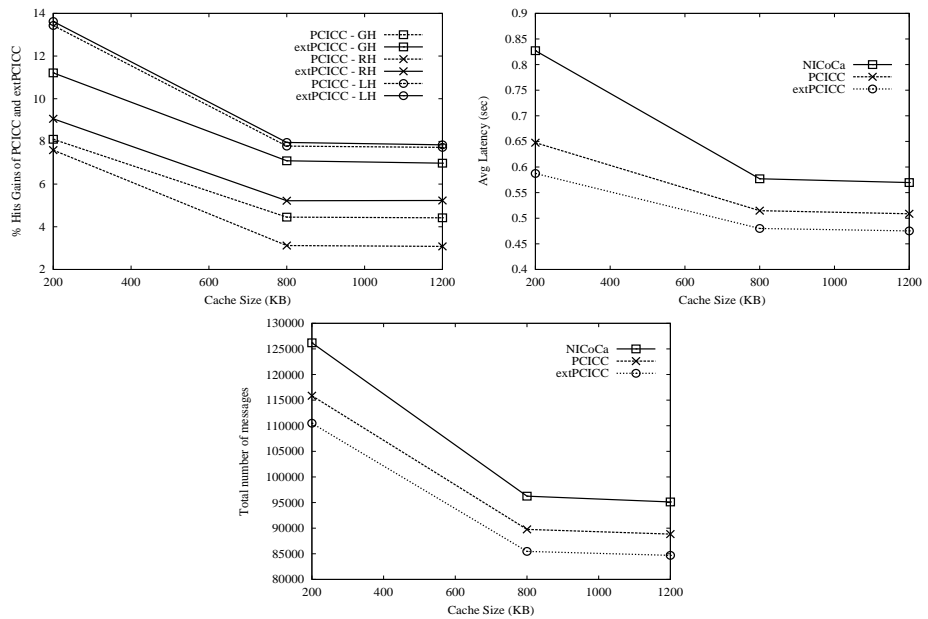
Οι μετρούμενες ποσότητες περιλαμβάνουν την καθυστέρηση πρόσβασης, τον αριθμό μηνυμάτων (ως μια άμεση μετρική της κατανάλωσης ενέργειας) και τον αριθμό των hits (local, remote και global) ως μια μετρική της αποτελεσματικότητας της συνεργασίας. Η καθυστέρηση πρόσβασης είναι η χρονική διάρκεια μεταξύ της αποστολής ενός ερωτήματος και της λήψης της απάντησης. Μια συχνά χρησιμοποιούμενη μετρική είναι ο συνολικός αριθμός των διαδιδόμενων μηνυμάτων λόγω των ερωτημάτων. Η μετρική των μηνυμάτων περιλαμβάνει όλα τα μηνύματα ερωτημάτων και απαντήσεων που διαχέονται στο δίκτυο ώστε να αποκτηθούν τα ζητούμενα δεδομένα. Στις μετρήσεις που πραγματοποιήσαμε αγνοούμε τον αριθμό των μηνυμάτων που οφείλονται στο πρωτόκολλο δρομολόγησης, επειδή είναι ίδιος για όλα τα εξεταζόμενα σχήματα. Είναι φανερό ότι ένας μικρός αριθμός από global hits συνεπάγεται μικρότερη συμφόρηση του δικτύου και συνεπώς λιγότερες συγκρούσεις και απώλειες πακέτων. Επιπλέον, ένας μεγάλος αριθμός από remote hits αποδεικνύει την αποτελεσματικότητα της συνεργασίας στη μείωση του αριθμού των global hits. Ένας μεγάλος αριθμός από local hits δεν συνεπάγεται μια αποτελεσματική πολιτική συνεργατικού caching, εκτός και αν συνδυάζεται με ένα μικρό αριθμό από global hits.

### 5.5.3 Αποτίμηση

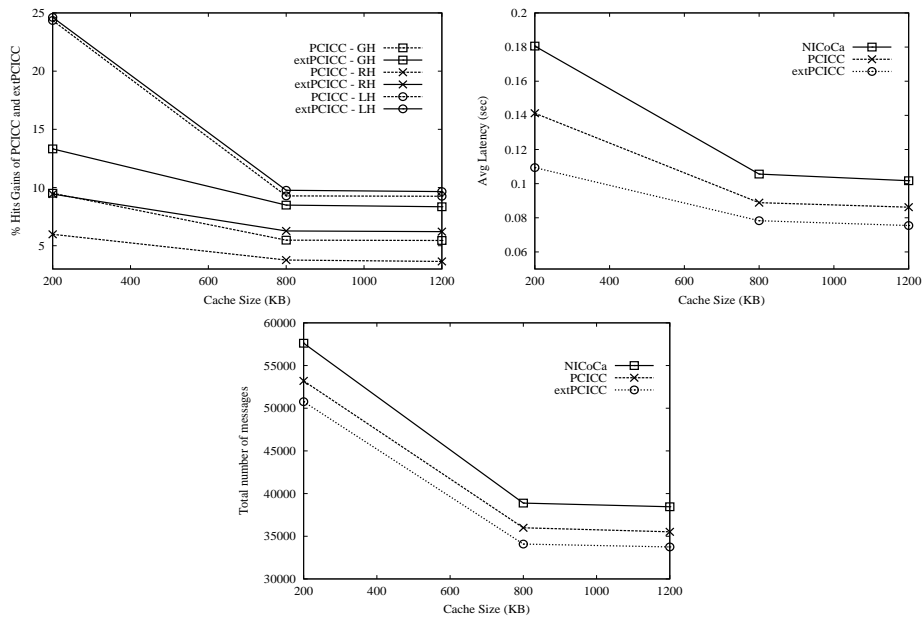
Η πειραματική αποτίμηση περιλαμβάνει ένα μεγάλο αριθμό πειραμάτων μεταβάλλοντας το μέγεθος του δικτύου (το πλήθος των αισθητήρων), μεταβάλλοντας τον τρόπο προσπέλασης από τους κόμβους και το μέγεθος της cache. Πιο συγκε-



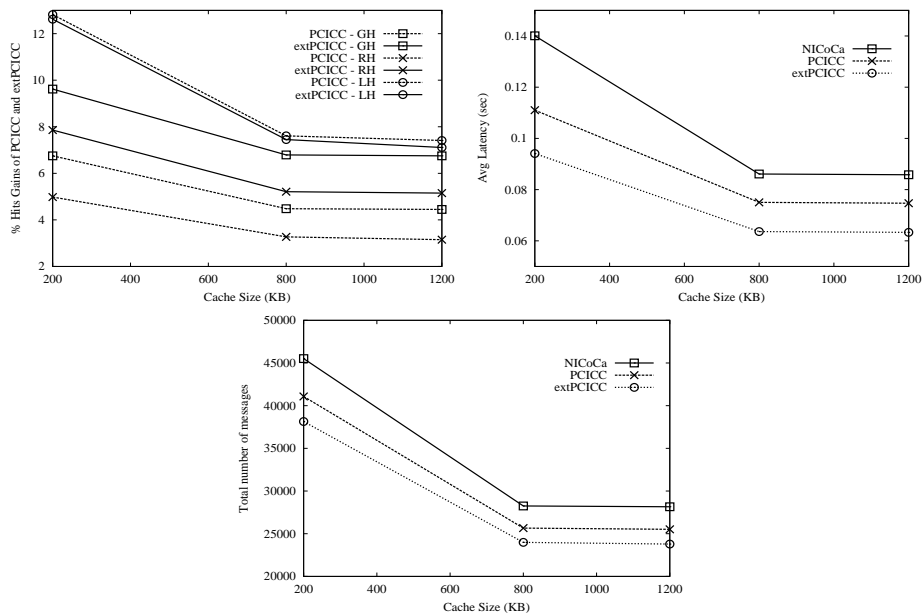
Σχήμα 5.7: Επίδραση του μεγέθους της cache ( $\theta=0.0$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΔΑ ( $d=4$ ) με 100 αισθητήρες.



Σχήμα 5.8: Επίδραση του μεγέθους της cache ( $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΔΑ ( $d=4$ ) με 100 αισθητήρες.



Σχήμα 5.9: Επίδραση του μεγέθους της cache ( $\theta=0.0$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΔΑ ( $d=10$ ) με 100 αισθητήρες.



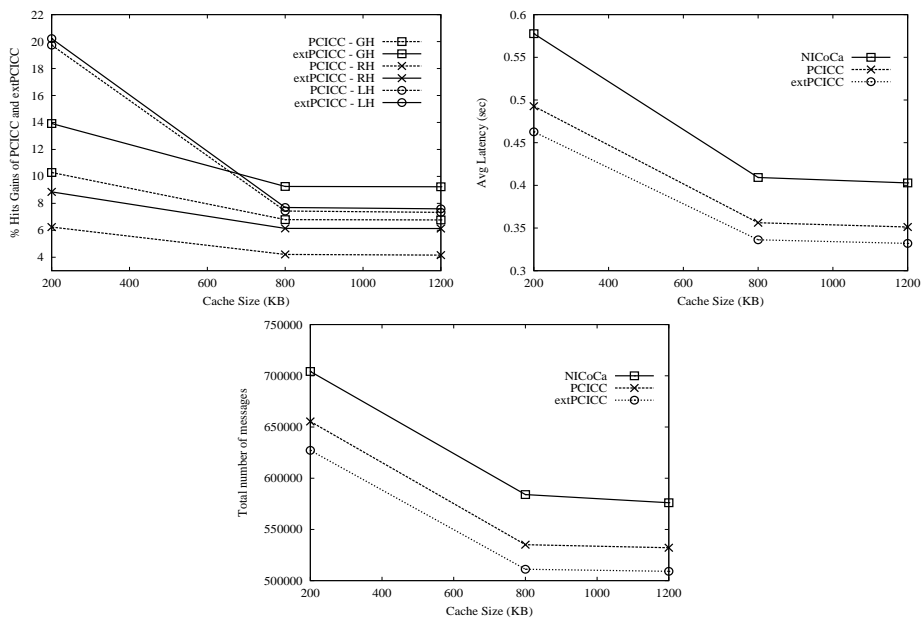
Σχήμα 5.10: Επίδραση του μεγέθους της cache ( $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΔΑ ( $d=10$ ) με 100 αισθητήρες.

κριμένα, πραγματοποιήσαμε πειράματα με 100 και 500 αισθητήρες, για μέγεθος cache ίσο με το 1%, 5% και 10% του συνολικού μεγέθους των δεδομένων, για πρότυπα προσπέλασης με το  $\theta$  να λαμβάνει τις τιμές από 0.0 (ομοιόμορφο πρότυπο πρόσβασης) μέχρι 1.0 (κυρτό πρότυπο πρόσβασης) και για μέσο όρο πυκνότητας των αισθητήρων ίσο με 4 (αραιό δίκτυο αισθητήρων) και 10 (πυκνό δίκτυο αισθητήρων). Τα δεδομένα έχουν μέγεθος λίγων KB. Για κάθε διαφορετικό σύνολο τιμών των παραμέτρων μετρήσαμε τον αριθμό των hits (local, remote, global), την καθυστέρηση πρόσβασης και το πλήθος των μηνυμάτων. Η καθυστέρηση πρόσβασης μετρήθηκε σε δευτερόλεπτα, το οποίο δεν αντιστοιχεί στη συνήθη μετρική του χρόνου αλλά σε ένα εσωτερικό ρολόι του προσομοιωτή. Τα Σχήματα 5.7-5.10 παρουσιάζουν τη σύγκριση απόδοσης των δύο πρωτοκόλλων για δίκτυα με 100 κόμβους αισθητήρων, ενώ τα Σχήματα 5.11-5.14 εμφανίζουν τα πειράματα που πραγματοποιήθηκαν για δίκτυα με 500 κόμβους.

Κάθε σχήμα περιλαμβάνει την επίδραση του μεγέθους της cache στον αριθμό των hits και την καθυστέρηση πρόσβασης (επάνω) και στον αριθμό των μηνυμάτων (κάτω). Όλα τα Σχήματα 5.7-5.14 στο επάνω γράφημα εμφανίζουν την απόδοση των πρωτοκόλλων extPCICC και PCICC συγκριτικά με το πρωτόκολλο NICoCa σε σχέση με τον αριθμό των hits. Τα γραφήματα θα πρέπει να ερμηνευθούν ως εξής: η γραμμή που αντιστοιχεί στα Global Hits παριστά την ποσοστιαία μείωση των global hits που επιτυγχάνεται από τα πρωτόκολλα extPCICC και PCICC, η γραμμή για τα Remote Hits παριστά την ποσοστιαία αύξηση των remote hits που επιτυγχάνεται από τα πρωτόκολλα extPCICC και PCICC, ενώ η γραμμή για τα Local Hits παριστά την ποσοστιαία μείωση των local hits που επιτυγχάνεται από τα πρωτόκολλα extPCICC και PCICC.

Όλα τα σχήματα συνεργατικού caching επιδεικνύουν καλύτερη απόδοση για όλες τις μετρικές καθώς αυξάνεται το μέγεθος της cache. Επομένως, το caching είναι μια χρήσιμη τεχνική ανεξάρτητα από τη δικτυακή τοπολογία. Μία δεύτερη γενική παρατήρηση είναι ότι τα προτεινόμενα πρωτόκολλα είναι ανώτερα από το ανταγωνιστικό για όλες τις δικτυακές τοπολογίες. Όσο αυξάνεται το μέγεθος της cache, τόσο περισσότερα δεδομένα μπορούν να φιλοξενηθούν και συνεπώς οι διαφορές στην απόδοση μεταξύ των πρωτοκόλλων αμβλύνονται.

Η απόδοση των πρωτοκόλλων για τη μέση καθυστέρηση πρόσβασης για διάφορα μεγέθη της cache τόσο για αραιά δίκτυα αισθητήρων όσο και για πυκνά παρουσιάζεται σε όλα τα Σχήματα 5.7-5.14 στο επάνω δεξιό γράφημα. Η κύρια παρατήρηση είναι ότι το caching είναι περισσότερο χρήσιμο για αραιές τοπολογίες, επειδή μπορεί να εξισορροπήσει τα (συγκριτικά) μεγαλύτερα μονοπάτια με τα δεδομένα που αυξάνουν την καθυστέρηση πρόσβασης. Η καθυστέρηση πρόσβα-

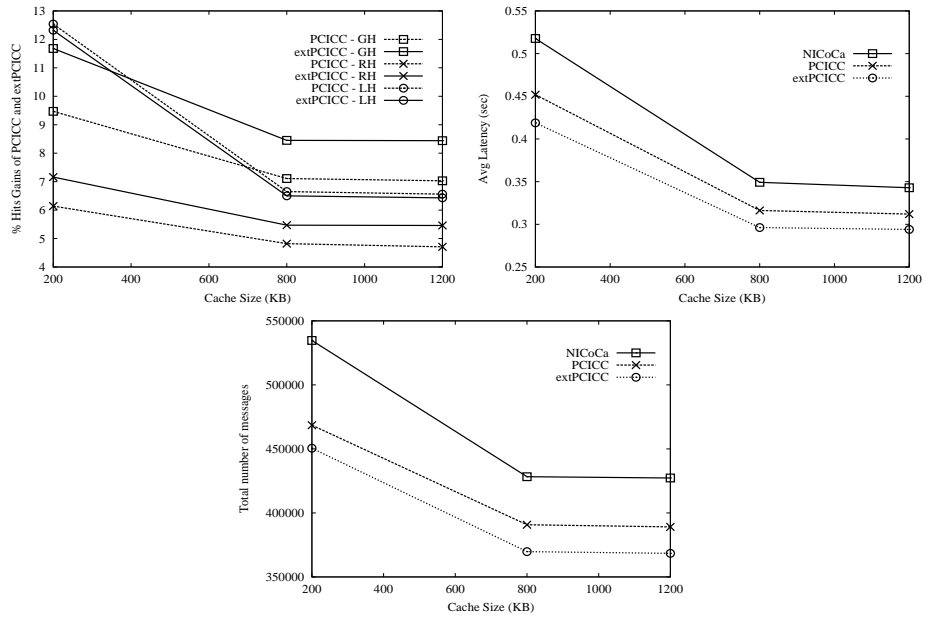


Σχήμα 5.11: Επίδραση του μεγέθους της cache ( $\theta=0.0$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΔΑ ( $d=4$ ) με 500 αισθητήρες.

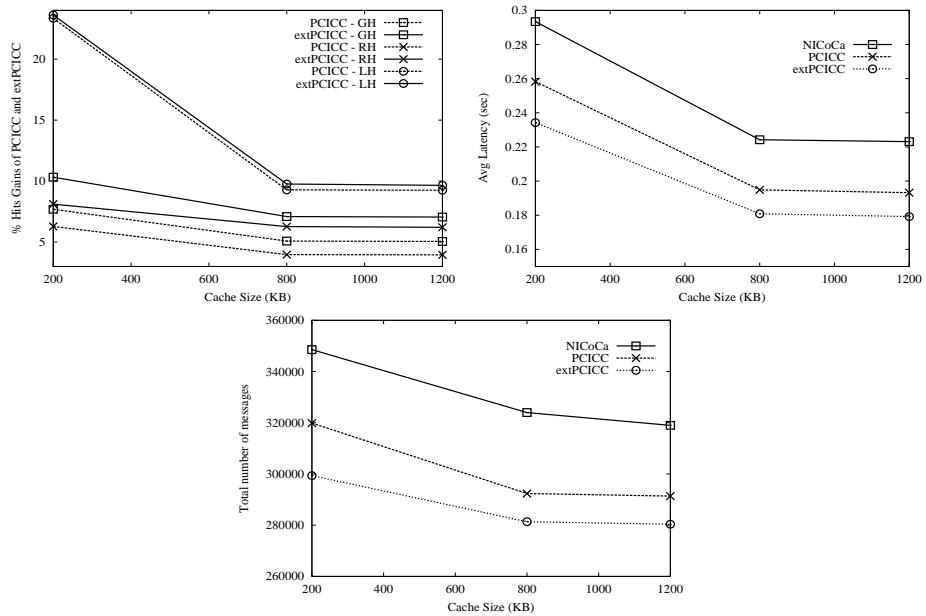
σης που επιτυγχάνεται με το πρωτόκολλο extPCICC είναι 8-16% μικρότερη από αυτή που επιτυγχάνεται με το πρωτόκολλο NICoCa. Εξαιτίας των κεντρικών σημείων όπου βρίσκονται οι κοινοτικοί cache κόμβοι, τα ερωτήματα εξυπηρετούνται ταχύτερα σε σχέση με το πρωτόκολλο NICoCa.

Επιπλέον, αποτιμήσαμε την απόδοση των πρωτοκόλλων με βάση τον αριθμό των μεταδιδόμενων μηνυμάτων τόσο σε αραιές όσο και σε πυκνές τοπολογίες. Τα αποτελέσματα εμφανίζονται σε όλα τα σχήματα στο κάτω γράφημα. Τα συγκριτικά αποτελέσματα ακολουθούν τις ίδιες τάσεις με αυτές που παρατηρήσαμε στα προηγούμενα πειράματα. Υπάρχει στενή σύνδεση μεταξύ του αριθμού των μηνυμάτων και της καθυστέρηση πρόσβασης, επειδή όσο μεγαλύτερος είναι ο αριθμός των μηνυμάτων τόσο μεγαλύτερος είναι και ανταγωνισμός μεταξύ των κόμβων για την απόκτηση του ασύρματου καναλιού και συνεπώς μεγαλύτερος είναι και ο αριθμός των συγκρούσεων που συμβαίνουν. Αυτό έχει ως αποτέλεσμα την αύξηση της καθυστέρηση πρόσβασης. Όταν ένας κόμβος εκπέμπει ένα μεγάλο αριθμό μηνυμάτων, τότε η κατανάλωση ενέργειας αυξάνεται. Έτσι, η μετρική του αριθμού των μηνυμάτων μπορεί να προσφέρει ασφαλή συμπεράσματα για την κατανάλωση ενέργειας.

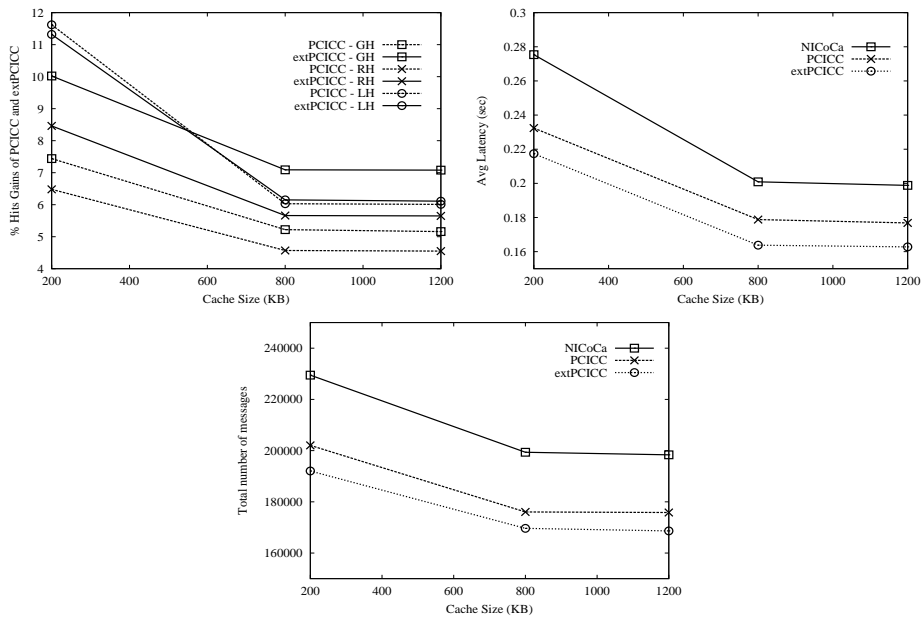
Είναι ενδιαφέρον να σημειώσουμε ότι το πρωτόκολλο extPCICC πετυχαίνει



Σχήμα 5.12: Επίδραση του μεγέθους της cache ( $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα αραιό ΑΔΑ ( $d=4$ ) με 500 αισθητήρες.



Σχήμα 5.13: Επίδραση του μεγέθους της cache ( $\theta=0.0$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΔΑ ( $d=10$ ) με 500 αισθητήρες.



Σχήμα 5.14: Επίδραση του μεγέθους της cache ( $\theta=0.8$ ) στα: (Επάνω) hits και καθυστέρηση πρόσβασης, (Κάτω) αριθμό μηνυμάτων σε ένα πυκνό ΑΔΑ ( $d=10$ ) με 500 αισθητήρες.

κατά μέσο όρο 13% λιγότερα global hits από ότι πετυχαίνει το πρωτόκολλο NiCoCa και περίπου 10% περισσότερα remote hits. Τα αποτελέσματα παρουσιάζονται στο επάνω αριστερά γράφημα όλων των σχημάτων. Η μείωση στα global hits και η ταυτόχρονη αύξηση των remote hits αποδεικνύει ότι το extPCICC πραγματοποιεί μια πιο επιτυχημένη συνεργασία, γεγονός που οφείλεται στην επιλογή των κόμβων CCN. Συμπερασματικά, η προτεινόμενη μετρική κεντρικότητας είναι πράγματι χρήσιμη και αναδεικνύει τους σημαντικούς κόμβους ενός δικτύου. Τα local hits που επιτυγχάνονται και από τα δύο πρωτόκολλα (extPCICC και PCICC) έχουν παρόμοια απόδοση, αλλά η διαφορετική πολιτική εισδοχής στην cache συνεπάγεται την ύπαρξη λιγότερων local hits και περισσότερων remote cache hits.

Γενικά, η απόδοση κάθε πρωτοκόλλου βελτιώνεται σε πυκνότερες τοπολογίες, επειδή περιορίζουμε περισσότερους κόμβους σε μικρότερη γεωγραφική περιοχή. Αυτό έχει ως αποτέλεσμα να δημιουργούμε περισσότερα αντίγραφα των ίδιων δεδομένων και να παρέχονται περισσότερες εναλλακτικές διαδρομές προς τα δεδομένα. Η βελτίωση της απόδοσης αντανακλάται στη καθυστέρηση πρόσβασης, στο hit ratio και στον αριθμό των μεταδιδόμενων μηνυμάτων.

Συνοψίζοντας, για όλες τις τοπολογίες το πρωτόκολλο extPCICC επιτυγχά-



νει περισσότερα remote hits και λιγότερα global hits από τα πρωτόκολλα PCICC και NICoCa. Η διαφορά στην απόδοση είναι ελαφρώς καλύτερη για το πρωτόκολλο extPCICC καθώς μετακινούμαστε από πυκνά σε αραιά ΑΔΑ. Τελικά, όλα τα πρωτόκολλα πετυχαίνουν σημαντικά οφέλη στην απόδοσή τους για κυρτό πρότυπο πρόσβασης ( $\theta=0.8$ ). Αυτό οφείλεται στο ότι περισσότεροι γειτονικοί κόμβοι αιτούνται για τα ίδια δεδομένα. Συνεπώς τα συνεργατικά πρωτόκολλα caching αποδίδουν καλύτερα.

## 5.6 Συμπεράσματα

Στο παρόν κεφάλαιο διαπραγματευθήκαμε το ζήτημα του συνεργατικού caching σε ΑΔΑ. Καταγράψαμε τα μειονεκτήματα των πολιτικών που έχουν προταθεί μέχρι στιγμής, και θέσαμε κάποιους όρους για τα χαρακτηριστικά που θα πρέπει να παρουσιάζει ένα αποδοτικό πρωτόκολλο συνεργατικού caching.

Η εξάπλωση των εφαρμογών που βασίζονται στα ΑΔΑ εξαρτάται κυρίως από την ικανότητα των πρωτοκόλλων να κλιμακώνονται σε μεγάλο πλήθος αισθητήρων, να διατηρούν την ενέργεια της μπαταρίας και να παρέχουν απαντήσεις με μικρή καθυστέρηση πρόσβασης. Το συνεργατικό caching δεδομένων έχει προταθεί ως μια αποδοτική, αποτελεσματική και αξιόπιστη τεχνική για να επιτευχθούν αυτοί οι στόχοι ταυτόχρονα. Τα κυριότερο όμως σημεία ενός πρωτοκόλλου συνεργατικού caching είναι η επιλογή των κόμβων που θα πραγματοποιήσουν το συντονισμό και τη συνεργασία μεταξύ των αισθητήρων, η πολιτική αντικατάστασης και η πολιτική ελέγχου εισδοχής στην cache.

Η επιλογή των σημαντικών κόμβων είναι μια ζωτικής σημασίας λειτουργία, καθώς οι κόμβοι αυτοί θα αναλάβουν ειδικούς ρόλους στο συντονισμό και την εποπτεία του caching και θα πάρουν αποφάσεις σχετικά με την προώθηση των αιτημάτων. Στο παρόν κεφάλαιο εισάγουμε δύο νέες μετρικές που μπορούν να χρησιμεύσουν στην επιλογή των σημαντικών κόμβων. Σκοπός των μετρικών είναι η ανεύρεση των κεντρικών γειτονικών κόμβων. Η πρώτη μετρική (*PCI*) αξιοποιεί την πληροφορία που σχετίζεται με την 1-hop γειτονιά ενός κόμβου, ενώ η δεύτερη μετρική (*extPCI*) αποτελεί βελτίωση της πρώτης και ο υπολογισμός της βασίζεται στη γνώση των συνδέσεων μεταξύ των κόμβων της 2-hop γειτονιάς.

Χρησιμοποιώντας τις δύο μετρικές δημιουργήσαμε δύο πρωτόκολλα συνεργατικού caching (PCICC και extPCICC), τα οποία μέσω πειραματικών προσομοιώσεων αποδείχθηκαν ανώτερα από το ανταγωνιστικό πρωτόκολλο. Τα πειραματικά αποτελέσματα εμφανίζουν μια βελτίωση στην απόδοση των προτεινόμενων πρω-

τοκόλλων έναντι του ανταγωνιστικού ως προς τον αριθμό των hits (local, remote και global), η καθυστέρηση πρόσβασης και τον αριθμό των μεταδιδόμενων μηνυμάτων. Επιπλέον, στο παρόν εισάγουμε μια νέα πολιτική ελέγχου εισδοχής στη cache και αντικατάστασης της cache, έτσι ώστε να βελτιώσουμε την απόδοση των πρωτοκόλλων μειώνοντας την καθυστέρηση πρόσβασης στα δεδομένα.

Συνοψίζοντας, τα πρωτόκολλα PCICC και extPCICC είναι πρακτικά, κλιμακώσιμα και αποδοτικά πρωτόκολλα συνεργατικού caching για τα ΑΔΑ.

## ΚΕΦΑΛΑΙΟ 6

# Ομαδοποίηση σε Ασύρματα Δίκτυα Αισθητήρων\*

### Περιεχόμενα

---

6.1	ΕΙΣΑΓΩΓΗ . . . . .	135
6.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	138
6.3	ΜΟΝΤΕΛΟ ΔΙΚΤΥΟΥ . . . . .	142
6.4	ΤΟ ΝΕΟ ΠΡΩΤΟΚΟΛΛΟ ΟΜΑΔΟΠΟΙΗΣΗΣ . .	142
6.5	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	147
6.6	ΕΦΑΡΜΟΓΕΣ ΟΜΑΔΟΠΟΙΗΣΗΣ . . . . .	155
6.7	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	159

---

### 6.1 Εισαγωγή

Στο Κεφάλαιο 2 είδαμε ότι τα ΑΔΑ έχουν αποκτήσει τεράστια δημοφιλία και ότι τα πρωτόκολλα και οι εφαρμογές που δημιουργούνται πρέπει να λαμβάνουν υπόψη την ενέργεια των αισθητήρων, έτσι ώστε να επιμηχυνθεί ο χρόνος ζωής του δικτύου. Άλλωστε η επαναφόρτιση της μπαταρίας των αισθητήρων είναι μια δύσκολη και ορισμένες φορές αδύνατη εργασία. Παράλληλα, διαπιστώσαμε ότι οι αισθητήρες έχουν περιορισμένες δυνατότητες και στέλνουν τις μετρημένες ποσότητες προς μια πηγή δεδομένων, η οποία ονομάζεται κόμβος sink. Η επικοινωνία

\*Το κεφάλαιο αντλεί τα περιεχόμενά του από την εργασία [42, 46].

μεταξύ δύο κόμβων όπου ο ένας δεν βρίσκεται μέσα στο εύρος μετάδοσης του άλλου επιτυγχάνεται μέσω των ενδιάμεσων κόμβων (multi-hop), οι οποίοι αναμεταδίδουν τα μηνύματα έτσι ώστε να δημιουργηθεί ένα κανάλι επικοινωνίας μεταξύ των κόμβων.

Υπάρχουν πολλά πεδία εφαρμογών που έχουν ωφεληθεί από τα ΑΔΑ, π.χ. παρακολούθηση κατοικίας [126], εποπτεία καταστροφών [133], παρακολούθηση οχημάτων [13] κ.α. Πολλές από αυτές τις εφαρμογές απαιτούν απλώς μια αθροιστική τιμή να δημοσιοποιηθεί στον κόμβο sink (παρατηρητής, σταθμός βάσης, κ.τ.λ.). Σε αυτές τις περιπτώσεις, οι αισθητήρες σε διαφορετικές περιοχές του πεδίου μπορούν να συνεργασθούν για να συναθροίσουν τη συλλεχθείσα πληροφορία. Για παράδειγμα, στις εφαρμογές παρακολούθησης κατοικίας μπορεί να απαιτείται η μέση θερμοκρασία, ενώ στις στρατιωτικές εφαρμογές η διαπίστωση ή όχι υψηλών επιπέδων ραδιενέργειας μπορεί να είναι η κρίσιμη πληροφορία. Είναι φανερό ότι με την οργάνωση των κόμβων σε ομάδες (clusters) μπορούμε να πετύχουμε σημαντικά κέρδη στην απόδοση του δικτύου. Στο παρόν κεφάλαιο εξετάζουμε την ομαδοποίηση (clustering) σε ΑΔΑ, η οποία επιτρέπει τη συνάθροιση της πληροφορίας και περιορίζει τη μετάδοση των δεδομένων κυρίως μέσα στη ομάδα. Συνεπώς, μειώνεται η κίνηση στο δίκτυο και ο συναγωνισμός για την απόκτηση του ασύρματου καναλιού.

Η διαδικασία της ομαδοποίησης περιλαμβάνει την εκλογή της *κεφαλής της ομάδας* (cluster head – CH) για κάθε ομάδα, ώστε αυτή να συντονίζει τους κόμβους. Η κεφαλή της ομάδας είναι υπεύθυνη για τη λήψη των μετρούμενων τιμών από τους κόμβους της ομάδας, για την συνάθροισή τους και την αποστολή της τελικής πληροφορίας μέσω άλλων κεφαλών ομάδας στους sink κόμβους. Διάφορες μελέτες [66, 151] δείχνουν ότι η ομαδοποίηση αυξάνει το χρονικό όριο ζωής του δικτύου. Αν και ο ορισμός του χρονικού ορίου ζωής εξαρτάται από τα εννοιολογικά των εφαρμογών, ένας ευρέως αποδεκτός ορισμός είναι ο χρόνος μέχρι ο πρώτος/τελευταίος κόμβος του δικτύου εξαντλήσει όλη την ενέργειά του [152].

Το ζήτημα της ομαδοποίησης των κόμβων εμφανίστηκε αρχικά στη μελέτη [9], επανεξετάστηκε στις μελέτες [57, 89] και βελτιώθηκε αργότερα στις μελέτες [6, 11, 27, 106, 124, 140] στα πλαίσια των MANET. Όλες οι προσπάθειες αναγνώριζαν τη σημασία της επιλογής των καταλληλότερων κόμβων ως κεφαλές ομάδων και αυτό συνέβη με τη χρήση των *κυρίαρχων συνόλων* (dominating sets – DS), δηλαδή υποσυνόλων κόμβων τέτοιων ώστε κάθε κόμβος του δικτυακού γράφου να ανήκει είτε στο κυρίαρχο σύνολο είτε να είναι γείτονας ενός κόμβου που ανήκει στο κυρίαρχο σύνολο. Μία παρουσίαση σχετικών μεθόδων μπορεί να βρεθεί στη μελέτη [12]. Το κύριο μειονέκτημα των αλγορίθμων είναι το γεγονός

ότι οι κόμβοι του κυρίαρχου συνόλου είναι υπεύθυνοι για τη διευθέτηση όλης της επικοινωνίας και επομένως μπορούν να διασπαθίσουν την ενέργεια τους πολύ σύντομα.

Εκτός της προαναφερθείσας οικογένειας αλγορίθμων, υπάρχει μια δεύτερη οικογένεια που παρέχει μηχανισμούς διευθέτησης του προβλήματος της κατανάλωσης ενέργειας εξαιτίας των επαναλαμβανόμενων μεταδόσεων από τους ίδιους κόμβους, π.χ., τις κεφαλές ομάδων. Αυτή η οικογένεια πρωτοκόλλων προτείνει τρόπους εναλλαγής του ρόλου της κεφαλής ομάδας μεταξύ των κόμβων της ομάδας, π.χ., το πρωτόκολλο SPAN [29], το πρωτόκολλο LEACH [66], και το πρωτόκολλο HEAD [151]. Οι προτεινόμενες μέθοδοι χρησιμοποιούν την εναπομείνουσα ενέργεια κάθε κόμβου για να αποφασίσουν εάν ένας κόμβος θα εκλεγεί ως κόμβος κεφαλή. Παρ' όλα αυτά, αυτή η οικογένεια μεθόδων αγνοεί τα τοπολογικά χαρακτηριστικά.

Όλοι οι μέχρι στιγμής γνωστοί αλγόριθμοι ομαδοποίησης παρουσιάζουν κάποιες αδυναμίες [1, 12, 152]. Κάποιες μέθοδοι βασίζονται στα αναγνωριστικά των κόμβων για να εξαλείψουν τις πλεονάζουσες εκπομπές [6, 11, 27, 124, 140]. Αυτές οι προσεγγίσεις υποφέρουν από το γεγονός ότι δεν μπορούν να ανιχνεύσουν όλες τις πλεονάζουσες εκπομπές επειδή αυτό εμποδίζεται από τη διαρρύθμιση με βάση το αναγνωριστικό του κόμβου (ή το βάρος του κόμβου). Συνεπώς οι μέθοδοι εμφανίζουν σημαντικά μεγάλο αριθμό επανα-μεταδόσεων. Μερικές μέθοδοι δεν αξιοποιούν πλήρως την υπάρχουσα πληροφορία [9, 57, 89, 106, 124]. Για παράδειγμα, η χρήση του βαθμού ενός κόμβου ως κριτηρίου για την επιλογή ως κεφαλή ομάδας μπορεί να μην έχει ως αποτέλεσμα την καλύτερη τοπική απόφαση. Τελικά, κάποιες μέθοδοι δημιουργούν μεγάλο αριθμό ομάδων [140], ή απαιτούν υπερβολικό κόστος επικοινωνίας [6].

Στο παρόν κεφάλαιο προτείνουμε ένα ενεργειακά αποδοτικό καταναμημένο πρωτόκολλο ομαδοποίησης για ΑΔΑ που ονομάζεται *GEodesic Sensor Clustering*, συντομογραφικά GESC. Το πρωτόκολλο αυτό λαμβάνει υπόψη τόσο την κατανάλωση ενέργειας όσο και τη δικτυακή τοπολογία. Το προτεινόμενο πρωτόκολλο αξιοποιεί την τοπική δομή του δικτύου και την εναπομείνουσα ενέργεια των γειτονικών κόμβων για να ορίσει ένα νέο τρόπο δυναμικού υπολογισμού των κεφαλών ομάδων. Το πρωτόκολλο GESC συγκρίνεται με το πρωτόκολλο LEACH [66], το οποίο αποτελεί το αποδοτικότερο πρωτόκολλο στο πεδίο των ΑΔΑ και με τα αποδοτικότερα πρωτόκολλα για MANET τα οποία αναφέρονται στις μελέτες [12, 152]. Το προτεινόμενο πρωτόκολλο GESC συμμορφώνεται με όλες τις απαιτήσεις που περιγράφονται στο [152]:

- Είναι τοπικό και επομένως καταναμημένο. Το πρωτόκολλο μπορεί να αξιοποιήσει την πληροφορία σχετικά με την 1-hop, 2-hop ή  $k$ -hop γειτονιά για να παρουσιάσει βελτιωμένη απόδοση και κόστος επικοινωνίας.
- Η χρονική διάρκεια ζωής του δικτύου επιμηχύνεται με την καταναμημένη κατανάλωση ενέργειας. Οι κεφαλές ομάδων υπολογίζονται δυναμικά με βάση τον κόμβο εκπομπής, ο οποίος επιθυμεί να μεταδώσει ένα μήνυμα. Έτσι οι κόμβοι κεφαλές δεν είναι στατικοί και συνεπώς αποφεύγεται το άδειασμα των μπαταριών τους.
- Περιγράφει ένα νέο τρόπο υπολογισμού της σπουδαιότητας ενός κόμβου. Οι κόμβοι σε μονοπάτια από όπου θα διέλθει η πλειονότητα των μεταδιδόμενων μηνυμάτων θεωρούνται ως σημαντικοί κόμβοι.
- Υπολογίζει τη σημαντικότητα ενός κόμβου σε γραμμικό χρόνο ως προς τον αριθμό των κόμβων και γραμμικό ως προς τον αριθμό των ακμών της δικτυακής γειτονιάς ενός κόμβου, ανεξάρτητα από το βαθμό κάθε κόμβου.
- Επιτρέπει τη γρήγορη ομαδοποίηση των κόμβων και συνεπώς είναι κατάλληλο για επαναληπτικές λειτουργίες ομαδοποίησης.

## 6.2 Σχετικές Εργασίες

Η ομαδοποίηση έχει διερευνηθεί ευρέως στα πλαίσια των MANET [6, 11, 12, 27, 106, 124, 140, 152]. Τα προτεινόμενα πρωτόκολλα είναι καταναμημένα και επιλέγουν τους σημαντικότερους κόμβους ως κεφαλές ομάδων. Η επίτευξη του στόχου πραγματοποιείται με τον υπολογισμό των κυρίαρχων συνόλων. Στη μελέτη [11], οι συγγραφείς θεωρούν τους κόμβους στατικούς με πραγματικές τιμές βαρών, ενώ ο αλγόριθμος Weighted Clustering Algorithm (WCA [27]) συνδυάζει πολλές ιδιότητες σε μια παράμετρο που χρησιμοποιεί για ομαδοποίηση. Ο αλγόριθμος Max-Min D-cluster [6] προτείνει μια νέα καταναμημένη διαδικασία εκλογής των κεφαλών των ομάδων όπου δεν υπάρχει κάποιος κόμβος σε απόσταση μεγαλύτερη από  $d$  hops (η τιμή  $d$  επιλέγεται από τον ευριστικό αλγόριθμο) από ένα CH.

Οι Wu&Li [140] πρότειναν έναν καταναμημένο αλγόριθμο εύρεσης συνδεδεμένων κυρίαρχων συνόλων (connected dominating set – CDS) για τη σχεδίαση αποδοτικών σχημάτων δρομολόγησης για MANET. Κάθε κόμβος ανταλλάσσει τη λίστα των γειτόνων του με όλους τους γείτονές του. Ένας κόμβος θέτει τον εαυτό του ως κυρίαρχο κόμβο εάν έχει τουλάχιστον δύο μη συνδεδεμένους γείτονες. Από τους συγγραφείς έχουν προταθεί κάποιοι κανόνες για να μειωθεί το

μέγεθος του CDS. Σύμφωνα με τον πρώτο κανόνα, ένας κόμβος διαγράφει τον εαυτό του από το CDS, όταν το κλειστό σύνολο γειτόνων του (όλοι οι άμεσοι γείτονες και ο εαυτός του) συμπεριλαμβάνεται εξολοκλήρου στο σύνολο γειτόνων ενός κυρίαρχου γειτονικού κόμβου και έχει μικρότερο αναγνωριστικό (ID) από το γειτονικό κόμβο. Σύμφωνα με το δεύτερο κανόνα, ένας κόμβος διαγράφει τον εαυτό του από το CDS, όταν το ανοιχτό σύνολο γειτόνων του (δηλαδή, οι άμεσοι γείτονές του) συμπεριλαμβάνεται εξολοκλήρου στα σύνολα γειτόνων δύο κυρίαρχων συνδεδεμένων γειτονικών κόμβων και έχει το μικρότερο αναγνωριστικό. Ο Stojmenović [124] πρότεινε έναν αλγόριθμο για τη βελτίωση της απόδοσης του πρωτοκόλλου των Wu&Li, σύμφωνα με τον οποίο οι κόμβοι ταξινομούνται ως εξής. Ένας κόμβος ονομάζεται ενδιάμεσος εάν υπάρχουν δύο γειτονικοί κόμβοι που δεν συνδέονται απευθείας. Δια-πύλη (intergateway) ονομάζεται ένας κόμβος που δεν διαγράφεται από τους κυρίαρχους κόμβους ύστερα από την εφαρμογή του Κανόνα 1 του πρωτοκόλλου των Wu&Li, ενώ πύλη (gateway) ονομάζεται ένας κόμβος που δεν διαγράφεται ύστερα από την εφαρμογή και του Κανόνα 2. Ο συγγραφέας αντικαθιστά τα αναγνωριστικά των κόμβων με μια εγγραφή που περιλαμβάνει το βαθμό του κόμβου και τις συντεταγμένες  $x$  και  $y$ . Οι μόνοι κόμβοι που επιτρέπεται να επαναμεταδώσουν ένα μήνυμα είναι οι πύλες και οι δια-πύλες. Τέλος, πριν ένας κόμβος μεταδώσει ένα μήνυμα υπολογίζει τον αριθμό των 1-hop γειτόνων του που έχουν καλυφθεί από την προηγούμενη μετάδοση. Στην περίπτωση που υπάρχουν κόμβοι που δεν καλύφθηκαν, τότε πραγματοποιείται η εκπομπή του μηνύματος.

Μεγάλος βαθμός τοπικότητας παρουσιάζεται από το πρωτόκολλο της μελέτης [106]. Οι συγγραφείς επικεντρώθηκαν στη μείωση των διπλότυπων μηνυμάτων καθώς τα μηνύματα προωθούνται προς τους κόμβους προορισμούς, έτσι ώστε να επιτευχθεί μια αποδοτική τεχνική πλημμυρίδας σε κινητά ασύρματα δίκτυα. Αρχικά ορίζονται τα σημεία αναμετάδοσης για ένα κόμβο  $u$ . Ένας κόμβος θεωρείται ότι έχει καλυφθεί εάν λαμβάνει ένα μήνυμα που δημιουργήθηκε από τον κόμβο  $u$  είτε απευθείας είτε μέσω αναμεταδόσεων από άλλους κόμβους. Τα σημεία αναμετάδοσης του  $u$  είναι οι 1-hop γείτονες του  $u$  που καλύπτουν όλους τους 2-hop γείτονες του  $u$ . Ο προτεινόμενος αλγόριθμος περιλαμβάνει τρεις φάσεις. Αρχικά, κάθε κόμβος  $u$  ξεκινά με ένα κενό σύνολο multipoint αναμεταδοτών. Στη δεύτερη φάση, ο κόμβος  $u$  επιλέγει ως multipoint αναμεταδότες εκείνους τους 1-hop γείτονες που είναι μοναδικοί γείτονες κάποιων κόμβων στη 2-hop γειτονιά του  $u$  και τους προσθέτει στο σύνολο των multipoint αναμεταδοτών. Στην τρίτη φάση εξετάζεται εάν υπάρχουν κόμβοι στην 2-hop γειτονιά που δεν έχουν καλυφθεί από το σύνολο των multipoint αναμεταδοτών. Σε αυτή την περίπτωση, για κάθε ένα

1-hop γείτονα που δεν περιλαμβάνεται στο σύνολο υπολογίζεται ο αριθμός των 2-hop γειτόνων που καλύπτει και δεν έχουν καλυφθεί ακόμα. Τέλος, προστίθεται στο σύνολο των multipoint αναμεταδοτών ο κόμβος με το μεγαλύτερο αριθμό.

Η ομαδοποίηση είναι μια αποδοτική προσέγγιση ελέγχου της τοπολογίας σε ΑΔΑ, η οποία μπορεί να βελτιώσει την κλιμακωσιμότητα και τη διάρκεια ζωής του δικτύου. Η ομαδοποίηση των κόμβων αισθητήρων είναι ένα πολύ σημαντικό πρόβλημα βελτιστοποίησης καθώς μπορεί αποδοτικά να επιμηκύνει τη διάρκεια ζωής των ΑΔΑ με τη περιορισμένη χρήση ενεργειακών πόρων των κόμβων αισθητήρων.

Το LEACH [66] είναι ένα ενεργειακά αποδοτικό πρωτόκολλο που σχεδιάστηκε για δίκτυα αισθητήρων με μηχανισμό συνεχούς παράδοσης δεδομένων και καθόλου κίνηση. Οι κόμβοι αισθητήρες εκλέγονται ως κεφαλές ομάδων με κάποια πιθανότητα και εκπέμπουν τις αποφάσεις τους. Οι υπόλοιποι κόμβοι προσχωρούν σε μια ομάδα με τον πλησιέστερο κόμβο κεφαλή με όρους ενεργειακού κόστους επικοινωνίας. Στη συνέχεια, ο ρόλος της κεφαλής εναλλάσσεται περιοδικά μεταξύ των κόμβων ώστε να εξισορροπηθεί η κατανάλωση ενέργειας επειδή οι κόμβοι κεφαλές επωμίζονται με μεταδόσεις μεγάλου εύρους. Έτσι, το LEACH εξισορροπεί το πρόβλημα της μη ομοιόμορφης κατανάλωσης ενέργειας με την εναλλαγή του ρόλου. Το πρωτόκολλο HEED [151] εισάγει μια μεταβλητή που είναι γνωστή ως ακτίνα ομάδας (cluster radius) και η οποία ορίζει την ισχύ μετάδοσης που πρέπει να χρησιμοποιηθεί για ενδο-ομαδικές (intracluster) εκπομπές. Η αρχική πιθανότητα για κάθε κόμβο να γίνει δοκιμαστικά κεφαλή της ομάδας εξαρτάται από την εναπομείνουσα ενέργεια, ενώ οι τελικοί κόμβοι κεφαλές επιλέγονται σύμφωνα με το ενδο-ομαδικό κόστος επικοινωνίας. Το πρωτόκολλο HEED βασίζεται στην θεώρηση ότι οι κόμβοι κεφαλές μπορούν να επικοινωνούν μεταξύ τους και να σχηματίζουν ένα συνδεδεμένο γράφο. Στη μελέτη [8], οι συγγραφείς χρησιμοποιούν ομαδοποίηση παρόμοια με αυτή του LEACH και multi-hop προώθηση τόσο για την ενδο-ομαδική όσο και για τη δια-ομαδική (intercluster) επικοινωνία. Παρέχουν επίσης μεθόδους για να υπολογισθούν εκ των προτέρων οι βέλτιστες τιμές των παραμέτρων του αλγορίθμου.

Οι Chang και Tassiulas [26] πρότειναν μεθόδους για να μεγιστοποιήσουν το χρόνο ζωής του δικτύου με τη δίκαιη κατανομή της κατανάλωσης ενέργειας. Οι κόμβοι ρυθμίζουν τα επίπεδα ισχύος των μεταδόσεών τους και επιλέγουν μονοπάτια ώστε να βελτιστοποιήσουν την απόδοση του συστήματος. Στη μελέτη [50] προτείνεται μια πολυεπίπεδη ιεραρχική δομή, όπου οι κόμβοι κεφαλές επιλέγονται με βάση την εναπομείνουσα ενέργειά τους. Ο Buttyan et al. [16] πρότειναν το πρωτόκολλο Position-based Aggregator Node Election (PANEL), που είναι



ένα ενεργειακά αποδοτικό πρωτόκολλο που ισορροπεί το φόρτο εργασίας και την κατανάλωση ενέργειας σε ΑΔΑ. Αυτό συμβαίνει με την εκλογή κάθε κόμβου ως κεφαλής ομάδας με ίση σχεδόν συχνότητα. Παρ' όλα αυτά, το PANEL χρησιμοποιεί την πληροφορία της γεωγραφική τοποθεσίας των κόμβων για να καθορίσει ποίοι από αυτούς θα πρέπει να είναι κεφαλές. Αυτό όμως αποτελεί ένα περιορισμό για τα ΑΔΑ, επειδή είναι δύσκολο να αποκτηθεί η γεωγραφική τοποθεσία χωρίς τη χρήση συσκευής GPS ή την ύπαρξη ενός κεντρικού συντονισμού.

Στη μελέτη [145] οι συγγραφείς προτείνουν μια νέα προσέγγιση ομαδοποίησης (EECS), η οποία είναι ενεργειακά αποδοτική και εφαρμόζεται σε single-hop ΑΔΑ. Το προτεινόμενο πρωτόκολλο είναι κατάλληλο για εφαρμογές περιοδικής συλλογής δεδομένων. Το πρωτόκολλο EECS επεκτείνει το πρωτόκολλο LEACH με το δυναμικό καθορισμό του μεγέθους των ομάδων, ο οποίος βασίζεται στην απόσταση μιας ομάδας από το σταθμό βάσης. Στη φάση εκλογής της κεφαλής της ομάδας, αντίθετα με το LEACH, η κεφαλή ομάδας εκλέγεται μέσω ενός τοπικού συναγωνισμού χωρίς την ύπαρξη επαναληπτικής διαδικασίας γεγονός που κάνει το πρωτόκολλο EECS να διαφέρει από το πρωτόκολλο HEED. Ο συναγωνισμός περιλαμβάνει την εκπομπή από τους υποψήφιους κόμβους της εναπομείνουσας ενέργειάς τους. Εάν κάποιος κόμβος δεν βρει έναν άλλο κόμβο με περισσότερη ενέργεια, τότε γίνεται κεφαλή της ομάδας. Παρ' όλα αυτά, το πρωτόκολλο EECS δεν λαμβάνει υπόψη τα δομικά τοπολογικά χαρακτηριστικά και συνεπώς οι κεφαλές εκλέγονται με βάση μόνο την εναπομείνουσα ενέργειά τους. Αντίθετα με το πρωτόκολλο GESC που σχεδιάστηκε για multi-hop δίκτυα, η στρατηγική που προτείνεται στη μελέτη [91] δεν είναι κλιμακώσιμη καθώς απαιτεί όλοι οι κόμβοι να βρίσκονται μέσα στο εύρος μετάδοσης του σταθμού βάσης. Οι συγγραφείς πρότειναν μια στρατηγική για τον περιορισμό της σπατάλης ενέργειας σε εφαρμογές συλλογής συνεχών δεδομένων σε ΑΔΑ. Αυτό πραγματοποιείται με την αξιοποίηση της χωροχρονικής συσχέτισης. Επομένως, ο sink κόμβος διαιρεί τους κόμβους με παρόμοιες μετρημένες τιμές σε ομάδες και οι κόμβοι που ανήκουν σε μια ομάδα ρυθμίζονται να εργάζονται εναλλακτικά για να μειώσουν την κατανάλωση ενέργειας. Οι Youssef et al. [154] πρότειναν τον κατανεμημένο αλγόριθμο Multi-hop Overlapping Clustering Algorithm (MOCA), ο οποίος χρησιμοποιείται για την οργάνωση των αισθητήρων σε επικαλυπτόμενες ομάδες. Η κύριος στόχος όμως της διαδικασίας ομαδοποίησης είναι να διασφαλίσει ότι κάθε κόμβος είτε είναι κεφαλή είτε βρίσκεται σε  $k$  hops απόσταση από μια κεφαλή ομάδας.

Λιγότερο σχετικές εργασίες με το πρωτόκολλο που προτείνουμε είναι τα σχήματα ελέγχου τοπολογίας, όπως είναι το GAF [141] και το SPAN [29], όπου οι

κόμβοι ταξινομούνται σε ισοδύναμες ομάδες σύμφωνα με τη γεωγραφική τοποθεσία τους. Τέλος, συμπληρωματικές ως προς την εργασία GESC είναι οι τεχνικές και τα πρωτόκολλα που προτάθηκαν στη μελέτη [52]. Οι συγγραφείς ερευνούν τη σχεδίαση τεχνικών και πρωτοκόλλων που οδηγούν σε αποδοτική συνάθροιση των δεδομένων χωρίς τη σαφή διατήρηση μιας δομής.

### 6.3 Μοντέλο Δικτύου

Στο παρόν κεφάλαιο υποθέτουμε ότι ένα ΑΔΑ αποτελείται από  $N$  αισθητήρες. Δεν γίνεται κάποια υπόθεση σχετικά με τη διάμετρο του δικτύου και την πυκνότητά του. Θεωρούμε επίσης τις επόμενες γενικές ιδιότητες ενός ΑΔΑ:

- Οι κόμβοι αισθητήρες είναι στατικοί και επικοινωνούν με multi-hop τρόπο.
- Αρχικά όλοι οι κόμβοι έχουν εφοδιασθεί με την ίδια ποσότητα ενέργειας.
- Οι σύνδεσμοι επικοινωνίας είναι διπλής κατεύθυνσης.
- Οι κόμβοι έχουν τις ίδιες υπολογιστικές και επικοινωνιακές δυνατότητες. Επιπλέον, δεν είναι εφικτή η επαναφόρτιση των μπαταριών των αισθητήρων. Για παράδειγμα, σε ένα πεδίο μάχης οι κόμβοι αισθητήρες διασκορπίζονται σε μια μεγάλη περιοχή, όπου η προσπάθεια προσέγγισης και φόρτισής τους είναι εξαιρετικά δύσκολη και επικίνδυνη. Αυτό απαιτεί να σχεδιάσουμε ένα πρωτόκολλο που θα λαμβάνει υπόψη την ενέργεια έτσι ώστε να επιμηκυνθεί ο χρόνος ζωής του δικτύου.
- Οι κόμβοι δεν χρειάζεται να είναι εφοδιασμένοι με υλικό τύπου GPS. Συνεπώς, δεν γνωρίζουν τη γεωγραφική τοποθεσία τους.
- Οι κόμβοι δεν γνωρίζουν την τοποθεσία των κόμβων sinks. Επιπρόσθετα, δεν ξέρουν πόσες sinks υπάρχουν.

### 6.4 Το Νέο Πρωτόκολλο Ομαδοποίησης

Το προτεινόμενο πρωτόκολλο είναι κατανεμημένο και οι κόμβοι λαμβάνουν αυτόνομα αποφάσεις χωρίς κάποιον κεντρικό έλεγχο. Επιπλέον, το πρωτόκολλο είναι ενεργειακά αποδοτικό καθώς αποφεύγει τη γρήγορη διασπάθιση της ενέργειας των κόμβων και το υπερβολικό κόστος επικοινωνίας. Το πρωτόκολλο ονομάζεται *GEodesic Sensor Clustering* (GESC) που αξιοποιεί τα τοπικά χαρακτηριστικά του δικτύου και την εναπομείνασα ενέργεια των γειτονικών κόμβων για

να επιτύχει την επιμήκυνση της διάρκειας ζωής του δικτύου. Ένα από τα κύρια χαρακτηριστικά του πρωτοκόλλου είναι ο υπολογισμός της σημαντικότητας των κόμβων με βάση την τοπολογία. Η βασική ιδέα είναι η ανακάλυψη εκείνων των ενεργειακά κατάλληλων κόμβων που βρίσκονται σε ένα σημαντικό κομμάτι των συντομότερων μονοπατιών που συνδέουν άλλους κόμβους. Αυτοί οι κόμβοι είναι η συντονιστές κάθε ομάδας για το πρωτόκολλο ομαδοποίησης.

Η σημαντικότητα των κόμβων έχει εξηγηθεί σε βάθος στην Παράγραφο 3.3.1 του Κεφαλαίου 3. Χρησιμοποιώντας την ανωτέρω μέθοδο μπορούμε να ορίσουμε τη σημαντικότητα των κόμβων στις τοπικές γειτονιές του δικτύου. Η Παράγραφος 6.4.1 παρουσιάζει το πρωτόκολλο ομαδοποίησης GESG.

#### 6.4.1 Το πρωτόκολλο ομαδοποίησης

Το πρωτόκολλο ομαδοποίησης διαιρείται σε δύο κύριες διαδικασίες: (α) τη διαδικασία ομαδοποίησης (Clustering Formation Procedure – CFP), και (β) τη διαδικασία λειτουργίας του δικτύου (Network Operation Procedure – NOP). Υποθέτουμε ότι  $T_{CFP}$  είναι το χρονικό κόστος για την ομαδοποίηση, ενώ η διάρκεια λειτουργίας του δικτύου  $T_{NOP}$  είναι το χρονικό διάστημα μεταξύ δύο διαδοχικών  $T_{CFP}$  διαστημάτων. Έτσι, η ομαδοποίηση διαιρείται σε γύρους, όπου στην αρχή κάθε γύρου η διαδικασία CFP για κάθε κόμβο εκτελείται για να επιλέξει τις κεφαλές των ομάδων. Η διαδικασία CFP ακολουθείται από τη διαδικασία λειτουργίας του δικτύου, όπου τα δεδομένα μεταφέρονται από τους κόμβους στις κεφαλές και μέσω multi-hop μονοπατιών στον κόμβο sink.

Υποθέτουμε ότι οι κόμβοι ανταλλάσσουν με τους γείτονές τους μηνύματα “Hello” (beacon messages), τα οποία περιέχουν τη λίστα των γειτόνων τους και την εναπομείνασα ενέργειά τους ( $E_{residual}$ ). Θεωρούμε ότι είμαστε σε θέση να καθορίσουμε τα χρονικά τμήματα (time slots) εκπομπής των κόμβων, έτσι ώστε να μην συμβαίνουν παρεμβολές, π.χ., δύο κόμβοι δεν μεταδίδουν στο ίδιο χρονικό τμήμα. Ένα τέτοιο σχήμα είναι ο αλγόριθμος D2-coloring που περιγράφεται στη μελέτη [56]. Έτσι, κάθε κόμβος είναι ικανός να σχηματίσει ένα γράφο που αντιστοιχεί στη 2-hop γειτονιά του. Επιπλέον, όταν ένας οποιοσδήποτε κόμβος λάβει ένα πακέτο μπορεί να διαπιστώσει από ποιον 1-hop γείτονα έχει σταλεί. Η διαδικασία ανταλλαγής beacon μηνυμάτων που περιέχουν τη λίστα των 1-hop γειτόνων πραγματοποιείται μόνο κατά τη διάρκεια του CFP του πρώτου γύρου λειτουργίας του δικτύου. Αυτό συμβαίνει επειδή οι κόμβοι δεν έχουν κίνηση και επομένως είναι αποδοτικότερο ο τοπικός γράφος να σχηματίζεται μια φορά και να μετασχηματίζεται μόνο όταν έχουμε αποτυχίες. Η αποτυχία ενός κόμβου

θεωρούμε ότι συμβαίνει όταν αδειάζει η μπαταρία. Κατά τη διάρκεια των επόμενων CFPs σε επακόλουθους γύρους, οι κόμβοι ανταλλάσσουν μόνο την τιμή  $E_{residual}$ .

#### 6.4.1.1 Διαδικασία ομαδοποίησης

Η ομαδοποίηση πραγματοποιείται κάθε  $T_{CFP} + T_{NOP}$  δευτερόλεπτα για να επιλέξει νέες κεφαλές ομάδων. Το CFP εκτελείται σε χρονικό διάστημα  $T_C$ , το οποίο θα πρέπει να είναι αρκετά μεγάλο ώστε κάθε κόμβος να λαμβάνει τα μηνύματα όλων των γειτονικών κόμβων. Το CFP συνδυάζει τα δομικά στοιχεία του τοπικού γράφου με την τιμή  $E_{residual}$  των γειτονικών κόμβων για να επιτύχει την καλύτερη επιλογή των κεφαλών ομάδων για κάθε κόμβο. Η διαδικασία CFP περιλαμβάνει τις εξής φάσεις.

*ΦΑΣΗ 1.* Υποθέτοντας ότι ο κόμβος  $v$  έχει μόλις συλλέξει τη λίστα των 1-hop και 2-hop γειτόνων του μέσω της ανταλλαγής των beacon μηνυμάτων, εκτελεί τη διαδικασία *CalculateNodeImportanceIndex* στο 2-hop γράφο  $LN_v$ . Εάν ο κόμβος  $v$  δεν έχει συνδέσμους προς όλους τους άλλους κόμβους του τοπικού δικτύου, τότε θα υπάρχει τουλάχιστον ένας κόμβος  $u$ , τέτοιος ώστε  $u \in N_{12}(v)$  αλλά  $u \notin N_1(v)$ . Επομένως, το εύρος μετάδοσης του  $v$  δεν καλύπτει ολόκληρη τη 2-hop γειτονιά.

*ΦΑΣΗ 2.* Στη συνέχεια εκτελείται ένα αλγόριθμος ταξινόμησης, έτσι ώστε να προκύψει μια φθίνουσα ταξινομημένη λίστα των γειτόνων ως προς την τιμή  $NI_v$ . Σε αυτό το σημείο σημειώνεται ότι η εκτέλεση κάθε αποδοτικού αλγορίθμου ταξινόμησης (π.χ., quicksort), δεν βλάπτει την υπολογιστική πολυπλοκότητα του πρωτοκόλλου, επειδή η πολυπλοκότητα ταξινόμησης είναι  $O(n \log n)$ , όπου  $n$  είναι ο αριθμός στοιχείων του συνόλου  $N_1(v)$ .

*ΦΑΣΗ 3.* Όσο η 2-hop γειτονιά δεν έχει καλυφθεί, εξετάζονται ένα προς ένα τα μέλη της λίστας που δημιουργήθηκε στη ΦΑΣΗ 2. Εάν ο τρέχων εξεταζόμενος 1-hop γείτονας  $u$  καλύπτει τουλάχιστον ένα 2-hop γείτονα, τότε ο 1-hop γείτονας υποδεικνύεται ως υποψήφιος κόμβος κεφαλή.

*ΦΑΣΗ 4.* Στη συνέχεια εκτελείται ένας αλγόριθμος ταξινόμησης για τη δημιουργία μιας λίστας υποψήφιων κόμβων κεφαλών της ομάδας, η οποία λίστα είναι φθίνουσα ως προς την τιμή  $E_{residual}$  των κόμβων. Αφού οι υποψήφιοι κόμβοι εξετασθούν ένας προς ένα, επιλέγονται ως κεφαλές οι κόμβοι με τη μεγαλύτερη εναπομείνασα ενέργεια και καλύπτουν την 2-hop γειτονιά. Στην περίπτωση που δύο ή περισσότεροι κόμβοι έχουν την ίδια τιμή  $E_{residual}$ , τότε επιλέγεται το ελάχιστο σύνολο 1-hop γειτόνων που καλύπτουν τη 2-hop γειτονιά. Αυτό μπορεί να επιτευχθεί εάν εξετασθούν οι κόμβοι με ισοδύναμη εναπομείνασα ενέργεια

σύμφωνα με την τιμή  $NI$ , ξεκινώντας από τον κόμβο με μέγιστη τιμή  $NI$  και συνεχίζοντας προς τον κόμβο με ελάχιστη τιμή  $NI$ . Εάν ο τρέχων εξεταζόμενος 1-hop γείτονας  $u$  καλύπτει τουλάχιστον ένα 2-hop γείτονα (που δεν έχει καλυφθεί ακόμη), τότε ο 1-hop γείτονας υποδεικνύεται ως κόμβος κεφαλή.

Κατά τη διάρκεια του πρώτου γύρου εκτελούνται και οι τέσσερις φάσεις της διαδικασίας CFP. Αρχικά όλοι οι κόμβοι έχουν την ίδια ποσότητα ενέργειας. Έτσι, στη Φάση 4 λαμβάνονται υπόψη μόνο οι τιμές  $NI$  των γειτονικών κόμβων. Στις επακόλουθες εκτελέσεις της διαδικασίας CFP (κατά τη διάρκεια των επόμενων γύρων) και μέχρι να πεθάνει ένα γειτονικός κόμβος (όταν τελειώσει η ενέργεια της μπαταρίας του) εκτελείται μόνο η Φάση 4 για να εκλεγούν οι νέες κεφαλές.

#### 6.4.1.2 Διαδικασία λειτουργίας του δικτύου

Αφού δημιουργηθούν οι ομάδες, κάθε κόμβος μπορεί να επικοινωνήσει με τον κόμβο sink. Για να επιτευχθεί αυτό, κάθε κόμβος μεταδίδει τα δεδομένα του στους εκλεγμένους κόμβους κεφαλές. Οι κόμβοι αυτοί λαμβάνουν τα δεδομένα από τους γειτονικούς τους και συναθροίζουν την πληροφορία. Στη συνέχεια, οι κόμβοι κεφαλές στέλνουν τη συλλεχθείσα πληροφορία στους δικούς τους κόμβους κεφαλές. Μέσω multi-hop επικοινωνίας τα δεδομένα από κάθε κόμβο του δικτύου φθάνουν τελικά στον κόμβο sink.

Κάθε φορά που μία κεφαλή πρόκειται να μεταδώσει ένα μήνυμα, ελέγχει πρώτα εάν υπάρχουν 1-hop γείτονες που έχουν ήδη εκπέμψει το μήνυμα. Στην περίπτωση μιας προηγούμενης εκπομπής του ίδιου μηνύματος η κεφαλή υπολογίζει το πλήθος των 2-hop γειτόνων που δεν έχουν καλυφθεί ακόμα. Η τρέχουσα κεφαλή επιλέγει ως επόμενους κόμβους κεφαλές μόνο εκείνους τους σημαντικούς κόμβους που καλύπτουν τη 2-hop γειτονιά που δεν έχει καλυφθεί. Κάθε κόμβος είναι εφοδιασμένος με μια τοπική cache, έτσι ώστε να αποφευχθούν βρόχοι μονοπατιών. Τα νέα μηνύματα που φθάνουν σε ένα κόμβο τοποθετούνται στη τοπική cache. Επομένως, τα διπλότυπα μηνύματα απορρίπτονται από τους κόμβους, επειδή έχει ήδη μεταδοθεί ένα αντίγραφο τους. Το μέγεθος της cache θεωρείται αρκετά μεγάλο και συνεπώς η αντικατάσταση των μηνυμάτων πραγματοποιείται σπάνια με βάση την τεχνική FIFO.

Κάθε κόμβος γνωρίζει τη 2-hop γειτονιά του μέσω των beacon μηνυμάτων που ανταλλάσσει με τους 1-hop γείτονές του. Έτσι, κάθε κόμβος μπορεί να αναγνωρίσει εάν ο sink κόμβος βρίσκεται σε 2-hop απόσταση. Κάθε κεφαλή που είναι 2-hop γείτονας του sink κόμβου, για να αναμεταδώσει το μήνυμα επιλέγει

τον 1-hop γείτονα με τη μέγιστη εναπομείνασα ενέργεια, που είναι επίσης 1-hop γείτονας του sink κόμβου. Αυτή η λειτουργία πετυχαίνει την ισορροπία στο φόρτο εργασίας και επομένως στην κατανάλωση ενέργειας μεταξύ των 1-hop γειτόνων του sink κόμβου. Ένας κόμβος θεωρείται “νεκρός” εάν έχει απωλέσει όλη την αρχική του ενέργεια. Κάθε κόμβος που έχει χάσει το 99.99% της αρχικής του ενέργειας θεωρεί το εαυτό του νεκρό και μεταδίδει ένα ανάλογο μήνυμα στους 1-hop γείτονες. Οι κόμβοι που λαμβάνουν το μήνυμα, διαγράφουν από τη λίστα των γειτόνων το νεκρό κόμβο και όσους 2-hop γείτονες καλύπτονται μοναδικά από αυτόν. Τελικά, οι κόμβοι εκτελούν τη διαδικασία CFP για να εκλέξουν νέες κεφαλές.

Το προτεινόμενο πρωτόκολλο ομαδοποίησης είναι δυναμικό και ενεργειακά αποδοτικό. Οι εκλεγμένες κεφαλές εξαρτώνται από την τοποθεσία του κόμβου πηγή που ξεκινά τη μετάδοση ενός μηνύματος, την εναπομείνασα ενέργεια και την εξέλιξη της λειτουργίας του δικτύου. Έτσι αποφεύγεται το φαινόμενο των hot-spots. Το πρωτόκολλο GESC είναι δυναμικό επειδή κάθε ενδιαμέσος κόμβος που έχει επιλεγεί να αναμεταδώσει ένα μήνυμα από ένα γειτονικό κόμβο, θα μεταδώσει το μήνυμα στις κεφαλές που μπορεί να είναι διαφορετικοί κόμβοι από αυτούς που έχουν επιλεγεί από τους γειτονικούς κόμβους. Αυτό οφείλεται στο γεγονός ότι κάθε κόμβος έχει διαφορετική όψη του τοπικού δικτύου και σύμφωνα με αυτή αποφασίζει ποιοι κόμβοι είναι οι σημαντικότεροι. Το πρωτόκολλο μπορεί να λειτουργήσει με ένα κεντρικό τρόπο εάν υποθέσουμε ότι κάθε κόμβος γνωρίζει ολόκληρο το δίκτυο και υπολογίζει τους σημαντικούς κόμβους σε όλο το δίκτυο. Η προηγούμενη προσέγγιση όμως έχει δύο μειονεκτήματα για να εφαρμοσθεί σε ένα ΑΔΑ. Το κόστος επικοινωνίας θα είναι εξαιρετικά υψηλό για να μπορέσει κάθε κόμβος να γνωρίζει όλους τους υπόλοιπους. Επομένως, η διασπάθιση της ενέργειας των κόμβων θα αυξηθεί και η επανα-ομαδοποίηση δεν θα είναι αποδοτική. Το δεύτερο μειονέκτημα σχετίζεται με τους κόμβους κεφαλές που θα είναι οι ίδιοι για όλους τους κόμβους και συνεπώς θα εξαντλήσουν την ενέργειά τους πολύ σύντομα.

**ΘΕΩΡΗΜΑ 6.1.** *Ο αλγόριθμος GESC είναι αξιόπιστος καθώς τα εκπεμπόμενα πακέτα μπορούν να διαδοθούν σε κάθε κόμβο ενός συνδεδεμένου δικτύου.*

**ΑΠΟΔΕΙΞΗ.** Θα αποδείξουμε το θεώρημα με μια αντίφαση. Στο τυχαίο δίκτυο θεωρήστε ένα κόμβο  $u$  που λαμβάνει ένα μήνυμα με τη χρήση της πλημμυρίδας, ενώ δεν το λαμβάνει με τη χρήση του αλγορίθμου GESC. Ο κόμβος που δημιουργεί το μήνυμα ονομάζεται κόμβος πηγή και ο κόμβος  $u$  κόμβος προορισμός. Επειδή ο κόμβος προορισμός έχει λάβει το μήνυμα με την τεχνική της πλημμυ-

ρίδας, θα υπάρχει ένα μονοπάτι από τον κόμβο πηγή προς τον κόμβο προορισμό. Το γεγονός ότι ο κόμβος προορισμός δεν έχει λάβει το μήνυμα με τη χρήση του GESG συνεπάγεται ότι δεν έχει εκλεγεί κάποιος από τους γειτονικούς κόμβους ως κόμβος κεφαλή. Έτσι, ακόμα και εάν λάβει το μήνυμα ένας γειτονικός κόμβος, αυτό δεν αναμεταδίδεται επειδή δεν είναι κεφαλή. Σύμφωνα όμως με τον αλγόριθμο GESG ένας κόμβος  $v$  εκλέγει ως κεφαλές τους 1-hop γείτονές του που καλύπτουν τη 2-hop γειτονιά (ΦΑΣΗ 3 και 4 της διαδικασίας σχηματισμού ομάδων – CFP). Αυτό συνεπάγεται ότι ο κόμβος προορισμός δεν ανήκει στη 2-hop γειτονιά κάποιου άλλου κόμβου ή δεν έχει λάβει το μήνυμα κάποιος 1-hop γείτονας του κόμβου προορισμού. Στην πρώτη περίπτωση ο κόμβος προορισμός θα είναι είτε 1-hop γείτονας του κόμβου πηγή ή το τοπικό δίκτυο που περιλαμβάνει τον κόμβο προορισμό και τους 1-hop γείτονες του θα είναι αποσυνδεδεμένο από το δίκτυο. Όμως και τα δύο είναι αντιφάσεις. Σύμφωνα με τη δεύτερη περίπτωση, δεν έχει λάβει το μήνυμα κάποιος από τους 1-hop γείτονες του κόμβου προορισμού. Αυτό συνεπάγεται ότι δεν έχει λάβει το μήνυμα κάποιος από τους 2-hop γείτονες, επειδή ορισμένοι από αυτούς θα είχαν εκλεγεί ως κεφαλές και θα είχαν πραγματοποιήσει την αναμετάδοση του μηνύματος. Συνεχίζοντας με παρόμοιο τρόπο μπορούμε να δείξουμε ότι και οι κεφαλές του κόμβου πηγή δεν έχουν επίσης λάβει το μήνυμα. Συνεπώς, δεν έχει μεταδοθεί στο δίκτυο κάποιο μήνυμα, το οποίο όμως είναι μια αντίφαση.  $\dashv$

## 6.5 Πειραματική Αποτίμηση

Στο κεφάλαιο αυτό θα παρουσιάσουμε την πειραματική αξιολόγηση των κυριότερων μεθόδων που έχουν προταθεί μέχρι στιγμής για την υλοποίηση της ομαδοποίησης σε MANET και AΔΑ. Πραγματοποιήσαμε διάφορα πειράματα για να αποτιμήσουμε την απόδοση του νέου πρωτοκόλλου και να το συγκρίνουμε με τα άλλα πρωτόκολλα. Όλα τα πρωτόκολλα υλοποιήθηκαν με το δικτυακό προσομοιωτή J-Sim [123]. Τα πειράματα εκτελέστηκαν σε δύο φάσεις. Στην πρώτη φάση συγκρίνουμε το πρωτόκολλο GESG με τα κυριότερα πρωτόκολλα που έχουν προταθεί στα πλαίσια των MANETs. Πρόθεση μας είναι να παρουσιάσουμε τη βελτίωση στην απόδοση που επιτυγχάνει το πρωτόκολλο GESG έναντι των άλλων πρωτοκόλλων και να αναδείξουμε ότι το νέο πρωτόκολλο υπερβαίνει τις αδυναμίες των υπολοίπων. Στα πλαίσια της πρώτης φάσης των πειραμάτων εξετάσαμε τα περισσότερα αποδοτικά πρωτόκολλα που αναφέρονται στις μελέτες [12, 152] και επομένως συγκρίναμε το GESG με το MPR [106], το WL [140] και το SSZ [124],

το οποίο είχε επιλεγεί ως “Fast Breaking Paper” στην Πληροφορική για τον Οκτώβριο του 2003. Στη δεύτερη φάση (Κεφάλαιο 6.6), συγκρίναμε το πρωτόκολλο GESC με μια multi-hop ενεργειακά αποδοτική έκδοση του πολύ γνωστού από τη βιβλιογραφία πρωτοκόλλου LEACH [66]. Το πρωτόκολλο LEACH έχει προταθεί για ΑΔΑ. Το LEACH προτάθηκε ως πρωτόκολλο ομαδοποίησης για συγκεκριμένη εφαρμογή. Επομένως, συγκρίνουμε το GESC με το LEACH στο επόμενο κεφάλαιο χρησιμοποιώντας μια παρόμοια πειραματική μεθοδολογία με αυτή που περιγράφεται στο [66]. Τα πειραματικά αποτελέσματα επιβεβαιώνουν ότι το πρωτόκολλο GESC, το οποίο λαμβάνει υπόψη τόσο τα τοπολογικά χαρακτηριστικά όσο και την εναπομείνασα ενέργεια, πετυχαίνει καλύτερη απόδοση από το πρωτόκολλο LEACH ως προς τη διάρκεια ζωής του δικτύου και τον αριθμό των μηνυμάτων που φθάνουν στο σταθμό βάση.

Το παρόν κεφάλαιο περιέχει τη σύγκριση του πρωτοκόλλου GESC με τα πρωτόκολλα των MANET που προτάθηκαν στις μελέτες [106, 124, 140]. Οι μετρικές απόδοσης περιλαμβάνουν τα εκπεμπόμενα μηνύματα κατά τη διάρκεια της δικτυακής λειτουργίας των πρωτοκόλλων, τα εκπεμπόμενα μηνύματα κατά τη διάρκεια σχηματισμού των ομάδων και τέλος την ενέργεια που καταναλώθηκε από τους αισθητήρες. Ο αριθμός των εκπεμπόμενων μηνυμάτων δείχνει την απόδοση των πρωτοκόλλων ομαδοποίησης προκειμένου να πετύχουν μια πλήρη κάλυψη του δικτύου. Αυτό συμβαίνει επειδή ένα πρωτόκολλο μπορεί να επιλέξει περισσότερο αποδοτικά τους κόμβους κεφαλές, οι οποίοι είναι υπεύθυνοι για την προώθηση των μηνυμάτων και συνεπώς να μειωθεί ο αριθμός των μεταδιδόμενων μηνυμάτων. Η μετρική της διασπάθισης ενέργειας αναδεικνύει την εναπομείνασα ενέργεια κάθε κόμβου όταν τερματισθεί η λειτουργία του δικτύου. Οι κόμβοι κεφαλές που ανήκουν στο DS είναι υπεύθυνοι για την διεκπεραίωση όλης της επικοινωνίας και συνεπώς μπορούν να ξεμείνουν από ενέργεια πολύ σύντομα. Επιπλέον, σύμφωνα με τη μετρική της ενέργειας μπορούμε να συμπεραίνουμε εάν ένα πρωτόκολλο είναι ενεργειακά αποδοτικό. Τέλος, η μετρική σχηματισμού των ομάδων δείχνει τον αριθμό των μηνυμάτων που έχουν ανταλλάξει οι κόμβοι ώστε να σχηματισθούν οι ομάδες και να επιλεγούν οι κεφαλές τους. Όσο μικρότερος είναι ο αριθμός των ομάδων, τόσο αποδοτικότερο είναι το πρωτόκολλο. Η λειτουργία δικτύου των πρωτοκόλλων ορίζεται ως η διασπορά ενός μηνύματος από έναν κόμβο πηγή σε όλους τους άλλους κόμβους μέχρι να επιτευχθεί η πλήρης κάλυψη του δικτύου. Αυτή η λειτουργία είναι επιθυμητή όταν οι υπάρχουν πολλοί sink κόμβοι σε ολόκληρο το δίκτυο και οι τοποθεσίες τους δεν είναι γνωστές.

Η τοπολογία του δικτύου αποτελείται από πολλά τετράγωνα τμήματα πλέγματος όπου τοποθετούνται ένας ή περισσότεροι κόμβοι. Ο αριθμός των τετράγωνων



τιμημάτων του πλέγματος εξαρτάται από τον αριθμό των κόμβων και την πυκνότητα τοποθέτησης τους. Οι τοπολογίες δημιουργούνται ως εξής: η τοποθεσία κάθε ενός από τους  $n$  κόμβους κατανέμεται ομοιόμορφα ανάμεσα στα σημεία  $(x=0, y=0)$  και  $(x=500, y=500)$ . Η μέση πυκνότητα  $d$  υπολογίζεται μέσω της ταξινόμησης όλων των  $n \times (n-1)/2$  ακμών του δικτύου με βάση το μήκος τους σε αύξουσα σειρά. Το μέγεθος κάθε τμήματος του πλέγματος το οποίο αντιστοιχεί στην τιμή του  $d$  ισούται με την τετραγωνική ρίζα του 2 πολλαπλασιαζόμενη με το μήκος της ακμής στη θέση  $n \times d/2$  της ταξινομημένης ακολουθίας. Δύο κόμβοι θεωρούνται γείτονες όταν τοποθετούνται μέσα στο ίδιο τμήμα του πλέγματος ή σε γειτονικά. Η περιοχή προσομοίωσης έχει μέγεθος  $500\text{m} \times 500\text{m}$  και διαιρείται σε ίσου μεγέθους τετραγωνικά τμήματα πλέγματος. Ξεκινώντας από το κάτω αριστερά άκρο, ονομάζουμε τα τμήματα κατά στήλες ως  $1, 2, \dots$

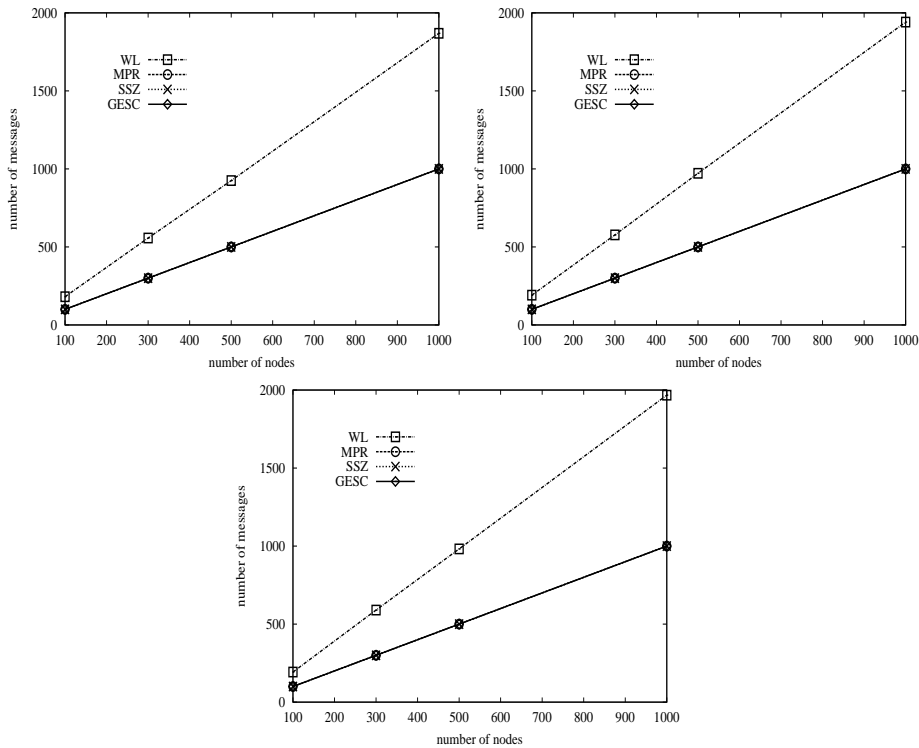
Το δίκτυο δημιουργείται σύμφωνα με όσα περιγράψαμε προηγουμένως με την προϋπόθεση ότι είναι συνδεδεμένο και οι κόμβοι είναι στατικοί. Κάθε κόμβος λαμβάνει ένα μοναδικό αναγνωριστικό και συντεταγμένες  $x, y$  μέσα στη περιοχή προσομοίωσης. Επιπλέον, τα χαρακτηριστικά της ραδιοεπικοινωνίας συνοψίζονται στον Πίνακα 4.2. Τέλος, η αρχική ενέργεια κάθε κόμβου αισθητήρα ισούται με  $E_0=2$  Joules.

### 6.5.1 Αποτίμηση

Η πειραματική αποτίμηση περιλαμβάνει ένα μεγάλο αριθμό πειραμάτων μεταβάλλοντας το μέγεθος του δικτύου (το πλήθος των αισθητήρων) και το βαθμό των κόμβων. Πιο συγκεκριμένα, πραγματοποιήσαμε πειράματα με 100, 300, 500 και 1000 αισθητήρες και για βαθμούς από 4 (πολύ αραιό δίκτυο) μέχρι και 10 (πυκνό δίκτυο). Για να επιτύχουμε μία ακριβέστερη αποτίμηση της δικτυακής λειτουργίας των πρωτοκόλλων εκτελέσαμε κάθε πρωτόκολλο τουλάχιστον 100 φορές για κάθε διαφορετικό βαθμό. Στην συνέχεια υπολογίσαμε το μέσο όρο μηνυμάτων που διαδίδονται στο δίκτυο και το μέσο όρο της ενέργειας που καταναλώθηκε. Σε κάθε εκτέλεση ενός πρωτοκόλλου επιλέγεται και ένας διαφορετικός κόμβος που ξεκινά τη μετάδοση των δεδομένων.

#### 6.5.1.1 Επίδραση του αριθμού των κόμβων

Το πρώτο σύνολο πειραμάτων αποτιμά την επίδραση του αριθμού των κόμβων ενός δικτύου στον αριθμό των μηνυμάτων που παράγονται κατά την ομαδοποίηση (Σχήμα 6.1) και κατά τη διάρκεια λειτουργίας του δικτύου (Σχήμα 6.2). Εξετάζοντας το δεύτερο σχήμα είναι φανερό πως υπάρχει μια γραμμική εξάρτηση μεταξύ

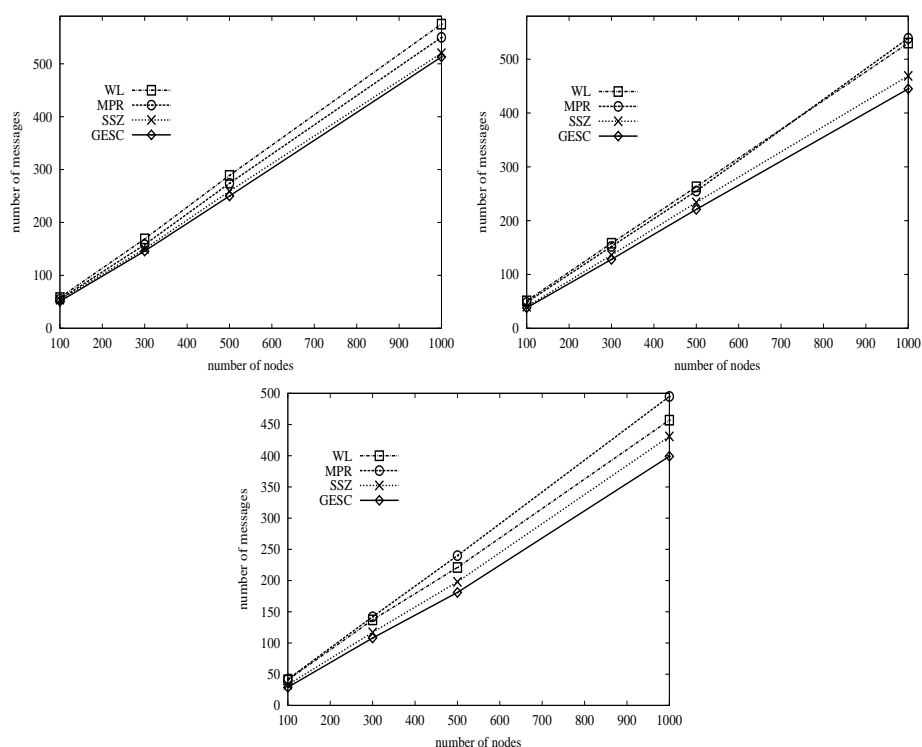


Σχήμα 6.1: Επίδραση του αριθμού των κόμβων στο σχηματισμό των ομάδων για: (Επάνω) βαθμός ίσος με 4 και 7, (Κάτω) βαθμός ίσος με 10.

των μεταδιδόμενων μηνυμάτων και της απόδοσης του πρωτοκόλλου GESC σχετικά με τη διαδικασία εκπομπής. Αυτό συμβαίνει επειδή δημιουργούνται λιγότερες ομάδες και συνεπώς κεφαλές που είναι υπεύθυνες για την προώθηση των μηνυμάτων. Το πρωτόκολλο GESC πετυχαίνει πάντα απόδοση που είναι κατά 4-15% καλύτερη από το δεύτερο αποδοτικότερο αλγόριθμο ομαδοποίησης (SSZ) ασχέτως του μεγέθους του δικτύου (πλήθος κόμβων).

Το επάνω αριστερά γράφημα του Σχήματος 6.2 απεικονίζει τον αριθμό των μεταδιδόμενων μηνυμάτων όταν ο μέσος βαθμός είναι ίσος με 4. Το πρωτόκολλο GESC αποδίδει ελαφρώς καλύτερα από το δεύτερο καλύτερο αλγόριθμο SSZ. Παρ' όλα αυτά, για πυκνές τοπολογίες δικτύου (επάνω δεξιά γράφημα Σχήματος 6.2) αυξάνεται η βελτίωση της απόδοσης του πρωτοκόλλου GESC έναντι του πρωτοκόλλου SSZ. Στο κάτω γράφημα του Σχήματος 6.2 παρατηρούμε ότι για πολύ πυκνές τοπολογίες το πρωτόκολλο GESC αποδίδει 15% καλύτερα από το πρωτόκολλο SSZ. Μπορούμε επίσης να παρατηρήσουμε ότι η διαφορά μεταξύ

των δύο πρωτοκόλλων διευρύνεται όσο αυξάνονται ο αριθμός των κόμβων που συμμετέχουν και ο βαθμός. Αυτό συμβαίνει επειδή όσο αυξάνεται ο βαθμός, τόσο αυξάνεται και ο αριθμός των συντομότερων μονοπατιών που μπορούν να αναγνωρισθούν στη 2-hop γειτονιά. Συνεπώς για πυκνές και πολύ πυκνές τοπολογίες το πρωτόκολλο GESC αναγνωρίζει αποτελεσματικότερα τους σημαντικούς κόμβους επειδή όσο αυξάνονται οι κόμβοι που συμμετέχουν στη τοπική γειτονιά, τόσο αυξάνονται και τα συντομότερα μονοπάτια που συμμετέχουν στον υπολογισμό του  $NI$ .



Σχήμα 6.2: Επίδραση του αριθμού των κόμβων στη μετάδοση για: (Επάνω) βαθμός ίσος με 4 και 7, (Κάτω) βαθμός ίσος με 10.

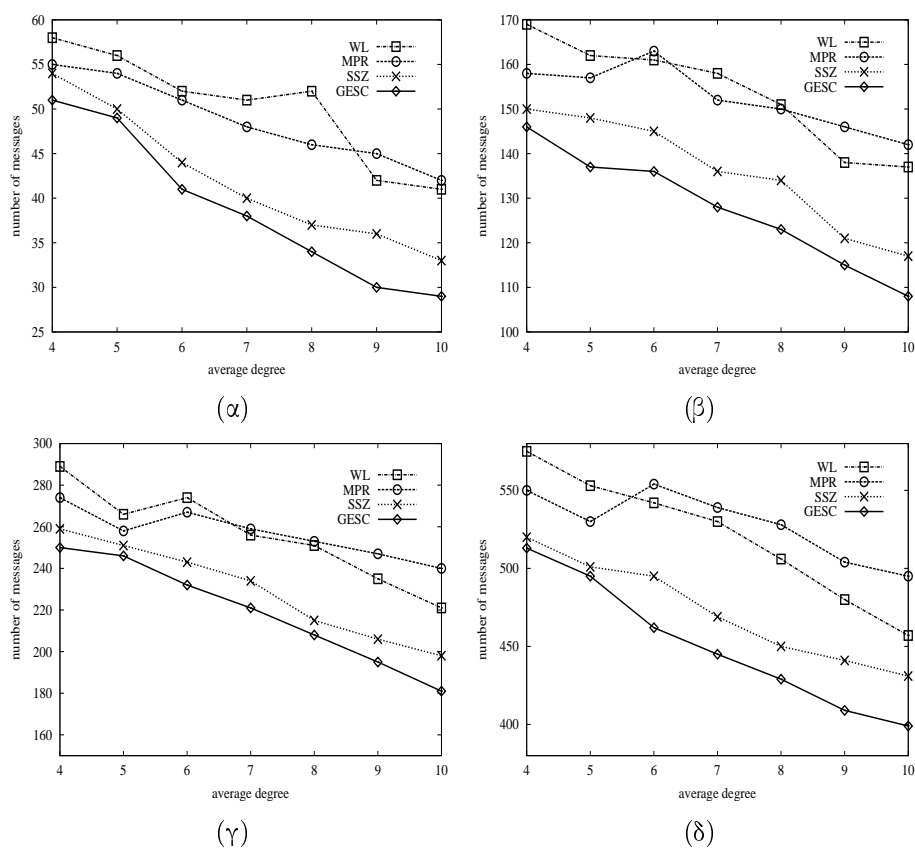
Επιπρόσθετα, παρατηρούμε ότι όλες οι μέθοδοι δημιουργούν τον ίδιο αριθμό μηνυμάτων κατά τη διάρκεια σχηματισμού των ομάδων εκτός από τη μέθοδο WL (Σχήμα 6.1). Αυτό οφείλεται στο γεγονός ότι οι τρεις μέθοδοι πρέπει να γνωρίζουν τους άμεσους 1-hop ή 2-hop γείτονες και προφανώς ο αριθμός αυτός δεν εξαρτάται από τη συνδεσιμότητα (μέσο βαθμό) του δικτύου.

Στα Σχήματα 6.4-6.6 (όρα Παράγραφο 6.5.1.3) παρουσιάζουμε την επίδραση

του αριθμού των κόμβων στη διασπάθιση της ενέργειας. Όσο αυξάνεται ο αριθμός των κόμβων, τόσο αυξάνεται και η καταναλισκόμενη ενέργεια. Σύμφωνα με την εγκατάσταση της προσομοίωσης, για ένα δίκτυο 100 κόμβων θα υπάρχουν 100 μηνύματα που διαχέονται στο δίκτυο, ενώ για ένα δίκτυο 1000 κόμβων τα μεταδιδόμενα μηνύματα αυξάνονται στα 1000. Το δεύτερο καλύτερο πρωτόκολλο (SSZ) από πλευράς απόδοσης διατηρεί τους ίδιους εκλεγμένους κόμβους κατά τη διάρκεια ολόκληρης της προσομοίωσης. Αυτό συνεπάγεται ότι οι ίδιοι κόμβοι αναλαμβάνουν όλο το φορτίο της επικοινωνίας πραγματοποιώντας συνεπώς όλες τις αναμεταδόσεις. Αντίθετα με το SSZ, το πρωτόκολλο GESC εξαιτίας της δυναμικής και ανεξάρτητης εκλογής των κεφαλών αποδίδει καλύτερα ως προς τη διασπάθιση ενέργειας. Κάθε κόμβος αναμεταδίδει ένα μήνυμα στις κεφαλές που εκλέγονται από κάθε κόμβο σύμφωνα με την τοπική όψη του δικτύου και την εναπομείνουσα ενέργεια. Το γεγονός όμως ότι γειτονικοί κόμβοι μπορούν να έχουν επιλέξει διαφορετικές κεφαλές έχει ως αποτέλεσμα τη μικρότερη κατανάλωση ενέργειας.

#### 6.5.1.2 Επίδραση του βαθμού των κόμβων

Το δεύτερο σύνολο πειραμάτων αποτιμά την επίδραση του μέσου βαθμού κόμβου στον αριθμό των μηνυμάτων κατά τη διάρκεια της διαδικασίας εκπομπής (Σχήμα 6.3). Τα αποτελέσματα περιλαμβάνουν ένα δίκτυο αισθητήρων με 100, 300, 500, και 1000 κόμβους. Ομοίως με τα συμπεράσματα που προέκυψαν στο προηγούμενο σύνολο πειραμάτων, παρατηρούμε ότι το πρωτόκολλο GESC είναι ο αποδοτικότερος αλγόριθμος με τα οφέλη να ανέρχονται μέχρι 15% ως προς τη δεύτερη καλύτερη μέθοδο. Σε όλα τα γραφήματα παρατηρούμε ότι τα πρωτόκολλα GESC και SSZ έχουν σημαντική διαφορά όταν ο βαθμός είναι μεγαλύτερος από 7. Οι αποδόσεις των πρωτοκόλλων είναι σχεδόν ισοδύναμες για μικρούς βαθμούς, αλλά το πρωτόκολλο GESC βελτιώνει την απόδοσή του σε σχέση με το SSZ για μεγάλους βαθμούς. Όταν ο βαθμός είναι μικρός (αραιές τοπολογίες), τότε είναι περιορισμένη η τοπική όψη του δικτύου που έχει ένας κόμβος. Αυτό έχει ως αποτέλεσμα το πρωτόκολλο GESC να εκλέγει πολλές κεφαλές. Επιπλέον, η διαδικασία εκπομπής περιλαμβάνει ένα μεγαλύτερο αριθμό αναμεταδόσεων συγκριτικά με τον αριθμό αναμεταδόσεων σε πυκνές και πολύ πυκνές τοπολογίες. Αν και ο αριθμός των αναμεταδόσεων του πρωτοκόλλου GESC αυξάνεται για αραιές τοπολογίες, παρ' όλα αυτά η απόδοσή του συνεχίζει να είναι καλύτερη από τα ανταγωνιστικά πρωτόκολλα. Στην περίπτωση όμως των πυκνών και πολύ πυκνών τοπολογιών η απόδοση του πρωτοκόλλου GESC είναι σημαντικά καλύ-

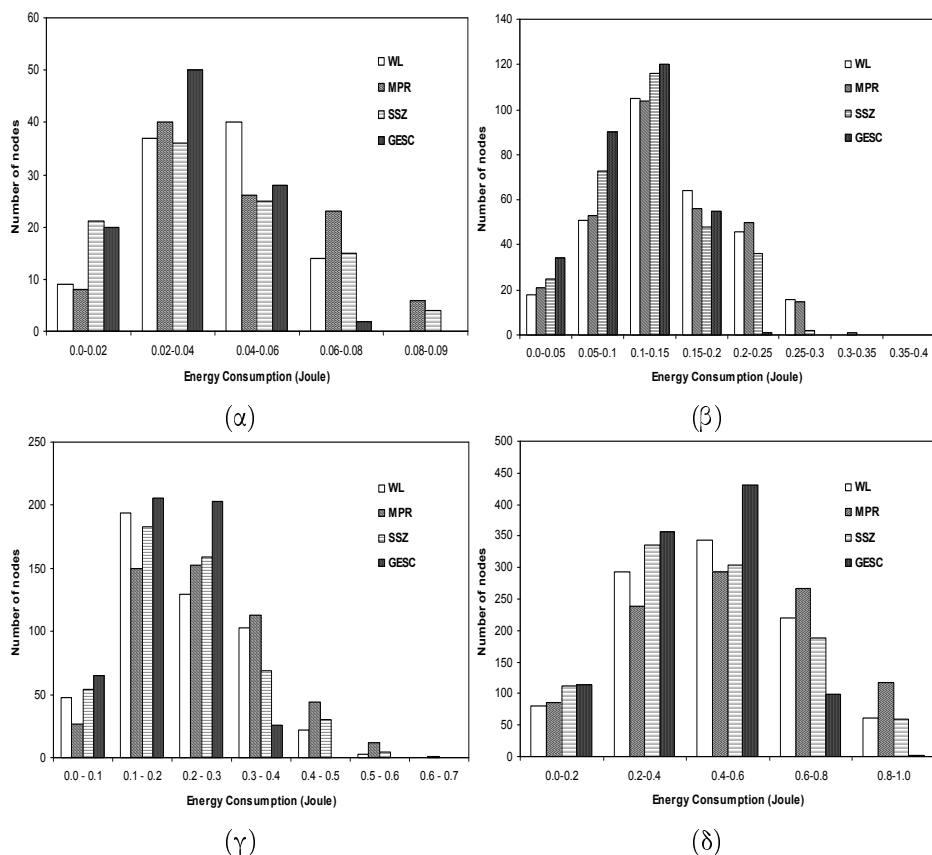


Σχήμα 6.3: Επίδραση του μέσου βαθμού κόμβου στο πλήθος μεταδόσεων για ένα δίκτυο: (α) 100 αισθητήρων, (β) 300 αισθητήρων, (γ) 500 αισθητήρων, (δ) 1000 αισθητήρων.

τερη από το πρωτόκολλο SSZ, επειδή η σημαντικότητα των κόμβων φανερώνεται καλύτερα. Τέλος, τα σχήματα παρουσιάζουν επίσης μια καλύτερη απόδοση του πρωτοκόλλου WL έναντι του πρωτοκόλλου MPR για μεγάλους βαθμούς.

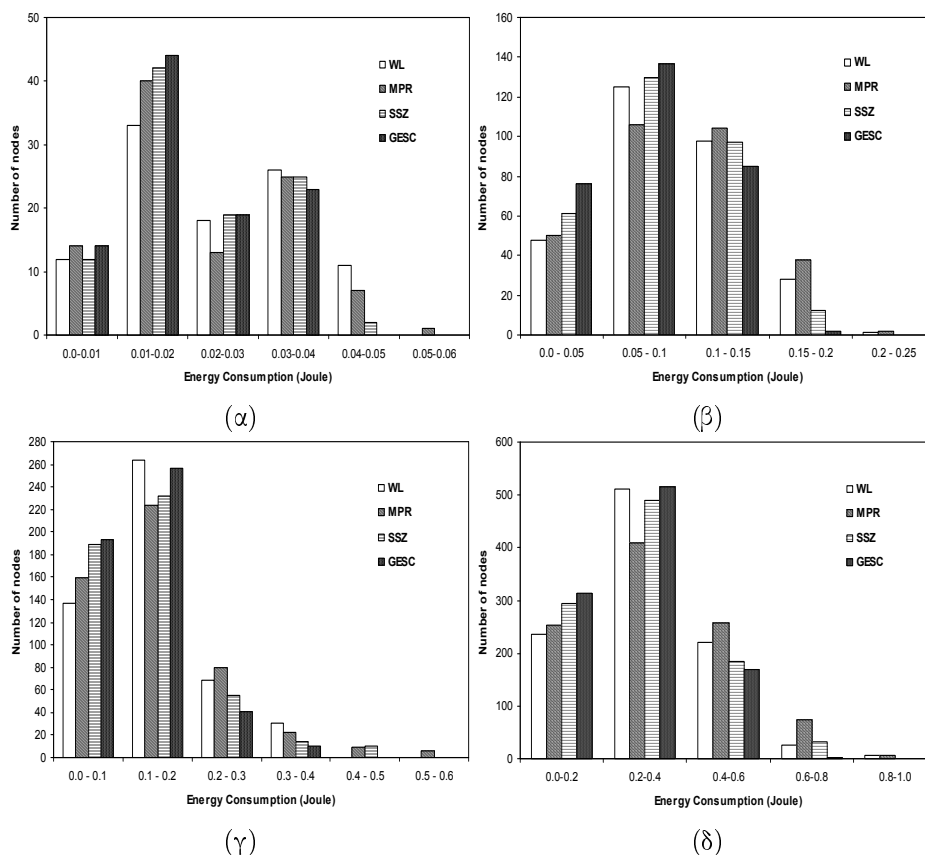
### 6.5.1.3 Επίδραση στην κατανάλωση ενέργειας

Το τελευταίο σύνολο πειραμάτων ερευνά την κατανάλωση ενέργειας. Αφού ολοκληρωθούν όλα τα πειράματα εξετάζουμε την εναπομείνουσα ενέργεια κάθε κόμβου. Στο σημείο αυτό παρουσιάζουμε τα αποτελέσματα για δίκτυα αισθητήρων με 100, 300, 500 και 1000 κόμβους και με βαθμό ίσο με 10, 7 και 4 (Σχήματα 6.4-6.6). Σε κάθε σχήμα, η πρώτη στήλη αντιστοιχεί στο πρωτόκολλο WL, η δεύτερη στήλη στο MPR, η τρίτη στήλη στο SSZ και η τέταρτη στήλη στο νέο πρωτόκολλο GESC. Για βαθμό ίσο με 10 (Σχήμα 6.4) παρατηρούμε τη μεγάλη



Σχήμα 6.4: Εναπομείνασα ενέργεια για βαθμό ίσο με 10 για ένα δίκτυο: (α) 100 αισθητήρων, (β) 300 αισθητήρων, (γ) 500 αισθητήρων, (δ) 1000 αισθητήρων.

κατανάλωση ενέργειας που προκαλείται σε κάποιους κόμβους των πρωτοκόλλων SSZ, WL και MPR, ενώ το πρωτόκολλο GESC επιτυγχάνει καλύτερη ισορροπία στην ενέργεια των κόμβων. Στο πρωτόκολλο GESC υπάρχουν πολλοί κόμβοι που έχουν καταναλώσει ένα μικρό ποσό της ενέργειάς τους, ενώ μόνο λίγοι από αυτούς έχουν καταναλώσει μια σημαντική ποσότητα ενέργειας. Αντίθετα, τα άλλα τρία πρωτόκολλα και ιδιαίτερα τα πρωτόκολλα WL και MPR προκαλούν σε πολλούς κόμβους μεγάλη διασπάθιση της ενέργειας. Αυτό συμβαίνει εξαιτίας της σχεδιαστικής αρχής του πρωτοκόλλου GESC όπου κάθε κόμβος εκλέγει με δυναμικό και ανεξάρτητο τρόπο τους κόμβους κεφαλές αξιοποιώντας την εναπομείνασα ενέργεια και τα δομικά χαρακτηριστικά του τοπικού δικτυακού γράφου. Επιπρόσθετα, οι εκλεγμένοι κόμβοι που αναμεταδίδουν τα μηνύματα δεν είναι πάντα οι ίδιοι. Όταν η τοπολογία δεν είναι πολύ πυκνή (Σχήμα 6.5) και ιδιαι-

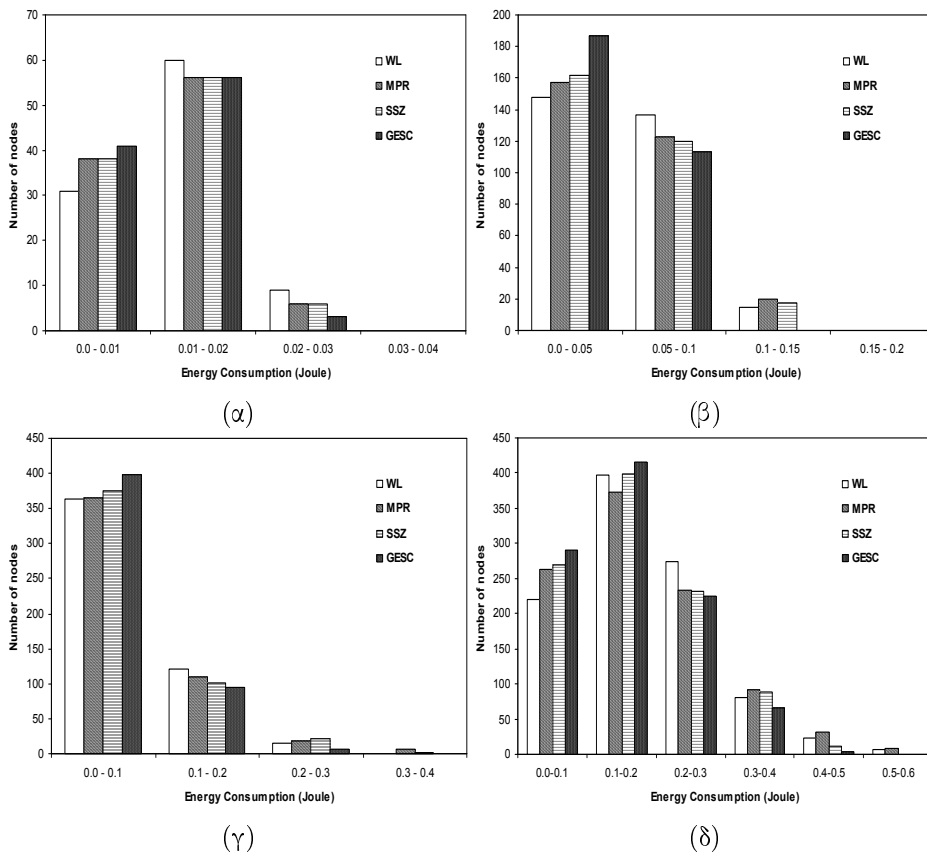


Σχήμα 6.5: Εναπομείνασα ενέργεια για βαθμό ίσο με 7 για ένα δίκτυο: (α) 100 αισθητήρων, (β) 300 αισθητήρων, (γ) 500 αισθητήρων, (δ) 1000 αισθητήρων.

τερα στην περίπτωση αραιού δικτύου (Σχήμα 6.6), το πρωτόκολλο GESZ αποδίδει επίσης καλύτερα από το πρωτόκολλο SSZ. Όμως η διαφορά στην κατανάλωση ενέργειας που προκαλείται στους κόμβους δεν είναι μεγάλη. Τέλος, μπορούμε να σημειώσουμε ότι το GESZ είναι κλιμακώσιμο με τον αριθμό των κόμβων και το βαθμό.

## 6.6 Εφαρμογές Ομαδοποίησης

Το LEACH [66] είναι ένας από τους δημοφιλέστερους αλγορίθμους ομαδοποίησης για ΑΔΑ, ο οποίος στοχεύει στην επιμήκυνση της διάρκειας ζωής του δικτύου. Το LEACH προτάθηκε για μια συγκεκριμένη εφαρμογή όπου οι κόμβοι είναι τυχαία κατανομημένοι σε μια περιοχή και συνεχώς ανιχνεύουν την περιοχή στόχο



Σχήμα 6.6: Εναπομείνασα ενέργεια για βαθμό ίσο με 4 για ένα δίκτυο: (α) 100 αισθητήρων, (β) 300 αισθητήρων, (γ) 500 αισθητήρων, (δ) 1000 αισθητήρων.

για να στείλουν μηνύματα στο σταθμό βάση (BS). Σε αυτό το κεφάλαιο συγκρίνουμε το πρωτόκολλο GESC με το πρωτόκολλο LEACH για να επιβεβαιώσουμε αρχικά ότι το πρωτόκολλο GESC μπορεί να χρησιμοποιηθεί αποδοτικά σε συγκεκριμένες εφαρμογές αισθητήρων, όπως αυτή που περιγράφεται στη μελέτη [66]. Επιπρόσθετα, σκοπός μας είναι να εκθέσουμε την ανωτερότητα του πρωτοκόλλου GESC ως προς τα πρωτόκολλα που χρησιμοποιούν την εναπομείνασα ενέργεια κάθε κόμβου για να εκλέξουν τις κεφαλές.

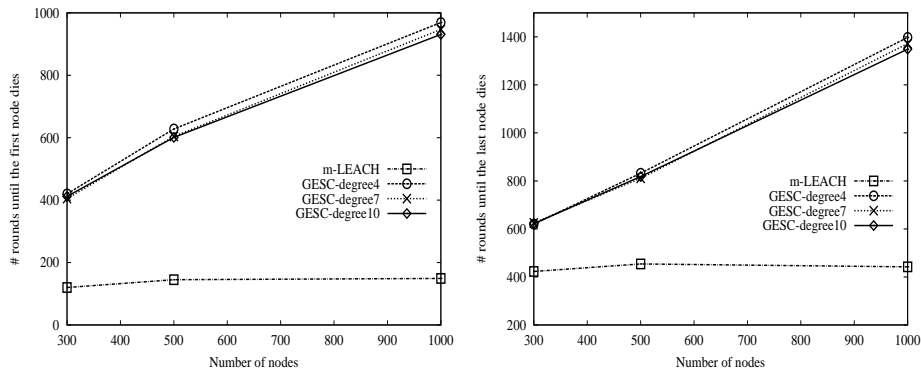
Το πρωτόκολλο LEACH σχηματίζει ομάδες βασιμμένος στη δύναμη του ληφθέντος σήματος και χρησιμοποιεί τις κεφαλές ως αναμεταδότες προς το σταθμό βάση. Όλη η επεξεργασία των δεδομένων όπως η συνάθροιση των δεδομένων γίνεται τοπικά στο επίπεδο της ομάδας. Το LEACH σχηματίζει τις ομάδες χρησιμοποιώντας ένα καταναμημένο αλγόριθμο όπου οι κόμβοι λαμβάνουν αυτόνομα



αποφάσεις χωρίς κάποιον κεντρικό έλεγχο και συντονισμό. Αρχικά ένας κόμβος αποφασίζει να είναι κεφαλή με πιθανότητα  $p$  και εκπέμπει αυτή την απόφασή του. Κάθε απλός κόμβος (κόμβος που δεν είναι κεφαλή) καθορίζει την ομάδα όπου θα ανήκει επιλέγοντας τον κόμβο κεφαλή με τον οποίο μπορεί να επικοινωνεί χρησιμοποιώντας τη μικρότερη δυνατή ενέργεια. Ο ρόλος της κεφαλής εναλλάσσεται περιοδικά μεταξύ των κόμβων της ομάδας για να υπάρξει ισορροπία στη κατανάλωση ενέργειας. Οι συγγραφείς υποθέτουν ότι η κεφαλή έχει μεγάλο εύρος επικοινωνίας ώστε να μπορεί να φθάσει απευθείας το σταθμό βάσης. Παρ' όλα αυτά, αυτή δεν είναι μια ρεαλιστική υπόθεση επειδή οι κεφαλές είναι συνηθισμένοι κόμβοι και ο σταθμός βάσης δεν είναι πάντα άμεσα προσβάσιμος από όλους τους κόμβους. Επιπρόσθετα, οι συγγραφείς θεωρούν μια single-hop επικοινωνία τόσο για την ενδο-ομαδική όσο και την δια-ομαδική επικοινωνία. Έτσι ένας κόμβος μπορεί να φθάσει απευθείας το σταθμό βάσης. Το πρωτόκολλο LEACH δεν είναι εφαρμόσιμο για ΑΔΑ που αναπτύσσονται σε μεγάλες περιοχές. Στις προσομοιώσεις που διεξάγουμε συγκρίνουμε το πρωτόκολλο ομαδοποίησης GESC με μια multi-hop έκδοση του LEACH που περιγράφεται στη μελέτη [66].

Η πειραματική αποτίμηση περιλαμβάνει ένα μεγάλο αριθμό πειραμάτων για διάφορες τοπολογίες δικτύου με 300, 500 και 1000 αισθητήρες και για βαθμό από 4 (πολύ αραιό δίκτυο) μέχρι και 10 (πυκνό δίκτυο). Η περιοχή προσομοίωσης έχει μέγεθος  $500 \times 500m$  και οι τοπολογίες του δικτύου δημιουργούνται όπως στο Κεφάλαιο 6.5. Επιπλέον, ο σταθμός βάσης βρίσκεται στη μέση της τοπολογίας ( $x=250, y=250$ ) και η αρχική ενέργεια κάθε κόμβου αισθητήρα είναι  $E_0=2$  Joules. Τα χαρακτηριστικά της ραδιοεπικοινωνίας είναι όμοια με αυτά που χρησιμοποιούνται στη μελέτη [66]. Κάθε πακέτο δεδομένων έχει μέγεθος 100 bytes, ενώ η κεφαλή του πακέτου έχει μέγεθος 25 bytes. Τέλος, η διάρκεια της δημιουργία των ομάδων και για τα δύο πρωτόκολλα είναι 5 δευτερόλεπτα, ενώ η διάρκεια λειτουργίας του δικτύου είναι 20 δευτερόλεπτα. Η παράμετρος  $k_{opt}$  για το multi-hop LEACH έχει επιλεγεί να είναι 11 για τοπολογίες δικτύου που περιλαμβάνουν 300 με 1000 κόμβους.

Ένας κόμβος θεωρείται “νεκρός” εάν χάσει ολόκληρη την αρχική του ενέργεια. Το Σχήμα 6.7 στο επάνω γράφημα συγκρίνει τη διάρκεια ζωής του δικτύου όταν εκτελείται το πρωτόκολλο GESC και όταν εκτελείται το multi-hop LEACH. Η διάρκεια ζωής του δικτύου είναι ο χρόνος μέχρι ο πρώτος κόμβος να εξαντλήσει την ενέργειά του. Η ομαδοποίηση με βάση το GESC βελτιώνει εμφανώς την επιμήκυνση του χρόνου ζωής του δικτύου σε σύγκριση με το multi-hop LEACH. Το multi-hop LEACH επιλέγει τυχαία τις κεφαλές γεγονός που συνεπάγεται τη μεγαλύτερη διασπάθιση ενέργειας για τους ίδιους κόμβους. Επιπλέον, οι κόμβοι της

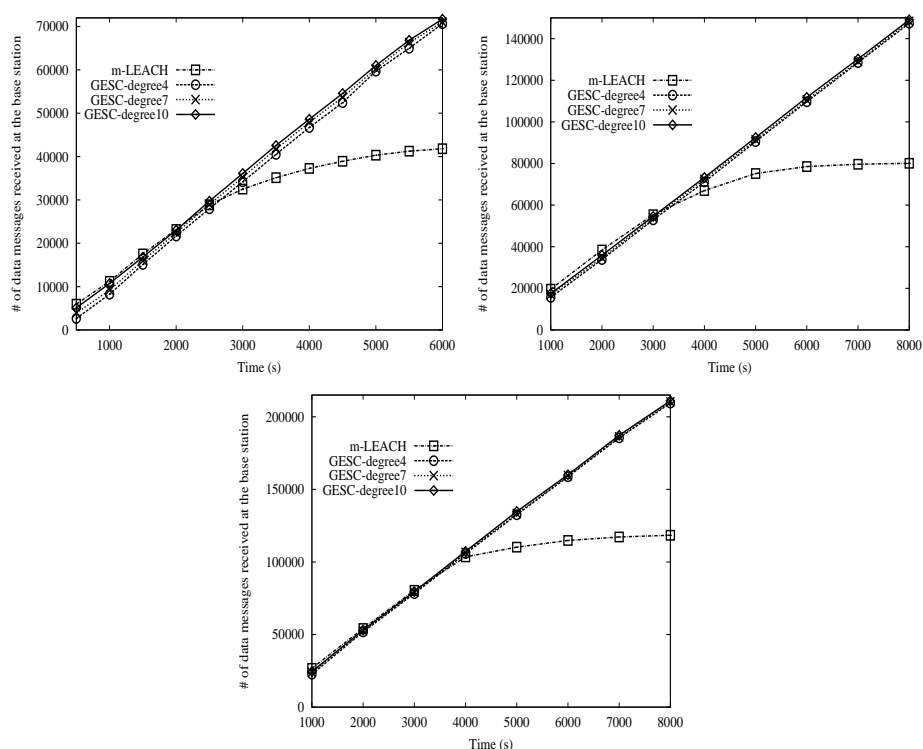


Σχήμα 6.7: Η απόδοση του πρωτοκόλλου GESC σε δικτυακές εφαρμογές. (Επάνω) πρώτος θάνατος κόμβου, (Κάτω) τελευταίος θάνατος κόμβου.

ομάδας και ιδιαίτερα οι κεφαλές πρέπει να μεταδώσουν το σήμα τους σε μεγάλες αποστάσεις και συνεπώς καταναλώνουν πολύ ενέργεια. Αυτό αποφεύγεται με το πρωτόκολλο GESC επειδή κάθε κόμβος μπορεί να επιλέξει διαφορετικές κεφαλές, ενώ οι κεφαλές είναι καλά κατανομημένες στο τοπικό δίκτυο. Επιπρόσθετα, κάθε κόμβος καταναλώνει πολύ λιγότερη ενέργεια για να επικοινωνήσει με τις κεφαλές. Το Σχήμα 6.7 στο κάτω γράφημα παρουσιάζει επίσης ότι το πρωτόκολλο GESC βελτιώνει το χρονικό όριο ζωής του δικτύου έναντι του LEACH.

Το Σχήμα 6.8 δείχνει το συνολικό αριθμό μηνυμάτων που λαμβάνονται στο σταθμό βάση για μια τοπολογία δικτύου με 300, 500 και 1000 κόμβους. Το Σχήμα 6.8 δείχνει ότι το πρωτόκολλο LEACH στέλνει περισσότερα μηνύματα στην αρχή του χρόνου προσομοίωσης σε σχέση με το πρωτόκολλο GESC. Αυτό συμβαίνει επειδή το πρωτόκολλο GESC πρέπει να μεταδώσει ένα μήνυμα μέσω περισσότερων hops από ότι το LEACH για να φθάσει στο σταθμό βάση. Όμως, το GESC αποδίδει καλύτερα από το LEACH κατά τη διάρκεια του υπόλοιπου χρόνου προσομοίωσης, επειδή οι κόμβοι εξαντλούν την ενέργειά τους γρηγορότερα στο LEACH από ότι στο GESC.

Οι συγγραφείς στη μελέτη [66] θεωρούν μια ιδανική συνάθροιση. Τα πλεονεκτήματα της χρήσης του GESC γίνονται περισσότερα όταν χαλαρώνει αυτή η θεώρηση. Επιπρόσθετα, υπάρχουν πολλές περιπτώσεις όπου οι κόμβοι προκειμένου να σώσουν ενέργεια μπορεί να χρειάζεται να μεταδώσουν δεδομένα μόνο αφού παρατηρήσουν κάποιο ενδιαφέρον συμβάν (π.χ., outliers [125]). Συνεπώς σε μια event-driven εφαρμογή το LEACH δεν μπορεί να χρησιμοποιήσει αποδοτικά το εύρος ζώνης, επειδή όλοι οι κόμβοι επικοινωνούν συνεχώς με τη κεφαλή.



Σχήμα 6.8: Ο συνολικός αριθμός από δεδομένα που έλαβε ο σταθμός βάσης για ένα δίκτυο: (Επάνω) 300 και 500 κόμβων, (Κάτω) 1000 κόμβων.

## 6.7 Συμπεράσματα

Στο Κεφάλαιο 6 εξετάσαμε το ζήτημα της ομαδοποίησης σε ΑΔΑ, η οποία επιτρέπει τη συνάθροιση της πληροφορίας, περιορίζει τη μετάδοση των δεδομένων κυρίως μέσα στην ομάδα, μειώνει το συναγωνισμό για την απόκτηση του ασύρματου καναλιού και ελαττώνει την κατανάλωση ενέργειας. Προτείναμε ένα νέο ενεργειακά αποδοτικό καταναμημένο πρωτόκολλο ομαδοποίησης για ΑΔΑ, το GESC. Το πρωτόκολλο GESC λαμβάνει υπόψη τόσο την κατανάλωση ενέργειας όσο και τα τοπολογικά στοιχεία σε αντίθεση με τα προηγούμενα πρωτόκολλα που ελάμβαναν υπόψη μόνο τον ένα από τους δύο παράγοντες. Το νέο πρωτόκολλο αξιοποιεί την τοπική δομή του δικτύου και την εναπομείνουσα ενέργεια των γειτονικών κόμβων για να ορίσει ένα νέο τρόπο δυναμικού υπολογισμού των κεφαλών. Το πρωτόκολλο αποδείχθηκε ότι είναι πολύ αποδοτικό και ότι συμβάλλει τόσο στη μείωση των μηνυμάτων που διαδίδονται στο δίκτυο, όσο και στη μείωση της διασπάθισης της ενέργειας.



## ΚΕΦΑΛΑΙΟ 7

# Καταλογοποίηση σε Ασύρματα Δίκτυα\*

### Περιεχόμενα

---

7.1	ΕΙΣΑΓΩΓΗ . . . . .	161
7.2	ΣΧΕΤΙΚΕΣ ΕΡΓΑΣΙΕΣ . . . . .	165
7.3	PARTIAL ORDERING BROADCAST INDEX . . . . .	170
7.4	ΠΕΙΡΑΜΑΤΙΚΗ ΑΠΟΤΙΜΗΣΗ . . . . .	172
7.5	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	179

---

### 7.1 Εισαγωγή

Η διατήρηση της ενέργειας των κινητών πελατών και η μικρή καθυστέρηση πρόσβασης στα δεδομένα είναι δύο θεμελιώδεις αλλά δυστυχώς αντιμαχόμενοι στόχοι για τα ασύρματα δίκτυα, όπου οι φτωχοί σε πόρους κόμβοι βασίζονται στην μετάδοση με εκπομπή για να ανακτήσουν πληροφορίες. Για να αντιμετωπισθεί το μειονέκτημα της μεγάλης κατανάλωσης ενέργειας εξαιτίας της ακολουθιακής αναζήτησης των σχετικών δεδομένων σε ασύρματα δίκτυα που βασίζονται στην μετάδοση με εκπομπή χωρίς παράλληλα να επιβαρύνεται η καθυστέρηση προσπέλασης των δεδομένων, έχει προταθεί η παρεμβολή ενός καταλόγου. Παρότι στη

---

\*Το κεφάλαιο αντλεί τα περιεχόμενά του από την εργασία [80].

σχετική βιβλιογραφία έχει προταθεί ένας μεγάλος αριθμός καταλόγων, εντούτοις παρουσιάζουν κάποια μειονεκτήματα. Στο παρόν κεφάλαιο προσπαθούμε να αντιμετωπίσουμε τα προβλήματα εισάγοντας ένα νέο μη-ισοζυγισμένο δενδρικό κατάλογο που αποτελεί γενίκευση δύο προϋπαρχόντων που αποτελούσαν τις καλύτερες σχετικές δομές. Η ανάπτυξη αυτού του καταλόγου εισαγάγει για πρώτη φορά στη σχετική βιβλιογραφία του κινητού υπολογισμού, το πρόβλημα της δημιουργίας καταλόγων για μερικώς διατεταγμένα δεδομένα.

Έστω το ακόλουθο σενάριο από τις εφαρμογές των δικτύων αισθητήρων, όπου ένας κόμβος με ποικίλες δυνατότητες ανίχνευσης (π.χ., θερμοκρασία, υγρασία, πίεση, συγκέντρωση διοξειδίου του άνθρακα), είναι σε θέση να ανιχνεύει το περιβάλλον σε τακτικά χρονικά διαστήματα (διαφορετικά για κάθε μετρούμενη ποσότητα). Οι μετρήσεις αυτές εκπέμπονται και συλλέγονται από τους κοντινούς κόμβους που υλοποιούν διάφορες εφαρμογές σχετιζόμενες είτε με θέματα θερμοκρασίας, είτε με συγκεντρώσεις διοξειδίου του άνθρακα, είτε με την υγρασία, κ.τ.λ. είτε με διάφορους συνδυασμούς όλων των προηγούμενων. Ένα τέτοιο σενάριο μπορεί να δημιουργηθεί με βάση την αρχιτεκτονική που περιγράφεται στη μελέτη [39]. Είναι φανερό πως η διάταξη των μετρήσεων για την ίδια ποσότητα (π.χ., θερμοκρασία, πίεση) στα πλαίσια ενός χρονικού διαστήματος έχει νόημα, αλλά η διάταξη των μετρήσεων διαφορετικών ποσοτήτων για την ίδια χρονική στιγμή δεν έχει ιδιαίτερη σημασία. Οι γειτονικοί κόμβοι, που εξυπηρετούν τις διαφορετικές εφαρμογές και έχουν περιορισμένη επάρκεια ενέργειας, θα απαιτήσουν ενεργειακά αποδοτική πρόσβαση στην εκπεμπόμενη πληροφορία (και όχι διαρκώς να “ακούν” το ασύρματο κανάλι εκπομπής). Το πρότυπο προσπέλασης των αισθητήρων στη διαδιδόμενη πληροφορία θα είναι ετερογενές π.χ. κυρτό, εξαιτίας των διαφορετικών προτιμήσεων των εφαρμογών στις μετρούμενες ποσότητες.

Σε ένα δεύτερο πιθανό σενάριο που συναντάται σε ασύρματα κυψελοειδή δίκτυα (π.χ., PCS), οι κινητές μονάδες με περιορισμένους πόρους, οι οποίες ανήκουν σε ένα ασύρματο κελί ανακτούν πληροφορίες από μια σχεσιακή βάση δεδομένων. Τα δεδομένα της βάσης εκπέμπονται από ένα σταθμό βάσης που εξυπηρετεί το κελί. Η πρόσφατη Smart Personal Objects Technology (SPOT) της Microsoft<sup>1</sup> αποδεικνύει το βιομηχανικό ενδιαφέρον σε τέτοιου είδους υπηρεσίες. Στη γενική περίπτωση, τα τμήματα της πληροφορίας αποτελούνται από προβολές (π.χ., στήλες) των πλειάδων του πίνακα για τις οποίες το κλειδί μπορεί να διαφέρει μόνο μεταξύ των πλειάδων και όχι μεταξύ των στηλών της ίδιας πλειάδας. Προφανώς,

<sup>1</sup><http://direct.msn.com/>

οι κινητοί χρήστες ενδιαφέρονται για διαφορετικούς συνδυασμούς των πλειάδων και των στηλών και όπως είναι φυσικό σε ρεαλιστικές καταστάσεις, το πρότυπο προσπέλασης των δεδομένων είναι κυρτό.

Τελικά, ας εξετάσουμε ένα τρίτο σενάριο εμπνευσμένο από τα σύγχρονα αυτοματοποιημένα πεδία μάχης όπου αναπτύσσεται ένα ad hoc δίκτυο. Ένας ισχυρός ενεργειακά κόμβος έχει σχεδιασθεί για να παρέχει πληροφορίες σχετικά με κινήσεις τακτικής στα φίλια στρατεύματα. Ο κόμβος αυτός εκπέμπει αναφορές σε μικρές κινητές μονάδες με περιορισμένη ενέργεια και τις οποίες φέρουν για παράδειγμα είτε απλοί (ή ομάδες από) στρατιώτες είτε ανώνυμα οχήματα εξερεύνησης. Προφανώς, όλοι η εκπεμπόμενη πληροφορία δεν είναι δυνατόν να ενδιαφέρει όλες τις μονάδες που λαμβάνουν τις εκπομπές. Επίσης υπάρχει η περίπτωση όπου το “περιγραφικό κλειδί” των τμημάτων της πληροφορίας δεν επιβάλλει κάποια συνολική διάταξη μεταξύ των τμημάτων της πληροφορίας.

Στα προηγούμενα σενάρια εφαρμογών είναι προφανές ότι:

- Οι καταναλωτές της πληροφορίας χρειάζεται να ανακτήσουν τα δεδομένα όσο το δυνατόν γρηγορότερα (π.χ., με μικρή καθυστέρηση πρόσβασης, που είναι ο χρόνος που έχει περάσει μεταξύ της στιγμής που ανέκυψε η ανάγκη για τα δεδομένα και η στιγμή που ανακτήθηκαν τα δεδομένα από το ασύρματο κανάλι).
- Οι καταναλωτές είναι κόμβοι με περιορισμένη ενέργεια. Επομένως, πρέπει να αποφεύγουν τη διαρκή παρακολούθηση του ασύρματου καναλιού, επειδή αυτό υποχρεώνει τον κόμβο στην κατανάλωση μιας μικρής ποσότητας ενέργειας.
- Δεν υπάρχει συνολική διάταξη μεταξύ των δεδομένων αλλά μόνο μερική.
- Το πρότυπο προσπέλασης στην εκπεμπόμενη πληροφορία εμφανίζει κυρτότητα. Αυτό σημαίνει ότι κάποια δεδομένα παρουσιάζουν μεγαλύτερο ενδιαφέρον και ζήτηση από ότι κάποια άλλα.

Η αποδοτικότητα στην προσπέλαση είναι ένας κοινός στόχος για πολλά συστήματα (π.χ., βάσεις δεδομένων), αλλά η διατήρηση της ενέργειας είναι ζωτικής σημασίας στόχος για την επιμήκυνση της διάρκειας ζωής ενός ασύρματου δικτύου αισθητήρων ή για την επίτευξη όσο το δυνατόν μεγαλύτερης ανεξαρτησίας των κινητών κόμβων. Οι κινητοί κόμβοι υποστηρίζουν δύο γενικές καταστάσεις λειτουργίας, την ενεργό κατάσταση που είναι μια πλήρως λειτουργική κατάσταση και η κατάσταση “ύπνου” (με βάση την ορολογία της μελέτης [74]), η οποία είναι

μια κατάσταση διατήρησης της ενέργειας. Ο ρυθμός κατανάλωσης της ενέργειας μεταξύ των δύο καταστάσεων παρουσιάζει πολύ μεγάλη διαφορά [134]. Παρόμοια, οι κόμβοι αισθητήρες μπορούν να βρίσκονται σε μία από τις τρεις ενεργές καταστάσεις  $\dot{\nu}$  μετάδοση, λήψη, αδρανής – ή σε κατάσταση ύπνου. Ένας αισθητήρας καταναλώνει 7-20 φορές λιγότερη ενέργεια όταν βρίσκεται σε κατάσταση ύπνου [53].

Είναι προφανές ότι όταν μεταδίδονται δεδομένα στο ασύρματο κανάλι, ο κόμβος δεν έχει άλλη εναλλακτική από το να παρακολουθεί συνεχώς το ασύρματο μέσο μέχρι να λάβει τα δεδομένα. Η διαρκής παρακολούθηση όμως του καναλιού έχει ως αποτέλεσμα τη μεγάλη κατανάλωση ενέργειας. Για να επιτρέψουμε στους κόμβους να βρίσκονται σε μια κατάσταση ύπνου έτσι ώστε να μη σπαταλούν ενέργεια, πρέπει με κάποιο τρόπο να ενσωματώσουμε καταλόγους μέσα στα δεδομένα εκπομπής ή πριν από αυτά. Χρησιμοποιώντας τον εναέριο κατάλογο (air indexing) [74], ο οποίος περιλαμβάνει πληροφορία σχετικά με το χρόνο άφιξης των δεδομένων, οι κόμβοι δεν χρειάζεται να παρακολουθούν διαρκώς το κανάλι. Αντίθετα, μπορούν να “ξυπνούν” σε συγκεκριμένες χρονικές στιγμές, όταν πρόκειται να μεταδοθούν τα ζητούμενα δεδομένα.

Οι κατάλογοι για δευτερεύουσα μνήμη είναι ένα θέμα που έχει διερευνηθεί διεξοδικά στην περιοχή των Βάσεων Δεδομένων [93]. Μερικές από αυτές τις τεχνικές έχουν προσαρμοσθεί στην ασύρματη περίπτωση, π.χ., κατακερματισμός [73, 142],  $B^+$ -δένδρο [74, 144, 147], όπου οι μέθοδοι που βασιζονται στο  $B^+$ -δένδρο παρουσιάζουν την καλύτερη συνολικά απόδοση όπως αποδεικνύεται άλλωστε στη μελέτη [144].

Παρ’ όλα αυτά, τα υπάρχοντα σχήματα καταλόγων (δες στη Παράγραφο 7.2.2) έχουν σχεδιασθεί υποθέτοντας ότι υπάρχει μια συνολική διάταξη μεταξύ των δεδομένων. Η μόνη εξαίρεση είναι ο δενδρικός κατάλογος που προτάθηκε στη μελέτη [31], ο οποίος δεν υποθέτει την ύπαρξη διάταξης μεταξύ των δεδομένων. Δυστυχώς, το σχήμα αυτό δεν είναι μια δομή καταλόγου, επειδή απαιτεί τη σάρωση ολόκληρου του δένδρου προκειμένου να ανακαλύψει ένα αντικείμενο. Αυτό συμβαίνει επειδή το δένδρο δεν προσφέρει κάποια καθοδήγηση (π.χ., ακολουθία διακλάδωσης). Επιπρόσθετα, τα περισσότερα σχήματα καταλόγων [74, 142, 144, 147] έχουν προταθεί για ομοιόμορφα πρότυπα προσπέλασης και συνεπώς αγνοούν την πιθανότητα κύρτωσης στο πρότυπο προσπέλασης. Μια λεπτομερής περιγραφή των σχετικών εργασιών καθώς και των ελαττωμάτων τους θα παρουσιασθούν στην Παράγραφο 7.2.2.

Στο παρόν κεφάλαιο προτείνουμε ένα νέο μη ισοζυγισμένο δενδρικό κατάλογο που ονομάζεται *Partial Ordering Broadcast Index*, συντομογραφικά POBI.



Ο νέος κατάλογος μπορεί να διαπλέκεται ή να προηγείται των εκπεμπόμενων δεδομένων έτσι ώστε να ελαχιστοποιηθεί η σπατάλη ενέργειας και η καθυστέρηση πρόσβασης των δεδομένων. Η προτεινόμενη δομή είναι φτιαγμένη για περιπτώσεις ρεαλιστικής κύρτωσης στο πρότυπο προσπέλασης των δεδομένων και όχι για μη ρεαλιστικά ομοιόμορφα πρότυπα. Η καινοτομία της νέας δομής είναι το γεγονός ότι μπορεί να διαχειρισθεί μερικώς διατεταγμένα δεδομένα, το οποίο δεν είχε παρουσιασθεί σε καμία προηγούμενη έρευνα. Το σημείο κλειδί είναι ότι ο νέος κατάλογος είναι μια φυσική γενίκευση των δύο αρτιότερων δομών καταλόγων, όπως το *Variant Fanout*, VF [31] και το *k-ary Alphabetic Tree*, kAT [119]. Στο τέλος του κεφαλαίου παρουσιάζεται μια λεπτομερής πειραματική ανάλυση της απόδοσης του καταλόγου. Έχει διεξαχθεί ένας μεγάλος αριθμός πειραμάτων για να συγκριθεί η νέα δομή με προηγούμενα σχήματα. Τα πειραματικά αποτελέσματα παρουσιάζουν ότι ο προτεινόμενος κατάλογος πετυχαίνει καλύτερη απόδοση και ευκαμψία από τα ανταγωνιστικά σχήματα.

## 7.2 Σχετικές Εργασίες

### 7.2.1 Προκαταρκτικά

Στο παρόν κεφάλαιο θα παρουσιάσουμε τα προκαταρκτικά στοιχεία της έρευνας και τις σχετικές εργασίες. Υποθέτουμε ένα γενικό σύστημα διάχυσης των δεδομένων, το οποίο εκπέμπει πληροφορίες στους πελάτες (καταναλωτές των πληροφοριών). Επιπρόσθετα, υπάρχει ένας κόμβος με μεγάλη αυτονομία και επάρκεια πόρων, ο εξυπηρετητής, ο οποίος εκπέμπει τις πληροφορίες μέσω του ασύρματου καναλιού. Κάθε κομμάτι πληροφορίας μπορεί να αναγνωρισθεί από μια τιμή κλειδιού. Η τιμή του κλειδιού δεν συνεπάγεται ότι τα δεδομένα είναι πλήρως διατεταγμένα, αλλά μόνο μερικώς. Οι κόμβοι του δικτύου είναι συσκευές με περιορισμένους πόρους (αντίθετα από τον εξυπηρετητή) και μπορούν να είναι για παράδειγμα κόμβοι ενός ασύρματου δικτύου αισθητήρων ή κινητοί κόμβοι που βρίσκονται μέσα στη περιοχή κάλυψης ενός συστήματος προσωπικών επικοινωνιών. Επίσης, θεωρούμε ότι ο εξυπηρετητής μπορεί να υπολογίζει τη δημοφιλία κάποιων τμημάτων της πληροφορίας. Οι κόμβοι συντονίζονται με το ασύρματο κανάλι εκπομπής και ανακτούν τα πακέτα που περιέχουν πληροφορίες καταλόγου μέχρι να λάβουν την πληροφορία για το χρόνο άφιξης του ζητούμενου δεδομένου. Αυτός ο επιλεκτικός μηχανισμός συντονισμού συνεισφέρει στη μείωση της κατανάλωσης της ενέργειας.

Συνακόλουθα, υποθέτουμε ότι υπάρχει ένα κανάλι εκπομπής. Αν και είναι

πιθανό να υπάρχουν πολλαπλά κανάλια, η χρησιμοποίησή τους σε MANET και AΔΑ δεν προτιμάται. Παρ' όλα αυτά, ο νέος κατάλογος POBI μπορεί με ελάχιστες σχεδιαστικές αλλαγές να προσαρμοσθεί στην ύπαρξη πολλαπλών καναλιών.

Αν και δεν αποτελεί κρίσιμο σημείο της μελέτης, υποθέτουμε μία *επίπεδη εκπομπή*, όπου κάθε δεδομένο εμφανίζεται ακριβώς μία φορά και όχι μια εκπομπή που παρουσιάζει κύρτωση όπου κάποια δεδομένα μπορεί να εμφανίζεται περισσότερες από μία φορές. Ο νέος κατάλογος POBI μπορεί να επεκταθεί για να υποστηρίζει και αντίγραφα δεδομένων. Εκτός όμως από τα αντίγραφα δεδομένων, υπάρχουν client-side τεχνικές που μπορούν επίσης να χρησιμοποιηθούν για να μειωθεί ο χρόνος προσπέλασης, όπως το caching [81] και το prefetching [83]. Οι τεχνικές αυτές όμως είναι ορθογώνιες στην εργασίας που παρουσιάζεται στο παρόν κεφάλαιο και δεν θα αναπτυχθούν περαιτέρω.

Στην παρούσα έρευνα υποθέτουμε ότι ένας εξυπηρετητής εκπέμπει  $n$  ίσου μεγέθους εγγραφές δεδομένων που συμβολίζονται με  $R_i$  ( $1 \leq i \leq n$ ) και έχουν πιθανότητα προσπέλασης  $Pr(R_i)$ , όπου  $\sum_i Pr(R_i) = 1$ . Θεωρούμε επίσης ότι  $I_{pb}(R_i)$  είναι ο αριθμός των διερευνήσεων του καταλόγου για να φθάσει ένας κόμβος στο  $R_i$  του δενδρικού καταλόγου και  $d(a_i)$  είναι ο παράγοντας διακλάδωσης (fanout) ενός κόμβου  $a_i$ . Επιπρόσθετα, συμβολίζουμε με  $Path(R_i)$  το σύνολο των κόμβων του καταλόγου από τη ρίζα του δένδρου μέχρι το  $R_i$ . Ομοίως με τη μελέτη [31], ορίζουμε το μέσο κόστος ενός δενδρικού καταλόγου ως εξής:

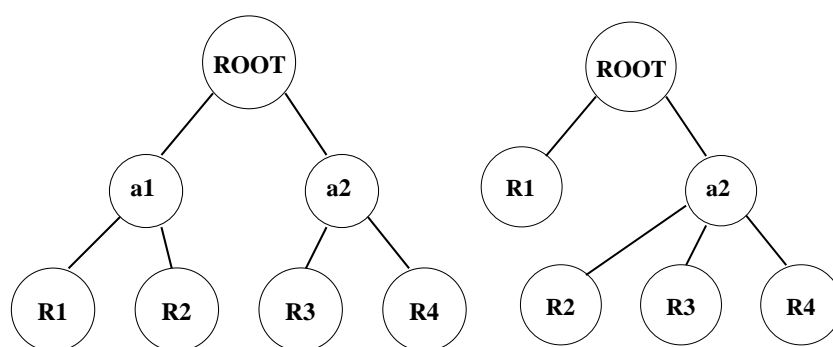
$$\sum_{1 \leq i \leq n} Pr(R_i) \times I_{pb}(R_i), \quad (7.1)$$

όπου το κόστος διερεύνησης του καταλόγου υπολογίζεται ως εξής:

$$I_{pb}(R_i) = \sum_{a_j \in Path(R_i)} d(a_j). \quad (7.2)$$

Αν και ο ορισμός εναλλακτικών μοντέλων κόστους είναι πιθανός, το μοντέλο που επιλέχθηκε είναι το περισσότερο ρεαλιστικό στη βιβλιογραφία. Έχοντας ορίσει τον τύπο κόστους, μπορούμε εύκολα να διαπιστώσουμε γιατί το μη ομοιόμορφο πρότυπο προσπέλασης συνδυασμένο με μη ισοζυγισμένες δενδρικές δομές προσφέρει δυνατότητες για καλύτερη απόδοση του συστήματος. Αρχικά, θεωρούμε τέσσερα αντικείμενα με ομοιόμορφο πρότυπο προσπέλασης, π.χ.,  $R_1=R_2=R_3=R_4=0.25$  (διατεταγμένα ως  $R_1 < R_2 < R_3 < R_4$ ) και υποθέτουμε ότι κατασκευάζουμε ένα ισορροπημένο δένδρο όπως αυτό της μελέτης [74] με παράγοντα διακλάδωσης  $d(a_i)=2$ . Στη συνέχεια, το παραγόμενο δένδρο θα είναι αυτό που απεικονίζεται στα αριστερά του Σχήματος 7.1 και το κόστος του θα είναι ίσο

με  $\sum_{1 \leq i \leq 4} Pr(R_i) \times I_{pb}(R_i) = \sum_{1 \leq i \leq 4} 0.25 \times 4 = 4$ . Στο σημείο αυτό πραγματοποιούμε μια διαφορετική υπόθεση, θεωρώντας ότι το πρότυπο προσπέλασης παρουσιάζει κυρτότητα και οι πιθανότητες είναι  $R_1=0.40, R_2=R_3=R_4=0.20$ . Στη συνέχεια το δένδρο στα αριστερά του Σχήματος 7.1 θα έχει ως αποτέλεσμα ένα κόστος ίσο με 4, ενώ το δένδρο στα δεξιά του Σχήματος 7.1 θα έχει ένα κόστος ίσο με 3.8. Συνεπώς παρουσιάζεται μια βελτίωση ίση με 5% ακόμα και για αυτό το πολύ μικρό παράδειγμα.



Σχήμα 7.1: Οφέλη από την αξιοποίηση της κυρτότητας στην προσπέλαση και του μη ισοζυγισμένου δένδρου.

Σκοπός της έρευνας που παρουσιάζεται στο παρόν κεφάλαιο είναι η δημιουργία ενός δενδρικού καταλόγου ελαχιστοποιώντας το κόστος της Εξίσωσης 7.1, υπό την προϋπόθεση της μη ομοιόμορφης προσπέλασης στα δεδομένα και της μερικής διάταξης των δεδομένων. Ο ορισμός του συνολικού προβλήματος θα δοθεί στο Κεφάλαιο 7.3.

### 7.2.2 Υπόβαθρο

Οι κατάλογοι δευτερεύουσας μνήμης (όπως,  $B^+$ -δένδρα, Κατακερματισμός,  $R^+$ -δένδρα, Γραμμικά Τετραδικά Δένδρα, Λίστες Παράλειψης) για παραδοσιακές και προηγμένες εφαρμογές είναι μια περιοχή που έχει ερευνηθεί διεξοδικά στο παρελθόν [93]. Μια προσαρμογή των  $B^+$ -δένδρων σε ασύρματα περιβάλλοντα περιγράφηκε πρώτη φορά στη μελέτη [74], όπου τα φύλλα του  $B^+$ -δένδρου αποθηκεύουν το χρόνο άφιξης κάθε δεδομένου στο ασύρματο κανάλι και όχι τις διευθύνσεις στο δίσκο. Παρομοίως, μια προσαρμογή του παραδοσιακού κατακερματισμού σε ασύρματα περιβάλλοντα παρουσιάστηκε νωρίτερα στη μελέτη [73], ενώ αργότερα περιγράφηκε μια γενίκευση στη μελέτη [142]. Υβριδικές τεχνικές μεταξύ των δύο προσεγγίσεων έχουν περιγραφεί στις μελέτες [116, 144], ενώ εφαρμογές των Δένδρων Υπογραφών ως καταλόγων αναφέρονται στη μελέτη [85].

Επιπρόσθετα, η προσαρμογή τέτοιων δομών καταλόγων (π.χ.,  $B^+$ -δένδρο), τα οποία απευθύνονται σε πολλαπλά κανάλια εκπομπής περιγράφονται στις μελέτες [70, 146]. Οι μελέτες αυτές δεν προτείνουν νέα σχήματα αλλά διαφορετικές μεθόδους κατανομής για τους κόμβους του δενδρικού καταλόγου. Σε όλες αυτές τις εργασίες έχει υποθεθεί ότι: (α) υπάρχει μια συνολική διάταξη μεταξύ των μεταδιδόμενων δεδομένων, και (β) το πρότυπο προσπέλασης είναι ομοιόμορφο γεγονός που σημαίνει ότι η πιθανότητα προσπέλασης είναι η ίδια για όλα τα δεδομένα.

Παρεκκλίνοντας από την ομοιόμορφη πιθανότητα προσπέλασης, παρατηρούμε ποικίλες εργασίες που λογαριάζουν την επίδραση της κυρτότητας στην προσπέλαση για το σχεδιασμό δομών καταλόγων. Ένα σχήμα προτάθηκε στη μελέτη [119], το οποίο είναι μια εκδοχή της  $k$ -ary εκδοχής του βασικού δυαδικού Alphabetic δένδρου [71], ενώ οι μελέτες [143, 146] προσαρμόζουν τη μέθοδο καταλόγου της μελέτης [74] για να διαχειρισθεί τη μη ομοιομορφία στην προσπέλαση. Αρκετές μέθοδοι έχουν βασισθεί στη δημιουργία ενός δυαδικού ή  $k$ -ary Alphabetic Tree για να αναπτύξουν σχήματα καταλόγων για πολλαπλά κανάλια εκπομπής [77, 92]. Οι μέθοδοι αυτές δεν παρέχουν νέους τύπους δενδρικών καταλόγων αλλά νέα μέθοδο κατανομής για τα Alphabetic trees στα πολλαπλά κανάλια. Όλες οι εργασίες [77, 92, 119, 143, 146] θεωρούν ότι υπάρχει μια συνολική διάταξη μεταξύ των μεταδιδόμενων δεδομένων.

Μόνο η εργασία [31] παρεκκλίνει και από τις δύο υποθέσεις: (α) της ομοιομορφίας, και (β) της συνολικής διάταξης. Παρ' όλα αυτά, γίνεται η υπόθεση της μη ύπαρξης διάταξης για τα δεδομένα και συνεπώς ο προτεινόμενος κατάλογος δεν αποτελεί δομή αναζήτησης.

Το ζήτημα των καταλόγων έχει λάβει μικρή προσοχή στα ΑΔΑ επειδή η πλειοψηφία της έρευνας έχει επικεντρωθεί σε θέματα όπως είναι η δρομολόγηση, η ομαδοποίηση, ο προγραμματισμός της κατάστασης ύπνου, έλεγχος πρόσβασης μέσου και caching. Συνεπώς, έχουν αναπτυχθεί κάποιοι μεμονωμένοι καταναμημένοι κατάλογοι που βρίσκονται στους αισθητήρες και δεν εκπέμπονται. Οι κατάλογοι αυτοί περικλείουν (με τον έναν ή τον άλλο τρόπο) προσαρμογές των παραδοσιακών δομών για δευτερεύουσα μνήμη και λαμβάνεται ιδιαίτερη φροντίδα για την επίτευξη τοπικής μόνο επικοινωνίας κατά τη δημιουργία τους και μικρό κόστος αποθήκευσης. Το GHT [109] είναι βασισμένο σε ένα (γεωγραφικό) σχήμα κατακερματισμού, το DIM [87], το DIFS [60] και το DIST [96] είναι βασισμένα στη δομή των τετραδικών δένδρων, και το TSAR [39] είναι βασισμένο σε Γράφους Παράλειψης (skip graphs), μια γενίκευση των skip lists για καταναμημένα περιβάλλοντα. Κανείς από τους προηγούμενους καταλόγους δεν εκπέμπεται στο

ασύρματο κανάλι ενώ όλα θεωρούν μια συνολική διάταξη των δεδομένων.

### 7.2.3 Κατάλογοι για μερικώς διατεταγμένα δεδομένα

Υποθέτουμε (όπως επισημαίνεται στη Παράγραφο 7.2.1) ότι δίνεται ένα σύνολο  $n$  δεδομένων ίσου μεγέθους, το κάθε ένα με πιθανότητα προσπέλασης  $R_i$  (όπου  $1 \leq i \leq n$ ). Επίσης, υποθέτουμε ότι τα  $n$  δεδομένα είναι κατανομημένα σε  $m$  bins, όπου κάθε bin  $B_j$  περιέχει  $|B_j|$  αντικείμενα (όπου  $1 \leq j \leq m$ ). Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι το bin  $B_1$  περιέχει τα αντικείμενα  $1, \dots, |B_1|$ , το bin  $B_2$  τα αντικείμενα  $|B_1| + 1, \dots, |B_1| + |B_2|$ , και το bin  $B_l$  τα αντικείμενα  $\sum_{1 \leq x \leq l-1} |B_x| + 1, \dots, \sum_{1 \leq x \leq l-1} |B_x| + |B_l|$ , όπου  $\sum_{1 \leq j \leq m} |B_j| = n$ . Επιπλέον, υποθέτουμε ότι για τα αντικείμενα ενός bin δεν υπάρχει κάποια διάταξη, οπότε κάθε αντικείμενο του bin  $B_i$  είναι μικρότερο από κάθε αντικείμενο του bin  $B_j$  για  $i < j$ . Επομένως, ο στόχος μας μπορεί να σχηματιστεί με μαθηματικούς όρους ως εξής:

**ΟΡΙΣΜΟΣ 7.1 (Partial Ordering Broadcast Index).** Δοθέντος του αριθμού  $n$  των δεδομένων, των πιθανοτήτων προσπέλασής τους  $R_i$ , του αριθμού  $m$  των bins  $B_j$  και μιας συνάρτησης συμμετοχής (που καθορίζει ποιο αντικείμενο ανήκει σε ποιο bin), στόχος μας είναι η κατασκευή ενός δενδρικού καταλόγου ελαχιστοποιώντας το κόστος της Εξίσωσης 7.1, έτσι ώστε στο παραγόμενο δένδρο, το αντικείμενο  $x$  να προηγείται του αντικειμένου  $y$  σε μια μη σειριακή διάσχιση του δένδρου, εάν  $x \in B_i$  και  $y \in B_j$  για  $i < j$ . Η ακολουθία των αντικειμένων μέσα σε ένα bin είναι αυθαίρετη.

### 7.2.4 Γενικεύοντας τους προηγούμενους καταλόγους

Προφανώς, ο ορισμός του προβλήματος υιοθετεί τα προβλήματα που ορίστηκαν στις μελέτες [119, 31]. Εάν θεωρήσουμε ότι υπάρχει μόνο ένα bin, δηλαδή  $m=1$ , τότε το πρόβλημά δεν περιορίζεται σε αυτό της μελέτης [31], επειδή δεν θα υπάρχει κάποια διάταξη μεταξύ των δεδομένων, και συνεπώς μπορεί να επιλυθεί με βάση τη δομή του (δυαδικού ή  $k$ -ary) Huffman δένδρου. Αντιθέτως, εάν υπάρχουν τόσα bins όσο και ο αριθμός των αντικειμένων ( $m=n$ ) και κάθε bin περιλαμβάνει μόνο ένα αντικείμενο, τότε το πρόβλημά μας ανάγεται σε αυτό της μελέτης [119], επειδή θα συνεπάγεται μια συνολική διάταξη μεταξύ όλων των αντικειμένων και συνεπώς το πρόβλημα μπορεί να επιλυθεί με βάση τη δομή του (δυαδικού ή  $k$ -ary) Alphabetic δένδρου.

### 7.3 Partial Ordering Broadcast Index

Μια αρχική απλοϊκή προσέγγιση για την αντιμετώπιση του προβλήματος είναι η δημιουργία κάθε πιθανού συνδυασμού των  $n$  αντικειμένων (με σεβασμό στη συμμετοχή στην ομάδα και τη διάταξη μεταξύ των ομάδων) και η κατασκευή ενός alphabetic δένδρου επάνω από αυτά, επιλέγοντας αυτό με το μικρότερο κόστος. Είναι φανερό ότι μία τέτοια μέθοδος είναι μη αποδοτική επειδή θα πρέπει να εξετάσουμε εκθετικά πολλούς συνδυασμούς, περίπου το γινόμενο των πιθανών συνδυασμών κάθε ομάδας. Πράγματι, διαπιστώνουμε ότι δεν μπορούμε να σχεδιάσουμε έναν αλγόριθμο πολυωνυμικού χρόνου για να επιλύσουμε το πρόβλημα. Συνεπώς, η αξιοποίηση έξυπνων προσεγγίσεων με μικρό χρόνο εκτέλεσης είναι προτιμητέα.

Μια φθηνή λύση είναι να επιλέξουμε μια τυχαία διάταξη για τα δεδομένα μέσα σε κάθε ομάδα και στη συνέχεια να εφαρμόσουμε σε όλα τα δεδομένα τον αλγόριθμο κατασκευής ενός δυαδικού ή  $k$ -ary alphabetic δένδρου. Παρ' όλα αυτά μπορούμε να επιτύχουμε καλύτερη λύση από την προηγούμενη. Το απλοϊκό αυτό σχήμα θα το αποκαλούμε kATr, όπου η παράμετρος  $k$  παριστά τον παράγοντα διακλάδωσης (για ένα δυαδικό δένδρο  $k=2$ ).

Μπορεί όμως η διάταξη των αντικειμένων μέσα σε κάθε ομάδα να παρέχει κάποια οφέλη; Μπορεί να αποδειχθεί ότι για την ειδική περίπτωση όπου όλες οι ομάδες, εκτός από την πρώτη και την τελευταία, περιέχουν μόνο ένα αντικείμενο, τότε το POBI μπορεί να πετύχει το χαμηλότερο κόστος εάν διατάξουμε τα αντικείμενα της πρώτης ομάδας σε μη αύξουσα σειρά και τα αντικείμενα της τελευταίας ομάδας σε μη φθίνουσα σειρά. Όμως στην περίπτωση όπου διάφορες ομάδες περιέχουν περισσότερα του ενός αντικείμενα, και όχι μόνο η πρώτη και η τελευταία, τότε η προηγούμενη διάταξη δεν μπορεί να οδηγήσει στη βέλτιστη λύση. Ωστόσο, με τη χρήση τέτοιων διατάξεων μπορούμε να σχεδιάσουμε κάποιες προσεγγίσεις στο πρόβλημά μας. Πιο συγκεκριμένα, θα μπορούσε να δημιουργηθεί μια ευρεστική ταξινομώντας τα αντικείμενα κάθε ομάδας σε μη φθίνουσα σειρά και στη συνέχεια κατασκευάζοντας ένα alphabetic δένδρο που ονομάζουμε kATi. Παρομοίως, με την ταξινόμηση των αντικειμένων κάθε ομάδας σε μη αύξουσα σειρά δημιουργούμε μια άλλη ευρεστική που ονομάζεται kATd.

Οι προηγούμενες προσεγγίσεις χρησιμοποιούν κάποια διάταξη με βάση τη δημοφιλία μεταξύ των αντικειμένων κάθε ομάδας. Η διάταξη αυτή προσπαθεί να ωθήσει τα λιγότερο δημοφιλή αντικείμενα βαθύτερα στο παραγόμενο δένδρο. Το μειονέκτημά τους όμως είναι ότι βασίζονται στον αλγόριθμο κατασκευής του alphabetic δένδρου για να σχηματίσουν το τελικό δένδρο. Λαμβάνοντας υπόψη τις

αδυναμίες των προηγούμενων απλοϊκών ευριστικών προσεγγίσεων, υιοθετούμε μια περισσότερο δομημένη προσέγγιση. Ο σκοπός παραμένει ο ίδιος και είναι η ώθηση των λιγότερων δημοφιλών αντικειμένων κάθε ομάδας βαθύτερα στο παραγόμενο δένδρο αναζήτησης. Η δομημένη προσέγγιση περιλαμβάνει στη δημιουργία υποδένδρων, όπου κάθε ένα αντιστοιχεί σε μια ομάδα. Στην συνέχεια, κάθε υποδένδρο αντιμετωπίζεται ως ένας κόμβος με βάρος το κόστος του υποδένδρου. Τέλος, εφαρμόζουμε έναν αλγόριθμο κατασκευής alphabetic δένδρων.

Το κύριο στοιχείο αυτής της προσέγγισης είναι η επινόηση ενός αλγορίθμου κατασκευής υποδένδρων. Για το σκοπό αυτό προτείνουμε τρεις γενικές στρατηγικές:

- τοποθέτησε το δημοφιλέστερο αντικείμενο στη ρίζα του δένδρου και μετά συνέχισε ομοίως στα κλαδιά της ρίζας,
- διάλεξε τη ρίζα ώστε να εξισορροπήσει το συνολικό βάρος των κλαδιών, και
- κατασκεύασε ένα δένδρο Huffman με μεταβλητό παράγοντα διακλάδωσης για τα αντικείμενα της κάθε ομάδας.

Οι προσεγγίσεις αυτές ονομάζονται MostPop, EqWeig και POBI, αντίστοιχα. Οι πρώτες δύο στρατηγικές είναι απλές στη κατανόηση και επομένως μόνο η τρίτη θα εξηγηθεί με περισσότερη λεπτομέρεια.

Η τρίτη στρατηγική λειτουργεί με τον ίδιο τρόπο για όλες τις ομάδες. Ξεκινά από έναν αρχικό κόμβο, τον  $x$ , ο οποίος έχει παιδιά του όλα τα αντικείμενα της ομάδας, δηλαδή  $c_1, c_2, \dots, c_y$ . Αρχικά, ταξινομεί τα αντικείμενα της ομάδας κατά αύξουσα δημοφιλία. Στη συνέχεια, εάν βρεθεί ένα αντικείμενο  $z$ , τέτοιο ώστε  $(y - z - 1) \times \sum_{1 \leq j \leq z} Pr(c_j) > \sum_{z+1 \leq j \leq y} Pr(c_j)$ , τότε δημιουργεί ένα νέο κόμβο  $nx$  ως πατέρα των κόμβων  $c_{z+1}, \dots, c_y$  και θέτει το νέο κόμβο ως παιδί του κόμβου  $x$ . Αυτή η διαμέλιση εφαρμόζεται αναδρομικά τόσο στο  $x$  όσο και στο  $nx$ , μέχρις να μη μπορεί να γίνει κάποια διαμέλιση.

### 7.3.1 Κατεύθυνση αναζήτησης στα υποδένδρα κάθε ομάδας

Για να χαρακτηρίσουμε μία δομή ως δένδρο αναζήτησης πρέπει να θέσουμε ένα ξεχωριστό στοιχείο, π.χ., οι εσωτερικοί κόμβοι πρέπει να παρέχουν καθοδήγηση για τη συνέχιση της αναζήτησης. Οι δενδρικές δομές όπως το  $B^+$ -δένδρο [74, 144], και το alphabetic δένδρο [119] θέτουν αυτό το στοιχείο και συνεπώς χαρακτηρίζονται ως δένδρα αναζήτησης. Από την άλλη όμως, το VF δένδρο [31] δεν είναι ένα δένδρο αναζήτησης, επειδή σε μια χωρίς σειρά διάσχιση του δένδρου, τα αντικείμενα δεν ανακτώνται με τη σειρά ταξινόμησης (λεξικογραφική ή αριθμητική).

Επομένως, δεν μπορούμε να χρησιμοποιήσουμε τους εσωτερικούς κόμβους ενός τέτοιου δένδρου για να “κλαδέψουμε” τμήματα του δένδρου κατά τη διάρκεια της αναζήτησης.

Η δενδρική δομή που δημιουργήσαμε αντιμετωπίζει ένα παρόμοιο πρόβλημα, αν και σε πιο περιορισμένη έκταση, επειδή χρειάζεται να παρέχουμε καθοδήγηση μόνο για τα αντικείμενα κάθε ομάδας. Υποθέστε ότι αναζητούμε ένα αντικείμενο  $i$  που ανήκει στην ομάδα  $g$ . Είναι εύκολο να ανακαλύψουμε το υποδένδρο που αντιστοιχεί στην ομάδα  $g$ . Αυτό πραγματοποιείται με τη σύγκριση του αντικειμένου  $i$  με άλλα αντικείμενα που δεν ανήκουν στην ομάδα  $g$ , επειδή υπάρχει μια διάταξη μεταξύ του  $i$  και των άλλων αντικειμένων. Στη ρίζα του υποδένδρου που αντιστοιχεί στην ομάδα  $g$ , πρέπει να λάβουμε μια απόφαση σχετικά με το κλαδί που πρέπει να ακολουθήσουμε προκειμένου να κατευθυνθούμε στο αντικείμενο  $i$ . Κάθε κόμβος είναι εφοδιασμένος με ένα Bloom filter [14], το οποίο παρέχει αυτή την ευχέρεια. Συνεπώς, είμαστε σε θέση να προσπεράσουμε την έλλειψη διάταξης μεταξύ των αντικειμένων μιας ομάδας.

## 7.4 Πειραματική Αποτίμηση

Για να αποτιμήσουμε την απόδοση των προτεινόμενων αλγορίθμων, αναπτύξαμε ένα σύστημα προσομοίωσης ενός περιβάλλοντος, όπου ένας αριθμός από κόμβους (πελάτες) προσπελαύνουν τα δεδομένα που παρέχονται από τον εξυπηρετητή μέσω του ασύρματου καναλιού εκπομπής. Κάθε κόμβος περιέχει μια cache για να αποθηκεύει τα δεδομένα που έχει προσπελάσει και να επιλέγει τα επόμενα προς προσπέλαση δεδομένα με βάση το προφίλ του. Κάθε φορά που ένας κόμβος χρειάζεται ένα αντικείμενο που δεν είναι αποθηκευμένο τοπικά, τότε συντονίζεται με το κανάλι εκπομπής ώστε να διαβάσει την πληροφορία καταλόγου και να εναλλάσσεται μεταξύ της ενεργούς κατάστασης και της κατάστασης ύπνου. Αυτό συνεχίζεται μέχρι ο κόμβος να λάβει την απαιτούμενη πληροφορία.

Επιπρόσθετα, υλοποιήσαμε ένα μοντέλο για την κατανομή αιτήσεων κόμβου, για κατανομή των μεγεθών των ομάδων, για κατανομή της δημοφιλίας αντικειμένου και για κατανομή της δημοφιλίας ομάδας. Όλα τα ανωτέρω βασίστηκαν στον Zipfian κανόνα, π.χ.,

$$p_x = \frac{(1/x)^\theta}{\sum_{x=1}^y (1/x)^\theta} \quad 1 \leq x \leq n \text{ (or } 1 \leq x \leq m) \quad (7.3)$$

όπου η παράμετρος  $\theta$  ελέγχει την κυρτότητα της κατανομής. Για  $\theta=0$  η Zipfian κατανομή μεταβάλλεται σε ομοιόμορφη κατανομή. Αντίθετα, μεγαλύτερες τιμές



του  $\theta$  έχουν ως αποτέλεσμα περισσότερο κυρτές κατανομές.

Ο σχηματισμός των ομάδων και οι δημοτικότητες των αντικειμένων δημιουργούνται ως εξής. Υποθέστε ότι έχουμε αποφασίσει να δημιουργήσουμε  $n$  αντικείμενα που ανατίθενται σε  $m$  ομάδες. Αρχικά, αποφασίζουμε τα σχετικά μεγέθη των ομάδων χρησιμοποιώντας την κατανομή Zipf (Εξίσωση 7.3). Όταν θέτουμε μικρές τιμές στο  $\theta$ , τότε όλες οι ομάδες έχουν παρόμοια μεγέθη, ενώ για μεγαλύτερες τιμές του  $\theta$  κάποιες ομάδες έχουν σημαντικά μεγαλύτερο αριθμό αντικειμένων. Στη συνέχεια, αναθέτουμε σε κάθε ομάδα μια σχετική δημοτικότητα χρησιμοποιώντας πάλι το κανόνα Zipf. Ανάλογα με το συνδυασμό, μπορούμε να δημιουργήσουμε μεγάλες ομάδες με μεγάλη δημοτικότητα, μικρές ομάδες με μεγάλη δημοτικότητα, ίσου μεγέθους ομάδες με σχεδόν ίδια δημοτικότητα κ.τ.λ. Έπειτα, σε μια τρίτη κλήση του κανόνα Zipf η δημοτικότητα κατανέμεται στα αντικείμενα της ομάδας. Η προτίμηση των χρηστών για αντικείμενα αποφασίζεται επίσης με το κανόνα Zipf.

Πραγματοποιήσαμε ένα μεγάλο αριθμό πειραμάτων για την αποτίμηση της επίδρασης του όγκου των δεδομένων, της κατανομής προσπέλασης των δεδομένων, της cache του πελάτη, της χρήσης του φίλτρου Bloom filters, κ.τ.λ. Στο παρόν κεφάλαιο θα παρουσιάσουμε μόνο την αναλυτική αποτίμηση των τριών προταθέντων δομών, όπως αυτές ορίστηκαν από την Εξίσωση 7.1. Πιο συγκεκριμένα, θα παρουσιάσουμε την απόδοση των αλγορίθμων με μεταβλητό αριθμό αντικειμένων, με μεταβλητό αριθμό ομάδων, με μεταβλητή κατανομή δημοτικότητας των ομάδων και με διαφορετικά σχετικά μεγέθη των ομάδων. Στον Πίνακα 7.1 παρουσιάζουμε τα κύρια σύμβολα που χρησιμοποιήθηκαν στην προσομοίωση μαζί με την περιγραφή τους και την προεπιλεγμένη τιμή τους. Η σημειογραφία  $nAg-BgsCgpD$  χρησιμοποιείται στον τίτλο των γραφημάτων για να συμβολίσει τις παραμέτρους όπου κάθε ένα από τα σύμβολα  $A$ ,  $B$ ,  $C$  ή  $D$  αντικαθίσταται είτε από μια αριθμητική τιμή ή από μία υπογράμμιση για να τονίσει ότι μεταβάλλεται. Για παράδειγμα, ο συμβολισμός  $n500g10gs01gp$  σημαίνει ότι οι διερευνούμενες τιμές για τη σχετική δημοτικότητα ομάδας για μια συλλογή 500 αντικειμένων η οποία διαιρείται σε 10 ομάδες, όπου η σχετική κυρτότητα στα μεγέθη ομάδας είναι 0.1.

Το κεφάλαιο εισάγει για πρώτη φορά το πρόβλημα του καταλόγου σε μερικώς διατεταγμένα αντικείμενα για ασύρματες εκπομπές και συνεπώς δεν υπάρχουν ανταγωνιστικές μέθοδοι στη βιβλιογραφία για να συγκριθούν με τη νέα μέθοδο. Το ρόλο των ανταγωνιστικών μεθόδων θα διαδραματίσουν κάποιες επεκτάσεις υπάρχοντων δενδρικών δομών που επινοήσαμε. Αναπτύξαμε μια επέκταση το VF δένδρου [31], χρησιμοποιώντας το ίδιο όνομα, για να λάβουμε υπόψη τη μερική διάταξη των αντικειμένων. Επίσης, υλοποιήσαμε μια επέκταση του πρωτότυπου

Παράμετρος	Συμβολισμός	Εξ ορισμού
$n$	πλήθος αντικειμένων	500
$g$	πλήθος ομάδων	10
$gs$	κυρτότητα για τη μέγεθος των ομάδων	0.1 (0.0 για ισομεγέθεις ομάδες)
$gp$	κυρτότητα για τη δημοφιλία των ομάδων	0.1 (0.0 για ισο-δημοφιλείς ομάδες)

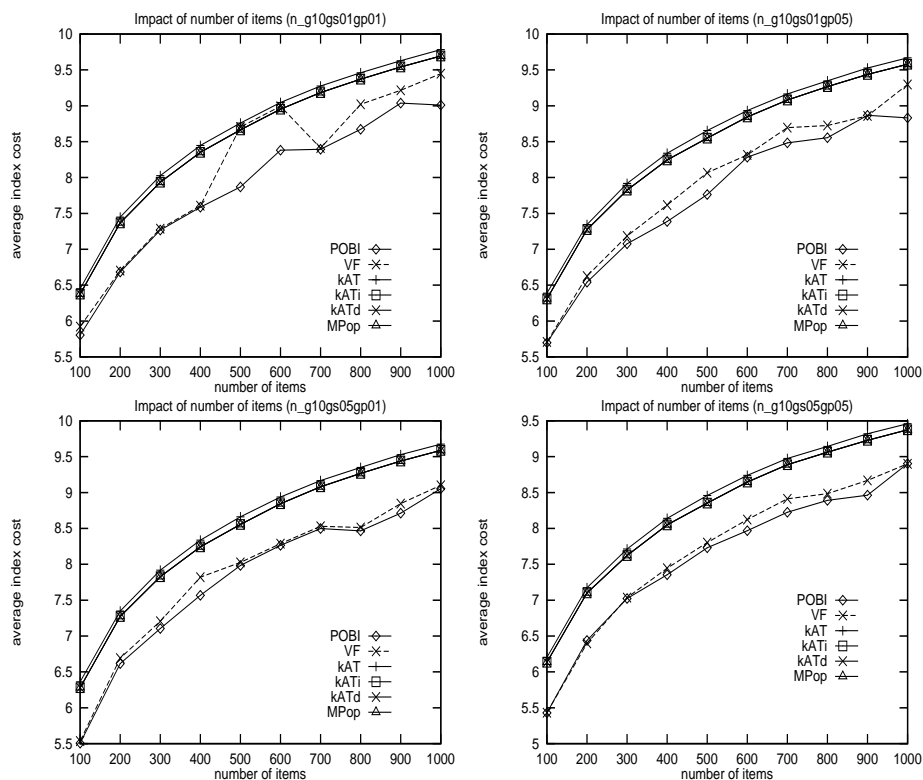
Πίνακας 7.1. Παράμετροι προσομοίωσης.

$k$ -ary Alphabetic δένδρου, χρησιμοποιώντας μία τυχαία διάταξη των αντικειμένων μέσα σε κάθε ομάδα και συμβολίζουμε τη διαφορετική εκδοχή ως kAT. Τέλος, υλοποιήσαμε τα προσεγγιστικά μας σχήματα που ονομάζονται kATi, kATd, MostPop, EqWeig και POBI. Στα γραφήματα των ακόλουθων παραγράφων εξαιρούμε από την απεικόνιση τη μέθοδο EqWeig, επειδή εμφανίζει την ίδια απόδοση με τη μέθοδο MostPop.

#### 7.4.1 Επίδραση του αριθμού των αντικειμένων

Το πρώτο πείραμα σκοπεύει να διερευνήσει την επίδραση του αριθμού των αντικειμένων στην απόδοση των αλγορίθμων. Μεταβάλλουμε τον αριθμό των αντικειμένων από 100 μέχρι 1000 και διερευνούμε το μέσο κόστος καταλόγου που προκαλείται από τους αλγορίθμους σε δύο διαστάσεις, δηλαδή μέγεθος ομάδας και δημοτικότητα ομάδας, θεωρώντας ίσα και αρκετά κυρτά τα μεγέθη των ομάδων και επίσης θεωρώντας ίσες και αρκετά κυρτές δημοτικότητες. Έτσι σχηματίζονται τέσσερις συνδυασμοί. Τα παραγόμενα γραφήματα απεικονίζονται στο Σχήμα 7.2.

Το κύριο στοιχείο που παρατηρούμε είναι η αύξηση του μέσου κόστους καταλόγου όταν αυξάνεται ο αριθμός των αντικειμένων. Αυτό οφείλεται στο γεγονός ότι όσο μεγαλύτερος είναι ο αριθμός των αντικειμένων, τόσο βαθύτερο είναι το παραγόμενο δένδρο. Το μέσο κόστος καταλόγου αυξάνεται λογαριθμικά επειδή το παραγόμενο δένδρο είναι αρκετά θαμνώδες και όχι ένα εκφυλισμένο αριστερό (ή δεξιό) βαθύ δένδρο. Αξιολογώντας την απόδοση κάθε αλγορίθμου, η πρώτη παρατήρηση είναι ότι η μέθοδος kAT είναι η χειρότερη από όλες, επειδή δεν πραγματοποιεί έξυπνες αποφάσεις στην κατασκευή του δένδρου αναζήτησης. Η απόδοση των μεθόδων που βασίζονται σε κάποια ταξινόμηση των αντικειμένων της κάθε ομάδας πριν τη κατασκευή του τελικού alphabetic δένδρου (π.χ., kATi,



Σχήμα 7.2. Επίδραση του αριθμού των αντικειμένων στην απόδοση του δένδρου.

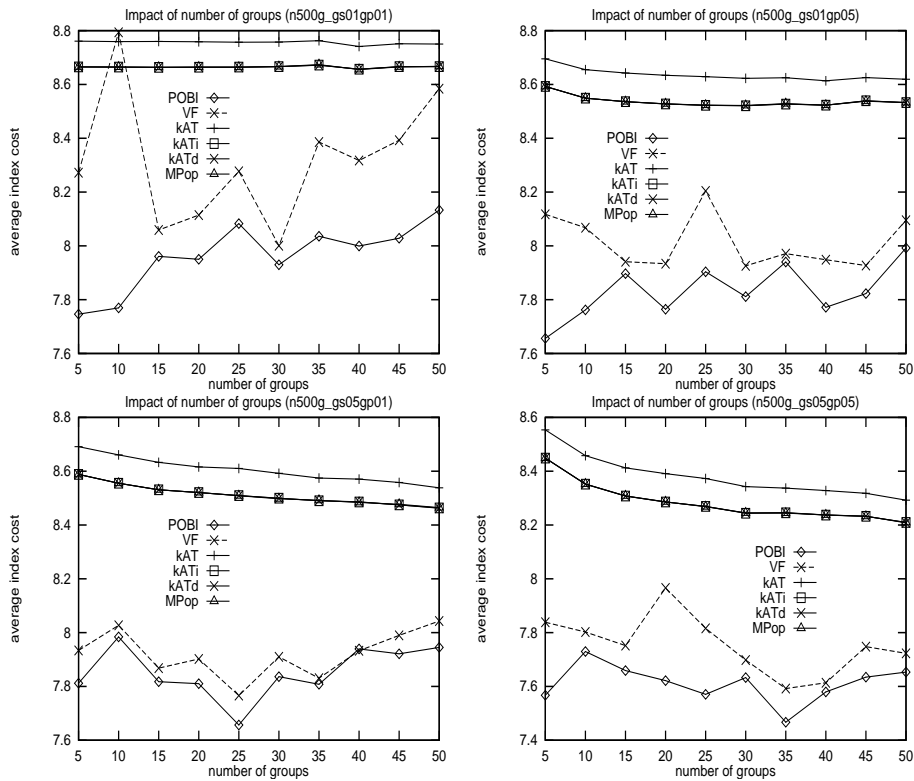
kATd, MostPop), είναι σχεδόν ισοδύναμη αλλά σημαντικά καλύτερη από αυτή της μεθόδου kAT.

Οι περισσότερο αποδοτικοί αλγόριθμοι είναι ο VF και ο POBI, με τον τελευταίο να είναι καλύτερος σε απόδοση σε όλες τις περιπτώσεις. Είναι ενδιαφέρον να τονίσουμε ότι αγνοώντας κάποια στατιστικά λάθη, η απόδοση του VF επιδεινώνεται συγκριτικά με αυτή του POBI με αύξοντα αριθμό αντικειμένων. Αυτό συμβαίνει λόγω των καλύτερων αποφάσεων που λαμβάνονται από τον αλγόριθμο POBI που αξιοποιεί τις συσχετίσεις των αντικειμένων μεταξύ των ομάδων.

#### 7.4.2 Επίδραση του αριθμού των ομάδων

Το δεύτερο πείραμα στοχεύει στη διερεύνηση της επίδρασης του αριθμού των ομάδων στην απόδοση των αλγορίθμων. Μεταβάλλοντας τον αριθμό των ομάδων από 10 μέχρι 50 και πάλι ερευνούμε το μέσο κόστος καταλόγου για δύο διαστάσεις, δηλαδή το συγκριτικό μέγεθος ομάδας και τη συγκριτική δημοτικότητα ομάδας,

θεωρώντας ίσα και αρκετά κυρτά μεγέθη για τις ομάδες και θεωρώντας επίσης ίσες και αρκετά κυρτές δημοτικότητες. Συνεπώς εξετάζουμε τέσσερις συνδυασμούς. Τα παραγόμενα γραφήματα απεικονίζονται στο Σχήμα 7.3.



Σχήμα 7.3. Επίδραση του αριθμού των ομάδων στην απόδοση του δένδρου.

Η κύρια τάση που παρατηρούμε στα αποτελέσματα δεν είναι όμοια και για τα τέσσερα γραφήματα. Αυτά του επάνω μέρος του σχήματος παριστούν την περίπτωση όπου οι ομάδες είναι ισομεγέθεις. Παρατηρούμε ότι όταν αυξάνεται ο αριθμός των ομάδων, τότε έχουμε μία σταδιακή αύξηση του μέσου κόστους καταλόγου. Αυτό οφείλεται στο γεγονός ότι όσο μεγαλύτερος είναι ο αριθμός των αντικειμένων, τόσο βαθύτερο είναι το παραγόμενο δένδρο. Στην περίπτωση όμως των κάτω γραφημάτων, παρατηρούμε ότι το μέσο κόστος καταλόγου είναι συγκριτικά σταθερό με την αύξηση του αριθμού των ομάδων. Η συμπεριφορά αυτή οφείλεται στο γεγονός ότι η κυρτότητα στα μεγέθη των ομάδων έχει ως αποτέλεσμα ένα μικρό μέρος των ομάδων να συγκεντρώνει το μεγαλύτερο πλήθος. Συνεπώς, οι λίγες αυτές ομάδες είναι υπεύθυνες για το μεγαλύτερο μέρος του κόστους καταλόγου, το οποίο παραμένει ανεπηρέαστο από το μεγάλο αριθμό

ομάδων.

Τα σχόλια για την απόδοση των αλγορίθμων kAT, kATi, kATd και MostPop είναι ανάλογα με αυτά που περιγράφηκαν, όταν εξετάσαμε την επίδραση του αριθμού των αντικειμένων. Το ενδιαφέρον εδώ βρίσκεται στη συγκριτική απόδοση των αλγορίθμων VF και POBI. Η γενική διαπίστωση είναι ότι ο VF είναι λιγότερο αποτελεσματικός από τον POBI, όταν οι ομάδες είναι ίσου μεγέθους και επιτυγχάνει να κλείσει την υστέρηση στην απόδοση όταν η κυρτότητα στα μεγέθη των ομάδων ή η κυρτότητα στη δημοτικότητα των ομάδων γίνεται σημαντική. Η απόδοση του αλγορίθμου POBI δεν επηρεάζεται από τον αριθμό των ομάδων εκτός από την περίπτωση των ίσου μεγέθους και ίσης δημοτικότητας ομάδων, όπου παρουσιάζει μια μικρή αύξηση στο μέσο κόστος καταλόγου.

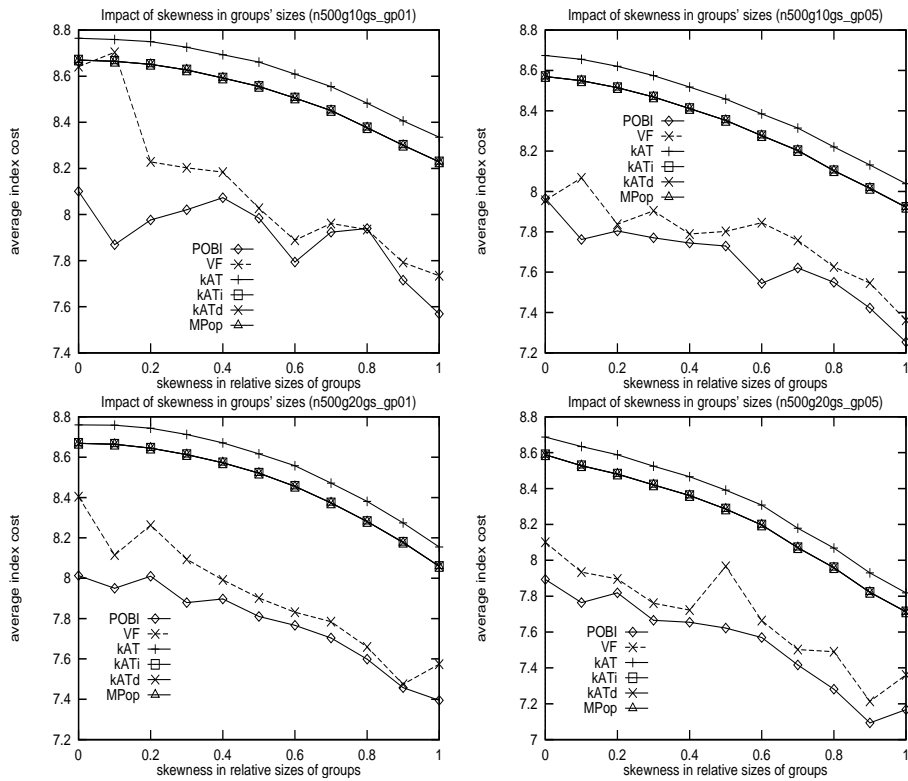
### 7.4.3 Επίδραση του μεγέθους της ομάδας

Το τρίτο πείραμα στοχεύει στη διερεύνηση της επίδρασης των μεγεθών των ομάδων στην απόδοση των αλγορίθμων. Μεταβάλλουμε την κυρτότητα στα μεγέθη των ομάδων από την τιμή 0.0 μέχρι την τιμή 1.0 και εξετάζουμε το μέσο κόστος καταλόγου, το οποίο προκαλείται από τους αλγορίθμους ως προς δύο διαστάσεις, δηλαδή τον αριθμό των ομάδων και τη συγκριτική δημοτικότητα ομάδας. Θεωρούμε την ύπαρξη 10 και 20 ομάδων και ίσες και αρκετά κυρτές δημοτικότητες. Συνεπώς, υπάρχουν τέσσερις συνδυασμοί. Τα αποτελέσματα απεικονίζονται στο Σχήμα 7.4.

Η γενική τάση που παρατηρείται στα αποτελέσματα είναι η μείωση του μέσου κόστους καταλόγου όταν αυξάνεται η κυρτότητα στη δημοτικότητα μιας ομάδας. Το συμπέρασμα αυτό είναι αναμενόμενο καθώς το μεγαλύτερο μέρος της δημοτικότητας συγκεντρώνεται σε λιγότερες ομάδες και συνεπώς αντικείμενα. Τα αντικείμενα τοποθετούνται επιτυχώς στα υψηλότερα τμήματα (κοντά στη ρίζα) των παραγόμενων δένδρων αναζήτησης. Επιπλέον, είναι φανερό ότι ο κατάλογος POBI είναι καλύτερος από το VF όταν δεν υπάρχει εξαιρετικά υψηλή κυρτότητα στη δημοτικότητα ομάδας, αλλά η κυρτότητα είναι είτε μέτρια είτε μικρή. Ο λόγος που συμβαίνει αυτό έχει επαρκώς εξηγηθεί στη Παράγραφο 7.4.2. Σε κάθε περίπτωση όμως, αποδίδει καλύτερα από το VF δένδρο.

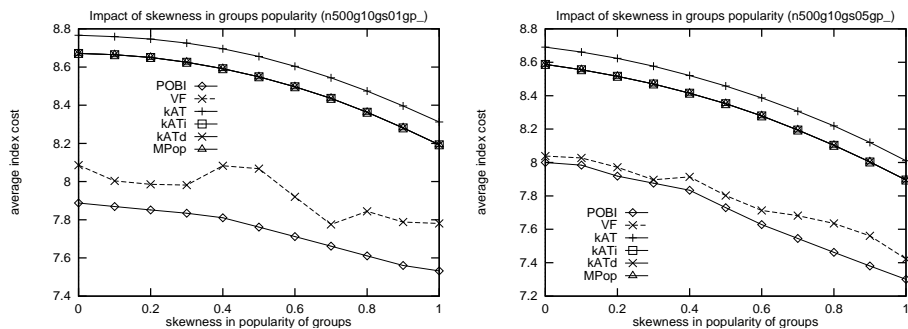
### 7.4.4 Επίδραση της δημοτικότητας της ομάδας

Τέλος, το τελευταίο πείραμα στοχεύει στη διερεύνηση της επίδρασης της συγκριτικής δημοτικότητας ομάδας στην απόδοση των μεθόδων. Τα συμπεράσματα που



Σχήμα 7.4. Επίδραση του μεγέθους της ομάδας στην απόδοση του δένδρου.

εξάγονται για όλες τις μεθόδους είναι σύμφωνα με όσα παρατηρήθηκαν στα προηγούμενα. Ο κατάλογος POBI αποκαλύπτει ακόμα μία χρησιμότητά του. Στην περίπτωση που υπάρχει μεγάλη κυρτότητα στη δημοτικότητα ομάδας, τότε είναι σε θέση να μειώσει το μέσο κόστος καταλόγου σε μεγαλύτερο βαθμό από ότι



Σχήμα 7.5. Επίδραση της δημοτικότητας της ομάδας.

συμβαίνει για μέτρια ή μικρή κυρτότητα. Αυτό είναι εμφανές από τη μεγαλύτερη κλίση που εμφανίζεται στη καμπύλη του POBI για τιμές κυρτότητας μεγαλύτερες του 0.5.

## 7.5 Συμπεράσματα

Στο κεφάλαιο αυτό εξετάσαμε για πρώτη φορά τη δημιουργία καταλόγων για εκπεμπόμενες πληροφορίες υπό το πρίσμα των μερικώς διατεταγμένων δεδομένων. Το κίνητρο για τη διερεύνηση της περίπτωσης των μερικώς διατεταγμένων δεδομένων προέρχεται από τις νέες αναπτυσσόμενες εφαρμογές που προορίζονται για AΔΑ και MANET. Ωστόσο, και οι συνηθισμένες εφαρμογές όπως αυτές που αναπτύχθηκαν για τα παραδοσιακά ασύρματα κυψελοειδή δίκτυα θα μπορούσαν να ωφεληθούν από τα θέματα που περιγράφονται σε αυτό το κεφάλαιο.

Αρχικά ορίσαμε το πρόβλημα κατασκευής καταλόγων για μερικώς διατεταγμένα δεδομένα. Στην συνέχεια παρουσιάσαμε ότι το πρόβλημα γενικεύει δύο προβλήματα που προτάθηκαν νωρίτερα στη βιβλιογραφία, τα οποία οδήγησαν στην ανάπτυξη δύο υψηλών αποδόσεων καταλόγων. Επιπρόσθετα, παρέχουμε προσεγγιστικούς αλγόριθμους για τη δημιουργία των δένδρων αναζήτησης εκπομπής, οι οποίοι επιλύουν το πρόβλημα που τέθηκε. Η σχεδίαση της δομής αναζήτησης λαμβάνει υπόψη την κυρτότητα στην κατανομή προσπέλασης των δεδομένων και συνεπώς προτείνουμε μια μη ισορροπημένη δομή για να δώσουμε λύση στην κυρτότητα που εμφανίζει η προσπέλαση των δεδομένων.

Για να αποτιμήσουμε την απόδοση των προτεινόμενων αλγορίθμων, υλοποιήσαμε ένα σύστημα που προσομοιώνει ένα περιβάλλον όπου ένας αριθμός από κόμβους (πελάτες) προσπελούν τα δεδομένα που παρέχονται από τον εξυπηρετητή μέσω του ασύρματου καναλιού εκπομπής. Με τη χρήση του περιβάλλοντος προσομοίωσης διερευνήσαμε την απόδοση των προτεινόμενων προσεγγιστικών αλγορίθμων και παρουσιάσαμε λεπτομερή πειράματα που ανέδειξαν την ανωτερότητα της μίας από τις μεθόδους, η οποία ονομάζεται POBI δενδρικός κατάλογος αναζήτησης. Ο αλγόριθμος POBI αποδείχθηκε ότι σαφώς παρουσιάζει την καλύτερη απόδοση συγκριτικά με όλες τις υπόλοιπες ανταγωνιστικές μεθόδους.





## ΚΕΦΑΛΑΙΟ 8

# Επίλογος

### Περιεχόμενα

---

8.1	ΣΥΜΠΕΡΑΣΜΑΤΑ . . . . .	181
8.2	ΔΡΟΜΟΙ ΜΕΛΛΟΝΤΙΚΗΣ ΕΡΕΥΝΑΣ . . . . .	184

---

### 8.1 Συμπεράσματα

Τα ΑΔΑ προκάλεσαν μεγάλο ενδιαφέρον στην ερευνητική κοινότητα και τη βιομηχανία κατά τη διάρκεια της τελευταίας δεκαετίας εξαιτίας των επιτευγμάτων στα χαμηλής κατανάλωσης κυκλώματα και την ανάπτυξη κατάλληλου λογισμικού. Καθώς διαρκώς αυξάνονται σε πλήθος οι εφαρμογές των ΑΔΑ και εξίσου μεγαλώνει ο αριθμός των αισθητήρων που επιθυμούν να προσπελάσουν τις πληροφορίες που παράγονται σε άλλους αισθητήρες, αυξάνεται σημαντικά η καθυστέρηση πρόσβασης στα προσφερόμενα δεδομένα και η κατανάλωση της ενέργειας κάθε αισθητήρα.

Στην παρούσα διατριβή, αναζητήσαμε τεχνικές για την ελάττωση της παρατηρούμενης καθυστέρησης και την ταυτόχρονη διατήρηση της ενέργειας των αισθητήρων. Εστίασαμε στην τεχνική που στηρίζεται στη αναζήτηση και ανάκτηση των δεδομένων από τη πηγή δημιουργίας τους προς τους καταναλωτές τους. Μεταξύ άλλων η τεχνική αυτή περιλαμβάνει τις μεθόδους του συνεργατικού caching, της ομαδοποίησης και των εναέριων καταλόγων, οι οποίες αποτέλεσαν αντικείμενο της διατριβής. Ενώ, οι τεχνικές αυτές έχουν μελετηθεί στο παρελθόν για διάφορες κατηγορίες των συστημάτων υπολογιστών όπως είναι οι Βάσεις Δεδομένων,

τα Λειτουργικά Συστήματα, ο Παγκόσμιος Ιστός, τα ασύρματα κυψελοειδή δίκτυα και τα MANET, το ασύρματο περιβάλλον των Δικτύων Αισθητήρων σε συνδυασμό με τους περιορισμένους πόρους των αισθητήρων παρουσιάζει μοναδικές προκλήσεις για τους αλγόριθμους που εμπίπτουν στις κατηγορίες αυτές.

Το μεγάλο πλήθος των αισθητήρων που αποτελούν ένα ασύρματο δίκτυο, οι περιορισμένες δυνατότητες τους σε ότι αφορά στην ενέργεια, την υπολογιστική ισχύ, την αποθηκευτική χωρητικότητα και την ακτίνα μετάδοσης, η multi-hop φύση των ΑΔΑ που λειτουργούν με βάση την τεχνική αποθήκευση-και-προώθηση εξαιτίας της απουσίας των σταθμών βάσης, η μη γνώση της γεωγραφικής τοποθεσίας των αισθητήρων, το περιορισμένο εύρος μετάδοσής τους, ο μεγάλος όγκος των δεδομένων και οι σημαντικές καθυστερήσεις πρόσβασης είναι οι σημαντικότεροι παράγοντες που καθιστούν ανεπαρκείς τις λύσεις που έχουν προταθεί μέχρι τώρα στο πλαίσιο των τεχνικών διαχείρισης των δεδομένων.

Στο Κεφάλαιο 3 ερευνήσαμε το πρόβλημα του συνεργατικού caching στην ερευνητική περιοχή των ΑΠΔΑ. Τα ΑΠΔΑ προσφέρουν μία νέα ώθηση και εντυπωσιακές δυνατότητες στις υπάρχουσες εφαρμογές των ΑΔΑ με αποτέλεσμα να δημιουργούνται πολλές νέες εφαρμογές. Λαμβάνοντας υπόψη τα μοναδικά χαρακτηριστικά των ΑΠΔΑ και την ανάγκη για Ποιότητα Υπηρεσίας σε επίπεδο εφαρμογής, ένα θέμα που είχε αγνοηθεί στα παραδοσιακά ΑΔΑ, αναπτύξαμε ένα πρωτόκολλο συνεργατικού caching, το οποίο ονομάζεται NiCoCa και είναι κατάλληλο για ΑΠΔΑ. Στα πλαίσια του συνεργατικού caching, πολλαπλοί κόμβοι αισθητήρες διαμοιράζονται τα αποθηκευμένα δεδομένα και συνεργάζονται ώστε να μειωθεί το κόστος επικοινωνίας, η κατανάλωση της ενέργειας, η καθυστέρηση πρόσβασης και, τελικά, να αξιοποιηθεί βέλτιστα ο αποθηκευτικός χώρος της cache κάθε συνεργαζόμενου αισθητήρα. Σχεδιάσαμε το πρωτόκολλο NiCoCa έτσι ώστε να ανιχνεύει τους κόμβους που βρίσκονται στα κεντρικότερα σημεία των γειτονιών του δικτύου και να τους αναθέτει το ρόλο του διαμεσολαβητή για να συντονίσουν τις αποφάσεις caching. Παράλληλα παρουσιάσαμε μια νέα πολιτική αντικατάστασης της cache, η οποία επιλέγει δυναμικά το αντικείμενο προς αντικατάσταση. Τα καινοτόμα στοιχεία του πρωτοκόλλου ήταν η επιλογή των κόμβων διαμεσολαβητών και η πολιτική αντικατάστασης της cache. Δείξαμε ότι το πρωτόκολλο είναι εξαιρετικά αποδοτικό επιτυγχάνοντας να μειώσει τα global hits κατά ένα μέσο ποσοστό 50% και να αυξήσει τα remote hits εξαιτίας της αποδοτικής συνεργασίας των αισθητήρων κατά ένα μέσο ποσοστό 40%.

Στο Κεφάλαιο 4 μελετήσαμε το πρόβλημα της συνέπειας της cache σε ΑΠΔΑ. Εξαιτίας των περιορισμένων πόρων των αισθητήρων, του εύρους μετάδοσης αλλά κυρίως της ανάγκης για ενημερωμένα και έγκυρα δεδομένα σχεδιάσαμε μια push-

pull πολιτική, η οποία σκοπό έχει τη μείωση της καθυστέρησης πρόσβασης των δεδομένων, τη μείωση στην κατανάλωση ενέργειας και τη διασφάλιση της εγκυρότητας των δεδομένων. Προτείναμε το πρωτόκολλο NICC που αποδείχτηκε καλύτερο από την ανταγωνιστική μέθοδο σχετικά με τον αριθμό των hits (local, remote, global), την καθυστέρηση πρόσβασης, τον αριθμό των διαδιδόμενων μηνυμάτων και την καταναλισκόμενη ενέργεια από τους αισθητήρες. Παράλληλα, στα πλαίσια του πρωτοκόλλου NICC προτείναμε μια νέα πολιτική αντικατάστασης της cache, η οποία λαμβάνει υπόψη το ρυθμό προσπέλασης και ενημέρωσης κάθε αντικειμένου, το μέγεθος του αντικειμένου και την απόσταση των κόμβων που ζητούν το αντικείμενο. Η νέα πολιτική αντικατάστασης αποδείχθηκε ότι υπερτερεί της ανταγωνιστικής πολιτικής αντικατάστασης και συμβάλλει στη βελτίωση της απόδοσης του πρωτοκόλλου NICC.

Στο Κεφάλαιο 5 μελετήσαμε το ζήτημα του συνεργατικού caching σε ΑΔΑ. Κίνητρό μας υπήρξε η βελτίωση της επιλογής των κόμβων που θα πραγματοποιήσουν το συντονισμό και τη συνεργασία μεταξύ των αισθητήρων και η ανάπτυξη μιας νέας πολιτικής ελέγχου εισδοχής στην cache. Προτείναμε δύο νέες μετρικές που μπορούν να χρησιμεύσουν στην επιλογή των σημαντικών κόμβων αισθητήρων. Σκοπός των μετρικών είναι η ανεύρεση των κεντρικών γειτονικών κόμβων. Η πρώτη μετρική *PCI* αξιοποιεί την πληροφορία που σχετίζεται με την 1-hop γειτονία ενός κόμβου, ενώ η δεύτερη μετρική *extPCI* αποτελεί βελτίωση της πρώτης και ο υπολογισμός της βασίζεται στη γνώση των συνδέσεων μεταξύ των κόμβων της 2-hop γειτονιάς. Χρησιμοποιώντας τις δύο μετρικές δημιουργήσαμε δύο πρωτόκολλα συνεργατικού caching το PCICC και το extPCICC, τα οποία μέσω πειραματικών προσομοιώσεων αποδείχθηκαν ανώτερα από το ανταγωνιστικό πρωτόκολλο.

Στο Κεφάλαιο 6 ασχοληθήκαμε με το ζήτημα της ομαδοποίησης σε ΑΔΑ, η οποία επιτρέπει τη συνάθροιση της πληροφορίας, περιορίζει τη μετάδοση των δεδομένων κυρίως μέσα στη ομάδα, μειώνει το συναγωνισμό για την απόκτηση του ασύρματου καναλιού και ελαττώνει την κατανάλωση ενέργειας. Προτείναμε ένα ενεργειακά αποδοτικό κατανεμημένο πρωτόκολλο ομαδοποίησης για ΑΔΑ, το οποίο ονομάζεται GESC. Το πρωτόκολλο GESC λαμβάνει υπόψη τόσο την κατανάλωση ενέργειας, όσο και τα τοπολογικά στοιχεία των κόμβων σε αντίθεση με τα πρωτόκολλα που είχαν προταθεί στο παρελθόν, τα οποία λάμβαναν υπόψη μόνο τον ένα από τους δύο παράγοντες. Το προτεινόμενο πρωτόκολλο αξιοποιεί την τοπική δομή του δικτύου και την εναπομείνασα ενέργεια των γειτονικών κόμβων για να ορίσει ένα νέο τρόπο δυναμικού υπολογισμού των κεφαλών ομάδων. Το πρωτόκολλο αποδείχθηκε ότι είναι πολύ αποδοτικό και συμβάλλει

τόσο στη μείωση των μηνυμάτων που διαδίδονται στο δίκτυο, όσο και στη μείωση της διασπάθισης ενέργειας.

Στο Κεφάλαιο 7 ερευνήσαμε για πρώτη φορά το ζήτημα της δημιουργίας καταλόγων για εκπεμπόμενες πληροφορίες υπό το πρίσμα των μερικώς διατεταγμένων δεδομένων. Το κίνητρο για τη διερεύνηση της περίπτωσης των μερικώς διατεταγμένων δεδομένων προέρχεται από τις νέες αναπτυσσόμενες εφαρμογές που προορίζονται για ΑΔΑ και MANET. Αρχικά ορίσαμε το πρόβλημα κατασκευής καταλόγων για μερικώς διατεταγμένα δεδομένα. Στην συνέχεια παρουσιάσαμε ότι το πρόβλημα γενικεύει δύο προβλήματα που προτάθηκαν νωρίτερα στη βιβλιογραφία, τα οποία οδήγησαν στην ανάπτυξη δύο υψηλών αποδόσεων δομών καταλόγων. Επιπρόσθετα, παρέχουμε προσεγγιστικούς αλγορίθμους για τη δημιουργία των δένδρων αναζήτησης εκπομπής, οι οποίοι επιλύουν το πρόβλημα που τέθηκε. Η σχεδίαση της δομής αναζήτησης λαμβάνει υπόψη την κυρτότητα στην κατανομή προσπέλασης των δεδομένων και συνεπώς προτείναμε μια μη ισορροπημένη δομή ως απάντηση στην κυρτότητα που εμφανίζει η προσπέλαση των δεδομένων. Διερευνήσαμε την απόδοση των προτεινόμενων προσεγγιστικών αλγορίθμων και η μέθοδος *POBI* σαφώς παρουσιάζει την καλύτερη απόδοση συγκριτικά με όλες τις υπόλοιπες ανταγωνιστικές μεθόδους.

## 8.2 Δρόμοι Μελλοντικής Έρευνας

Στην ενότητα αυτή παρουσιάζουμε σύντομα μερικές περιοχές και ιδέες για μελλοντική έρευνα.

Στο Κεφάλαιο 3 παρουσιάσαμε το πρωτόκολλο NiCoCa και αξιολογήσαμε την επίδοσή του με την προϋπόθεση ότι οι κόμβοι αισθητήρες είναι στατικοί, ενώ το πρωτόκολλο δρομολόγησης που χρησιμοποιήθηκε είναι το AODV. Ως συνέχεια της αξιολόγησής του επιθυμούμε να υλοποιήσουμε κάποια μοντέλα κίνησης και να χρησιμοποιήσουμε το πρωτόκολλο δρομολόγησης GPSR, έτσι ώστε να μελετήσουμε την απόδοσή του. Επιπρόσθετα, σκοπεύουμε να μελετήσουμε την απόδοσή του σε Δίκτυα Ανεκτικά στην Καθυστέρηση (Delay Tolerant Networks), καθώς η τεχνική του συνεργατικού caching είναι μια πολύ σημαντική μέθοδος, ώστε να αντιμετωπισθούν τα προβλήματα συνδεσιμότητας μεταξύ των κόμβων και να γίνει αποτελεσματικότερη η δρομολόγηση των δεδομένων.

Στο Κεφάλαιο 4 ασχοληθήκαμε μόνο με το ζήτημα της συνέπειας της cache, σε αντίθεση με το Κεφάλαιο 3 όπου είχαμε ασχοληθεί με το συνεργατικό caching και είχαμε θεωρήσει ότι τα δεδομένα δεν ενημερώνονται. Συνεπώς, η συνέχεια

της αξιολόγησης του πρωτοκόλλου NICC που προτείναμε είναι η ενσωμάτωσή του στο πρωτόκολλο NiCoCa και η πειρατική αποτίμησή του για μεγαλύτερες τοπολογίες δικτύου. Παράλληλα σκοπεύουμε να αξιολογήσουμε το πρωτόκολλο NICC για ομοιόμορφη κατανομή στην προσπέλαση των δεδομένων.

Στο Κεφάλαιο 5 παρουσιάσαμε τους αλγόριθμους PCICC και extPCICC και δείξαμε ότι η μετρική *PCI* όπου βασίζεται ο αλγόριθμος PCICC έχει μικρή πολυπλοκότητα και μπορεί να υπολογισθεί πολύ γρήγορα. Οι δύο αλγόριθμοι αποτιμήθηκαν υποθέτοντας ότι οι αισθητήρες παραμένουν ακίνητοι. Είναι ενδιαφέρον να εξετάσουμε και να συγκρίνουμε τις επιδόσεις των δύο αλγορίθμων για διάφορα μοντέλα κίνησης στο πλαίσιο των MANET. Επιπλέον, σκοπεύουμε να εξετάσουμε τους δύο αλγόριθμους και κυρίως τον PCICC σε ad hoc δίκτυα οχημάτων (Vehicular Ad Hoc Networks). Τα ad hoc δίκτυα οχημάτων περιλαμβάνουν κίνηση, ενώ η τοπολογία του δικτύου μεταβάλλεται διαρκώς. Η επιλογή των σημαντικών κόμβων που θα συντονίζουν το caching θα πρέπει να πραγματοποιείται γρήγορα με χαμηλό κόστος, κάτι που μπορεί να επιτευχθεί με τη χρήση της μετρικής *PCI*. Επομένως, με βάση τη μετρική *PCI* και λαμβάνοντας υπόψη τα μοναδικά χαρακτηριστικά των ad hoc δικτύων οχημάτων, όπως για παράδειγμα ότι η κίνηση δεν είναι τυχαία αλλά περιορίζεται στα όρια των δρόμων, σκοπεύουμε να ερευνήσουμε το ζήτημα του caching σε ad hoc δίκτυα οχημάτων.

Στο Κεφάλαιο 6 ασχοληθήκαμε με το ζήτημα της ομαδοποίησης σε ΑΔΑ και παρουσιάσαμε τον αλγόριθμο GESG. Δείξαμε ότι είναι καλύτερος από τον αλγόριθμο LEACH, που όμως θεωρεί ότι κάθε αισθητήρας είναι σε θέση να ρυθμίζει το επίπεδο μετάδοσης. Σκοπός μας είναι υλοποιήσουμε τον αλγόριθμο GESG λαμβάνοντας υπόψη τα διαφορετικά επίπεδα μετάδοσης, τα οποία μπορεί να πραγματοποιεί ανεξάρτητα ο κάθε αισθητήρας έτσι ώστε να μελετήσουμε την αποτελεσματικότητα και την απόδοση του αλγορίθμου. Ο αλγόριθμος GESG αναδεικνύει με κατανομημένο τρόπο ως κεφαλές ομάδας τους κόμβους από όπου διέρχονται τα περισσότερα συντομότερα μονοπάτια του τοπικού δικτύου. Σκοπός της μελλοντικής έρευνας στο ζήτημα της ομαδοποίησης είναι η μετατροπή του GESG ώστε κάθε μονοπάτι να μην εμπεριέχει μόνο την έννοια της συντομότερης απόστασης αλλά και της ενεργειακά αποδοτικότερης. Αυτό θα πραγματοποιηθεί αν σε κάθε ακμή του γράφου τοποθετηθεί ως βάρος το ημίθροισμα των ενεργειών των κόμβων που συνδέει. Συνεπώς, η επέκταση του αλγορίθμου GESG θέτει ως κεφαλές ομάδας τους κόμβους που βρίσκονται επάνω σε ένα μεγάλο αριθμό συντομότερων και πλούσιων ενεργειακά μονοπατιών.

Στο Κεφάλαιο 7 ερευνήσαμε το ζήτημα της δημιουργίας καταλόγων για μερικώς διατεταγμένα δεδομένα. Για να αποτιμήσουμε την απόδοση των προτεινό-

μενων αλγορίθμων, υλοποιήσαμε ένα σύστημα που προσομοιώνει ένα περιβάλλον όπου ένας αριθμός κόμβων προσπελούν τα δεδομένα που παρέχονται από τον εξυπηρετητή μέσω του ασύρματου καναλιού εκπομπής. Σκοπός μελλοντικής ενασχόλησης με το ζήτημα είναι η υλοποίηση των προτεινόμενων αλγορίθμων στο περιβάλλον ενός δικτυακού προσομοιωτή για τη βέλτιστη και αποτελεσματικότερη αποτίμησή τους.

Στο ερευνητικό τοπίο που ξεκίνησε να αναδύεται πριν από τρία χρόνια, διαβλέπουμε δυο μεγάλες περιοχές, που δημιουργούν νέες προκλήσεις σχετικά με τις τεχνικές που αναπτύχθηκαν στα πλαίσια της παρούσας διατριβής. Η πρώτη περιοχή εστιάζεται γύρω από τα ad hoc δίκτυα οχημάτων. Τα ad hoc δίκτυα οχημάτων παρουσιάζουν κάποια ιδιαίτερα χαρακτηριστικά, κυριότερα από τα οποία είναι οι συχνές μεταβολές στην τοπολογία του δικτύου εξαιτίας της γρήγορης κίνησης των οχημάτων, οι συχνές αποσυνδέσεις του δικτύου ιδιαίτερα στην περίπτωση της αραιής τοπολογίας, η συνάθροιση των δεδομένων που απαιτείται εξαιτίας του περιορισμένου εύρους ζώνης και το γεγονός ότι η ενέργεια δεν αποτελεί σημαντικό ζήτημα. Η τεχνική του caching αλλά και της ομαδοποίησης μπορούν να αποτελέσουν μια λύση έτσι ώστε να βελτιωθεί η ποιότητα της επικοινωνίας μεταξύ των οχημάτων και να μειωθεί ο χρόνος προσπέλασης των δεδομένων. Η έρευνα στον τομέα αυτό είναι ακόμη σε αρχικό στάδιο, ενώ οι σημαντικότερες προσπάθειες που έχουν γίνει είναι οι εργασίες [41, 54]. Παράλληλα στα πλαίσια των ad hoc δικτύων οχημάτων σημαντική έρευνα έχει διενεργηθεί στην περιοχή της διάχυσης, ανάκτησης και αποδοτικής δρομολόγησης των δεδομένων όπως αναφέρεται στις εργασίες [55, 120, 122, 132, 157, 158]. Στόχος μας είναι η ανάπτυξη νέων τεχνικών διάχυσης της πληροφορίας.

Η δεύτερη περιοχή εστιάζεται γύρω από τα Δίκτυα Ανεκτικά στην Καθυστέρωση. Η υιοθέτηση των ασύρματων δικτύων που είναι ανεκτικά στην καθυστέρηση δημιουργεί σημαντικές προκλήσεις που σχετίζονται με τις μεθόδους συνεργατικού (cooperative) και προσαρμοζόμενου (adaptation-aware ή transcoding) caching και broadcasting. Δύο χαρακτηριστικές εργασίες είναι οι [67, 79].

# ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] I. Abbasi and M. Younis. A survey on clustering algorithms for wireless sensor networks. *Computer Communications*, 30(14–15):2826–2841, 2007.
- [2] C. Aggrawal, J. Wolf, and P. Yu. Caching on the World Wide Web. *IEEE Transactions on Knowledge and Data Engineering*, 11(1):94–107, 1999.
- [3] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury. Wireless Multimedia Sensor Networks: A Survey. *IEEE Wireless Communications Magazine*, 14(6):32–39, 2007.
- [4] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on wireless sensor networks. *IEEE Communications magazine*, 40(8):102–116, 2002.
- [5] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless Sensor Networks: a survey. *Computer Networks*, 38(4):393–422, 2002.
- [6] A. D. Amis, R. Prakash, T. H. P. Vuong, and D. T. Huynh. Max-min d-cluster formation in wireless ad hoc networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, pages 32–41, 2000.
- [7] A. Balamash and M. Krunz. An Overview of Web Caching Replacement Algorithms. *IEEE Communications Surveys & Tutorials*, 6(2):44–56, 2004.
- [8] S. Bandyopadhyay and E. Coyle. An Energy-Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, pages 1713–1723, 2003.

- [9] D. Baker and A. Ephremides. The architectural organization of a mobile radio network via a distributed algorithm. *IEEE Transactions on Communications*, 29(11):1694–1701, 1981.
- [10] D. Barbara and T. Imielinski. Sleepers and Workaholics: Caching Strategies for Mobile Environments. *VLDB Journal*, 4(4):567–602, 1995.
- [11] S. Basagni. Distributed clustering for ad hoc networks. In *Proceedings of the International Symposium on Parallel Architectures, Algorithms and Networks (I-SPAN)*, pages 310–315, 1999.
- [12] S. Basagni, M. Mastrogiovanni, A. Panconesi, and C. Petrioli. Localized protocols for ad hoc clustering and backbone formation: A performance comparison. *IEEE Transactions on Parallel and Distributed Systems*, 17(4):292–306, 2006.
- [13] P. K. Biswas and S. Phoha. Self-organizing sensor networks for integrated target surveillance. *IEEE Transactions on Computers*, 55(8):1033–1047, 2006.
- [14] B. H. Bloom. Space/time trade-offs in hash coding with allowable errors. *Communications of the ACM*, 13(7):422–426, 1970.
- [15] U. Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.
- [16] L. Buttyan and P. Schaffer. PANEL: Position-based Aggregator Node Election in Wireless Sensor Networks. In *Proceedings of the IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS)*, pages 1–9, 2007.
- [17] G. Cao. A Scalable Low-Latency Cache Invalidation Strategy for Mobile Environment. *IEEE Transaction on Knowledge and Data Engineering*, 15(5):1251–1265, 2003.
- [18] P. Cao and S. Irani. Cost-aware WWW proxy caching algorithms. In *Proceedings of USENIX Symposium on Internet Technology and Systems*, December, 1997.
- [19] P. Cao and C. Liu. Maintaining strong cache consistency in the World Wide Web. *IEEE Transactions on Computers*, 47(4):445–457, 1998.



- [20] G. Cao, L. Yin, and C. Das. Cooperative cache based data access framework for ad hoc networks. *IEEE Computer*, 37(2):32–39, 2004.
- [21] J. Cao, Y. Zhang, and G. Cao. Consistency of cooperative caching in mobile peer-to-peer systems over MANET. In *Proceedings of the International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pages 573–579, 2005.
- [22] J. Cao, Y. Zhang, G. Cao and L. Xie. Data consistency for cooperative caching in mobile environments. *IEEE Computer*, pages 60–66, Apr. 2007.
- [23] N. Chand, R. C. R. C. Joshi, and M. Misra. A zone co-operation approach for efficient caching in mobile ad hoc networks. *International Journal of Communication Systems*, 19:1009–1028, 2006.
- [24] N. Chand, R. C. R. C. Joshi, and M. Misra. Cooperative caching strategy in mobile ad hoc networks based on clusters. *Wireless Personal Communications*, 43(1):41–63, 2007.
- [25] M. F. Chang and Y.-B. Lin. Performance of a Weakly Consistent Wireless Web Access Mechanism. *ACM SIGMETRICS Performance Evaluation Review*, 28(3):12–20, 2000.
- [26] J. Chang and L. Tassiulas. Energy conserving routing in wireless ad hoc networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, pages 22–31, 2000.
- [27] M. Chatterjee, S. K. Das, and D. Turgut. WCA: A weighted clustering algorithm for mobile ad hoc networks. *ACM/Kluwer Cluster Computing*, 5(2):193–204, 2002.
- [28] H. Che, Y. Tung, and Z. Wang. Hierarchical Web caching systems: Modeling, desing and experimental results. *IEEE Journal on Selected Areas in Communications*, 20(7):1305–1314, 2002.
- [29] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris. SPAN: An energy efficient coordination algorithm for topology maintenance in ad hoc networks. *ACM/Kluwer Wireless Networks*, 8(5):481–494, 2002.
- [30] Y.-Z. P. Chen and A. L. Liestman. Approximating Minimum Size Weakly-Connected Dominating Sets for Clustering Mobile Ad Hoc Networks. In

- 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 165–172, 2002.
- [31] M. S. Chen, K. L. Wu, and P.S. Yu. Optimizing index allocation for sequential data broadcasting in wireless mobile computing. *IEEE Transactions on Knowledge and Data Engineering*, 15(1):161–173, 2003.
- [32] H. Chen and Y. Xiao. Cache Access and Replacement for Future Wireless Internet. *IEEE Communications Magazine*, 44(5):113–123, 2006.
- [33] C.-Y. Chow, H. V. Leong, and A. T. S. Chan. Peer-to-peer cooperative caching in mobile environments. In *Proceedings of the IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pages 528–533, 2004.
- [34] C.-Y. Chow, H. V. Leong, and A. T. S. Chan. GroCoca: Group-based peer-to-peer cooperative caching in mobile environment. *IEEE Journal on Selected Areas in Communications*, 25(1):179–191, 2007.
- [35] B. Das, R. Sivakumar and V. Bhargavan. Routing in Ad Hoc Networks Using a Spine. In *Proceedings of the IEEE International Conference of Computers Communication and Networks*, pages 1–20, 1997.
- [36] B. Das and V. Bhargavan. Routing in ad-hoc networks using minimum connected dominating sets. In *Proceedings of the IEEE International Conference on Communications (ICC)*, pages 376–380, 1997.
- [37] S. M. Das, H. Pucha, and Y. C. Hu. Mitigating the gateway bottleneck via transparent cooperative caching in wireless mesh networks. *Ad Hoc Networks*, 5(6):680–703, 2007.
- [38] K. Dasgupta, K. Kalpakis and P. Namjoshi. An efficient clustering-based heuristic for data gathering and aggregation in sensor networks. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1948–1953, 2003.
- [39] P. Desnoyers, D. Ganesan, and P. Shenoy. TSAR: A two tier sensor storage architecture using interval skip graphs. In *Proceedings of the ACM International Conference on Embedded Networked Sensor Systems (SenSys)*, pages 39–50, 2005.

- [40] Y. Diao, D. Ganesan, G. Mathur, and P. Shenoy. Rethinking data management for storage-centric sensor networks. In *Proceedings of the Conference on Innovative Data Systems Research (CIDR)*, pages 22–31, 2007.
- [41] M. D. Dikaiakos, S. Iqbal, T. Nadeem, and L. Iftode. Vitp: an information transfer protocol for vehicular computing. In *Proceedings of the 2nd ACM international workshop on Vehicular ad hoc networks*, pages 30–39, 2005.
- [42] N. Dimokas, D. Katsaros, and Y. Manolopoulos. Node Clustering in Wireless Sensor Networks by Considering Structural Characteristics of the Network Graph. In *Proceedings of the International Conference on Information Technology : New Generations (ITNG)*, pages 122–127, 2007.
- [43] N. Dimokas, D. Katsaros, and Y. Manolopoulos. Cooperative caching in wireless multimedia sensor networks. In *Proceedings of the ACM SIGMM International Mobile Multimedia Communications Conference (MobiMedia)*, pages 377–382, 2007.
- [44] N. Dimokas, D. Katsaros, and Y. Manolopoulos. Cooperative caching in wireless multimedia sensor networks. *ACM Mobile Networks and Applications*, 13(3–4):337–356, 2008.
- [45] N. Dimokas, D. Katsaros, L. Tassiulas and Y. Manolopoulos. High Performance, Low Overhead Cooperative Caching for Wireless Sensor Networks. In *Proceedings of the IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Jun. 2009.
- [46] N. Dimokas, D. Katsaros, and Y. Manolopoulos. Energy-Efficient Distributed Clustering in Wireless Sensor Networks. *Journal of Parallel and Distributed Computing*, accepted, August, 2009.
- [47] N. Dimokas, D. Katsaros, and Y. Manolopoulos. Cache Consistency in Wireless Multimedia Sensor Networks. *Ad Hoc Networks*, 8(2):214–240, 2010.
- [48] Y. Eisenberg, C. E. Luna, T. N. Pappas, R. Berry, and A. K. Katsagelos. Joint source coding and transmission power management for energy efficient wireless video communications. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(6):411–424, 2002.

- [49] V. Erramilli, M. Crovella, A. Chaintreau, and C. Diot. Delegation forwarding. In *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking & Computing (MobiHoc)*, pages 251–259, 2008.
- [50] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar. Next Century Challenges: Scalable Coordination in Sensor Networks. In *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM)*, 1999.
- [51] L. Fan, P. Cao, and A. Z. Almeida, J. M. Broder. Summary cache: A scalable wide-area web cache sharing protocol. *IEEE/ACM Transactions on Networking*, 8(3):281–293, 2000.
- [52] K. W. Fan, S. Liu, and P. Sinha. Structure-Free Data Aggregation in Sensor Networks. *IEEE Transactions on Mobile Computing*, 6(8):929–942, 2007.
- [53] L. M. Feeney. Energy efficient communication in ad hoc wireless networks. In *Mobile Ad Hoc Networking*, (S. Basagni, M. Conti, S. Giordano, and I. Stojmenovic, editors), pages 301–327. IEEE/Wiley, 2004.
- [54] M. Fiore, C. Casetti, and C.-F. Chiasserini. On-demand content delivery in vehicular wireless networks. In *Proceedings of the 8th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, pages 87–94, 2005.
- [55] M. Fiore, C. Casetti, and C.-F. Chiasserini. Efficient Retrieval of User Contents in MANETs. In *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, pages 10–18, 2007.
- [56] R. Gandhi and S. Parthasarathy. Fast distributed well connected dominating sets for ad hoc networks. Technical Report CS-TR-4559, Computer Science Department, University of Maryland at College Park, 2004.
- [57] M. Gerla and J. C. Tsai. Multicluster, mobile, multimedia radio network. *ACM/Baltzer Wireless Networks*, 1(3):255–265, 1995.
- [58] B. Girod, M. Kalman, Y. J. Liang, and R. Zhang. Advances in channel-adaptive video streaming. *Wireless Communications and Mobile Computing*, 2(6):573–584, 2002.

- [59] C. Gray and D. Cheriton. Leases: An efficient fault-tolerant mechanism for distributed file cache consistency. In *Proceedings of the 12th ACM International Symposium on Operating Systems Principles (SOSP)*, pages 202–210, 1989.
- [60] B. Greenstein, S. Ratnasamy, S. Shenker, R. Govindan, and D. Estrin. DIFS: A distributed index for features in sensor networks. *Ad Hoc Networks*, 1(2–3):333–349, 2003.
- [61] S. Guha and S. Khuller. Approximation Algorithms for Connected Dominating Sets. *Springer-Verlag New York, LLC*, ISSN: 0178-4617, 1998.
- [62] T. Hara. Cooperative caching by mobile clients in push-based information systems. In *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM)*, pages 186–193, 2002.
- [63] T. Hara. Replica allocation methods in ad hoc networks with data update. *ACM/Kluwer Mobile Networks and Applications*, 8(4):343–354, 2003.
- [64] T. Hara and S. K. Madria. Data replication for improving data accessibility in ad hoc networks. *IEEE Transactions on Mobile Computing*, 5(11):1515–1532, 2006.
- [65] H. Hayashi, T. Hara, and S. Nishio. Updated data dissemination methods for updating old replicas in ad hoc networks. *Personal and Ubiquitous Computing*, 9(5): 273–283, 2005.
- [66] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan. An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(4):660–670, 2002.
- [67] D. Henriksson, T. F. Abdelzaher and R. K. Ganti. A Caching-Based Approach to Routing in Delay-Tolerant Networks. In *Proceedings of the 16th International Conference on Computer Communications and Networks (ICCCN)*, pages 69–74, 2007.
- [68] Y. T. Hou, Y. Shi and H. D. Sherali. On energy provisioning and relay node placement for wireless sensor networks. *IEEE Transactions on Wireless Communications*, 4(5):2579–2590, 2005.

- [69] J. Howard, M. Kazar, S. Menees, D. Nichols, M. Satyanarayanan, R. Sidebotham, and M. West. Scale and performance in a distributed file system. *ACM Transactions On Computer Systems*, 6(1):51–81, 1988.
- [70] C.-H. Hsu, G. Lee, and A. L. P. Chen. Index and data allocation on multiple broadcast channels considering data access frequencies. In *Proceedings of the Conference on Mobile Data Management (MDM)*, pages 87–93, 2002.
- [71] T. Hu and A. Tucker. Optimal computer search trees and variable-length alphabetical codes. *SIAM Journal on Applied Mathematics*, 21(4):514–532, 1971.
- [72] G. T. Huang. Casting the wireless sensor net. *Technology Review*, pages 51–56, Jul. 2003.
- [73] T. Imielinski, S. Viswanathan, and B. R. Badrinath. Power efficient filtering of data on air. In *Proceedings of the Conference on Extending Data Base Technology (EDBT)*, pages 245–258, 1994.
- [74] T. Imielinski, S. Viswanathan, and B. R. Badrinath. Data on air: Organization and access. *IEEE Transactions on Knowledge and Data Engineering*, 9(3):353–372, 1997.
- [75] J. Jing, A. Elmagarmid, A. Helal and R. Alonso. Bit-Sequences: an Adaptive Cache Invalidation Method in Mobile Client/Server Environments. *Mobile Networks and Applications*, 2(2):115–127, 1997.
- [76] G. Jolly and M. Younis. An energy efficient, scalable and collision less MAC layer protocol for wireless sensor networks. *Wireless Communications and Mobile Computing*, 5(3):285–304, 2005.
- [77] S. Jung, B. Lee, and S. Pramanik. A tree-structured index allocation method with replication over multiple broadcast channels in wireless environments. *IEEE Transactions on Knowledge and Data Engineering*, 17(3):311–325, 2005.
- [78] H. Karl and A. Willig. *Protocols and Architectures for Wireless Sensor Networks*. John Wiley & Sons, 2006.
- [79] G. Karlsson, V. Lenders and M. May. Delay-Tolerant Broadcasting. *IEEE Transactions on Broadcasting*, 53(1):369–381, 2007.

- [80] D. Katsaros, N. Dimokas and Y. Manolopoulos. Generalized Indexing for Energy-Efficient Access to Partially Ordered Broadcast Data in Wireless Networks. In *Proceedings of the International Database Engineering and Applications Symposium (IDEAS)*, pages 89–96, 2006.
- [81] D. Katsaros and Y. Manolopoulos. Web caching in broadcast mobile wireless environments. *IEEE Internet Computing*, 8(3):37–45, 2004.
- [82] D. Katsaros and Y. Manolopoulos. Caching in Web memory hierarchies. In *Proceedings of the ACM Symposium on Applied Computing*, pages 1109–1113, 2004.
- [83] D. Katsaros and Y. Manolopoulos. The geodesic broadcast scheme for wireless ad hoc networks. In *Proceedings of the IEEE International Symposium on a World of Wireless Mobile Multimedia (WoWMoM)*, pages 571–575, 2006.
- [84] P. Kulkarni, D. Ganesan, P. Shenoy, and Q. Lu. SensEye: A multi-tier camera sensor network. In *Proceedings of the ACM International Conference on Multimedia (MM)*, pages 229–238, 2005.
- [85] W.-C. Lee and D. L. Lee. Using signature techniques for information filtering in wireless and mobile environments. *Distributed and Parallel Databases*, 4(3):205–227, 1996.
- [86] W. Li, E. Chan, and D. Chen. Energy-efficient cache replacement policies for cooperative caching in mobile ad hoc network. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, pages 3349–3354, 2007.
- [87] X. Li, Y.-J. Kim, R. Govindan, and W. Hong. Multi-dimensional range queries in sensor networks. In *Proceedings of the ACM Conference on Embedded Networked Sensor Systems (SenSys)*, pages 63–75, 2003.
- [88] S. Lim, W.-C. Lee, G. Cao, and C. R. Das. A novel caching scheme for improving internet-based mobile ad hoc networks performance. *Ad Hoc Networks*, 4(2):225–239, 2006.
- [89] C. R. Lin and M. Gerla. Adaptive clustering for mobile wireless networks. *IEEE Journal on Selected Areas in Communications*, 15(7):1265–1275, 1997.

- [90] Y.-B. Lin, W.-R. Lai and J.-J. Chen. Effects of Cache Mechanism on Wireless Data Access. *IEEE Transaction on Wireless Communications*, 2(6):1240–1246, 2003.
- [91] C. Liu, K. Wu and J. Pei. An Energy-Efficient Data Collection Framework for Wireless Sensor Networks by Exploiting Spatiotemporal Correlation. *IEEE Transactions on Parallel and Distributed Systems*, 18(7):1010–1023, 2007.
- [92] S. C. Lo and A. L. P. Chen. Optimal index and data allocation in multiple broadcast channels. In *Proceedings of the IEEE Conference on Data Engineering (ICDE)*, pages 293–302, 2000.
- [93] Y. Manolopoulos, Y. Theodoridis, and V. J. Tsotras. *Advanced Database Indexing*, volume 17 of *The Kluwer International Series on Advances in Databases Systems*. Kluwer Academic Publishers, Boston, 1999.
- [94] G. Mathur, P. Desnoyers, D. Ganesan, and P. Shenoy. Ultra-low power data storage for sensor networks. In *Proceedings of the ACM International Conference on Information Processing in Sensor Networks (IPSN)*, pages 374–381, 2006.
- [95] N. Megiddo and D. S. Modha. ARC: A self-tuning, low overhead replacement cache. In *Proceedings of the USENIX Conference on File and Storage Technologies (FAST)*, pages 115–130, 2003.
- [96] A. Meka and A. K. Singh. DIST: A distributed spatio-temporal index structure for sensor networks. In *Proceedings of the ACM Conference on Information and Knowledge Management (CIKM)*, pages 139–146, 2005.
- [97] J. Mena, and V. Kalogeraki. Dynamic Relay Node Placement in Wireless Sensor Networks. In *Proceedings of the 2008 International Symposium on Applications and the Internet (SAINT 2008)*, pages 8–17, 2008.
- [98] S. Nath and A. Kansal. FlashDB: Dynamic self-tuning database for NAND flash. In *Proceedings of the ACM International Conference on Information Processing in Sensor Networks (IPSN)*, pages 410–419, 2007.
- [99] E. O’Neil, P. O’Neil, and G. Weikum. The LRU-K page replacement algorithm for database disk buffering. In *Proceedings of the ACM International Conference on Management of Data (SIGMOD)*, pages 297–306, 1993.



- [100] M. Nelson, B. Welch, and J. Ousterhout. Caching in the Sprite network file system. *ACM Transactions On Computer Systems*, 6(1):134–154, 1988.
- [101] G. Pallis, A. Vakali, and J. Pokorny. A clustering-based prefetching scheme on a Web cache environment. *Computers & Electrical Engineering*, 34(4):309–323, 2007.
- [102] M. Papadopouli and H. Schulzrinne. Effects of power conservation, wireless coverage and cooperation on data environments. In *Proceedings of ACM Symposium on Mobile Ad Hoc Networking and Computing (MOBI-HOC)*, pages 117–127, 2001.
- [103] C. Perkins. Ad Hoc Networks. *Addison-Wesley*, Reading, MA, 2000.
- [104] C. E. Perkins and E. Royer. Ad hoc On-demand Distance Vector routing. In *Proceedings of the IEEE Workshop on Mobile Computing Systems and Applications*, pages 90–100, 1999.
- [105] K. S. Prabh and T. F. Abdelzaher. Energy-conserving data cache placement in sensor networks. *ACM Transactions On Sensor Networks*, 1(2):178–203, 2005.
- [106] A. Qayyum, L. Viennot, and A. Laouiti. Multipoint relaying: An efficient technique for flooding in mobile wireless networks. Technical Report 3898, INRIA, March 2000.
- [107] M. Rahimi, R. Baer, O. I. Iroezi, J. C. Garcia, J. Warrior, D. Estrin, and M. Srivastava. Cyclops: In situ image sensing and interpretation in wireless sensor networks. In *Proceedings of the ACM International Conference on Embedded Networked Sensor Systems (SenSys)*, pages 192–204, 2005.
- [108] R. Rajaraman. Topology Control and Routing in Ad Hoc Networks: A Survey. *ACM SIGACT News*, 33(2):60–73, 2002.
- [109] S. Ratnasamy, B. Karp, S. Shenker, D. Estrin, R. Govindan, L. Yin, and F. Yu. Data-centric storage in sensor networks with GHT: A geographic hash table. *ACM Mobile Networks and Applications*, 8(4):427–442, 2003.
- [110] J. T. Robinson and M. V. Devarakonda. Data cache management using frequency-based replacement. In *Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, pages 134–142, 1990.

- [111] A. Rousskov and D. Wessels. Cache digests. *Computer Networks and ISDN Systems*, 30(22–23):2155–2168, 1998.
- [112] J.-H. Ryu, S. Song and D.-H. Cho. New Clustering Schemes for Energy Conservation in Two-Tiered Mobile Ad Hoc Networks. *Proceedings of the International Conference on Communications*, pages 862–866, 2001.
- [113] F. Sailhan and V. Issarny. Energy-aware Web caching for mobile terminals. In *Proceedings of the IEEE International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pages 820–825, 2002.
- [114] F. Sailhan and V. Issarny. Cooperative caching in ad hoc networks. In *Proceedings of the IEEE International Conference on Mobile Data Management (MDM)*, pages 13–28, 2003.
- [115] M. Satyanarayanan. Fundamental Challenges in Mobile Computing. In *Proceedings of the Annual ACM Symposium on Principles of Distributed Computing*, pages 1–7, 1996.
- [116] A. Seifert and J.-J. Hung. FlexInd: A flexible and parameterizable air-indexing scheme for data broadcast systems. In *Proceedings of the Conference on Extending Data Base Technology (EDBT)*, volume 3820 of *Lecture Notes in Computer Science*, pages 902–920, 2006.
- [117] H. Shen, S. K. Das, M. Kumar, and Z. Wang. Cooperative caching with optimal radius in hybrid wireless networks. In *Proceedings of the International IFIP-TC6 Networking Conference (NETWORKING)*, ser. Lecture Notes on Computer Science, 3042:841–853, 2004.
- [118] H. Shen, M. S. Joseph, M. Kumar, and S. K. Das. PReCinCt: A scheme for cooperative caching in mobile peer-to-peer systems. In *Proceedings of the International Parallel and Distributed Processing Symposium (IPDPS)*, 2005.
- [119] N. Shivakumar and S. Venkatasubramanian. Efficient indexing for broadcast based wireless systems. *ACM/Baltzer Mobile Networks and Applications*, 1(4):433–446, 1996.
- [120] M. L. Sichitiu and M. Kihl, Inter-vehicle communication systems: a survey. *IEEE Communications Surveys & Tutorials*, 10(2):88–105, 2008.

- [121] J. Shim, P. Scheuermann, and R. Vingralek. Proxy cache algorithms: Design, implementation and performance. *IEEE Transactions on Knowledge and Data Engineering*, 11(4):549–562, 1999.
- [122] A. Skordylis and N. Trigoni. Delay-bounded Routing in Vehicular Ad-hoc Networks. In *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 341–350, 2008.
- [123] A. Sobeih, J. C. Hou, L.-C. Kung, N. Li, H. Zhang, W.-P. Chen, H.-Y. Tian, and H. Lim. J-Sim: A simulation and emulation environment for wireless sensor networks. *IEEE Wireless Communications magazine*, 13(4):104–119, 2006.
- [124] I. Stojmenovic, M. Seddigh, and J. Zunic. Dominating sets and neighbor elimination-based broadcasting algorithms in wireless networks. *IEEE Transactions on Parallel and Distributed Systems*, 13(1):14–25, 2002.
- [125] S. Subramaniam, T. Palpanas, D. Papadopoulos, V. Kalogeraki and D. Gunopulos. Online outlier detection in sensor data using non-parametric models. In *Proceedings of the International Conference on Very Large Data Bases (VLDB)*, pages 187–198, 2006.
- [126] R. Szewczyk, E. Osterweil, J. Polastre, M. Hamilton, A. Mainwaring, and D. Estrin. Habitat monitoring with sensor networks. *Communications of the ACM*, 47(6):34–40, 2004.
- [127] M. Takaaki and H. Aida. Cache data access system in ad hoc networks. In *Proceedings of the IEEE Spring Semiannual Vehicular Technology Conference (VTC)*, 2:1228–1232, 2003.
- [128] B. Tang, S. Das, and H. Gupta. Cache placement in sensor networks under update cost constraint. In *Proceedings of the (ADHOC-NOW)*, ser. Lecture Notes on Computer Science, 3738, 334–348, 2005.
- [129] B. Tang and H. Gupta. Cache placement in sensor networks under an update cost constraint. *Journal of Discrete Algorithms*, 5(3):422–435, 2007.
- [130] B. Tang, H. Gupta and S. R. Das. Benefit-Based Data Caching in Ad Hoc Networks. *IEEE Transactions on Mobile Computing*, 7(3):289–304, 2008.

- [131] L. Tassiulas and C. Su. Optimal memory management strategies for a mobile user in a broadcast data delivery system. *IEEE Journal on Selected Areas in Communications*, 15(7):1226–1238, 1997.
- [132] Y. Toor, P. Muhlethaler and A. Laouiti. Vehicle Ad Hoc networks: applications and related technical issues. *IEEE Communications Surveys & Tutorials*, 10(3):74–88, 2008.
- [133] Y.-C. Tseng, M.-S. Pan, and Y.-Y. Tsai. Wireless sensor networks for emergency navigation. *IEEE Computer*, 39(7):55–62, Aug. 2006.
- [134] M. A. Viredaz, L. S. Brakmo, and W. R. Hamburger. Energy management on handheld devices. *ACM Queue*, 1(7):44–52, 2003.
- [135] J. Wang. A Survey of Web Caching Schemes for the Internet. *ACM SIGCOMM Computer Communication Review*, 29(5):36–46, 1999.
- [136] D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393:440–442, 1998.
- [137] D. Wessels and K. Claffy. ICP and the Squid Web cache. *IEEE Journal on Selected Areas in Communications*, 16(3):345–357, 1998.
- [138] S. Williams, M. Abrams, C. R. Standrige, G. Abdulla and E. A. Fox. Removal Policies in Network Caches for World-Wide-Web Documents. In *Proceedings on Applications, Technologies, Architectures, and Protocols for Computer Communications*, pages 293–305, 1996.
- [139] J. Wu and H. Li. On Calculating Connected Dominating Sets for Efficient Routing in Ad Hoc Wireless Networks. In *Proceedings of the International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, pages 7–14, 1999.
- [140] J. Wu and H. Li. A dominating-set-based routing scheme in ad hoc wireless networks. *Telecommunication Systems*, 18(1–3):13–36, 2001.
- [141] Y. Xu, J. Heidemann, and D. Estrin. Geography-Informed Energy Conservation for Ad Hoc Routing. In *Proceedings of the ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM)*, pages 1713–1723, 2001.

- [142] J. Xu, W.-C. Lee, X. Tang, Q. Gao, and S. Li. An error-resilient and tunable distributed indexing scheme for wireless data broadcast. *IEEE Transactions on Knowledge and Data Engineering*, 18(3):392–404, 2006.
- [143] J. Xu and K.-L. Tan. An analysis of selective tuning schemes for nonuniform broadcast. *Data and Knowledge Engineering*, 22(3):319–344, 1997.
- [144] X. Yang and A. Bouguettaya. Adaptive data access in broadcast-based wireless environments. *IEEE Transactions on Knowledge and Data Engineering*, 17(3):326–338, 2005.
- [145] M. Ye, C. Li, G. Chen, and J. Wu. An Energy Efficient Clustering Scheme in Wireless Sensor Networks. *Ad Hoc and Sensor Wireless Networks*, 3:99–119, 2006.
- [146] W. G. Yee and S. B. Navathe. Efficient data access to multi-channel broadcast programs. In *Proceedings of the ACM Conference on Information and Knowledge Management (CIKM)*, pages 153–160, 2003.
- [147] W. G. Yee, S. B. Navathe, E. Omiecinski, and C. Jermaine. Bridging the gap between response time and energy-efficiency in broadcast schedule design. In *Proceedings of International Conference on Extending Data Base Technology (EDBT)*, pages 572–589, 2002.
- [148] J. Yin, L. Alvisi, M. Dahlin, and C. Lin. Volume Leases for consistency in large-scale systems. *IEEE Transactions on Knowledge and Data Engineering*, 11(4):563–576, 1999.
- [149] L. Yin and G. Cao. Supporting cooperative caching in ad hoc networks. *IEEE Transactions on Mobile Computing*, 5(1):77–89, 2006.
- [150] N. Young. The  $k$ -server dual and loose competitiveness for paging. *Algorithmica*, 11(6):525–541, 1994.
- [151] O. Younis and S. Fahmy. HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing*, 3(4):366–379, 2004.
- [152] O. Younis, M. Krunz, and S. Ramasubramanian. Node clustering in wireless sensor networks: Recent developments and deployment challenges. *IEEE Network*, 20(3):20–25, May/June. 2006.

- [153] M. Younis, M. Youssef and K. Arisha. Energy-aware management in cluster-based sensor networks. *Computer Networks*, 43(5):649–668, 2003.
- [154] A. Youssef, M. Younis, M. Youssef, and A. Agrawala. Distributed formation of overlapping multi-hop clusters in wireless sensor networks. In *Proceedings of the 49th Annual IEEE Global Communication Conference (Globecom)*, pages 1–6, 2006.
- [155] X. Yu. Distributed cache updating for the dynamic source routing protocol. *IEEE Transactions on Mobile Computing*, 5(6):609–626, 2006.
- [156] J. Y. Yu and P. H. Chong. A Survey of Clustering Schemes for Mobile Ad Hoc Networks. *IEEE Communications Surveys*, 7(1):32–48, 2005.
- [157] J. Zhao and G. Cao. VADD: Vehicle-Assisted Data Delivery in Vehicular Ad Hoc Networks. In *Proceedings of the Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, pages 1–12, 2006.
- [158] J. Zhao, Y. Zhang, G. Cao. Data Pouring and Buffering on the Road: A New Data Dissemination Paradigm for Vehicular Ad Hoc Networks. *IEEE Transactions on Vehicular Technology*, 56(6):3266–3277, 2007.

## ΠΑΡΑΡΤΗΜΑ Α

# Λίστα Ερευνητικών Εργασιών

### Περιοδικά διεθνή με κριτές

1. N. Dimokas, D. Katsaros and Y. Manolopoulos: “Cooperative Caching in Wireless Multimedia Sensor Networks”, *ACM Mobile Networks and Applications*, Vol.13, No.3-4, pp.337-356, Aug. 2008.
2. N. Dimokas, D. Katsaros and Y. Manolopoulos: “Energy-Efficient Distributed Clustering in Wireless Sensor Networks”, *Journal of Parallel and Distributed Computing (Academic Press/Elsevier)*, accepted, August, 2009.
3. N. Dimokas, D. Katsaros and Y. Manolopoulos: “Cache Consistency in Wireless Multimedia Sensor Networks”, *Ad Hoc Networks (Elsevier)*, Vol.8, No.2, pp.214–240, March, 2010.

### Κεφάλαια σε βιβλία με κριτές

1. N. Dimokas, D. Katsaros and Y. Manolopoulos: “Cooperative Caching in Wireless Multimedia Sensor Networks”, chapter in book *Cooperative Wireless Communications*, CRC Press (Taylor & Francis Group), chapter 16, pp.371–390, 2009.
2. N. Dimokas, D. Katsaros, P. Bozanis and Y. Manolopoulos: “Predictive Location Tracking in Cellular and in Ad Hoc Wireless Networks”, chapter in book *Mobile Intelligence: when Computational Intelligence Meets Mobile Paradigm*, (Laurence T. Yang and Agustinus Borgy Waluyo, eds.), John Wiley & Sons, Inc, chapter 8, pp.163–190, 2010.

### Συνέδρια διεθνή με κριτές

1. D. Katsaros, N. Dimokas and Y. Manolopoulos: “Generalized Indexing for Energy-Efficient Access to Partially Ordered Broadcast Data in Wireless Networks”, *Proceedings 10th International Database Engineering and Applications Symposium (IDEAS)*, Delhi, India, pp.89-96, Dec 2006.
2. N. Dimokas, D. Katsaros and Y. Manolopoulos: “Node Clustering in Wireless Sensor Networks by Considering Structural Characteristics of the Network Graph”, *Proceedings 4th International Conference on Information Technology: New Generations (ITNG)*, Las Vegas, NV, pp.122-127, Apr 2007.
3. N. Dimokas, D. Katsaros and Y. Manolopoulos: “Cooperative Caching in Wireless Multimedia Sensor Networks”, *Proceedings 3rd ICST/ACM International Mobile Multimedia Communications Conference (MobiMedia)*, Nafpaktos, Greece, pp.377-382, Aug 2007.
4. N. Dimokas, D. Katsaros, L. Tassioulas and Y. Manolopoulos: “High Performance, Low Overhead Cooperative Caching for Wireless Sensor Networks”, *Proceedings 10th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Kos Island, Greece, June 2009.

### Συνέδρια εθνικά με κριτές

1. N. Dimokas, N. Mittas, A. Nanopoulos and L. Angelis: “A Prototype System for Educational Data Warehousing and Mining”, *Proceedings 12th Panhellenic Conference on Informatics (PCI)*, Samos Island, Greece, pp.199-203, Aug 2009.

### Εργασίες υποβληθείσες ή υπό προετοιμασία

1. N. Dimokas, D. Katsaros, L. Tassioulas and Y. Manolopoulos: “High Performance Cooperative Caching for Wireless Sensor Networks”, under major review for journal publication, 2009.
2. N. Dimokas, D. Katsaros and Y. Manolopoulos, “An Efficient Data Dissemination and Caching Policy in Vehicular Ad Hoc Networks”, under preparation, 2009.