



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

**Διαχείριση Πληροφορίας Πλαισίου για το Διάχυτο
Υπολογισμό**

Οδυσσέας Λ. Σέκκας

**ΑΘΗΝΑ
ΦΕΒΡΟΥΑΡΙΟΣ 2010**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Διαχείριση Πληροφορίας Πλαισίου για το Διάχυτο Υπολογισμό

Οδυσσέας Λ. Σέκκας

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ:

Ευστάθιος Χατζηευθυμιάδης, Επίκουρος Καθηγητής ΕΚΠΑ

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ ΠΑΡΑΚΟΛΟΥΘΗΣΗΣ:

Ευστάθιος Χατζηευθυμιάδης, Επίκουρος Καθηγητής ΕΚΠΑ

Λάζαρος Μεράκος, Καθηγητής ΕΚΠΑ

Δρακούλης Μαρτάκος, Αναπληρωτής Καθηγητής ΕΚΠΑ

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:

Ευστάθιος Χατζηευθυμιάδης
Επίκουρος Καθηγητής ΕΚΠΑ

Λάζαρος Μεράκος
Καθηγητής ΕΚΠΑ

Δρακούλης Μαρτάκος
Αναπληρωτής Καθηγητής ΕΚΠΑ

Γεώργιος Ξυλωμένος
Επίκουρος Καθηγητής ΟΠΑ

Ευάγγελος Ζέρβας
Καθηγητής ΤΕΙ Αθήνας

Αλέξης Δελής
Καθηγητής ΕΚΠΑ

Ηλίας Μανωλάκος
Αναπληρωτής Καθηγητής ΕΚΠΑ

Ημερομηνία Εξέτασης 04/02/2010

ΠΕΡΙΛΗΨΗ

Τα συστήματα διάχυτου υπολογισμού καλούνται να διαχειριστούν την πληροφορία πλαισίου (context) που χαρακτηρίζει την περιρρέουσα κατάσταση διαφόρων οντοτήτων (π.χ., χρήστες, υπολογιστές, περιβάλλον, κτλ.). Στην παρούσα διατριβή εξετάζονται θέματα που αφορούν τη διαχείριση πλαισίου (context management) καθώς και την ικανότητα ενός συστήματος διάχυτου υπολογισμού να προσαρμόζεται στις ανάγκες των χρηστών λαμβάνοντας αυτόνομα αποφάσεις, ελαχιστοποιώντας έτσι την παρέμβασή τους. Έμφαση δίνεται στην αναπαράσταση και το συμπερασμό πλαισίου, καθώς και στη χρήση σχημάτων προσεγγιστικού συμπερασμού γνώσης (Θεωρία Ασαφών Συνόλων - ΘΑΣ). Ιδιαίτερης σημασίας αποτελεί και η διαχείριση σε χαμηλότερο επίπεδο της πληροφορίας περιβάλλοντος που προέρχεται από δίκτυα αισθητήρων. Μέσω των τεχνικών σύντηξης δεδομένων/αποφάσεων που προτείνονται καθίσταται δυνατή η επεξεργασία αυτού του είδους πλαισίου. Επιπλέον, μελετήθηκαν θέματα συνεργατικής επίγνωσης, συμπερασμού και βιο-μιμητικής διάδοσης πλαισίου. Προτάθηκαν κατάλληλοι αλγόριθμοι προσαρμοστικής συμπεριφοράς ώστε να καταστεί δυνατή η αποδοτικότερη διάδοσή της πληροφορίας σε κατανεμημένα περιβάλλοντα. Ο στόχος είναι η ελαχιστοποίηση του κόστους επικοινωνίας και παράλληλα η μεγιστοποίηση της «ποιότητας» του διαδιδόμενου πλαισίου. Κατά συνέπεια, στα πλαίσια της διατριβής αναδείχθηκαν βασικά ζητήματα όπως η ανακάλυψη, η αναπαράσταση και ο συμπερασμός πλαισίου, η σύντηξη πλαισίου και η συνεργατική συμπεριφορά οντοτήτων σε κατανεμημένα περιβάλλοντα.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Διάχυτος Υπολογισμός

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Κινητός Υπολογισμός, Διαχείριση Πληροφορίας Πλαισίου, Επίγνωση Πληροφορίας Πλαισίου, Συμπερασμός Πλαισίου, Επιδημικοί Αλγόριθμοι

ABSTRACT

Pervasive Computing systems have to deal with the contextual information (context), which characterizes the current situation of the involved entities (e.g., users, mobile devices, environment, etc.). This thesis studies context management issues related to the capability of a pervasive system on adapting its behavior to the involved entities' context. Specifically, the interaction between the user and such system has to be less intruding as long as the latter recognizes the current user situation and adapts its functions accordingly. Such issues comprise the concept of Context Awareness. The thesis focuses on context knowledge representation and reasoning as well as on approximate reasoning (Fuzzy Sets Theory). The management of lower level environmental information that emanates from sensors is also of great importance and is achieved with novel data fusion and decision fusion techniques that are proposed. Moreover, have been studied issues regarding collaborative context awareness and reasoning as well as bio-mimetic contextual dissemination. Adaptive algorithms are proposed so that is rendered possible the efficient dissemination of information in distributed environments. The objective is the minimization of communicating costs and the enhancement of "context quality". Consequently, in this thesis have been designated issues such as context discovery, context representation and inference, context fusion and collaborative context awareness.

SUBJECT AREA: Pervasive Computing

KEYWORDS: Mobile Computing, Context Management, Context-Awareness, Context Reasoning, Epidemical Algorithms

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

| | |
|--|-----------|
| ΠΡΟΛΟΓΟΣ | 13 |
| ΚΕΦΑΛΑΙΟ 1 | 15 |
| ΕΙΣΑΓΩΓΗ | 15 |
| 1.1 Διάχυτος Υπολογισμός | 15 |
| 1.1.1 Χαρακτηριστικά του διάχυτου υπολογισμού..... | 16 |
| 1.2 Πληροφορία Πλαισίου..... | 18 |
| 1.3 Επίγνωση Πληροφορίας Πλαισίου | 19 |
| 1.4 Συνεργατική Επίγνωση Πληροφορίας Πλαισίου | 21 |
| 1.5 Συνεισφορά Διδακτορικής Διατριβής..... | 22 |
| ΚΕΦΑΛΑΙΟ 2 | 27 |
| ΕΠΙΓΝΩΣΗ ΤΗΣ ΚΑΤΑΣΤΑΣΗΣ ΤΟΥ ΧΡΗΣΤΗ | 27 |
| 2.1 Εισαγωγή..... | 27 |
| 2.2 Πιθανο-θεωρητική Σύντηξη..... | 28 |
| 2.2.1 Προσδιορισμός χωρικής πληροφορίας πλαισίου (θέσης) του χρήστη..... | 30 |
| 2.2.2 Αξιολόγηση συστήματος προσδιορισμού θέσης | 43 |
| 2.3 Πιθανο-θεωρητική Σύντηξη Βασισμένη στην Αξιοπιστία των Πηγών | 47 |
| 2.3.1 Αξιοπιστία πηγής και βαθμός εμπιστοσύνης..... | 48 |
| 2.3.2 Ενσωμάτωση της αξιοπιστίας των πηγών στη διαδικασία σύντηξης..... | 49 |
| 2.3.3 Αξιολόγηση Συστήματος | 52 |
| 2.4 Συμπεράσματα | 56 |
| ΚΕΦΑΛΑΙΟ 3 | 57 |
| ΕΠΙΓΝΩΣΗ ΤΗΣ ΚΑΤΑΣΤΑΣΗΣ ΤΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ | 57 |
| 3.1 Εισαγωγή..... | 57 |
| 3.2 Διαδικασία Σύντηξης για την Επίγνωση Καταστάσεων του Περιβάλλοντος | 58 |
| 3.2.1 Γενικά..... | 58 |
| 3.2.2 Πρώτο επίπεδο σύντηξης – Σύντηξη δεδομένων | 59 |
| 3.2.3 Δεύτερο επίπεδο σύντηξης – Σύντηξη πληροφορίας | 67 |
| 3.3 Χρήση της Δι-επίπεδης Σύντηξης για την Πυρανήχνευση..... | 71 |
| 3.3.1 Ανίχνευση με βάση τους αισθητήρες θερμοκρασίας και υγρασίας | 74 |
| 3.3.2 Σενάρια αξιολόγησης με τη χρήση της τεχνικής CUSUM | 75 |
| 3.3.3 Σενάρια αξιολόγησης με τη χρήση της θεωρίας DST | 79 |
| 3.4 Συμπεράσματα | 82 |
| ΚΕΦΑΛΑΙΟ 4 | 83 |
| ΒΙΟ-ΜΙΜΗΤΙΚΗ ΣΥΝΕΡΓΑΤΙΚΗ ΕΠΙΓΝΩΣΗ ΠΛΑΙΣΙΟΥ | 83 |
| 4.1 Εισαγωγή..... | 83 |
| 4.2 Προσαρμογή ρυθμού μόλυνσης βάσει των χαρακτηριστικών του δικτύου..... | 86 |
| 4.2.1 Μη προσαρμοστικός αλγόριθμος $\Delta 0$ | 86 |
| 4.2.2 Προσαρμοστικός αλγόριθμος $\Delta 1$ | 87 |
| 4.2.3 Προσαρμοστικός αλγόριθμος $\Delta 2$ | 88 |

| | |
|--|------------|
| 4.2.4 Πειραματική αξιολόγηση..... | 89 |
| 4.3 Προσαρμογή ρυθμού μόλυνσης βάσει των χαρακτηριστικών της πληροφορίας πλαisiού..... | 96 |
| 4.3.1 Προσαρμοστικό Επιδημικό Σχήμα Διάδοσης Πληροφορίας L1 | 97 |
| 4.3.2 Προσαρμοστικό Επιδημικό Σχήμα Διάδοσης Πληροφορίας L2 | 107 |
| 4.4 Συμπεράσματα | 111 |
| ΚΕΦΑΛΑΙΟ 5 | 113 |
| ΣΥΝΕΡΓΑΤΙΚΗ ΕΠΙΓΝΩΣΗ ΠΛΑΙΣΙΟΥ ΓΙΑ ΤΟΝ ΣΥΜΠΕΡΑΣΜΟ ΤΗΣ ΑΞΙΟΠΙΣΤΙΑΣ ΤΩΝ ΚΟΜΒΩΝ ΣΕ AD-HOC ΔΙΚΤΥΑ..... | 113 |
| 5.1 Εισαγωγή..... | 113 |
| 5.2 Γενική αρχιτεκτονική του πλαισιού ATF | 115 |
| 5.2.1 Αισθητήρες Αξιοπιστίας..... | 117 |
| 5.2.2 Δημιουργός Αξιοπιστίας..... | 117 |
| 5.2.3 Διαχειριστής Φήμης..... | 118 |
| 5.2.4 Αισθητήρας Αξιοπιστίας της Υπηρεσίας Σύστασης..... | 118 |
| 5.3 Η λειτουργία των Αισθητήρων Αξιοπιστίας | 119 |
| 5.4 Ο υπολογισμός της αξιοπιστίας | 120 |
| 5.4.1 Η ποιοτική άποψη | 120 |
| 5.4.2 Τα ιστορικά δεδομένα..... | 122 |
| 5.4.3 Η ποσοτική άποψη..... | 125 |
| 5.5 Πολιτική Οικοδόμησης Εμπιστοσύνης | 128 |
| 5.6 Επιλογή των Αξιόπιστων Συστηνόντων Κόμβων | 129 |
| 5.7 Πειραματική αξιολόγηση της αρχιτεκτονικής ATF | 131 |
| 5.7.1 Περιβάλλον προσομοίωσης | 131 |
| 5.7.2 Αποτελέσματα προσομοιώσεων | 132 |
| 5.8 Συμπεράσματα | 135 |
| ΚΕΦΑΛΑΙΟ 6 | 137 |
| ΣΧΕΤΙΚΗ ΕΡΕΥΝΑ..... | 137 |
| 6.1 Εισαγωγή..... | 137 |
| 6.2 Σχετική Έρευνα στην Επίγνωση Πλαισιού για τον Χρήστη | 137 |
| 6.3 Σχετική Έρευνα στην Επίγνωση Πλαισιού για το Περιβάλλον..... | 139 |
| 6.4 Σχετική έρευνα στην Βιο-μιμητική Συνεργατική Επίγνωση Πλαισιού..... | 141 |
| 6.5 Σχετική έρευνα στην Συνεργατική Επίγνωση Πλαισιού για τον συμπερασμό της αξιοπιστίας των κόμβων | 142 |
| ΚΕΦΑΛΑΙΟ 7 | 145 |
| ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΚΑΤΕΥΘΥΝΣΕΙΣ | 145 |
| 7.1 Συμπεράσματα | 145 |
| 7.1.1 Συμπεράσματα από την Επίγνωση της Κατάστασης του Χρήστη..... | 146 |
| 7.1.2 Συμπεράσματα από την Επίγνωση της Κατάστασης του Περιβάλλοντος | 147 |
| 7.1.3 Συμπεράσματα από την Βιο-μιμητική Συνεργατική Επίγνωση Πλαισιού | 148 |
| 7.1.4 Συμπεράσματα από τον συμπερασμό της αξιοπιστίας με βάση τη Συνεργατική Επίγνωση Πλαισιού..... | 149 |

| | |
|---------------------------------------|------------|
| 7.2 Μελλοντικές Κατευθύνσεις..... | 151 |
| ΚΕΦΑΛΑΙΟ 8 | 155 |
| ΠΑΡΑΡΤΗΜΑ..... | 155 |
| 8.1 Δίκτυα Bayes | 155 |
| 8.2 Ασαφής Λογική | 157 |
| 8.3 Φάροι υπερύθρων | 162 |
| ΟΡΟΛΟΓΙΑ..... | 165 |
| ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ | 167 |
| ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ..... | 169 |

ΠΡΟΛΟΓΟΣ

Η παρούσα διδακτορική διατριβή εκπονήθηκε στο Τμήμα Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών και υποστηρίχθηκε από το «Πρόγραμμα Ενίσχυσης Ερευνητικού Δυναμικού (ΠΕΝΕΔ 2003)».

Στα πλαίσια της διατριβής μελετώνται διάφορα θέματα που σχετίζονται με τη διαχείριση πληροφορίας πλαισίου στο διάχυτο υπολογισμό. Συγκεκριμένα, εξετάζονται προβλήματα που αφορούν τη αναπαράσταση και το συμπερασμό πλαισίου καθώς και την ικανότητα ενός συστήματος επίγνωσης πλαισίου να προσαρμόζεται στις ανάγκες των χρηστών λαμβάνοντας αυτόνομα αποφάσεις, ελαχιστοποιώντας έτσι την παρέμβασή τους. Τεχνικές σύντηξης δεδομένων, σύντηξης αποφάσεων καθώς και χρήση σχημάτων προσεγγιστικού συμπερασμού γνώσης (Θεωρία Ασαφών Συνόλων), συνέβαλλαν στην αντιμετώπιση των προαναφερθέντων προβλημάτων. Ιδιαίτερης σημασίας αποτελεί και η διαχείριση σε χαμηλότερο επίπεδο της πληροφορίας περιβάλλοντος προερχόμενη από δίκτυα αισθητήρων. Επιπλέον, μελετήθηκαν θέματα συνεργατικής επίγνωσης πλαισίου για τον συμπερασμό της αξιοπιστίας των κόμβων ενός ad-hoc δικτύου καθώς και θέματα βιο-μιμητικής διάδοσης πλαισίου. Προτάθηκαν κατάλληλοι αλγόριθμοι προσαρμοστικής συμπεριφοράς ώστε να καταστεί δυνατή η αποδοτικότερη διάδοση της πληροφορίας σε κατανομημένα περιβάλλοντα και η αντιμετώπιση ορισμένων προβλημάτων που σχετίζονται με την συνεργατική επίγνωση πλαισίου όπως κόστος επικοινωνίας, κτλ.).

Ολοκληρώνοντας την εκπόνησή της διδακτορικής μου διατριβής, νιώθω την ανάγκη να ευχαριστήσω εκείνους τους ανθρώπους, οι οποίοι συνέβαλαν στην ολοκλήρωση της. Πρωτίστως, ευχαριστώ θερμά τους καθηγητές της τριμελούς συμβουλευτικής επιτροπής, τον Επίκουρο Καθηγητή Ευστάθιο Χατζηευθυμιάδη, τον Καθηγητή Λάζαρο Μεράκο και τον Αναπληρωτή Καθηγητή Δρακούλη Μαρτάκο για την υποστήριξη και τις συμβουλές που μου παρείχαν.

Ιδιαίτερα, ευχαριστώ θερμά τον Επίκουρο Καθηγητή Ευστάθιο Χατζηευθυμιάδη, για την πολύ σημαντική καθοδήγησή του και τις εύστοχες επισημάνσεις του καθ' όλη την διάρκεια εκπόνησης της διατριβής. Η συνεχής στήριξή του σε όλα τα επίπεδα και ο ανθρωποκεντρικός του χαρακτήρας συνέβαλλε καθοριστικά στην επίλυση των προβλημάτων που ανέκυπταν και οι συζητήσεις μαζί του υπήρξαν πηγή έμπνευσης για νέες ιδέες.

Ως μέλος της ερευνητικής ομάδας του διάχυτου υπολογισμού (<http://p-comp.di.uoa.gr/>), είχα την ευκαιρία να συνεργαστώ με ανθρώπους, οι οποίοι ήταν πάντα πρόθυμοι να με συμβουλέψουν και να με βοηθήσουν. Ευχαριστώ ιδιαίτερα τους: Ευάγγελο Ζέρβα (Καθηγητή ΤΕΙ Αθηνών), Χρήστο Αναγνωστόπουλο, Βασίλειο Τσέτσο, Γιώργο Αλυφαντή, Ιωάννη Πρίγγουρη, Ευάγγελο Νομικό, Βασίλη Παπαταξιάρχη, Βασίλειο Μπαούση και Μίλτο Κυριακάκο. Τέλος, ευχαριστώ την οικογένειά μου και τους κοντινούς μου ανθρώπους για την αμέριστη υποστήριξη και συμπαράστασή τους όλα αυτά τα χρόνια.

ΚΕΦΑΛΑΙΟ 1

ΕΙΣΑΓΩΓΗ

Η παρούσα διδακτορική διατριβή εστιάζει στη διερεύνηση και μελέτη των επιστημονικών περιοχών που σχετίζονται με τη συλλογή και αναπαράσταση της πληροφορίας περιβάλλοντος και χρήστη (context modelling) καθώς και την επίγνωση πλαισίου (context-awareness) σε περιβάλλοντα διάχυτου υπολογισμού (pervasive computing). Έμφαση δίνεται στα προβλήματα της αναπαράστασης και του συμπερασμού πλαισίου καθώς και στην ικανότητα ενός συστήματος επίγνωσης πλαισίου να προσαρμόζεται στις ανάγκες των χρηστών λαμβάνοντας αυτόνομα αποφάσεις και ελαχιστοποιώντας την παρέμβασή τους. Τεχνικές σύντηξης δεδομένων (data fusion), σύντηξης αποφάσεων (decision fusion) καθώς και χρήση σχημάτων προσεγγιστικού συμπερασμού γνώσης (Fuzzy Sets Theory-Θεωρία Ασαφών Συνόλων), συνέβαλλαν στην αντιμετώπιση των προαναφερθέντων θεμάτων.

Μελετήθηκαν επίσης θέματα συνεργατικής επίγνωσης πλαισίου κυρίως σε ασύρματα δίκτυα ομοτίμων οντοτήτων (ad-hoc) όπως ο συμπερασμός του βαθμού αξιοπιστίας ενός κόμβου βάσει των άμεσων παρατηρήσεων και των συστάσεων από αξιόπιστους συστήνοντες, ελαχιστοποιώντας έτσι τον αριθμό των περιττών μηνυμάτων. Τέλος, με τη χρήση επιδημικών αλγορίθμων και την προσαρμοστική διάχυση πληροφορίας πλαισίου κατέστη δυνατή η αντιμετώπιση ορισμένων προβλημάτων που σχετίζονται με την εξοικονόμηση ενέργειας των κόμβων του δικτύου αλλά και τη μεγιστοποίηση της «χρονικής ποιότητας» του πλαισίου.

Πριν όμως αναφερθούμε στη συνεισφορά της παρούσας διατριβής στα επιμέρους επιστημονικά πεδία που εξετάσαμε, κρίνεται σκόπιμο να περιγράψουμε βασικές έννοιες όπως, διάχυτος υπολογισμός, πληροφορία πλαισίου, επίγνωση και συνεργατική επίγνωση πλαισίου.

1.1 Διάχυτος Υπολογισμός

Ο διάχυτος υπολογισμός (pervasive computing) αποτελεί την εξέλιξη του κινητού υπολογισμού (mobile computing) και ξεκινάει από τα μέσα της δεκαετίας του 1970. Οραματιστής του θεωρείται ο Mark Weiser. Πίστευε, ότι ο διάχυτος υπολογισμός έχει ως στόχο την εκτεταμένη χρήση πολλών υπολογιστών, καθιστώντας τους διαθέσιμους σε όλο το φυσικό περιβάλλον. Οι διάφορες τεχνολογίες θα μπορούν να χάνονται στο παρασκήνιο, ώστε ο άνθρωπος να μην έχει επίγνωση της χρήσης τους. Έτσι, οι χρήστες δεν θα πρέπει να εκπαιδεύονται στη χρήση εξειδικευμένων συσκευών, αλλά θα

πρέπει οι τεχνολογίες που τους περιβάλλουν να προσαρμόζονται σε αυτούς και να τους υποστηρίζουν [1].

Όταν συνελήφθη για πρώτη φορά το όραμα του διάχυτου υπολογισμού θεωρείτο σχεδόν εξωπραγματικό, γιατί η τεχνολογία υλικού (hardware technology) εκείνης της εποχής δεν ήταν δυνατόν να το ενσαρκώσει. Σήμερα όμως, μετά από πολλά χρόνια τεχνολογικών εξελίξεων, πολλά απαραίτητα συστατικά του διάχυτου υπολογισμού είναι διαθέσιμα ακόμη και ως εμπορικά προϊόντα (π.χ. υπολογιστές χειρός, laptop, ασύρματα δίκτυα, ετικέτες RFID, αισθητήρες, κ.α.).

Οι δύο σημαντικές περιοχές στις οποίες βασίζεται ο διάχυτος υπολογισμός είναι τα κατανεμημένα συστήματα (distributed systems) και ο κινητός υπολογισμός. Ο τομέας των κατανεμημένων συστημάτων προέκυψε από τον συνδυασμό των προσωπικών υπολογιστών και των τοπικών δικτύων. Η έρευνα που ακολούθησε από τα μέσα της δεκαετίας του 1970 έως τις αρχές της δεκαετίας του 1990 δημιούργησε ένα εννοιολογικό πλαίσιο και μια αλγοριθμική βάση που έχει διαχρονική αξία σε οποιαδήποτε εργασία εμπλέκονται δύο ή περισσότεροι υπολογιστές (σταθεροί ή κινητοί, ενσύρματοι ή ασύρματοι). Η γνώση αυτή εκτείνεται σε πολλές περιοχές οι οποίες είναι θεμελιώδεις για τον διάχυτο υπολογισμό. Η εμφάνιση των φορητών υπολογιστών (laptop) και των ασύρματων δικτύων (wireless LAN) στις αρχές της δεκαετίας του 1990 οδήγησε τους ερευνητές να αντιμετωπίσουν τα προβλήματα που ανακύπτουν κατά τη δημιουργία ενός κατανεμημένου συστήματος με κινητούς χρήστες. Έτσι, εμφανίστηκε ο κινητός υπολογισμός.

1.1.1 Χαρακτηριστικά του διάχυτου υπολογισμού

Ένα περιβάλλον διάχυτου υπολογισμού είναι κορεσμένο με υπολογιστικές και επικοινωνιακές δυνατότητες και χαρακτηρίζεται δίκαια πολλές φορές σαν «τεχνολογία που εξαφανίζεται». Δεδομένου ότι η κίνηση είναι ένα αναπόσπαστο τμήμα της καθημερινής ζωής, ένα τέτοιο περιβάλλον πρέπει να υποστηρίζει την κινητικότητα, διαφορετικά ένας χρήστης θα συνειδητοποιεί έντονα την ύπαρξη της τεχνολογίας (λόγω της απουσίας της) όταν αυτός θα κινείται. Ως εκ τούτου, η έρευνα που είναι σχετική με τον διάχυτο υπολογισμό συμπεριλαμβάνει εκτός από τα ερευνητικά πεδία του κινητού υπολογισμού και επιπλέον σημαντικά θέματα που περιγράφονται παρακάτω.

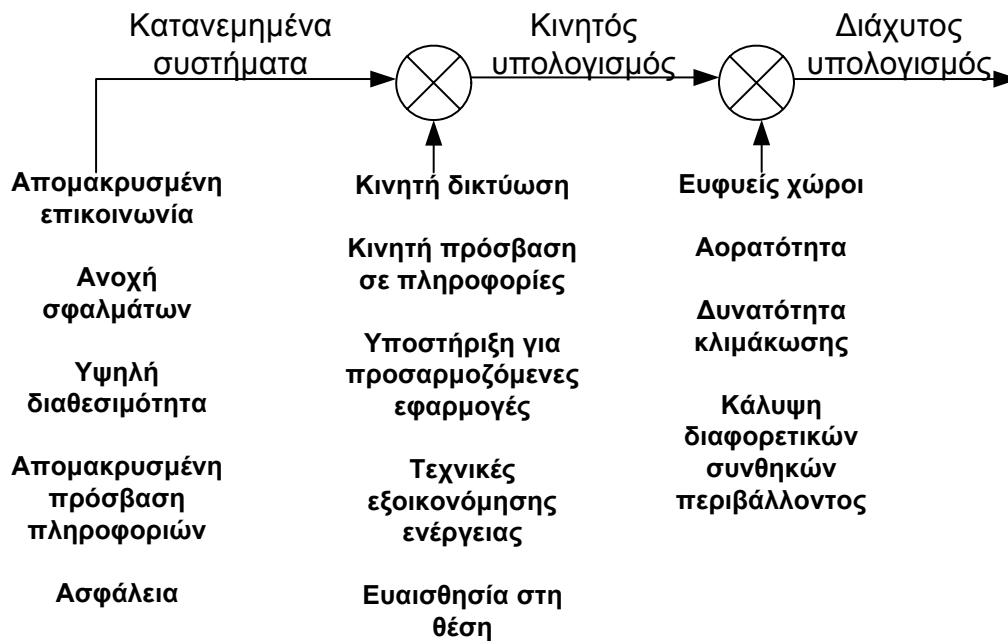
- *Αποτελεσματική χρήση ευφυών χώρων (smart spaces)*. Η πρώτη ερευνητική ώθηση του διάχυτου υπολογισμού είναι η αποτελεσματική χρήση των ευφυών (έξυπνων)

χώρων. Ένας χώρος μπορεί να είναι μέρος ενός κτιρίου (μια αίθουσα συνεδριάσεων, ένας διάδρομος, κτλ.) ή μπορεί να είναι μια καθορισμένη ανοικτή περιοχή όπως ένα προαύλιο ή ένας υπαίθριος χώρος. Η ενσωμάτωση υπολογιστικής υποδομής στην κτιριακή υποδομή είναι αυτό που χαρακτηρίζεται σαν «ευφυής χώρος» και αποτελεί ουσιαστικά τη συνένωση δύο διαφορετικών περιοχών [2].

- *Αορατότητα.* Το ιδανικό που εκφράστηκε από τον M. Weiser είναι η πλήρης εξαφάνιση της τεχνολογίας του διάχυτου υπολογισμού από την επίγνωση του χρήστη. Στην πράξη, μια λογική προσέγγιση σε αυτό το ιδανικό είναι η ελάχιστη απόσπαση της προσοχής των χρηστών. Εάν ένα περιβάλλον διάχυτου υπολογισμού ικανοποιεί συνεχώς τις προσδοκίες των χρηστών και σπάνια τους παρουσιάζει εκπλήξεις, τους επιτρέπει μία ξεκούραστη αλληλεπίδραση σχεδόν σε υποσυνείδητο επίπεδο [3]. Συγχρόνως, μία αναμονή μπορεί να είναι απαραίτητη για την αποφυγή μία μεγάλης και δυσάρεστης έκπληξης αργότερα.
- *Δυνατότητα κλιμάκωσης.* Καθώς οι ευφυείς χώροι εξελίσσονται συνεχώς οι αλληλεπιδράσεις μεταξύ του εξοπλισμού ενός χρήστη (κινητό τηλεκονσόλ, κτλ.) και του περιβάλλοντός του αυξάνονται. Αυτό έχει σημαντικές επιπτώσεις στο εύρος ζώνης, αλλά και στην ενέργεια των τηλεκονσολών. Όπως είναι φυσικό η παρουσία παραπάνω χρηστών περιπλέκει το πρόβλημα. Έτσι, η κλιμάκωση, υπό την ευρύτερη έννοια, είναι μία πτυχή που πρέπει να λαμβάνουμε σοβαρά υπ' όψιν στον διάχυτο υπολογισμό.
- *Κάλυψη (απόκρυψη) διαφορετικών συνθηκών περιβάλλοντος.* Ο βαθμός διείσδυσης του διάχυτου υπολογισμού στις υποδομές μπορεί να ποικίλει αρκετά και εξαρτάται από πολλούς μη τεχνικούς παράγοντες όπως η οργανωτική δομή και τα οικονομικά και επιχειρησιακά πρότυπα. Η ομοιόμορφη διείσδυση απέχει πολλά έτη ή δεκαετίες μακριά. Στο μεσοδιάστημα, θα υπάρχουν τεράστιες διαφορές «ευφυΐας» των διαφορετικών περιβαλλόντων. Ο διαθέσιμος εξοπλισμός σε γραφεία, αίθουσες συσκέψεων ή σε τάξεις μπορεί να είναι περισσότερο περίπλοκος απ' ό,τι σε άλλες θέσεις. Αυτή η μεγάλη δυναμική περιοχή της «ευφυΐας» μπορεί να είναι ενοχλητική σε έναν χρήστη, μειώνοντας έτσι την αορατότητα του διάχυτου υπολογισμού. Η πλήρης αορατότητα μπορεί να είναι αδύνατη, αλλά η μειωμένη μεταβλητότητα είναι κάτι που μπορεί να επιτευχθεί.

Στο Σχήμα 1.1 διακρίνεται πώς τα ερευνητικά προβλήματα στον διάχυτο υπολογισμό είναι σχετικά με εκείνα του κινητού υπολογισμού και των καταναμημένων συστημάτων.

Τα νέα προβλήματα εμφανίζονται καθώς κινούμαστε από τα αριστερά προς τα δεξιά. Επιπλέον, η επίλυση πολλών προβλημάτων που αντιμετωπίσαμε προηγουμένως τώρα γίνεται πιο σύνθετη. Η αύξηση της πολυπλοκότητας είναι πολλαπλασιαστική παρά προσθετική. Είναι πολύ δυσκολότερο να σχεδιαστεί και να εφαρμοστεί ένα σύστημα διάχυτου υπολογισμού από ένα απλό καταμεμημένο σύστημα της ίδιας ευρωστίας και αποδοτικότητας.



Σχήμα 1.1: Η εξέλιξη των τεχνολογιών.

1.2 Πληροφορία Πλαισίου

Η ετυμολογία της λέξης «πλαίσιο», στην Αγγλική γλώσσα «context», αναφέρεται σε μια «κατάσταση όπου διάφορα συμβάντα λαμβάνουν χώρα». Η έννοια αυτή σχετίζεται με την τρέχουσα/περιρρέουσα κατάσταση που γίνεται αντιληπτή από τον άνθρωπο λόγω των παρατηρούμενων συμβάντων. Ως πληροφορία πλαισίου ορίζεται «οποιαδήποτε πληροφορία μπορεί να περιγράψει την κατάσταση μιας οντότητας. Οντότητα μπορεί να αποτελεί μια συσκευή, μια εφαρμογή, ή και ένας άνθρωπος» [4].

Η έννοια «πληροφορία πλαισίου» είναι ισοδύναμη με την έννοια «πλαίσιο», εφόσον παρουσιάζει ενδιαφέρον η πληροφορία που περιγράφει την περιρρέουσα κατάσταση. Παράλληλα εμφανίζεται και η ανάγκη για τον κατάλληλο προσδιορισμό του πλαισίου σε συσχέτιση με τις ανάγκες κάθε εφαρμογής. Η έρευνα που διεξάγεται γύρω από την έννοια της «πληροφορίας πλαισίου» (contextual information) είναι τεράστια.

Μελετούνται θέματα σχετικά με την αξιοποίησή της στην ανάπτυξη «εφαρμογών διάχυτου υπολογισμού» και στα δίκτυα του μέλλοντος. Το πλαίσιο είναι ουσιαστικά ό,τι μας περιβάλλει και ο όρος αυτός χρησιμοποιείται κυρίως σε σχέση με το φυσικό κόσμο που περιβάλλει μια κινητή συσκευή, μια εφαρμογή ή ένα ολόκληρο σύστημα.

Η έννοια του πλαισίου αρχικά αποτελούνταν από τρεις πτυχές: που βρίσκεται ο χρήστης, ποιοι είναι οι γύρω του χρήστες και ποια είναι τα γειτονικά αντικείμενα. Αν και η πληροφορία θέσης είναι βασικό συστατικό για την περιγραφή της κατάστασης του χρήστη, δεν μπορεί όμως να εντοπίσει τυχόν αλλαγές στις συνθήκες του περιβάλλοντος ή δυναμικές ανακατατάξεις των γειτονικών αντικειμένων του χρήστη. Έτσι, η έννοια του πλαισίου επεκτείνεται σε μια πιο γενική θεώρηση της πληροφορίας που μπορεί να περιγράψει την κατάσταση του χρήστη συμπεριλαμβανομένου εκτός της πληροφορίας θέσης και πληροφορία που έχει σχέση και με το ίδιο το περιβάλλον όπως, για παράδειγμα, θερμοκρασία, φωτεινότητα, επίπεδο θορύβου, διαθεσιμότητα πρόσβασης σε δίκτυο, κτλ.

Η χρησιμότητα του πλαισίου είναι έκδηλη σε διάφορα επίπεδα επεξεργασίας και αξιοποίησης μέσα σε ένα σύστημα διάχυτου υπολογισμού. Σε μεγάλα κατακεντρωμένα συστήματα μπορεί να αξιοποιηθεί στη διαχείριση ενέργειας και πόρων. Σε μια κινητή εφαρμογή μπορεί να αξιοποιηθεί για προσαρμοστικότητα και αυτονομία αποφάσεων σε διάφορες συνθήκες και απαιτήσεις του χρήστη.

1.3 Επίγνωση Πληροφορίας Πλαισίου

Η επίγνωση πληροφορίας πλαισίου ή επίγνωση πλαισίου (context awareness) μπορεί να οριστεί ως: «η ικανότητα ενός συστήματος να ανακαλύπτει, να ερμηνεύει, να συμπεραίνει, να αξιοποιεί και να συλλογίζεται βάσει της περιρρέουσας πληροφορίας ώστε να λαμβάνει αποφάσεις, να προβαίνει σε προκαθορισμένες ενέργειες και να προσαρμόζεται σε διάφορες καταστάσεις» [5].

Έτσι λοιπόν ένα σύστημα διάχυτου υπολογισμού που προσπαθεί να είναι ελάχιστα παρεισφρητικό πρέπει να είναι context-aware. Με άλλα λόγια, πρέπει να έχει γνώση της κατάστασης του χρήστη και του περιβάλλοντος χώρου ώστε να τροποποιεί κατάλληλα την συμπεριφορά του. Η πληροφορία που το επηρεάζει είναι αχανής και εκτείνεται από χωρικές και χρονικές παραμέτρους ως και παράγοντες που δεν είναι πάντα τόσο προφανείς αλλά μπορεί να αποδειχθούν κρίσιμοι. Οι πληροφορίες δε που αφορούν ένα χρήστη (πληροφορίες πλαισίου) μπορεί να είναι αρκετά πλούσιες και να αποτελούνται από ιδιότητες όπως φυσική θέση, φυσική κατάσταση (θερμοκρασία του σώματος και

παλμοί της καρδιάς), συναισθηματική κατάσταση (θυμωμένος ή ήρεμος), καθημερινή συμπεριφορά, κτλ. Εάν δινόταν αυτές οι πληροφορίες σε έναν άνθρωπο, θα λάμβανε αποφάσεις με προδραστικό τρόπο (proactivity), προλαμβάνοντας και προβλέποντας τις ανάγκες των χρηστών. Στη λήψη αυτών των αποφάσεων, ο άνθρωπος δεν θα ενοχλούσε το χρήστη εκτός και αν υπήρχε μια επείγουσα περίπτωση. Αν ένα σύστημα λοιπόν μπορέσει να φτάσει στο σημείο να «προσαρμόζεται» κάθε φορά στο περιβάλλον του κι ακόμη στους ίδιους τους χρήστες με αξιοπιστία και προνοητικότητα, τότε θα μπορούμε να μιλάμε για ένα πραγματικά εύχρηστο, χρήσιμο, κατά το δυνατόν διαθέσιμο και ανθρωποκεντρικό προϊόν τεχνολογίας.

Ένας πρωταρχικός στόχος για τα συστήματα διάχυτου υπολογισμού είναι η πρόσληψη και εκμετάλλευση της πληροφορίας που σχετίζεται με το πλαίσιο μιας συσκευής έτσι ώστε να παρέχονται υπηρεσίες κατάλληλες σε συγκεκριμένους χρήστες, καταστάσεις και γεγονότα. Η πληροφορία αυτή είναι αξιοποιήσιμη μόνο όταν μπορεί να ερμηνευθεί με χρήσιμο τρόπο. Αντικείμενο έρευνας είναι και το ζήτημα της «αναπαράστασης» (representation) και «ερμηνείας» (interpretation) της πληροφορίας πλαισίου. Η πληροφορία που συλλέγεται από πολλές (διαφορετικές) πηγές ερμηνεύεται διαφορετικά από κάθε «εφαρμογή επίγνωσης πλαισίου» (context-aware application). Επομένως, αναπτύσσονται μηχανισμοί μέσω των οποίων το πλαίσιο ελέγχεται και ερμηνεύεται κατάλληλα βάσει της μελλοντικής του χρήσης και αξιοποίησης.

Μεγάλο κομμάτι της έρευνας έχει ασχοληθεί με τη χρήση αισθητήρων, όπως ανιχνευτές θέσης χρήστη (εσωτερικών ή εξωτερικών χώρων), τοποθεσίας της κινητής συσκευής, φυσικών ή περιβαλλοντολογικών παραμέτρων (π.χ., θερμοκρασία, ταχύτητα κίνησης, πίεση, φωτεινότητα, κατεύθυνση, ήχος, υγρασία, ταχύτητα ανέμου) και κάμερες. Οι αισθητήρες θέσης παρέχουν πληροφορία που σχετίζεται με την τοποθεσία σαν ανεξάρτητο αλλά και ιδιαίτερα χρήσιμο δεδομένο. Συχνά, βέβαια, δεν είναι η τοποθεσία που συνιστά μόνη της άμεσα αξιοποιήσιμο δεδομένο αλλά σε συνδυασμό με επιπρόσθετη πληροφορία που μπορεί να παραχθεί (π.χ. οι πόροι που είναι διαθέσιμοι σε κάποιο σημείο). Επίσης, οι διαφορές μετρήσεις από αισθητήρες μπορούν να ορίσουν μια πιο σύνθετη αναπαράσταση και επίγνωση του πλαισίου (π.χ., η κατάσταση πυρκαγιάς επάγεται από μετρήσεις αισθητήρων καπνού, θερμοκρασίας και υγρασίας αλλά και οπτικών αισθητήρων).

Η δημιουργία λοιπόν ενός context-aware συστήματος απαιτεί την αντιμετώπιση πολλών ζητημάτων όπως:

- Πώς μοντελοποιούνται όλες αυτές οι πληροφορίες πλαισίου που αφορούν το χρήστη και το περιβάλλον; Πώς αυτές οι πληροφορίες συνδυάζονται με την κατάσταση του συστήματος και των εφαρμογών; Πού αποθηκεύονται οι πληροφορίες; Ποιες είναι οι σχετικές δομές δεδομένων και οι αλγόριθμοι επεξεργασίας τους;
- Πόσο συχνά πρέπει να συμβουλευόμαστε αυτές τις πληροφορίες; Ποιο είναι το κόστος αν τις λαμβάνουμε υπόψη; Ποιες τεχνικές μπορεί κάποιος να χρησιμοποιήσει για να ελαχιστοποιήσει αυτό το κόστος;
- Ποιες είναι οι ελάχιστες υπηρεσίες που ένα περιβάλλον πρέπει να παρέχει ώστε να θεωρείται context-aware; Είναι οι ιστορικές πληροφορίες χρήσιμες;
- Πως μπορούν να διαδοθούν οι αποδοτικά οι πληροφορίες αυτές ώστε να ενημερωθούν γειτονικοί κόμβοι;

1.4 Συνεργατική Επίγνωση Πληροφορίας Πλαισίου

Η συνεργατική επίγνωση πληροφορίας πλαισίου (collaborative context awareness) αποτελεί την κατανεμημένη εκδοχή της επίγνωσης πλαισίου και έχει μεγάλη σημασία για τα περιβάλλοντα διάχυτου υπολογισμού. Γενικά, οι χρήστες συγκεντρώνονται σε ομάδες, για παράδειγμα σε μουσεία, σε αίθουσες διδασκαλίας, σε συσκέψεις, κτλ. Τα άτομα αυτά μοιράζονται (τουλάχιστον μέσα σε όμοια χρονικά πλαίσια) κοινά ενδιαφέροντα και προτιμήσεις (π.χ., άτομα που ενδιαφέρονται για την ίδια έκθεση σε ένα μουσείο). Τα μέλη μιας ομάδας λοιπόν βιώνουν παρόμοιες καταστάσεις καθώς πολλοί από τους φορητούς υπολογιστές / κόμβους (nodes) που φέρουν ανιχνεύουν, εντοπίζουν και περιεργάζονται πανομοιότυπη πληροφορία πλαισίου (π.χ., θέση, υγρασία, θερμοκρασία, περιβαλλοντικές συνθήκες) [6]. Οι κόμβοι που κινούνται προς διάφορες θέσεις μπορούν σε συνεργασία να διαδώσουν αυτήν την πληροφορία. Έτσι κάποιος μπορεί να συλλέξει την πληροφορία πλαισίου από τους γειτονικούς κόμβους όχι μόνο για τον εαυτό του, αλλά και για την μοιραστεί περαιτέρω με άλλους κόμβους που θα βρεθούν στην εμβέλειά του αργότερα. Προκύπτει λοιπόν η συνεργατική επίγνωση πληροφορίας πλαισίου, που υποδηλώνει την κατανόηση της πληροφορίας πλαισίου από όλα τα μέλη μιας ομάδας παρέχοντας κατά συνέπεια μια πιο αναβαθμισμένη πληροφορία πλαισίου για κάθε μέλος της ίδιας ομάδας ξεχωριστά [7],[8]. Η συνεργατική πληροφορία πλαισίου είναι η πληροφορία που αποκτάται μέσω του δικτύου (π.χ. εγγύτητα) μεταξύ των αισθητήρων και των συσκευών υψηλότερου

επιπέδου και μπορεί να χρησιμοποιηθεί για να αυξήσει την κοινή κατανόηση για τη γειτονιά κάθε κόμβου, να βελτιώσει τη διαθεσιμότητα των προσφερόμενων υπηρεσιών, ή να ενισχύσει την αξιοπιστία της πληροφορίας μέσω των πρόσθετων πληροφοριών από τους γείτονες.

1.5 Συνεισφορά Διδακτορικής Διατριβής

Η ερευνητική συνεισφορά της διατριβής μπορεί συνοψισθεί σε δύο βασικές θεματικές ενότητες:

1. Αναπαράσταση, μοντελοποίηση και συμπερασμός πλαισίου.
2. Συνεργατική επίγνωση πλαισίου.

Πιο συγκεκριμένα στην πρώτη θεματική ενότητα παρουσιάζεται ένας μηχανισμός συμπερασμού της κατάστασης του χρήστη που αποτελεί βασικό παράγοντα για τις εφαρμογές και τις προσφερόμενες υπηρεσίες σε περιβάλλοντα διάχυτου υπολογισμού (ΚΕΦΑΛΑΙΟ 2). Η κατάσταση του χρήστη συνήθως μεταβάλλεται με το χρόνο και οι πληροφορίες του παρελθόντος μπορούν να επηρεάσουν άμεσα την παρούσα κατάσταση. Στα πλαίσια της έρευνας για την επίγνωση πλαισίου προτείνουμε μια αρχιτεκτονική πιθανο-θεωρητικής σύντηξης δεδομένων που χρησιμοποιεί τις μετρήσεις από αισθητήρες διαφορετικών τεχνολογιών (Wi-Fi, IR beacons, κ.τ.λ.). Το σύστημα βασίζεται σε μία στρωματοποιημένη αρχιτεκτονική επιδιώκοντας την διευκόλυνση στην ενσωμάτωση νέων αισθητήρων ώστε να αυξηθεί η απόδοσή του. Για την μοντελοποίηση της χρονικής μεταβολής αλλά και της χρονικής εξάρτησης του πλαισίου η προτεινόμενη αρχιτεκτονική βασίζεται στη χρήση των Δυναμικών Δικτύων Bayes, δίνοντας έτσι τη δυνατότητα ενσωμάτωσης στη διαδικασία της σύντηξης και ιστορικών δεδομένων. Το προτεινόμενο σχήμα αξιολογείται ως προς τον προσδιορισμό χωρικού πλαισίου (θέση του χρήστη).

Ως επέκταση του προηγούμενου συστήματος προτείνεται μια καινοτόμος αρχιτεκτονική που διαμορφώνει και μοντελοποιεί τα δεδομένα από τους αισθητήρες καθώς επίσης αποφασίζει και συμπεραίνει για την τρέχουσα κατάσταση του χρήστη. Τα δεδομένα αυτά είναι συνήθως ανακριβή (π.χ., θόρυβος των μετρήσεων των αισθητήρων). Το σύστημα που βασίζεται στα Δυναμικά Δίκτυα Bayes (DBNs) και την Θεωρία Ασαφών Συνόλων (Fuzzy Set Theory) λαμβάνει υπόψη του και την αξιοπιστία των πηγών της πληροφορίας πλαισίου την οποία ενσωματώνει στον συμπερασμό για την κατάσταση του χρήστη.

Εν συνεχεία, στο δεύτερο μέρος της πρώτης ενότητας (ΚΕΦΑΛΑΙΟ 3) εξετάζεται το θέμα της σύντηξης πολλαπλών αισθητήρων βάση της θεωρίας ενδείξεων (evidence theory) αξιοποιώντας δεδομένα και μετρήσεις που έχουν ήδη επεξεργαστεί. Η προσέγγιση που ακολουθείται είναι μία δι-επίπεδη αρχιτεκτονική σύντηξης. Στο πρώτο επίπεδο ανεπεξέργαστα δεδομένα από τους αισθητήρες συνδυάζονται με την τεχνική του σωρευτικού αθροίσματος (CUSUM) για την ανίχνευση αλλαγής στην κατανομή τους. Στο δεύτερο επίπεδο με τη βοήθεια της θεωρίας που θεμελίωσαν οι Dempster και Shafer, συντήκεται η πληροφορία πλαισίου που προέρχεται από διαφορετικούς και ανεξάρτητους ειδικούς του πρώτου επιπέδου. Το σύστημα αξιολογήθηκε για τον συμπερασμό σχετικά με την εμφάνιση πυρκαγιών σε μια γεωγραφική περιοχή η οποία καλύπτεται από ένα δίκτυο αισθητήρων θερμοκρασίας, υγρασίας αλλά και οπτικών αισθητήρων. Η υιοθέτηση του δι-επίπεδου σχήματος έχει το πλεονέκτημα της έγκαιρης ανίχνευσης πυρκαγιών ενώ παράλληλα μειώνει την πιθανότητα λανθασμένης ανίχνευσης (false alarm).

Στην δεύτερη θεματική ενότητα της διατριβής εξετάστηκαν θέματα που αφορούν τη συνεργατική επίγνωση πλαισίου. Στο πρώτο μέρος της ενότητας παρουσιάζεται ένας μηχανισμός που βασίζεται στη βιο-μιμητική διάχυση της πληροφορίας μεταξύ των κόμβων ενός ασύρματου δικτύου (ΚΕΦΑΛΑΙΟ 4). Μία σημαντική παράμετρος στα ασύρματα δίκτυα είναι η ενέργεια των κόμβων η οποία μειώνεται με τον αριθμό των μεταδόσεων. Υιοθετείται το επιδημικό μοντέλο (epidemic model) ως βασικός μηχανισμός για μετάδοση πληροφορίας πλαισίου. Υπάρχει μία αναλογία μεταξύ της διάδοσης του πλαισίου και της επιδημίας ενός ιού. Η στοχαστική φύση του επιδημικού μοντέλου δεν εγγυάται την πλήρη κάλυψη του χώρου με πληροφορία πλαισίου συγκρινόμενο με την τεχνική της πλημμύρας (flooding). Όμως, δείχνεται ότι με την επιδημική διάδοση πληροφορίας επιτυγχάνεται μια ικανοποιητική κάλυψη του χώρου ακόμη και με χαμηλούς ρυθμούς μετάδοσης. Η πιθανοτική αυτή προσέγγιση (δηλαδή, πιθανότητα διάδοσης ενός μηνύματος μεταξύ δύο κόμβων) οδηγεί στην εξοικονόμηση ενέργειας σε σχέση με την τεχνική της πλημμύρας. Στα πλαίσια της διατριβής μελετήθηκαν προβλήματα που σχετίζονται με την επιδημική διάδοση της πληροφορίας και προτάθηκαν αρχιτεκτονικές που λαμβάνουν υπόψη την τοπολογία του δικτύου, την κινητικότητα και την πυκνότητα των κόμβων ούτως ώστε να καθορίσουν την πιθανότητα διάδοσης της πληροφορίας μεταξύ των κόμβων (ρυθμός «μόλυνσης»). Επιπλέον, παρουσιάζεται ένα καινοτόμο σχήμα προσαρμογής του ρυθμού «μόλυνσης» ανεξάρτητα σε κάθε κόμβο με βάση τις αλλαγές που εμφανίζονται στις τιμές της πληροφορίας πλαισίου που δειγματοληπτούνται. Υψηλή μεταβλητότητα της

πληροφορίας πλαισίου οδηγεί σε αύξηση του ρυθμού «μόλυνσης» της γειτονιάς του κόμβου που εκτελεί την δειγματοληψία. Αντιθέτως, η χαμηλή μεταβλητότητα διατηρεί τον ρυθμό μόλυνσης σε χαμηλά επίπεδα. Το σχήμα αξιολογήθηκε όσον αφορά το σφάλμα της διαδεδομένης πληροφορίας που υπεισέρχεται στον πιο απομακρυσμένο κόμβο του δικτύου και τον αριθμό των μηνυμάτων που εξέπεμψε η πηγή. Το σχήμα αυτό στη συνέχεια επεκτείνεται προσαρμόζοντας πλέον το ρυθμό μόλυνσης όχι μόνο με βάση τη μεταβλητότητα αλλά και την χρονική εγκυρότητα του πλαισίου. Έτσι οι παρωχημένες πληροφορίες πλαισίου δεν υπάρχει λόγος να μεταδίδονται, εξοικονομώντας επιπλέον ενέργεια αλλά και αυξάνοντας τη «χρονική ποιότητα» του πλαισίου που υπάρχει στο δίκτυο. Τα πειράματα πραγματοποιήθηκαν με πραγματικά δεδομένα θερμοκρασίας και ταχύτητας ανέμου που διαφέρουν σημαντικά ως προς τη μεταβλητότητά τους. Τα αποτελέσματα των πειραμάτων οδηγούν στο συμπέρασμα ότι η προσαρμοστική διάχυση πληροφορίας πλαισίου συντελεί σημαντικά στην εξοικονόμηση ενέργειας των κόμβων του δικτύου αλλά και στην μεγιστοποίηση της «χρονικής ποιότητας» του πλαισίου.

Στο δεύτερο μέρος της ενότητας (ΚΕΦΑΛΑΙΟ 5) παρουσιάζεται ένα σύστημα συμπερασμού της αξιοπιστίας (trust) των κόμβων ενός ασύρματου δικτύου. Η συνεργασία μεταξύ των κόμβων και η ανταλλαγή πληροφοριών πλαισίου αποτελεί σημαντικό παράγοντα στην εύρυθμη λειτουργία του δικτύου. Η προτεινόμενη κατανομημένη αρχιτεκτονική (Ad-hoc Trust Framework) συμπεραίνει την αξιοπιστία των κόμβων ως προς διάφορες υπηρεσίες που αυτοί προσφέρουν (προώθηση πακέτων, δρομολόγηση, κτλ.). Ενσωματώνει πληροφορίες πλαισίου σχετικά με την συμπεριφορά των κόμβων, τις υποδείξεις ορισμένων κόμβων προς άλλους και την ιστορία των ενεργειών των κόμβων. Ειδικά οι υποδείξεις λαμβάνονται μόνο από κόμβους με αυξημένη αξιοπιστία ελαχιστοποιώντας έτσι το κόστος επικοινωνίας με πολλούς κόμβους. Η αρχιτεκτονική του συστήματος επιτρέπει σε έναν κόμβο που θεωρείται μη αξιόπιστος («απομονωμένος») να αρχίσει να συνεργάζεται έτσι ώστε να αυξήσει την αξιοπιστία του. Η αξιολόγηση του συστήματος έδειξε ότι οι ιδιοτελείς κόμβοι (selfish) εντοπίζονται γρήγορα και με ακρίβεια.

Τέλος, στο ΚΕΦΑΛΑΙΟ 6 αναφέρεται η βασικότερη σχετική έρευνα καθώς και η σχετική σύγκριση με όλα τα ερευνητικά πεδία που εστίασε η διδακτορική διατριβή ενώ στο ΚΕΦΑΛΑΙΟ 7 συνοψίζεται η ερευνητική συνεισφορά της διατριβής και αποτυπώνονται τα ζητήματα που αντιμετωπίστηκαν και οι δοθείσες λύσεις. Παράλληλα, παρουσιάζονται οι βασικότερες μελλοντικές κατευθύνσεις της έρευνας που

Διαχείριση Πληροφορίας Πλαισίου για το Διάχυτο Υπολογισμό

διενεργήθηκε κατά την εκπόνηση της διατριβής και τα θέματα που μπορούν να προκύψουν με αφετηρία κάθε ένα από τα επιμέρους προβλήματα που μελετήθηκαν.

ΚΕΦΑΛΑΙΟ 2

ΕΠΙΓΝΩΣΗ ΤΗΣ ΚΑΤΑΣΤΑΣΗΣ ΤΟΥ ΧΡΗΣΤΗ

2.1 Εισαγωγή

Στο διάχυτο υπολογισμό η κατάσταση (situation) που βρίσκεται ο χρήστης αποτελεί ίσως τη σημαντικότερη πληροφορία πλαισίου. Η επίγνωση της κατάστασης αυτής (θέση χρήστη, ενέργειες, δραστηριότητα, κτλ.) δίνει τη δυνατότητα σε ένα σύστημα διάχυτου υπολογισμού να παρέχει υπηρεσίες οι οποίες χάνονται στο παρασκήνιο και συνεπικουρούν το χρήστη. Στο κεφάλαιο αυτό μελετάμε θέματα σχετικά με την σύντηξη πληροφορίας πλαισίου (context fusion) για την επίγνωση της κατάστασης του χρήστη. Η πληροφορία αυτή εξάγεται από διάφορους αισθητήρες, ιστορικά στοιχεία ή ακόμα και από τις ίδιες τις ενέργειες του χρήστη. Ουσιαστικά η σύντηξη πρόκειται για μια διαδικασία συμπερασμού πλαισίου (context reasoning).

Η κατάσταση του χρήστη συνήθως μεταβάλλεται με το χρόνο και οι πληροφορίες του παρελθόντος μπορούν να επηρεάσουν άμεσα την παρούσα κατάσταση. Στα πλαίσια της έρευνας για την επίγνωση πλαισίου στα περιβάλλοντα διάχυτου υπολογισμού προτείνουμε μια αρχιτεκτονική πιθανο-θεωρητικής σύντηξης δεδομένων που χρησιμοποιεί τις μετρήσεις από αισθητήρες διαφορετικών τεχνολογιών. Για την μοντελοποίηση της χρονικής μεταβολής αλλά και της χρονικής εξάρτησης του πλαισίου η προτεινόμενη αρχιτεκτονική βασίζεται στη χρήση των Δυναμικών Δικτύων Bayes, δίνοντας έτσι τη δυνατότητα ενσωμάτωσης στη διαδικασία της σύντηξης και ιστορικών δεδομένων. Το προτεινόμενο σχήμα αξιολογείται ως προς τον προσδιορισμό χωρικού πλαισίου (θέση του χρήστη).

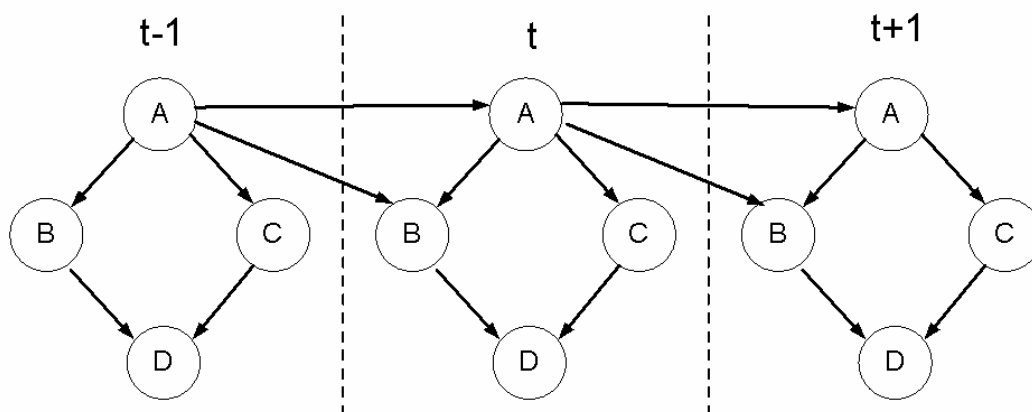
Στην συνέχεια προτείνεται η επέκταση της συγκεκριμένης αρχιτεκτονικής όπου μαζί με τα δεδομένα των αισθητήρων και το προκύπτον πλαίσιο (inferred context) συμπεριλαμβάνεται στη διαδικασία σύντηξης και η αξιοπιστία των πηγών (αισθητήρων, κτλ.). Αυτό επιτυγχάνεται μέσω της Θεωρίας των Ασαφών Συνόλων. Επιπλέον, εισήχθη και μία μέθοδος για τον υπολογισμό του βαθμού αξιοπιστίας των πηγών που θα παρουσιαστεί αναλυτικά στις παρακάτω υπο-ενότητες. Σύμφωνα με αυτή τη μέθοδο, που βασίζεται σε στατιστικά μεγέθη υπολογίζουμε την αξιοπιστία κάθε αισθητήρα χρησιμοποιώντας τις πραγματικές πληροφορίες που μας παρέχει. Αν οι μετρήσεις από τον αισθητήρα παρέχουν «σημαντική» πληροφορία τότε αυτός θεωρείται αξιόπιστος. Εισάγουμε ουσιαστικά μία έννοια παρόμοια με την εντροπία μιας πηγής.

2.2 Πιθανο-θεωρητική Σύντηξη

Η επίγνωση πλαισίου είναι σύνθετη διεργασία που απαιτεί σωστή αναπαράσταση της διαθέσιμης πληροφορίας. Τα περιβάλλοντα διάχυτου υπολογισμού είναι δυναμικά με συχνές αλλαγές των συνθηκών που επηρεάζουν άμεσα τους χρήστες. Έτσι, απαιτείται ένας μηχανισμός αναπαράστασης της χρονικής μεταβολής αλλά και της χρονικής εξάρτησης του πλαισίου. Ο μηχανισμός που προτείνεται και αναλύεται στις επόμενες ενότητες βασίζεται στα Δυναμικά Δίκτυα Bayes και χρησιμοποιείται για τον συμπερασμό της κατάστασης του χρήστη.

Τα Δίκτυα Bayes (Bayesian Networks) είναι ένα στατιστικό εργαλείο που έχει γίνει πολύ δημοφιλές σε περιοχές όπως η πιθανοτική μάθηση και η τεχνητή νοημοσύνη ([10],[11]). Είναι κατάλληλα για πρόβλεψη (prediction) και συμπερασμό (inference) διότι μοντελοποιούν αιτιώδης επιδράσεις (causal influences) μεταξύ τυχαίων μεταβλητών. Αποτελούνται από έναν κατευθυνόμενο ακυκλικό γράφο. Οι κόμβοι (nodes) του γράφου αναπαριστούν μεταβλητές (π.χ., τυχαίες συνεχείς ή διακεκριμένες μεταβλητές), ενώ οι σύνδεσμοι (arcs) που ενώνουν τους κόμβους αναπαριστούν μια πιθανοτική σχέση μεταξύ αυτών των μεταβλητών. Πλεονεκτούν από άποψη απόδοσης σε σχέση με τις παραδοσιακές πιθανοτικές μεθόδους διότι δεν είναι απαραίτητοι οι σύνθετοι υπολογισμοί για να ληφθεί μία απόφαση. Μπορούν λοιπόν να χρησιμοποιηθούν και για την επίγνωση πλαισίου σε περιβάλλοντα διάχυτου υπολογισμού. Περισσότερες πληροφορίες σχετικά με τα Δίκτυα Bayes παρέχονται στο ΠΑΡΑΡΤΗΜΑ.

Τα Δυναμικά Δίκτυα Bayes (ΔΔΒ) είναι μία ειδική κατηγορία Δικτύων Bayes (ΔΒ) που απεικονίζουν ένα προσωρινό μοντέλο πιθανοτήτων και επαναλαμβάνονται στο χρόνο. (Σχήμα 2.1). Αν παρατηρήσουμε μεμονωμένα ένα ΔΔΒ σε μια χρονική στιγμή t αυτό αποτελεί ένα απλό ΔΒ.



Σχήμα 2.1: Δυναμικό Δίκτυο Bayes (εξέλιξη στο χρόνο).

Οι τυχαίες μεταβλητές (κόμβοι) ενός ΔΔΒ σε μια χρονική στιγμή t επηρεάζονται άμεσα και από μεταβλητές προηγούμενων χρονικών στιγμών $(t-1, t-2, \dots)$. Συνήθως όμως για απλότητα υποθέτουμε ότι οι γονείς ενός κόμβου είναι στην ίδια ή στην προηγούμενη χρονοθυρίδα $(t-1)$ (αλυσίδα Markov 1^{ης} τάξης). Στο δίκτυο που απεικονίζεται στο Σχήμα 2.1 η τυχαία μεταβλητή A τη χρονική στιγμή $t-1$ είναι άμεσος πρόγονος των τυχαίων μεταβλητών A και B τη χρονική στιγμή t . Γενικά ότι ισχύει για τα απλά (στατικά) Δίκτυα Bayes ισχύει και για τα Δυναμικά Δίκτυα Bayes με τη μόνη διαφορά ότι στα τελευταία υπάρχει χρονική εξέλιξη.

Στον προτεινόμενο μηχανισμό για την επίγνωση της κατάστασης του χρήστη στα περιβάλλοντα διάχυτου υπολογισμού υιοθετούμε τα ΔΔΒ. Οι τυχαίες μεταβλητές για το ΔΔΒ τώρα θα έχουν την παρακάτω μορφή:

- βασική συνιστώσα πλαισίου $p(0) \in \mathbf{P}(0)$ (π.χ., μετρήσεις αισθητήρων) και
- πλαίσιο $p(k) \in \mathbf{P}(k)$ με επίπεδο αφαίρεσης $k > 0$.

Το πλαίσιο $p(k)$ την χρονική στιγμή t μπορεί να επηρεάσει (όσον αφορά την πιθανότητα ενδεχομένου) σε ένα ΔΔΒ το πλαίσιο $p(m)$ με επίπεδο αφαίρεσης $m < k$ καθώς και την συνιστώσα $p(0)$ όπως απεικονίζεται στο Σχήμα 2.2.

Σε ένα ΔΔΒ η πιθανότητα ένα πλαίσιο p να έχει την τιμή $v(t)$ την χρονική στιγμή t εξαρτάται

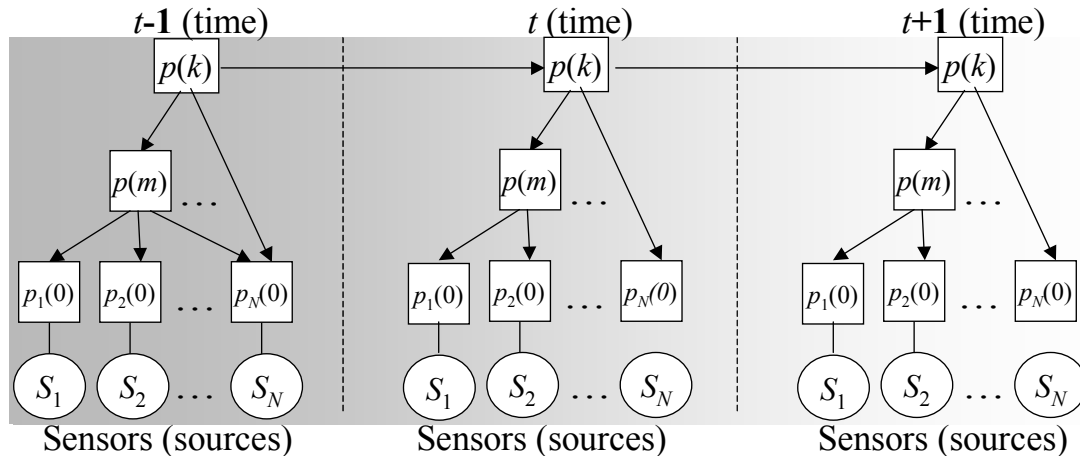
- από την πιθανότητα το ίδιο πλαίσιο p να έχει την τιμή $v(t-1)$ την χρονική στιγμή $t-1$ και
- από την πιθανότητα οι εξαρτημένες συνιστώσες του πλαισίου (χαμηλότερου επιπέδου πλαίσια που συμπεραίνουν το πλαίσιο p) να έχουν τιμή $v_i(t)$ την χρονική στιγμή t .

Για κάθε πηγή (π.χ. αισθητήρα) S_i εκτιμούμε την κατανομή πιθανότητας

$$P(p_i(0) \text{ is } v_i | p_j(k) \text{ is } v_j), p_j(k) \in \mathbf{P}(k) \text{ με } k > 0.$$

Επίσης, για κάθε πλαίσιο, στο ΔΔΒ, η κατανομή πιθανότητας είναι

$$P(p_i(m) \text{ is } v_i | p_j(k) \text{ is } v_j), p_i(k) \in \mathbf{P}(k) \text{ και } p_j(m) \in \mathbf{P}(m) \text{ με } k > m.$$



Σχήμα 2.2: Δυναμικό Δίκτυο Bayes για τον συμπερασμό καταστάσεων (πλαίσιο).

Το αποτέλεσμα της πιθανο-θεωρητικής σύντηξης αναφέρεται στην τιμή του $p = p(k, t)$ όπου μεγιστοποιεί την συνδυασμένη πιθανότητα όπως αναφέρεται στη σχέση (2.1), όπου N_k είναι ο αριθμός των συνιστωσών που επάγουν το πλαίσιο επιπέδου k .

$$p^*(k, t) = \operatorname{argmax}_{i \in N_k} \{ \operatorname{Prob}(p_i(k, t) | p(k, t-1), p(k-1, t), p(k-2, t), \dots, p(0, t)) \} \quad (2.1)$$

2.2.1 Προσδιορισμός χωρικής πληροφορίας πλαισίου (θέσης) του χρήστη

Στην ενότητα αυτή θα παρουσιάσουμε την εφαρμογή του μηχανισμού συμπερασμού πλαισίου που βασίζεται στα ΔΔΒ για τον εντοπισμό του χρήστη σε ένα περιβάλλον διάχυτου υπολογισμού. Το σύστημα χρησιμοποιεί πολλαπλές τεχνολογίες και τεχνικές σύντηξης συνδυάζοντας δεδομένα από ετερογενείς αισθητήρες. Εκτός από αυτά τα δεδομένα χρησιμοποιεί επίσης και ιστορικά δεδομένα κίνησης του χρήστη. Όλες οι πληροφορίες συνδυάζονται με τη χρήση των ΔΔΒ.

Ένα σύστημα εντοπισμού χαρακτηρίζεται από την ορθότητα (accuracy) και την ακρίβεια ή διαθεσιμότητα (precision or availability) που μπορεί να προσφέρει. Θα πρέπει να αναφέρει τις θέσεις σωστά από μέτρηση σε μέτρηση.

- η μετρική της ορθότητας δηλώνει την απόσταση μέσα στην οποία το σύστημα έχει την ικανότητα να εντοπίσει τον χρήστη (π.χ., σε ακτίνα 1-5 μέτρα), και,
- η μετρική της ακρίβειας ή διαθεσιμότητας δηλώνει το ποσοστό του χρόνου που το σύστημα παρέχει συγκεκριμένη ορθότητα (π.χ., 80% του χρόνου το σύστημα παρέχει ορθότητα σε ακτίνα 1-5 μέτρων).

Προφανώς, αν δεχτούμε λιγότερη ορθότητα, μπορούμε να έχουμε περισσότερη ακρίβεια.

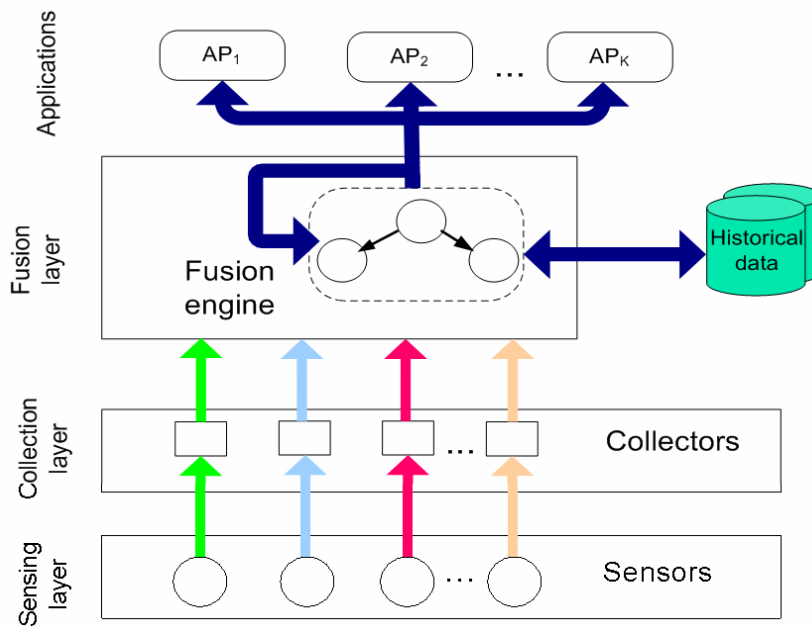
Η αρχιτεκτονική του συστήματος προσδιορισμού θέσης απεικονίζεται στο Σχήμα 2.3. Όπως φαίνεται, διαιρείται σε τρία επίπεδα (στρώματα):

1. το επίπεδο των αισθητήρων,
2. το επίπεδο της συλλογής των δεδομένων και
3. το επίπεδο της σύντηξης των δεδομένων.

Στο Σχήμα 2.3 απεικονίζονται επίσης οι εφαρμογές υπηρεσιών θέσης (location-aware applications) που εκμεταλλεύονται την πληροφορία θέσης των χρηστών με σκοπό να τους παράσχουν υπηρεσίες (location based services-LBS). Έχοντας μία τέτοια πληροφορία οι εφαρμογές αυτές μπορούν να εκτελούν εργασίες στο παρασκήνιο με σκοπό πάντα να ελαχιστοποιήσουν την απόσπαση της προσοχής των χρηστών (η αρχή της αορατότητας του διάχυτου υπολογισμού).

Οι βάσεις δεδομένων που διακρίνονται στην εικόνα αποτελούν μέσα αποθήκευσης του προσωπικού προφίλ των χρηστών και των ιστορικών στοιχείων (δεδομένων) της συμπεριφοράς τους. Ένα σύστημα διάχυτου υπολογισμού πρέπει συχνά να έχει πρόσβαση στο προφίλ και τις προτιμήσεις ενός χρήστη προκειμένου να παρασχεθούν υπηρεσίες και πληροφορίες που προσαρμόζονται σ' αυτόν. Κάθε χρήστης χαρακτηρίζεται από μερικές ιδιότητες (π.χ. όνομα, ηλικία, φύλο, e-mail, κ.λ.π.) και έχει διάφορες προτιμήσεις. Τα ιστορικά δεδομένα της κίνησης των χρηστών είναι επίσης πολύ σημαντικό στοιχείο για τον προσδιορισμό της παρούσας θέσης του.

Η στρωματοποιημένη προσέγγιση της αρχιτεκτονικής στοχεύει στη διευκόλυνση της προσθήκης νέων στοιχείων (access points, φάρων, κτλ.). Είναι προφανές ότι η προσθήκη παραπάνω στοιχείων βελτιώνει την ορθότητα και την ακρίβεια που μπορεί να προσφέρει το σύστημα. Στις επόμενες ενότητες παρέχουμε μια λεπτομερέστερη περιγραφή της αρχιτεκτονικής και εξετάζουμε αναλυτικά το ρόλο του κάθε επιπέδου.



Σχήμα 2.3: Αρχιτεκτονική του συστήματος προσδιορισμού θέσης.

2.2.1.1 Επίπεδο αισθητήρων

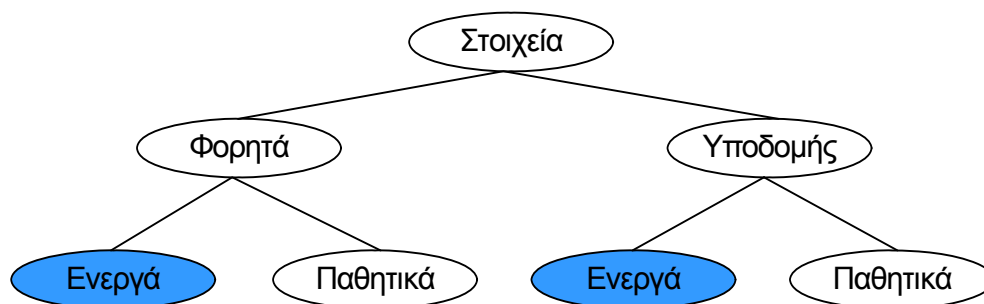
Το επίπεδο των αισθητήρων είναι το χαμηλότερο επίπεδο της αρχιτεκτονικής του συστήματος. Όπως το όνομά του υποδηλώνει, αυτό το επίπεδο είναι υπεύθυνο για την συλλογή των ανεπεξεργαστων δεδομένων (raw data) που θα χρησιμοποιηθούν στην διαδικασία προσδιορισμού θέσης των χρηστών. Σε αυτό το σημείο κρίνεται απαραίτητη μία κατηγοριοποίηση των συσκευών που συναντούμε σε εσωτερικά περιβάλλοντα (σταθμοί βάσης, φάροι, Radio Frequency Identification Tags - RFIDs, κλπ.) με σκοπό την κατανόηση της λειτουργίας τους αλλά και την ενσωμάτωσή τους στην προτεινόμενη αρχιτεκτονική.

Όλα τα στοιχεία (οι συσκευές) των διαφορετικών τεχνολογιών μπορούν να βρεθούν σε εσωτερικούς χώρους, είτε εγκατεστημένα στην κτιριακή υποδομή (building infrastructure), είτε προσαρτημένα σε κινητές συσκευές, είτε φερόμενα από τον ίδιο το χρήστη (π.χ. στα ρούχα του). Μερικά από αυτά εκπέμπουν πληροφορίες (δεδομένα) και άλλα ανιχνεύουν (διαβάζουν) πληροφορίες. Ανάλογα λοιπόν με τη θέση στην οποία βρίσκονται και τη λειτουργία τους τα στοιχεία αυτά μπορούν να ταξινομηθούν ως εξής:

1. Φορητά στοιχεία ονομάζονται αυτά που φέρονται από τους χρήστες (πάνω στα ρούχα τους) ή είναι προσαρτημένα στις κινητές συσκευές τους. Σε αυτή την κατηγορία εμπίπτουν οι ετικέτες RFID, οι προσαρμογείς (κάρτες) Wi-Fi, οι θύρες υπερύθρων, κλπ.

2. Στοιχεία υποδομής είναι αυτά που είναι εγκατεστημένα σε σταθερά (γνωστά σημεία) ενός κτιρίου. Σε αυτή την κατηγορία ανήκουν οι σταθμοί βάσης (access points), οι φάροι υπερύθρων (IR beacons), οι αναγνώστες RFID ετικετών, κλπ.
3. Ενεργά στοιχεία είναι αυτά που ανιχνεύουν ένα φαινόμενο ή κάνουν μία μέτρηση. Στοιχεία αυτής της κατηγορίας είναι οι αναγνώστες RFID ετικετών, οι κάρτες Wi-Fi, κλπ.
4. Παθητικά στοιχεία ονομάζονται τα στοιχεία που εκπέμπουν πληροφορίες οι οποίες ανιχνεύονται από τα ενεργά στοιχεία. Σε αυτή την κατηγορία ανήκουν οι σταθμοί βάσης Wi-Fi, οι φάροι υπερύθρων, οι Bluetooth ετικέτες, κλπ.

Το Σχήμα 2.4 παρουσιάζει σχηματικά την κατηγοριοποίηση που μόλις περιγράψαμε. Είναι εμφανές ότι ένα στοιχείο υποδομής παραδείγματος χάριν μπορεί να είναι ενεργό ή παθητικό. Το ίδιο συμβαίνει και με τα φορητά στοιχεία. Τα ενεργά στοιχεία είναι αυτά που αποκαλούνται και «αισθητήρες». Το χαμηλότερο επίπεδο λοιπόν της αρχιτεκτονικής του συστήματος αποτελείται από αισθητήρες, είτε αυτοί είναι φορητοί, είτε είναι εγκατεστημένοι στην υποδομή. Στη συνέχεια παρουσιάζουμε μερικούς από τους βασικούς αισθητήρες που χρησιμοποιούνται στην διαδικασία προσδιορισμού θέσης.



Σχήμα 2.4: Κατηγοριοποίηση στοιχείων.

Φορητοί αισθητήρες

Ένας προσαρμογέας (adapter) Wi-Fi, που παρέχει ασύρματη πρόσβαση στους χρήστες και είναι ενσωματωμένος σχεδόν σε όλους τους νέους φορητούς υπολογιστές, μπορεί να χρησιμοποιηθεί επίσης για να μετρήσει τη λαμβανόμενη ισχύ του σήματος (Received Signal Strength - RSS) από ένα σταθμό βάσης Wi-Fi (παθητικό στοιχείο υποδομής). Ένα τέτοιος αισθητήρας λοιπόν ρυθμίζεται κατάλληλα και ανά τακτά χρονικά διαστήματα

μπορεί επίσης να παρέχει δεδομένα που αποτελούνται από παρατηρήσεις (μετρήσεις) όλων των διαθέσιμων σταθμών βάσης που βρίσκονται στην περιοχή.

Ομοίως, η θύρα υπερύθρων (IR port) μιας φορητής συσκευής χρησιμοποιείται ως αναγνώστης (reader) για τις εκπομπές των φάρων υπερύθρων που είναι εγκατεστημένοι στο κτίριο. Οι φάροι εκπέμπουν το αναγνωριστικό τους. Ο αισθητήρας (θύρα υπερύθρων) σε αυτή την περίπτωση δεν μετράει κάποιο μέγεθος αλλά αντιλαμβάνεται ένα γεγονός (το γεγονός ότι βρίσκεται στην εγγύτητα ενός φάρου).

Για να εκτελεστούν οι παραπάνω διαδικασίες από τους αισθητήρες (μετρήσεις, κτλ) θα πρέπει αυτοί να συνοδεύονται απαραίτητα από το αντίστοιχο λογισμικό (drivers). Εκτός από την «ενσωματωμένη» έκδοση του προσαρμογέα Wi-Fi και της θύρας IR, αυτοί οι αισθητήρες θα μπορούσαν να προστεθούν επίσης στη συσκευή μέσω μιας κάρτας ή μέσω της θύρας USB παρέχοντας τις ίδιες πληροφορίες.

Αισθητήρες υποδομής

Σε αυτή την κατηγορία ανήκουν οι αναγνώστες ετικετών RFID οι οποίοι είναι εγκατεστημένοι στο κτίριο. Μπορούν να εντοπίζουν ετικέτες και να διαβάζουν το αναγνωριστικό τους όταν αυτές βρίσκονται κοντά (τεχνική της εγγύτητας). Οι χρήστες φέρουν πάνω τους τέτοιες ετικέτες (φορητά παθητικά στοιχεία) οι οποίες έχουν μοναδικό αναγνωριστικό που σχετίζεται με τον κάτοχο τους. Ένας τέτοιος αισθητήρας λοιπόν, αντιλαμβάνεται ένα γεγονός (παρουσία του χρήστη).

Διαφορετική είναι η λειτουργία μίας συσκευής υπερηχητικών σημάτων. Η συσκευή αυτή εκτελεί μετρήσεις αποστάσεων από γνωστά σημεία του χώρου, οπότε η πληροφορία που μας παρέχει δεν είναι ένα απλό γεγονός αλλά μία αριθμητική μέτρηση. Η εκτίμηση της απόστασης μπορεί να προκύψει είτε με την αντανάκλαση των υπερηχητικών σημάτων πάνω στο χρήστη, είτε με μία κατά την οποία και ο χρήστης φέρει πάνω του μία παρόμοια συσκευή υπερηχητικών σημάτων.

Επίσης, οι αναγνώστες Bluetooth που εντοπίζουν εκπομπές από ετικέτες Bluetooth (φορητά παθητικά στοιχεία) ανήκουν σε αυτή την κατηγορία. Η λειτουργία αυτών των αναγνωστών μπορεί να είναι πιο περίπλοκη καθώς μπορούν να κάνουν μετρήσεις λαμβανόμενης ισχύος (στη ζώνη των 2.4 GHz) αλλά και να αντιλαμβάνονται την παρουσία κάποια ετικέτας. Ανάλογα με τον τύπο και την κατασκευή τους οι συσκευές αυτές ρυθμίζονται κατάλληλα ώστε να παρέχουν πληροφορίες με τη μία ή την άλλη μορφή, ή ακόμα και με τις δύο.

2.2.1.2 Επίπεδο συλλογής δεδομένων

Το επίπεδο αυτό αποτελείται από συστατικά λογισμικού (software components) που ονομάζονται συλλέκτες (collectors). Για κάθε αισθητήρα υπάρχει και ο αντίστοιχος συλλέκτης. Έτσι έχουμε έναν συλλέκτη Wi-Fi, έναν Bluetooth συλλέκτη, κ.ο.κ. Ο κύριος ρόλος ενός τέτοιου συστατικού είναι η αλληλεπίδραση του με τον υποκείμενο αισθητήρα (ενεργό στοιχείο) για τον οποίο είναι υπεύθυνο. Η αλληλεπίδραση αυτή περιλαμβάνει τη συλλογή των διαθέσιμων μετρήσεων ή γεγονότων.

Οι αισθητήρες όπως αναφέρθηκε προηγουμένως μπορεί να παράγουν δεδομένα που εξαρτώνται από τον τύπο και τη λειτουργία τους. Όταν αυτά τα δεδομένα φτάνουν στο επίπεδο της συλλογής, ακολουθούν δύο διαδικασίες που έχουν σαν σκοπό την ενοποίηση τους σε μία κοινή μορφή για την διευκόλυνση της περαιτέρω επεξεργασίας τους. Οι διαδικασίες αυτές (*προεπεξεργασία δεδομένων* και *δημιουργία διανυσμάτων*) περιγράφονται αναλυτικά στις επόμενες δύο ενότητες.

Προεπεξεργασία δεδομένων

Ας υποθέσουμε ότι μία νέα μέτρηση σήματος φτάνει από τον προσαρμογέα Wi-Fi. Τότε ο αντίστοιχος συλλέκτης (Wi-Fi collector) κβαντίζει την ληφθείσα τιμή σε N διακριτές στάθμες:

$$(S_1, S_2 \dots S_N)$$

Αν για παράδειγμα η ισχύς από το σταθμό βάσης (access point) με αναγνωριστικό AP_2 είναι μεταξύ -70 dBm και -60 dBm τότε προσδίδεται σε αυτόν το σταθμό βάσης (παθητικό στοιχείο υποδομής) η τιμή " S_1 ".

Πρέπει να αναφέρουμε εδώ ότι ο αριθμός των σταθμών (levels) εξαρτάται από τα όρια εκπομπής του σταθμού βάσης και τις συνθήκες του περιβάλλοντος. Αναλόγως των συνθηκών μπορούμε να προσαρμόσουμε τον αριθμό των σταθμών αλλά και το εύρος τους ώστε να επιτύχουμε μία ικανοποιητική λειτουργία του αισθητήρα αλλά και να λαμβάνουμε δεδομένα που αντικατοπτρίζουν την πραγματική εικόνα της ισχύος εκπομπής.

Ας αναλύσουμε τώρα την λειτουργία ενός συλλέκτη IR Beacon (φάρου υπερύθρων). Κατά τη διαδικασία της προεπεξεργασίας λειτουργεί με έναν διαφορετικό και πιο απλό

τρόπο από τον συλλέκτη Wi-Fi. Οι δύο καταστάσεις που μπορεί να βρίσκεται ένας φάρος υπερύθρων είναι:

(Visible, Not_Visible)

δηλαδή ορατός και μη ορατός, ανάλογα με τη θέση που βρισκόμαστε. Ας υποθέσουμε ότι ένας δέκτης υπερύθρων (φορητό ενεργό στοιχείο) βρίσκεται στην εμβέλεια ενός φάρου με αναγνωριστικό *IRB3*. Η κατάσταση αυτή θα προξενήσει ένα γεγονός εγγύτητας (proximity event) που θα ανιχνευθεί και έτσι ο συλλέκτης θα προσδώσει στο φάρο την τιμή “*Visible*” που σημαίνει ότι είναι ορατός. Ένας συλλέκτης RFID αναγνωστών έχει παρόμοια λειτουργία σε αυτή τη φάση καθώς αυτοί οι αναγνώστες εντοπίζουν επίσης γεγονότα εγγύτητας (proximity events).

Σε αυτό το στάδιο μπορεί να εκτελεστεί και μία εξομάλυνση (smoothing) των τιμών που λαμβάνουμε από τους αισθητήρες. Οι ξαφνικές αλλαγές στις τιμές μπορεί να οφείλονται σε πολλούς παράγοντες όπως ο θόρυβος του περιβάλλοντος, η ταχύτητα μετακίνησης και ο προσανατολισμός του χρήστη, κάποια βλάβη στους αισθητήρες, κ.α. Ένα τέτοιο γεγονός μπορεί να προκαλέσει εσφαλμένα συμπεράσματα για το ποια πραγματικά είναι η τιμή της μέτρησης, κάτι το οποίο θα έχει άμεσο αντίκτυπο στην διαδικασία προσδιορισμού της θέσης του χρήστη. Αυτά τα δυσάρεστα αποτελέσματα μπορούν να ελαχιστοποιηθούν με την εφαρμογή ενός φίλτρου το οποίο επιτρέπει την εξομάλυνση των τιμών (π.χ. ένα φίλτρο Kalman).

Δημιουργία διανύσματος δεδομένων

Μετά την προεπεξεργασία των ακατέργαστων δεδομένων που φθάνουν από το επίπεδο των αισθητήρων, κάθε συλλέκτης (collector) σχηματίζει ένα διάνυσμα του τύπου:

διάνυσμα = (user_ID, IE_ID, value)

όπου το πεδίο *user_ID* είναι το μοναδικό αναγνωριστικό (identifier) ενός χρήστη, το πεδίο *IE_ID* είναι το μοναδικό χαρακτηριστικό ενός στοιχείου υποδομής και το πεδίο *value* είναι μία μέτρηση ή ένα γεγονός.

Για παράδειγμα, ένας συλλέκτης Wi-Fi μπορεί να σχηματίσει το παρακάτω διάνυσμα:

διάνυσμα1 = (userA, AP1, S1)

που υποδηλώνει ότι ο προσαρμογέας Wi-Fi (φορητό ενεργό στοιχείο) του κινητού υπολογιστή του χρήστη με αναγνωριστικό *userA*, μετράει την ισχύ του σήματος (RSS)

Διαχείριση Πληροφορίας Πλαισίου για το Διάχυτο Υπολογισμό

από το σταθμό βάσης με αναγνωριστικό *AP1* (παθητικό στοιχείο υποδομής) και μετά την προεπεξεργασία η κβαντισμένη τιμή του σήματος είναι *S1*.

Ένα πιθανό διάνυσμα που μπορεί να σχηματιστεί από τον συλλέκτη RFID αναγνωστών (RFID reader collector) είναι το εξής:

$$\text{διάνυσμα2} = (\text{userB}, \text{RFR1}, \text{Visible})$$

Η παραπάνω έκφραση υποδηλώνει ότι μία ετικέτα RFID την οποία φέρει ο χρήστης με αναγνωριστικό *userB* βρίσκεται στην εγγύτητα του αναγνώστη RFID με αναγνωριστικό *RFR1*.

Κάθε ένα από αυτά τα διανύσματα χρονοσημαίνεται (timestamped) ούτως ώστε να προωθηθούν στο παραπάνω επίπεδο. Η χρονοσήμανση έχει ιδιαίτερη σημασία και αποτελεί απαραίτητη προϋπόθεση για την ομαλή λειτουργία του συστήματος διότι όπως θα δείξουμε παρακάτω χωρίς αυτήν δεν θα ήταν δυνατή η αποτελεσματική σύντηξη των δεδομένων.

Παρατήρηση: Τα διανύσματα δεδομένων περιέχουν αναγνωριστικά μόνο των στοιχείων υποδομής (IE_ID) και όχι αναγνωριστικά των φορητών στοιχείων. Αυτό έχει να κάνει με το γεγονός ότι τα στοιχεία υποδομής (ενεργά ή παθητικά) βρίσκονται σε σταθερές και γνωστές θέσεις της κτιριακής υποδομής ενώ τα φορητά στοιχεία δεν έχουν μία γνωστή θέση καθώς κινούνται μαζί με τους χρήστες. Θα ήταν λοιπόν δύσκολη αλλά και ανούσια η μοντελοποίηση των φορητών στοιχείων.

Παράδειγμα: Υποθέτουμε ότι όλα τα διαθέσιμα στοιχεία υποδομής ενός κτιρίου (τα αναγνωριστικά τους) είναι τα παρακάτω:

- *AP1, AP2* (Σταθμοί βάσης Wi-Fi)
- *IRB1* (Υπέρυθρος φάρος)
- *RFR1, RFR2* (Αναγνώστες ετικετών RFID)
- *US1, US2* (Συσκευές υπερήχων)

Στον παρακάτω πίνακα παρουσιάζουμε μερικές πιθανές εισόδους και εξόδους των συλλεκτών. Θεωρούμε ότι έχουμε ενδείξεις (μετρήσεις) μόνο από τα παρακάτω στοιχεία:

$$\{AP1, AP2, IRB1, RFR1, US1\}$$

Όπως φαίνεται στον παρακάτω πίνακα από την έξοδο των συλλεκτών παραλείψαμε το αναγνωριστικό του χρήστη και την χρονοσήμανση για λόγους ευκολίας.

Πίνακας 2-1: Έξοδος των συλλεκτών μετά την προεπεξεργασία και το σχηματισμό των διανυσμάτων

| Συλλέκτης | Είσοδος | Έξοδος |
|-----------------------|----------------------------------|----------------------|
| Wi-Fi (Σταθμοί βάσης) | (AP1, -65 dBm) (AP2, -30 dBm) | AP1 = S1 AP2 = S3 |
| Φάροι (IR Beacon) | (IRB1, γεγονός εγγύτητας) | IRB1 = Visible |
| Αναγνώστες RFID | (RFR1, γεγονός εγγύτητας) | RFR1=Visible |
| Συσκευές υπερήχων | (US1, 8 μέτρα) | US1 = D1 |

Έτσι λοιπόν μέχρι εδώ γνωρίζουμε τις τιμές που σχετίζονται με κάθε χρήστη των στοιχείων υποδομής (ενεργών και παθητικών). Στην επόμενη ενότητα θα περιγράψουμε πώς αυτές οι τιμές χρησιμοποιούνται για τον προσδιορισμό της θέσης ενός χρήστη.

2.2.1.3 Επίπεδο σύντηξης δεδομένων (data fusion)

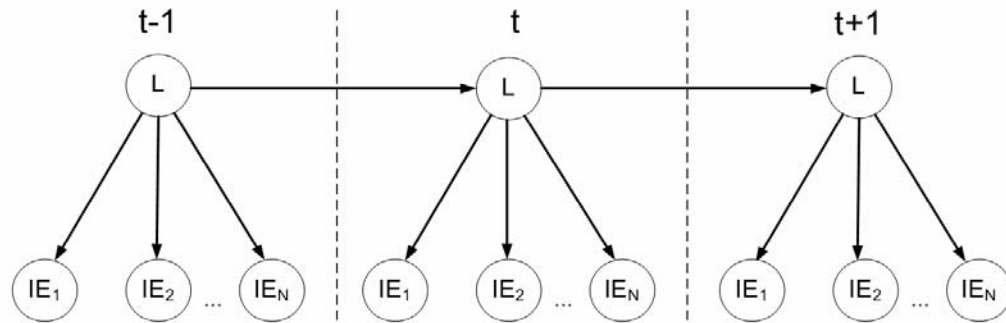
Αυτό το επίπεδο αποτελεί τον πυρήνα του συστήματος όπου η διαδικασία της σύντηξης των δεδομένων λαμβάνει χώρα. Περιέχει την μηχανή σύντηξης (fusion engine) ή αλλιώς συμπερασμού η οποία βασίζεται σε ένα Δυναμικό Δίκτυο Bayes (ΔΔΒ).

Ενσωμάτωση ενός ΔΔΒ στο σύστημα προσδιορισμού θέσης

Το ΔΔΒ που χρησιμοποιήθηκε στο σύστημα προσδιορισμού θέσης παρουσιάζεται στο Σχήμα 2.5. Όπως φαίνεται οι διακριτές τυχαίες μεταβλητές (κόμβοι) που χρησιμοποιήθηκαν για την κατασκευή του είναι:

- η θέση L του χρήστη, η οποία μπορεί να πάρει τιμές από ένα σύνολο K προκαθορισμένων θέσεων $\{L1, L2, \dots, LK\}$. Βέβαια, αυτό το σύνολο μπορεί να αποτελείται και από συμβολικές θέσεις π.χ. $\{room1, room2, corridorA, entrance, \dots\}$.

- τα N στοιχεία υποδομής IE_1, IE_2, \dots, IE_N . Το πεδίο ορισμού αυτών των τυχαίων μεταβλητών εξαρτάται από τον τύπο τους. Για παράδειγμα ένας σταθμός βάσης (access point) μπορεί να πάρει τιμές $\{S_1, S_2, \dots, S_N\}$ ανάλογα με την ισχύ του σήματος που λαμβάνουμε. Ένας φάρος υπερύθρων μπορεί να πάρει τιμές από το σύνολο $\{Visible, Not_Visible\}$ ανάλογα με το αν είμαστε μέσα στην εμβέλειά του ή όχι. Αντίστοιχες τιμές με τον φάρο λαμβάνει και ένας αναγνώστης ετικετών RFID.



Σχήμα 2.5: Δυναμικό Δίκτυο Bayes του συστήματος προσδιορισμού θέσης.

Στη συνέχεια θα περιγράψουμε αναλυτικά το δίκτυο της παραπάνω εικόνας, τις εξαρτήσεις μεταξύ των μεταβλητών και θα αναφέρουμε τους λόγους που καταλήξαμε σε μία τέτοια δομή.

Η τυχαία μεταβλητή L (θέση) τη χρονική στιγμή t , $L^{(t)}$, επηρεάζεται άμεσα από την τυχαία μεταβλητή L (θέση) τη χρονική στιγμή $t-1$, $L^{(t-1)}$. Έτσι η μεταβλητή $L^{(t-1)}$ είναι το αίτιο και η μεταβλητή $L^{(t)}$ είναι το αποτέλεσμα. Τα παραπάνω εκφράζονται μαθηματικά ως εξής:

$$L^{(t-1)} = \text{Parent}(L^{(t)})$$

Αυτό είναι λογικό καθώς η θέση ενός χρήστη εξαρτάται από την προηγούμενη θέση του (τη χρονική στιγμή $t-1$). Η πληροφορία αυτή μας δίνει τη δυνατότητα να αποκλείσουμε μερικές υποψήφιες θέσεις για το που μπορεί να βρίσκεται κάποιος, λόγω της αδυναμίας μετάβασης μεταξύ δύο θέσεων για ένα μέσο χρήστη σε συγκεκριμένο χρονικό διάστημα.

Επίσης οι τιμές που λαμβάνουμε από τους αισθητήρες τη χρονική στιγμή t εξαρτώνται από τη θέση $L^{(t)}$. Δηλαδή, η θέση μας επηρεάζει τις τιμές που μπορεί να έχουν τα

στοιχεία υποδομής. Για παράδειγμα η ισχύς του σήματος που μετράμε από ένα σταθμό βάσης, εξαρτάται από το πού βρισκόμαστε (κοντά, μακριά).

Οι πιθανοτικές κατανομές που σχετίζονται με κάθε κόμβο του δικτύου προσδιορίζονται με τεχνικές μάθησης (learning) Δικτύων Bayes. Συγκεκριμένα για κάθε στοιχείο υποδομής (IE_1, IE_2, \dots, IE_N) προσδιορίζουμε την πιθανοτική κατανομή $P(IE_i | L)$. Αυτό επιτυγχάνεται αν λάβουμε υπόψη τις σταθερές θέσεις των στοιχείων, τα μοντέλα διάδοσης των ραδιοκυμάτων και της υπέρυθρης ακτινοβολίας, το χρόνο διάδοσης και τις ανακλάσεις των υπερήχων, την εμβέλεια των συστημάτων RFID, κτλ.

Μία πιο απλή τεχνική μάθησης, η οποία ακολουθήθηκε, είναι η μέθοδος της δειγματοληψίας. Σε κάθε προκαθορισμένη θέση του χώρου $\{L_1, L_2, \dots, L_K\}$, εκτελούμε δειγματοληψία για κάποιο χρονικό διάστημα και προσδιορίζουμε τις τιμές των στοιχείων υποδομής λαμβάνοντας μετρήσεις με τη βοήθεια των αισθητήρων. Ανάλογα με την συχνότητα εμφάνισης των τιμών αυτών είμαστε σε θέση να δημιουργήσουμε τους πίνακες που απεικονίζουν τις κατανομές. Η τεχνική αυτή ονομάζεται επίσης και *Maximum Likelihood Estimation* [12].

Στον Πίνακα 2-2 παρουσιάζουμε μία τέτοια κατανομή που δημιουργήθηκε ύστερα από δειγματοληψία. Το στοιχείο υποδομής αποτελεί ένα σταθμό βάσης (access point) με (αναγνωριστικό AP_1) και όπως φαίνεται παρουσιάζονται οι πιθανότητες των τιμών που μπορεί να πάρει ανάλογα με την θέση του χρήστη μέσα στο κτίριο. Για παράδειγμα,

$$P(AP_1=S_1 | L=L_1) = 0.5$$

Η παραπάνω έκφραση σημαίνει ότι η πιθανότητα το AP_1 να έχει τιμή ισχύος σήματος (κβαντισμένη) S_1 , δεδομένου ο χρήστης βρίσκεται στην θέση L_1 , είναι 0.5.

Πίνακας 2-2: Πιθανοτική κατανομή τιμών για ένα σταθμό βάσης

| | L_1 | L_2 | ... |
|-------|-------|-------|-----|
| S_1 | 0.5 | 0.0 | ... |
| S_2 | 0.3 | 0.8 | ... |
| S_3 | 0.1 | 0.2 | ... |
| ... | ... | ... | ... |

Επιπλέον, οι πιθανοτικές κατανομές $P(L^t | L^{t-1})$ για την μετάβαση μεταξύ των θέσεων μέσα στο κτίριο δημιουργούνται αν λάβουμε υπόψη τη δομή του κτιρίου, την απόσταση μεταξύ των θέσεων και το χρόνο που κάνει ένας μέσος χρήστης για να διανύσει την κάθε απόσταση. Π.χ. είναι αδύνατο για ένα άτομο να διανύσει μία απόσταση 30 μέτρων από τη θέση A στη θέση B σε 2 μόνο δευτερόλεπτα. Άρα η πιθανότητα μετάβασης από τη θέση A στη θέση B είναι 0.

Ο Πίνακας 2-3 παρουσιάζει την πιθανοτική κατανομή μεταξύ των θέσεων ενός κτιρίου. Μπορούμε να διακρίνουμε ότι

$$P(L=L_2 | L=L_1) = 0.1$$

που σημαίνει ότι η πιθανότητα να βρισκόμαστε στη θέση L_2 δεδομένου ότι πριν βρισκόμασταν στη θέση L_1 είναι 0.1.

Πίνακας 2-3: Πιθανοτική κατανομή μετάβασης μεταξύ των θέσεων ενός κτιρίου

| | | | |
|-------|-------|-------|-----|
| | L_1 | L_2 | ... |
| L_1 | 0.5 | 0.0 | ... |
| L_2 | 0.1 | 0.8 | ... |
| L_3 | 0.1 | 0.05 | ... |
| ... | ... | ... | ... |

Επερωτήσεις προσδιορισμού θέσης

Όπως αναφέρθηκε τα Δίκτυα Bayes χρησιμοποιούνται εκτενώς σε διαδικασίες συμπερασμού καταστάσεων. Η κατασκευή λοιπόν του Δυναμικού Δικτύου Bayes που περιγράψαμε προηγουμένως μας δίνει τη δυνατότητα να εκτελέσουμε επερωτήσεις για να συμπεράνουμε (προσδιορίσουμε) τη θέση ενός χρήστη μέσα σε ένα κτίριο. Μία τέτοια επερώτηση θα μπορούσε να ήταν η εξής:

«Ποια είναι η θέση του χρήστη A δεδομένου της προηγούμενης του θέσης και δεδομένων των τιμών (παρατηρήσεων) των στοιχείων υποδομής;».

Για να απαντήσουμε σε αυτήν την επερώτηση υπολογίζουμε για κάθε μία από τις K θέσεις $\{L_1, L_2, \dots, L_K\}$ την ακόλουθη υπό συνθήκη πιθανότητα:

$$P(L^{(t)} | L^{(t-1)}, O^{(t)}) \quad (2.2)$$

που είναι η μαθηματική αναπαράσταση της παραπάνω επερώτησης και υποδηλώνει γενικώς την πιθανότητα να βρισकेσαι στη θέση $L^{(t)}$ τη χρονική στιγμή t (η ζητούμενη θέση) δεδομένου της ήδη γνωστής προηγούμενης θέσης $L^{(t-1)}$ και δεδομένων των τιμών (παρατηρήσεων) από τα N στοιχεία υποδομής τη χρονική στιγμή t , $O^{(t)}$. Για λόγους ευκολίας ορίζουμε:

$$O^{(t)} = \{IE_1^{(t)}, IE_2^{(t)}, \dots, IE_N^{(t)}\} \quad (2.3)$$

Γενικά η από κοινού πιθανότητα δύο ενδεχομένων E_1 και E_2 προσδιορίζεται από τη σχέση:

$$P(E_1, E_2) = P(E_1) * P(E_2 | E_1)$$

η οποία μπορεί να εκφραστεί και σαν:

$$P(E_2 | E_1) = \frac{P(E_1, E_2)}{P(E_1)} \quad (2.4)$$

Έτσι από τη σχέση (2.2) και λαμβάνοντας υπόψη την (2.4) έχουμε:

$$P(L^{(t)} | L^{(t-1)}, O^{(t)}) = \frac{P(L^{(t)}, L^{(t-1)}, O^{(t)})}{P(L^{(t-1)}, O^{(t)})} \quad (2.5)$$

Καθώς ο παρονομαστής της (2.5) δεν εξαρτάται από την τυχαία μεταβλητή $L^{(t)}$, μπορεί να θεωρηθεί σαν παράγοντας κανονικοποίησης. Μπορούμε να αντικαταστήσουμε τον παρονομαστή με $1/a$ και από την (4) θα λάβουμε:

$$P(L^{(t)} | L^{(t-1)}, O^{(t)}) = a * P(L^{(t)}, L^{(t-1)}, O^{(t)}) \quad (2.6)$$

Λαμβάνοντας επίσης υπόψη ότι κάθε μεταβλητή (κόμβος) του Δυναμικού Δικτύου Bayes που χρησιμοποιήσαμε στο σύστημά μας είναι υπό συνθήκη ανεξάρτητη δεδομένων των άμεσων προγόνων της, μπορούμε να υπολογίσουμε την από κοινού πιθανότητα (joint probability) που εμφανίζεται στο δεξιό μέρος της σχέσης (2.6). Άρα έχουμε,

$$P(L^{(t)}, L^{(t-1)}, O^{(t)}) = P(L^{(t)} | L^{(t-1)}) * P(O^{(t)} | L^{(t)}) \quad (2.7)$$

Αν τώρα συνδυάσουμε τις (2.5), (2.6) και (2.7) λαμβάνουμε την παρακάτω σχέση:

$$P(L^{(t)} | L^{(t-1)}, O^{(t)}) = a * P(L^{(t)} | L^{(t-1)}) * P(O^{(t)} | L^{(t)}) \quad (2.8)$$

Ο παράγοντας κανονικοποίησης a μπορεί να υπολογιστεί εύκολα καθώς όλες οι πιθανότητες (οι πιθανότητες για όλες τις K θέσεις) πρέπει να αθροίζονται στο 1. Οπότε είναι:

$$\sum_{i=1}^K P(L_i^{(t)} | L^{(t-1)}, O^{(t)}) = a * \sum_{i=1}^K P(L_i^{(t)} | L^{(t-1)}) * P(O^{(t)} | L_i^{(t)}) = 1 \quad (2.9)$$

Τελικά από τις (2.3), (2.8) και (2.9) λαμβάνουμε:

$$P(L^{(t)} | L^{(t-1)}, O^{(t)}) = \frac{P(L^{(t)} | L^{(t-1)}) * P(IE_1^{(t)} | L^{(t)}) * \dots * P(IE_N^{(t)} | L^{(t)})}{\sum_{i=1}^K P(L_i^{(t)} | L^{(t-1)}) * P(IE_1^{(t)} | L_i^{(t)}) * \dots * P(IE_N^{(t)} | L_i^{(t)})} \quad (2.10)$$

Οι πιθανοτικές κατανομές $P(IE_i | L)$ και $P(L^{(t)} | L^{(t-1)})$ που εμφανίζονται στην παραπάνω σχέση είναι γνωστές από τη διαδικασία μάθησης του δικτύου (learning), έτσι μπορούμε να υπολογίσουμε τις πιθανότητες για κάθε θέση $\{L1, L2, \dots, LK\}$. Το πρόβλημα του προσδιορισμού της θέσης του χρήστη είναι να βρούμε εκείνη τη θέση L_i , όπου η πιθανότητα μεγιστοποιείται δηλαδή,

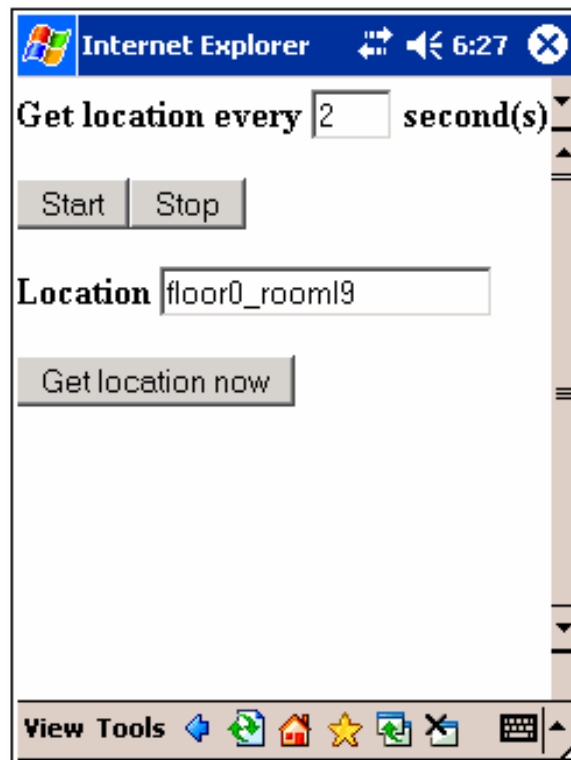
$$\max \{P(L_i^{(t)} | L^{(t-1)}, O^{(t)})\}$$

Η θέση με τη μέγιστη πιθανότητα αποθηκεύεται στην βάση δεδομένων και το προφίλ του χρήστη ενημερώνεται. Μετά από αυτό το σύστημα προχωράει στον επόμενο προσδιορισμό (συμπερασμό) λαμβάνοντας υπόψη την προηγούμενη θέση του χρήστη και τις τρέχουσες τιμές των αισθητήρων $\{IE1, IE2, \dots, IEN\}$.

2.2.2 Αξιολόγηση συστήματος προσδιορισμού θέσης

Το πεδίο δοκιμών στη διαδικασία της αξιολόγησης ήταν το ισόγειο και ο πρώτος όροφος του κτιρίου του Τμήματος Πληροφορικής και Τηλεπικοινωνιών στο ΕΚΠΑ. Κάθε όροφος έχει διαστάσεις περίπου 30X100 μέτρα. Ένας χρήστης εξοπλισμένος με μία κινητή

συσκευή (PDA) περιφερόταν στους δύο ορόφους. Στη συσκευή εκτελούνταν μία απλή εφαρμογή η οποία ρυθμιζόταν κατάλληλα να επιστρέφει τη θέση του χρήστη ανά τακτά χρονικά διαστήματα (Σχήμα 2.6).



Σχήμα 2.6: Εφαρμογή υπηρεσιών θέσης που εκτελείται σε υπολογιστή χειρός (PDA) και επιστρέφει τη θέση του χρήστη κάθε 2 δευτερόλεπτα.

Η αξιολόγηση του συστήματος εκτελέστηκε χρησιμοποιώντας δύο τεχνολογίες:

- σταθμούς βάσης (Wi-Fi access points) του ασύρματου δικτύου που είναι εγκατεστημένοι στο κτίριο
- φάρους υπερύθρων (IR beacons) της εταιρίας Lesswire AG (βλ. ΠΑΡΑΡΤΗΜΑ)

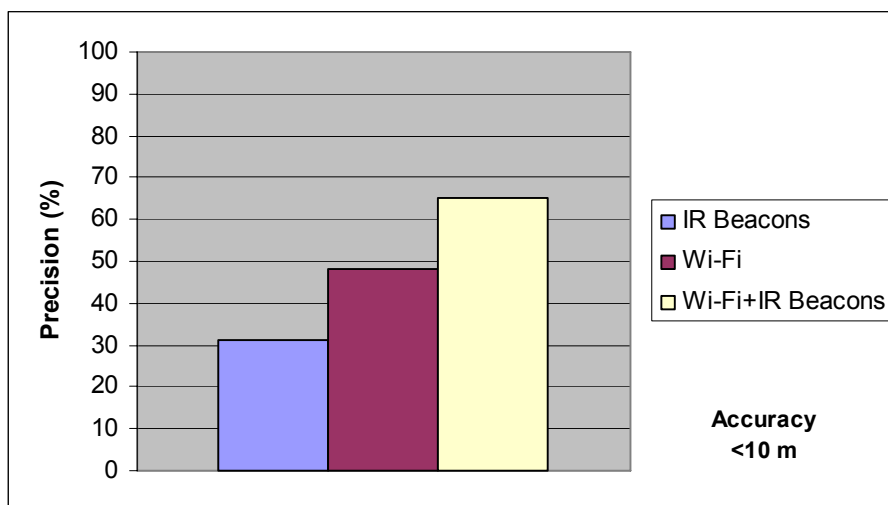
Οι σταθμοί βάσης της ασύρματης υποδομής του κτιρίου του Τμήματος είναι 4 (2 στο ισόγειο και 2 στον πρώτο όροφο). Η κάλυψη που παρέχουν για ασύρματη πρόσβαση είναι αρκετά ικανοποιητική, αλλά μόνο σε ελάχιστα σημεία του κτιρίου υπάρχει κάλυψη από δύο σταθμούς ταυτόχρονα. Αυτό επηρεάζει αρνητικά την απόδοση του συστήματος, καθώς όσες περισσότερες ενδείξεις από διαφορετικούς σταθμούς βάσης έχουμε σε ένα συγκεκριμένο σημείο, τότε τόσο καλύτερα αποτελέσματα θα επιστρέψει η διαδικασία της σύντηξης.

Οι φάροι υπερύθρων που χρησιμοποιήθηκαν ήταν 5. Οι 2 από αυτούς εγκαταστάθηκαν στο ισόγειο και οι άλλοι 3 στον πρώτο όροφο. Τα σημεία εγκατάστασης επιλέχθηκαν με τέτοιο τρόπο ούτως ώστε η συνολική κάλυψη που παρέχεται (σταθμοί βάσης και φάροι) να είναι η μέγιστη δυνατή.

Χρησιμοποιήσαμε επίσης 35 συμβολικές τοποθεσίες (*floor0_room19*, *floor0_room12*, *floor1_corridorA*,...). Οι 20 από αυτές βρίσκονται στο ισόγειο και οι 15 στον πρώτο όροφο. Η τεχνική μάθησης (learning) του Δυναμικού Δικτύου Bayes που χρησιμοποιήθηκε για τον προσδιορισμό των πιθανοτικών κατανομών, ήταν η δειγματοληψία σε κάθε μία από τις παραπάνω τοποθεσίες (στην ενότητα 2.2.1.3 περιγράφουμε αναλυτικά αυτή την τεχνική).

Η μηχανή σύντηξης (fusion engine) εκτελούνταν σε έναν υπολογιστή με επεξεργαστή Athlon 1800+ και μνήμη 512 MB. Το λειτουργικό σύστημα του υπολογιστή ήταν Windows 2000. Στην πλευρά του χρήστη χρησιμοποιήθηκε ένα iPAQ™ Pocket PC εξοπλισμένο με μία ασύρματη κάρτα δικτύου τύπου Orinoco™. Η κινητή συσκευή του χρήστη είχε επίσης ενσωματωμένη θύρα υπερύθρων.

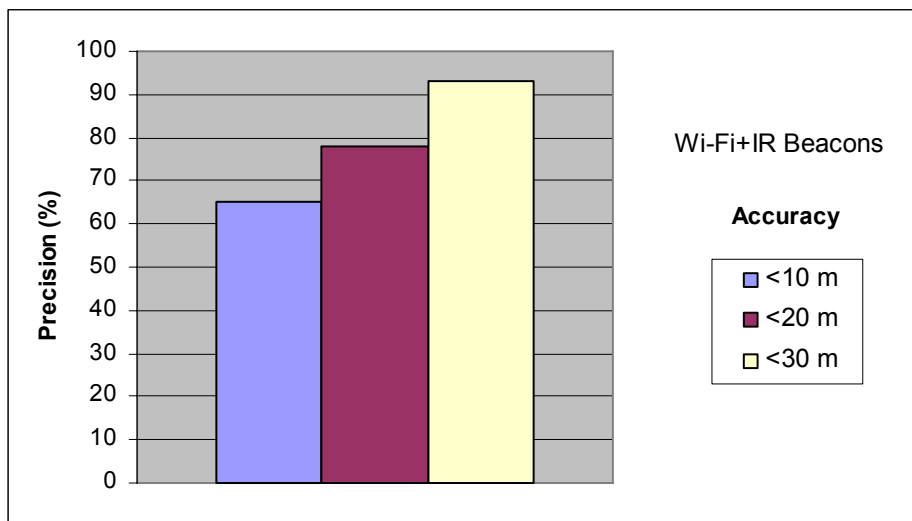
Τα πειράματα εκτελέστηκαν για συνολικό διάστημα μιας εβδομάδας σε διαφορετικές ώρες της ημέρας έτσι ώστε να έχουμε μία πιο ρεαλιστική εικόνα για τις δυνατότητες του συστήματος. Όπως αναφέρθηκε η ορθότητα (accuracy) και η διαθεσιμότητα (ή ακρίβεια) αποτελούν ίσως τα πιο σημαντικά γνωρίσματα ενός συστήματος προσδιορισμού θέσης. Στο Σχήμα 2.7 παρουσιάζεται η διαθεσιμότητα που λαμβάνουμε από το σύστημα για ορθότητα μικρότερη από 10 μέτρα, αν χρησιμοποιήσουμε μόνο φάρους υπερύθρων (IR Beacons), μόνο σταθμούς βάσης (access points) Wi-Fi (απόδοση συστήματος Nibble[13]) και τέλος το συνδυασμό αυτών των δύο τεχνολογιών.



Σχήμα 2.7: Ακρίβεια διαφορετικών τεχνολογιών για ορθότητα <10 μέτρων.

Η ακρίβεια που παρέχει το σύστημα για την πρώτη περίπτωση είναι 31%. Αν χρησιμοποιήσουμε μόνο σταθμούς βάσης η ακρίβεια ανεβαίνει στο 48%. Τέλος ο συνδυασμός και των δύο τεχνολογιών δίνει ακρίβεια 65%. Είναι προφανές λοιπόν από τα παραπάνω αποτελέσματα ότι η χρήση πολλαπλών τεχνολογιών αυξάνει την απόδοση του συστήματος.

Το Σχήμα 2.8 παρουσιάζει την ακρίβεια του συστήματος για διαφορετικές τιμές ορθότητας (<10, <20 και <30 μέτρα).



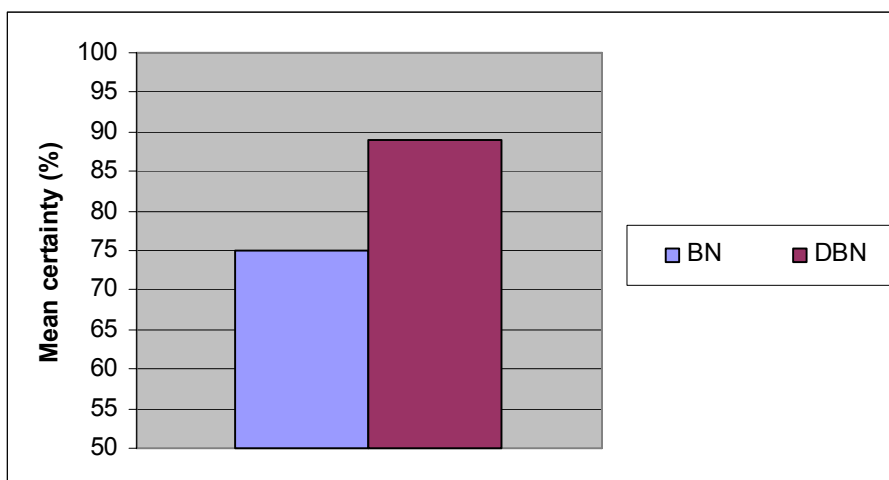
Σχήμα 2.8: Ακρίβεια του συστήματος για διαφορετικές τιμές ορθότητας.

Όπως έχει αναφερθεί η αύξηση της ακρίβειας ενός συστήματος είναι εις βάρος της ορθότητας. Έτσι, στο παραπάνω σχήμα φαίνεται καθαρά ότι αν θέλουμε περισσότερη ακρίβεια πρέπει να μειώσουμε την ορθότητα που παρέχει το σύστημα. Για ορθότητα μικρότερη από 10 μέτρα η ακρίβεια είναι 65% ενώ για ορθότητα μικρότερη από 30 μέτρα μπορούμε να έχουμε ακρίβεια 92%.

Το σύστημα υπολογίζει τις πιθανότητες για κάθε θέση και στη συνέχεια συμπεραίνει τη θέση ενός χρήστη με βάση τη μέγιστη πιθανότητα. Ορίζουμε σαν *βεβαιότητα* του συστήματος αυτή τη μέγιστη πιθανότητα που επιστρέφει σε κάθε υπολογισμό. Π.χ. αν το σύστημα επιστρέφει 60% τότε η *βεβαιότητά* του για τη θέση του χρήστη είναι 60%.

Το Σχήμα 2.9 παρουσιάζει τη μέση *βεβαιότητα* του συστήματος χρησιμοποιώντας ένα απλό στατικό Δίκτυο Bayes και ένα Δυναμικό Δίκτυο Bayes. Στην πρώτη περίπτωση (στατικό Δίκτυο) δεν λαμβάνουμε υπόψη την προηγούμενη θέση του χρήστη για να συμπεράνουμε την τρέχουσα, αλλά εκμεταλλευόμαστε μόνο τις ενδείξεις που έχουμε

από τους αισθητήρες. Η μέση *βεβαιότητα* του συστήματος σε αυτήν την περίπτωση είναι 75%.



Σχήμα 2.9: Μέση βεβαιότητα του συστήματος χρησιμοποιώντας στατικά Δίκτυα Bayes και Δυναμικά Δίκτυα Bayes.

Για να λάβουμε υπόψη την προηγούμενη θέση για τον παραπάνω υπολογισμό πρέπει να χρησιμοποιήσουμε ένα Δυναμικό Δίκτυο Bayes. Σε αυτήν την περίπτωση η μέση *βεβαιότητα* του συστήματος ανέβηκε στο 89%. Είναι αδιαμφισβήτητο λοιπόν ότι η χρήση Δυναμικών Δικτύων Bayes αντί στατικών Δικτύων και η χρήση πολλαπλών τεχνολογιών για τον προσδιορισμό της θέσης, αυξάνει τη *βεβαιότητα* του συστήματος για το που μπορεί να βρίσκεται ένας χρήστης και ως εκ' τούτου αυξάνει και την απόδοσή του.

2.3 Πιθανο-θεωρητική Σύντηξη Βασισμένη στην Αξιοπιστία των Πηγών

Στην προηγούμενη ενότητα παρουσιάσαμε ένα σχήμα πιθανο-θεωρητικής σύντηξης δεδομένων σε ένα περιβάλλον διάχυτου υπολογισμού το οποίο λαμβάνει υπ' όψιν την πληροφορία πλαισίου (δεδομένα) που παρέχουν οι αισθητήρες (πηγές). Τα δεδομένα αυτά λαμβάνονταν χωρίς να αξιολογηθούν για την ποιότητα τους. Ως εκ τούτου θα μπορούσε η δυσλειτουργία ενός ή περισσότερων αισθητήρων να επηρεάσει άμεσα το αποτέλεσμα της σύντηξης δεδομένων παρέχοντας λανθασμένες μετρήσεις. Προκειμένου να αποφευχθούν τέτοιες καταστάσεις ορίζουμε την αξιοπιστία πηγής δηλαδή το πόσο βέβαιο είμαστε για την τιμή μιας συνιστώσας πληροφορίας πλαισίου.

2.3.1 Αξιοπιστία πηγής και βαθμός εμπιστοσύνης

Η αξιοπιστία της πηγής (source reliability) h , συσχετίζει έναν βαθμό αξιοπιστίας σε κάθε πηγή S_i από ένα σύνολο πηγών πληροφορίας πλαισίου $S = \{S_1, \dots, S_N\}$ με $h: S \rightarrow [0, 1]$. Μια μηδενική τιμή για την h υποδηλώνει ότι οι τιμές από την συγκεκριμένη πηγή πληροφορίας δεν είναι αξιόπιστες ενώ μια τιμή $h = 1$ υποδηλώνει ότι η πηγή είναι πλήρως αξιόπιστη. Αξίζει να σημειωθεί ότι ο βαθμός αξιοπιστίας μιας πηγής δεν είναι σταθερός αλλά μπορεί να αλλάζει, π.χ., η ακρίβεια μιας μέτρησης από έναν αισθητήρα μπορεί να ποικίλει σε διαφορετικές κλιματολογικές συνθήκες.

Θεωρήστε την τιμή v_i για μια συνιστώσα πλαισίου p_i που μετρήθηκε από τις πηγές S_i . Τότε, ο βαθμός εμπιστοσύνης (degree of confidence) $conf$ των τιμών που συμπεραίνουν το πλαίσιο p ορίζεται ως η δυαδική σύντηξη των επιμέρους βαθμών αξιοπιστίας από κάθε πηγή, έτσι ώστε

$$conf = \max_u \left[\min_{(i,j) \in N \times N} (h(S_i), h(S_j))_u \right], u = 1.. \binom{N}{2} \quad (2.11)$$

Ο βαθμός $conf$ είναι η μέγιστη από τις ελάχιστες τιμές $h(S_i), h(S_j)$ για κάθε ζευγάρι S_i και S_j πηγών. Συγκεκριμένα υπάρχουν $\binom{N}{2}$ ζευγάρια S_i και S_j πηγών στο $S, i, j = 1, \dots, N$. Έτσι, η εμπιστοσύνη $conf$ είναι η μέγιστη τιμή των $\binom{N}{2}$ ελαχίστων βαθμών h για κάθε ζευγάρια πηγών.

Παράδειγμα: Έστω $S = \{S_1, S_2, S_3\}$ με τιμές αξιοπιστίας $\{0.2, 0.4, 0.8\}$, αντίστοιχα, τότε, ο βαθμός εμπιστοσύνης για την τιμή μιας συνιστώσας που μετράται και από τις τρεις πηγές είναι $conf = \max\{\min\{0.2, 0.4\}, \min\{0.2, 0.8\}, \min\{0.4, 0.8\}\} = 0.4$.

Δεν επιλέγουμε η τιμή του $conf$ να είναι η $\min\{h(S_i)\}, i = 1, \dots, N$ εφόσον μια τέτοια σύντηξη είναι περιοριστική επειδή ο βαθμός αξιοπιστίας μιας πηγής για ένα επίπεδο αφαίρεσης πλαισίου είναι μικρότερος από το προηγούμενο επίπεδο αφαίρεσης πλαισίου. Έτσι η συνολική εμπιστοσύνη εξαρτάται μόνο από μία πηγή και όχι από των συνδυασμό των άλλων επιπέδων αφαίρεσης.

Ο τελεστής $conf$ είναι ένας δυαδικός τελεστής ασαφούς σύντηξης πληροφορίας έχοντας ως βαθμό συγγένειας την ποσότητα h , δεδομένου ότι υπάρχει τουλάχιστον μια S_i για την οποία $h(S_i) = 1$. Συγκεκριμένα, ο τελεστής αυτός βασίζεται στην αρχή της επέκτασης,

Διαχείριση Πληροφορίας Πλαισίου για το Διάχυτο Υπολογισμό

που στην περίπτωση μας αν $a_i = S_i$, $\mu_{a_i}(v_i) = h(S_i)$ για την μέτρηση v_i και ο τελεστής min είναι δυαδικός τότε $\mu_{f(a_1, \dots, a_N)} = conf(v_i)$ (βλ. ΠΑΡΑΡΤΗΜΑ).

Βάσει του βαθμού εμπιστοσύνης, επεκτείνουμε τον κλασικό κανόνα προσδιορισμού πλαισίου θέτοντας την πρόταση ' p_i is $conf_i$ ' για κάθε συνιστώσα του πλαισίου, δηλαδή,

$$[(p_i \text{ is } v_i) \wedge (p_i \text{ is } conf_i)] \wedge \dots \wedge [(p_k \text{ is } v_k) \wedge (p_k \text{ is } conf_k)] \rightarrow [(p \text{ is } v) \wedge (p \text{ is } conf)] \quad (2.12)$$

Η πρόταση ' p_i is $conf_i$ ' στον κανόνα προσδιορισμού πλαισίου υποδηλώνει την εμπιστοσύνη στην παρατήρηση της τιμής v_i για την συνιστώσα p_i προκειμένου να επαχθεί το πλαίσιο p . Εάν $p \in P(0)$, δηλαδή δεν έχουμε επίπεδα αφαίρεσης του πλαισίου, τότε η ποσότητα $conf$ εκφυλίζεται στην ποσότητα h .

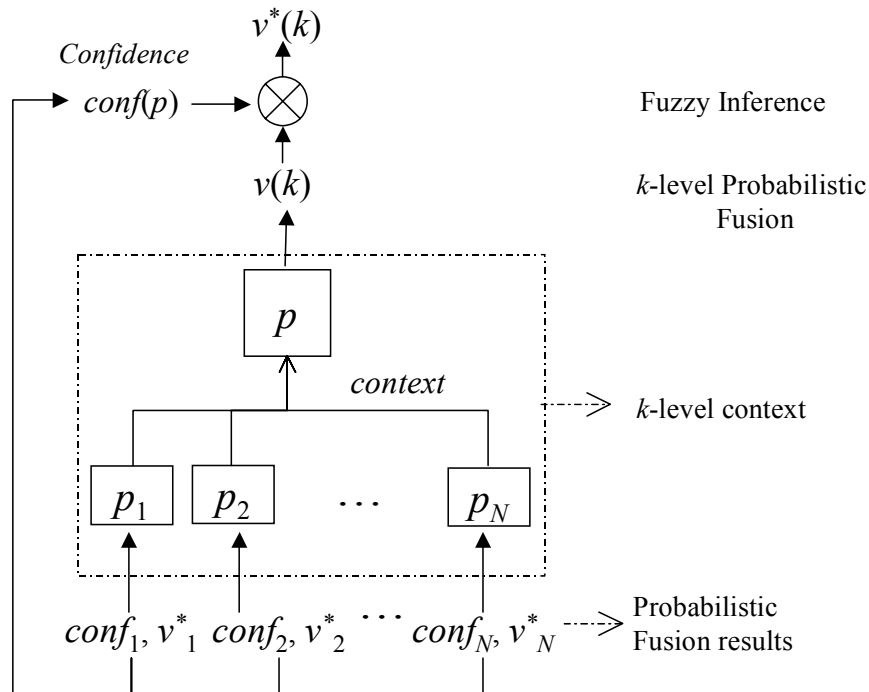
Η ποσότητα $conf$ για το πλαίσιο p υπολογίζεται βάσει της (2.11) εφόσον οι ποσότητες $conf_i$ έχουν καθοριστεί προηγουμένως. Η πιο πιθανή εμφάνιση του πλαισίου p που υπολογίζεται από τον κανόνα προσδιορισμού πλαισίου βάσει του βαθμού εμπιστοσύνης (2.12) αναφέρεται στο αποτέλεσμα της σύντηξης πληροφορίας πλαισίου (context fusion result). Στόχος της ενότητας αυτής είναι ο βαθμός ισχύς της συνεπαγωγής του κανόνα προσδιορισμού του πλαισίου (2.12) λαμβάνοντας υπ' όψιν τον βαθμό εμπιστοσύνης.

2.3.2 Ενσωμάτωση της αξιοπιστίας των πηγών στη διαδικασία σύντηξης

Η πιθανο-θεωρητική σύντηξη στην σχέση (2.1) δεν λαμβάνει υπ' όψιν τον βαθμό εμπιστοσύνης για την τιμή που μεγιστοποιεί την πιθανότητα του ενδεχομένου εμφάνισης του πλαισίου p . Για παράδειγμα, θεωρήστε το γεγονός ότι το $p^*(k, t)$ δίνεται με μεγάλη πιθανότητα αλλά οι μετρήσεις για να επαχθεί το p δεν ήταν έμπιστες, άρα με χαμηλό βαθμό $conf(p)$. Αυτό μπορεί να συντελέσει σε έναν εσφαλμένο προσδιορισμό της πιθανότητας εμφάνισης του p , $Prob(p)$.

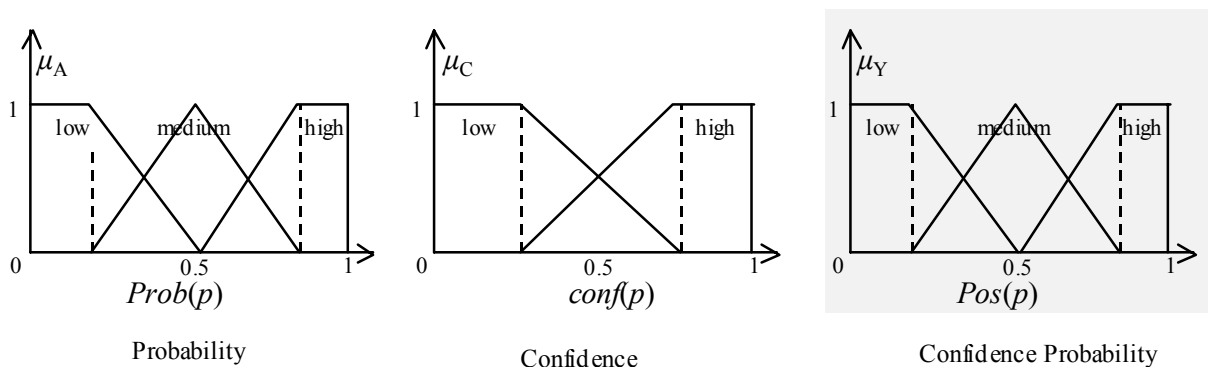
Για να ενσωματώσουμε την γνώση που ανακτάται από τον βαθμό εμπιστοσύνης στον υπολογισμό της πιθανότητας εμφάνισης του p υιοθετούμε την Θεωρία Ασαφών Συνόλων. Συγκεκριμένα, μετατρέπουμε τις τιμές της πιθανότητας και του βαθμού εμπιστοσύνης σε ασαφή σύνολα. Έτσι, μέσω ενός συνόλου κανόνων ασαφούς συμπερασμού επάγουμε την συνδυασμένη πιθανότητα εμφάνισης του πλαισίου p . Η επαγόμενη πιθανότητα αυτή καλείται πιθανότητα εμπιστοσύνης (*confidence probability*) $Pos(p)$ για το πλαίσιο p .

Στο Σχήμα 2.10 απεικονίζεται ο τρόπος εφαρμογής του προτεινόμενου ασαφούς συστήματος για τον υπολογισμό της $Pos(p)$. Συγκεκριμένα, ο ασαφής τελεστής σύντηξης \otimes είναι μοναδικός για όλα τα επίπεδα αφαίρεσης του πλαισίου και εφαρμόζεται στο ανώτατο επίπεδο ιεραρχίας του πλαισίου.



Σχήμα 2.10: Εφαρμογή Ασαφούς Λογικής στο συμπερασμό πλαισίου βάσει αξιοπιστίας πηγών.

Τα ασαφή σύνολα του τελεστή \otimes εκφράζουν την πιθανότητα και τον βαθμό εμπιστοσύνης με έναν λεκτικό τρόπο. Λεκτικοί όροι υιοθετούνται για να καθορίσουν τα ασαφή σύνολα για κάθε μια ασαφή μεταβλητή που υποδηλώνει την πιθανότητα $Prob(p)$, τον βαθμό $conf(p)$ και την επαγόμενη $Pos(p)$, όπως απεικονίζονται στο Σχήμα 2.11.



Σχήμα 2.11: Ασαφή σύνολα αναπαράστασης βασικής πιθανότητας, βαθμού εμπιστοσύνης και πιθανότητας εμφάνισης πλαισίου.

Επίσης, στο Σχήμα 2.12 απεικονίζονται οι κανόνες ασαφούς συμπερασμού πάνω από τα ασαφή σύνολα υποδηλώνοντας μια λογική ενσωμάτωσης του βαθμού εμπιστοσύνης στον υπολογισμό της πιθανότητας εμφάνισης του πλαισίου. Για την πιθανότητα εμφάνισης του πλαισίου p ορίζονται τρεις λεκτικοί όροι που χαρακτηρίζουν την τιμή της πιθανότητας αυτής από το σύνολο $C = \{high, medium, low\}$. Συγκεκριμένα, η τιμή $low Prob(p)$ υποδηλώνει ότι το επαγόμενο πλαίσιο p παράγεται με χαμηλή πιθανότητα παρατήρησης ενώ η τιμή $high Prob(p)$ ότι ένα μεγάλος βαθμός βεβαιότητας υποστηρίζει την ύπαρξη του πλαισίου p . Η μέση τιμή $medium Prob(p)$ υποδηλώνει την αβεβαιότητα για την εμφάνιση του πλαισίου p .

1. *if Prob(p) is low then Pos(p) is low*
2. *if Prob(p) is medium and conf(p) is low then Pos(p) is very low*
3. *if Prob(p) is medium and conf(p) is high then Pos(p) is somewhat high*
4. *if Prob(p) is high and conf(p) is low then Pos(p) is medium*
5. *if Prob(p) is high and conf(p) is high then Pos(p) is high*

Σχήμα 2.12: Ασαφείς κανόνες συνδυασμού του βαθμού εμπιστοσύνης στον υπολογισμό της πιθανότητας εμφάνισης.

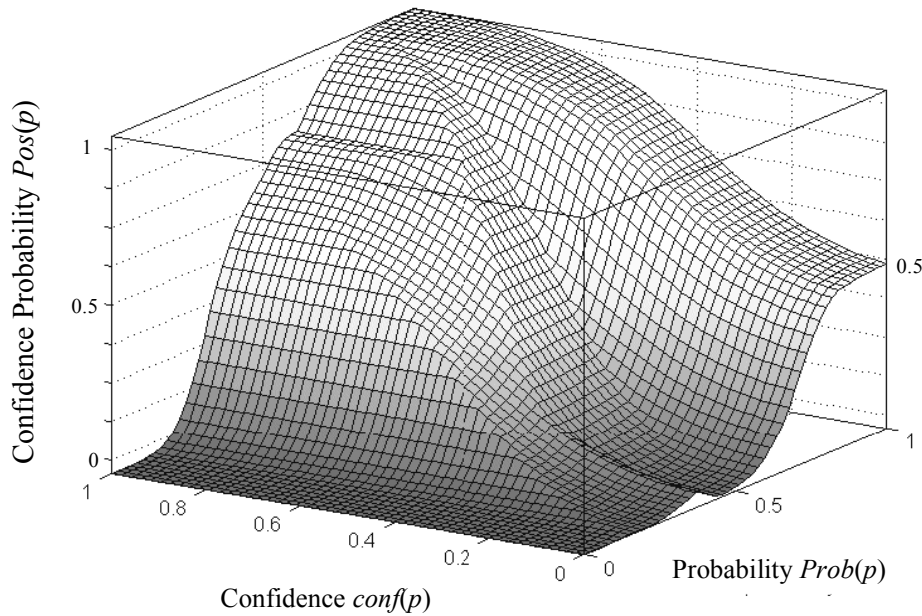
Όμοια, δύο ασαφή σύνολα που αντιστοιχούν στους λεκτικούς όρους $A = \{high, low\}$ υιοθετούνται για τον χαρακτηρισμό της τιμής του βαθμού εμπιστοσύνης $conf(p)$. Μια $low conf(p)$ τιμή υποδηλώνει ότι το επαγόμενο πλαίσιο p παρήχθη από μη έμπιστους αισθητήρες ή πηγές πληροφορίας πλαισίου. Συνεπώς, η τιμή του p δεν θεωρείται έμπιστη. Μια ανάθεση $high conf(p)$ υποδηλώνει ότι το επαγόμενο p προήλθε από έμπιστες πηγές.

Η συνδυασμένη πιθανότητα εμπιστοσύνης $y = Pos(p)$ παράγεται από την σχέση των ασαφών συνόλων που χαρακτηρίζουν τις τιμές του $Prob(p)$ και $conf(p)$ βάσει των ασαφών κανόνων. Τα ασαφή σύνολα που χαρακτηρίζουν την πιθανότητα εμπιστοσύνης είναι $Y = \{high, medium, low\}$. Έτσι, βάσει των κανόνων, η τιμή της $Pos(p)$ παράγεται με αποσαφήνιση όπως,

$$Y(y) = \bigvee_{1 \leq i \leq m} [C_{li}(Prob(p(t))) \wedge A_{li}(conf_p) \wedge Y_{li}(y)]$$

Οπότε, τώρα η (2.1) γράφεται ώστε να επιλέγεται εκείνο το πλαίσιο $p^*(k, t)$ που μεγιστοποιεί την πιθανότητα εμπιστοσύνης λόγω αξιοπιστίας των μετρήσεων, άρα,

$$p^*(k, t) = \operatorname{argmax}_{i \in N_k} \{Pos(p_i(k, t))\} \tag{2.13}$$



Σχήμα 2.13: Συμπεριφορά ασαφούς τελεστή σύντηξης.

Οι ασαφείς κανόνες στο Σχήμα 2.12 δεν περιγράφουν το πλαίσιο στο οποίο οι τιμές της πιθανότητας και της εμπιστοσύνης ταυτόχρονα λαμβάνουν χαμηλές τιμές, για παράδειγμα ένας τέτοιος κανόνας θα ήταν της μορφής *if Prob(p) is low and conf(p) is low then Pos(p) is high*. Αντιθέτως, στην περίπτωση αυτή, η τιμή της $Pos(p)$ εξαρτάται μόνο από την τιμή της $Prob(p)$ (βλέπε πρώτο κανόνα στο Σχήμα 2.12).

Το Σχήμα 2.13 απεικονίζει την συμπεριφορά του ασαφούς τελεστή σύντηξης. Συνεπώς, όταν ο βαθμός εμπιστοσύνης λαμβάνει μηδενική τιμή, που σημαίνει ότι οι μετρήσεις των αισθητήρων δεν είναι αξιόπιστες, τότε η πιθανότητα $Pos(p)$ λαμβάνει τιμή 0.5, δίνοντας έτσι την μεγαλύτερη τιμή για την $Prob(p)$, δηλαδή πλήρης αβεβαιότητα. Αυτό σημαίνει ότι το σύστημα είναι εξίσου αβέβαιο και βέβαιο για να αποφανθεί για το τελικό αποτέλεσμα της σύντηξης. Αξίζει να σημειωθεί ότι τα σχήματα των καμπυλών των ασαφών συνόλων στο Σχήμα 2.11 μπορούν να επαχθούν από πραγματικά δεδομένα καθώς πολλές τεχνικές μπορούν να υιοθετηθούν για την κατασκευή τους ώστε να εστιάζουν σε ακόμη καλύτερα αποτελέσματα ασαφούς σύντηξης (π.χ., neuro-fuzzy ταξινομητές [14]).

2.3.3 Αξιολόγηση Συστήματος

Η επίδοση του συστήματος που περιγράφηκε στις προηγούμενες ενότητες εξετάζεται για χωρική πληροφορία πλαισίου (θέση) με παρεμφερή πειραματικά δεδομένα όπως στον σύστημα εντοπισμού που περιγράφηκε στην ενότητα 2.2. Δυο τεχνολογίες

χρησιμοποιήθηκαν ξανά για την ανίχνευση της θέσης ενός χρήστη: σταθμοί βάσης Wi-Fi και φάροι υπερώθρων (IR Beacons). Χρησιμοποιήθηκαν $M = 35$ συμβολικοί χώροι μόνο στον 1^ο όροφο του Τμήματος Πληροφορικής και Τηλεπικοινωνιών ΕΚΠΑ. Μια φορητή PDA συσκευή λάμβανε σήματα από τις δύο τεχνολογίες εντοπισμού θέσης και ένα ΔΔΒ κατασκευάστηκε με βάση τις μετρήσεις με δειγματοληψία ενός δευτερολέπτου. Οι συνιστώσες του χωρικού πλαισίου είναι το Received Signal Strength (RSS) από τα κοντινά WiFi Access Points (AP) και το Infrared Radiation (IR) detection οπότε π.χ., το πλαίσιο p επάγεται από τον κανόνα προσδιορισμού:

$$AP1_RSS \text{ is } -60 \text{ dBm} \wedge IRB1 \text{ is visible} \wedge AP3_RSS \text{ is } -30\text{dBm} \rightarrow p$$

όσον αφορά την θέση του χρήστη π.χ. στο διάδρομο του 1^{ου} ορόφου.

Προκειμένου να ποσοτικοποιήσουμε τον βαθμό h_i για κάθε αισθητήρα S_i , χρησιμοποιούμε τις κατανομές πιθανότητας στις διαφορετικές συμβολικές θέσεις L_i για κάθε αισθητήρα όπως απεικονίζονται στον Πίνακα 2-1 που αποκομίζονται από την φάση εκπαίδευσης του ΔΔΒ.

Είναι φανερό ότι ο εάν ο αριθμός των μετρήσεων ενός αισθητήρα κατά την διάρκεια της εκπαίδευσης είναι κατανεμημένος ομοιόμορφα μεταξύ μιας χαμηλής και μιας υψηλής τιμής για κάποια θέση L_i τότε οι πιθανότητες για την συγκεκριμένη θέση κατανέμονται εξίσου ομοιόμορφα. Όμως, η συνθήκη της ομοιόμορφης κατανομής των πιθανοτήτων δεν αποσκοπεί σε κάποια πραγματική πληροφορία για τον αισθητήρα εφόσον κάθε τιμή του έχει ίση περίπου πιθανότητα να εμφανιστεί.

Πίνακας 2-4: Κατανομή πιθανοτήτων συμβολικών θέσεων

| Θέση / Πιθανότητα | | L_1 | L_2 | ... | L_M |
|-------------------|-------|-------|-------|-----|-------|
| | v_1 | 0.5 | 0.0 | ... | 0.1 |
| | v_2 | 0.3 | 0.8 | ... | 0.1 |
| | ... | ... | ... | ... | ... |
| | v_k | 0.0 | 0.0 | ... | 0.05 |

Για να έχουμε καλύτερα αποτελέσματα στην εκτίμηση του χωρικού πλαισίου τα δείγματα δεν πρέπει να κατανέμονται ομοιόμορφα ως προς την πιθανότητά τους. Έστω $V(L_i)$, $i = 1, \dots, M$, η διακριτή τυχαία μεταβλητή που λαμβάνει τιμές από την στήλη L_i του Πίνακα

2-1, όπου M είναι ο αριθμός των συμβολικών θέσεων. Καθώς η πιθανότητα για μια θέση L_i κατανέμεται ομοιόμορφα σε όλες τις μετρήσεις $v_j, j = 1, \dots, k$, δηλαδή, $V(L_i) = k^{-1}$ όπου k είναι ο αριθμός των διαφορετικών τιμών μέτρησης τότε αυτό υποδηλώνει πλήρη άγνοια σε αυτές τις τιμές. Έτσι, όσο μεγαλύτερη είναι η διακύμανση $Var[V]$ της διακριτής τυχαίας μεταβλητής $V(L_i)$ τόσο περισσότερη πληροφορία μπορούμε να λάβουμε για μια συγκεκριμένη θέση L_i , δηλαδή ο αισθητήρας για την θέση αυτή είναι αξιόπιστος.

Ο βαθμός αξιοπιστίας τότε για τον αισθητήρα υπολογίζεται με την μέση τιμή της διακύμανσης του σε όλες τις θέσεις, δηλαδή,

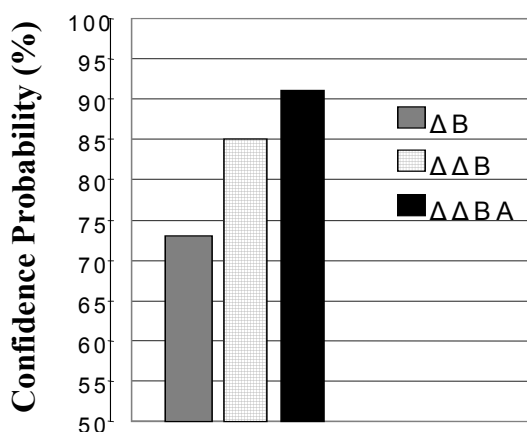
$$h = \frac{1}{M} \sum_{i=1}^M \beta \cdot Var[V(L_i)] \quad (2.14)$$

όπου β είναι ένας παράγοντας κανονικοποίησης εφόσον $h \in [0,1]$. Με την εκτέλεση των πειραμάτων παρατηρήσαμε ότι οι φάροι (IR Beacons) παρουσιάζονται περισσότερο αξιόπιστοι για τον προσδιορισμό θέσης από ότι οι σταθμοί βάσης (WLAN APs) όπως παρουσιάζεται και στον παρακάτω πίνακα. Διαισθητικά, αυτό θεωρείται σωστό δεδομένου ότι οι φάροι έχουν μικρότερη ακτίνα εκπομπής που βελτιώνει έτσι την ακρίβεια της εκτίμησης θέσης.

Πίνακας 2-5: Ο βαθμός αξιοπιστίας (h) για κάθε αισθητήρα(A=Access Point (AP), B = IR Beacon)

| ID | A1 | A2 | A3 | A4 | B1 | B3 | B3 | B4 | B5 |
|----|------|------|------|-----|------|------|------|------|----|
| h | 0.66 | 0.79 | 0.77 | 0.8 | 0.82 | 0.89 | 0.93 | 0.95 | 1 |

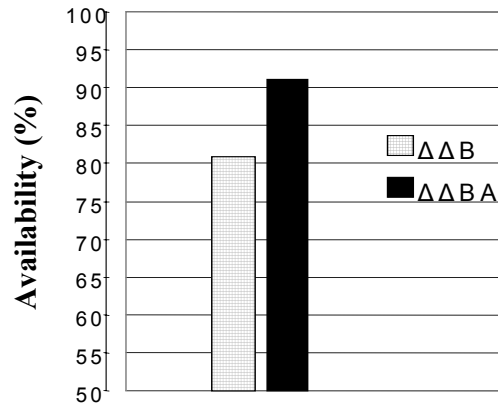
Το Σχήμα 2.14 απεικονίζει την πιθανότητα αξιοπιστίας για τις τεχνικές σύντηξης: (1) πιθανο-θεωρητική σύντηξη με χρήση στατικών Bayes Δικτύων (ΔB), (2) πιθανο-θεωρητική σύντηξη με χρήση Δυναμικών Bayes Δικτύων (ΔΔB) και (3) πιθανο-θεωρητική σύντηξη με χρήση Δυναμικών Bayes Δικτύων βάσει Αξιοπιστίας Πηγών (ΔΔBA).



Σχήμα 2.14: Πιθανότητα αξιοπιστίας για τις τεχνικές σύντηξης με χρήση στατικών Bayes Δικτύων (ΔB), με χρήση Δυναμικών Bayes Δικτύων (ΔΔB) και με χρήση Δυναμικών Bayes Δικτύων βάσει Αξιοπιστίας Πηγών (ΔΔBA).

Στην πρώτη περίπτωση η μέση τιμή της πιθανότητα εμπιστοσύνης (στην οποία η τιμές των πιθανοτήτων την χρονική στιγμή $t-1$ δεν λαμβάνονται υπό όψιν) είναι 73%, ενώ για τα ΔΔB αυξάνει στα 85%. Στην περίπτωση των ΔΔBA λαμβάνεται η τιμή 91% σηματοδοτώντας ότι το σύστημα είναι εν γένει πιο βέβαιο για τον εντοπισμό θέσης. Συγκεκριμένα, η τεχνική σύντηξης που βασίζεται στον ασαφή συμπερασμό παράγει καλύτερα αποτελέσματα όταν η πιθανότητα εμφάνισης πλαισίου έχει τιμή στο 0.5, δηλαδή το σύστημα είναι πλήρως αβέβαιο. Ο βαθμός εμπιστοσύνης υποστηρίζει την πιθανο-θεωρητική σύντηξη ώστε να γίνει το σύστημα είτε (1) πιο βέβαιο για το ενδεχόμενο εμφάνισης ενός πλαισίου είτε (2) να γίνει το σύστημα πιο βέβαιο για το ενδεχόμενο μη-εμφάνισης του πλαισίου άμεσα, και όχι έμμεσα λόγω άρνησης του πρώτου ενδεχομένου.

Επίσης, για να εξετάσουμε την επίδοση του ασαφούς τελεστή σύντηξης υιοθετούμε τις μετρικές ορθότητα (accuracy) και ακρίβεια ή διαθεσιμότητα (precision or availability) που υφίστανται στα συστήματα προσδιορισμού θέσης και περιγράψαμε στην ενότητα 2.2.



Σχήμα 2.15: Διαθεσιμότητα του συστήματος για ορθότητα λιγότερο από 5 μέτρα για τις τεχνικές σύντηξης ΔΔΒ και ΔΔΒΑ

Το Σχήμα 2.15 απεικονίζει την διαθεσιμότητα του συστήματος για λιγότερο από 5 μέτρα ορθότητα για τις τεχνικές σύντηξης ΔΔΒ και ΔΔΒΑ. Στην ΔΔΒ περίπτωση έχουμε διαθεσιμότητα 82% ενώ, ενσωματώνοντας τον βαθμό αξιοπιστίας στην ακρίβεια υπολογισμού θέσης, λαμβάνεται η διαθεσιμότητα σε τιμή 91%. Είναι πασιφανές ότι η αξιοπιστία των πηγών αυξάνει την βεβαιότητα στην εκτίμηση της θέσης του χρήστη κατά 11%.

2.4 Συμπεράσματα

Η βασική συνεισφορά της προαναφερθείσας έρευνας στο πεδίο επίγνωσης πλαισίου (context awareness) και σύντηξης πληροφορίας είναι:

- Μηχανισμός σύντηξης δεδομένων ετερογενών αισθητήρων με χρήση Δυναμικών Δικτύων Bayes και ενσωμάτωση των ιστορικών δεδομένων για τον προσδιορισμό της κατάστασης ενός χρήστη.
- Ορισμός ενός γενικού πλαισίου (framework) για την επίγνωση της κατάστασης του χρήστη σε περιβάλλοντα διάχυτου υπολογισμού με χρήση Δυναμικών Δικτύων Bayes και Θεωρίας Ασαφών Συνόλων.
- Μηχανισμός σύντηξης πληροφορίας πλαισίου με χρήση Δυναμικών Δικτύων Bayes και Θεωρίας Ασαφών Συνόλων ενσωματώνοντας τον βαθμό αξιοπιστίας των πηγών.

Τα θέματα που αναλύθηκαν στο κεφάλαιο αυτό αναφέρονται στις δημοσιεύσεις: [26], [27], [28], [29], [30], [31].

ΚΕΦΑΛΑΙΟ 3

ΕΠΙΓΝΩΣΗ ΤΗΣ ΚΑΤΑΣΤΑΣΗΣ ΤΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ

3.1 Εισαγωγή

Ένας βασικός στόχος για τα συστήματα διάχυτου υπολογισμού είναι η πρόσληψη και εκμετάλλευση της πληροφορίας που σχετίζεται με το χρήστη έτσι ώστε να παρέχονται κατάλληλες υπηρεσίες. Συχνά όμως η πληροφορία αυτή (π.χ, θέση, δραστηριότητα, κτλ.) δεν συνιστά από μόνη της ένα άμεσα αξιοποιήσιμο δεδομένο. Η επιπλέον πληροφορία σχετικά με την κατάσταση του περιβάλλοντος μέσα στο οποίο κινείται ο χρήστης και οι διαφορές μετρήσεις από αυτούς τους αισθητήρες μπορούν να ορίσουν μια πιο σύνθετη αναπαράσταση και επίγνωση του πλαισίου. Μεγάλο κομμάτι της έρευνας έχει ασχοληθεί με τη χρήση αισθητήρων φυσικών ή περιβαλλοντολογικών παραμέτρων (π.χ., θερμοκρασία, πίεση, φωτεινότητα, υγρασία, ταχύτητα ανέμου). Παραδείγματος χάριν, μία κατάσταση πυρκαγιάς μπορεί να επαχθεί από μετρήσεις αισθητήρων καπνού, θερμοκρασίας και υγρασίας αλλά και οπτικών αισθητήρων. Η επίγνωση λοιπόν της κατάστασης του περιβάλλοντος είναι σημαντική για ένα σύστημα διάχυτου υπολογισμού διότι οι καταστάσεις αυτές επηρεάζουν έμμεσα και τις ενέργειες των χρηστών.

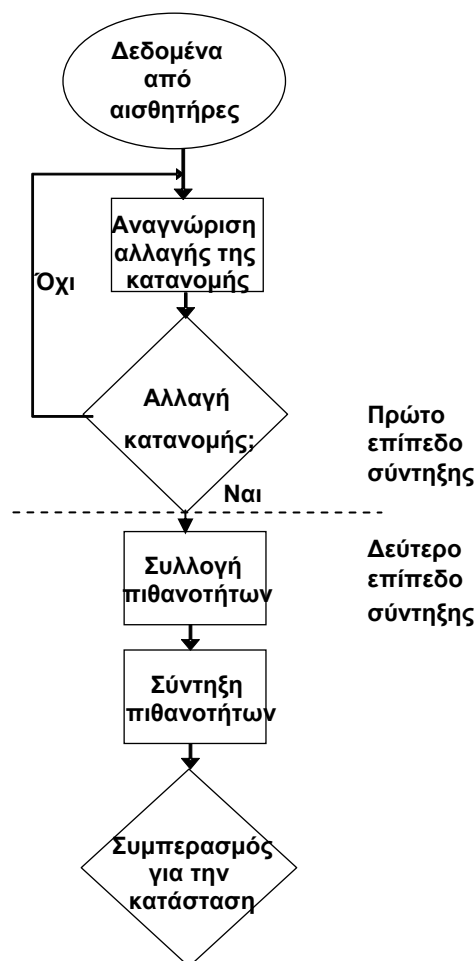
Στο κεφάλαιο αυτό μελετούμε θέματα που σχετίζονται με τη διαχείριση δεδομένων που προέρχονται από αισθητήρες ώστε να συμπεράνουμε την ύπαρξη μίας *κατάστασης* του περιβάλλοντος (εμφάνιση ενός γεγονότος - event). Για την διαδικασία αυτή μπορούν να χρησιμοποιηθούν δεδομένα/μετρήσεις από αισθητήρες εγκατεστημένους στην υπό έλεγχο περιοχή (in-field) και/ή δεδομένα από απομακρυσμένους αισθητήρες (out-field) που ελέγχουν την ίδια περιοχή από μια απόσταση. Τα δεδομένα αυτά αφού επεξεργαστούν συνδυάζονται σε ένα δι-επίπεδο σχήμα σύντηξης, βελτιώνοντας κατά συνέπεια την αξιοπιστία του συστήματος όσον αφορά στην έγκαιρη αλλά και στη σωστή ανίχνευση της κατάστασης. Η μεταφορά των δεδομένων από τους αισθητήρες μπορεί να γίνεται ασύρματα ή ενσύρματα. Στο πρώτο επίπεδο σύντηξης (σύντηξη δεδομένων) υιοθετούμε την τεχνική του σωρευτικού αθροίσματος (CUSUM) για την πιθανή ανίχνευση αλλαγής της κατανομής των δεδομένων. Έτσι, αναθέτουμε σε κάθε ένα από τους αισθητήρες (κόμβους) μία πιθανότητα ανίχνευσης της κατάστασης. Στο δεύτερο επίπεδο της σύντηξης (σύντηξη πληροφορίας) οι πιθανότητες για την εμφάνιση της κατάστασης από το πρώτο επίπεδο συνδυάζονται μέσω της Θεωρίας Ενδείξεων Dempster-Shafer. Η υιοθέτηση ενός τέτοιου σχήματος δι-επίπεδης σύντηξης έχει το

πλεονέκτημα της έγκαιρης επίγνωσης της κατάστασης μειώνοντας ταυτόχρονα το ποσοστό των ψευδών συναγερμών (false alarm rate).

3.2 Διαδικασία Σύντηξης για την Επίγνωση Καταστάσεων του Περιβάλλοντος

3.2.1 Γενικά

Τα δεδομένα από τους αισθητήρες επεξεργάζονται συνεχώς σε μία διαδικασία σύντηξης δεδομένων (πρώτο επίπεδο σύντηξης) προκειμένου να ανιχνευθεί οποιαδήποτε αλλαγή στον περιβάλλοντα χώρο. Το σύστημα παρακολουθεί συνεχώς την κατανομή των δεδομένων που παράγονται στο χρόνο. Η ανίχνευση κάποιας αλλαγής στην κατανομή απεικονίζεται σε συγκεκριμένη «μετρική». Αυτή η «μετρική» μεταφράζει τον αντίκτυπο της ένδειξης σε ένα ορισμένο ποσό πεποίθησης (belief) για μια τρέχουσα υπόθεση. Οι πιθανότητες από πολλαπλούς ετερογενείς αισθητήρες συλλέγονται στο δεύτερο επίπεδο και συνδυάζονται μέσω της Θεωρίας Ενδείξεων Dempster-Shafer ώστε να εξαχθεί ένα ασφαλές συμπέρασμα για το εάν έχει εμφανιστεί η κατάσταση ή όχι. Το Σχήμα 3.1 απεικονίζει τη ροή πληροφορίας σχετικά με την διαδικασία της σύντηξης.



Σχήμα 3.1: Διαδικασία της δι-επίπεδης σύντηξης.

3.2.2 Πρώτο επίπεδο σύντηξης – Σύντηξη δεδομένων

Ας υποθέσουμε ότι η $\{X_i\}$ αποτελεί μια ακολουθία τυχαίων μεταβλητών, δηλ., μια ακολουθία ανεξάρτητων μετρήσεων ενός αισθητήρα. Θεωρούμε ότι η X_i έχει συνάρτηση πυκνότητας πιθανότητας $f(x_i; \mu_0, \sigma)$ για $i = 1, \dots, t-1$ και $f(x_i; \mu_F, \sigma)$ για $i \geq t$, όπου η παράμετρος μ_0 είναι γνωστή ενώ οι παράμετροι μ_F και σ είναι άγνωστες. Ο δείκτης t επισημαίνει τη χρονική στιγμή στην οποία εμφανίζεται μια αλλαγή στην κατανομή των μετρήσεων από τον αισθητήρα. Η παράμετρος μ_0 υποδηλώνει τη μέση τιμή των δεδομένων, και εκτιμάται κάθε T_0 min (π.χ. $T_0 = 20$ min) με βάση τις μετρήσεις των αισθητήρων που βρίσκονται στην περιοχή. Αυτό το χρονικό παράθυρο είναι γενικά μεταβλητό και ενδείκνυται η μείωσή του κατά τη διάρκεια των περιόδων που χαρακτηρίζονται από μεγάλες διακυμάνσεις της παραμέτρου που μετράται (π.χ. από 5:00 πμ μέχρι 12:00 πμ η θερμοκρασία περιβάλλοντος έχει μεγάλες διακυμάνσεις). Η παράμετρος μ_F δείχνει τη μέση τιμή των δεδομένων που μετρώνται σε περίπτωση εμφάνισης της κατάστασης και θεωρείται γενικά άγνωστη. Ομοίως η σ^2 αποτελεί την άγνωστη διακύμανση των μετρήσεων. Για παράδειγμα, για τον αισθητήρα θερμοκρασίας Sensirion SHT11 παρέχεται από τον κατασκευαστή η ακρίβειά του που είναι ± 2.5 °C στο διάστημα από -40 °C έως -120 °C. Προσθέτοντας ένα περιθώριο για τυχόν διακυμάνσεις μπορούμε να υποθέσουμε ότι $\sigma = 5.5$ °C. Εντούτοις, η τυπική απόκλιση σ είναι γενικά άγνωστη. Εάν εμφανιστεί μία κατάσταση η παράμετρος t είναι ο χρονικός δείκτης που δείχνει μια αλλαγή της κατανομής.

Ένας από τους πιο διαδεδομένους αλγορίθμους για ανίχνευση αλλαγών είναι η τεχνική CUSUM (Cumulative Sum) που παρουσιάζεται στην [43]. Παραδείγματος χάριν, εάν η παράμετρος ενδιαφέροντος που δείχνει την πιθανότητα εμφάνισης μιας κατάστασης είναι η μέση τιμή, μπορούμε να ελέγξουμε τα μερικά αθροίσματα

$$S_n - \min S_k, \quad n = 1, 2, \dots \text{ και } 1 \leq k \leq n$$

όπου $S_n = \sum_{i=1}^n X_i$ και να συμπεράνουμε ότι μια αλλαγή από την αρχική μέση τιμή μ_0

στην τιμή μ_F εμφανίζεται τη χρονική στιγμή n . Στην εργασία [44] χρησιμοποιείται η παραπάνω τεχνική για ανίχνευση αλλαγής παρουσία ενοχλητικών παραμέτρων (nuisance parameters). Προτείνονται στατιστικές που βασίζονται στα αποδοτικά αποτελέσματα (efficient score), στην εκτίμηση μέγιστης πιθανοφάνειας (maximum likelihood estimator) ή στους λόγους λογαριθμικής πιθανοφάνειας (log-likelihood ratio). Το διάνυσμα του αποδοτικού αποτελέσματος ορίζεται ως:

$$V_k(\mu, \sigma) = \sum_{i=1}^k \nabla_v \log f(X_i; \mu, \sigma) \quad (3.1)$$

Όπως μπορεί να αποδειχθεί, εάν η πυκνότητα $f(\cdot)$ ανήκει στην εκθετική οικογένεια, π.χ., Gaussian, τότε υπό ορισμένες συνθήκες, κατά την μηδενική υπόθεση - null hypothesis υπάρχει μία διαδικασία Wiener $W(t)$ που προσεγγίζει το:

$$W_k = \Gamma^{-1/2}(\mu_0, \sigma) V_k(\mu_0, \tilde{\sigma}_k) \quad (3.2)$$

όπου $\tilde{\sigma}_k$ είναι η εκτίμηση μέγιστης πιθανοφάνειας του σ και $\Gamma(\mu_0, \sigma)$ είναι ο πίνακας πληροφορίας του Fisher (Fisher information matrix). Το στατιστικό ελέγχου (test statistic) W_k στη σχέση (3.2) μπορεί να χρησιμοποιηθεί για να ελέγξει εάν έχει εμφανιστεί κάποια αλλαγή στις πυκνότητες για χρόνο $t \leq k$. Με βάση την εναλλακτική υπόθεση (το γεγονός εμφανίστηκε το χρόνο t) το στατιστικό αυτό μετατοπίζεται για $k \geq t$. Το μέγεθος της μετατόπισης είναι ανάλογο με το ρυθμό τον οποίο το στατιστικό μετατοπίζεται στην κατεύθυνση της εναλλακτικής πυκνότητας. Επιπλέον, προκειμένου να ληφθούν αποφάσεις μετά από n παρατηρήσεις χρησιμοποιούμε το ακόλουθο αποτέλεσμα [45]:

$$\lim_{n \rightarrow \infty} P\{a(\log(n)) \max_{1 < k \leq n} k^{-1/2} W_k \leq t + b(\log(n))\} = \exp(-e^{-t}) \quad (3.3)$$

όπου $a(x) = (2\log(x))^{1/2}$ και $b(x) = 2\log(x) + 0.5\log(\log(x)) - 0.5\log(\pi)$. Για να χρησιμοποιήσουμε αυτό το αποτέλεσμα θέτουμε ως ποσοστό ψευδώς θετικών συναγερμών (false alarm rate) f , π.χ. $f = 0.001$, όπου $1 - f = \exp(-e^{-t})$ και υπολογίζουμε το κατώτατο όριο (threshold):

$$T(f) = (2\log(\log(n)))^{-1/2} [-\log(-\log(1-f)) + 2\log(\log(n)) + 0.5\log(\log(\log(n))) - 0.5\log(\pi)] \quad (3.4)$$

Κατόπιν, καταλήγουμε στο συμπέρασμα ότι η εναλλακτική υπόθεση, δηλ., γεγονός εμφάνισης της κατάστασης, υποστηρίζεται από τα στοιχεία στα πρώτα k , εάν

$$k^{-1/2} W_k \geq T(f) \quad (3.5)$$

Διαχείριση Πληροφορίας Πλαισίου για το Διάχυτο Υπολογισμό

Για $n=900$ και δύο ενδεικτικές τιμές του $f=0.01$ και $f=0.001$ λαμβάνουμε $T(f)=4.1$ και $T(f)=5.3$ αντίστοιχα. Στις επόμενες παραγράφους υποθέτουμε ότι όλες οι μετρήσεις $X_i \geq 1$ είναι ανεξάρτητες τυχαίες μεταβλητές. Σε αυτήν την περίπτωση το στατιστικό ελέγχου στη σχέση (3.3) απλοποιείται αρκετά.

Υποθέτουμε ότι η κατανομή των δεδομένων που προέρχονται από τις μετρήσεις ενός αισθητήρα ακολουθούν την παρακάτω κατανομή κατά την μηδενική υπόθεση (μή εμφάνιση της κατάστασης)

$$f(x_i; \mu_0, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x_i - \mu_0)^2}{2\sigma^2}}$$

και με βάση την εναλλακτική υπόθεση (εμφάνιση της κατάστασης) έχουμε:

$$f(x_i; \mu_F, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x_i - \mu_F)^2}{2\sigma^2}}$$

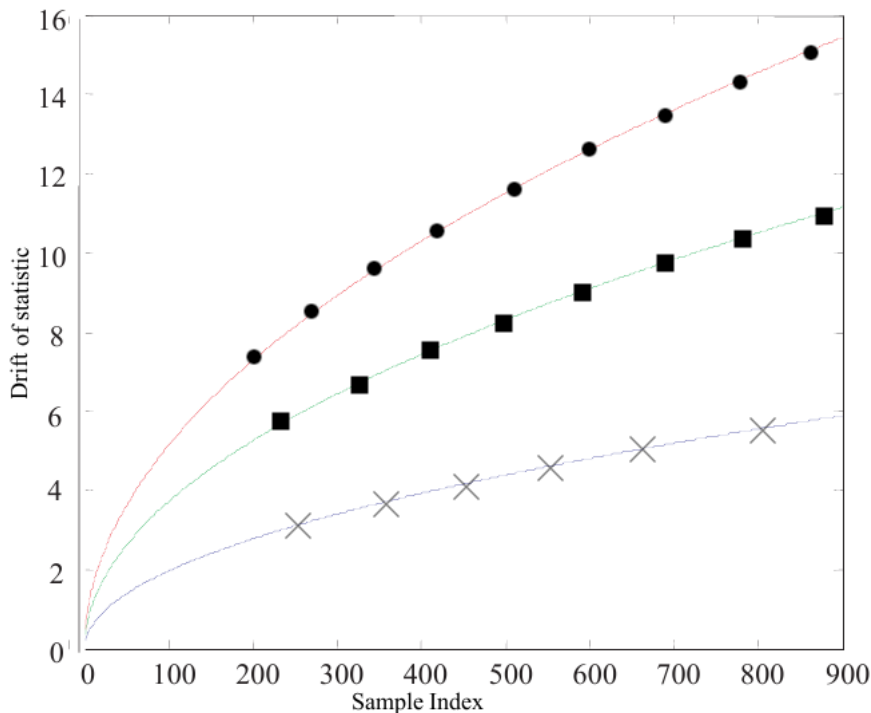
όπου θεωρούμε ότι $\mu_F > \mu_0$. Το μόνο γνωστό στοιχείο είναι η μ_0 που είναι η μέση τιμή της παραμέτρου που μετράται ελλείπει του γεγονότος εμφάνισης της κατάστασης. Επίσης θεωρούμε ότι $Y_i = X_i - \mu_0$ και $\mu_d = \mu_F - \mu_0$. Είναι σαφές ότι στην περίπτωση έλλειψης ενός τέτοιου γεγονότος ισχύει $Y_i \sim N(0, \sigma^2)$ ενώ με βάση την εναλλακτική υπόθεση ισχύει $Y_i \sim N(\mu_d, \sigma^2)$. Σε αυτήν την περίπτωση το στατιστικό ελέγχου είναι:

$$k^{-1/2} W_k = k^{-1/2} \frac{\sum_{i=1}^k Y_i}{\sqrt{\sum_{i=1}^k Y_i^2 / k}} \quad (3.6)$$

Με βάση την εναλλακτική υπόθεση, η μετατόπιση του $k^{-1/2} W_k$ μετά από μία αλλαγή στο χρόνο τ είναι:

$$\text{Μετατόπιση} = k^{-1/2} \frac{(k - (\tau - 1)) \mu_d}{\left(\sigma^2 + \frac{k - (\tau - 1)}{k} \mu_d^2\right)^{1/2}} \quad (3.7)$$

Το Σχήμα 3.2 παρουσιάζει τη μετατόπιση για $\tau = 1$, $n = 900$, $\sigma = 5$, $\mu_d = 1$ (σταυρός), $\mu_d = 2$ (τετράγωνο), και $\mu_d = 3$ (κύκλος). Όπως παρατηρείται, όσο μεγαλύτερη είναι η διαφορά της μέσης τιμής ($\mu_d = \mu_F - \mu_0$) τόσο μεγαλύτερη είναι η κλίση της μετατόπισης.



Σχήμα 3.2: Μετατόπιση του στατιστικού ελέγχου για τις διάφορες τιμές της μ_d .

Ο Πίνακας 3-1 παρουσιάζει τις χρονικές στιγμές που το στατιστικό ελέγχου διασταυρώνεται με τις τιμές κατωφλίου $T(f) = 4.1$ και $T(f) = 5.3$ υπό τον όρο ότι η αλλαγή εμφανίστηκε τη χρονική στιγμή $\tau = 1$.

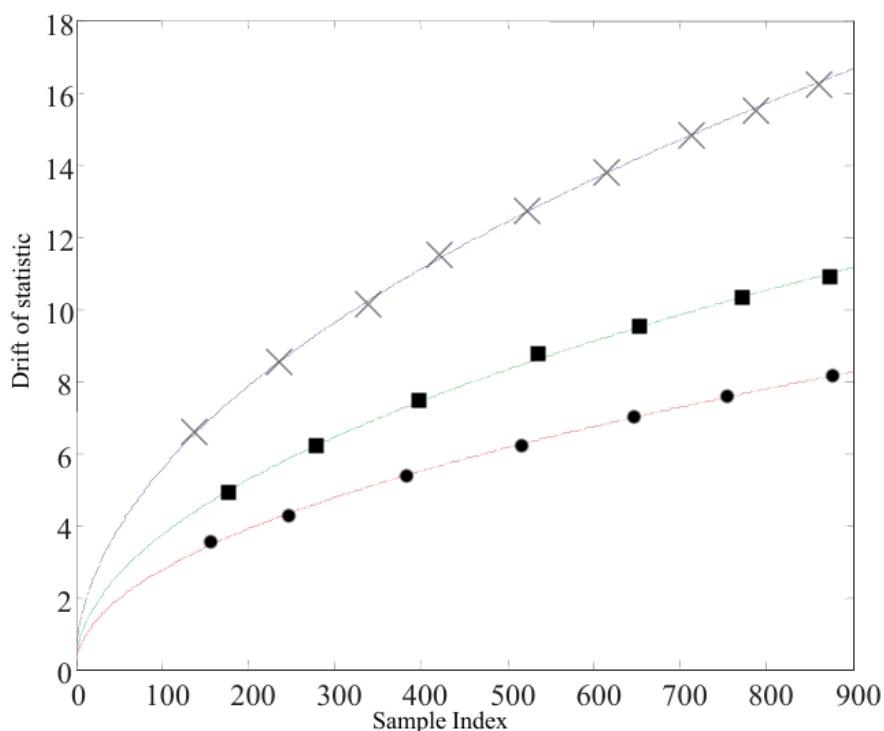
Πίνακας 3-1: Χρονικές στιγμές διασταύρωσης για το στατιστικό ελέγχου.

| $T(f)$ | $\mu_d = 1$ | $\mu_d = 2$ | $\mu_d = 3$ |
|--------|-------------|-------------|-------------|
| 4.1 | 435 | 120 | 65 |
| 5.3 | 740 | 205 | 100 |

Όπως παρατηρείται από τον πίνακα, εάν η μέση τιμή για την εναλλακτική πυκνότητα είναι $\mu_d = 3$, η αλλαγή αυτή θα ανιχνευθεί μετά από 100 δείγματα με ποσοστό ψευδώς θετικών συναγερωμών $f=0.001$.

Στο Σχήμα 3.3 απεικονίζεται η μετατόπιση (drift) για $\mu_d = 2$ και διάφορες τιμές του σ . Οι τιμές είναι οι: $\sigma = 3$ (σταυρός), $\sigma = 5$ (τετράγωνο), $\sigma = 7$ (κύκλος). Όπως αναμένεται, οι

χαμηλότερες τιμές του σ καθιστούν τις δύο πυκνότητες πιο διακριτές και επομένως δημιουργούνται μεγαλύτερες μετατοπίσεις. Αυτή η παρατήρηση είναι σημαντική για την περίπτωση που οι μετρήσεις των αισθητήρων δεν παρουσιάζουν καμία διακύμανση. Σε αυτήν την περίπτωση, δηλαδή για $\sigma = 0$ ακόμα και μια μικρή αύξηση της μέσης τιμής μπορεί να ενεργοποιήσει το μηχανισμό ανίχνευσης με μερικά μόνο δείγματα. Η λύση σε αυτό το πρόβλημα είναι να προστεθεί μια σταθερή τιμή στον παρονομαστή της (3.6) που θα ενεργήσει ως εικονική τυπική απόκλιση των μετρήσεων. Αυτή η σταθερά ρυθμίζει την ευαισθησία του μηχανισμού ανίχνευσης στην παράμετρο μ_d .



Σχήμα 3.3: Μετατόπιση του στατιστικού ελέγχου για τις διάφορες τιμές της τυπικής απόκλισης σ .

Η παράμετρος μ_0 εκτιμάται κάθε T_0 min και βασίζεται σε όλους τους αισθητήρες σε μία συγκεκριμένη περιοχή. Η περίοδος T_0 πρέπει να είναι αρκετά μεγάλη ώστε να εφαρμοστεί η διαδοχική ανίχνευση αλλαγών με όσο το δυνατόν περισσότερα δείγματα αλλά και αρκετά μικρή προκειμένου να ανιχνευθούν και οι συχνές αλλαγές των συνθηκών.

Διάφορα ζητήματα ανακύπτουν όταν υιοθετείται η διαδικασία ανίχνευσης που περιγράφηκε παραπάνω. Καταρχήν, η μέθοδος υποθέτει ότι όλοι οι αισθητήρες είναι σωστά βαθμονομημένοι. Εάν ένας από τους αισθητήρες (S_i) παρουσιάζει μία σχετικά μεγάλη (θετική) απόκλιση σε σχέση με τους υπόλοιπους αισθητήρες, θα υπάρξουν

λανθασμένες εκτιμήσεις. Αυτό που συμβαίνει σε αυτήν την περίπτωση είναι ότι η μέση τιμή των δεδομένων μ_0 που μετριέται στην έναρξη του χρονικού παραθύρου T_0 θα είναι συνεχώς μικρότερη από τις μετρήσεις του συγκεκριμένου αισθητήρα και, επομένως, αυτός θα ανιχνεύσει ψευδώς ένα γεγονός εμφάνισης της κατάστασης μετά από κάποιο χρόνο. Ο χρόνος αυτός εξαρτάται από το μέγεθος της απόκλισης. Η λύση που προτείνουμε για αυτό το πρόβλημα είναι η περιοδική βαθμονόμηση των αισθητήρων. Κατά τη διάρκεια ορισμένων περιόδων όταν οι εξωγενείς παράμετροι δεν επηρεάζουν τις μετρήσεις, οι αποκλίσεις των αισθητήρων μπορούν να υπολογιστούν και να ληφθούν υπόψη στη διαδικασία ανίχνευσης. Κατά συνέπεια, εάν ένας αισθητήρας παρουσιάσει μία μ_{off} συγκρινόμενος με τη μέση τιμή, τότε κατά τη διαδικασία ανίχνευσης της κατάστασης σε αυτόν τον αισθητήρα θα χρησιμοποιηθεί η τιμή $\mu_0 + \mu_{off}$ αντί μ_0 .

Ένα δεύτερο ζήτημα είναι ο συσχετισμός των μετρήσεων από τους αισθητήρες. Η σχέση (3.6) ορίστηκε με βάση την υπόθεση ότι οι μετρήσεις είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν την Gaussian κατανομή. Εντούτοις, στην πραγματικότητα οι μετρήσεις συσχετίζονται και αυτό μπορεί να προκαλέσει πρόβλημα όπως φαίνεται στο ακόλουθο σενάριο. Υποθέστε ότι στην έναρξη της περιόδου T_0 , όταν υπολογίζεται η μέση τιμή μ_0 διάφορες παράμετροι οδηγούν σε υποτίμηση της μ_0 . Όταν οι παράγοντες αυτοί εκλείψουν, η μέση τιμή των δεδομένων θα αυξηθεί και θα παραμείνει σε υψηλότερα επίπεδα από την αρχική τιμή της για μεγάλο αριθμό δειγμάτων. Ανάλογα με τη σχετική αύξηση της τιμής και το παράθυρο συσχετισμού τα κατώφλια για την ανίχνευση της κατάστασης μπορούν να διασταυρωθούν και έτσι να υπάρχει λανθασμένη εκτίμηση εμφάνισης ενός γεγονότος. Προκειμένου να ποσοτικοποιηθεί και να προσομοιωθεί η προαναφερθείσα κατάσταση θεωρούμε το ακόλουθο μοντέλο. Υποθέτουμε ότι μετρήσεις αισθητήρων X_i , μπορούν να εκφραστούν ως:

$$X_i = \mu_0 + z_i + r_i \quad (3.8)$$

όπου η τυχαία μεταβλητή z_i αντιπροσωπεύει το θόρυβο λόγω των ηλεκτρονικών του αισθητήρα και μπορεί να εκφραστεί ως μια Gaussian διαδικασία με μέση τιμή μηδέν και διακύμανση σ_z^2 . Η τυχαία μεταβλητή r_i είναι το δείγμα τη χρονική στιγμή i μιας διαδικασίας $r(t)$ και μοντελοποιεί τις διακυμάνσεις των μετρήσεων λόγω διαφόρων εξωγενών παραγόντων. Υποθέτουμε ότι αυτή η διαδικασία είναι Gaussian έχοντας συνάρτηση αυτοσυσχέτισης της μορφής

$$R_r(\tau) = E[r(t)r(t+\tau)] = \sigma_m^2 e^{-a|\tau|} \quad (3.9)$$

Όσο μικρότερη είναι η σταθερά α τόσο μεγαλύτερος είναι ο συσχετισμός μεταξύ των διαδοχικών δειγμάτων. Η διαδικασία $r(t)$ μπορεί να παραχθεί με τη διαβίβαση λευκού Gaussian θορύβου $w(t)$ μέσω ενός συστήματος με έναν πόλο στο α , δηλαδή:

$$\frac{dr(t)}{dt} = -\alpha r(t) + w(t) \quad (3.10)$$

όπου η συνάρτηση αυτοσυσχέτισης της $w(t)$ είναι η $R_w(\tau) = 2\alpha\sigma_m^2\delta(\tau)$. Από τη σχέση (3.10) προκύπτει:

$$\frac{d}{dt}(e^{\alpha t} r(t)) = e^{\alpha t} w(t) \Rightarrow \int_t^{t+T_s} \frac{d}{dt}(e^{\alpha t} r(t)) = \int_t^{t+T_s} e^{\alpha \tau} w(\tau) d(\tau)$$

ή αλλιώς

$$e^{\alpha(t+T_s)} r(t+T_s) - e^{\alpha t} r(t) = \int_t^{t+T_s} e^{\alpha \tau} w(\tau) d(\tau).$$

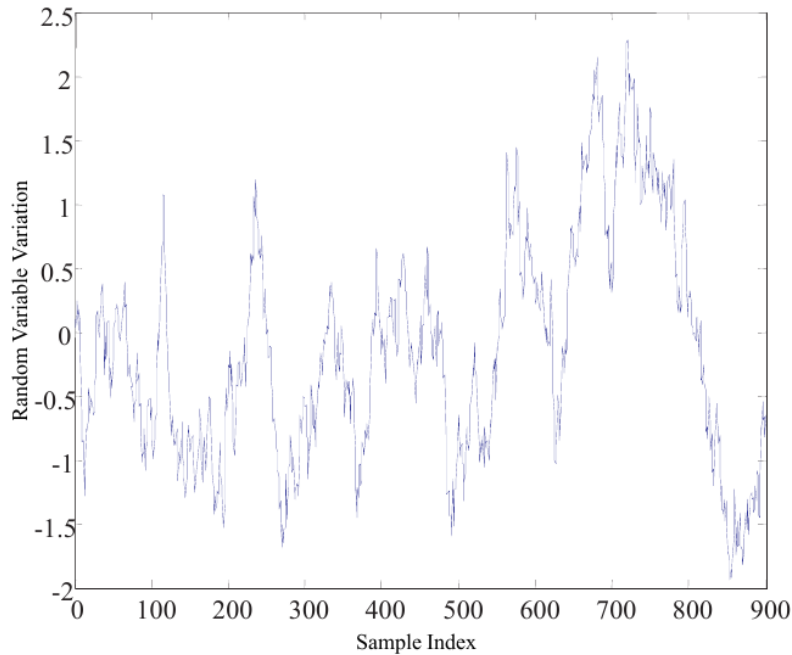
Υπολογίζοντας την προηγούμενη σχέση για $t=iT_s$ λαμβάνουμε

$$r_{i+1} = e^{-\alpha T_s} r_i + w_i \quad (3.11)$$

όπου $w_i = \int_{iT_s}^{(i+1)T_s} e^{-\alpha(i+1)T_s-\tau} w(\tau) d(\tau)$.

Η τυχαία μεταβλητή w_i είναι Gaussian διαδικασία με μέση τιμή μηδέν και διακύμανση $\sigma_{w_i}^2 = E[w_i^2] = \sigma_m^2 (1 - e^{-2\alpha T_s})$.

Το Σχήμα 3.4 απεικονίζει μία συνάρτηση του r_i που προέκυψε για $\alpha = 1/120\text{sec}$, $T_s=2\text{sec}$ και $\sigma_m=1$. Όπως μπορεί να φανεί από το σχήμα, ακόμη και μικρές τιμές του σ_m^2 μπορούν να προκαλέσουν μεγάλες αποκλίσεις από τη μηδενική μέση τιμή. Η επιλογή της σταθεράς α υποδεικνύει ένα μέσο παράθυρο συσχετισμού 120 sec, που σημαίνει ότι οι αποκλίσεις εμμένουν για 60 και πλέον δείγματα.



Σχήμα 3.4: Συνάρτηση της τυχαίας μεταβλητής r_i που μοντελοποιεί τις αλλαγές στην τιμή μιας μεταβλητής λόγω εξωγενών παραγόντων.

Το πρόβλημα που εισάγεται από τις συσχετισμένες μετρήσεις μπορεί να παρακαμφθεί με έναν από τους ακόλουθους τρόπους:

1. Μια μέθοδος είναι να αυξηθούν τα κατώτατα όρια (κατώφλια) έτσι ώστε οι διασταυρώσεις λόγω του συσχετισμού να αποφευχθούν. Μια σωστή πρακτική είναι η εξαγωγή των κατωφλίων από πραγματικά δεδομένα. Εντούτοις, αυτή η αύξηση των κατωφλίων μπορεί να αναβάλει την ανίχνευση μιας κατάστασης ή ακόμα και να προκαλέσει την μη ανίχνευση όταν η διασταύρωση βρεθεί έξω από το χρονικό παράθυρο T_0 . Η αύξηση των κατωφλίων λειτουργεί κατάλληλα μόνο εάν η υποτιθέμενη μέση τιμή μ_d είναι αρκετά μεγάλη.
2. Μια πιο ελπιδοφόρος λύση είναι να στηριχθούμε στην συνεργασία των γειτονικών αισθητήρων ούτως ώστε να ελαχιστοποιήσουμε την συσχέτιση. Δεδομένου ότι οι μετρήσεις των γειτονικών αισθητήρων υποβάλλονται στις ίδιες αλλαγές από εξωγενής παράγοντες ο όρος r_i μπορεί να υπολογιστεί από τους γειτονικούς κόμβους και να αφαιρεθεί από την X_i .

Με βάση τη δεύτερη προσέγγιση προκειμένου να εξεταστούν οι συσχετισμένες μετρήσεις, υποθέτουμε έναν αισθητήρα S_i και τους γείτονές του S_j , $j=1, \dots, |S_i|$, όπου το $|S_i|$ δείχνει τον αριθμό των στοιχείων του συνόλου (cardinality) από τους γείτονες του

Διαχείριση Πληροφορίας Πλαισίου για το Διάχυτο Υπολογισμό

αισθητήρα. Υποθέτουμε επίσης ότι ο αισθητήρας S_i αντιλαμβάνεται μία αύξηση στη μέση τιμή της μεταβλητής που μετράει η οποία είναι

$$X_i = \mu_0 + \mu_d + z_i + r_i \quad (3.12)$$

Για τους γειτονικούς αισθητήρες υποθέτουμε ότι οι μετρήσεις τους είναι της μορφής

$$X_j = \mu_0 + z_j + r_{i-Dj} \quad (3.13)$$

όπου ο όρος θορύβου z_j είναι ανεξάρτητος από τον z_i , και ο όρος r_{i-Dj} εκφράζει την ίδια διακύμανση r_i που υπεισέρχεται στις μετρήσεις του αισθητήρα S_i , καθυστερημένη ή προηγμένη κατά Dj . Έπειτα, για τον αισθητήρα S_i εφαρμόζουμε το προτεινόμενο στατιστικό ελέγχου στα δεδομένα.

$$Y_i = X_i - \frac{1}{|S_i|} \sum_{j=1}^{|S_i|} X_j = \mu_d + z_i + r_i - \frac{1}{|S_i|} \sum_{j=1}^{|S_i|} z_j - \frac{1}{|S_i|} \sum_{j=1}^{|S_i|} r_{i-Dj} \quad (3.14)$$

Ο όρος $\frac{1}{|S_i|} \sum_{j=1}^{|S_i|} z_j$ θα είναι κοντά στο μηδέν ενώ ο όρος $\frac{1}{|S_i|} \sum_{j=1}^{|S_i|} r_{i-Dj}$ ενεργεί ως πρόβλεψη

για το r_i και επομένως σχεδόν ακυρώνει αυτόν τον όρο. Σημειώστε ότι στην εφαρμογή του στατιστικού ελέγχου όσον αφορά τα στοιχεία Y_i δεν είναι απαραίτητο να αφαιρέσουμε την τιμή μ_0 καθώς ο όρος έχει ήδη ακυρωθεί.

3.2.3 Δεύτερο επίπεδο σύντηξης – Σύντηξη πληροφορίας

Τα δεδομένα από τους αισθητήρες υποβάλλονται συνεχώς σε επεξεργασία προκειμένου να ανιχνευθεί οποιαδήποτε σημαντική αλλαγή στην κατάσταση του περιβάλλοντος. Το σύστημα ελέγχει τακτικά την κατανομή των δεδομένων που παράγονται. Εάν εμφανιστεί μια αλλαγή στην κατανομή, αυτό απεικονίζεται σε συγκεκριμένη «μετρική». Αυτή η «μετρική» μεταφράζει τον αντίκτυπο της ένδειξης σε ένα ορισμένο ποσό πεποίθησης (belief) για μια τρέχουσα υπόθεση (η υπόθεση του γεγονότος εμφάνισης μιας κατάστασης). Η πιθανότητα αυτή υπολογίζεται για κάθε

τοπικό αισθητήρα. Η «μετρική» αυτή ονομάζεται Βασική Πιθανότητα Καταχώρισης – ΒΠΚ (Basic Probability Assignment - BPA).

Υπάρχουν διάφοροι τρόποι να οριστεί η BPA σύμφωνα με τις μετρήσεις των αισθητήρων. Ένας τρόπος είναι να χρησιμοποιήσουμε τις πυκνότητες πιθανότητας για κάθε αισθητήρα όπως φαίνεται παρακάτω.

$$m(f) = \frac{f(x_i; \mu_F, \sigma)}{f(x_i; \mu_F, \sigma) + f(x_i; \mu_0, \sigma)} \quad (3.15)$$

Η παράμετρος σ εξάγεται από τον παρονομαστή στην σχέση (3.6), ενώ η μ_F τίθεται ίση με $\mu_0 + \mu_d$, όπου μ_d είναι μια αρκετά μεγάλη τιμή που υπερβαίνει τις πιθανές διακυμάνσεις μέσα στο αντίστοιχο χρονικό παράθυρο T_0 . Αυτές οι διακυμάνσεις μπορούν να ληφθούν μέσω ιστορικών δεδομένων, των καθημερινών δεδομένων από τις μετρήσεις των αισθητήρων, κ.λπ. Εναλλακτικά, μπορούμε να χρησιμοποιήσουμε μια αύξουσα συνάρτηση $g(\cdot)$ που απεικονίζει το διάστημα $[\mu_0, \mu_0 + \mu_d]$ στο διάστημα $[0, 1]$. Πρέπει να σημειώσουμε ότι η έκφραση $m(f)$ δείχνει το βαθμό εμπιστοσύνης που έχουμε στην παρατήρηση ενός γεγονότος εμφάνισης της κατάστασης που παρακολουθούμε. Οι τιμές της πιθανότητας καταχώρισης ΒΠΚ (BPA) στους αισθητήρες συνδυάζονται μεταξύ τους. Οι τιμές αυτές μεταφράζονται σε μία Πιθανότητα Απόφασης (Pignistic Probability)¹. Τέλος, η τιμή αυτή συγκρίνεται με ένα κατώφλι που δείχνει εάν μία κατάσταση έχει εμφανιστεί ή όχι.

Η Θεωρία Ενδείξεων του Dempster – Shafer (DST) [47] και η εξέλιξή της, δηλαδή το Μοντέλο Μεταβιβάσιμης Πεποίθησης (Transferable Belief Model ή αλλιώς TBM) [48], είναι οι δύο προσεγγίσεις που χρησιμοποιούνται στο προτεινόμενο σύστημα. Η Θεωρία του Dempster - Shafer εξετάζει ένα περιοριστικό πιθανοτικό μοντέλο χωρίς πρωθύστερες γνωστές (a-priori) πιθανότητες. Συγκεκριμένα, η κύρια ιδέα της είναι να χρησιμοποιηθούν μόνο διαθέσιμες πληροφορίες που λαμβάνονται από τους αισθητήρες για το συνδυασμό των ενδείξεων. Υποθέστε μια συγκεκριμένη κατάσταση του συστήματός μας όπου οι πεποιθήσεις για τα γεγονότα εμφάνισης μιας κατάστασης αξιολογούνται και συνδυάζονται. Θεωρήστε επίσης έναν αισθητήρα S_i . Υπάρχουν δύο πιθανές ενδείξεις για τον S_i : είτε ανιχνεύει μια κατάσταση είτε όχι. Ως εκ τούτου, το πεπερασμένο σύνολο που αντιπροσωπεύει το χώρο των υποθέσεων για τον S_i είναι το

¹ Η πιθανότητα απόφασης (pignistic probability) είναι η πιθανότητα που ένας ειδικός (expert) αναθέτει σε μία επιλογή όταν απαιτείται η λήψη μιας απόφασης.

σύνολο $\Theta = \{\text{εμφάνιση κατάστασης-event, μη εμφάνιση κατάστασης-no event}\}$. Αυτό το σύνολο καλείται επίσης και Πλαίσιο Διάκρισης (Frame of Discernment).

Το δυναμοσύνολο (power set) $\Omega = 2^\Theta$ αντιπροσωπεύει όλα τα ατομικά - μεμονωμένα και μη ατομικά γεγονότα που χρησιμοποιούνται κατά τη διάρκεια της διαδικασίας συνδυασμού, δηλαδή $\Omega = \{\text{NULL, event, no event, event or no event}\}$. Το κενό σύνολο (NULL) δείχνει ότι δεν μπορούμε να υπολογίσουμε - αξιολογήσουμε οποιαδήποτε ένδειξη στο Ω (δηλ., το φαινόμενο των συγκρουόμενων προτάσεων στο συνδυασμό ενδείξεων). Το στοιχείο “event or no event” αντιπροσωπεύει πλήρη αβεβαιότητα κατά την οποία καμία κατάσταση δεν ανιχνεύεται (δεν μπορεί να υπολογιστεί το “event”) ούτε μπορούμε να συμπεράνουμε γεγονός “no event”. Η πλήρης αβεβαιότητα συνδέεται ιδιαίτερα με τον περιορισμό που έχουμε λόγω της αξιοπιστίας των μετρήσεων από τους αισθητήρες.

Η θεωρία DST παρέχει μια προσέγγιση που μας δίνει τη δυνατότητα να καταχωρίσουμε – αναθέσουμε αβεβαιότητα στα γεγονότα. Αυτή η προσέγγιση αρχίζει με μια συνάρτηση πεποιθήσης (μερικές φορές αποκαλούμενη και σαν συνάρτηση υποστήριξης ή αξιοπιστίας). Λαμβάνοντας υπόψη το σύνολο Θ των πιθανών γεγονότων και το σύνολο $U \subseteq \Theta$, η πίστη στο U που εκφράζεται από το $Bel(U)$, είναι ένας αριθμός στο διάστημα $[0,1]$. Μια συνάρτηση αξιοπιστίας Bel που καθορίζεται σε ένα χώρο Θ πρέπει να ικανοποιεί τις ιδιότητες: $Bel(\text{NULL}) = 0$ και $Bel(\Theta) = 1$. Εντούτοις, οι συναρτήσεις αξιοπιστίας είναι μέρος της θεωρίας των ενδείξεων. Οι ενδείξεις παρέχουν κάποιο βαθμό υποστήριξης για κάθε υποσύνολο του Θ . Το συνολικό «ποσό» υποστήριξης είναι ίσο με 1. Η πεποίθηση ότι ισχύει το U , $Bel(U)$ εκφράζει το άθροισμα των πεποιθήσεων για τα υποσύνολα του U . Μια συνάρτηση μάζας πεποιθήσης ή πιθανότητας (βασική πιθανότητα καταχώρισης - ΒΠΚ) στο Θ είναι ουσιαστικά μια συνάρτηση απεικόνισης από τα μέλη του δυναμοσυνόλου του Ω στο διάστημα $[0,1]$ και τη συμβολίζουμε με το λατινικό γράμμα m , δηλαδή $m: 2^\Theta \rightarrow [0,1]$. Η συνάρτηση αυτή ικανοποιεί τις ακόλουθες ιδιότητες:

$$m(\text{NULL}) = 0$$

$$m(U) \geq 0, \forall U \subseteq \Theta$$

$$\sum_{U \subseteq \Theta} m(U) = 1$$

Η $m(U)$ περιγράφει το βαθμό στον οποίο οι ενδείξεις υποστηρίζουν το U , δηλαδή η $m(U)$ μπορεί να θεωρηθεί ως πιθανότητα παρατήρησης του γεγονότος που εκφράζεται από

το U . Συγκεκριμένα, η $m(U)$ εκφράζει το βαθμό γνώσης ότι μία ένδειξη (π.χ., ένα γεγονός– event) ανήκει στο U . Λαμβάνοντας υπόψη μια συνάρτηση μάζας πεποίθησης m η συνάρτηση αξιοπιστίας Bel είναι

$$Bel(U) = \sum_{U': U' \subseteq U} m(U')$$

δηλαδή το άθροισμα των μαζών όλων των συνόλων που είναι υποσύνολα του U . Η πεποίθηση Bel θεωρείται ως η χαμηλότερη τιμή πιθανότητας στη θεωρία DST.

Η διαδικασία συμπερασμού που βασίζεται στην θεωρία DST δείχνει ότι λαμβάνουμε γνώση μόνο την εμφάνιση μιας κατάστασης (event) καθώς τα δεδομένα των αισθητήρων (πολλαπλοί εμπειρογνώμονες) υποστηρίζουν γνώση μόνο για της εμφάνιση της κατάστασης. Αυτό σημαίνει ότι η βασική πιθανότητα καταχώρησης για το γεγονός μη ύπαρξης της κατάστασης (no event) δεν παρέχεται. Ως εκ τούτου, η $Bel(\text{no event})$ συνάγεται σιωπηρά από την $m(\text{event})$. Θεωρήστε την $m_i(U)$ που είναι η τρέχουσα πιθανότητα εμφάνισης μίας ένδειξης στο U για τον εμπειρογνώμονα S_i . Ο συνδυασμός των αποφάσεων των εμπειρογνωμόνων S_i και S_j παρέχεται από:

$$m_{i,j}(U) = (m_i \oplus m_j) = \frac{1}{1-K} \sum_{B \cap C = U \neq \emptyset} m_i(B)m_j(C) \quad (3.16)$$

όπου

$$K = \sum_{B \cap C = \emptyset} m_i(B)m_j(C) \quad (3.17)$$

Το K αντιπροσωπεύει ένα μέτρο σύγκρουσης μεταξύ των δύο συνόλων. Εφ' όσον η πεποίθηση για την εμφάνιση μιας κατάστασης (event) και το γεγονός μη εμφάνισης της κατάστασης (no event) υπολογίζεται από την (3.16), θα πρέπει να μετασχηματίσουμε αυτές τις πεποιθήσεις (μάζες πιθανότητας) σε πιθανότητες απόφασης (pignistic probabilities). Για αυτό το λόγο, υιοθετούμε το Μοντέλο Μεταβιβάσιμης Πεποίθησης (TBM) [48] με το οποίο παράγουμε την πιθανότητα του γεγονότος (event) χρησιμοποιώντας τον ακόλουθο μετασχηματισμό

$$BetP(x) = \sum_{x \in U \subseteq \Omega} \frac{m(U)}{|U|} \quad (3.18)$$

όπου $|U|$ είναι το πλήθος των υποθέσεων του Ω στο U . Στη σχέση (3.18) λαμβάνουμε την πιθανότητα του γεγονότος θέτοντας $x=\{\text{event}\}$. Η άγνοια για το γεγονός μη

εμφάνισης της κατάστασης (no event) δεν αντιμετωπίζεται ως σύγκρουση ενδείξεων καθώς κάθε αισθητήρας παράγει πιθανότητες μόνο για το γεγονός εμφάνισης της κατάστασης.

Η βασική ιδέα για το δι-επίπεδο σχήμα σύντηξης είναι η εκπλήρωση δύο αντικρουόμενων πτυχών στη διαδικασία ανίχνευσης ενός γεγονότος. Οι δύο αυτές πτυχές είναι η ελαχιστοποίηση του ποσοστού των ψευδώς θετικών συναγερμών και η έγκαιρη ανίχνευση του γεγονότος. Η μείωση του ποσοστού ψευδώς θετικών συναγερμών μπορεί να επιτευχθεί με το συνδυασμό των δεδομένων από όλους τους διαθέσιμους αισθητήρες. Εντούτοις, αυτή η ενέργεια έχει μια δυσμενή συνέπεια στην έγκαιρη ανίχνευση, δεδομένου ότι οι αισθητήρες μπορεί να μην ανιχνεύσουν ταυτόχρονα μία κατάσταση. Επομένως οι συγκρουόμενες παρατηρήσεις των αισθητήρων, θα αναγκάσουν τον αλγόριθμο που βασίζεται στην θεωρία Dempster - Shafer να παράγει μικρές τιμές πεποίθησης (belief) στο γεγονός εμφάνισης μιας κατάστασης (event). Η προσέγγισή μας για την αντιμετώπιση αυτού του προβλήματος είναι να λάβουμε μόνο τις M υψηλότερες πιθανότητες από τους αισθητήρες και να τις συνδυάσουμε στο δεύτερο επίπεδο. Σημειώστε ότι, εάν το ελάχιστο των πιθανοτήτων είναι μεγαλύτερο από 0.5, τότε το σύστημα θεωρεί οπωσδήποτε ότι έχει εμφανιστεί μία κατάσταση.

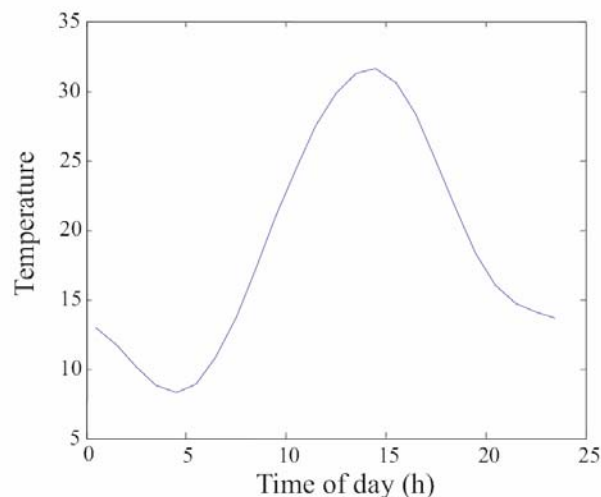
3.3 Χρήση της Δι-επίπεδης Σύντηξης για την Πυρανίχνευση

Στα πλαίσια της αξιολόγησης του προτεινομένου συστήματος χρησιμοποιήθηκαν τοπικοί αισθητήρες θερμοκρασίας και σχετικής υγρασίας και απομακρυσμένοι οπτικοί αισθητήρες για την έγκαιρη ανίχνευση της κατάστασης εμφάνισης πυρκαγιάς. Οι τοπικοί αισθητήρες αναπτύσσονται στην περιοχή που μας ενδιαφέρει να ελέγξουμε ενώ οι οπτικοί αισθητήρες ρυθμίζονται να παρατηρούν την ίδια περιοχή δίνοντας έτσι τη δυνατότητα παρουσίας πολλαπλών συμπληρωματικών πηγών πληροφορίας για το ίδιο φαινόμενο.

Το γεγονός κατά το οποίο η συνήθης θερμοκρασία του περιβάλλοντος αυξάνεται με έναν απότομο τρόπο, θα μπορούσε να υποδηλώνει πιθανή ύπαρξη πυρκαγιάς. Στην περίπτωση αυτή θεωρούμε ότι τα δεδομένα από τους αισθητήρες θερμοκρασίας έχουν μέση τιμή μ_0 σε κατάσταση μη ύπαρξης πυρκαγιάς και μ_F σε κατάσταση ύπαρξης πυρκαγιάς. Έτσι χρησιμοποιούμε την τεχνική για την ανίχνευση αλλαγής της κατανομής των δεδομένων από αισθητήρες που παρουσιάσαμε στην ενότητα 3.2.2. Κατά τη διαδικασία αυτή λαμβάνουμε υπόψη τις διακυμάνσεις για τους αισθητήρες που παρέχει

ο κατασκευαστής και θεωρούμε τα αντίστοιχα κατώφλια για την ανίχνευση της πυρκαγιάς. Επίσης οι συσχετισμένες χρονικές διακυμάνσεις λόγω εξωγενών παραγόντων (π.χ., σύννεφων) που αλλάζει τη θερμοκρασία από τη μέση τιμή για κάποια χρονική περίοδο, αντιμετωπίζονται με την τεχνική που περιγράφηκε ώστε να ακυρωθεί αυτή η αλλαγή.

Για τον προσδιορισμό της μέσης θερμοκρασίας περιβάλλοντος σε περίπτωση μη ύπαρξης πυρκαγιάς έχουν αναπτυχθεί πολλά μοντέλα αναπαράστασης. Το Σχήμα 3.5 παρουσιάζει μία χαρακτηριστική γραφική παράσταση για τις ωριαίες αλλαγές της θερμοκρασίας στη διάρκεια της ημέρας και βασίζεται στο μοντέλο που παρουσιάζεται στην [46]. Η ελάχιστη τιμή είναι $T_{min} = 12 \text{ }^\circ\text{C}$, η μέγιστη τιμή είναι $T_{max} = 35 \text{ }^\circ\text{C}$ και ο καθημερινός μέσος όρος είναι $T_{ave} = 20 \text{ }^\circ\text{C}$. Όπως παρατηρείται υπάρχουν διαστήματα στα οποία η διακύμανση της θερμοκρασίας είναι μικρή, δηλ. 2:00μ.μ. έως 3:00μ.μ. ή η θερμοκρασία μειώνεται, δηλ. 3:00μ.μ. έως 3:00π.μ. Για αυτά τα χρονικά διαστήματα μπορούμε να υποθέσουμε μια σχετικά μεγάλη τιμή του T_0 , παραδείγματος χάριν $T_0 = 30\text{min}$. Με ένα ρυθμό δειγματοληψίας $F_s = 0.5\text{Hz}$, που είναι ένα δείγμα κάθε 2 sec, το χρονικό παράθυρο για να φθάσει κανείς σε μια απόφαση είναι $n=30 \times 60 \times 0.5 = 900$ δείγματα. Για τα χρονικά διαστήματα που συντελείται απότομη αύξηση της θερμοκρασίας, δηλ. στις 5:00 π.μ. έως 12:00 π.μ, μπορούμε να υποθέσουμε μια μικρότερη τιμή, παραδείγματος χάριν $T_0 = 15\text{min}$ και να αυξήσουμε το ποσοστό δειγματοληψίας $F_s = 1\text{Hz}$ προκειμένου να ληφθεί ο ίδιος αριθμός δειγμάτων $n=900$.

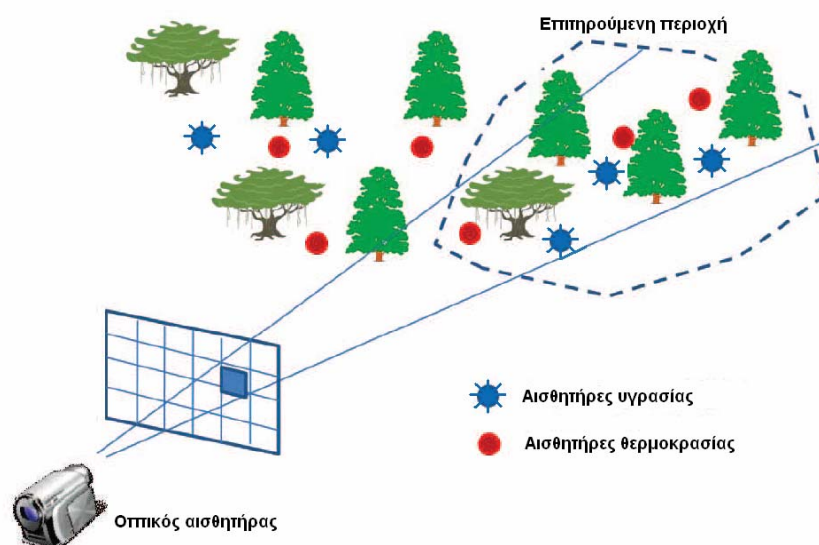


Σχήμα 3.5: Τυπικές αλλαγές της θερμοκρασίας που παρουσιάζονται στη διάρκεια της ημέρας.

Στην περίπτωση όμως που έχουμε παραπάνω από έναν αισθητήρες θερμοκρασίας στην περιοχή προσδιορίζουμε την μέση τιμή μ_0 συνδυάζοντας τις μετρήσεις από αυτούς τους αισθητήρες.

Πέραν από τους αισθητήρες θερμοκρασίας οι αισθητήρες υγρασίας μπορούν να χρησιμοποιηθούν επίσης για την πυρανίχνευση διότι η σχετική υγρασία μειώνεται σε περίπτωση εμφάνισης πυρκαγιάς. Σε αυτήν την περίπτωση η τιμή μ_0 δείχνει την σχετική υγρασία του περιβάλλοντος που επαναυπολογίζεται κάθε $T_0 \text{ min}$, και βασίζεται σε όλους του αισθητήρες υγρασίας στην περιοχή. Το ίδιο προτεινόμενο στατιστικό ελέγχου όπως στην περίπτωση της θερμοκρασίας μπορεί να εφαρμοστεί, αλλά τώρα τα κατώφλια $T(f)$ είναι αρνητικά, δηλ., $T(f)=-4.1$ για $f=0.01$ και $T(f)=-5.3$ για $f=0.001$. Όπως και στην περίπτωση των αισθητήρων θερμοκρασίας λαμβάνονται υπόψη και οι διακυμάνσεις στις μετρήσεις. Για τον αισθητήρα σχετικής υγρασίας SHT11, η ακρίβεια είναι ± 3 για τις τιμές $[20-80]\%$ και ± 5 για τις τιμές της σχετικής υγρασίας έξω από το προαναφερθέν διάστημα. Εάν υπάρχουν συσχετισμένες χρονικές διακυμάνσεις λόγω, π.χ., ενός ρεύματος κρύου αέρα που αλλάζει την υγρασία από την μέση τιμή κάποια χρονική περίοδο, τότε μπορεί να εφαρμοστεί η ίδια τεχνική όπως και στους αισθητήρες θερμοκρασίας που ακυρώνει αυτήν την αλλαγή.

Οι απομακρυσμένοι αισθητήρες σε περίπτωση πυρκαγιάς μεταδίδουν πληροφορίες σχετικά με την πιθανότητα καπνού ή φλόγας που αντιστοιχούν στις περιοχές που καλύπτονται και που συνήθως είναι μακριά. Οποιαδήποτε σημαντική αλλαγή στην αντίθεση ή στη φωτεινότητα της ελεγχόμενης σκηνής μεταφράζεται μέσω συγκεκριμένων αλγορίθμων σε μια πιθανότητα για εμφάνιση πυρκαγιάς. Το Σχήμα 3.6 απεικονίζει την κατανομή των αισθητήρων (τοπικών και απομακρυσμένων) στην περιοχή.



Σχήμα 3.6: Κατανομή των αισθητήρων στην περιοχή.

3.3.1 Ανίχνευση με βάση τους αισθητήρες θερμοκρασίας και υγρασίας

Ας υποθέσουμε δύο καταστάσεις: NOTIFY και ALERT. Στην κατάσταση NOTIFY το σύστημα ειδοποιεί για μια πιθανή αλλαγή στη μέση τιμή της θερμοκρασίας επισημαίνοντας έτσι μια πιθανή απειλή από πυρκαγιά. Στην κατάσταση ALERT υπάρχει ικανοποιητική πεποίθηση για ένα γεγονός πυρκαγιάς. Ως εκ τούτου, μπορούμε να χρησιμοποιήσουμε τα αποτελέσματα από τον Πίνακα 3-1 και να επιλέξουμε το ποσοστό των ψευδώς θετικών συναγεργμών να είναι ίσο με $f=0.01$ για την κατάσταση NOTIFY και $f=0.001$ για την κατάσταση ALERT. Όταν υπάρχουν και οι δύο τύποι αισθητήρων (θερμοκρασία και υγρασία), μπορούμε να συνδυάσουμε τις αποφάσεις σχετικά με ένα γεγονός πυρκαγιάς προκειμένου να έχουμε μία πιο ολοκληρωμένη εικόνα για το γεγονός. Συγκεκριμένα, υποθέτουμε ότι τα αντίστοιχα κατώτατα όρια για την μετάπτωση στην κατάσταση NOTIFY και ALERT είναι αντίστοιχα $T_{TN}(f)$, $T_{TA}(f)$ και $T_{HN}(f)$, $T_{HA}(f)$. Παραδείγματος χάριν, για $f=0.01$, $T_{TN}(f)=4.1$ και $T_{HN}(f)=-4.1$ ενώ για $f=0.001$, $T_{TA}(f)=5.3$ και $T_{HA}(f)=-5.3$. Η προτεινόμενη σύντηξη στο πρώτο επίπεδο βασίζεται στην εφαρμογή διάφορων τελεστών στα αντίστοιχα κατώφλια (thresholds).

Ανίχνευση για την κατάσταση NOTIFY

Μπορούμε να συνδυάσουμε τα χαμηλότερα κατώτατα όρια μέσω του ελάχιστου τελεστή (min) σε μία προσπάθεια να αποφασίσουμε πότε ο κόμβος θα μεταπέσει στην κατάσταση NOTIFY. Θεωρήστε ότι το t_N δείχνει τη χρονική στιγμή που το σύστημα μεταπίπτει στην κατάσταση NOTIFY. Επίσης τα $k^{-1/2} W_k^T$ και $k^{-1/2} W_k^H$ υποδεικνύουν το στατιστικό ελέγχου που εφαρμόζεται στα δεδομένα θερμοκρασίας και υγρασίας αντίστοιχα. Τότε έχουμε:

$$t_N = \arg \left\{ \min_k \left\{ k^{-1/2} W_k^T > T_{TN}(f) \text{ ή } k^{-1/2} W_k^H < T_{HN}(f) \right\} \right\} \quad (3.19)$$

Με την υιοθέτηση αυτής της επιλογής το σύστημα θα μεταπέσει στην κατάσταση NOTIFY τη στιγμή που είτε ο αισθητήρας θερμοκρασίας είτε ο αισθητήρας υγρασίας υπερβούν το αντίστοιχο κατώτατο όριο.

Ανίχνευση για την κατάσταση ALERT

Προκειμένου να συμπεράνουμε για την κατάσταση ALERT διάφορα κριτήρια μπορεί να εφαρμοστούν ως εξής:

- Εάν οι μετρήσεις σχετικής υγρασίας είναι πιο αξιόπιστες από τις μετρήσεις θερμοκρασίας (δηλ., οι αισθητήρες είναι καλύτεροι) μπορούμε να βασίσουμε τη διαδικασία ανίχνευσης μόνο στο στατιστικό ελέγχου $k^{-1/2} W_k^H$. Ως εκ τούτου, εάν ισχύει

$$k^{-1/2} W_k^H < T_{HA}(f) \quad (3.20)$$

τότε το σύστημα μεταπίπτει στην κατάσταση ALERT.

- Ένα ασφαλέστερο κριτήριο είναι να χρησιμοποιηθούν και τα δύο στατιστικά ελέγχου δηλαδή εάν ισχύει

$$k^{-1/2} W_k^T > T_{TA}(f) \text{ \& } k^{-1/2} W_k^H < T_{HA}(f) \quad (3.21)$$

τότε το σύστημα μεταπίπτει στην κατάσταση ALERT.

- Εναλλακτικά, μπορούμε να χρησιμοποιήσουμε το κριτήριο

$$k^{-1/2} (W_k^T - W_k^H) > a_1 T_{TA}(f) - a_2 T_{HA}(f) \quad (3.22)$$

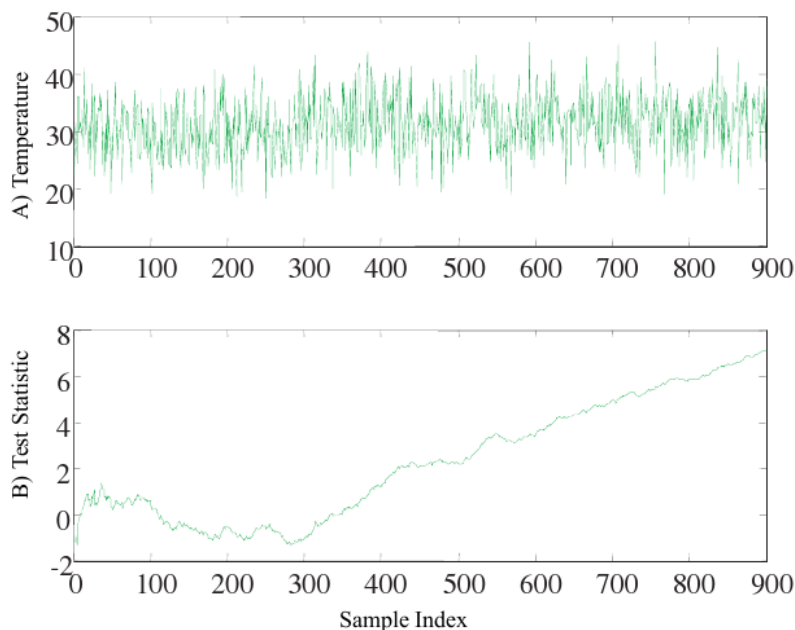
όπου τα a_1 και a_2 είναι σταθερές στάθμισης. Για $a_1 = a_2 = 1$ η παραπάνω σχέση βρίσκεται μεταξύ του κριτηρίου (3.20) και (3.21).

3.3.2 Σενάρια αξιολόγησης με τη χρήση της τεχνικής CUSUM

Στις επόμενες παραγράφους θα παρουσιάσουμε μερικά αποτελέσματα προσομοίωσης βασισμένα σε σενάρια, που υπογραμμίζουν τη δυνατότητα που παρέχει η τεχνική CUSUM για την έγκαιρη ανίχνευση της πυρκαγιάς με βάση μετρήσεις από δεδομένα θερμοκρασίας. Υποθέτουμε ο ρυθμός δειγματοληψίας είναι $F_s=0.5\text{Hz}$, δηλαδή ένα δείγμα κάθε 2 sec. Η εκτίμηση για την τιμή της μέσης θερμοκρασίας γίνεται κάθε 30 λεπτά και επομένως το χρονικό παράθυρο για να ληφθεί μια απόφαση είναι $n = 30 \times 60 \times 0.5 = 900$ δείγματα.

Σενάριο 1

Σε αυτό το σενάριο, υποθέτουμε ότι μια πυρκαγιά εμφανίζεται 10 λεπτά ($\tau = 300$) μετά την εκτίμηση της μέσης θερμοκρασίας του περιβάλλοντος μ_0 . Η πυρκαγιά προκαλεί μια μέση αύξηση της θερμοκρασίας από $\mu_0 = 30^\circ\text{C}$ στους $\mu_F = 32^\circ\text{C}$ στις μετρήσεις ενός αισθητήρα και η τυπική απόκλιση θεωρείται $\sigma = 5$. Σημειώστε ότι η μ_F είναι μια άγνωστη παράμετρος που επηρεάζει την κλίση της μετατόπισης του στατιστικού ελέγχου. Το Σχήμα 3.7α παρουσιάζει τα δείγματα των μετρήσεων, ενώ το Σχήμα 3.7β απεικονίζει την εξέλιξη του στατιστικού ελέγχου.



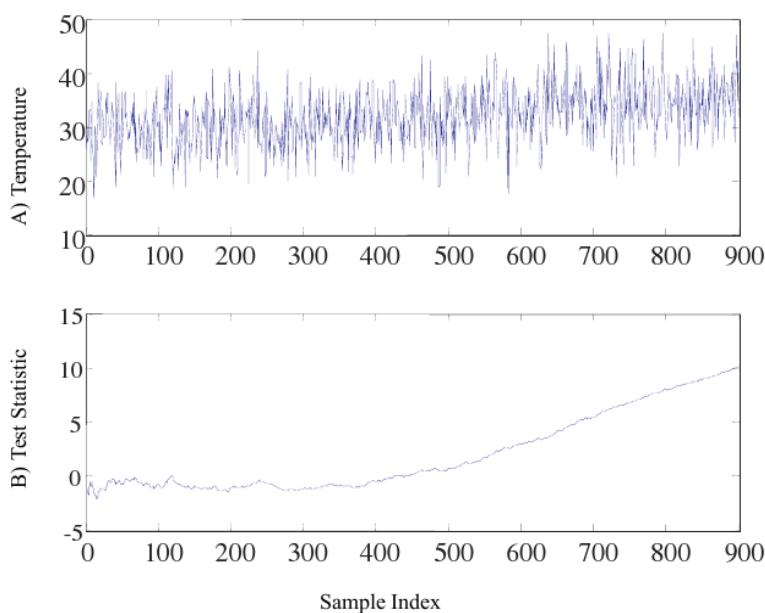
Σχήμα 3.7: (α) Τιμές θερμοκρασίας σε $^\circ\text{C}$, (β) η εξέλιξη του στατιστικού ελέγχου 10 λεπτά μετά την εμφάνιση της πυρκαγιάς ($\tau = 300$) με τιμή $\mu_d = \mu_F - \mu_0 = 2$.

Χρησιμοποιώντας τα προαναφερθέντα ποσοστά ψευδών θετικώς συναγερμών $f = 0.01$ (κατώφλι $T(f) = 4.1$) και $f = 0.001$ (κατώφλι $T(f) = 5.3$), το σύστημα θα μεταπέσει στην κατάσταση NOTIFY μετά από περίπου 300 δείγματα (10 λεπτά) και στην κατάσταση ALERT μετά από περίπου 400 δείγματα (13,5 λεπτά). Σημειώστε, επίσης, ότι αν και η θερμοκρασία έχει μεγάλη διακύμανση ($20^\circ\text{C} - 40^\circ\text{C}$), το στατιστικό ελέγχου που χρησιμοποιείται δε επηρεάζεται από τις στιγμιαίες αλλαγές.

Σενάριο 2

Σε αυτό το σενάριο, υποθέτουμε ότι το μέτωπο της πυρκαγιάς πλησιάζει έναν αισθητήρα θερμοκρασίας. Η πυρκαγιά εμφανίζεται 10 λεπτά ($\tau = 300$) μετά την εκτίμηση

της μέσης θερμοκρασίας του περιβάλλοντος μ_0 και προκαλεί μια βαθμιαία αύξηση της θερμοκρασίας με ρυθμό 0.01°C ανά δείγμα. Αυτό σημαίνει ότι η μέση θερμοκρασία αυξάνει 1°C κάθε 3,3 λεπτά. Η αρχική θερμοκρασία και η τυπική απόκλιση τίθενται όπως στο προηγούμενο σενάριο σε $\mu_0 = 30^\circ\text{C}$ και $\sigma = 5$. Το Σχήμα 3.8α παρουσιάζει τα δείγματα των μετρήσεων, ενώ το Σχήμα 3.8β απεικονίζει την εξέλιξη του στατιστικού ελέγχου.



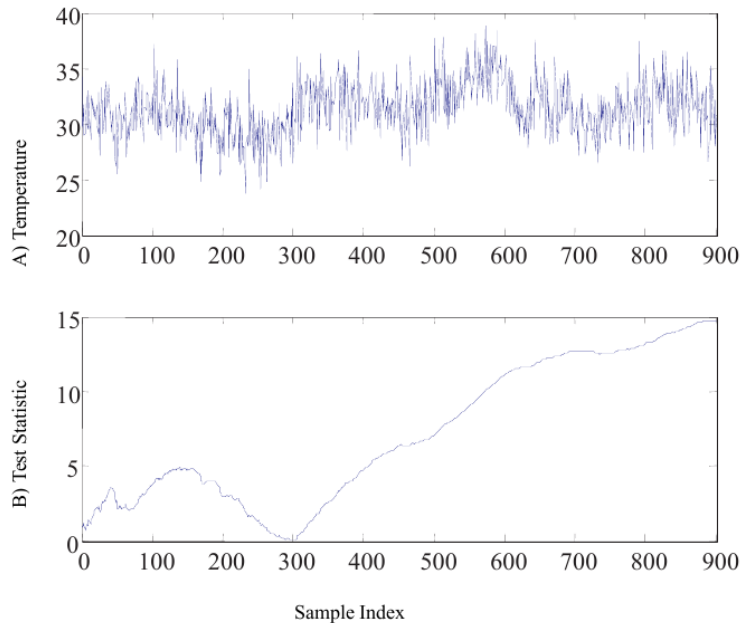
Σχήμα 3.8: (α) Τιμές θερμοκρασίας σε $^\circ\text{C}$, (β) η εξέλιξη του στατιστικού ελέγχου 10 λεπτά μετά την εμφάνιση της πυρκαγιάς ($\tau = 300$) με βαθμιαία αύξηση της θερμοκρασίας 0.01°C /δείγμα.

Δεδομένου ότι ο ρυθμός αλλαγής της θερμοκρασίας είναι μικρός, το σύστημα θα εντοπίσει την αλλαγή (και θα μεταπέσει σε κατάσταση NOTIFY) μετά από 350 δείγματα, δηλαδή 12 λεπτά. Εντούτοις, η βαθμιαία αύξηση της θερμοκρασίας θα αναγκάσει το σύστημα να μεταπέσει στην κατάσταση ALERT αρκετά γρήγορα (μετά από 700 δείγματα).

Σενάριο 3

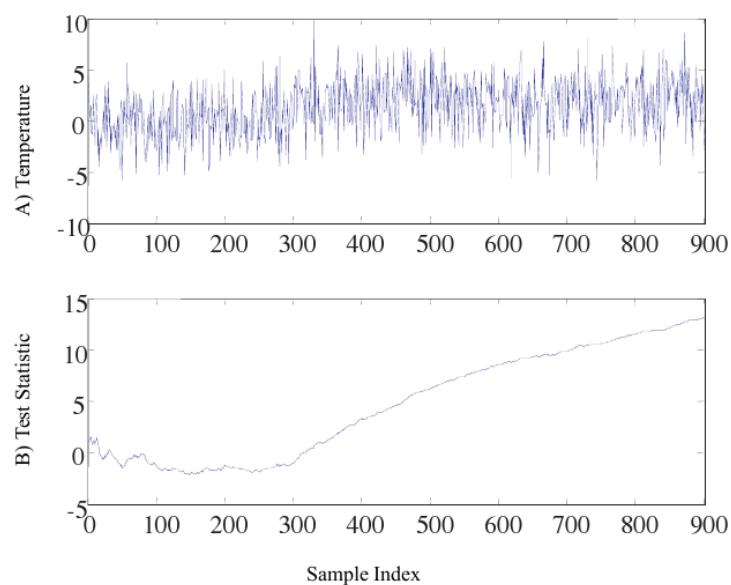
Το σενάριο αυτό προσομοιώνει την περίπτωση των συσχετισμένων μετρήσεων. Το Σχήμα 3.9α απεικονίζει μία συνάρτηση της διαδικασίας που περιγράφεται από τη σχέση (3.12). Η μέση τιμή μ_0 τέθηκε ως 30°C , η $\sigma_z = 2$, η $\sigma_m = 1$ και η παράμετρος $\alpha = (120\text{sec})^{-1}$. Μια αλλαγή των πυκνοτήτων εμφανίζεται στο χρόνο $\tau = 300$ με διαφορά μέσης τιμής $\mu_d = 2^\circ\text{C}$. Αυτή η τιμή χρησιμοποιείται μόνο ενδεικτικά για την προσομοίωση. Υψηλότερες τιμές, οδηγούν το στατιστικό ελέγχου να μετακινηθεί γρηγορότερα προς τις

τιμές των κατωφλίων που έχουν τεθεί χρησιμοποιώντας λιγότερα δείγματα. Το Σχήμα 3.9β παρουσιάζει την εξέλιξη του στατιστικού ελέγχου. Όπως παρατηρείται αρχίζει να μετατοπίζεται σημαντικά μετά από $\tau = 300$ αλλά υπάρχει περίπτωση να διασχίσει τις τιμές των κατωφλίων που έχουν τεθεί αρκετά νωρίτερα παράγοντας κατά συνέπεια ψευδείς συναγερμούς για την εμφάνιση πυρκαγιάς.



Σχήμα 3.9: (α) Τιμές θερμοκρασίας της διαδικασίας X_i που περιγράφεται από τη σχέση (3.12), (β) η εξέλιξη του στατιστικού ελέγχου.

Το Σχήμα 3.10 παρουσιάζει τα αποτελέσματα προσομοίωσης για την τεχνική που προτείνεται έτσι ώστε να μετριάσουν οι συσχετισμένες μετρήσεις. Υποθέτουμε ότι ο αισθητήρας S_i έχει τρεις γείτονες με τις αντίστοιχες καθυστερήσεις, που μετριοούνται σε δείγματα, $D_1=-3$ (χρονική προπορεία), $D_2=2$ (καθυστέρηση) και $D_3=5$ (καθυστέρηση). Ο θόρυβος z_j για κάθε αισθητήρα είναι ανεξάρτητος Gaussian κατανομής με μέση τιμή μηδέν και τυπική απόκλιση $\sigma_z=2$. Επίσης υποθέτουμε ότι $\mu_d = 2^\circ\text{C}$. Μια αλλαγή των πυκνοτήτων εμφανίζεται στο χρόνο $\tau=300$ και όπως απεικονίζεται στο Σχήμα 3.10 δεν υπάρχει κανένας ψευδώς θετικός συναγερμός πριν από το τ .



Σχήμα 3.10: (α) Τιμές θερμοκρασίας της διαδικασίας Y_i που περιγράφεται από τη σχέση (3.14), (β) η εξέλιξη του στατιστικού ελέγχου συμπεριλαμβάνοντας τα δεδομένα από τους γειτονικούς κόμβους.

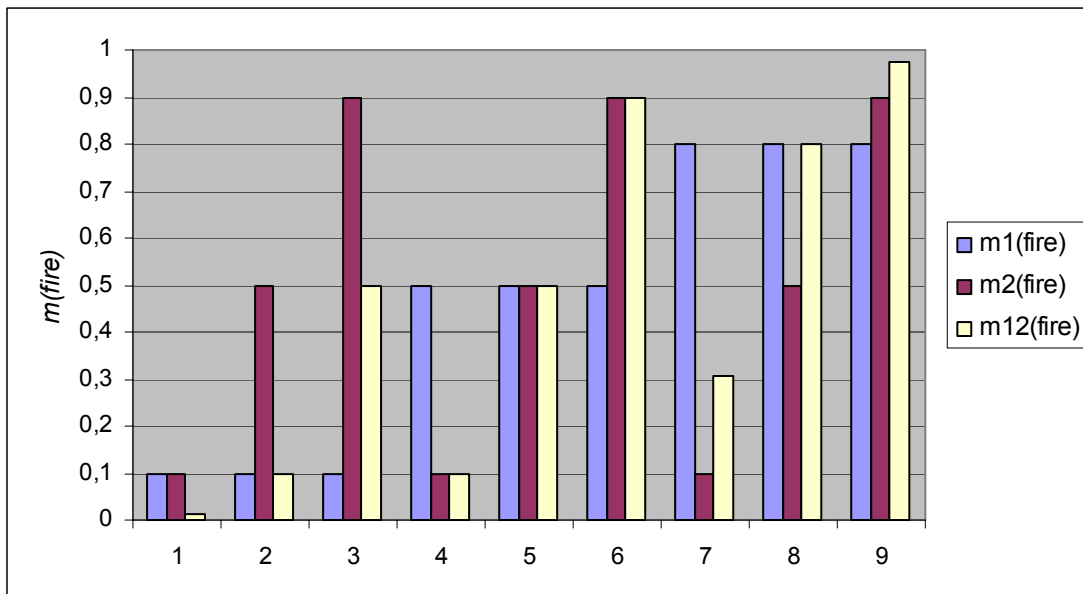
3.3.3 Σενάρια αξιολόγησης με τη χρήση της θεωρίας DST

Σενάριο 1

Σε αυτό το σενάριο υποθέτουμε ότι $M = 1$ (προκειμένου να ελαχιστοποιηθεί το ποσοστό ψευδώς θετικών συναγερμών), που σημαίνει ότι συνδυάζεται η μέγιστη πιθανότητα που λαμβάνεται από τους τοπικούς αισθητήρες με την πιθανότητα για ενδεχόμενο πυρκαγιάς από τον απομακρυσμένο οπτικό αισθητήρα. Υποθέτουμε τρεις διαφορετικές πιθανότητες πυρκαγιάς για τον οπτικό αισθητήρα (0.1, 0.5, 0.9) και η καθεμία συνδυάζεται με τρεις διαφορετικές πιθανότητες (0.1, 0.5, 0.8) που προκύπτουν από του τοπικούς αισθητήρες. Ο Πίνακας 3-2 και το παρουσιάζει τα αποτελέσματα του συνδυασμού του προτεινομένου αλγορίθμου που βασίζεται στη θεωρία Dempster-Shafer. Όπως παρατηρείται η τελική πιθανότητα διατηρείται σε σχετικά μικρές τιμές εκτός αν και οι δύο πιθανότητες που συνδυάζονται υπερβαίνουν την τιμή 0.5. Κατά συνέπεια, ένας αισθητήρας που δυσλειτουργεί δεν είναι ικανός να προκαλέσει έναν συναγερμό πυρκαγιάς. Επίσης οι υψηλές τιμές πιθανότητας και των δύο αισθητήρων, θα ενισχύσουν την εμπιστοσύνη μας για ένα γεγονός πυρκαγιάς.

Πίνακας 3-2: Η εφαρμογή του αλγορίθμου DS στη σύντηξη των πιθανοτήτων ενός τοπικού αισθητήρα και ενός απομακρυσμένου αισθητήρα σχετικά με το γεγονός πυρκαγιάς.

| Τοπικός αισθητήρας (Sensor 1) $m_1(\text{fire})$ | Οπτικός αισθητήρας (Sensor V) $m_2(\text{fire})$ | Σύγκρουση Ενδείξεων (Conflict- K) | Συνδυασμένες πιθανότητες (Sensor 1 + Sensor V) $m_{12}(\text{fire})$ |
|---|---|-----------------------------------|---|
| 0.1 | 0.1 | 0.18 | 0.0122 |
| 0.1 | 0.5 | 0.50 | 0.1000 |
| 0.1 | 0.9 | 0.82 | 0.5000 |
| 0.5 | 0.1 | 0.50 | 0.1000 |
| 0.5 | 0.5 | 0.50 | 0.5000 |
| 0.5 | 0.9 | 0.50 | 0.9000 |
| 0.8 | 0.1 | 0.74 | 0.3077 |
| 0.8 | 0.5 | 0.50 | 0.8000 |
| 0.8 | 0.9 | 0.26 | 0.9730 |



Σχήμα 3.11: Η εφαρμογή του αλγορίθμου DS στη σύντηξη των πιθανοτήτων.

Σενάριο 2

Σε αυτό το σενάριο υποθέτουμε ότι $M = 2$ που σημαίνει ότι συνδυάζουμε τις πιθανότητες που προέκυψαν από το πρώτο σενάριο - $m_{12}(\text{fire})$ – με τη πιθανότητα που προκύπτει από ένα άλλο τοπικό αισθητήρα (Sensor 3). Για αυτόν τον αισθητήρα υποθέτουμε δύο τιμές (0.2 και 0.6) και τα αποτελέσματα απεικονίζονται στον και αντίστοιχα. Όπως παρατηρείται από τις καταχωρήσεις στον Πίνακα 3-3 (σειρές 3 και 5), η τιμή 0.5 όταν συνδυάζεται με τη μικρή πιθανότητα του δεύτερου τοπικού αισθητήρα δηλαδή 0.2 παράγει τη χαμηλότερη τιμή, η οποία είναι 0.2. Αντίθετα, όπως παρατηρείται από (σειρές 3 και 5) του Πίνακα 3-4, η τιμή 0.5 όταν συνδυάζεται με μια υψηλότερη πιθανότητα, δηλαδή 0.6, αυτό παράγει τη μεγαλύτερη τιμή, δηλαδή 0.6. Εάν όλες οι πιθανότητες είναι μεγαλύτερες από 0.5 (δύο τελευταίες σειρές στον Πίνακα 3-4), η πεποίθησή μας για ένα γεγονός πυρκαγιάς ενισχύεται όπως υποδεικνύεται από τις υψηλές τιμές της συνδυασμένης πιθανότητας $m_{123}(\text{fire})$.

Πίνακας 3-3: Η εφαρμογή του αλγορίθμου DS στη σύντηξη των πιθανοτήτων με τη συμμετοχή ενός δεύτερου τοπικού αισθητήρα (πιθανότητα 0.2)

| Συνδυασμένες πιθανότητες (Sensor 1 + Sensor V) $m_{12}(\text{fire})$ | Τοπικός αισθητήρας (Sensor 3) $m_3(\text{fire})$ | Σύγκρουση Ενδείξεων (Conflict- K) | Συνδυασμένες πιθανότητες $m_{123}(\text{fire})$ |
|---|---|---|---|
| 0.0122 | 0.2 | 0.2073 | 0.0031 |
| 0.1000 | 0.2 | 0.2600 | 0.0270 |
| 0.5000 | 0.2 | 0.5000 | 0.2000 |
| 0.1000 | 0.2 | 0.2600 | 0.0270 |
| 0.5000 | 0.2 | 0.5000 | 0.2000 |
| 0.9000 | 0.2 | 0.7400 | 0.6923 |
| 0.3077 | 0.2 | 0.3846 | 0.1000 |
| 0.8000 | 0.2 | 0.6800 | 0.5000 |
| 0.9730 | 0.2 | 0.7837 | 0.9000 |

Πίνακας 3-4: Η εφαρμογή του αλγορίθμου DS στη σύντηξη των πιθανοτήτων με τη συμμετοχή ενός δεύτερου τοπικού αισθητήρα (πιθανότητα 0.6)

| Συνδυασμένες πιθανότητες (Sensor 1 + Sensor V) $m_{12}(\text{fire})$ | Τοπικός αισθητήρας (Sensor 3) $m_3(\text{fire})$ | Σύγκρουση Ενδείξεων (Conflict- K) | Συνδυασμένες πιθανότητες $m_{123}(\text{fire})$ |
|---|---|-----------------------------------|---|
| 0.0122 | 0.6 | 0.5975 | 0.0182 |
| 0.1000 | 0.6 | 0.5800 | 0.1429 |
| 0.5000 | 0.6 | 0.5000 | 0.6000 |
| 0.1000 | 0.6 | 0.5800 | 0.1429 |
| 0.5000 | 0.6 | 0.5000 | 0.6000 |
| 0.9000 | 0.6 | 0.4200 | 0.9310 |
| 0.3077 | 0.6 | 0.5384 | 0.4000 |
| 0.8000 | 0.6 | 0.4400 | 0.8571 |
| 0.9730 | 0.6 | 0.4054 | 0.9818 |

3.4 Συμπεράσματα

Η έρευνα που διεξήχθη στο πεδίο της επίγνωσης κατάστασης του περιβάλλοντος στο διάχυτο υπολογισμό συνεισέφερε στα παρακάτω βασικά σημεία:

- Πρόταση μηχανισμού για την επίγνωση της κατάστασης του περιβάλλοντος που βασίζεται στη σύντηξη δεδομένων από πολλαπλούς αισθητήρες και υλοποιείται σε ένα δι-επίπεδο σχήμα σύντηξης.
- Χρήση της τεχνικής CUSUM για την αξιόπιστη ανίχνευση αλλαγής στα δεδομένα των αισθητήρων.
- Τεχνική συνδυασμού των μετρήσεων γειτονικών αισθητήρων για την αποφυγή ψευδών συναγερμών.
- Εφαρμογή της Θεωρίας Ενδείξεων DST για άμεση ανταπόκριση του συστήματος με όσο το δυνατόν χαμηλότερο ποσοστό ψευδών συναγερμών.

Τα θέματα που αναλύθηκαν στο κεφάλαιο αυτό αναφέρονται στις δημοσιεύσεις: [50], [51], [52].

ΚΕΦΑΛΑΙΟ 4

ΒΙΟ-ΜΙΜΗΤΙΚΗ ΣΥΝΕΡΓΑΤΙΚΗ ΕΠΙΓΝΩΣΗ ΠΛΑΙΣΙΟΥ

4.1 Εισαγωγή

Στην δεύτερη θεματική ενότητα της διατριβής εξετάστηκαν θέματα που αφορούν τη συνεργατική επίγνωση πλαισίου στο διάχυτο υπολογισμό. Η συνεργατική επίγνωση πλαισίου (ΣΕΠ) αποτελεί την κατανεμημένη εκδοχή της επίγνωσης πλαισίου και υποδηλώνει την κατανόηση της πληροφορίας από όλα τα μέλη μιας ομάδας παρέχοντας κατά συνέπεια μια πιο αναβαθμισμένη πληροφορία για κάθε μέλος της ίδιας ομάδας ξεχωριστά. Η διάταξη των κινητών κόμβων σε ένα δίκτυο εισάγει την ανάγκη να αντιμετωπιστεί η επίγνωση πληροφορίας πλαισίου με συνεργατικό τρόπο. Μια τέτοιου είδους συνεργασία υποδηλώνει τη σύμπραξη μεταξύ των γειτονικών κόμβων για την αντίληψη, την ερμηνεία και τη διανομή της πληροφορίας πλαισίου (π.χ., ο κόμβος Α έλαβε μια συγκεκριμένη πληροφορία που παράχθηκε και επεξεργάσθηκε στον κόμβο Β, ο κόμβος Γ συμπεραίνει επιπλέον πληροφορίες από αυτές που διαδόθηκαν από τον κόμβο Δ). Η συνεργατική πληροφορία πλαισίου είναι η πληροφορία που αποκτάται μέσω του δικτύου (π.χ. εγγύτητα) μεταξύ των αισθητήρων και των συσκευών υψηλότερου επιπέδου και μπορεί να χρησιμοποιηθεί για:

1. να αυξήσει την κοινή κατανόηση για τη γειτονιά κάθε κόμβου
2. να βελτιώσει τη διαθεσιμότητα των προσφερόμενων υπηρεσιών
3. να ενισχύσει την αξιοπιστία της πληροφορίας μέσω των πρόσθετων / συμπληρωματικών πληροφοριών από τους γείτονες.

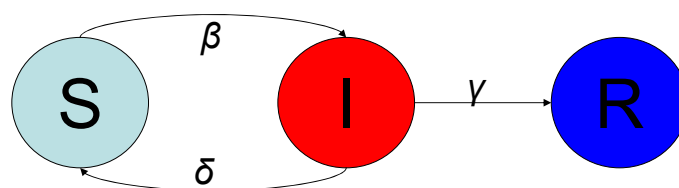
Η διάδοση της πληροφορίας πλαισίου σε δίκτυα κινητών κόμβων πρέπει να πραγματοποιηθεί μέσω ενός απλού, πρωτοκόλλου ανταλλαγής μηνυμάτων ώστε να ικανοποιηθούν οι επιθυμητές απαιτήσεις διάδοσης πληροφορίας, δηλαδή να ικανοποιηθούν οι χρονικοί και χωρικοί περιορισμοί. Η πληροφορία πρέπει να φθάσει στους πιθανούς παραλήπτες μέσα σε ένα ορισμένο χρονικό διάστημα (π.χ., σε λιγότερο από 2 λεπτά). Επιπλέον, η πληροφορία θα μπορούσε να θεωρηθεί έγκυρη μόνο μέσα σε μια ορισμένη γεωγραφική περιοχή. Αυτό σημαίνει ότι, δεν πρέπει να διαδοθεί παντού. Αντ' αυτού, τα λεπτομερή δεδομένα θέσης που συλλαμβάνονται από έναν κόμβο μπορούν να χρησιμοποιηθούν μόνο από τους γειτονικούς / κοντινούς τού κόμβους.

Το γεγονός ότι οι κόμβοι σε ένα περιβάλλον διάχυτου υπολογισμού κινούνται συνεχώς και δημιουργούν στοχαστικές συνδέσεις μεταξύ τους, μειώνει τη δυνατότητα εφαρμογής των αλγορίθμων ανακαλύψεως της γειτονιάς (neighbourhood discovery schemes) και

των σύνθετων πρωτοκόλλων δρομολόγησης. Τυπικά οι κόμβοι αντλούν την ενέργειά τους από μπαταρίες, ως εκ τούτου το σχήμα διάδοσης πληροφορίας που ακολουθείται πρέπει να είναι όσο το δυνατόν αποδοτικότερο. Αν και ένα απλό σχήμα πλημμύρας (flooding) θα μπορούσε να υποστηρίξει τις βασικές προαναφερθείσες απαιτήσεις (χρονικές και χωρικές), δεν προστατεύει την ενεργειακή αποδοτικότητα των κόμβων και τελικά προκύπτουν πολλές, ανεπιθύμητες μεταδόσεις. Άλλα σχήματα περισσότερο εξελιγμένα από αυτό της πλημμύρας είναι λιγότερο επιθυμητά δεδομένου ότι αποτυγχάνουν να λειτουργήσουν στις ιδιαίτερα δυναμικές τοπολογίες δικτύων. Η τοπολογία του δικτύου είναι γενικά άγνωστη και η συνοχή αυτού μπορεί να είναι αρκετά μεταβλητή ανάλογα με το περιβάλλον.

Υπάρχει μια ενδιαφέρουσα αναλογία μεταξύ της διάδοσης της πληροφορίας πλαισίου ανάμεσα στους κόμβους ενός δικτύου και της διάδοσης των μολυσματικών ασθενειών μεταξύ των ατόμων. Και οι δύο είναι διαδικασίες στις οποίες, κάποια μορφή επαφής συμβάλλει στο να μεταδοθεί «κάτι». Η διαδεδομένη εμφάνιση μιας ασθένειας, που μολύνει τα άτομα, αναφέρεται ως *επιδημία* [73]. Η φύση της επιδημικής διάδοσης [74] υιοθετείται ως τεχνική για τη μετάδοση πληροφορίας πλαισίου σε ένα δίκτυο κινητών κόμβων. Τα επιδημικά μοντέλα υποθέτουν ότι οι μεμονωμένοι κόμβοι μεταπίπτουν μεταξύ των καταστάσεων (μολυσμένοι και εύτρωτοι/επιρρεπείς) με ένα ορισμένο σταθερό ρυθμό. Μετά από κάποια χρονική περίοδο, οι *μολυσμένοι* κόμβοι μπορούν να ανακτήσουν από την επιδημία και να διέλθουν έπειτα στην κατάσταση «εύτρωτος» / «επιρρεπής» με ένα ρυθμό δ . Σε αυτήν την κατάσταση, μπορούν να μολυνθούν πάλι, με ένα ρυθμό μόλυνσης β .

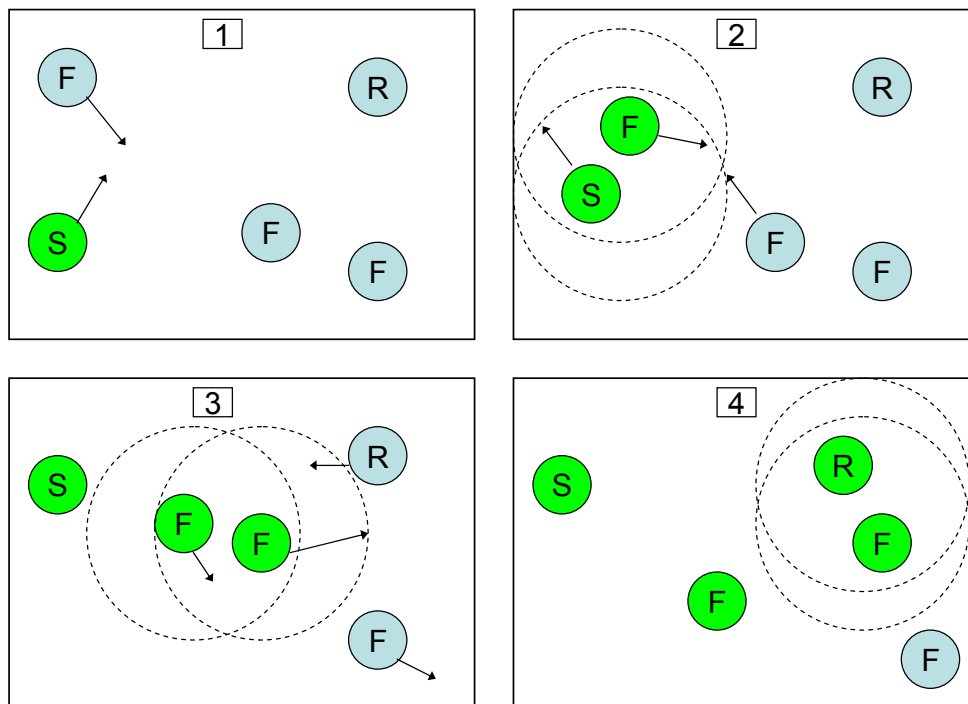
Σύμφωνα με το μοντέλο των Kermack και McKendrick [75], ένας κόμβος μπορεί να βρίσκεται σε τρεις καταστάσεις: μολυσμένος (infected), «εύτρωτος» / «επιρρεπής» (susceptible), δηλαδή είναι επιρρεπής στη μόλυνση από την επιδημία και τέλος στην κατάσταση που έχει αποκτήσει ανοσία (recovered). Αυτό το είδος του μοντέλου αναφέρεται συνήθως ως SIR (Susceptible-Infected-Recovered) (Σχήμα 4.1).



Σχήμα 4.1: Το μοντέλο SIR (Susceptible-Infected-Recovered).

Μια απλουστευμένη περίπτωση του SIR μοντέλου είναι το SIS (Susceptible-Infected-Susceptible) στο οποίο ένας κόμβος μπορεί να βρίσκεται μόνο σε δύο καταστάσεις, μολυσμένος και «εύρωτος» / «επιρρεπής». Αυτό σημαίνει ότι ένας κόμβος δεν αποκτά ποτέ ανοσία μετά από τη μόλυνσή του με την επιδημία. Επιπλέον, μία ακόμα πιο απλουστευμένη μορφή είναι το μοντέλο SI (Susceptible-Infected) σύμφωνα με το οποίο ο κόμβος δεν επιστρέφει ποτέ στην κατάσταση «εύρωτος» / «επιρρεπής» αν έστω και μία φορά έχει μολυνθεί, δηλαδή παραμένει για πάντα μολυσμένος.

Η πληροφορία πλαισίου διαδίδεται φαινομενικώς από τον κόμβο αποστολέα κατά τρόπο επαναλαμβανόμενο που «μολύνει» τους γειτονικούς κόμβους (Σχήμα 4.2). Τα πλεονεκτήματα αυτού του σχήματος είναι ότι (α) δεν υπάρχει καμία ανάγκη για τη γνώση της τοπολογίας του δικτύου ή ακόμα και της θέσης των κόμβων και (β) η επικοινωνία μεταξύ των αποσυνδεδεμένων κόμβων είναι δυνατή. Έχει αποδειχθεί επίσης ότι ακόμα και χωρίς τη σπατάλη πολλών πόρων η παράδοση των μηνυμάτων είναι εγγυημένη σε πεπερασμένο χρόνο [92].



Σχήμα 4.2: Ο κόμβος πηγή (S) παράγει πληροφορία προοριζόμενη για τον κόμβο (R), που προωθείται μέσω των κόμβων (F) με στοχαστικό τρόπο, μεταφέροντας τελικά το μήνυμα από την πηγή στο προορισμό.

Αν και τα επιδημικά σχήματα αποτελούν μια αποτελεσματικότερη λύση από ενεργειακής άποψης σε σύγκριση με άλλες τεχνικές διάδοσης (π.χ. πλημμύρα), η επιλογή του ρυθμού μόλυνσης β είναι εξαιρετικά σημαντική και μπορεί να βελτιώσει περαιτέρω την ενεργειακή απόδοση του σχήματος. Στα πλαίσια της διατριβής μελετήθηκαν θέματα που σχετίζονται με την προσαρμογή του ρυθμού μόλυνσης β ανάλογα με τις συνθήκες που επικρατούν στο δίκτυο (κινητικότητα, πυκνότητα) αλλά και την φύση της μεταδιδόμενης πληροφορίας πλαισίου (variability - μεταβλητότητα). Στις επόμενες ενότητες προτείνουμε και αξιολογούμε τεχνικές οι οποίες προσαρμόζουν το ρυθμό μόλυνσης β καθιστώντας έτσι δυνατή την εξοικονόμηση ενέργειας αλλά ταυτόχρονα συντελούν στην επαρκή διάδοση της πληροφορίας πλαισίου.

4.2 Προσαρμογή ρυθμού μόλυνσης βάσει των χαρακτηριστικών του δικτύου

Σε αυτήν την ενότητα προτείνουμε και υλοποιούμε τρεις αλγόριθμους διάδοσης πλαισίου, τον $\Delta 0$, τον $\Delta 1$ και τον $\Delta 2$. Ο αλγόριθμος $\Delta 0$ υιοθετεί μια μη-προσαρμοστική τεχνική διάδοσης, ενώ οι $\Delta 1$ και $\Delta 2$ παρουσιάζουν μια προσαρμοστική συμπεριφορά στη διάδοση πλαισίου.

4.2.1 Μη προσαρμοστικός αλγόριθμος $\Delta 0$

Ο αλγόριθμος $\Delta 0$ υποθέτει ένα στατικό περιβάλλον δικτύου, δηλ., με μηδενική κινητικότητα των κόμβων. Οι παράμετροι και οι κανόνες του $\Delta 0$ αλγορίθμου είναι:

- Ένα μήνυμα m χαρακτηρίζεται από τη χρονική παράμετρο α . Η παράμετρος αυτή (Time-To-Live ή TTL) δηλώνει τη διάρκεια μεταξύ του χρόνου παραγωγής του μηνύματος και του χρόνου μέσα στον οποίο η πληροφορία πλαισίου που φέρει το μήνυμα m θεωρείται ακόμα έγκυρη.
- Ένας κόμβος αποφασίζει εάν θα διαβιβάσει ένα μήνυμα m ή όχι, βάσει του ρυθμού μόλυνσης $\beta \in [0,1]$. Ο κόμβος που παράγει το μήνυμα στην πρώτη προσπάθεια το στέλνει με πιθανότητα $\beta = 1$. Η τιμή του β είναι σταθερή παράμετρος εισόδου στον $\Delta 0$.
- Ένας κόμβος καθυστερεί στην αποστολή του πρώτου αντιγράφου ενός λαμβανόμενου μηνύματος m . Αυτό μετριάζει τον κίνδυνο συστηματικών συγκρούσεων. Ο χρόνος μεταξύ δύο διαδοχικών προσπαθειών αποστολής του ίδιου μηνύματος εκφράζεται από την παράμετρο καθυστέρησης γ που είναι ομοιόμορφα κατανεμημένη στο διάστημα $[0, 1]$.

- Όταν ένας κόμβος λαμβάνει ένα μήνυμα m στο επίπεδο δικτύου, αποφασίζει εάν θα το διαβιβάσει στο επίπεδο εφαρμογής ή όχι. Η πολιτική για την παρούσα απόφαση είναι να διαβιβαστεί το μήνυμα μόνο μία φορά στο επίπεδο εφαρμογής. Τα άλλα αντίγραφα του μηνύματος απορρίπτονται.

Με βάση τη μελέτη των προσομοιώσεων του Δ0 αλγορίθμου, οδηγούμαστε σε μερικές ενδείξεις για το σχεδιασμό των πολιτικών προσαρμογής του ρυθμού μόλυνσης β για τους αλγόριθμους Δ1 και Δ2. Συγκεκριμένα,

Ένδειξη-1: ο ρυθμός μόλυνσης β πρέπει να μειώνεται με την πυκνότητα d του δικτύου.

Ένδειξη-2: μια επιθυμητή κάλυψη με ένα χαμηλό ενεργειακό κόστος επιτυγχάνεται μόλις αυξήσουμε την τιμή του β αλλά και το χρόνο επανεκπομπής του ίδιου μηνύματος.

Ένδειξη-3: ο αλγόριθμος πρέπει επίσης να παρατηρεί ή να ενημερωθεί για την κατάσταση του MAC επιπέδου. Όταν απορρίπτονται πακέτα στο επίπεδο MAC ο ρυθμός μόλυνσης β πρέπει να μειωθεί έως ότου ο αριθμός των απορριφθέντων μηνυμάτων να γίνει αμελητέος. Το συμπέρασμα αυτό προήλθε από σενάριο προσομοίωσης με αυξημένη κίνηση στο δίκτυο.

Προφανώς, ο ρυθμός β πρέπει να ρυθμιστεί κατά τη διάρκεια της διαδικασίας διάδοσης προκειμένου να επιτευχθούν οι στόχοι της διάδοσης της πληροφορίας πλαισίου. Εκμεταλλευόμαστε αυτές τις ενδείξεις προκειμένου να καθοριστεί ένα σύνολο πολιτικών προσαρμογής για τους αλγόριθμους Δ1 και Δ2 που συζητούνται στις επόμενες παραγράφους.

4.2.2 Προσαρμοστικός αλγόριθμος Δ1

Ο Δ1 αποτελεί έναν προσαρμοστικό αλγόριθμο διάδοσης πλαισίου. Η διαφορά μεταξύ του Δ1 και του Δ0 είναι ότι ο ρυθμός μόλυνσης β ρυθμίζεται δυναμικά σε κάθε κόμβο τη χρονική στιγμή t με βάση την Ένδειξη-1 (βλ. 4.2.1). Υποθέτουμε ότι το $V(i)$ υποδηλώνει το σύνολο των γειτονικών κόμβων ενός κόμβου i , που είναι μέσα στην ακτίνα εκπομπής του. Ειδικά, το σύνολο $V(i)$ υπολογίζεται εύκολα, δεδομένου ότι υπάρχουν κόμβοι που στέλνουν μηνύματα στον κόμβο i . Αυτό σημαίνει ότι, μόλις λάβει ο κόμβος i ένα μήνυμα από έναν κόμβο j ο κόμβος j υποτίθεται ότι είναι ένας γειτονικός κόμβος του κόμβου i . Καθορίζουμε την πυκνότητα $d(i)$ του δικτύου για έναν κόμβο i ως τον αριθμό των γειτονικών κόμβων του, δηλαδή $d(i) = |V(i)|$. Ο κόμβος i προσαρμόζει την τιμή του β σύμφωνα με την τιμή $d(i)$. Πειραματικά, η πολιτική προσαρμογής για έναν κόμβο i είναι: όσο χαμηλότερη είναι η τιμή $d(i)$ τόσο υψηλότερος ο ρυθμός μόλυνσης β . Αυτό δείχνει

ότι, προκειμένου να επιτευχθεί υψηλή κάλυψη της πληροφορίας πλαισίου (δηλ., η μερίδα των κόμβων που κατέχουν ένα έγκυρο κομμάτι του πλαισίου να εκφράζουν ένα μεγάλο ποσοστό) ο κόμβος i πρέπει να διαδώσει το πλαίσιο με υψηλότερη πιθανότητα στο $V(i)$. Αυτό συμβαίνει εξαιτίας του γεγονότος ότι το $V(i)$ αποτελείται από έναν μικρό αριθμό κόμβων. Αφ' ετέρου, ο κόμβος i διαδίδει το πλαίσιο με μια χαμηλή τιμή του β όταν οι γειτονικοί κόμβοι του είναι πολλοί. Όπως συμπεραίνεται από τα αποτελέσματα προσομοίωσης ο αλγόριθμος Δ1 επιτρέπει την πλήρη κάλυψη ακόμα και όταν ο αριθμός των μεταδόσεων είναι μόνο μία ανά μήνυμα ανά κόμβο. Πειραματικά, η τιμή του β καθορίζεται από την τιμή του $d(i)$ όπως διευκρινίζεται στον Πίνακα 4-1.

Πίνακας 4-1: Τιμές του ρυθμού μόλυνσης β σε σχέση με τον αριθμό των γειτονικών κόμβων.

| $ V(i) $ | 0...3 | 4, 5 | 6, 7 | 8, 9 | 10, 11 | 12, 13 | 14, 15 | 16, 17 | 18, 19 | ≥ 20 |
|----------|-------|------|------|------|--------|--------|--------|--------|--------|-----------|
| β | 1.0 | 0.9 | 0.8 | 0.7 | 0.6 | 0.5 | 0.4 | 0.3 | 0.2 | 0.1 |

4.2.3 Προσαρμοστικός αλγόριθμος Δ2

Στη ενότητα αυτή θα περιγράψουμε άλλη μία προσαρμοστική προσέγγιση για την επίτευξη της διάδοσης πληροφορίας πλαισίου. Αρχικά παρουσιάζουμε τις παραμέτρους και τους κανόνες για τον προσαρμοστικό αλγόριθμο Δ2. Υποθέτουμε ότι το $m(t)$ είναι ένα μήνυμα που λαμβάνεται από έναν κόμβο i στο χρόνο t . Η παράμετρος $a(m(t)) < a(m(0)) = a$ δείχνει την υπόλοιπη χρονική ισχύ του μηνύματος m που λήφθηκε τη χρονική στιγμή t . Όταν ο κόμβος i λάβει το μήνυμα m πρέπει να το αποστείλει μέσα στο χρονικό διάστημα $[t, t + a(m(t))]$. Το μήνυμα m μπορεί να μεταδοθεί $f(m)$ φορές μέσα σε αυτό το χρονικό πλαίσιο, επομένως, η περίοδος μετάδοσης $q(m)$ του μηνύματος είναι $a(m(t))/f(m)$. Ο κόμβος i , στην αρχή κάθε μετάδοσης, αποφασίζει στοχαστικά εάν θα μεταδώσει ή όχι το μήνυμα με πιθανότητα $\beta(t)$. Η πιθανότητα αυτή ισοδυναμεί με το ρυθμό μόλυνσης. Η τιμή του ρυθμού μόλυνσης β αρχίζει από μια αρχική τιμή β_0 και ρυθμίζεται σύμφωνα με την πυκνότητα $d(i)$ του δικτύου. Μια ελάχιστη τιμή β_{min} υιοθετείται επίσης για να αποφευχθεί το φαινόμενο που όλοι οι κόμβοι δεν μεταδίδουν τίποτα. Επιπλέον, η τιμή του β καθορίζεται από την κινητικότητα του κόμβου i . Για να ποσοτικοποιήσουμε την κινητικότητα, καθορίζουμε έναν παράγοντα $M(i)$ στο χρόνο t ως το ρυθμό αλλαγής στην τιμή του $|V(i)|$, δηλαδή ο αριθμός των κόμβων που είναι γείτονες του κόμβου i . Μια αλλαγή στο $V(i)$ ανιχνεύεται όταν ο κόμβος i λαμβάνει ή "ακούει" ένα μήνυμα από έναν κόμβο j . Ο παράγοντας $M(i)$ εξαρτάται από τον αριθμό των αλλαγών στο $V(i)$ και έχει τιμή:

$$M(i) = \Delta(d(i))/\Delta t - d^*(i)/T.$$

Η παράμετρος $d^*(i)$ είναι η μέση τιμή των αλλαγών που παρατηρήθηκαν στο $V(i)$ για χρόνο T . Όσο μεγαλύτερη είναι η τιμή του $M(i)$ τόσο μεγαλύτερη είναι η κινητικότητα στη γειτονία του κόμβου i . Έτσι, βάσει της πυκνότητας του δικτύου $d(i)$ και του παράγοντα της κινητικότητας $M(i)$ ο κόμβος i προσαρμόζει την τιμή του ρυθμού μόλυνσης β σύμφωνα με τον παρακάτω κανόνα:

$$\beta(t + 1) = \beta(t) + M^* \cdot M(i, t) + D^* \cdot d(i, t)$$

με $\beta(0) = \beta_0$. Οι παράμετροι M^* και D^* ελέγχουν τον αντίκτυπο της κινητικότητας $M(i, t)$ και της πυκνότητας $d(i, t)$ στην προσαρμογή του β τη χρονική στιγμή t . Τυπικά θεωρούμε ότι $(M^*, D^*) = (0.1, 0.1)$. Ο αλγόριθμος Δ2 εκμεταλλεύεται την κινητικότητα και την πυκνότητα προκειμένου να ρυθμίσει το β . Αφ' ετέρου, ο Δ1 στηρίζεται στον αριθμό των γειτόνων για να ρυθμίσει κατάλληλα το β .

4.2.4 Πειραματική αξιολόγηση

Πραγματοποιήσαμε εκτενείς συνόδους προσομοίωσης για τους σκοπούς της αξιολόγησης των αλγορίθμων Δ0, Δ1 και Δ2. Στη συνέχεια παρέχουμε τα αποτελέσματα προσομοίωσης για κάθε έναν από τους τρεις αλγορίθμους και συζητάμε τις εκβάσεις των πειραμάτων. Όλες οι προσομοιώσεις έχουν εκτελεσθεί χρησιμοποιώντας τον προσομοιωτή Omnet++[94]. Η βασική οργάνωση περιλαμβάνει $N = 25$ κόμβους που τοποθετούνται σε ένα τετραγωνικό πλέγμα πλευράς w . Ενώ ο αριθμός και οι αρχικές σχετικές θέσεις των κόμβων παραμένουν σταθερά, το μέγεθος w της πλευράς του πλέγματος ποικίλλει ούτως ώστε να αλλάξει η πυκνότητα του δικτύου. Για τα χαμηλότερα επίπεδα του δικτύου, χρησιμοποιήσαμε το πρότυπο IEEE 802.15.4 και το πρωτόκολλο Carrier Sense Multiple Access (CSMA) για το επίπεδο MAC (Media Access Control).

4.2.4.1 Αποτελέσματα για τον αλγόριθμο Δ0

Παρουσιάζουμε τα αποτελέσματα προσομοίωσης του μη-προσαρμοστικού αλγορίθμου Δ0. Αυτές οι αρχικές προσομοιώσεις αξιοποιούνται για το σχεδιασμό των Δ1 και Δ2. Στην περίπτωση αυτή το πλήθος των μεταδόσεων $f(m)$ ενός μηνύματος m και η πιθανότητα μετάδοσης β ενός τέτοιου μηνύματος για κάθε κόμβο παραμένουν σταθερά

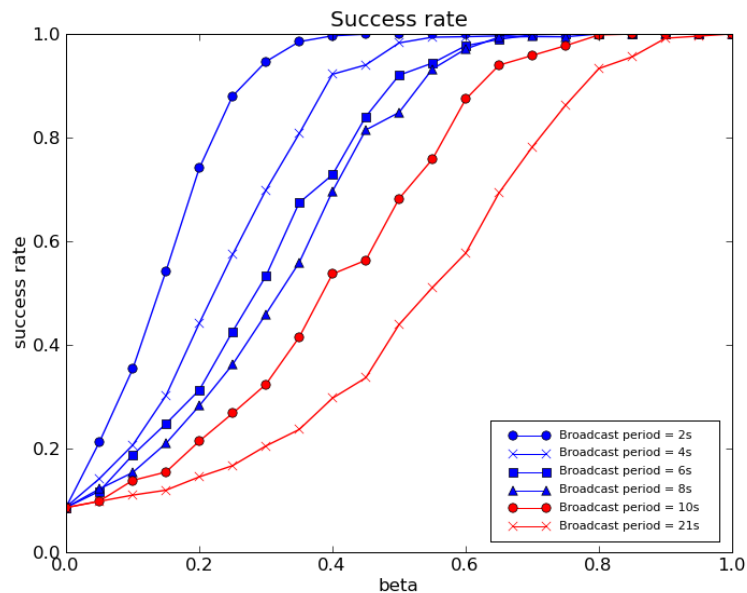
κατά τη διάρκεια της προσομοίωσης. Ο Πίνακας 4-2 συνοψίζει τις παραμέτρους της προσομοίωσης.

Πίνακας 4-2: Παράμετροι προσομοίωσης για τον αλγόριθμο Δ0.

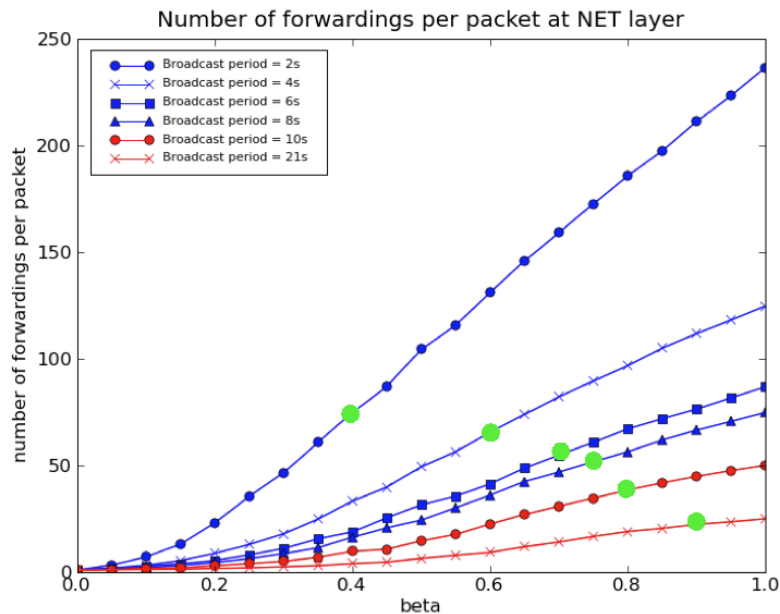
| Όνομα παραμέτρου και περιγραφή | Τιμή παραμέτρου |
|---|------------------------------------|
| N : Αριθμός κόμβων | 25 |
| P_0 : Αριθμός των πακέτων που αποστέλλονται από τον πάνω αριστερά κόμβο σε μια τοπολογία τετραγωνικού πλέγματος (οι υπόλοιποι κόμβοι δεν παράγουν μηνύματα) | 50 |
| Παραγωγή μηνυμάτων | Εκθετική (10s) |
| a (TTL) [s] | 20 |
| γ (καθυστέρηση)[s] | Ομοιόμορφα κατανεμημένη στο [0, 1] |
| Αριθμός εκτελέσεων ανά β | 5 |
| w : μέγεθος [m] | 50, 100, 200, 300, 400 |
| β : ρυθμός μόλυνσης (πιθανότητα εκπομπής) | Από 0 έως 1 με βήματα 0.05 |
| $q(m)$: περίοδος μετάδοσης [s] | 2, 4, 6, 8, 10, 21 |

Υποθέτουμε ότι μόνο ο πάνω αριστερός κόμβος του πλέγματος στέλνει τα πακέτα. Οι μετρικές που υιοθετούνται για την αξιολόγηση του αλγορίθμου είναι:

- *Ποσοστό επιτυχίας(success rate) ή κάλυψη (coverage)*. Η μετρική αυτή ισούται με τον ολικό αριθμό των ληφθέντων πακέτων προς τον ολικό αριθμό των αναμενόμενων πακέτων. Ο συνολικός αριθμός των αναμενόμενων πακέτων ορίζεται ως $(N-1) \cdot P_0$.
- *Αριθμός προωθήσεων*. Η μετρική αυτή δηλώνει το μέσο αριθμό των μεταδόσεων στο δίκτυο ανά μήνυμα.



(α)



(β)

Σχήμα 4.3:(α) Η κάλυψη σε σχέση με τον ρυθμό β για διαφορετικές τιμές των περιόδων μετάδοσης. (β) Αριθμός μεταδόσεων πακέτων ανά μήνυμα στο δίκτυο. Το μέγεθος της πλευράς του πεδίου είναι $w=300m$.

Στο Σχήμα 4.3 (α) παρουσιάζεται η κάλυψη για τις διαφορετικές περιόδους μετάδοσης σε ένα πλέγμα πλευράς $w=300m$. Παρατηρούμε τα φαινόμενα διείσδυσης (percolation) δηλαδή υπάρχει πλήρης κάλυψη του δικτύου μετά από μια ορισμένη τιμή του β . Μόλις επιτευχθεί αυτό το κατώτατο όριο, δεν υπάρχει νόημα για περαιτέρω αύξηση του β . Αυτή η τιμή διείσδυσης (δηλ., η ελάχιστη πιθανότητα μετάδοσης β_{min}) μειώνεται με την πυκνότητα του δικτύου και αυξάνεται με την περίοδο μετάδοσης (δηλ., μειώνεται με τον

αριθμό των μεταδόσεων), όπως αναμενόταν. Μόλις καταστούν γνωστές οι συνθήκες διείσδυσης, μπορούμε να εξετάσουμε τον παράγοντα κόστος, ο οποίος είναι ο αριθμός των προωθήσεων για κάθε μήνυμα. Υπολογίζουμε το κόστος από τον αριθμό των μεταδόσεων που απαιτούνται μέσα σε ολόκληρο το δίκτυο προκειμένου να ληφθεί μια δεδομένη κάλυψη.

Στο Σχήμα 4.3 (β) έχουμε προσθέσει τα διευρυμένα (πράσινα) σημεία για να δείξουμε τις πιθανότητες διείσδυσης. Τα διευρυμένα σημεία δείχνουν σε ποια τιμή του β λαμβάνουμε την πλήρη κάλυψη του δικτύου. Παρατηρούμε ότι: *για να φθάσουμε στην πλήρη κάλυψη, είναι φτηνότερο, από άποψη αριθμού μεταδόσεων, να σταλεί ένα μήνυμα μία φορά με υψηλή πιθανότητα από να το σταλεί αρκετές φορές με χαμηλότερη πιθανότητα.*

Επιπλέον, προσομοιώθηκε ένα σενάριο συμφόρησης, όπου όλοι οι κόμβοι στέλνουν 50 πακέτα αλλάζοντας ταυτόχρονα και την εκθετική παράμετρο παραγωγής μηνυμάτων από έναν μέσο όρο 10 sec σε 1 sec (πιο πολλά μηνύματα). Εισάγουμε επίσης μία πρόσθετη μετρική που είναι ο αριθμός των πακέτων που απορρίπτονται στο επίπεδο MAC λόγω της συμφόρησης. Από αυτό το σενάριο, είμαστε σε θέση να συμπεράνουμε την Ενδειξη-3, που παρουσιάσαμε στην Παράγραφο 4.2.1.

4.2.4.2 Αποτελέσματα για τον αλγόριθμο Δ1

Για την αξιολόγηση του αλγορίθμου Δ1, υιοθετήσαμε τις ίδιες παραμέτρους προσομοίωσης όπως για τον Δ0 (βλ. Πίνακα 4-2). Η περίοδος μετάδοσης $q(m)$ τίθεται ως $TTL + 1$ ώστε να εξασφαλιστεί ότι κάθε κόμβος εκτελεί μόνο μια μετάδοση ανά μήνυμα. Εξετάζουμε τη συμπεριφορά του β που προέρχεται από τους κόμβους στους οποίους μόνο ο πάνω αριστερός κόμβος στέλνει 50 πακέτα. Επιπλέον, το μέσο χρονικό διάστημα παραγωγής μεταξύ δύο πακέτων είναι 10secs, η περίοδος μετάδοσης είναι 20 + 1 secs και η αρχική τιμή του β είναι 0.9. Όπως μπορούμε να συμπεράνουμε από τον Πίνακα 4-3 ο προσαρμοστικός αλγόριθμος επιτρέπει πλήρη ή σχεδόν πλήρη κάλυψη, ακόμα και όταν έχουμε μόνο μία μετάδοση.

Πίνακας 4-3: Κάλυψη και αριθμός προωθήσεων για τον αλγόριθμο Δ1.

| Μέγεθος w [m] | 50 | 100 | 200 | 300 | 400 |
|--|-------|-------|-------|-------|-------|
| Κάλυψη | 1.000 | 0.999 | 0.996 | 0.980 | 0.055 |
| Αριθμός μεταδόσεων ανά μήνυμα ² | 14.9 | 15.7 | 20.7 | 22.0 | 2.2 |

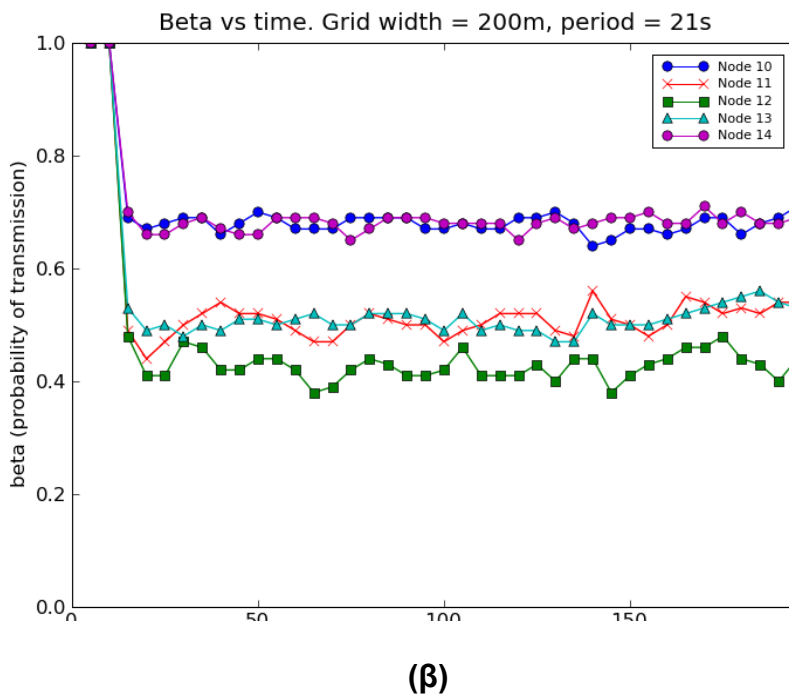
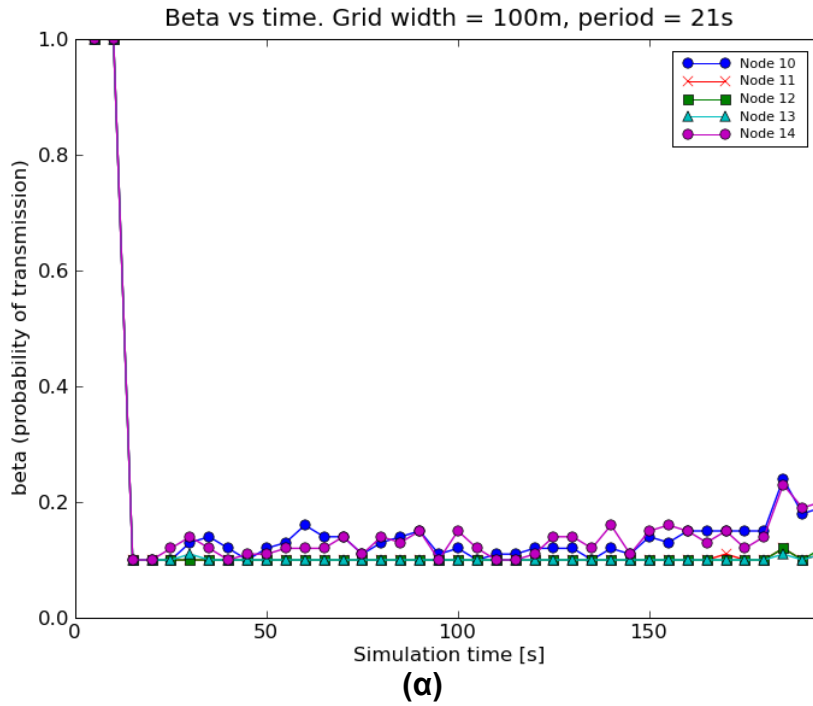
² Ο μέσος αριθμός των φορών που ένα μήνυμα προωθείται, συμπεριλαμβανομένης της πρώτης εκπομπής όταν δημιουργείται. Αν $\beta = 1.0$ και η περίοδος μετάδοσης $> TTL$ (μια μετάδοση σε κάθε κόμβο), τότε ισχύει: μέσος αριθμός αποστολής ανά μήνυμα = αριθμός κόμβων.

Μόλις η πυκνότητα του δικτύου είναι υψηλή, παρατηρούμε ότι ο μέσος αριθμός μεταδόσεων ανά μήνυμα είναι υψηλός. Στη συνέχεια επανεκτελούμε τον αλγόριθμο και θέτουμε και τους 25 κόμβους να στείλουν 50 πακέτα. Ο Πίνακας 4-4 παρουσιάζει τα αποτελέσματα ενώ το απεικονίζει την τιμή του ρυθμού μόλυνσης β σε σχέση με το χρόνο για διαφορετικές πυκνότητες δικτύου.

Πίνακας 4-4: Κάλυψη και αριθμός προωθήσεων για τον αλγόριθμο Δ1 όταν όλοι οι κόμβοι στέλνουν 50 πακέτα.

| Μέγεθος w [m] | 50 | 100 | 200 | 300 | 400 |
|-------------------------------|-------|-------|-------|-------|-------|
| Κάλυψη | 0.996 | 0.962 | 0.952 | 0.964 | 0.111 |
| Αριθμός μεταδόσεων ανά μήνυμα | 3.9 | 5.0 | 15.9 | 20.0 | 3.6 |

Όπως αναμένεται, ο αλγόριθμος Δ1 είναι αποδοτικότερος στην παρουσία αρκετής κυκλοφορίας στο δίκτυο. Παρατηρήσαμε επίσης ένα φαινόμενο όπου *οι κόμβοι που είναι κοντά στα σύνορα του πεδίου προσομοίωσης(πλέγματος) ανιχνεύουν λιγότερους γείτονες και συμπεραίνουν χαμηλότερη πυκνότητα*. Εκτελέσαμε επίσης ένα σενάριο κινητικότητας με τον Δ1 αλγόριθμο. Παρατηρήσαμε ότι η αύξηση του αριθμού των προσπαθειών μετάδοσης (μείωση της περιόδου μετάδοσης) επιτυγχάνει καλύτερη κάλυψη στο δίκτυο. Η χρήση της κινητικότητας για να ελέγξει τον αριθμό προσπαθειών μετάδοσης οδηγεί στην πλήρη ή τη σχεδόν πλήρη κάλυψη διατηρώντας τον αλγόριθμο απλό.



Σχήμα 4.4: Η τιμή β σε σχέση με το χρόνο για μέγεθος (a) $w = 100m$ και (b) $w = 200m$. Και οι 25 κόμβοι παράγουν και διαβιβάζουν 50 μηνύματα. Η αρχική τιμή του β είναι 1.0.

4.2.4.3 Αποτελέσματα για τον αλγόριθμο Δ2

Για την αξιολόγηση του αλγορίθμου Δ2, η προσομοίωση περιλαμβάνει $N = 25$ κινούμενους κόμβους που αρχικά τοποθετήθηκαν τυχαία στο πεδίο. Οι κόμβοι κινούνται τυχαία σύμφωνα με το ακόλουθο σενάριο:

- Στην αρχή, οι κόμβοι τοποθετούνται τυχαία στο πεδίο προσομοίωσης.
- Κάθε κόμβος αλλάζει την ταχύτητα και την κατεύθυνσή του πολλές φορές. Το χρονικό διάστημα μεταξύ των αλλαγών ακολουθεί την κανονική κατανομή ($\sim N(10, 0.5)$ secs.)
- Σε κάθε αλλαγή, κάθε κόμβος γυρίζει κατά μια ορισμένη γωνία ($\sim N(0, 30)$, σε μοίρες)
- Σε κάθε αλλαγή, κάθε κόμβος επιλέγει τη νέα ταχύτητά του ($\sim N(0.1, 0.1)$, σε m/s.)
- Όταν ένας κόμβος φθάνει σε μια άκρη του πεδίου παιχνιδιού, «αναπηδά» σε αυτό.

Ο Πίνακας 4-5 περιγράφει τις παραμέτρους προσομοίωσης για τον αλγόριθμο Δ2.

Πίνακας 4-5: Παράμετροι προσομοίωσης για τον αλγόριθμο Δ2.

| Όνομα παραμέτρου και περιγραφή | Τιμή παραμέτρου |
|---|------------------------------------|
| N : Αριθμός κόμβων | 25 |
| P_0 : Αριθμός των πακέτων που αποστέλλονται από τον πάνω αριστερά κόμβο σε μια τοπολογία τετραγωνικού πλέγματος (οι υπόλοιποι κόμβοι δεν παράγουν μηνύματα) | 50 |
| Παραγωγή μηνυμάτων | Εκθετική (10s) |
| a (TTL) [s] | 20 |
| γ (καθυστέρηση)[s] | Ομοιόμορφα κατανεμημένη στο [0, 1] |
| Μέση ταχύτητα κόμβου (m/s) | 5,10,15, 20 |
| w : μέγεθος [m] | 200, 300, 400 |
| β_0 | 0.5, 0.7, 0.9 |
| $q(m)$: περίοδος μετάδοσης [s] | 10, 21 |

Μόνο ένας κόμβος στέλνει τα πακέτα και συγκεκριμένα ο πάνω αριστερός σε μια τοπολογία τετραγωνικού πλέγματος. Οι παράμετροι των προσομοιώσεων σχετικά με το μήνυμα m καθώς επίσης και την τιμή των μετρικών που υιοθετείται (*κάλυψη* και *αριθμός προωθήσεων ανά μήνυμα*) είναι η ίδιες με εκείνες της Παραγράφου 4.2.4.1. Παρατηρούμε ότι για χαμηλή πυκνότητα, ο μέσος αριθμός μεταδόσεων ανά μήνυμα καθώς επίσης και η κάλυψη μειώνεται. Ο συνοψίζει τα συμπεράσματά μας. Τα αποτελέσματα παρουσιάζονται για περιόδους μετάδοσης $q(m) = 10s$ και $21s$ (μια ή δύο μεταδόσεις το πολύ) και $\beta_0=0.5$. Μπορούμε να παρατηρήσουμε την αύξηση στην κάλυψη καθώς η μέση ταχύτητα των κόμβων αυξάνεται. Αυτή η παρατήρηση γίνεται εμφανέστερη όταν το δίκτυο είναι αραιό στην περίπτωση δηλαδή που $w = 400m$. Αυτό σημαίνει ότι η κινητικότητα βοηθά τη διαδικασία διάδοσης πληροφορίας πλαισίου. Οι

προσομοιώσεις δείχνουν ότι οι τιμές του β εξαρτώνται σε μεγάλο βαθμό από την τιμή β_0 .

Πίνακας 4-6: Κάλυψη και αριθμός μεταδόσεων σε σχέση με την μέση ταχύτητα και το μέγεθος w . Η τιμή β_0 ισούται με 0.5. Σε κάθε κελί, η πρώτη τιμή αντιστοιχεί για περίοδο μετάδοσης 10s και η δεύτερη για περίοδο μετάδοσης 21s.

| Μέση ταχύτητα | Μέγεθος w [m] | 200 | 300 | 400 |
|---------------|-------------------------------|-----------------------|-----------------------|-----------------------|
| 5 m/s | Κάλυψη | 0.993583, 0.901000 | 0.771667, 0.554000 | 0.417667, 0.294167 |
| | Αριθμός μεταδόσεων ανά μήνυμα | 23.312, 11.326 | 15.932, 7.422 | 8.572, 4.408 |
| 10 m/s | Κάλυψη | 0.991500, 0.897250 | 0.823500, 0.561167 | 0.487250, 0.287500 |
| | Αριθμός μεταδόσεων ανά μήνυμα | 22.566, 11.264 | 16.81, 7.51 | 9.708, 4.306 |
| 15 m/s | Κάλυψη | 0.996583, 0.921167 | 0.891333, 0.595583 | 0.505583, 0.292250 |
| | Αριθμός μεταδόσεων ανά μήνυμα | 22.86, 11.64 | 18.276, 8.034 | 9.504, 4.566 |
| 20 m/s | Κάλυψη | 0.993000, 0.937417 | 0.893083, 0.599833 | 0.550333, 0.305917 |
| | Αριθμός μεταδόσεων ανά μήνυμα | 22.498, 11.72 | 18.146, 8.324 | 18.146, 8.324 |

4.3 Προσαρμογή ρυθμού μόλυνσης βάσει των χαρακτηριστικών της πληροφορίας πλαισίου

Όταν υιοθετείται ένα επιδημικό μοντέλο διάδοσης της πληροφορίας πλαισίου μία υψηλή τιμή του ρυθμού μόλυνσης β οδηγεί σε αυξανόμενο αριθμό μηνυμάτων που ανταλλάσσονται μεταξύ των κόμβων του δικτύου. Ως εκ τούτου, η κατανάλωση ενέργειας για τις αντίστοιχες μεταδόσεις των μηνυμάτων αυξάνεται. Αφ' ετέρου, μια χαμηλή τιμή του ρυθμού μόλυνσης β οδηγεί στην κατάσταση που η πληροφορία πλαισίου στο δίκτυο είναι απαρχαιωμένη. Με βάση αυτό το γεγονός σε αυτήν την ενότητα προτείνονται ορισμένα προσαρμοστικά επιδημικά μοντέλα έτσι ώστε να ρυθμίζεται ο ρυθμός μόλυνσης β σύμφωνα με τη μεταβλητότητα της πληροφορίας πλαισίου αλλά και την ηλικία αυτής της πληροφορίας (δηλ., το χρόνο που μεσολάβησε από τη στιγμή της δημιουργίας της πληροφορίας). Αξιολογούμε τα προτεινόμενα επιδημικά σχήματα όσον αφορά:

- Το σφάλμα της πληροφορίας πλαισίου στον πιο απομακρυσμένο από την πηγή κόμβο
- Τον αριθμό των μηνυμάτων που έστειλε η πηγή στο δίκτυο, και,
- Την απόδοση σχετικά με την χρονική ποιότητα (ηλικία) της πληροφορίας πλαισίου

Υποθέτουμε λοιπόν ένα ασύρματο δίκτυο που αποτελείται από N συνεργατικούς κόμβους. Κάθε κόμβος i έχει εμβέλεια μετάδοσης $R(i)=r$, επομένως έχει $N(i)$ από τους N συνολικούς κόμβους ως άμεσους γείτονες. Ο κόμβος i είναι σε θέση να μολύνει τους γειτονικούς του κόμβους που βρίσκονται εντός της εμβέλειας μετάδοσης r . Ένας συγκεκριμένος κόμβος που καλείται ως πηγή (*source*) έχει προσαρμοσμένο αισθητήρα και παράγει σε κάθε χρονική στιγμή t την πληροφορία πλαισίου $V(t)$ (π.χ. $V(t) =$ (π.χ. θερμοκρασία, ταχύτητα ανέμου)). Αυτές οι τιμές, $V(t) = V(0), V(1), \dots$, διαδίδονται στο δίκτυο χρησιμοποιώντας το επιδημικό μοντέλο SIS (Susceptible-Infected-Susceptible). Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι ο ρυθμός ίασης δ για το μοντέλο είναι $\delta = 0$. Αυτό σημαίνει ότι όταν ένας κόμβος μολύνεται θα παραμείνει μολυσμένος για πάντα. Αυτή η υπόθεση υποδηλώνει ότι ένας κόμβος, που έχει λάβει τη συγκεκριμένη πληροφορία πλαισίου $V(t)$, την χρησιμοποιεί (εκμεταλλεύεται) μέχρι να λάβει ένα μήνυμα που κομίζει νέα πληροφορία $V(t')$ με $t' > t$. Στις επόμενες παραγράφους του κεφαλαίου μελετάμε τη διαδικασία διάδοσης της πληροφορίας πλαισίου όταν οι κόμβοι αντικαθιστούν το υπάρχον πλαίσιο με το ένα πιο ενημερωμένο πλαίσιο, δηλαδή όταν μολύνεται ένας κόμβος με το πλαίσιο $V(t_1)$, αυτό το πλαίσιο παραμένει έγκυρο έως ότου ο κόμβος μολυνθεί με το πλαίσιο $V(t_2)$, με $t_2 > t_1$.

4.3.1 Προσαρμοστικό Επιδημικό Σχήμα Διάδοσης Πληροφορίας L1

4.3.1.1 Περιγραφή Σχήματος

Στην αρχή της διαδικασίας ($t = 0$) όλοι οι κόμβοι είναι στην κατάσταση «εύρωτος» / «επιρρεπής» (susceptible - S) εκτός από τον κόμβο-πηγή που είναι μολυσμένος (I). Προφανώς, η πηγή παραμένει μολυσμένη (και μολυσματική) για πάντα, καθώς είναι ο δημιουργός της πληροφορίας πλαισίου. Ορίζουμε ως $\beta_s \in [0,1]$ τον ρυθμό μόλυνσης των γειτόνων της πηγής και ως $\beta_r \in [0,1]$ τον ρυθμό μόλυνσης εκείνων των κόμβων του δικτύου που δεν είναι στην εμβέλεια μετάδοσης της πηγής (ενδιάμεσοι κόμβοι). Ο ρυθμός β_s υποδηλώνει την πιθανότητα με την οποία η πηγή θα μεταδώσει την πληροφορία πλαισίου που έχει δημιουργήσει σε κάθε έναν από τους γείτονές της. Η τιμή του β_s είναι προσαρμόσιμη σε ένα υπο-διάστημα του $[0,1]$ και εξαρτάται από τη

μεταβλητότητα της πληροφορίας πλαισίου. Όσο υψηλότερη είναι η τιμή του ρυθμού μόλυνσης β_s , τόσο μεγαλύτερος είναι ο αριθμός των κόμβων του δικτύου που θα έχει ενημερωμένες τιμές οι οποίες βρίσκονται κοντά στην πραγματική τιμή που παράχθηκε στην πηγή, ακόμα κι αν ο ρυθμός β_r παραμένει σταθερός. Αυτό αποδίδεται στο μεγάλο αριθμό κόμβων που η πηγή μολύνει καθώς ο ρυθμός β_s είναι υψηλός. Στη συνέχεια οι υπόλοιποι κόμβοι αναμεταδίδουν την πληροφορία με σταθερό ρυθμό β_r . Εντούτοις, σε αυτήν την περίπτωση, ο αριθμός των μεταδιδόμενων μηνυμάτων στο δίκτυο αυξάνεται. Αφ' ετέρου, εάν ο ρυθμός β_s είναι χαμηλός, η δικτυακή κυκλοφορία μειώνεται αλλά ο αριθμός κόμβων που έχουν ενημερωμένη πληροφορία πλαισίου επίσης μειώνεται.

Υπάρχουν πολλές τεχνικές στη βιβλιογραφία που εξετάζουν την ανίχνευση και την ποσοτικοποίηση της αλλαγής σε ένα ρεύμα δεδομένων (change detection in data streams). Οι συγγραφείς στην [44] υιοθετούν την τεχνική CUSUM (Cumulative Sum) της σελίδας για την ανίχνευση αλλαγής ενώ στην [96] το ίδιο πρόβλημα αντιμετωπίζεται χρησιμοποιώντας δύο κινούμενα παράθυρα W_1 και W_2 στο ίδιο ρεύμα δεδομένων που παράγονται από τις κατανομές P_1 και P_2 . Ως εκ τούτου, το πρόβλημα ανίχνευσης της αλλαγής καταλήγει στο να συμπεράνουμε από τα W_1 και W_2 εάν $P_1 = P_2$ (δηλ. εάν μια αλλαγή έχει εμφανιστεί). Διατηρώντας ένα σύνολο από διαθέσιμες μεθόδους ανίχνευσης αλλαγής ο κόμβος-πηγή μπορεί να επιλέξει μία από αυτές ανάλογα με τις υπολογιστικές ικανότητές του (μνήμη, CPU, κτλ). Έτσι σε κάθε χρονική στιγμή t ο ρυθμός μόλυνσης θα είναι $\beta_s(t) = f(c(t))$, όπου c είναι η ανιχνευθείσα αλλαγή στο ρεύμα δεδομένων.

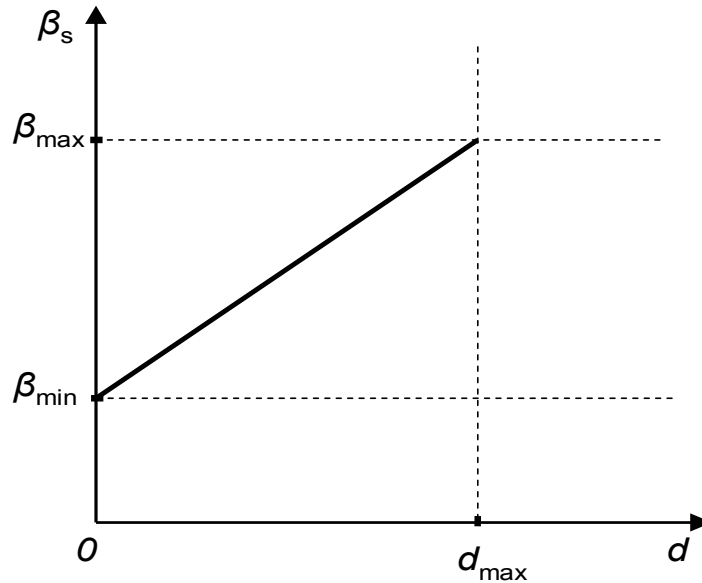
Μια απλούστερη μέθοδος για την ανίχνευση της αλλαγής, που χρησιμοποιείται επίσης στα πειράματά μας, είναι ο καθορισμός της διαφοράς d μεταξύ δύο διαδοχικών τιμών του ρεύματος δεδομένων, δηλαδή

$$c(t) = d(t) = ||V(t) - V(t-1)|| \quad (4.1)$$

Ο ρυθμός μόλυνσης β_s είναι εξαρτώμενος από το χρόνο, δηλαδή $\beta_s = \beta_s(t)$, και έχει γραμμική σχέση με την ανιχνευθείσα διαφορά d . Ως εκ τούτου, ο ρυθμός β_s τη χρονική στιγμή t προσαρμόζεται με βάση την παρακάτω σχέση (4.2):

$$\beta_s(t) = \begin{cases} \frac{\beta_{\max} - \beta_{\min}}{d_{\max}(t)} d(t) + \beta_{\min} & d_{\max}(t) \neq 0 \\ \beta_{\min} & d_{\max}(t) = 0 \end{cases} \quad (4.2)$$

όπου $\beta_{\max}, \beta_{\min} \in [0,1]$ είναι το ανώτερο και κατώτερο όριο του ρυθμού β_s , αντίστοιχα. Στο Σχήμα 4.5, παρουσιάζουμε τη συμπεριφορά του β_s ως συνάρτηση της διαφοράς d .



Σχήμα 4.5: Η συμπεριφορά του ρυθμού μόλυνσης β_s ως συνάρτηση της διαφοράς d .

Ο παρονομαστής d_{\max} στη σχέση (4.2) μπορεί να είναι ίσος με $\|V_{\max} - V_{\min}\|$ όπου V_{\max} είναι το διάνυσμα με τη μέγιστη τιμή από το ρεύμα δεδομένων και V_{\min} το διάνυσμα με την κατώτερη τιμή. Όμως, η πηγή δεν γνωρίζει εκ των προτέρων τα ανώτερα και χαμηλότερα όρια του ρεύματος δεδομένων V . Ως εκ τούτου, ο παράγοντας d_{\max} είναι επίσης προσαρμόσιμος ανάλογα με την ιστορία των τιμών που έχουν παρουσιαστεί μέχρι το παρόν και υπολογίζεται από τις k πιο πρόσφατες διαφορές όπως φαίνεται στη σχέση (4.3).

$$d_{\max}(t) = \begin{cases} \max[d(1), \dots, d(t)] & 1 \leq t \leq k \\ \max[d(t-k+1), \dots, d(t)] & t > k \end{cases} \quad (4.3)$$

Επιπλέον, με αυτήν την τεχνική υπολογισμού του d_{\max} ο ρυθμός β_s δεν επηρεάζεται από τις τυχόν υψηλές τιμές της διαφοράς d στο παρελθόν. Αυτή η ανεπιθύμητη κατάσταση μπορεί να εμφανιστεί παραδείγματος χάριν όταν οι αρχικές τιμές στο ρεύμα δεδομένων έχουν υψηλή μεταβλητότητα και κατόπιν σταθεροποιούνται χωρίς ιδιαίτερη διακύμανση. Ο προσαρμοστικός επιδημικός αλγόριθμος A1 που εφαρμόζεται μόνο στον κόμβο-πηγή, παρουσιάζεται στο Σχήμα 4.6. Η μέθοδος **infect** ($j, Vi(t), \beta_s$) στον αλγόριθμο δείχνει ότι ο γειτονικός κόμβος j (που είναι γείτονας του κόμβου πηγής) μολύνεται με την τιμή $Vi(t)$ με πιθανότητα β_s .

```

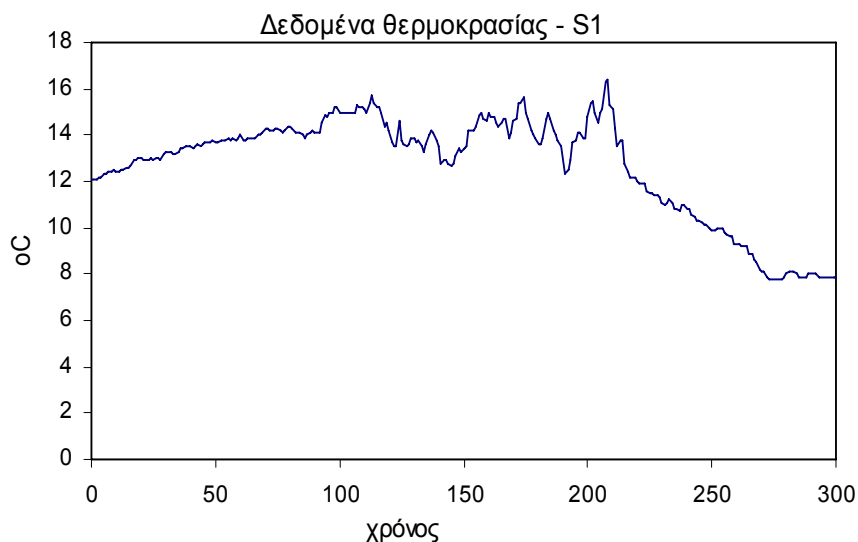
Αλγόριθμος A1: Προσαρμοστική Επιδημική Διάδοση τη χρονική στιγμή
t για τον κόμβο-πηγή i
input:  $\beta_{max}$ ,  $\beta_{min}$ ,  $k$ , Neighbourhood  $N(i)$ , data stream  $V$ 
output:  $\beta_s$ 
Begin:
  While(true)
    /*calculation of difference d */
     $d(t) = ||(V_i(t) - V_i(t-1))||$ 
    /*calculation of  $d_{max}$  */
     $d_{max} = \max (d(u)) \quad u = 1, \dots, k$ 
    /*calculation of  $\beta_s$  */
    if ( $d_{max} == 0$ )
       $\beta_s = \beta_{min}$ 
    else
       $\beta_s = (\beta_{max} - \beta_{min}) / d_{max} * d(t) + \beta_{min}$ 
      /*infect neighbourhood nodes */
      for each node  $j \in N(i)$  do
        infect ( $j, V_i(t), \beta_s$ )
      end for
    End While
End.

```

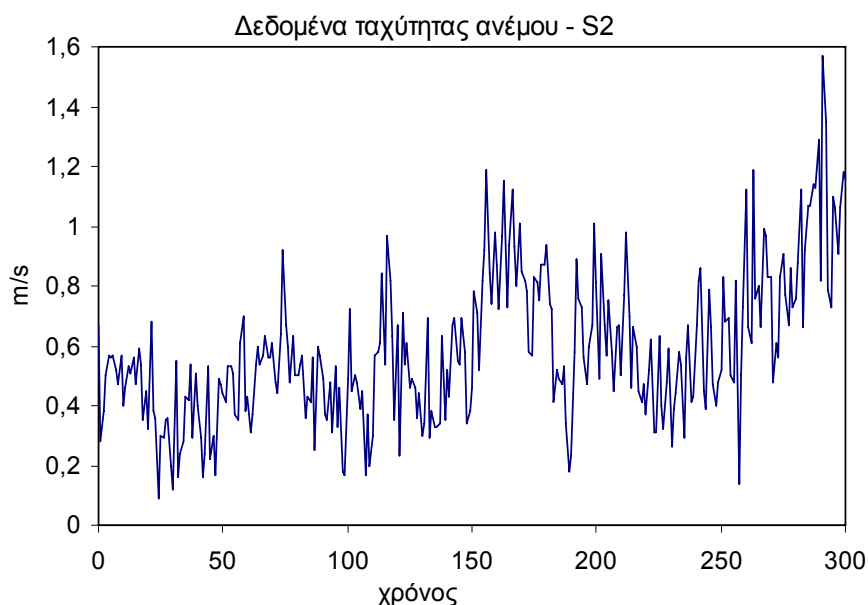
Σχήμα 4.6: Κόμβος πηγή: προσαρμοστικός επιδημικός αλγόριθμος για τη διάδοση πληροφορίας πλαισίου.

4.3.1.2 Προσομοίωση και Απόδοση του Σχήματος L1

Χρησιμοποιούμε μια περιοχή δύο διαστάσεων (2D) για τη μελέτη της συμπεριφοράς του προσαρμοστικού επιδημικού σχήματος L1 που περιγράφηκε στην προηγούμενη ενότητα. Ένας κόμβος i σε αυτήν την περιοχή αντιπροσωπεύεται από τις συντεταγμένες του (x, y) , όπου x, y είναι ακέραιοι αριθμοί και $1 \leq x, y \leq L$ με $L = 100$. Υποθέτουμε ένα δίκτυο με $N = 100$ κόμβους ομοιόμορφα κατανεμημένους στην περιοχή. Θεωρούμε τον κόμβο πηγή να βρίσκεται στην θέση $(0,0)$ και τον κόμβο προορισμό να βρίσκεται στην θέση $(100,100)$. Ο αλγόριθμος που χρησιμοποιείται παρουσιάζεται στο Σχήμα 4.6 και αξιολογήθηκε χρησιμοποιώντας δύο ρεύματα πραγματικών δεδομένων $S1$ και $S2$ (θερμοκρασία και ταχύτητα ανέμου) που προέκυψαν από τις μετρήσεις των αντίστοιχων αισθητήρων εγκατεστημένων στη Σταμάτα Αττικής. Οι $T = 300$ τιμές για κάθε ρεύμα δεδομένων που διακρίνονται στο Σχήμα 4.7 καταγράφηκαν κατά τη διάρκεια ενός χρονικού διαστήματος περίπου 2.5 ωρών με συχνότητα 0.033 Hz (2 δείγματα ανά λεπτό). Ο προσαρμοστικός επιδημικός αλγόριθμος A1 προσομοιώνεται για ένα διάστημα T χρονικών στιγμών μέσω 1000 εκτελέσεων.



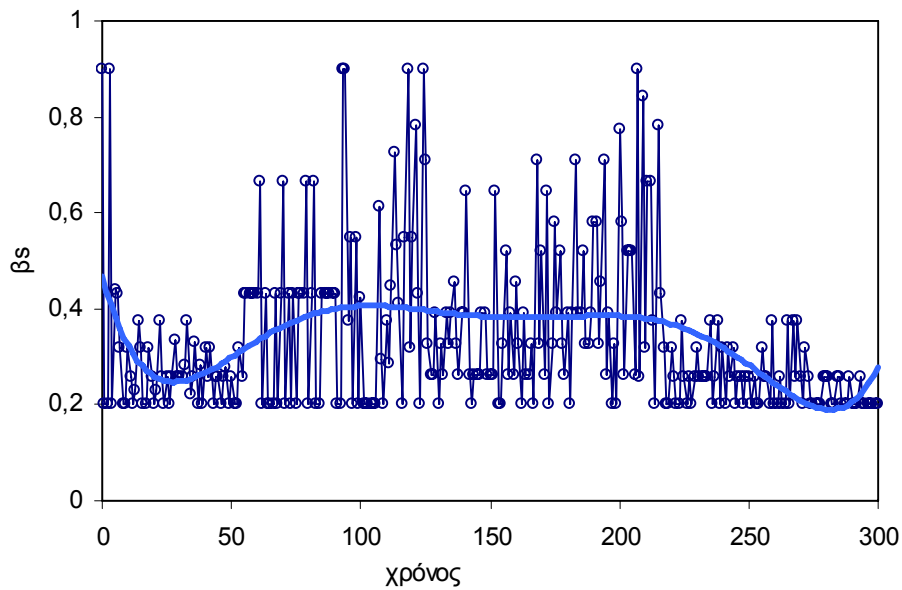
(α)



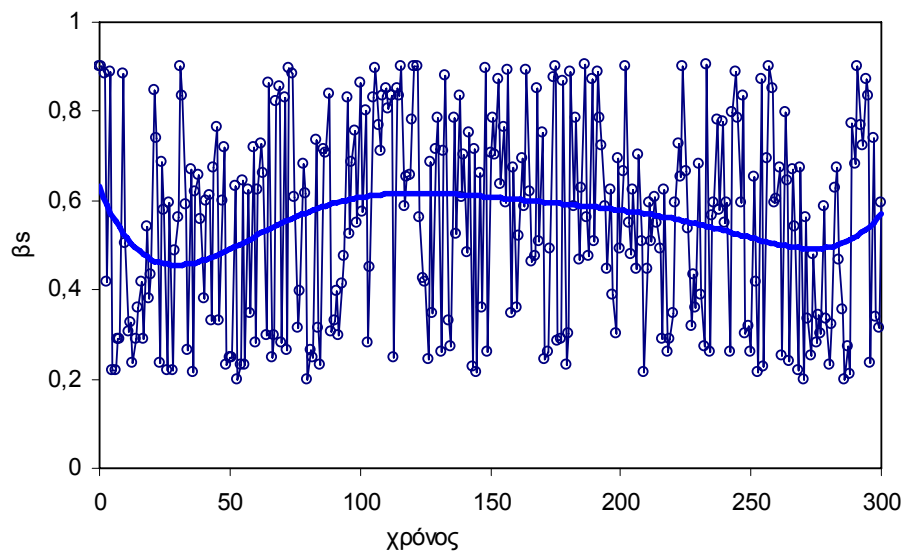
(β)

Σχήμα 4.7:(α) Ρεύμα δεδομένων θερμοκρασίας, S1 και (β) Ρεύμα δεδομένων ταχύτητας ανέμου, S2.

Τα δεδομένα που επιλέχθηκαν παρουσιάζουν διαφορετική μεταβλητότητα (το ρεύμα δεδομένων S1 είναι "ομαλότερο" από το S2) προκειμένου να αξιολογηθεί η αποδοτικότητα του προσαρμοστικού αλγορίθμου A1. Το Σχήμα 4.8 απεικονίζει το ρυθμό μόλυνσης β_s που προκύπτει για τα δύο ρεύματα δεδομένων σύμφωνα με τη σχέση (4.2).



(α)



(β)

Σχήμα 4.8: Ο ρυθμός μόλυνσης β_s για τα δεδομένα (α) θερμοκρασίας S1 και (β) ταχύτητας ανέμου S2.

Όπως φαίνεται, ο ρυθμός μόλυνσης για την περίπτωση των μετρήσεων της ταχύτητας του ανέμου (β_s^{S2}) έχει σε γενικές γραμμές υψηλότερες τιμές από το ρυθμό μόλυνσης για τα δεδομένα θερμοκρασίας (β_s^{S1}). Η μέση τιμή του ρυθμού μόλυνσης για το ρεύμα δεδομένων S2 είναι $\sum_{t=0}^T \beta_s^{S2}(t) = 0.562$ ενώ για το ρεύμα S1 είναι $\sum_{t=0}^T \beta_s^{S1}(t) = 0.335$. Το γεγονός αυτό είναι το αποτέλεσμα των γρήγορων αλλαγών στο ρεύμα δεδομένων S2. Ως εκ τούτου, ο ρυθμός β_s^{S2} λαμβάνει υψηλότερες τιμές ώστε η πηγή να διαδώσει όσο

το δυνατόν πιο ενημερωμένη πληροφορία πλαισίου, που στην προκειμένη περίπτωση είναι τα δεδομένα της ταχύτητας του ανέμου.

Στο τέλος της προσομοίωσης ($T=300$) καταγράφουμε τις τιμές που έχει λάβει ο πιο απομακρυσμένος κόμβος (destination). Κάθε τιμή που παράχθηκε από την πηγή έχει χρονοσημανθεί οπότε είμαστε σε θέση να γνωρίζουμε επακριβώς ποιες τιμές λάβαμε από την πηγή και ποιες όχι. Αυτό που μας ενδιαφέρει είναι από τις τιμές που έχει λάβει ο κόμβος προορισμός να μπορεί να «ανασυστήσει» την αρχική πληροφορία πλαισίου. Έτσι γίνεται κατανοητό ότι σε περιόδους που τα δεδομένα της πηγής δεν έχουν μεγάλη διακύμανση είναι περιττό να αποσταλούν όλες οι τιμές. Το σφάλμα της πληροφορίας πλαισίου που έλαβε ο κόμβος προορισμός σε σχέση με αυτό που έστειλε η πηγή ορίζεται ως η τιμή της διανυσματικής διαφοράς $V_s(t)$ και $V_d(t)$, δηλαδή

$$error(t) = ||V_s(t) - V_d(t)|| \quad (4.4)$$

όπου $V_s(t)$ είναι η τιμή της πληροφορίας πλαισίου (ή αλλιώς ρεύματος δεδομένων – data stream) που έχει παράγει η πηγή τη χρονική στιγμή t και $V_d(t)$, η τιμή που έχει λάβει ο κόμβος προορισμός για τη χρονική στιγμή t . Αν για κάποια χρονική στιγμή t' ο προορισμός δεν έχει λάβει τιμή τότε ως τιμή $V_d(t')$, λαμβάνεται η αμέσως προηγούμενη τιμή $V_d(t)$ που έχει λάβει με $t' > t$. Προφανώς όσες περισσότερες τιμές έχει λάβει ο κόμβος προορισμός τόσο μικρότερο θα είναι το σφάλμα που υπεισέρχεται και τόσο πιο εύκολα θα μπορεί να γίνει η ανασύσταση του αρχικού ρεύματος δεδομένων. Η πρόθεσή μας είναι να ελαχιστοποιήσουμε αυτό το σφάλμα ενώ συγχρόνως να διατηρήσουμε σε χαμηλά επίπεδα την δικτυακή κυκλοφορία. Με αυτό τον τρόπο, ο προτεινόμενος αλγόριθμος πρέπει να ρυθμίσει αποτελεσματικά την τιμή του ρυθμού μόλυνσης β_s ώστε να υπολογίζεται με βάση την αλλαγή των τιμών στο ρεύμα δεδομένων $V_s(t)$ που παράγεται στην πηγή (ή λαμβάνεται από ένα αισθητήρα προσαρμοσμένο στην πηγή).

Η προτεινόμενη μέθοδος αξιολογείται στο τέλος της προσομοίωσης όσον αφορά το μέσο σφάλμα στην πληροφορία πλαισίου στον πιο απομακρυσμένο κόμβο (προορισμό) όπως αυτό υπολογίζεται από τη σχέση (4.4), και τον συνολικό αριθμό των μηνυμάτων που μεταδόθηκαν στο δίκτυο από την πηγή. Η μέση τιμή του σφάλματος (av_error) συναρτήσει του χρόνου υπολογίζεται όπως στην (4.5)

$$av_error = \frac{1}{T} \sum_{t=0}^T error(t) \quad (4.5)$$

όπου T ο συνολικός χρόνος της προσομοίωσης. Επιπλέον, καθορίζουμε τον αριθμό $M(t)$ των μηνυμάτων που η πηγή έχει μεταδώσει σε κάθε χρονική στιγμή. Στο τέλος των πειραμάτων ο συνολικός αριθμός των μηνυμάτων M που μεταδόθηκαν από την πηγή δίνεται από την παρακάτω σχέση:

$$M = \sum_{t=0}^T M(t) \quad (4.6)$$

Για να συγκρίνουμε το προτεινόμενο προσαρμοστικό σχήμα σε σχέση με τα μη προσαρμοστικά σχήματα ορίζουμε την μετρική em που εκφράζει την αποδοτικότητα του επιδημικού σχήματος ως

$$em = \sqrt{(av_error_n)^2 + (M_n)^2} \quad (4.7)$$

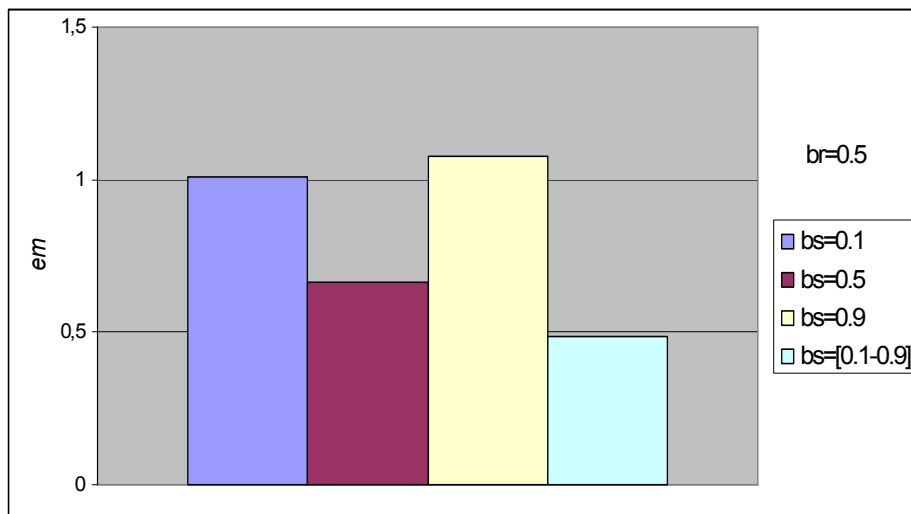
όπου av_error_n είναι η κανονικοποιημένη μορφή στο διάστημα $[0,1]$ του μέσου σφάλματος av_error με παράγοντα κανονικοποίησης το μέγιστο σφάλμα που παρατηρείται σε όλα τα επιδημικά σχήματα (προσαρμοστικό και μη προσαρμοστικά). Ο παράγοντας M_n εκφράζει τον κανονικοποιημένο αριθμό των μηνυμάτων στο διάστημα $[0,1]$ που εξέπεμψε η πηγή σε όλα τα επιδημικά σχήματα. Η κανονικοποίηση εκτελείται ώστε τα δύο μεγέθη (μέσο σφάλμα και αριθμός μηνυμάτων) να είναι συγκρίσιμα μεταξύ τους. Μπορεί παραδείγματος χάριν ανάλογα με την υπό μετάδοση πληροφορία πλαισίου ο αριθμός των μηνυμάτων να είναι τάξεις μεγέθους μεγαλύτερες από το μέσο σφάλμα.

Στο πρώτο πείραμά μας συγκρίνουμε ως προς την μετρική em τα επιδημικά σχήματα για το ρεύμα δεδομένων S1 θέτοντας σταθερές τιμές στο ρυθμό μόλυνσης $\beta_s = 0.1$, $\beta_s = 0.5$, $\beta_s = 0.9$ και μεταβλητές τιμές με $\beta_s \in [0.1 - 0.9]$, σύμφωνα με τα προαναφερθέντα για την προσαρμογή του β_s . Επίσης η αξιολόγηση γίνεται για εμβέλεια μετάδοσης r , $r = \alpha L$ με $\alpha = 0.15$. Ο ρυθμός μόλυνσης για τους κόμβους που δεν ανήκουν στη γειτονιά της πηγής έχει σταθερή τιμή $\beta_r = 0.5$. Η παράμετρος k της σχέσης (4.3), που υποδεικνύει τις πιο πρόσφατες διαφορές, τίθεται στην τιμή $k = 10$. Τα αποτελέσματα της προσομοίωσης παρουσιάζονται στον Πίνακα 4-7 και απεικονίζονται στο Σχήμα 4.9.

Πίνακας 4-7: Μέσο σφάλμα και αριθμός συνολικών μηνυμάτων που μεταδόθηκαν από την πηγή για τα δεδομένα θερμοκρασίας (S1).

| Ρυθμός μόλυνσης (β_s) | Μέσο σφάλμα (av_error) | Αριθμός μηνυμάτων (M) |
|-------------------------------|-----------------------------|---------------------------|
| 0.1 | 2.4141 | 40 |
| 0.5 | 0.8873 | 153 |
| 0.9 | 0.9540 | 278 |
| [0.1-0.9] | 1.046 | 62 |

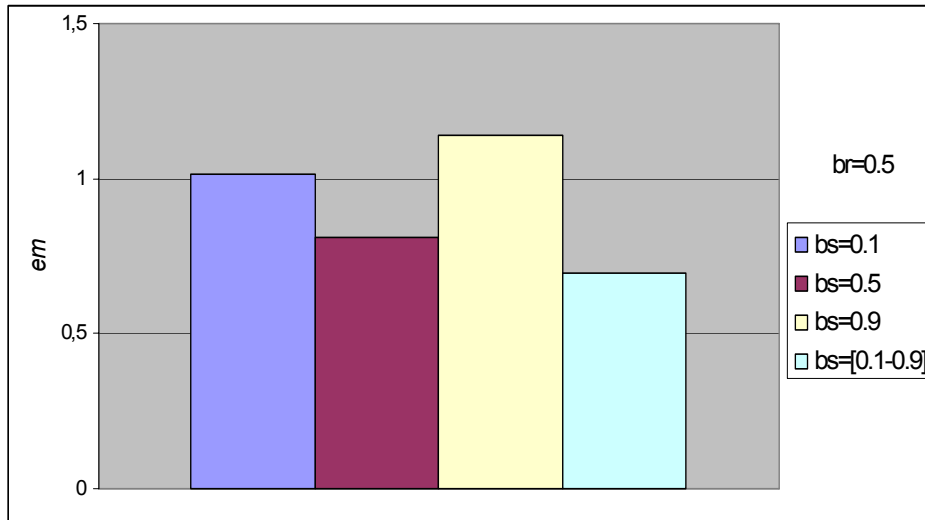
Από τα αποτελέσματα της προσομοίωσης παρατηρείται ότι το προσαρμοστικό επιδημικό μοντέλο συμπεριφέρεται αποδοτικότερα σε σχέση με τα μη προσαρμοστικά σχήματα ως προς το σφάλμα της πληροφορίας πλαισίου στον κόμβο προορισμό αλλά και το αριθμό των μηνυμάτων που εξέπεμψε η πηγή. Μπορεί σε σχέση με μικρές τιμές του β_s να μειονεκτεί ως προς τον αριθμό των μηνυμάτων που έστειλε η πηγή αλλά πλεονεκτεί ως προς το μέσο σφάλμα. Αντιθέτως σε σύγκριση με τις υψηλές τιμές του β_s μειονεκτεί στην μέση τιμή του σφάλματος (παρουσιάζει μεγαλύτερο σφάλμα) αλλά πλεονεκτεί ως προς τον αριθμό των μηνυμάτων.



Σχήμα 4.9: Η μετρική em των επιδημικών σχημάτων με σταθερό ρυθμό μόλυνσης $\beta_s=0.1$, $\beta_s = 0.5$, $\beta_s = 0.9$ και μεταβλητό ρυθμό $\beta_s = [0.1-0.9]$ για το ρεύμα δεδομένων S1 (θερμοκρασία). Για όλες τις περιπτώσεις ισχύει $\beta_r=0.5$.

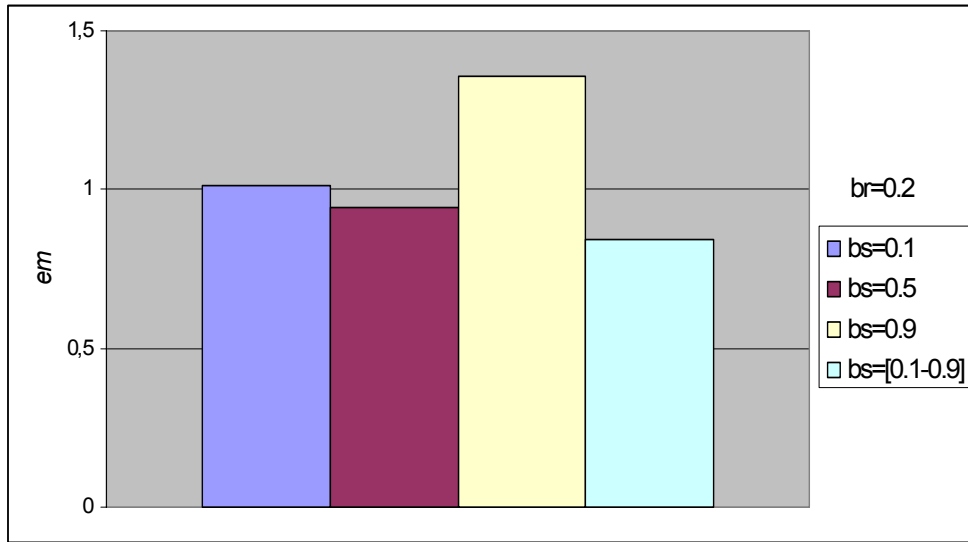
Με τις ίδιες τιμές του ρυθμού μόλυνσης β_s , της εμβέλειας μετάδοσης r και των k πιο πρόσφατων διαφορών όπως για το πρώτο πείραμα, εκτελέσαμε το δεύτερο πείραμα για το ρεύμα δεδομένων S2 (δεδομένα ταχύτητας ανέμου). Στο Σχήμα 4.10, παρουσιάζουμε τα αποτελέσματα της προσομοίωσης. Είναι αξιοπρόσεκτο ότι στην περίπτωση του S2 η μετρική em έχει γενικά υψηλότερες τιμές σε σχέση με το S1 κάτι

που οφείλεται στις τιμές του $S1$. Εντούτοις το προσαρμοστικό επιδημικό σχήμα είναι αποδοτικότερο σε σχέση με τα άλλα αφού η μετρική em είναι η χαμηλότερη.

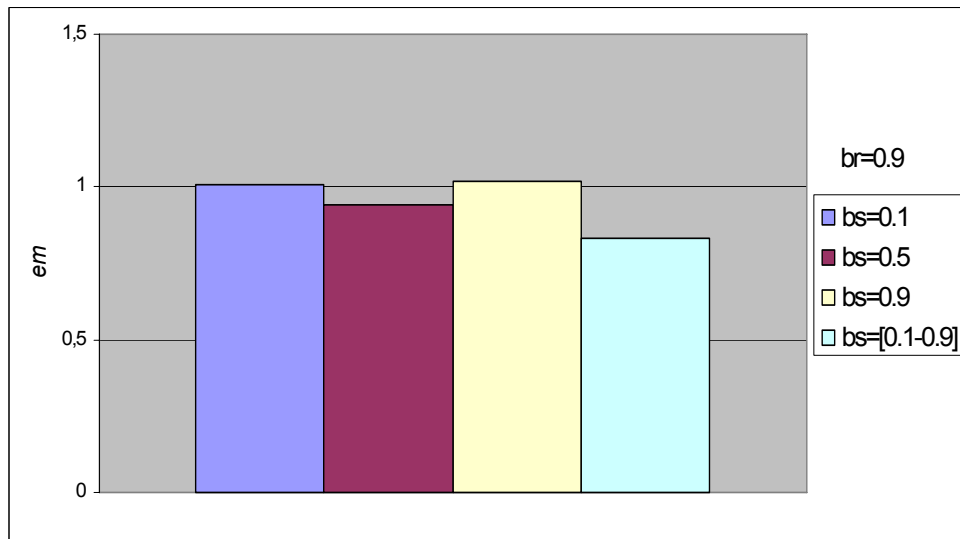


Σχήμα 4.10: Η μετρική em των επιδημικών σχημάτων με σταθερό ρυθμό μόλυνσης $\beta_s=0.1$, $\beta_s = 0.5$, $\beta_s = 0.9$ και μεταβλητό ρυθμό $\beta_s = [0.1-0.9]$ για το ρεύμα δεδομένων $S2$ (ταχύτητα ανέμου). Για όλες τις περιπτώσεις ισχύει $\beta_r=0.5$.

Στη συνέχεια μελετήσαμε τη συμπεριφορά των επιδημικών αλγορίθμων για διαφορετικές τιμές του ρυθμού μόλυνσης των ενδιάμεσων κόμβων $\beta_r=0.5$. Με τις ίδιες τιμές του ρυθμού μόλυνσης β_s , της εμβέλειας μετάδοσης r και των k πιο πρόσφατων διαφορών όπως προηγουμένως, εκτελέσαμε επιπλέον πειράματα για $\beta_r=0.2$ και $\beta_r=0.9$ για το ρεύμα δεδομένων $S1$ (θερμοκρασία). Στο Σχήμα 4.11 παρουσιάζουμε τα αποτελέσματα της προσομοίωσης. Όπως παρατηρούμε και σε περιπτώσεις χαμηλού ή υψηλού ρυθμού μόλυνσης β_r το προτεινόμενο προσαρμοστικό επιδημικό σχήμα είναι προτιμότερο από τα υπόλοιπα σχήματα επιδημικής διάδοσης. Επίσης για την περίπτωση του $\beta_r=0.9$ οι διαφορές που παρατηρούνται στη μετρική Em μεταξύ των επιδημικών σχημάτων είναι μικρότερες σε σχέση με τις περιπτώσεις $\beta_r=0.2$ και $\beta_r=0.5$ και αυτό οφείλεται στο γεγονός ότι η πληροφορία πλαισίου που μεταδίδει η πηγή (έστω και με χαμηλό ρυθμό μόλυνσης) φτάνει με μεγάλη πιθανότητα στον κόμβο προορισμό.



(α)



(β)

Σχήμα 4.11: Η μετρική em των επιδημικών σχημάτων με σταθερό ρυθμό μόλυνσης $\beta_s=0.1$, $\beta_s = 0.5$, $\beta_s = 0.9$ και μεταβλητό ρυθμό $\beta_s = [0.1-0.9]$ για το ρεύμα δεδομένων S1 (θερμοκρασία). (α) $\beta_r=0.2$ και (β) $\beta_r=0.9$.

4.3.2 Προσαρμοστικό Επιδημικό Σχήμα Διάδοσης Πληροφορίας L2

4.3.2.1 Περιγραφή Σχήματος

Στην περίπτωση του σχήματος L1 που περιγράφηκε στις προηγούμενες ενότητες ο ρυθμός μόλυνσης εξαρτάται μόνο από τα χαρακτηριστικά του ρεύματος δεδομένων. Θεωρούμε ότι ένας κόμβος χρησιμοποιεί (εκμεταλλεύεται) την πληροφορία πλαισίου που έχει λάβει και μολύνει τους γείτονές του με αυτήν μέχρι την λήψη ενός μηνύματος

που φέρει μία πιο πρόσφατη πληροφορία. Υπό αυτούς τους όρους, μπορούν να διαδοθούν στο δίκτυο απαρχαιωμένες πληροφορίες που δεν έχουν ουσιαστικά καμία αξία αφού η χρονική εγκυρότητά τους έχει πάψει να υφίσταται. Στις επόμενες παραγράφους εξετάζουμε μια εναλλακτική προσέγγιση για τη διάδοση πληροφορίας πλαισίου.

Στο προτεινόμενο σχήμα L2, υιοθετούμε μεταβλητό ρυθμό μόλυνσης για όλους τους κόμβους στο δίκτυο. Αν και οι κόμβοι μπορούν να χρησιμοποιήσουν απαρχαιωμένες πληροφορίες πλαισίου ελλείψει πιο πρόσφατων, η διάδοση αυτών των πληροφοριών (π.χ. μετρήσεις αισθητήρων) επηρεάζεται από την ηλικία τους, δηλαδή, τον χρόνο που έχει περάσει από τότε που έγινε η μέτρηση στο κόμβο πηγή. Οι πιο πρόσφατες τιμές των δεδομένων θα πρέπει να διαδοθούν γρήγορα σε έναν μεγάλο αριθμό κόμβων ενώ οι παλιότερες τιμές πρέπει να διαδοθούν αργά. Αυτό μπορεί να επιτευχθεί με τη ρύθμιση του ρυθμού μόλυνσης β_r στους ενδιάμεσους κόμβους. Ως εκ τούτου, ο προσαρμοστικός ρυθμός μόλυνσης β_r' για το σχήμα L2 εξαρτάται από την ηλικία των τιμών των επιμέρους μετρήσεων. Επομένως, κάθε χρονική στιγμή t , ισχύει $\beta_r'(t) = f(t)$. Στο σχήμα διάδοσης L2 ο ρυθμός μόλυνσης β_s για τη γειτονιά του κόμβου πηγή εξαρτάται μόνο από την μεταβλητότητα του ρεύματος δεδομένων λόγω του ότι αυτός ο κόμβος δημιουργεί σε κάθε χρονική στιγμή νέα πληροφορία πλαισίου.

Ορίζουμε ως $age(V_i(t)) = t - generationTime(V_i(t))$ την ηλικία της τιμής V_i που έχει ο κόμβος i τη χρονική στιγμή t . Ο μεταβλητός ρυθμός μόλυνσης β_r' , που περιλαμβάνει τον παράγοντα γήρανσης της πληροφορία πλαισίου, υπολογίζεται με βάση ένα σταθερό ρυθμό β_r (και αυξάνεται ή μειώνεται ανάλογα με την ηλικία της τιμής που θα μεταδοθεί. Η εκτίμηση του β_r' δίνεται από τη σχέση (4.8):

$$\beta_r' = \beta_r + \left[\beta_r - \frac{2\beta_r}{age_{max}} age(V_i(t)) \right] = 2\beta_r \left(1 - \frac{age(V_i(t))}{age_{max}} \right) \quad \begin{array}{l} \text{αν } \beta_r' > 1 \text{ τότε } \beta_r' = 1 \\ \text{αν } \beta_r' < 0 \text{ τότε } \beta_r' = 0 \end{array} \quad (4.8)$$

όπου η παράμετρος age_{max} είναι μια ορισμένη χρονική περίοδος που πέραν αυτής η τιμή $V_i(t)$ θεωρείται πλήρως απαρχαιωμένη και δεν πρέπει εν γένει να διαδοθεί περαιτέρω (π.χ., η τιμή της θερμοκρασίας πριν από 30 λεπτά θεωρείται μη ουσιαστική για μια εφαρμογή). Ως εκ τούτου, για $age(V_i(t)) = age_{max}$ η τιμή του ρυθμού είναι ίση με $\beta_r' = 0$, δηλαδή δεν υπάρχει κανένας λόγος να διαδοθούν χρονικά ξεπερασμένες τιμές της πληροφορίας πλαισίου. Ο προσαρμοστικός επιδημικός αλγόριθμος A2 που εφαρμόζεται στους ενδιάμεσους κόμβους προκειμένου να ρυθμιστεί ο ρυθμός μόλυνσης σύμφωνα με την σχέση (4.8) παρουσιάζεται στο Σχήμα 4.12.

```

Αλγόριθμος A2: Προσαρμοστική Επιδημική Διάδοση τη χρονική στιγμή
t για τον ενδιαμέσο κόμβο i
input:  $\beta_r$ ,  $age_{max}$ , Neighbourhood  $N(i)$ , time step t, value  $V_i(t)$ 
output:  $\beta_r'$ 
Begin:
While (true)
    /*calculation of value age*/
     $age(V_i(t)) = t - generationTime(V_i(t))$ 
    /*calculation of infection rate  $\beta_r'$ */
     $\beta_r' = 2 * \beta_r (1 - age(V_i(t)) / age_{max})$ 
    if ( $\beta_r' < 0$ )
         $\beta_r' = 0$ 
    if ( $\beta_r' > 1$ )
         $\beta_r' = 0$ 
    /*infect neighbourhood nodes */
    for each node j  $\in N(i)$  do
        infect (j,  $V_i(t)$  ,  $\beta_r'$ )
    end for
End While
End.

```

Σχήμα 4.12: Ενδιάμεσοι κόμβοι: προσαρμοστικός επιδημικός αλγόριθμος για τη διάδοση πληροφορίας πλαισίου

4.3.2.2 Προσομοίωση και Απόδοση του Σχήματος L2

Αξιολογούμε το προτεινόμενο σχήμα διάδοσης L2 σε σχέση με την ηλικία της πληροφορίας πλαισίου σε κάθε κόμβο κάθε χρονική στιγμή *t*. Αυτή η νέα μετρική γήρανσης της πληροφορίας πλαισίου (*context ageing*) στην κανονικοποιημένη της μορφή δίνεται από την σχέση (4.9):

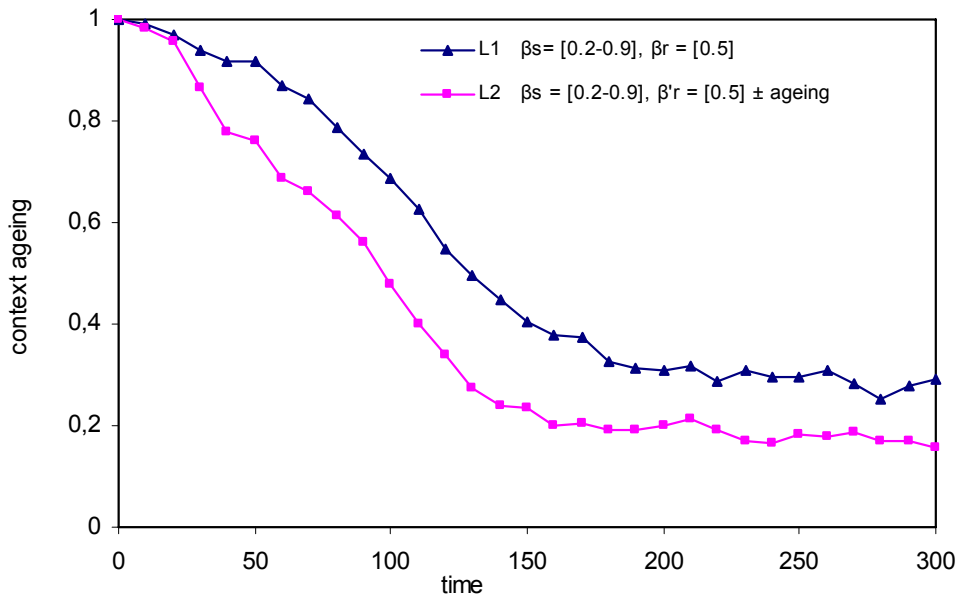
$$context_ageing(t) = \frac{\sum_{i=1}^N age(V_i(t))}{\max\{\sum_{i=1}^N age(V_i(t))\}} \quad (4.9)$$

Η μετρική της γήρανσης της πληροφορίας πλαισίου υποδηλώνει τη μέση τιμή του πόσο ανανεωμένη είναι η πληροφορία στο δίκτυο. Συγκεκριμένα, μια χαμηλή τιμή της μετρικής αυτής δείχνει «φρέσκια» πληροφορία στο δίκτυο ενώ μια υψηλή τιμή δείχνει ότι το δίκτυο έχει κατακλυσθεί από απαρχαιωμένη μη ουσιώδη πληροφορία πλαισίου.

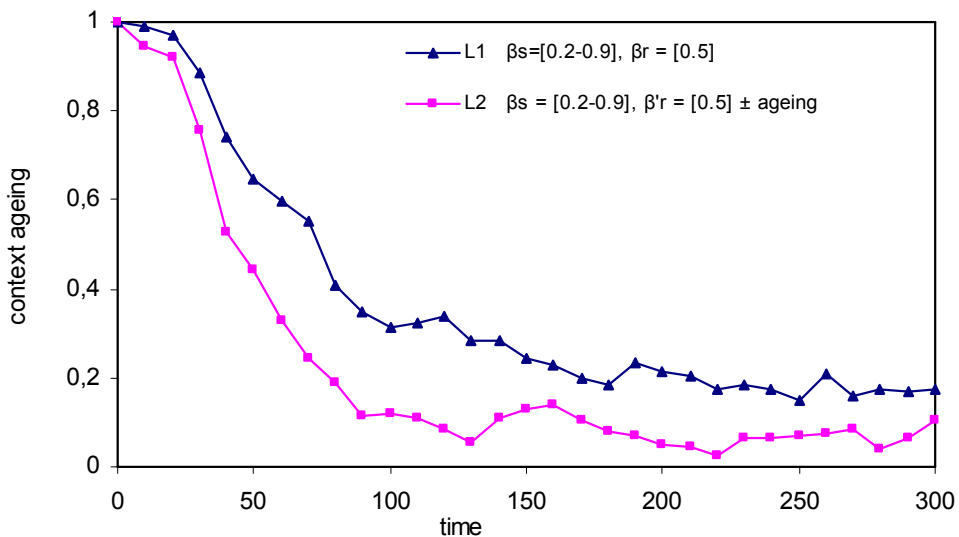
Συγκρίνουμε τα δύο επιδημικά σχήματα L1 και L2. Οι αλγόριθμοι κάθε προσαρμοστικού επιδημικού σχήματος προσομοιώθηκαν για χρόνο $T = 300$ μέσω 1000 εκτελέσεων με εμβέλεια μετάδοσης $r = \alpha L$ ($\alpha = 0.15$ και $L=100$). Για το σχήμα L1 εφαρμόζουμε σταθερό ρυθμό μόλυνσης για όλους τους ενδιαμέσους κόμβους, $\beta_r = 0.5$. Για το σχήμα L2, ο ρυθμός μόλυνσης της γειτονιάς των ενδιαμέσων κόμβων υπολογίζεται με βάση τον

αλγόριθμο A2 που περιγράφεται στο Σχήμα 4.12 της προηγούμενης ενότητας. Η παράμετρος age_{max} τίθεται στην τιμή 10 και για την περίπτωση που ένας κόμβος i δεν έχει μολυνθεί ακόμα με κάποια πληροφορία θεωρούμε ότι $age(V_i(t))=300$.

Τα αποτελέσματα της προσομοίωσης όσον αφορά τη γήρανση της πληροφορίας πλαισίου (*context ageing*) απεικονίζονται στο Σχήμα 4.13 για τα ρεύματα S1 και S2.



(α)



(β)

Σχήμα 4.13: Γήρανση πληροφορίας για το σχήμα L1 ($\beta_s = [0.2-0.9]$, $\beta_r = 0.5$) και το L2 ($\beta_s = [0.2-0.9]$, $\beta'_r = [0.5] \pm ageing$) για (α) S1 και (β) S2. Εμβέλεια μετάδοσης $r = 0.15L$.

Στην περίπτωση της γήρανσης της πληροφορίας πλαισίου το σχήμα L2 είναι αποδοτικότερο από το L1 και για τα δύο ρεύματα δεδομένων. Κατά συνέπεια η εφαρμογή του συγκεκριμένου σχήματος οδηγεί σε πιο επικαιροποιημένη πληροφορία πλαισίου στο δίκτυο.

4.4 Συμπεράσματα

Η έρευνα που διεξήχθη στο πεδίο της βιο-μιμητικής συνεργατικής επίγνωσης πλαισίου στο διάχυτο υπολογισμό συνεισέφερε στα παρακάτω βασικά σημεία:

- Προτάθηκαν και αξιολογήθηκαν αποδοτικά σχήματα για την επιδημική διάδοση προσαρμόζοντας το ρυθμό μόλυνσης β των κόμβων ανάλογα με την πυκνότητα του δικτύου και την κινητικότητά τους.
- Προτάθηκαν και αξιολογήθηκαν αποδοτικά σχήματα για την επιδημική διάδοση προσαρμόζοντας το ρυθμό μόλυνσης β των κόμβων ανάλογα με την μεταβλητότητα αλλά και την «ηλικία» του μεταδιδόμενου πλαισίου.
- Η διάδοση πληροφορίας πλαισίου που επιτυγχάνουν τα παραπάνω σχήματα δεν επιβαρύνει το δίκτυο με παραπάνω (μη ουσιώδη) μηνύματα λειτουργώντας έτσι σαν ένας μηχανισμός εξοικονόμησης ενέργειας.
- Η παρωχημένη πληροφορία που βρίσκεται στο δίκτυο δεν μεταδίδεται περαιτέρω, αναβαθμίζοντας έτσι την χρονική ποιότητα του πλαισίου στο δίκτυο.

Τα θέματα που αναλύθηκαν στο κεφάλαιο αυτό αναφέρονται στις δημοσιεύσεις: [76], [77],[78].

ΚΕΦΑΛΑΙΟ 5

ΣΥΝΕΡΓΑΤΙΚΗ ΕΠΙΓΝΩΣΗ ΠΛΑΙΣΙΟΥ ΓΙΑ ΤΟΝ ΣΥΜΠΕΡΑΣΜΟ ΤΗΣ ΑΞΙΟΠΙΣΤΙΑΣ ΤΩΝ ΚΟΜΒΩΝ ΣΕ ad-hoc ΔΙΚΤΥΑ

5.1 Εισαγωγή

Στο κεφάλαιο αυτό χρησιμοποιήσαμε τις ιδιότητες της Συνεργατικής Επίγνωσης Πλαισίου για την δημιουργία ενός συστήματος οικοδόμησης εμπιστοσύνης (trust) μεταξύ των κόμβων ενός ασύρματου δικτύου. Η αξιοπιστία ενός κόμβου (για μια συγκεκριμένη υπηρεσία που προσφέρει, π.χ. προώθηση πακέτων) αποτελεί πληροφορία πλαισίου για αυτόν τον κόμβο. Ο συμπερασμός λοιπόν της αξιοπιστίας ενός κόμβου γίνεται βάσει των άμεσων παρατηρήσεων της δραστηριότητας του κόμβου (π.χ. αν προωθεί τα πακέτα τα οποία του στείλαμε) και των συστάσεων από «αξιόπιστους» συστήνοντες. Οι άμεσες παρατηρήσεις της δραστηριότητας του κόμβου και οι συστάσεις αποτελούν και αυτές πληροφορία πλαισίου που συνδυάζονται για να οδηγηθούμε στον συμπερασμό της αξιοπιστίας του κόμβου.

Λόγω της μετακίνησης των κινητών κόμβων, η δομή ενός ad-hoc δικτύου (Mobile Ad hoc Network ή αλλιώς MANET) αναδιοργανώνεται και επομένως η τοπολογία του ποικίλλει με το χρόνο. Οι κινητοί κόμβοι συνεχώς συνδέονται ή χωρίζονται, σύμφωνα με τη θέση τους και τις ρυθμίσεις των άλλων κόμβων. Το κίνητρο ενός νεοεισερχόμενου κόμβου που επιθυμεί να συνδεθεί στο δίκτυο είναι η προσφορά υπηρεσιών όπως π.χ. δρομολόγηση πακέτων δεδομένων στους υπάρχοντες κόμβους, οι οποίοι με τη σειρά τους προσφέρουν ισοδύναμες υπηρεσίες-λειτουργίες στα νέα μέλη. Αυτή η συνεργασία εμπεριέχει ένα σημαντικό παράγοντα για την εύρυθμη λειτουργία του ad hoc δικτύου, την αξιοπιστία που διέπει τους κόμβους.

Οι γειτονικοί κόμβοι αναπτύσσουν με το χρόνο κάποιο βαθμό εμπιστοσύνης μεταξύ τους έτσι ώστε να είναι εφικτή η συνεργασία τους και να βελτιωθεί η ασφάλεια, η συνδετικότητα και η ποιότητα των παρεχόμενων υπηρεσιών από το δίκτυο. Από την άλλη πλευρά η τιμή της αξιοπιστίας που έχει δημιουργήσει ένας κόμβος για τους γειτονικούς του κόμβους ελαττώνεται όταν αυτοί απομακρύνονται λόγω της κινητικότητας. Αυτό οφείλεται στο γεγονός ότι πιθανώς αναξιόπιστοι κόμβοι παρεμβάλλονται στο μονοπάτι της μεταξύ τους επικοινωνίας. Κατά συνέπεια, η εμπιστοσύνη που καθιερώνεται μεταξύ δύο κόμβων για τη δρομολόγηση και την αποστολή πακέτων μπορεί να χαθεί με το χρόνο, επηρεάζοντας έτσι την απόδοση του δικτύου. Επιπλέον, κανείς δεν μπορεί να υποθέσει ότι οι κόμβοι ενός τέτοιου δικτύου συμπεριφέρονται λογικά [57]. Οι εγωιστικοί (selfish) και οι κακόβουλοι (malicious)

κόμβοι μπορεί να έχουν αρχικά συνεργατική συμπεριφορά προκειμένου να συνδεθούν στο δίκτυο, αλλά αργότερα οι προθέσεις τους να αλλάξουν.

Ένας εγωιστικός κόμβος είναι απρόθυμος να ξοδέψει τους πόρους του (ενέργεια, υπολογιστική ισχύ, κα.) προκειμένου να εξυπηρετηθούν οι διαδικασίες του δικτύου και να μεγιστοποιηθεί η «ευημερία» (π.χ., προώθηση πακέτων πληροφορίας πλαισίου που δεν προορίζονται για αυτόν). Αντιθέτως απαιτεί την εκτέλεση εκείνων των διαδικασιών του δικτύου που συμβάλουν ώστε να μεγιστοποιήσει το κέρδος του (π.χ., ζητά την παράδοση πακέτων που προέρχονται ή που προορίζονται για αυτόν). Αντίστοιχα οι κακόβουλοι κόμβοι έχουν ως κύριο στόχο να βλάψουν τη λειτουργία του δικτύου, παρά να εξοικονομήσουν πόρους (π.χ., ενέργεια μπαταρίας). Η τεχνική της πλημμύρας (flooding) και η τεχνική της στέρησης ύπνου (sleep deprivation torture) [58] χρησιμοποιούνται συνήθως από τους κακόβουλους κόμβους. Με την πρώτη τεχνική οι κόμβοι αυτοί πλημμυρίζουν το δίκτυο με ψευδή πακέτα δρομολόγησης προκειμένου να καταναλωθούν οι πόροι των υπολοίπων κόμβων και να δυσλειτουργήσει η διαδικασία της δρομολόγησης (routing paths). Η δεύτερη τεχνική στοχεύει στην κατανάλωση μπαταριών ενός συγκεκριμένου κόμβου στέλνοντας του συνεχώς πακέτα για δρομολόγηση. Τέλος, οι κόμβοι χάκερ (hacker nodes) έχουν σαν στόχο την υποκλοπή των πληροφοριών που ανταλλάσσονται μεταξύ των κόμβων, ακυρώνοντας κατά συνέπεια τη συνεργασία που υπάρχει στο δίκτυο. Οι κόμβοι αυτοί συνήθως υλοποιούν τις προθέσεις τους μέσω επιθέσεων τύπου Sybil (Sybil attack) [59].

Οι αυτοεξελισσόμενοι (self evolving) μηχανισμοί που βασίζονται στη διάδοση φήμης (reputation based) θεωρούνται κατάλληλοι για την καθιέρωση εμπιστοσύνης μεταξύ των κόμβων ενός δικτύου όπως τα MANET, όπου τα κέντρα διανομής πιστοποιητικών (certificates) είναι σπάνια και οι πόροι των κόμβων (π.χ., ενέργεια, μνήμη) είναι λιγοστοί. Τα σχήματα αυτά βασίζονται στον προσδιορισμό της αξιοπιστίας των κόμβων, σχετικά με τις υπηρεσίες που προσφέρουν. Ένας από τους βασικούς στόχους των «λογικών» κόμβων είναι να συνεργαστούν μεταξύ τους προκειμένου να αποφύγουν, ή ακόμα και να απομονώσουν τους κακόφημους κόμβους (δηλ., τους εγωιστικούς, κακόβουλους ή χάκερ). Αυτή η συνεργασία απαιτεί την ανταλλαγή πληροφορίας πλαισίου μεταξύ των κόμβων όπως συστάσεων (recommendations), και τον προσδιορισμό των αξιόπιστων/έμπιστων συστηνόντων (trusted recommenders). Επιπρόσθετοι στόχοι περιλαμβάνουν την ελαχιστοποίηση των αποτελεσμάτων που εισήχθησαν στις κανονικές διαδικασίες από τους κακόφημους κόμβους.

Στη διάρκεια της παρούσας διδακτορικής διατριβής αναπτύχθηκε ένα πλαίσιο για την αυτοεξελισσόμενη καθιέρωση εμπιστοσύνης (self evolving trust) μεταξύ των κόμβων ενός δικτύου που ονομάζεται πλαίσιο εμπιστοσύνης για τα αδόμητα δίκτυα (Ad hoc Trust Framework ή αλλιώς ATF). Το πλαίσιο αυτό βασίζεται σε μία κατανεμημένη αρχιτεκτονική και στη συνεργασία μεταξύ των κόμβων. Κάθε συστατικό (module) του ATF (που βρίσκεται σε όλους τους κόμβους) εκτελεί ένα καθορισμένο σύνολο ενεργειών για να αξιολογηθεί η αξιοπιστία κάποιου συγκεκριμένου κόμβου ή να συστηθεί αυτός σε τρίτους. Για να αξιολογηθεί ο βαθμός αξιοπιστίας των ομότιμων κόμβων (peers) ενσωματώνονται άμεσα αποδεικτικά στοιχεία (direct evidence - *DE*), έμμεσες συστάσεις (second hand recommendations), η υποκειμενική κρίση (subjective judgment) και ιστορικά στοιχεία. Όλα αυτά αποτελούν πληροφορία πλαισίου που οδηγεί στο συμπέρασμα της αξιοπιστίας ενός κόμβου. Για την ανίχνευση αυτής της σημασιολογίας, προτείνεται ένας καινοφανής μηχανισμός που βασίζεται επίσης στη συμπεριφορά του φυσικού χρήστη (τον κάτοχο του κόμβου) μέσω μιας πολιτικής εμπιστοσύνης (Trust Policy) που καθορίζεται από τον ίδιο.

Το πλαίσιο ATF αποσυνδέεται από τις παραδοσιακές μεθόδους οικοδόμησης εμπιστοσύνης (trust building) όπως η συμμετρική ή η δημόσια κρυπτογραφία, αποφεύγοντας έτσι τους σύνθετους υπολογισμούς και την κατανάλωση πόρων (ενέργεια, CPU και μνήμη). Υπό αυτή την έννοια, το προτεινόμενο πλαίσιο υπολογισμού αξιοπιστίας μέσω της συνεργατικής επίγνωσης πλαισίου μπορεί να χρησιμοποιηθεί σε διαφορετικούς τύπους κατανεμημένων συστημάτων ή συστημάτων ομοτίμων (ad hoc δίκτυα, υπηρεσίες διάχυτου υπολογισμού, αυτόνομα συστήματα), αν και στα πλαίσια της διατριβής αξιοποιείται πρωτίστως στα ad hoc δίκτυα.

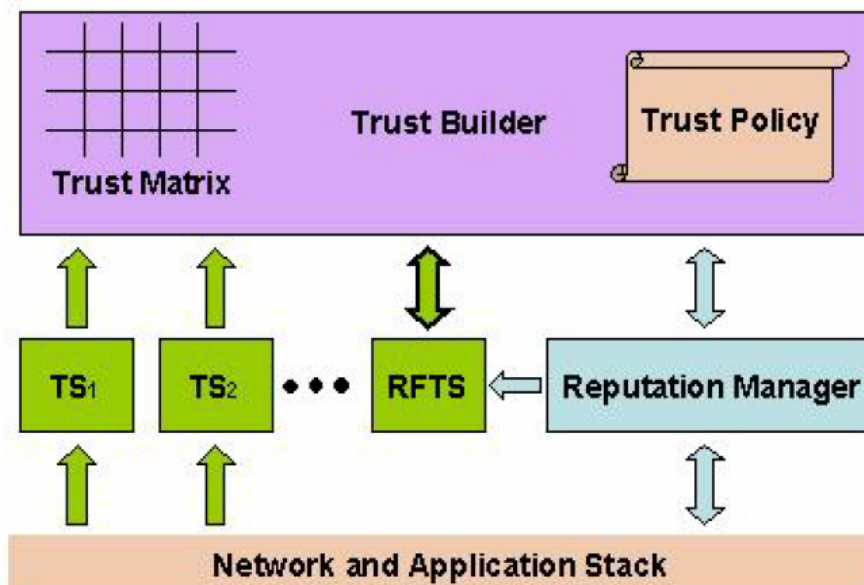
5.2 Γενική αρχιτεκτονική του πλαισίου ATF

Το πλαίσιο ATF χρησιμοποιεί μια στρωματοποιημένη αρχιτεκτονική, η οποία απεικονίζεται στο Σχήμα 5.1 και αποτελείται από τα ακόλουθα συστατικά:

- *Αισθητήρες Αξιοπιστίας (Trust Sensors ή αλλιώς TS)*
- *Δημιουργός Αξιοπιστίας (Trust Builder ή αλλιώς TB)*
- *Διαχειριστής Φήμης (Reputation Manager ή αλλιώς RM)*

Το προτεινόμενο πλαίσιο είναι πλήρως κατανεμημένο. Κάθε κόμβος είναι «εξοπλισμένος» με αυτά τα συστατικά και παρέχει διάφορες λειτουργίες (υπηρεσίες) όπως δρομολόγηση, προώθηση πακέτων, κλπ. Επιπλέον, κάθε κόμβος εφαρμόζει μια

ειδική λειτουργία, αποκαλούμενη και ως *Λειτουργία (υπηρεσία) Σύστασης (Recommendation Function)*. Αποτελεί μία απλή υπηρεσία που παρέχει συστάσεις προς άλλους, τρίτους κόμβους όταν αυτό ζητηθεί. Το πλαίσιο ATF υιοθετεί τον ορισμό που εισάγεται στην [60] για τη φήμη μίας υπηρεσίας και καθορίζεται ως εξής: *Φήμη = {Αναγνωριστικό κόμβου(NodeId), Υπηρεσία (Function), Βαθμός αξιοπιστίας (Trust Value)}*. Κατά συνέπεια, η φήμη για μια συγκεκριμένη λειτουργία f ενός κόμβου n ορίζεται ως $R(n,f) = \{n, f, TV_{n,f}\}$ όπου το $TV_{n,f}$ είναι ο βαθμός αξιοπιστίας για την υπηρεσία f του κόμβου n . Ο βαθμός αξιοπιστίας TV ενημερώνεται μέσω των άμεσων αποδεικτικών στοιχείων, των συστάσεων, της υποκειμενικής κρίσης και των ιστορικών δεδομένων.



Σχήμα 5.1: Αρχιτεκτονική του πλαισίου ATF.

Θεωρούμε απαραίτητο να αναφερθεί εδώ η ονοματολογία που θα χρησιμοποιηθεί κατωτέρω. Χρησιμοποιούμε τον όρο *ανιχνευτής (detector)* να ορίσουμε έναν κόμβο που παρακολουθεί άμεσα τη συμπεριφορά ενός άλλου κόμβου ο οποίος καλείται *στόχος (target)*. Σε μία τέτοια περίπτωση ο *ανιχνευτής* συλλέγει τα άμεσα αποδεικτικά στοιχεία (*DE*) σχετικά με την αξιοπιστία μιας συγκεκριμένης υπηρεσίας που παρέχει ο κόμβος *στόχος*. Ένας *αιτών (requestor)* είναι ένας κόμβος που αιτείται συστάσεις για άλλους κόμβους. Ένας *συστήνων (recommender)* παρέχει απαντήσεις σε τέτοιες αιτήσεις. *Γειτονικοί* ονομάζονται οι κόμβοι που βρίσκονται ο ένας στην εμβέλεια του άλλου (*one-hop*). Ένα *μονοπάτι* είναι μια ακολουθία από ασύρματες συνδέσεις (*hops*) που οδηγούν από τον κόμβο πηγή σε έναν κόμβο προορισμού.

Γενικά, ένας μηχανισμός οικοδόμησης εμπιστοσύνης θα μπορούσε να σχεδιαστεί βασιζόμενος σε δύο διαφορετικές αρχιτεκτονικές κατευθύνσεις. Η πρώτη στηρίζεται σε ένα σύστημα φήμης *κατόπιν αίτησης* (*on demand*), ενώ η δεύτερη σε ένα σύστημα *οδηγούμενο από τα γεγονότα* (*event driven*). Η διαφορά μεταξύ των δύο αυτών προσεγγίσεων έγκειται στον τρόπο με τον οποίο οι κόμβοι ενημερώνονται για τις αλλαγές στις τιμές του βαθμού αξιοπιστίας (*Trust Values - TVs*). Στην πρώτη περίπτωση (*on demand*), ένας *συστήνων* κόμβος ενημερώνει τον ενδιαφερόμενο κόμβο (δηλαδή τον αιτούντα) για την τιμή της *TV μιας υπηρεσίας* ενός συγκεκριμένου κόμβου στόχου, μόνο όταν ο αιτών ζητεί ρητά αυτήν την τιμή. Στη δεύτερη προσέγγιση (*event driven*), οποτεδήποτε ένας *ανιχνευτής* επαναυπολογίζει μια τιμή *TV* για έναν κόμβο στόχο, αυτός ενημερώνει αμέσως τους άλλους κόμβους, μέσω ενός μηχανισμού broadcast (εκπομπής), multicast (σημείο προς πολλά σημεία) ή unicast (σημείο προς σημείο). Η αρχιτεκτονική του προτεινόμενου πλαισίου *ATF* είναι σχεδιασμένη να παρέχει *συστάσεις* (*recommendations*) *κατόπιν αίτησης* (*on demand*). Στις επόμενες παραγράφους παρουσιάζουμε αναλυτικά τα συνιστώμενα μέρη της αρχιτεκτονικής καθώς και το ρόλο τους.

5.2.1 Αισθητήρες Αξιοπιστίας

Για κάθε υπηρεσία που προσφέρεται από έναν γειτονικό κόμβο υπάρχει ένας *Αισθητήρας Αξιοπιστίας* (*Trust Sensor-TS*) που ελέγχει την εκτέλεση της συγκεκριμένης υπηρεσίας. Ουσιαστικά αντιστοιχεί τα άμεσα αποδεικτικά στοιχεία σε μία αριθμητική τιμή και διαβιβάζει αυτήν την τιμή στον *Δημιουργό Αξιοπιστίας* (*Trust Builder-TB*) για περαιτέρω υπολογισμούς. Ο μηχανισμός που συλλαμβάνει τα άμεσα αποδεικτικά στοιχεία για μια υπηρεσία ενός γειτονικού κόμβου είναι μια γενίκευση του μηχανισμού που περιγράφεται στην [61]. Επιπλέον, αυτά τα στοιχεία μπορούν να ελέγχονται μέσω μετρήσεων πραγματικού χρόνου αλλά και της στατιστικής ανάλυσης των καταγεγραμμένων δεδομένων (logs). Οι κόμβοι αποκτούν ένα σύνολο από τιμές *TSS* και χρησιμοποιούν ένα επιθυμητό υποσύνολο σύμφωνα με τις ανάγκες τους.

5.2.2 Δημιουργός Αξιοπιστίας

Αυτό το συστατικό υπολογίζει το βαθμό αξιοπιστίας *TV* των υπηρεσιών άλλων κόμβων. Οι τιμές *TV* αποθηκεύονται σε μια *Μήτρα Αξιοπιστίας* (*Trust Matrix*), όπου ο *Δημιουργός Αξιοπιστίας* (*TB*) διατηρεί και ενημερώνει, και χρησιμοποιούνται όταν υπάρχει πρόθεση

για συνεργασία ή αλληλεπίδραση με άλλους κόμβους. Η τιμή TV είναι μοναδική για κάθε υπηρεσία ανά κόμβο. Ένας κόμβος, παραδείγματος χάριν, μπορεί να είναι αξιόπιστος για να εκτελέσει προώθηση πακέτων αλλά αναξιόπιστος όσον αφορά τη δρομολόγηση. Ο υπολογισμός της τιμής TV εξαρτάται από διάφορους παράγοντες, όπως η συμπεριφορά του εξεταζόμενου κόμβου στο παρελθόν (ιστορία), οι συστάσεις από αξιόπιστους τρίτους κόμβους, και τα άμεσα αποδεικτικά στοιχεία. Κάθε παράγοντας από τους προαναφερθέντες συμβάλλει με ένα συγκεκριμένο βάρος. Η ιστορία και τα άμεσα στοιχεία μπορούν να εισαγάγουν υψηλή βεβαιότητα για τη συμπεριφορά ενός κόμβου, ενώ οι συστάσεις από τρίτους μπορεί να έχουν λιγότερη συμβολή. Επιπλέον, κάθε κόμβος ακολουθεί την δική του *Πολιτική Οικοδόμησης Εμπιστοσύνης (Trust Policy)*. Η πολιτική αυτή εισάγει στο μοντέλο την υποκειμενικότητα του χρήστη (π.χ., δύσπιστος, ευκολόπιστος), εκφρασμένη με μία χρονικά μεταβαλλόμενη συνάρτηση.

5.2.3 Διαχειριστής Φήμης

Ο κύριος ρόλος του *Διαχειριστή Φήμης (RM)* είναι η παροχή συστάσεων από τρίτους κόμβους στο *Δημιουργό Αξιοπιστίας (TB)* ούτως ώστε ο τελευταίος να υπολογίσει τις τιμές του βαθμού αξιοπιστίας TV . Με ένα μηχανισμό κατόπιν αίτησης (on demand) ο TB ζητά μια σύσταση για ένα κόμβο στόχο. Στη συνέχεια ο RM επιλέγει τους κόμβους που θα επικοινωνήσει για να λάβει τις ζητούμενες τιμές. Αυτοί θα πρέπει να είναι όσο το δυνατόν αξιόπιστοι και σε μικρή εμβέλεια ούτως ώστε να περιοριστεί το κόστος επικοινωνίας. Έτσι ο RM λαμβάνει υπόψη το βαθμό αξιοπιστίας της υπηρεσίας συστάσεων (recommendation function) των κόμβων στους οποίους κάνει την αίτηση αλλά και την απόστασή τους. Όταν ένας συστήνων κόμβος δεχθεί ένα αίτημα για σύσταση, ο δικός του αντίστοιχος RM επικοινωνεί με τον *Δημιουργό Αξιοπιστίας (TB)* και λαμβάνει την TV για τη ζητούμενη υπηρεσία του κόμβου στόχου εάν αυτή υπάρχει. Έπειτα, ο συστήνων κόμβος επιστρέφει αυτήν την τιμή στον αιτούντα κόμβο. Ο TB χρησιμοποιεί τις συστάσεις που συλλέγονται από τον RM και ο τελευταίος χρησιμοποιεί τις τιμές TV που έχουν ήδη υπολογιστεί προκειμένου να ενημερώσει άλλους κόμβους.

5.2.4 Αισθητήρας Αξιοπιστίας της Υπηρεσίας Σύστασης

Ο *Αισθητήρας Αξιοπιστίας της Υπηρεσίας Σύστασης (Recommendation Function Trust Sensor - RFTS)* αποτελεί έναν ειδικό αισθητήρα που αξιολογεί την αξιοπιστία ενός κόμβου σχετικά με την υπηρεσία σύστασης που αυτός προσφέρει. Ο $RFTS$ όπως

οποιοδήποτε άλλος αισθητήρας TS , ταξινομεί κάθε άμεση παρατήρηση ως *επιτυχία* ή *αποτυχία*. Όπως αναφέρθηκε προηγουμένως ο *Διαχειριστής Φήμης RM* ενός κόμβου ανιχνευτή λαμβάνει από τους συστήνοντες κόμβους τις συστάσεις που αντιστοιχούν σε μία συγκεκριμένη υπηρεσία ενός κόμβου στόχο. Μια σύσταση αποστέλλεται μόνο όταν ο συστήνων κόμβος έχει επαρκή άμεσα αποδεικτικά στοιχεία DE για το στόχο ούτως ώστε να μειωθεί η διάδοση λανθασμένης φήμης. Αυτές οι τιμές προωθούνται στη συνέχεια στον *Δημιουργό Αξιοπιστίας TB* και τον αισθητήρα $RFTS$. Για κάθε συστήνοντα κόμβο ο $RFTS$ αποθηκεύει σε έναν πίνακα τις τιμές που λαμβάνει από τον *Διαχειριστή Φήμης RM*. Οποτεδήποτε ο κόμβος ανιχνευτής ολοκληρώνει μια άμεση αλληλεπίδραση με τον κόμβο στόχο, ο αισθητήρας TS για τη συγκεκριμένη υπηρεσία επιστρέφει *επιτυχία* ή *αποτυχία* στο *Δημιουργό Αξιοπιστίας TB* σύμφωνα με την παρατήρησή του. Ο TB έπειτα υπολογίζει την τιμή DE (που αφορά την άμεση παρατήρηση) για τον κόμβο στόχο και επιστρέφει αυτήν την τιμή στον αισθητήρα $RFTS$ για την αξιολόγηση της υπηρεσίας σύστασης των συστηνόντων κόμβων. Ο $RFTS$ συγκρίνει την τιμή DE με την τιμή TV_{RF} που είναι αποθηκευμένη για κάθε συστήνοντα κόμβο. Εάν οι δύο οι τιμές δεν παρουσιάζουν ιδιαίτερη απόκλιση (ανάλογα με την *Πολιτική Οικοδόμησης Εμπιστοσύνης* που ακολουθείται), ο αισθητήρας $RFTS$ επιστρέφει *επιτυχία* ή *αποτυχία* στον *Δημιουργό Αξιοπιστίας TB*. Ένας κόμβος ανιχνευτής (δηλαδή ο εκτιμητής) πρέπει να είναι σε θέση να ελέγξει την τιμή TV της υπηρεσίας σύστασης εκείνων των κόμβων που προσφέρουν συστάσεις (συστηνόντων). Κατά συνέπεια, ένας μηχανισμός ελέγχου είναι απαραίτητος. Ανά τακτά χρονικά διαστήματα ο κόμβος εκτιμητής ζητά συστάσεις για τις υπηρεσίες των κόμβων στόχων για τις οποίες διατηρεί ο ίδιος άμεσα αποδεικτικά στοιχεία (DE) και ειδικά για εκείνους τους κόμβους με τους οποίους έχει έναν μεγάλο αριθμό αλληλεπιδράσεων στο πρόσφατο παρελθόν. Όταν οι συστάσεις φθάνουν, ο εκτιμητής συγκρίνει κάθε λαμβανόμενη τιμή με την τιμή DE που αυτός διατηρεί. Ανάλογα με την απόκλιση αυτών των δύο τιμών (άμεση τιμή και τιμή σύστασης) ο εκτιμητής αυξάνει ή μειώνει την τιμή του βαθμού αξιοπιστίας TV για την υπηρεσία σύστασης ορισμένων κόμβων.

5.3 Η λειτουργία των Αισθητήρων Αξιοπιστίας

Η πλειοψηφία των προτεινόμενων συστημάτων διαχειρίσεις φήμης θεωρεί ότι ο σημαντικότερος παράγοντας για την οικοδόμηση εμπιστοσύνης μεταξύ των κόμβων είναι τα άμεσα αποδεικτικά στοιχεία (direct evidence). Στο έργο SECURE [65], αυτός ο έλεγχος των άμεσων αποδεικτικών στοιχείων εκτελείται ανεξάρτητα από κάθε κόμβο

μέσω ενός συστήματος ελέγχου [66] το οποίο καταγράφει κάθε δραστηριότητα και την αποθηκεύει. Στην [61] προτείνεται ένας μηχανισμός ελέγχου που έχει σαν σκοπό την παρατήρηση της συμπεριφοράς των κόμβων του δικτύου ως προς την υπηρεσία της δρομολόγησης. Η αρχιτεκτονική *ATF* βασίζεται στους *Αισθητήρες Αξιοπιστίας (Trust Sensors-TS)* για να ανιχνεύσει τα άμεσα αποδεικτικά στοιχεία. Ένας αισθητήρας αξιοπιστίας *TS* λειτουργεί όπως οποιοσδήποτε κοινός αισθητήρας, δηλαδή μεταφράζει ένα (φυσικό) φαινόμενο σε μια μορφή ερμηνεύσιμη από τους υπολογιστές. Στην περίπτωση μας αυτό το φαινόμενο είναι η αξιοπιστία ενός κόμβου όσον αφορά την ποιότητα υπηρεσίας που αυτός προσφέρει. Ένας αισθητήρας *TS* παρακολουθεί τη συμπεριφορά ενός άλλου κόμβου και την συγκρίνει με μια προκαθορισμένη συμπεριφορά αναφοράς, (δηλαδή την αναμενόμενη λειτουργία). Υπό αυτήν την έννοια, το προτεινόμενο σχήμα *ATF* χρησιμοποιεί τους αισθητήρες αξιοπιστίας για να βοηθήσει έναν κόμβο ώστε να καθορίσει την αξιοπιστία των υπόλοιπων συνεργαζόμενων κόμβων. Ο προτεινόμενος μηχανισμός αφορά τα εξής:

- Ένα εννοιολογικό μοντέλο και καθορισμός της αναμενόμενης λειτουργίας ενός κόμβου. Αυτό το μοντέλο σχετίζεται με τις μεθόδους παρατήρησης που επιλέγονται (όπως αναφέρεται παρακάτω). Παραδείγματος χάριν, εάν η μέθοδος παρατήρησης είναι η ανάλυση προτύπων (*pattern recognition*), η αναμενόμενη λειτουργία του κόμβου θα εκφραζόταν με βάση τα αποδεκτά πρότυπα.
- *Οι μηχανισμοί παρατήρησης*. Περιλαμβάνουν την ανάλυση προτύπων στα ιστορικά δεδομένα, την παρακολούθηση των γειτονικών κόμβων σε σχέση με την προώθηση πακέτων, τους κωδικούς επιστροφής από την εκτέλεση απομακρυσμένων διαδικασιών, κ.α. Η μέθοδος παρατήρησης εξαρτάται από την υπηρεσία που θα ελέγχει ο αισθητήρας αξιοπιστίας *T*.
- *Η ποσοτικοποίηση της διαφοράς μεταξύ των παρατηρήσεων και της αναμενόμενης λειτουργίας*. Μία διαισθητική προσέγγιση σε αυτό το ζήτημα είναι η κατηγοριοποίηση των παρατηρήσεων σε *επιτυχίες* και *αποτυχίες* σχετικά με την αναμενόμενη λειτουργία [67].

5.4 Ο υπολογισμός της αξιοπιστίας

5.4.1 Η ποιοτική άποψη

Σε αυτή την ενότητα περιγράφουμε τις ιδιότητες του *Δημιουργού Αξιοπιστίας (Trust Builder)*. Το προτεινόμενο σχέδιο ενσωματώνει αρκετούς συντελεστές (βάρη) που

καθορίζονται από το χρήστη και εξαρτώνται από το χρόνο. Η χρονική εξάρτηση είναι σημαντικός παράγοντας, δεδομένου ότι επιτρέπει τη διαμόρφωση χρονικά μεταβαλλόμενων στρατηγικών που μπορούν να ακολουθηθούν από τους κόμβους. Επιπλέον, τα βάρη καθορίζονται χωριστά για κάθε κόμβο στην *Πολιτική Οικοδόμησης Εμπιστοσύνης* που ακολουθεί. Στην αρχιτεκτονική *ATF* ο χρόνος αντιμετωπίζεται ως μία διακριτή ακολουθία από άμεσες αλληλεπιδράσεις μεταξύ των κόμβων. Κατά συνέπεια, ο χρόνος για κάθε κόμβο παρέρχεται με διαφορετικό ρυθμό. Χρησιμοποιούμε μόνο τις άμεσες αλληλεπιδράσεις ως χρονική αναφορά, αφού αυτές θεωρούνται σημαντικότερες από τις έμμεσες (συστάσεις από άλλους κόμβους) για την διαδικασία οικοδόμησης εμπιστοσύνης. Επιπλέον, οι αλληλεπιδράσεις είναι ταξινομημένες σε θετικές και αρνητικές (σύμφωνα με το πρότυπο *επιτυχία/αποτυχία* που ενσωματώνεται από κάθε αισθητήρα *TS*).

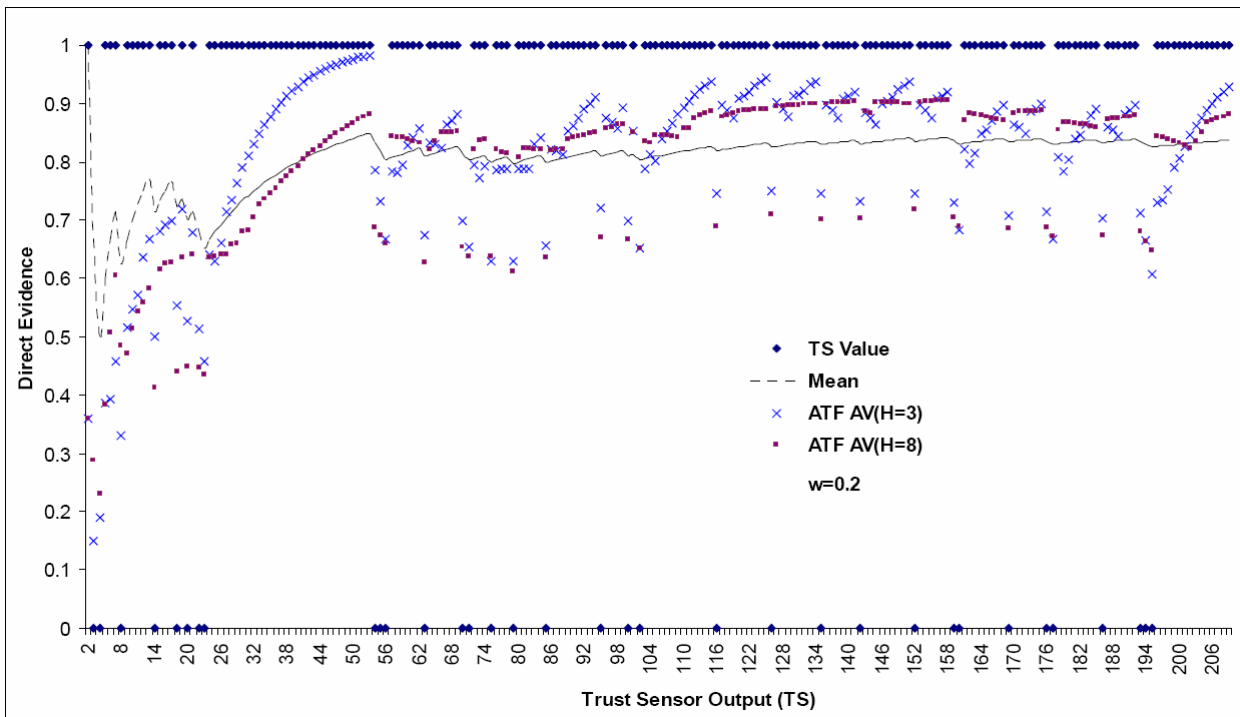
Η πλειοψηφία των αρχιτεκτονικών οικοδόμησης εμπιστοσύνης χρησιμοποιεί δύο κύρια συστατικά στην εκτίμηση της αξιοπιστίας των κόμβων: τα άμεσα αποδεικτικά στοιχεία (*DE*) και τις συστάσεις (*RECs*) από τρίτους. Η τιμή του *DE* υπολογίζεται από τα δεδομένα που στέλνουν οι αισθητήρες *TS*, όπως περιγράφεται στις προηγούμενες παραγράφους και είναι χρήσιμα για την αξιολόγηση των γειτονικών κόμβων. Οι συστάσεις (*RECs*) ανταλλάσσονται μεταξύ των κόμβων που συμμετέχουν σύμφωνα με ένα πρωτόκολλο διάδοσης φήμης, που εφαρμόζεται στον *Διαχειριστή Φήμης (RM)*.

Πολλές προσεγγίσεις για την διαδικασία οικοδόμησης εμπιστοσύνη [68] υπαγορεύουν ότι ο υπολογισμός της αξιοπιστίας πρέπει επίσης να περιλαμβάνει ένα υποκειμενικό συστατικό. Κάθε κόμβος έχει έναν μοναδικό, υποκειμενικό τρόπο να εμπιστεύεται τους άλλους. Στην προτεινόμενη αρχιτεκτονική υιοθετούμε αυτήν την προσέγγιση και ένας *Υποκειμενικός Παράγοντας (SUB)* εισάγεται στο μοντέλο του υπολογισμού της αξιοπιστίας. Αυτό το συστατικό είναι χρονικά εξαρτώμενο ώστε να επιτραπεί η αλλαγή της συμπεριφοράς των κόμβων (π.χ., ένας κόμβος μπορεί να εμπιστεύεται έναν νεοφερμένο μέχρι ένα σημείο, έως ότου οι δύο κόμβοι καθιερώσουν έναν συγκεκριμένο αριθμό επιτυχών αλληλεπιδράσεων). Ο παράγοντας *SUB* υπαγορεύεται από την *Πολιτική Οικοδόμησης Εμπιστοσύνης (TP)* κάθε κόμβου και διαμορφώνει χαρακτήρες εμπιστοσύνης όπως απρόσεκτος, ύποπτος, δυσκολόπιστος, κ.λπ. Αυτό το συστατικό παρέχει ευελιξία στην εφαρμογή της στρατηγικής που ακολουθείται χωρίς να εισάγει σημαντική πολυπλοκότητα στον υπολογισμό της αξιοπιστίας των κόμβων.

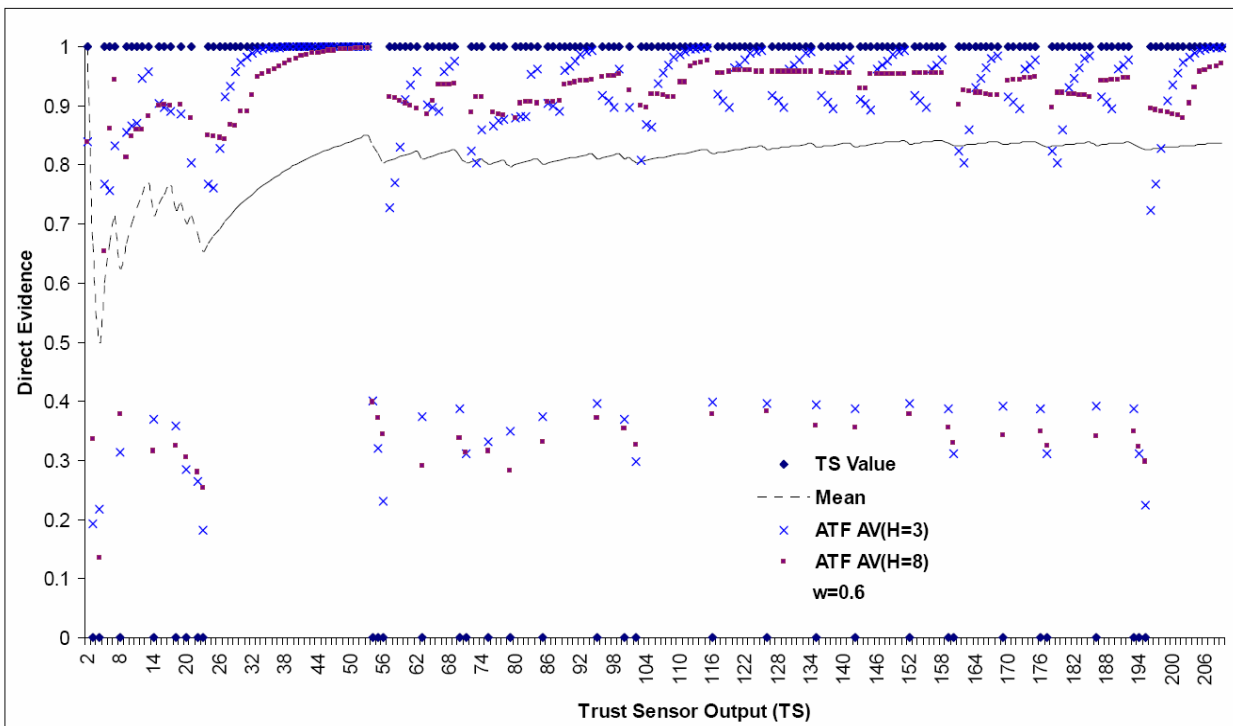
5.4.2 Τα ιστορικά δεδομένα

Τα ιστορικά δεδομένα είναι μια πρόσθετη έννοια που παίζει σημαντικό ρόλο στα συστήματα οικοδόμησης εμπιστοσύνης. Διάφοροι ερευνητές χρησιμοποιούν αυτά τα δεδομένα ως βασικό συστατικό στον υπολογισμό της αξιοπιστίας των κόμβων. Μερικοί ορίζουν συγκεκριμένο συντελεστή (βάρος) στις άμεσες παρατηρήσεις του παρελθόντος ή στις συστάσεις προκειμένου να παρέχουν μία ομαλή διακύμανση στο βαθμό αξιοπιστίας TV [63]. Το σύστημα CORE χρησιμοποιεί μία χρονικά μεταβαλλόμενη συνάρτηση που δίνει υψηλότερη σχετικότητα στις προηγούμενες παρατηρήσεις για να αξιολογήσει την αξιοπιστία της υπηρεσίας ενός κόμβου στόχου [69]. Αυτή η συνάρτηση λαμβάνει υπόψη μόνο τις πρόσφατες παρατηρήσεις, και αντιπροσωπεύεται μέσω ενός *Φίλτρου Πεπερασμένης Κρουστικής Απόκρισης (FIR)*.

Οι συγγραφείς στην [69] θεωρούν ότι αν δοθεί περισσότερο βάρος στο απώτερο παρελθόν τότε οι σποραδικές συμπεριφορές των κόμβων στο πρόσφατο παρελθόν δεν θα επηρεάσουν την αξιολόγηση της αξιοπιστίας. Στην [70], μια διαφορετική προσέγγιση χρησιμοποιείται. Οι παλαιότερες παρατηρήσεις γίνονται λιγότερο σημαντικές από τις νεότερες και οι παρατηρήσεις που είναι παλιότερες από ένα συγκεκριμένο κατώφλι αγνοούνται. Στην [71] μια παρόμοια προσέγγιση ακολουθείται για να επιτραπεί η εξασθένηση της αξιοπιστίας. Μέσω μιας τροποποιημένης διαδικασίας δίνεται λιγότερο βάρος στις παρατηρήσεις του παρελθόντος. Στην αρχιτεκτονική ATF συνδυάζονται και οι δύο προσεγγίσεις. Πιστεύουμε ότι ένα σημαντικό βάρος πρέπει να δοθεί στην τρέχουσα παρατήρηση. Αυτό θα επιτρέψει σε ένα κόμβο εκτιμητή να προσδιορίσει γρήγορα την αλλαγή συμπεριφοράς (προς το χειρότερο) ενός κόμβου στόχου. Επιπλέον, το πλήθος των ιστορικών δεδομένων που θα εξεταστούν για την αξιολόγηση της αξιοπιστίας ενός κόμβου εξαρτάται από την υποκειμενικό παράγοντα του κόμβου εκτιμητή. Κατά συνέπεια ένας δύσπιστος κόμβος μπορεί να χρειαστεί τις παλιότερες προηγούμενες παρατηρήσεις, ενώ ένας ευκολόπιστος κόμβος να χρειαστεί μόνο τα νεότερα δεδομένα. Έτσι όταν εμφανίζεται μία νέα παρατήρηση (δηλ., $TS(n, f)$) ένας δύσπιστος κόμβος λαμβάνει υπόψη την ιστορία h_1 από τα άμεσα αποδεικτικά στοιχεία όπως φαίνεται στο Σχήμα 5.2 ενώ ο ευκολόπιστος κόμβος λαμβάνει υπόψη την ιστορία h_2 .



Σχήμα 5.3: Διακύμανση της τιμής DE για δύο διαφορετικά παράθυρα (H ίσο με 3 και 8) και $w=0.2$ ως συντελεστής της τρέχουσας παρατήρησης.



Σχήμα 5.4: Διακύμανση της τιμής DE για δύο διαφορετικά παράθυρα (H ίσο με 3 και 8) και $w=0.6$ ως συντελεστής της τρέχουσας παρατήρησης.

Κάθε σχήμα απεικονίζει τον αριθμητικό μέσο όρο ($\frac{1}{N} \sum_{i=1}^N TS$), τις απαντήσεις του αισθητήρα TS σε κάθε παρατήρηση (δηλ., 0 ή 1), και δύο διακριτές διαδικασίες $DE_{n,f}(t)$ που λαμβάνουν υπόψη διαφορετικό ιστορικό παράθυρο, δηλ., $h_1 = 3$ και $h_2 = 8$. Όπως παρατηρείται στα σχήματα, μέτριες τιμές του συντελεστή w (π.χ., 0.2) εξασφαλίζουν ότι οι νέες τιμές του DE που υπολογίζονται ανήκουν σε μια σειρά που δεν παρεκκλίνει από την αρχική μας υπόθεση (δηλ., ο κόμβος n είναι δίκαιος στην εκτέλεση της υπηρεσίας f). Κατά συνέπεια, λαμβάνοντας υπόψη μόνο τις τιμές μετά από μία αρχική περίοδο εκπαίδευσης (δηλ., τιμές μετά από την 30ή παρατήρηση), οι νέες τιμές του DE στο Σχήμα 5.3 ($w = 0.2$) δεν πέφτουν κάτω από 0.6, ενώ στο Σχήμα 5.4 ($w = 0.6$) οι τιμές του DE πλησιάζουν την τιμή 0.3. Έτσι, μέτριες τιμές του w στη σχέση (5.1) εξασφαλίζουν ότι τα άμεσα αποδεικτικά στοιχεία θα «συλλάβουν» την πραγματική συμπεριφορά του κόμβου στόχου, χωρίς σημαντικές αποκλίσεις λόγω των σποραδικών εγωιστικών συμπεριφορών ή των διάφορων λαθών. Επιπλέον, τα ιστορικά δεδομένα στη σχέση (5.1) εισάγουν διαφορετική διακύμανση στον υπολογισμό της τιμής DE . Όταν απαιτούνται μόνο οι πρόσφατες παρατηρήσεις (π.χ., $H=3$, ευκολόπιστος κόμβος), το DE προσεγγίζει γρηγορότερα το ελάχιστο ή το μέγιστο εισάγοντας έτσι ξαφνικές διακυμάνσεις. Αντιθέτως όταν ο κόμβος εκτιμητής λαμβάνει υπόψη μία πιο μακροχρόνια ιστορία (π.χ., $H=8$, δυσκολόπιστος κόμβος) το DE πλησιάζει αργότερα τις τιμές των κατωφλίων, εισάγοντας έτσι ομαλότερες διακυμάνσεις.

5.4.3 Η ποσοτική άποψη

Η ενότητα αυτή περιγράφει τους μαθηματικούς τύπους για το προτεινόμενο μοντέλο υπολογισμού αξιοπιστίας. Ο χρόνος αξιοπιστίας (*trust time*) μετρά τις άμεσα παρατηρήσιμες αλληλεπιδράσεις. Το προτεινόμενο πρότυπο ορίζει τον Θετικό Χρόνο (*Positive Time - PT*), τον Αρνητικό Χρόνο (*Negative Time - NT*) και το Συνολικό Χρόνο (*Total Time - TT*). Ο τελευταίος ορισμός προσθέτει απλά τα άλλα δύο χρονικά διαστήματα. Ελέγχουμε τη χρονική εξέλιξη κάθε υπηρεσίας όλων των κόμβων έτσι ώστε αυτά τα χρονικά διαστήματα να απεικονίζονται ως πίνακες μεγέθους $N \times F$, όπου το N είναι ο αριθμός των κόμβων στο δίκτυο, και το F αντιστοιχεί στο συνολικό αριθμό των υπηρεσιών που προσφέρονται. Δηλαδή έχουμε:

$$\begin{aligned} PT &\equiv (PT_{n,f}) \in N \\ NT &\equiv (NT_{n,f}) \in N \\ TT &\equiv (TT_{n,f}) \in N \end{aligned}$$

όπου $n = 1 \dots N$, $f = 1 \dots F$.

Κάθε κόμβος πρέπει να έχει τουλάχιστον ένα πίνακα χρόνου. Κάθε χρονική στιγμή όταν παρατηρείται μια αλληλεπίδραση (*επιτυχία*, ή *αποτυχία*) με την υπηρεσία ενός κόμβου το αντίστοιχο στοιχείο του χρονικού πίνακα αυξάνεται κατά μία μονάδα. Έτσι στο εξής το $T_{n,f}$ θα υποδεικνύει τον $PT_{n,f}$, $NT_{n,f}$ ή τον $TT_{n,f}$. Κάθε κόμβος ανιχνευτής διατηρεί ένα πίνακα $N \times F$ (*Trust Matrix-TM*) που αντιπροσωπεύει την τιμή της αξιοπιστίας TV που ο κόμβος υπολογίζει για κάθε υπηρεσία των κόμβων στόχων. Κάθε στοιχείο $TM_{n,f}$ ($1 \leq n \leq N$ και $1 \leq f \leq F$) αναφέρεται σε μια συγκεκριμένη υπηρεσία f ενός κόμβου n και αυτό μπορεί να αλλάξει με την πάροδο του χρόνου. Οι εξισώσεις για τον πίνακα TM και την τιμή TV είναι:

$$TV' \equiv TV'(n, f, t) = (a * DE_{n,f} + b * REC_{n,f}) * SUB_{n,f}(t)$$

$$\text{όπου } t = T_{n,f}$$

$$TV(n, f, t) = TV' * u(1 - TV') + u(TV' - 1)$$

$$TM \equiv (TM_{n,f})$$

$$TM_{n,f} = TV(n, f, t) \in [0,1]$$

Όπως αναφέρουμε και στις επόμενες παραγράφους ισχύει: $DE_{n,f} \in [0,1]$, $REC_{n,f} \in [0,1]$ και $SUB_{n,f}(t) \in [0,2]$. Έτσι προκύπτει ότι η τιμή $TV' \in [0,2]$. Οι παράμετροι a και b είναι συναρτήσεις βηματικής αύξησης ή μείωσης. Προκειμένου να αντιστοιχίσουμε τις τιμές TV στο διάστημα $[0,1]$ χρησιμοποιούμε τη συνάρτηση μοναδιαίου βήματος $u(t)$ της σχέσης (5.2) για την κανονικοποίηση της τιμής TV' στην τελική τιμή TV . Οι τιμές που μπορεί να λάβει η $TV(n,f,t)$ ανήκουν στο διάστημα $[0,1]$ όπου η τιμή 0 σημαίνει ότι ο κόμβος ανιχνευτής θεωρεί πλήρως αναξιόπιστο τον κόμβο n για την υπηρεσία f που αυτός προσφέρει και η τιμή 1 σημαίνει ότι τον εμπιστεύεται πλήρως.

$$u(t) = \begin{cases} 0, & t < 0 \\ 1/2, & t = 0 \\ 1, & t > 0 \end{cases} \quad (5.2)$$

Η $DE_{n,f}$ είναι η τιμή των άμεσων αποδεικτικών στοιχείων για ένα κόμβο n και της υπηρεσίας του f , όπως αυτός παρατηρείται από τον αντίστοιχο αισθητήρα αξιοπιστίας TS του κόμβου ανιχνευτή. Ο πίνακας $N \times F$ για την τιμή DE που αποθηκεύεται σε κάθε κόμβο ορίζεται ως $DE \equiv (DE_{n,f}) \in [0,1]$, όπου τα στοιχεία του πίνακα $DE_{n,f}$

υπολογίζονται σύμφωνα με την σχέση (5.1) (η παράμετρος *CurrentValue* ισούται με *TS*) ως εξής: $DE_{n,f}(t) = WA_H(n,f,t)$ και $TS(n,f) \in \{0,1\}$.

Η χαμηλότερη τιμή (δηλ. 0) δηλώνει *αποτυχία*, ενώ η υψηλότερη (δηλ. 1) δηλώνει *επιτυχία*. Ο ορισμός $REC_{n,f}$ υποδηλώνει τις συσσωρευμένες συστάσεις που έχουμε για την υπηρεσία *f* του κόμβου *n* από τρίτους κόμβους. Αυτές οι συστάσεις είναι είτε άμεσα αποδεικτικά στοιχεία *DEs* είτε συστάσεις *RECs* από τους τρίτους κόμβους. Θα μπορούσε να ειπωθεί ότι επιτρέποντας σε έναν τρίτο να αποστείλει συστάσεις που λαμβάνει από άλλους (συστάσεις δεύτερου επιπέδου), οδηγούμαστε σε συνθήκες διάδοσης μη επαρκούς πληροφορίας (φημών). Εντούτοις, αυτό είναι απαραίτητο σε περιπτώσεις που δεν έχουμε άλλα στοιχεία για την αξιοπιστία ενός κόμβου. Οι συστάσεις αυτές αποθηκεύονται. Κατά συνέπεια, κάθε κόμβος διατηρεί έναν πίνακα $N \times F$ που ορίζεται ως $REC \equiv (REC_{n,f}) \in [0,1]$, όπου τα στοιχεία του πίνακα $REC_{n,f}$ υπολογίζονται σύμφωνα με την σχέση (5.3).

$$REC_{n,f}(t) = WA_H(n, f, t) \quad (5.3)$$

Ο υποκειμενικός παράγοντας *SUB* στον τύπο υπολογισμού της τιμής *TV* ενσωματώνει την υποκειμενικότητα του κόμβου. Αυτή η υποκειμενικότητα είναι βασικός διαφοροποιητής μεταξύ των κόμβων [68]. Ο παράγοντας *SUB* είναι ένας πίνακας $N \times F$ με στοιχεία στο πεδίο $\{f : T \rightarrow [0,2]\}$. Κατά συνέπεια, τα στοιχεία του είναι χρονικές συναρτήσεις. Το διάστημα $[0,2]$ επιτρέπει στον ανιχνευτή να μη εμπιστεύεται καθόλου τον κόμβο στόχο (δηλ. τιμή ίση με 0), να τον εμπιστεύεται (δηλ. τιμή ίση με 1), να είναι «ενθουσιώδης» για τον κόμβο στόχο (τιμή ίση με 2) ή να αναπτύξει οποιαδήποτε άλλη ενδιάμεση μορφή υποκειμενική στρατηγικής. Επιλέξαμε την τιμή 2 ως άνω όριο για να επιτραπεί η ύπαρξη «ενθουσιωδών» κόμβων αλλά όχι σε τέτοιο βαθμό που θα διακινδύνευε η ορθή λειτουργία του δικτύου. Με άλλα λόγια επιθυμούμε να έχουμε κόμβους με διαφορετικές στρατηγικές οικοδόμησης εμπιστοσύνης, αλλά θέλουμε ταυτόχρονα να περιορίσουμε την διακύμανση αυτής της ποικιλομορφίας. Μία χρονική συνάρτηση υποκειμενικής στρατηγικής θα μπορούσε να οριστεί ως εξής:

$$SUB_{n,f}(t) = u(t - 2), t = PT_{n,f}$$

Η συνάρτηση αυτή δείχνει πως ανεξάρτητα από τις άμεσες παρατηρήσεις *DEs* ή τις συστάσεις *RECs* που έχει ένας κόμβος για μια συνάρτηση ενός κόμβου στόχου (*n,f*), δεν θα εμπιστευθεί τον τελευταίο αν δεν έχει δύο επιτυχείς (θετικές) άμεσες αλληλεπιδράσεις. Στην περίπτωση που ο προαναφερθέν παράγοντας *SUB* χρησιμοποιείται αδιακρίτως για όλους τους κόμβους στόχους και όλες τις παρεχόμενες

υπηρεσίες τότε θα είναι ίδιος για όλα τα στοιχεία του πίνακα SUB του κόμβου ανιχνευτή. Το σύνολο των χρονικών συναρτήσεων SUB καθορίζεται στην πολιτική που ακολουθεί κάθε κόμβος και μπορεί να ρυθμιστεί ανάλογα με το είδος της υπηρεσίας. Εντούτοις, στην πράξη, είναι ιδιαίτερα απίθανο ότι ένας κόμβος ανιχνευτής να έχει $N \times F$ διαφορετικές χρονικές συναρτήσεις. Αντί αυτού, ένας ανιχνευτής συνήθως χρησιμοποιεί την ίδια υποκειμενική στρατηγική για κάθε στόχο ή κάθε υπηρεσία.

5.5 Πολιτική Οικοδόμησης Εμπιστοσύνης

Όπως ήδη αναφέρθηκε, για κάθε κόμβο ορίζεται μια *Πολιτική Οικοδόμησης Εμπιστοσύνης (Trust Policy-TP)*. Αυτή η πολιτική παρέχει ένα σύνολο παραμέτρων που καθορίζουν πλήρως τη λειτουργία του Διαχειριστή Φήμης (*Reputation Manager*) και του Δημιουργού Αξιοπιστίας (*Trust Builder*) και παρουσιάζονται στον Πίνακα 5-1. Η πολιτική ανιχνεύει την υποκειμενική συμπεριφορά της οντότητας (π.χ., του τελικού χρήστη). Ο Πίνακας 5-1 περιγράφει επίσης τη σημασιολογία τους (σειρά, συνθήκες που πρέπει να ισχύουν, ειδικές τιμές, κ.λπ.). Η παράμετρος HFI χρησιμοποιείται από τον *Διαχειριστή Φήμης* και η εφαρμογή της περιγράφεται στην επόμενη ενότητα. Άλλοι παράγοντες και παράμετροι μπορούν να περιληφθούν στις προδιαγραφές, όπως η αρχική τιμή της DE και των συστάσεων REC για του νεοεισερχόμενους κόμβους.

Πίνακας 5-1: Οι παράμετροι της Πολιτικής Οικοδόμησης Εμπιστοσύνης.

| Παράμετρος | Σημασιολογία |
|-----------------------------|--|
| MI | Οι ελάχιστες άμεσες αλληλεπιδράσεις που απαιτούνται για να υπάρξει βεβαιότητα για την αξιοπιστία TV ενός κόμβου στόχου. |
| a | Η επίδραση (βάρος) των άμεσων αποδεικτικών στοιχείων DE στον υπολογισμό της αξιοπιστίας TV ($0 \leq a \leq 1, a+b=1$). Αποτελεί αύξουσα βηματική συνάρτηση στο MI . |
| b | Η επίδραση (βάρος) των συστάσεων REC στον υπολογισμό της αξιοπιστίας TV ($0 \leq b \leq 1, a+b=1$). Αποτελεί φθίνουσα βηματική συνάρτηση στο MI . |
| w | Η επίδραση των ιστορικών δεδομένων στον υπολογισμό της DE (ή των συστάσεων REC) ($0 \leq w \leq 1$). |
| $TV_{RF} \text{ threshold}$ | Η ελάχιστη επιτρεπόμενη αξιοπιστία TV_{RF} που πρέπει να έχει ένας κόμβος προκειμένου να γίνει συστήνων (recommender). |
| $SUB_{n,f}(t)$ | Ένα σύνολο χρονικών συναρτήσεων που καθορίζουν την υποκειμενική στρατηγική ενός κόμβου. Η τιμή $SUB(0)$ είναι η στρατηγική για τους νεοεισαχθέντες δηλ., ορίζει μια αρχική TV σε έναν νεοεισαχθέντα. |
| HFI | Honorable Friend Index. Ο ελάχιστος αριθμός αξιόπιστων συστηγόντων (trusted recommenders) που απαιτείται για να ερωτηθούν από έναν αιτούντα ώστε να αξιολογηθεί επαρκώς η αξιοπιστία ενός άγνωστου κόμβου. |

Οι παράμετροι a και b είναι βηματικές αύξουσες ή φθίνουσες συναρτήσεις. Αυτή η πολιτική επιλέχτηκε διότι όταν ένας κόμβος ανιχνευτής αντιλαμβάνεται την ύπαρξη ενός νεοεισαχθέντα στο δίκτυο, τότε μόνο οι συστάσεις από αξιόπιστους κόμβους πρέπει να χρησιμοποιηθούν (υψηλές τιμές για την παράμετρο b). Κατά συνέπεια, στην αρχική φάση οι συστάσεις *RECs* είναι ουσιαστικές. Με την πάροδο του χρόνου όταν ο ανιχνευτής αρχίζει να αλληλεπιδρά με το νέο κόμβο τα άμεσα στοιχεία (*DE*) γίνονται σημαντικότερα, και αυτό συμβαίνει μόνο όταν η τιμή *MI* ξεπερνιέται.

5.6 Επιλογή των Αξιόπιστων Συστηνόντων Κόμβων

Όπως απεικονίζεται στη σχέση (5.3) ένας αιτών κόμβος βασίζεται στις συστάσεις για να υπολογίσει την αρχική τιμή αξιοπιστίας *TV* για έναν νεοεισαχθέντα κόμβο στο δίκτυο. Εντούτοις, μόνο οι αξιόπιστοι κόμβοι (δηλαδή οι κόμβοι που παρουσιάζουν μια υψηλή τιμή *TV* για την υπηρεσία σύστασης) πρέπει να προβούν σε αυτές τις συστάσεις. Αυτό θα ελαχιστοποιήσει τα αρνητικά αποτελέσματα της διάδοσης φήμης και θα αποφευχθούν επιθέσεις άρνησης της υπηρεσίας *DoS* [64]. Επιπλέον, οι επιλεγμένοι συστήνοντες θα πρέπει να είναι όσο το δυνατόν πιο κοντά στον αιτούντα ώστε να περιοριστεί το κόστος της επικοινωνίας. Οι Buchegger και Le Boudec [62] πρότειναν τη χρήση ενός διαχειριστή μονοπατιών (*path manager*) που ταξινομεί μονοπάτια σύμφωνα με τις μετρικές ασφάλειας (π.χ. φήμη των κόμβων σε ένα μονοπάτι) και διαγράφει μονοπάτια που περιέχουν εγωιστικούς ή κακόβουλους κόμβους. Μία αντίστοιχη αρχιτεκτονική (*Pathrater*) προτείνεται στην [61] για να μετριάσει τις επιπτώσεις από μία δρομολόγηση που δεν λειτουργεί σωστά. Ο αλγόριθμος *Pathrater* συνεπικουρείται από έναν μηχανισμό που προσδιορίζει την εγωιστική συμπεριφορά των κόμβων οι οποίοι δεν προωθούν τα πακέτα. Ο κόμβος στον οποίο υλοποιείται ο αλγόριθμος χρησιμοποιεί αυτήν την γνώση για να επιλέξει το κατάλληλο μονοπάτι ώστε να παραδοθούν τα πακέτα με μεγάλη πιθανότητα.

Στα πλαίσια της παρούσας διατριβής αναπτύχθηκε μία καινοτόμος μέθοδος που ονομάζεται *TrustSpan*. Ο μηχανισμός αυτός χρησιμοποιείται από τον *Διαχειριστή Φήμης (RM)*, και επιτρέπει στον αιτούντα κόμβο να συμβουλευτείται μόνο αξιόπιστους όταν απαιτούνται συστάσεις για την αξιολόγηση των νέων κόμβων στο δίκτυο. Θεωρούμε ως νεοεισαχθέντα κόμβο αυτόν για τον οποίο ο αιτών δεν έχει κανένα στοιχείο ως προς την αξιοπιστία του ή έχει ανεπαρκή αριθμό αλληλεπιδράσεων (δηλ., πολύ λίγες προηγούμενες άμεσες αλληλεπιδράσεις). Όταν ο νέος κόμβος γίνεται γείτονας σε έναν αιτούντα, ο τελευταίος δεν γνωρίζει εκ των προτέρων (*a priori*) εάν είναι αξιόπιστος και

επομένως εκτελεί αυτόν τον αλγόριθμο προκειμένου να συλλέξει συστάσεις από αξιόπιστους συστήνοντες. Ένας αιτών χαρακτηρίζει έναν κόμβο ως *αξιόπιστο συστήνοντα* (*trusted recommender*) εάν η τιμή TV για την υπηρεσία σύστασής του είναι υψηλή. Υπενθυμίζουμε εδώ ότι η υπηρεσία σύστασης (RF) ελέγχεται και αξιολογείται ακριβώς όπως οποιαδήποτε άλλη υπηρεσία μέσω του αισθητήρα $RFTS$.

Ο αλγόριθμος *TrustSpan* παρουσιάζεται στο Σχήμα 5.5. Η παράμετρος HFI , η TV_{RF} και τα βάρη στη γραμμή 7 καθορίζονται μέσω της πολιτικής που ακολουθεί κάθε κόμβος (TP). Ο πίνακας απόστασης (DM) περιλαμβάνει τις αποστάσεις άλλων κόμβων. Η HFI εκφράζει τον ελάχιστο αριθμό αξιόπιστων συστηνότων (*trusted recommenders*) που απαιτείται να ερωτηθούν από έναν αιτούντα ώστε να αξιολογηθεί επαρκώς η αξιοπιστία ενός άγνωστου κόμβου στόχου. Η παράμετρος αυτή είναι ανεξάρτητη από τον κόμβο στόχο ή τις υπηρεσίες του.

```

1: input: HFI, DM, TM,  $TV_{RF}$  threshold
2: output: IDs of trusted recommenders
3: procedure TrustSpan () {
4:   for (j=0; j<numberOfNodes; j++)
5:      $TV_{RF}[j] = TM[j][RF]$ ;
6:   for (i=0; i<length( $TV_{RF}$ ); i++)
7:      $Candidate\_Recommenders[i] = 0.75 * TV_{RF}[i] + 0.2 * (1/DM[i])$ ;
8:   int trusted_IDS[] = new int[HFI];
9:   for (k=0; k<HFI; k++) {
10:    if ( $\max(Candidate\_Recommenders) \geq TV_{RF\_threshold}$ ) {
11:      int maxID =  $\text{indexOfmax}(Candidate\_Recommenders)$ ;
12:       $trusted\_IDs[k] = maxID$ ;
13:       $Candidate\_Recommenders[maxID] = 0$ ;
14:    } else break;
15:  }
16:   return trusted_IDS;
17: }
```

Σχήμα 5.5: Ο αλγόριθμος TrustSpan.

Σύμφωνα με τον αλγόριθμο *TrustSpan*, κάθε φορά που ένας αιτών κόμβος ρωτά για συστάσεις ως προς μια υπηρεσία f του κόμβου n , εξετάζει μόνο τους αξιόπιστους κόμβους. Κατά συνέπεια, προσπαθεί να ελαχιστοποιήσει και την καθυστέρηση στη μετάδοση των ζητούμενων συστάσεων αλλά και το γενικό κόστος επικοινωνίας. Ο αλγόριθμος *TrustSpan* εκτελείται μόνο όταν ένας ανιχνευτής έχει ανεπαρκή αριθμό

άμεσων αλληλεπιδράσεων με ένα κόμβο στόχο (νεοεισαχθέντα). Αυτός ο αριθμός καθορίζεται στην πολιτική (*TP*) που ακολουθεί κάθε κόμβος μέσω της παραμέτρου *MI*. Η παράμετρος *MI* επηρεάζει επίσης το ποιοι κόμβοι θα απαντήσουν στα αιτήματα σύστασης. Ειδικότερα, οι συστήνοντες κόμβοι που επιλέγονται από το αλγόριθμο *TrustSpan* και έχουν λιγότερες από *MI* άμεσες αλληλεπιδράσεις με έναν κόμβο στόχο, δεν θα δώσουν συστάσεις για αυτόν.

Όταν ο αλγόριθμος *TrustSpan* έχει επιστρέψει τα αναγνωριστικά των αξιόπιστων συστηνόντων (trusted recommenders), ο *Διαχειριστής Φήμης (RM)* ζητά τις συστάσεις (απευθείας) από τον αντίστοιχο διαχειριστή φήμης των συστηνόντων. Ένα αίτημα περιλαμβάνει το αναγνωριστικό (ταυτότητα) του κόμβου στόχου και τη συγκεκριμένη υπηρεσία την οποία ο αιτών θέλει να αξιολογήσει. Όταν οι απαραίτητες συστάσεις αφιχθούν ή κάποια χρονική περίοδος περάσει, ο *RM* του αιτούντος διαβιβάζει τις εισερχόμενες πληροφορίες στον *Δημιουργό Αξιοπιστίας (TB)* όπου εκτελείται ο υπολογισμός της αξιοπιστίας του κόμβου *n* για τη συνάρτηση *f*.

5.7 Πειραματική αξιολόγηση της αρχιτεκτονικής ATF

5.7.1 Περιβάλλον προσομοίωσης

Αξιολογήσαμε την αρχιτεκτονική *ATF* χρησιμοποιώντας τον προσομοιωτή J-SIM [72]. Χρησιμοποιήσαμε το 802.11 πρωτόκολλο MAC και πηγές σταθερού ρυθμού (Constant Bit Rate-CBR) για να παράγουμε δεδομένα. Το πρωτόκολλο δρομολόγησης IP που χρησιμοποιήθηκε ήταν το AODV (Ad hoc On-Demand Distance Vector). Η διαμόρφωση του πεδίου ήταν 50 κόμβοι σε μία έκταση 300m X 300m. Η εμβέλεια της ασύρματης μετάδοσης κάθε κόμβου τέθηκε ίση με 50m. Αρχικά, οι κόμβοι κατανεμήθηκαν τυχαία στην περιοχή. Η κινητικότητα των κόμβων προσομοιώθηκε σύμφωνα με το μοντέλο Random Waypoint στο οποίο κάθε κόμβος κινείται σε μια τυχαία επιλεγμένη θέση με μια ταχύτητα (π.χ 2 m/s) και έπειτα ακινητοποιείται για μικρό χρόνο (π.χ.5 sec), πριν επιλέξει τυχαία άλλη θέση. Κατά τη διάρκεια των προσομοιώσεων, το προτεινόμενο πλαίσιο *ATF* υπολογίζει την *τιμή αξιοπιστίας (TV)* για την υπηρεσία της *προώθησης πακέτων*. Κατά συνέπεια, οι τιμές που αφορούν τις άμεσες αλληλεπιδράσεις *DEs* και τις συστάσεις *RECs* χρησιμοποιήθηκαν ή ανταλλάχθηκαν ως προς την υπηρεσία *προώθησης πακέτων* για κάθε έναν από τους 50 κόμβους στην περιοχή. Ο αριθμός των εγwisτικών κόμβων τέθηκε ίσος με 10, και κάθε ένας από αυτούς ρυθμίστηκε ώστε να απορρίπτει ένα ποσοστό (30%) των πακέτων που λαμβάνει αλλά δεν προωθεί περαιτέρω. Για την παραγωγή δεδομένων (πακέτων) χρησιμοποιήσαμε 20 CBR

συνόδου. Κάθε σύνοδος παράγει πακέτα με ρυθμό 4 πακέτα/sec Η διάρκεια κάθε συνόδου είναι ίση με διάρκεια της προσομοίωσης. Ο Πίνακας 5-2 συνοψίζει τις παραμέτρους του περιβάλλοντος προσομοίωσης.

Πίνακας 5-2:Οι παράμετροι της προσομοίωσης.

| | |
|------------------------------|-------------|
| Αριθμός κόμβων | 50 |
| Αριθμός εγωιστικών κόμβων | 10 |
| Μέγιστη ταχύτητα | 2 m/sec |
| Χρόνος παύσης | 5 sec |
| Διαστάσεις περιοχής | 300m x 300m |
| Είδος επικοινωνίας | CBR |
| Ζεύγη επικοινωνίας | 20 |
| Ρυθμός επικοινωνίας | 4 pkts/sec |
| Εμβέλεια ασύρματης μετάδοσης | 30 m |

Κατά τη διάρκεια των προσομοιώσεων μετρήσαμε:

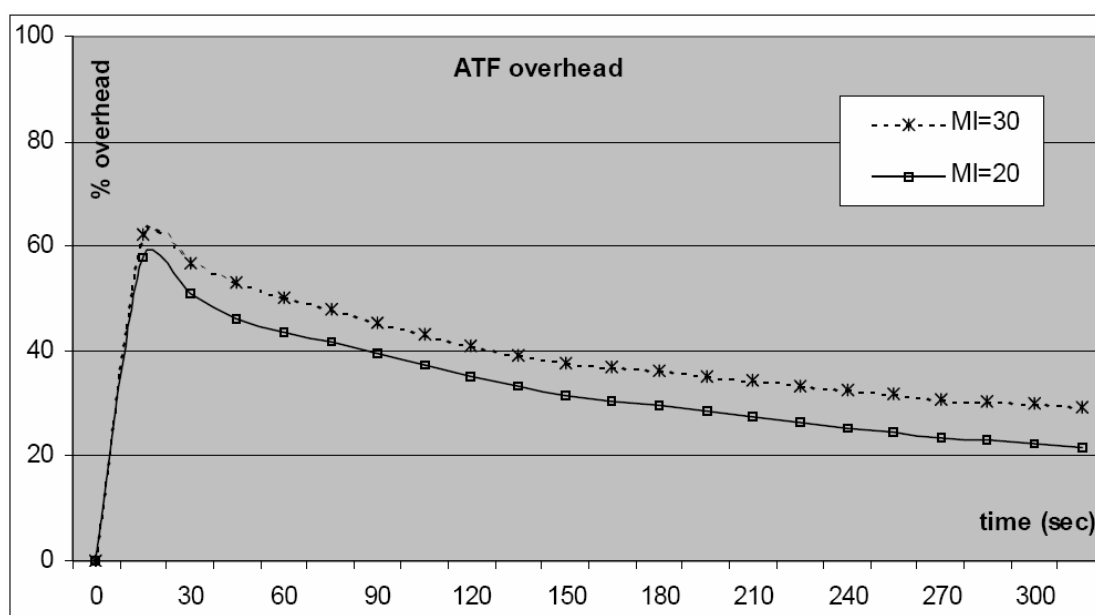
- Το κόστος της επικοινωνία που εισάγεται από τις αιτήσεις σύστασης και τις αντίστοιχες απαντήσεις.
- Το χρόνο που απαιτείται από κάθε κόμβο για να προσδιορίσει την πραγματική συμπεριφορά ενός άλλου κόμβου (δηλ., πόσο γρήγορα το προτεινόμενο πλαίσιο *ATF* δίνει τη δυνατότητα στους κόμβους να προσδιορίσουν την πραγματική τιμή αξιοπιστίας *TV* που αντιστοιχεί στην υπηρεσία ενός εγωιστικού κόμβου).
- Την ακρίβεια του *ATF*, δηλ., τη δυνατότητα των κόμβων να προσδιορίσουν σωστά τη συμπεριφορά των άλλων.

Ακόμα κι αν οι παράμετροι a και b (βλ. τους τύπους για τον υπολογισμό της *TV*) είναι αύξουσες και φθίνουσες αντίστοιχα συναρτήσεις στο *MI*, στις προσομοιώσεις χρησιμοποιήσαμε προκαθορισμένες σταθερές τιμές.

5.7.2 Αποτελέσματα προσομοιώσεων

Το επιπλέον κόστος επικοινωνίας (φόρτος-overhead) που εισάγει το προτεινόμενο πλαίσιο *ATF* λόγω των συστάσεων που ανταλλάσσονται απεικονίζεται στο Σχήμα 5.6. Για αυτό το σενάριο οι τιμές των παραμέτρων είναι οι εξής: $H = 4$, $w = 0,8$, $HFI = 3$, και $a = 0.9$. Ο φόρτος μετριέται για δύο διαφορετικούς δείκτες *MI* (20 και 30 άμεσες

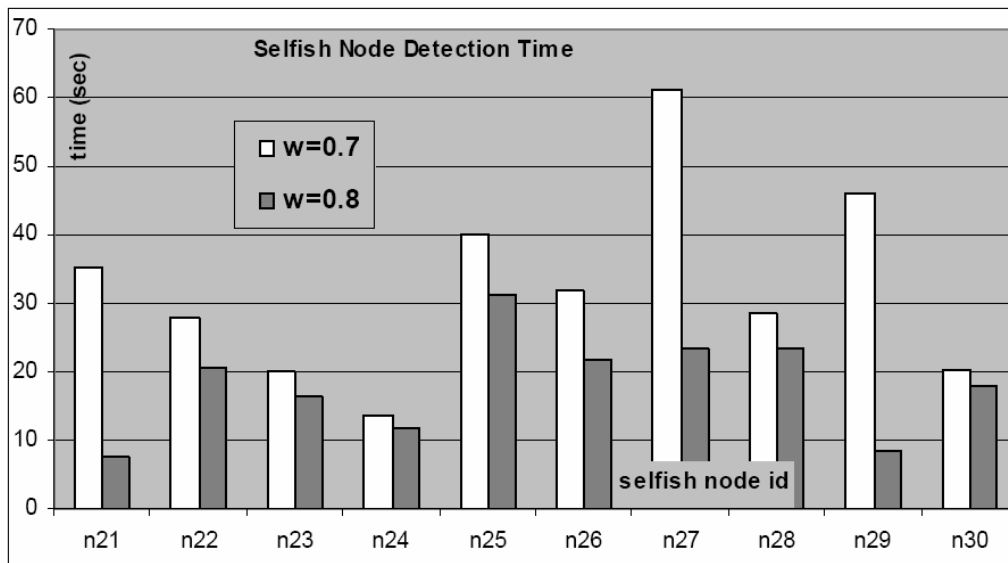
αλληλεπιδράσεις). Όπως φαίνεται στο Σχήμα 5.6, ο φόρτος επικοινωνίας μειώνεται με το χρόνο, και φθάνει τελικά στο 20% της πραγματικής κυκλοφορίας των δεδομένων στο δίκτυο (IP κυκλοφορία). Αυτό συμβαίνει επειδή οι άμεσες αλληλεπιδράσεις μεταξύ των κόμβων αυξάνονται με το χρόνο, και έτσι λιγότερες συστάσεις απαιτούνται από τους συστήνοντες κόμβους. Επιπλέον, για υψηλότερες τιμές του MI το επιπλέον κόστος επικοινωνίας είναι υψηλότερο, καθώς για υψηλότερο MI οι ελάχιστες άμεσες αλληλεπιδράσεις που απαιτούνται για την βεβαιότητα της αξιοπιστίας TV ενός κόμβου στόχου είναι υψηλότερη και κατά συνέπεια απαιτούνται περισσότερες συστάσεις.



Σχήμα 5.6: Το επιπλέον κόστος επικοινωνίας που εισάγει το πλαίσιο ATF λόγω των συστάσεων που ανταλλάσσονται.

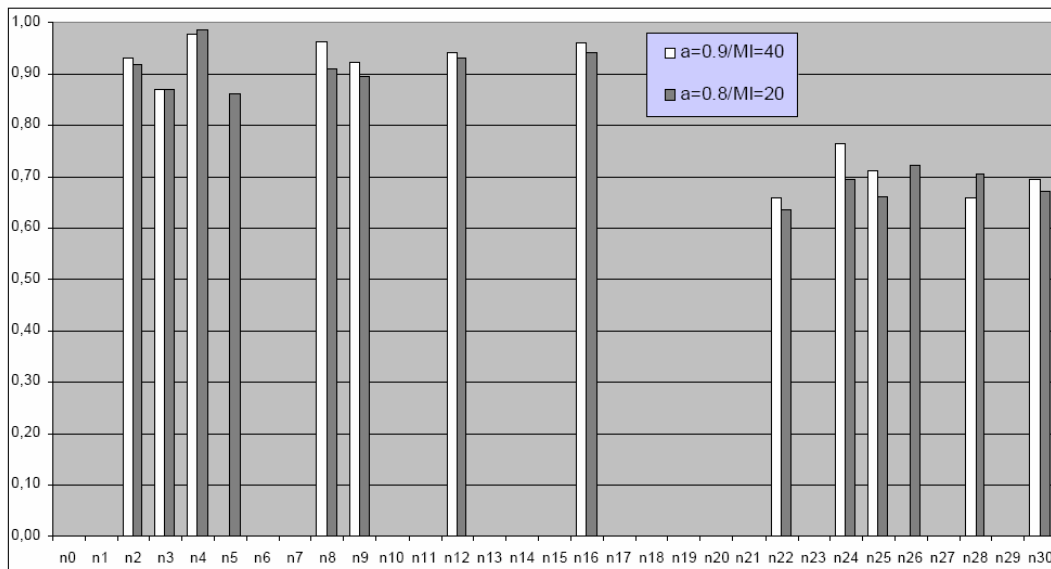
Στο Σχήμα 5.7 απεικονίζεται η ταχύτητα σύγκλισης του πλαισίου ATF , δηλ., ο χρόνος που απαιτείται για τους «δίκαιους» κόμβους (40 κόμβους) να ανιχνεύσουν έναν εγωιστικό κόμβο (10 αναξιόπιστοι κόμβοι). Αυτός είναι ο μέσος χρόνος που απαιτείται για έναν κόμβο να υπολογίσει την τιμή TV για την υπηρεσία προώθησης πακέτων των εγωιστικών κόμβων ($n_{21}, n_{22}, \dots, n_{30}$) η οποία είναι ίση με 0.7 δεδομένου ότι υποθέσαμε ότι κάθε ένας εγωιστικός κόμβος απορρίπτει το 30% των λαμβανόμενων πακέτων. Στη χειρότερη περίπτωση απαιτείται χρόνος 60 δευτερόλεπτων για την ανίχνευση ενός κόμβου που συμπεριφέρεται εγωιστικά. Αυτή η χειρότερη περίπτωση εμφανίζεται όταν οι εγωιστικοί κόμβοι ($n_{21}, n_{22}, \dots, n_{30}$) δεν παράγουν αρκετές άμεσες αλληλεπιδράσεις με τους «δίκαιους» κόμβους, λόγω του τρόπου κινητικότητάς τους. Κατά συνέπεια, ακόμα κι αν η παράμετρος MI τέθηκε σε 20 αλληλεπιδράσεις, οι εγωιστικοί κόμβοι n_{27}, n_{29} και n_{25} δεν παρήγαγαν μεγάλο αριθμό άμεσων αλληλεπιδράσεων κατά τη διάρκεια της

αρχικής περιόδου προσομοίωσης. Όπως φαίνεται στο Σχήμα 5.7 οι υψηλότερες τιμές της παραμέτρου w επιτρέπουν τη γρηγορότερη σύγκλιση στις πραγματικές τιμές αξιοπιστίας (δηλ., 0.7). Αυτό συμβαίνει επειδή οι χαμηλές τιμές του w , σχέση (5.1), εξασφαλίζουν ότι οι άμεσες αλληλεπιδράσεις θα ανιχνεύσουν την πραγματική συμπεριφορά του κόμβου στόχου, χωρίς σημαντικές αποκλίσεις λόγω σποραδικών εγωιστικών συμπεριφορών ή λαθών. Για τα αποτελέσματα στο Σχήμα 5.7 ισχύουν τα εξής: $H=4$, $MI=20$, $HFI=3$, και $a=0,9$.



Σχήμα 5.7: Ταχύτητα σύγκλισης του πλαισίου *ATF*.

Το Σχήμα 5.8 απεικονίζει την ακρίβεια του πλαισίου *ATF*, δηλαδή τη μέση τιμή αξιοπιστίας TV για την υπηρεσία της προώθησης πακέτων όπως μετρήθηκε στο τέλος της προσομοίωσης, και μόνο από τους δίκαιους κόμβους που έχουν τουλάχιστον MI αλληλεπιδράσεις με τους καθένα από τους άλλους κόμβους. Εδώ χρησιμοποιήσαμε $H=4$, $MI=\{20,40\}$, $HFI=3$, $w=0.3$, και $a=\{0.9, 0.8\}$. Για τα αποτελέσματα που απεικονίζονται στο Σχήμα 5.8 υπολογίσαμε κατά μέσο όρο τις τιμές αξιοπιστίας TV που τρία ζεύγη κόμβων πηγής-προορισμού υπολογίζουν για την υπηρεσία προώθησης πακέτων των 30 κόμβων. Από αυτούς τους 30 κόμβους, οι πρώτοι 20 ρυθμίστηκαν έτσι ώστε να είναι αξιόπιστοι και οι τελευταίοι 10 (δηλ., $n_{21}, n_{22}, \dots, n_{30}$) ρυθμίστηκαν να είναι εγωιστικοί. Αυτοί οι εγωιστικοί κόμβοι απορρίπτουν το 30% από τα πακέτα δεδομένων που λαμβάνουν. Εάν ένας από τους κόμβους εκτιμητές (μια πηγή ή ένας προορισμός των τριών ζευγών) δεν έχει τουλάχιστον MI αλληλεπιδράσεις με κάποιον από τους από τους 30 κόμβους η τιμή TV δεν υπολογίζεται. Κατά συνέπεια, για τον κόμβο n_0 δεν εξάγαμε κάποιο αποτέλεσμα ενώ για τον κόμβο n_5 έχουμε εκτίμηση της αξιοπιστίας του TV μόνο στην περίπτωση των $MI = 20$.



Σχήμα 5.8: Ακρίβεια του πλαισίου ATF.

Το Σχήμα 5.8 δείχνει ότι το πλαίσιο ATF είναι ακριβές, δεδομένου ότι καταφέρνει να προβλέψει την πραγματική συμπεριφορά των κόμβων. Για την υπηρεσία προώθησης πακέτων των δίκαιων κόμβων (n_2 , n_3 , n_4 , n_5 , n_8 , n_9 , n_{12} , και n_{16}) όταν υπήρχε ικανοποιητικός αριθμός αλληλεπιδράσεων η τιμή TV υπολογίστηκε περίπου ίση με ένα που είναι η πραγματική τιμή. Για την υπηρεσία προώθησης πακέτων των εγωιστικών κόμβων που υπήρχαν αρκετές αλληλεπιδράσεις (δηλ., οι κόμβοι n_{22} , n_{24} , n_{25} , n_{26} , n_{28} , n_{30}) η τιμή TV υπολογίστηκε κοντά στο 0.7, μια εκτίμηση που προσεγγίζει την πραγματική συμπεριφορά (δεδομένου ότι κάθε εγωιστικός κόμβος απέρριψε το 30% των πακέτων).

5.8 Συμπεράσματα

Τα συμπεράσματα από την έρευνα που περιγράφηκε στο κεφάλαιο αυτό είναι:

- Εισαγάγαμε μία τεχνική για τον υπολογισμό της αξιοπιστίας των κόμβων σε ένα ad hoc δίκτυο μέσω της συνεργασίας των κόμβων. Το πλαίσιο αυτό (ATF) υλοποιήθηκε μέσω μιας καινοτόμου και κατανεμημένης αρχιτεκτονικής.
- Προτάθηκε ένας μηχανισμός για την ανταλλαγή των συστάσεων ο οποίος λαμβάνει υπόψη μόνο τους αξιόπιστους συστήνοντες χωρίς να δημιουργεί περιττή κίνηση στο δίκτυο. Η αρχιτεκτονική σχεδιάστηκε κατά τέτοιο τρόπο ώστε να ενσωματώνει πολλές παραμέτρους που καθορίζονται από τους χρήστες μέσω μιας πολιτικής. Αυτό επιτρέπει

πιθανές επεκτάσεις και εναλλακτικές λύσεις και βρίσκεται σε συμφωνία με το όραμα του διάχυτου υπολογισμού.

- Το πλαίσιο *ATF* εισάγει την έννοια των αξιόπιστων συστηνόντων (*trusted recommenders*). Οι εξειδικευμένοι αισθητήρες αξιοπιστίας *RFTS* εξετάζουν, ελέγχουν, και αξιολογούν την αξιοπιστία των κόμβων όσον αφορά την υπηρεσία σύστασης και έτσι συμβάλλουν στην αποφυγή φαινομένων κακόβουλης διάδοσης φήμης.
- Ο προτεινόμενος μηχανισμός *TrustSpan* ελαχιστοποιεί τα κόστη επικοινωνίας και επιταχύνει τη διαδικασία ανίχνευσης της αξιοπιστίας, δεδομένου ότι μόνο οι αξιόπιστοι κόμβοι ερωτούνται για συστάσεις. Τα ιστορικά δεδομένα ενσωματώνονται ομαλά στο μοντέλο υπολογισμού αξιοπιστίας, διευκολύνοντας τον ταχύ προσδιορισμό των εγωιστικών κόμβων, χωρίς να «τιμωρούν» εκείνους τους κόμβους που συμπεριφέρθηκαν εγωιστικά λόγω τυχαίων λαθών στο δίκτυο.
- Η προτεινόμενη αρχιτεκτονική παρέχει έναν μηχανισμό δεύτερης ευκαιρίας, όπου οι απομονωμένοι κόμβοι έχουν πάντα την ευκαιρία να συμμετέχουν στις διαδικασίες του δικτύου, εάν αυτοί αρχίζουν να συνεργάζονται.

Τα θέματα που αναλύθηκαν στο κεφάλαιο αυτό αναφέρονται στις δημοσιεύσεις: [53], [54], [55], [56].

ΚΕΦΑΛΑΙΟ 6 ΣΧΕΤΙΚΗ ΕΡΕΥΝΑ

6.1 Εισαγωγή

Στο κεφάλαιο αυτό αναφέρεται η βασικότερη σχετική έρευνα καθώς και σχετική σύγκριση με όλα τα ερευνητικά πεδία που εστίασε η διδακτορική διατριβή. Συγκεκριμένα, για κάθε κεφάλαιο αναφέρεται και η αντίστοιχη σχετική έρευνα. Η σειρά των υπο-παραγράφων του κεφαλαίου αυτού είναι σε αντιστοιχία με την σειρά των κεφαλαίων της διατριβής.

6.2 Σχετική Έρευνα στην Επίγνωση Πλαισίου για τον Χρήστη

Στην παράγραφο αυτή παρουσιάζουμε τη σχετική έρευνα στην επίγνωση πλαισίου για τα κατάσταση του χρήστη. Τα συστήματα προσδιορισμού θέσης αποτελούν έναν ενεργό ερευνητικό τομέα από το 1992 με το σύστημα Active Badge [1]. Στο σύστημα αυτό μία ετικέτα που φέρεται από τους χρήστες εκπέμπει ένα μοναδικό χαρακτηριστικό μέσω υπέρυθρης ακτινοβολίας (IR). Αισθητήρες που τοποθετούνται σε γνωστές θέσεις μέσα σε ένα κτήριο ανιχνεύουν τα σήματα αυτά και τα αναμεταδίδουν στο κατάλληλο λογισμικό ώστε να προσδιοριστεί η θέση του χρήστη. Από τότε, διάφορα συστήματα θέσης εσωτερικού έχουν προταθεί. Πολλά από αυτά χρησιμοποιούν την τεχνολογία 802.11 (Wifi) για να προσδιοριστεί η θέση των χρηστών.

Το σύστημα RADAR [16] βασίζεται στα ραδιοκύματα (RF) για να εντοπίσει τους χρήστες στο εσωτερικού ενός κτιρίου. Λειτουργεί με την καταγραφή και την επεξεργασία των λαμβανόμενων πληροφοριών που σχετίζονται με την ισχύ του σήματος (RSS) από πολλαπλούς σταθμούς βάσης τοποθετημένους έτσι ώστε να παρέχουν επαρκή κάλυψη στο χώρο ενδιαφέροντος. Στη συνέχεια συνδυάζει τις εμπειρικές μετρήσεις με το μοντέλο διάδοσης των σημάτων για να καθορίσει τη θέση χρηστών. Η μέθοδος RSS χρησιμοποιείται επίσης από το εμπορικό σύστημα EkaHau [17]. Το σύστημα EkaHau χρησιμοποιεί τεχνικές διόρθωσης για τον προσδιορισμό χωρικής πληροφορίας πλαισίου. Συγκεκριμένα, η θέση του χρήστη υπολογίζεται βάσει στατιστικών μεγεθών που εκμεταλλεύονται μετρήσεις σήματος ισχύος στον δέκτη (Received Signal Strength). Αυτό το σύστημα βασίζεται στην εκπαίδευση (training) σε κάθε κτήριο που εγκαθίσταται και χρησιμοποιεί μια προσέγγιση όπου η θέση υπολογίζεται με τη βοήθεια των μετρήσεων RSS στην πλευρά των χρηστών. Ένα άλλο σύστημα εντοπισμού που βασίζεται στην ασύρματη τεχνολογία LAN περιγράφεται στην [18].

Τα συστήματα CRICKET [19] και Active Bat [20] είναι δύο συστήματα που χρησιμοποιούν την τεχνολογία υπερηχητικών σημάτων. Προσδιορίζοντας την απόσταση από γνωστά σημεία σε ένα κτήριο μέσω των σημάτων υπερήχων μπορούν να παρασχεθούν πληροφορίες θέσης. Ως πλεονέκτημα τους θεωρείται η ακρίβεια αλλά σαν κύριο μειονέκτημά τους η κλιμάκωση που μπορούν να προσφέρουν αλλά και το υψηλό κόστος. Για αυτούς του λόγους τα συστήματα αυτά είναι απρόσιτα στην πλειοψηφία των χρηστών.

Μια άλλη κατηγορία συστημάτων θέσης χρησιμοποιεί πολλαπλούς αισθητήρες (Wifi, RFIDs) και τεχνικές σύντηξης δεδομένων για να υπολογιστεί η θέση ενός χρήστη. Το σύστημα Location Stack [21] υιοθετεί τέτοιες τεχνικές ώστε να συνδυαστούν οι αναγνώσεις από τους αισθητήρες. Είναι ανάλογο με το σχήμα OSI (Open System Interconnection) για τις υπηρεσίες εξαρτώμενες από τη θέση LBS (Location Based Services). Μια άλλη παρόμοια προσέγγιση περιγράφεται στην [22]. Το μειονέκτημα αυτών των συστημάτων είναι η ανικανότητά τους στην υποστήριξη των κινητών συσκευών με τις περιορισμένες δυνατότητες (CPU, μνήμη) επειδή οι πολύπλοκοι υπολογισμοί για την εκτίμηση θέσης εκτελούνται κυρίως στα τερματικά των χρηστών πελατών, ως εκ τούτου συσκευές αναλαμβάνουν το κόστος των σύνθετων αυτών υπολογισμών.

Στα πλαίσια της διδακτορικής διατριβής προτάθηκε ένα σύστημα εκτίμησης θέσης που χρησιμοποιεί τα δεδομένα από τους αισθητήρες διαφορετικών τεχνολογιών για να καθορίσει τη θέση ενός χρήστη. Το προτεινόμενο σύστημα διαφέρει από τις προηγούμενες προσεγγίσεις σε διάφορες πτυχές. Κατ' αρχάς, χρησιμοποιούμε τα Δυναμικά Δίκτυα Bayes (DBNs) για τον προσδιορισμό της θέσης. Με τη χρήση DBNs επιτυγχάνουμε καλύτερα αποτελέσματα εκτίμησης θέσης διότι μαζί με τα ετερογενή στοιχεία που λαμβάνουμε από τους αισθητήρες μπορούμε επίσης να συνδυάσουμε πληροφορίες του παρελθόντος που σχετίζονται με το χρήστη (προηγούμενη θέση και ιστορικά στοιχεία) και έτσι δίνεται η δυνατότητα για προβλέψεις όσον αφορά τη μελλοντική συμπεριφορά του χρήστη.

Δεύτερον, το σύστημά μας μπορεί να υποστηρίξει ποικίλες κινητές συσκευές (PDAs, palmtops) με χαμηλές υπολογιστικές ικανότητες, καθώς ο υπολογισμός θέσης πραγματοποιείται σε έναν κεντρικό υπολογιστή. Οι κινητές συσκευές στέλνουν τις παρατηρήσεις από τους αισθητήρες σε αυτόν τον κεντρικό υπολογιστή όποτε αυτό απαιτείται και όχι συνεχώς. Κατά συνέπεια, το κόστος επικοινωνίας μειώνεται. Επίσης η αρχιτεκτονική του συστήματος που υιοθετούμε έχει το πλεονέκτημα της εύκολης

διαχείρισης. Παραδείγματος χάριν, η εγκατάσταση ενός νέου σταθμού βάσης είναι εντελώς διαφανής στους χρήστες. Λαμβάνοντας επίσης υπόψη ότι μπορούμε να χρησιμοποιήσουμε την ήδη υπάρχουσα υποδομή του τοπικού ασύρματου LAN ενός κτηρίου και ποικίλους φθηνούς, απλούς αισθητήρες, το προτεινόμενο σύστημα εκτίμησης θέσης προσφέρει μια αξιοπρόσεκτη λύση για τον εσωτερικό προσδιορισμό θέσης.

Το μοντέλο εκτίμησης θέσης που περιγράφηκε στις προηγούμενες παραγράφους βασίζεται σε Δυναμικά Δίκτυα Bayes. Παρόλα αυτά, γενικά η επίγνωση πλαισίου δεν στηρίζεται μόνο στον εντοπισμό χωρικού πλαισίου (θέσης). Ένας χρήστης χαρακτηρίζεται από κάποια κατάσταση σε κάποια χρονική στιγμή (π.χ. «συμμετοχή σε συνάντηση»). Για αυτό, η σύντηξη πολλών διαφορετικών συνιστωσών πλαισίου κρίνεται απαραίτητη. Στα πλαίσια της παρούσας διατριβής προτάθηκε πλαίσιο για την επίγνωση της κατάστασης του χρήστη που αποτελεί την γενίκευση του συστήματος προσδιορισμού θέσης. Η έρευνα [24] ασχολείται με την σύντηξη πολλαπλών δεδομένων που αντιστοιχούν σε συνιστώσες πλαισίου για τον συμπερασμό της τρέχουσας πληροφορίας πλαισίου. Το μοντέλο όμως αυτό δεν λαμβάνει υπό όψιν την πιστότητα των μετρήσεων – δεδομένων κατά τον συμπερασμό του πλαισίου. Επίσης, η εκτίμηση του πλαισίου στην [25] χρησιμοποιώντας ταξινόμηση κατά Naive-Bayes επιτυγχάνεται χωρίς την ανακριβή φύση του πλαισίου συντελώντας έτσι σε μη ικανοποιητικά αποτελέσματα. Το προτεινόμενο πλαίσιο χρησιμοποιώντας κανόνες Ασαφούς Λογικής έχει τη δυνατότητα ενσωμάτωσης της αξιοπιστίας των πηγών.

6.3 Σχετική Έρευνα στην Επίγνωση Πλαισίου για το Περιβάλλον

Στην παράγραφο αυτή παρουσιάζουμε τη σχετική έρευνα στην επίγνωση πλαισίου για τα κατάσταση του περιβάλλοντος. Εστιάζουμε κυρίως στην ανίχνευση καταστάσεων για το περιβάλλον που αποτελούν κίνδυνο, όπως οι πυρκαγιές. Οι προγενέστερες ερευνητικές δραστηριότητες στην πυρανίχνευση χρησιμοποιούν τους αισθητήρες θερμοκρασίας και υγρασίας, τους ανιχνευτές καπνού και τους υπέρυθρους οπτικούς αισθητήρες (κάμερες). Συγκεκριμένα, στην [32] προτείνεται ένα σύστημα πυρανίχνευσης βασισμένο στην τεχνολογία πολλαπλών αισθητήρων και τα νευρωνικά δίκτυα. Τα δεδομένα που ανακτώνται από τους αισθητήρες είναι η περιβαλλοντική θερμοκρασία, η πυκνότητα καπνού και η πυκνότητα του μονοξειδίου του άνθρακα (CO). Εντούτοις, το προτεινόμενο μοντέλο χρειάζεται έναν μεγάλο αριθμό δειγμάτων ώστε να εκπαιδευτεί (train) λόγω της ταξινόμησης που χρησιμοποιείται στα νευρωνικά δίκτυα. Επιπλέον, τα

πιθανά γεγονότα της πυρκαγιάς πρέπει να είναι γνωστά (a-priori) από το νευρωνικό δίκτυο. Ως εκ τούτου, για ένα απαραίτητο γεγονός ολόκληρη η δομή του δικτύου πρέπει να ανασχηματιστεί (το δίκτυο δεν υποστηρίζει τη λειτουργία της τοπικής πρότυπης ταξινόμησης). Στην [33] και [34], οι συγγραφείς παρουσιάζουν ένα σύστημα που είναι βασισμένο σε διαφορετικούς τύπους αισθητήρων και χρησιμοποιεί (πιθανοκρατικά) νευρωνικά δίκτυα. Παρόλα αυτά τα συγκεκριμένα συστήματα απαιτούν τη χρήση δεδομένων εκπαίδευσης (training data) και τα περισσότερα αξιολογούνται στο εσωτερικό των κτιρίων όπου οι περιβαλλοντικές συνθήκες είναι πλήρως ελεγχόμενες και απολύτως διαφορετικές σε σύγκριση με εκείνες που παρατηρούνται σε εξωτερικό χώρο. Ένα σύστημα πυρανίχνευσης που χρησιμοποιεί ένα ασύρματο δίκτυο αισθητήρων (WSN) για τη συλλογή της θερμοκρασίας, τη σχετική υγρασία και τη βαρομετρική πίεση περιγράφεται στην [35]. Οι συγγραφείς στην [36] και στην [37] προτείνουν συστήματα που βασίζονται στην υπέρυθη τεχνολογία (IR) για την ανίχνευση των πυρκαγιών. Επιπλέον, στις εργασίες [38] και [39] έχουν προταθεί λύσεις στις οποίες οι αισθητήρες ρίπτονται από αεροσκάφος. Στην [40] προτείνεται ένας καταναμημένος αλγόριθμος για την εξισορρόπηση του φόρτου σε ένα δίκτυο αισθητήρων. Εντούτοις, αυτά τα συστήματα χρησιμοποιούν είτε τους τοπικούς αισθητήρες ή τους απομακρυσμένους αισθητήρες, και επομένως είναι επιρρεπής σε ψευδείς συναγερμούς. Επιπλέον, οι εναέριες ή δορυφορικές εικόνες χρησιμοποιούνται συχνά για την κάλυψη περιοχών με στόχο την προστασία από τις πυρκαγιές. Στην εργασία [41] παρουσιάζεται μια αρχιτεκτονική που προσπαθεί να συνδυάσει πρόσφατες μετρήσεις (από τις δορυφορικές εικόνες και άλλους αισθητήρες) με ένα σύστημα πρόβλεψης. Τα δεδομένα των αισθητήρων χρησιμοποιούνται για να ελαχιστοποιήσουν το σφάλμα του μοντέλου πρόβλεψης. Αυτή η μέθοδος είναι γνωστή και ως Δυναμικό Σύστημα Εφαρμογής Οδηγούμενο από τα Δεδομένα (Dynamic Data Driven Application System - DDDAS) και παρουσιάζει μεγάλο ενδιαφέρον για την επιστημονική κοινότητα. Η πυρανίχνευση που βασίζεται σε δορυφορικές εικόνες προτείνεται επίσης και στην [42]. Όμως, η περίοδος ανίχνευσης και η χαμηλή ανάλυση των δορυφορικών εικόνων καθιστούν αυτές τις μεθόδους ανίκανες για την έγκαιρη (real time) πυρανίχνευση. Η προσέγγισή μας είναι παρόμοια με την [35] δεδομένου ότι χρησιμοποιείται ένα ασύρματο δίκτυο για τη συλλογή των δεδομένων. Εντούτοις, υιοθετούμε τη χρήση των απομακρυσμένων (οπτικών) αισθητήρων για τη διαδικασία της πυρανίχνευσης. Επιπλέον, η προτεινόμενη αρχιτεκτονική της δι-επίπεδης σύντηξης, μειώνει το ποσοστό των ψευδών συναγερμών.

6.4 Σχετική έρευνα στην Βιο-μιμητική Συνεργατική Επίγνωση Πλαισίου

Σχήματα βασισμένα στην πιθανοτική διάδοση πληροφορίας ([76], [80]) έχουν προταθεί προκειμένου να αντιμετωπιστεί το ζήτημα της αξιοπιστίας που προκύπτει κατά την μετάδοση των μηνυμάτων καθώς απομακρυνόμαστε από την περίπτωση του σχήματος της πλημμύρας. Οι αλγόριθμοι που βασίζονται στην διάδοση φήμης (gossip based) συνήθως «ανταλλάσσουν» την παρεχόμενη αξιοπιστία με την κλιμάκωση που μπορεί να επιτευχθεί. Ο αλγόριθμος που περιγράφεται στην [76] αποτελεί ένα νέο αλγόριθμο βασισμένο στη διάδοση φήμης που συμπληρώνει την κλιμάκωση των παραδοσιακών πιθανοτικών αλγορίθμων διάδοσης πληροφορίας με μια εξελικτική τεχνική διαχείρισης της μνήμης σε κάθε κόμβο του δικτύου. Εντούτοις, η έννοια της επίγνωσης μιας κατάστασης (awareness) μπορεί να ισχύσει και για τις ιδιότητες που σχετίζονται με το δίκτυο, όπως η λαμβανόμενη ισχύς του σήματος RSSI (Received Signal Strength Indicator) [81]. Η περίπτωση αυτή ναι μεν μας παρέχει πληροφορίες που αφορούν κυρίως την κατάσταση του δικτύου, δεν αποτελεί όμως επίγνωση που βασίζεται στη σημασιολογία όπως τα πρωτόκολλα SPIN [82], [83]. Παρόλα αυτά, στην [81] παρουσιάζεται μία αρχιτεκτονική η οποία χρησιμοποιώντας πληροφορίες της κατάστασης του δικτύου καταφέρνει την εξοικονόμηση ενέργειας στους κινητούς κόμβους και ταυτόχρονα την ικανοποιητική διάδοση της πληροφορίας στο δίκτυο.

Στην [84] μελετήθηκαν θέματα διάδοσης πληροφορίας σε ad-hoc δίκτυα χρησιμοποιώντας επιδημικούς αλγορίθμους. Στην [85] οι συγγραφείς μελέτησαν την επιδημική διάδοση σε πολύπλοκα / σύνθετα δίκτυα και η αρχιτεκτονική που παρουσιάζεται στην [86] χρησιμοποιεί την κινητικότητα των χρηστών για τη μεταφορά δεδομένων μεταξύ των δικτύων. Στην [87] αναλύεται ένα πλαίσιο (framework) που βασίζεται σε μία διαδικασία Markov χρησιμοποιώντας ταυτόχρονα το μοντέλο SIS για την διάδοση της πληροφορίας και μελετάται επίσης ο αντίκτυπος της τοπολογίας του δικτύου στη διάδοση. Στην εργασία [88] εξετάζεται η εφαρμογή του SI μοντέλου για την διάδοση σε ad-hoc δίκτυα. Οι συγγραφείς στην [89] προτείνουν ένα πρότυπο για τη διάδοση σε αυθαίρετες τοπολογίες δικτύων, ενώ η αρχιτεκτονική που μελετάται στην [90] περιλαμβάνει μια προσέγγιση της συνεργατικής διάδοσης πληροφορίας πλαισίου μεταξύ ομάδων από κινητούς κόμβους. Επιπλέον, στην [91] προτείνεται μια μέθοδος για την βελτίωση της βασική λειτουργίας της επιδημικής διάδοσης στην οποία η ιστορία της συνδετικότητας και της θέσης των κόμβων χρησιμοποιείται για να ληφθεί η κατάλληλη απόφαση προώθησης των μηνυμάτων.

Στις προαναφερθείσες προσεγγίσεις που εξετάζονται αλγόριθμοι πιθανοτικής διάδοσης και επιδημικοί αλγόριθμοι, ο ρυθμός μόλυνσης β των κινητών κόμβων με νέα,

πρόσφατη πληροφορία πλαισίου παραμένει σταθερός. Στην [92] προτείνεται μια προσαρμογή του ρυθμού μόλυνσης β και του ρυθμού ίασης δ σύμφωνα με τα χαρακτηριστικά της γειτονιάς του κόμβου, ενώ οι συγγραφείς στην [93] εισάγουν την έννοια της χρησιμοποίησης μηνυμάτων αναγνώρισης (acknowledgment messages) για να «θεραπεύσουν» μια επιδημία με στόχο την μείωση της κατανάλωσης των πόρων (ενέργεια, εύρος ζώνης, κτλ.). Εντούτοις, ούτε αυτές οι προσεγγίσεις λαμβάνουν υπόψη τα χαρακτηριστικά της πληροφορίας πλαισίου, έναν σημαντικό παράγοντα που οδηγεί στην υιοθέτηση ενός προσαρμοστικού σχήματος διάδοσης πληροφορίας.

6.5 Σχετική έρευνα στην Συνεργατική Επίγνωση Πλαισίου για τον συμπερασμό της αξιοπιστίας των κόμβων

Παρόμοιες αρχιτεκτονικές διαχείρισης αξιοπιστίας παρουσιάζονται στις εργασίες [57] και [62]. Στην [57] προτείνεται μια αναδρομική ιεραρχία στη δημιουργία αξιοπιστίας που είναι βασισμένη σε ιστορικά δεδομένα. Η παραγόμενη αξιοπιστία μπορεί να χρησιμοποιηθεί για την περαιτέρω δημιουργία αξιοπιστίας. Στην [62] οι συγγραφείς προτείνουν την αρχιτεκτονική CONFIDANT που αποτελείται από το σύστημα ελέγχου (monitor), το σύστημα φήμης (reputation system), το διαχειριστή μονοπατιών (path manager) και το διαχειριστή αξιοπιστίας (trust manager). Το σύστημα ελέγχου ανιχνεύει κακόβουλες συμπεριφορές (DoS, κ.λπ.). Ο διαχειριστής αξιοπιστίας εξετάζει εισερχόμενα και εξερχόμενα μηνύματα συναγερμού (ALARM) που στέλνονται από ένα κόμβο για να προειδοποιήσουν για την ύπαρξη κακόβουλων κόμβων. Το σύστημα φήμης διαχειρίζεται έναν πίνακα που αποτελείται από τις καταχωρήσεις για τους κόμβους και την εκτίμηση τους. Η εκτίμηση αυτή τροποποιείται μόνο όταν υπάρχουν επαρκή στοιχεία για την κακόβουλη συμπεριφορά. Ο διαχειριστής μονοπατιών ταξινομεί μονοπάτια σύμφωνα με τις μετρικές ασφάλειας (π.χ. φήμη των κόμβων σε ένα μονοπάτι) και διαγράφει μονοπάτια που περιέχουν κακόβουλους κόμβους. Η κύρια διαφορά μεταξύ της προτεινόμενης αρχιτεκτονικής ATF και του CONFIDANT είναι ότι η πρώτη είναι ικανή να υποστηρίξει συστάσεις κατόπιν αίτησης (on demand) ενώ η δεύτερη βασίζεται στα μηνύματα ALARM που μεταδίδονται από ένα κόμβο για να προειδοποιήσει για την ύπαρξη άλλων κακόβουλων κόμβων. Επιπλέον, στο σχήμα ATF κάθε κόμβος χαρακτηρίζεται από μία *Πολιτική Οικοδόμησης Εμπιστοσύνης (Trust Policy)* που λαμβάνει υπόψη την υποκειμενική κρίση και τα ιστορικά δεδομένα για να αξιολογήσει το επίπεδο αξιοπιστίας των υπηρεσιών που προσφέρουν οι άλλοι κόμβοι.

Η εργασία στην [57] θεωρείται ως βάση για την προτεινόμενη αρχιτεκτονική ATF. Η έννοια της γνώσης στην [57] ενσωματώνεται στο μοντέλο μας μέσω των ιστορικών

δεδομένων, των υποκειμενικών παραγόντων αλλά και της *Πολιτικής Οικοδόμησης Εμπιστοσύνης (Trust Policy)*. Οι συστάσεις κατόπιν αίτησης και τα άμεσα αποδεικτικά στοιχεία (direct evidence - *DE*) που αποθηκεύονται υποστηρίζουν την αναδρομική δημιουργία αξιοπιστίας καθώς χρησιμοποιούνται για μελλοντικές εκτιμήσεις. Η εργασία στην [57] καθορίζει έναν αφηρημένο μηχανισμό αξιοπιστίας για να αξιολογήσει τα επίπεδα αξιοπιστίας. Στην αρχιτεκτονική *ATF* προτείνουμε έναν πιο εξελιγμένο μηχανισμό τον *Δημιουργό Αξιοπιστίας (Trust Builder- TB)*. Όπως αναφέρεται στην [63] μπορεί να εμφανιστούν επιθέσεις άρνησης της υπηρεσίας (Denial of Service – *DoS*) αν διαδίδονται αρνητικές φήμες για διάφορους κόμβους. Επιπλέον, οι συγγραφείς στην [64] αναφέρουν ότι η χρήση μόνο θετικών συστάσεων ελαχιστοποιεί την επίδραση των φαινομένων διάδοσης φήμης. Στην αρχιτεκτονική *ATF* λαμβάνουμε υπόψη τα άμεσα αποδεικτικά στοιχεία, καθώς επίσης και τις συστάσεις που παρέχονται μόνο από τους αξιόπιστους κόμβους. Επιπλέον, προτείνουμε έναν καινοτόμο μηχανισμό ελέγχου για την αξιολόγηση της αξιοπιστίας των συστημάτων (*recommenders*). Τέλος υποθέτουμε την ύπαρξη ενός πρωτοκόλλου που μεταδίδει τις συστάσεις μεταξύ των κόμβων. Ένα τέτοιο πρωτόκολλο περιγράφεται στην [60].

ΚΕΦΑΛΑΙΟ 7 ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΚΑΤΕΥΘΥΝΣΕΙΣ

7.1 Συμπεράσματα

Στη διατριβή αυτή παρουσιάστηκε ένα σύνολο από μελέτες που αφορούν τη διαχείριση πλαισίου σε συστήματα διάχυτου υπολογισμού. Ο διάχυτος υπολογισμός αναφέρεται στην εξελιγμένη κλάση συστημάτων του κινητού υπολογισμού. Τα συστήματα αυτά καλούνται να διαχειριστούν την πληροφορία που χαρακτηρίζει την περιρρέουσα κατάσταση των επενεργουσών οντοτήτων (π.χ., χρήστες, κινητοί υπολογιστές). Η έννοια «πληροφορία πλαισίου» ή «πλαίσιο» αναφέρεται στην πληροφορία που χαρακτηρίζει την περιρρέουσα, αντιλαμβανόμενη από το σύστημα, κατάσταση μιας οντότητας. Τα συστήματα διάχυτου υπολογισμού καλούνται:

- να προσαρμόζονται καταλλήλως στη συμπεριφορά των οντοτήτων αυτών βάσει της κατάστασής τους, καθώς και,
- να λαμβάνουν αυτόνομα αποφάσεις ελαχιστοποιώντας τη συμβολή, διαμεσολάβηση και παρέμβαση των οντοτήτων.

Η διαχείριση πλαισίου συντελεί στην αξιοποίηση ενέργειας και πόρων σε κατανεμημένα συστήματα, και στην προσαρμοστικότητα της συμπεριφοράς της εφαρμογής κινητού υπολογισμού σύμφωνα με την τρέχουσα κατάσταση του χρήστη. Εξετάστηκαν διάφορα προβλήματα σε διαφορετικά και συμπληρωματικά επίπεδα. Βασικός άξονας της μελέτης που διενεργήθηκε αποτελεί η μεθοδολογία που ακολουθήθηκε για την μοντελοποίηση, αλλά κι επίλυση των εν λόγω προβλημάτων, στα οποία ενέχονται στοιχεία αναπαράστασης και συμπερασμού γνώσης, καθώς και προσαρμοστικότητας καταστάσεων.

Η παρούσα διατριβή χωρίζεται σε δύο μέρη: στην «Αναπαράσταση, Μοντελοποίηση και Συμπερασμός Πλαισίου» και στην «Συνεργατική Επίγνωση Πλαισίου». Η συνεισφορά και τα αποτελέσματα για κάθε ένα μέρος αναφέρονται παρακάτω.

- Σημεία εστίασης στην Αναπαράσταση και Μοντελοποίηση Πλαισίου: η αναπαράσταση / μοντελοποίηση πλαισίου για ένα σύνολο οντοτήτων καθώς και η μοντελοποίηση των προβλημάτων αναπαράστασης, ερμηνείας και συμπερασμού πλαισίου με τη χρήση τεχνικών σύντηξης και θεωρίας ασαφούς συλλογιστικής.
- Σημεία εστίασης στην Συνεργατική Επίγνωση Πλαισίου: η οικοδόμηση εμπιστοσύνης μεταξύ των κόμβων ενός ασύρματου δικτύου, η αξιοπιστία των

κόμβων, η βιο-μιμητική διάχυση της πληροφορίας μεταξύ των κόμβων ενός ασύρματου δικτύου.

Από τη μελέτη των παραπάνω προσεγγίσεων, προέκυψαν τα κάτωθι σημαντικά συμπεράσματα.

7.1.1 Συμπεράσματα από την Επίγνωση της Κατάστασης του Χρήστη

Στα πλαίσια της έρευνας για την επίγνωση της κατάστασης του χρήστη στα περιβάλλοντα διάχυτου υπολογισμού παρουσιάσαμε μια στρωματοποιημένη αρχιτεκτονική σύντηξης δεδομένων που εκμεταλλεύεται τις πληροφορίες από αισθητήρες διαφορετικών τεχνολογιών για να υπολογίσει τη θέση ενός χρήστη. Μια βασική διαφορά από παρόμοια συστήματα είναι η χρήση των Δυναμικών Δικτύων Bayes για τον προσδιορισμό της θέσης. Η χρήση DBNs βελτιώνει τις εκτιμήσεις μας. Μαζί με τις πληροφορίες αισθητήρων λαμβάνουμε υπόψη και την προηγούμενη θέση του χρήστη που βελτιώνει έτσι την απόδοση του συστήματος. Επιπλέον, το σύστημα υποστηρίζει ποικίλες κινητές συσκευές συμπεριλαμβανομένων εκείνων με περιορισμένες υπολογιστικές ικανότητες (PDAs, κ.λπ....) δεδομένου ότι δεν υφίστανται το φόρτο των σύνθετων υπολογισμών θέσης. Τέλος, η αξιολόγηση του συστήματος σε πραγματικές συνθήκες απέδειξε την καταλληλότητά του για τον προσδιορισμό θέσης.

Επεκτείναμε την έρευνα για τον εντοπισμό θέσης λαμβάνοντας υπόψη την αξιοπιστία των πηγών (αισθητήρων, κτλ.). Μαζί με τα δεδομένα αισθητήρων, το προκύπτον πλαίσιο (inferred context) και η αξιοπιστία των πηγών συμπεριλαμβάνονται στη διαδικασία σύντηξης. Οι κανόνες της Ασαφούς Λογικής χρησιμοποιούνται προκειμένου να προκύψει ένα πιο συγκεντρωτικό αποτέλεσμα σύντηξης βασισμένο στην πιθανοτική κατανομή του πλαισίου. Η πειραματική αξιολόγηση του προτεινόμενου μοντέλου απέδειξε την ικανότητά του για την σύντηξη πλαισίου. Επιπλέον, εισήχθη και μια μέθοδος για τον υπολογισμό του βαθμού αξιοπιστίας των πηγών. Σύμφωνα με αυτή τη μέθοδο, που βασίζεται σε στατιστικά μεγέθη υπολογίζουμε την αξιοπιστία κάθε αισθητήρα χρησιμοποιώντας τις *πραγματικές* πληροφορίες που μας παρέχει, δηλαδή εισάγουμε την έννοια της *εντροπίας* της πηγής.

7.1.2 Συμπεράσματα από την Επίγνωση της Κατάστασης του Περιβάλλοντος

Σε αυτό το πεδίο έρευνας προτάθηκε ένας μηχανισμός για την επίγνωση της κατάστασης του περιβάλλοντος που βασίζεται στη σύντηξη δεδομένων από πολλαπλούς αισθητήρες. Ο μηχανισμός χρησιμοποιήθηκε για την ανίχνευση πυρκαγιών. Τα δεδομένα λαμβάνονται χρησιμοποιώντας ένα ασύρματο δίκτυο από κόμβους (αισθητήρες) θερμοκρασίας και υγρασίας και ένα απομακρυσμένο οπτικό αισθητήρα που ελέγχει την ίδια γεωγραφική περιοχή. Χρησιμοποιήθηκε μία τεχνική σωρευτικού αθροίσματος (Cumulative Sum -CUSUM) που συνδυάζει τα δεδομένα από γειτονικούς κόμβους για την ανίχνευση αλλαγών στις μετρήσεις που λαμβάνονται. Οι αλλαγές αυτές στη συνέχεια συγκρίνονται έναντι κατάλληλα επιλεγμένου κατωφλίου, το οποίο όταν ξεπερνιέται το σύστημα μεταπίπτει σε κατάσταση ALERT ή NOTIFY. Η τιμή του κατωφλίου επιλέγεται αναλόγως με το επιθυμητό ποσοστό ψευδών συναγερμών (false alarm rate). Επιπλέον, ορίζονται πιθανότητες ύπαρξης πυρκαγιάς στις παρατηρήσεις κάθε τοπικού αισθητήρα οι οποίες συνδυάζονται σε ένα πιο υψηλό επίπεδο με τις πιθανότητες που προκύπτουν από έναν απομακρυσμένο οπτικό αισθητήρα που ελέγχει την ίδια περιοχή.

Η σύντηξη πληροφορίας στο δεύτερο επίπεδο περιλαμβάνει δύο βήματα: επιλέγονται καταρχήν οι M υψηλότερες πιθανότητες για το γεγονός πυρκαγιάς από τους τοπικούς αισθητήρες και στη συνέχεια ο προτεινόμενος αλγόριθμος του δεύτερου επιπέδου τις συνδυάζει με την πιθανότητα πυρκαγιάς που παρέχει ο απομακρυσμένος οπτικός αισθητήρας. Το πρώτο βήμα είναι απαραίτητο για την άμεση ανταπόκριση του συστήματος στην διαδικασία πυρανίχνευσης, ενώ το δεύτερο βήμα χρησιμοποιεί συμπληρωματικά δεδομένα αισθητήρων για να αντιμετωπίσει πιθανές συγκρούσεις ενδείξεων που προκαλούνται από έκτοπες τιμές (outliers) ή δυσλειτουργία κάποιου αισθητήρα.

Η τεχνική CUSUM δεν έχει εφαρμοστεί για την πυρανίχνευση στο παρελθόν. Επιπλέον, θεωρούμε ότι η σύντηξη πιθανοτήτων πυρκαγιάς από έναν τοπικό αισθητήρα, με τις αντίστοιχες πιθανότητες που προέρχονται από τους απομακρυσμένους οπτικούς αισθητήρες αποτελεί μία πρωτοπόρα εφαρμογή της Θεωρίας Ενδείξεων DST. Τα πειραματικά αποτελέσματα που παρουσιάζονται επιβεβαιώνουν την επιλογή των προτεινόμενων τεχνικών. Μια κατάλληλη αλλαγή της μέσης τιμής στα δεδομένα που ακολουθούν την Gaussian κατανομή, ωθεί τον αλγόριθμο CUSUM στην ανίχνευση αλλαγής που σημάνει κατά πάσα πιθανότητα την ύπαρξη πυρκαγιάς.

7.1.3 Συμπεράσματα από την Βιο-μιμητική Συνεργατική Επίγνωση Πλαισίου

Στο συγκεκριμένο πεδίο έρευνας υιοθετούμε το βιο-μιμητικό μοντέλο επιδημικής διάδοσης για τη διάχυση της πληροφορίας αυτής σε ένα αδόμετο δίκτυο κινητών κόμβων. Μια παράμετρος στα ασύρματα δίκτυα είναι η κατανάλωση ενέργειας δεδομένου ότι η πλειοψηφία των κόμβων τροφοδοτείται με μπαταρίες. Αν και τα επιδημικά μοντέλα διάδοσης που προτείνονται στην βιβλιογραφία αποτελούν μια αποτελεσματικότερη λύση από ενεργειακής άποψης σε σύγκριση με άλλες τεχνικές διάδοσης (π.χ. σχήμα πλημμύρας), η επιλογή του ρυθμού μόλυνσης β που υιοθετεί το μοντέλο είναι εξαιρετικά σημαντική και μπορεί να βελτιώσει περαιτέρω την ενεργειακή απόδοση του επιδημικού σχήματος.

Στο κεφάλαιο αυτό εξετάσαμε αποδοτικά σχήματα για την επιδημική διάδοση προσαρμόζοντας το ρυθμό μόλυνσης β των κόμβων. Τα αποτελέσματα προσομοίωσης δείχνουν ότι για μια δεδομένη πυκνότητα στο δίκτυο, υπάρχει ένας ελάχιστος ρυθμός μόλυνσης β_{min} (πιθανότητα διείσδυσης), που επιτυγχάνει την πλήρη κάλυψη του δικτύου. Δοκιμάζοντας διαφορετικές περιόδους μετάδοσης, ανακαλύψαμε ότι η πιθανότητα διείσδυσης μειώνεται με τον αριθμό των μεταδόσεων. Εντούτοις, από άποψη αποδοτικότητας, χρησιμοποιώντας έναν χαμηλό αριθμό προσπαθειών και υψηλή πιθανότητα μετάδοσης οδηγούμαστε σε λιγότερες μεταδόσεις. Επιπλέον, οι προσομοιώσεις μας παρουσίασαν τη σημασία ενός μηχανισμού ελέγχου συμφόρησης. Εάν το δίκτυο «δημιουργήσει» πάρα πολλά μηνύματα, ο μηχανισμός δεν θα επιτρέψει την πλήρη κάλυψη αλλά θα βοηθήσει στη μέγιστη πιθανή κάλυψη δεδομένης της πραγματικής παραγόμενης κυκλοφορίας.

Προτείνουμε επίσης έναν προσαρμοστικό πιθανολογικό αλγόριθμο μετάδοσης ($\Delta 1$) βασισμένο στην πυκνότητα του δικτύου που προκύπτει από την εκτίμηση του αριθμού των γειτόνων. Τα αποτελέσματα προσομοίωσης αποδεικνύουν ότι μια καλή κάλυψη μπορεί να ληφθεί για διαφορετικούς όρους πυκνότητας, μεγέθους και κινητικότητας των κόμβων. Ένα άλλο πλεονέκτημα αυτού του σχήματος είναι ότι δεν απαιτεί τη σηματοδότηση. Ως εκ τούτου το κόστος που επιφέρει είναι ελάχιστο. Επιπλέον προτείνεται και ο προσαρμοστικός αλγόριθμος $\Delta 2$. Η κύρια διαφορά με το $\Delta 1$ είναι ότι προσπαθεί να συμπεράνει την κινητικότητα και την λαμβάνει υπόψη κατά την υπολογισμό του ρυθμού μόλυνσης β . Επιπλέον, ο $\Delta 2$ αρχίζει από μια τιμή αναφοράς β_0 και αυξάνει ή μειώνει το ρυθμό σύμφωνα με την αλλαγή της πυκνότητας του δικτύου. Οι προσομοιώσεις δείχνουν ότι οι προσδιορισμένες τιμές για το β εξαρτώνται σε μεγάλο βαθμό από το β_0 .

Στην συνέχεια μελετάμε την προσαρμογή του ρυθμού μόλυνση με βάση την μεταβλητότητα του πλαισίου που μεταδίδεται. Στο πρώτο προτεινόμενο προσαρμοστικό σχήμα διάδοσης (L1) ο ρυθμός μόλυνσης ρυθμίζεται σύμφωνα με τις αλλαγές στις τιμές της πληροφορίας πλαισίου που παράγονται από τον κόμβο πηγή. Οι γρήγορες αλλαγές στις τιμές του ρεύματος δεδομένων (υψηλή μεταβλητότητα) θα αυξήσουν το ρυθμό μόλυνσης (από την πηγή προς στους γείτονές της) ενημερώνοντας έτσι περισσότερους κόμβους στο δίκτυο για αυτήν την αλλαγή. Αξιολογούμε την προτεινόμενη μέθοδο σε σχέση με το σφάλμα της πληροφορίας πλαισίου τον αριθμό των μηνυμάτων εξέπεμψε η πηγή. Η προσομοίωση εκτελέσθηκε χρησιμοποιώντας δύο ρεύματα δεδομένων από πραγματικές μετρήσεις (θερμοκρασία και ταχύτητα ανέμου) που διαφέρουν σημαντικά όσον αφορά την μεταβλητότητά τους. Από τα πειράματα που πραγματοποιήθηκαν αποδείχθηκε ότι το σχήμα με τον προσαρμοστικό ρυθμό μόλυνσης β_s είναι αποδοτικότερο σε σχέση με τα σχήματα που έχουν σταθερό ρυθμό μόλυνσης. Τέλος, προτείνεται, το σχήμα L2 που είναι βασισμένο στο L1, προκειμένου να αντιμετωπιστεί η κατάσταση κατά την οποία στο δίκτυο διαδίδεται απαρχαιωμένη, μη ουσιώδης χρονικά πληροφορία πλαισίου. Σε αυτό το επιδημικό σχήμα διάδοσης αν και οι κόμβοι μπορούν να χρησιμοποιήσουν την απαρχαιωμένη πληροφορία πλαισίου, η διάδοσή της επηρεάζεται από την ηλικία της. Η αξιολόγηση του L2 σχετικά με την παλαιότητα (ηλικία) της πληροφορίας στο δίκτυο δείχνει ότι το σχήμα L2 είναι αποδοτικότερο του L1.

7.1.4 Συμπεράσματα από τον συμπερασμό της αξιοπιστίας με βάση τη Συνεργατική Επίγνωση Πλαισίου

Εισαγάγαμε ένα πλαίσιο για τον συμπερασμό της αξιοπιστίας των κόμβων σε ένα ad hoc δίκτυο. Το πλαίσιο αυτό (ATF) υλοποιήθηκε μέσω μίας καινοτόμου και κατανεμημένης αρχιτεκτονικής που αποτελείται από διάφορα συστατικά για τον συμπερασμό, την ανταλλαγή συστάσεων και την ανίχνευση της συμπεριφοράς των κόμβων. Προτάθηκε επίσης ένας μηχανισμός για την ανταλλαγή των συστάσεων ο οποίος λαμβάνει υπόψη μόνο τους αξιόπιστους συστήνοντες χωρίς να δημιουργεί περιττή κίνηση στο δίκτυο. Η αρχιτεκτονική σχεδιάστηκε κατά τέτοιο τρόπο ώστε να ενσωματώνει πολλές παραμέτρους που καθορίζονται από τους χρήστες μέσω μιας πολιτικής. Αυτό επιτρέπει πιθανές επεκτάσεις και εναλλακτικές λύσεις και βρίσκεται σε συμφωνία με το όραμα του διάχυτου υπολογισμού.

Μια βασική διαφορά από παρόμοια συστήματα είναι εισαγωγή του υποκειμενικού παράγοντα, ο οποίος εισάγει τη φυσική συμπεριφορά των τελικών χρηστών και

μοντελοποιεί διάφορες στρατηγικές οικοδόμησης εμπιστοσύνης. Επιπλέον, το πλαίσιο *ATF* εισάγει την έννοια των αξιόπιστων συστηνότων (*trusted recommenders*). Οι εξειδικευμένοι αισθητήρες αξιοπιστίας *RFTS* εξετάζουν, ελέγχουν, και αξιολογούν την αξιοπιστία των κόμβων όσον αφορά την υπηρεσία σύστασης και έτσι συμβάλλουν στην αποφυγή φαινομένων κακόβουλης διάδοσης φήμης. Επιπλέον, ο προτεινόμενος μηχανισμός *TrustSpan* ελαχιστοποιεί τα κόστη επικοινωνίας και επιταχύνει τη διαδικασία ανίχνευσης της αξιοπιστίας, δεδομένου ότι μόνο οι αξιόπιστοι κόμβοι ερωτούνται για συστάσεις. Τα ιστορικά δεδομένα ενσωματώνονται ομαλά στο μοντέλο υπολογισμού αξιοπιστίας, διευκολύνοντας τον ταχύ προσδιορισμό των εγwissτικών κόμβων, χωρίς να «τιμωρούν» εκείνους τους κόμβους που συμπεριφέρθηκαν εγwissτικά λόγω τυχαίων λαθών στο δίκτυο. Τέλος, η προτεινόμενη αρχιτεκτονική παρέχει έναν μηχανισμό δεύτερης ευκαιρίας, όπου οι απομονωμένοι κόμβοι έχουν πάντα την ευκαιρία να συμμετέχουν στις διαδικασίες του δικτύου, εάν αυτοί αρχίζουν να συνεργάζονται.

Μια πιθανή απειλή που το *ATF* μπορεί να αντιμετωπίσει (ως ένα ορισμένο βαθμό) είναι η διάδοση των ψευδών συστάσεων ενορχηστρωμένες από κακόβουλους κόμβους. Η μέθοδος οικοδόμησης εμπιστοσύνης του *ATF* μπορεί να δώσει αρκετά λιγότερο βάρος στις συστάσεις απ' ό,τι στα άμεσα αποδεικτικά στοιχεία, εάν η παράμετρος a τεθεί κατάλληλα. Κατά συνέπεια, ακόμα και όταν οι λαμβανόμενες συστάσεις παρέχουν ψευδείς πληροφορίες για την αξιοπιστία ενός κόμβου *στόχου* σε έναν κόμβο *ανιχνευτή*, η τιμή αυτή θα ρυθμιστεί γρήγορα κατά την πρώτη αλληλεπίδραση των δύο κόμβων. Επιπλέον, όταν ένας *ανιχνευτής* έχει τον επαρκή αριθμό άμεσων αλληλεπιδράσεων με ένα κόμβο *στόχο* μπορεί να σταματήσει την λήψη των συστάσεων και επιπλέον να τιμωρήσει τους κακόβουλους συστήνοντες αποφεύγοντας τους για μελλοντικές αναφορές. Εντούτοις, είναι πάντα δυνατό για έναν κόμβο να απομονωθεί από κακόβουλους κόμβους.

Η αξιολόγηση της απόδοσης του πλαισίου *ATF* μέσω της διάφορων πειραμάτων προσομοίωσης δείχνει ότι οι αιτήσεις για συστάσεις και οι αντίστοιχες απαντήσεις εισάγουν μικρό φόρτο επικοινωνίας. Αυτό συμβαίνει επειδή οι κόμβοι αιτούνται συστάσεις μόνο από αξιόπιστους κόμβους. Επιπλέον, με την πάροδο του χρόνου οι κόμβοι αρχίζουν να διατηρούν άμεσες αλληλεπιδράσεις, και έτσι οι συστάσεις δεν είναι ουσιαστικές. Οι προσομοιώσεις έδειξαν επίσης ότι το πλαίσιο *ATF* επιτρέπει στους κόμβους να προσδιορίζουν γρήγορα τις τιμές αξιοπιστίας των άλλων κόμβων. Κατά συνέπεια, το πλαίσιο *ATF* παρέχει ικανοποιητικά μέσα στους δίκαιους κόμβους για γρήγορο προσδιορισμό και απομόνωση των εγwissτικών ή κακόβουλων κόμβων.

Τέλος, το ATF εξασφαλίζει υψηλό ποσοστό επιτυχίας στην πρόβλεψη της συμπεριφοράς των κόμβων, δεδομένου ότι κάθε αξιόπιστος κόμβος προσδιορίζει έναν αναξιόπιστο με μεγάλη πιθανότητα.

7.2 Μελλοντικές Κατευθύνσεις

Μελλοντικές κατευθύνσεις της έρευνας που διενεργήθηκε στα πλαίσια της παρούσας διατριβής μπορούν να προκύψουν με αφετηρία κάθε ένα από τα επιμέρους προβλήματα που μελετήθηκαν. Συγκεκριμένα:

Στο πεδίο της επίγνωσης πλαισίου όσον αφορά το χρήστη ως μελλοντική κατεύθυνση ερευνούμε δύο ζητήματα που θα επηρεάσουν άμεσα το προτεινόμενο σύστημα στην απόδοση και την κλιμάκωσή του. Το πρώτο ζήτημα είναι η χρήση της τεχνικής στίγματος εξ αναμετρήσεως (dead reckoning - DR) για να βελτιωθεί η ακρίβεια και η ορθότητα που το σύστημα παρέχει. Γενικά η τεχνική dead reckoning είναι η διαδικασία υπολογισμού της θέσεως ενός κινούμενου αντικειμένου και βασίζεται στην ταχύτητα και την κατεύθυνση του αντικειμένου και του χρόνου που παρήλθε από την τελευταία γνωστή του θέση. Στην περίπτωση μας, παραδείγματος χάριν, η κινητή συσκευή ενός χρήστη που είναι εξοπλισμένη με μια ηλεκτρονική πυξίδα και ένα επιταχυνσιόμετρο θα μπορούσε να παρέχει τις πληροφορίες για την κατεύθυνση και την ταχύτητα του κατόχου του. Λαμβάνοντας επίσης υπόψη την τελευταία γνωστή θέση του χρήστη και το χρόνο που παρήλθε από τότε, μπορούμε να προβλέψουμε την τρέχουσα θέση. Φυσικά για τον ακριβέστερο υπολογισμό της θέσης θα πρέπει να συμπεριλάβουμε τις τρέχουσες τιμές από όλους τους αισθητήρες.

Το δεύτερο ζήτημα το οποίο ερευνούμε είναι η υιοθέτηση μιας κατανεμημένης αρχιτεκτονικής για το σύστημα. Όπως αναφέραμε προηγουμένως η εκτίμηση θέσης πραγματοποιείται σε έναν κεντρικό υπολογιστή. Στην προτεινόμενη κατανεμημένη αρχιτεκτονική το κτήριο διαιρείται σε περιοχές που το μέγεθός τους εξαρτάται από τη διάσταση ολόκληρου του κτηρίου και τον αριθμό των αισθητήρων. Για κάθε μία περιοχή υπάρχει ένας κεντρικός υπολογιστής αρμόδιος για την εκτίμηση θέσης. Οι κεντρικοί υπολογιστές των παρακείμενων περιοχών είναι διασυνδεδεμένοι προκειμένου να ανταλλάσσουν πληροφορίες για τους χρήστες (handovers μεταξύ των περιοχών, κ.λπ.). Η κατανεμημένη προσέγγιση ενισχύει την απόδοση του συστήματος, βελτιώνει την κλιμάκωσή του και τον καθιστά πιο εύρωστο σε περίπτωση αποτυχίας κάποιων κεντρικών υπολογιστών.

Επίσης, ως μελλοντική εργασία ερευνούμε την γενίκευση του συστήματος που βασίζεται στην Ασαφή Λογική έτσι ώστε η διαδικασία σύντηξης να περιλαμβάνει περαιτέρω πληροφορίες ακόμα και το πώς επιδρούν οι ίδιοι οι χρήστες με τα χαρακτηριστικά τους στην επίγνωση πλαισίου. Η περαιτέρω έρευνα στοχεύει στην εξέταση της απόδοσης του συστήματος που συμπεραίνει ταυτόχρονα πολλαπλά «κομμάτια» του πλαισίου (θέση, κατάσταση χρήστη, κλπ.). Τέλος, η έρευνά μας εστιάζεται και στην ολοκλήρωση πρόσθετων πληροφοριών για τον υπολογισμό της αξιοπιστίας των αισθητήρων, π.χ., χρόνος λειτουργίας αισθητήρων, τιμές από γειτονικούς αισθητήρες, κτλ.

Στο πεδίο της επίγνωσης κατάστασης του περιβάλλοντος ως μελλοντική κατεύθυνση προτείνουμε την βελτίωση των προτεινομένων αλγορίθμων με εναλλακτικούς κανόνες συνδυασμού, π.χ. [49], και την υιοθέτηση της Ασαφούς Λογικής (Fuzzy Set Theory) για να εξετάσουμε την αβεβαιότητα, την ανακρίβεια και την ημιτέλεια (incompleteness) των δεδομένων που λαμβάνουμε από τους αισθητήρες. Επίσης προς αυτήν την κατεύθυνση είναι και η εξέταση του προτεινόμενου συστήματος πυρανίχνευσης μέσω των πραγματικών δεδομένων. Τέλος, θα πραγματοποιηθούν δοκιμές με ελεγχόμενη πυρκαγιά έτσι ώστε να υπολογιστούν διάφορες παράμετροι όπως τα κατώφλια, το ποσοστό ψευδών συναγερμών, κτλ.

Στο πεδίο της βιο-μιμητικής διάδοσης της πληροφορίας πλαισίου ως μελλοντική κατεύθυνση της έρευνας αποτελεί η εκτέλεση προσομοιώσεων προκειμένου να βελτιστοποιηθούν ορισμένα μέρη των αλγορίθμων, όπως ο έλεγχος συμφόρησης και η προσαρμογή του αριθμού των μεταδόσεων με βάση την κινητικότητα. Στα πλαίσια της μελλοντικής έρευνας συγκαταλέγεται και η αξιολόγηση των προτεινόμενων σχημάτων διάδοσης με διαφορετικές τεχνικές ανίχνευσης αλλαγής στα ρεύματα δεδομένων σύμφωνα με τις υπολογιστικές ικανότητες κάθε κόμβου. Επιπλέον, μελετάται μια ενισχυμένη έκδοση των αλγορίθμων προσαρμοστικής επιδημικής διάδοσης όπου ο ρυθμός μόλυνσης β εξαρτάται επίσης και από την κινητικότητα των κόμβων. Τέλος, σε αυτήν την έκδοση η εμβέλεια μετάδοσης r είναι ρυθμιζόμενη ανάλογα με την αλλαγή που παρατηρείται στο ρεύμα δεδομένων αυξάνοντας ή μειώνοντας τον αριθμό γειτόνων που μολύνονται με την πληροφορία πλαισίου.

Ως μελλοντική εργασία στο πεδίο της οικοδόμησης εμπιστοσύνης μεταξύ των κόμβων ενός δικτύου, συγκαταλέγονται σενάρια προσομοίωσης χρησιμοποιώντας μία διαφορετική οργάνωση δικτύων, δηλ., λιγότερο πυκνό δίκτυο, άλλες μορφές κινητικότητας και αλλαγή των παραμέτρων ώστε να προσδιορίσουμε πώς το

προτεινόμενο σύστημα *ATF* αποδίδει κάτω από αυτά τα σενάρια. Τέλος, αυτές οι προσομοιώσεις στοχεύουν στην μέτρηση της επίδραση του αισθητήρα *RFTS*, στην αξιολόγηση του φόρτου επικοινωνίας που αυτός εισάγει και στη μέτρηση της κατανάλωσης ενέργειας.

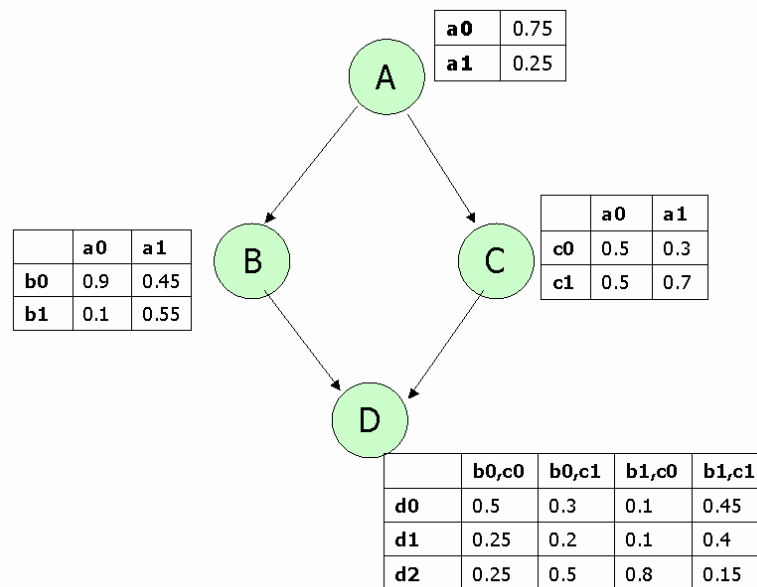
ΚΕΦΑΛΑΙΟ 8 ΠΑΡΑΡΤΗΜΑ

8.1 Δίκτυα Bayes

Τα Δίκτυα Bayes (Bayesian Networks) είναι ένα στατιστικό εργαλείο που έχει γίνει πολύ δημοφιλές τα τελευταία χρόνια σε περιοχές όπως και η πιθανοτική μάθηση και η τεχνητή νοημοσύνη ([10], [11]). Είναι κατάλληλα για πρόβλεψη (prediction) και συμπερασμό (inference) διότι μοντελοποιούν αιτιώδης επιδράσεις (causal influences) μεταξύ τυχαίων μεταβλητών.

Ένα Δίκτυο Bayes (Σχήμα 8.1) αποτελείται από δύο μέρη:

Το πρώτο μέρος είναι ένας κατευθυνόμενος ακυκλικός γράφος (directed acyclic graph-DAG). Οι κόμβοι (nodes) του γράφου αναπαριστούν μεταβλητές (π.χ., τυχαίες συνεχείς ή διακεκριμένες μεταβλητές), ενώ οι σύνδεσμοι (arcs) που ενώνουν τους κόμβους αναπαριστούν μια πιθανοτική σχέση μεταξύ αυτών των μεταβλητών.



Σχήμα 8.1: Δίκτυο Bayes (γράφος και πιθανοτικές κατανομές).

Αν υπάρχει σύνδεσμος από τον κόμβο A προς τον κόμβο B τότε λέμε ότι ο B επηρεάζεται άμεσα από τον A και ο A είναι πατέρας (ή άμεσος πρόγονος) του B ³. Οι άμεσοι πρόγονοι ενός κόμβου X συμβολίζονται ως $Parents(X)$. Στο δίκτυο που απεικονίζεται στο Σχήμα 8.1 είναι:

$$Parents(D) = \{B, C\}$$

³ Μία άλλη έκφραση που χρησιμοποιείται είναι ότι ο A αποτελεί το αίτιο και ο B το αποτέλεσμα.

Ένα άλλο χαρακτηριστικό ενός Δικτύου Bayes είναι ότι κάθε κόμβος του είναι υπό συνθήκη ανεξάρτητος (conditionally independent) από οποιονδήποτε άλλον, δεδομένων των άμεσων προγόνων του.

Π.χ. για το δίκτυο στο Σχήμα 8.1 ισχύει:

$$P(D | A, B, C) = P(D | B, C)$$

Το δεύτερο μέρος ενός Δικτύου Bayes αποτελείται από ένα σύνολο υπό συνθήκη πιθανοτικών κατανομών (conditional probability distributions) που σχετίζονται με κάθε κόμβο του δικτύου (κάθε τυχαία μεταβλητή). Αυτές οι κατανομές περιγράφουν την πιθανότητα όλων των αποτελεσμάτων μιας τυχαίας μεταβλητής, δεδομένων όλων των πιθανών τιμών των άμεσων προγόνων της. Οι παράμετροι αυτών των κατανομών προσδιορίζονται με βάση ήδη υπάρχοντα στατιστικά στοιχεία και δεδομένα στα οποία εφαρμόζονται τεχνικές μάθησης (Bayesian Network learning techniques) [11].

Ο κατευθυνόμενος ακυκλικός γράφος και οι πιθανοτικές κατανομές μαζί ορίζουν την από κοινού πιθανοτική κατανομή (joint probability distribution). Έτσι γενικά για ένα Δίκτυο Bayes που έχει τους εξής κόμβους:

$$X_1, X_2, \dots, X_N,$$

η από κοινού πιθανοτική κατανομή (ή αλλιώς από κοινού πιθανότητα) περιγράφεται σαν το γινόμενο των επιμέρους πιθανοτικών κατανομών:

$$P[X_i | Parents(X_i)] \text{ για } i=1 \dots N$$

Αν το X δεν έχει προγόνους τότε η πιθανοτική κατανομή του λέγεται ότι είναι απόλυτη (unconditional). Η από κοινού πιθανοτική κατανομή για το δίκτυο της Εικόνας 4.3 υπολογίζεται ως εξής:

$$P(A, B, C, D) = P(A) * P(B | A) * P(C | A) * P(D | B, C)$$

Μπορούμε να δούμε ότι οι σχέσεις ανεξαρτησίας μας επιτρέπουν να παρουσιάσουμε αυτήν την κατανομή με μια πιο συμπαγή μορφή. Εδώ οι κόμβοι (μεταβλητές) είναι λίγοι, αλλά γενικά, εάν είχαμε n δυαδικούς κόμβους, ο υπολογισμός της πλήρους από κοινού πιθανότητας (full joint probability) θα απαιτούσε χώρο $O(2^n)$. Έχοντας όμως γνώση των εξαρτήσεων μεταξύ των μεταβλητών απαιτείται χώρος $O(n2^k)$, όπου το k είναι το μέγιστο fan-in (μέγιστος αριθμών εισόδων) ενός κόμβου. Έτσι λοιπόν γίνεται κατανοητό ότι η χρησιμοποίηση ενός Δικτύου Bayes πλεονεκτεί σε σχέση με τις παραδοσιακές πιθανοτικές μεθόδους διότι δεν είναι απαραίτητος ο υπολογισμός όλων των πιθανοτήτων για να ληφθεί μία απόφαση.

Πως γίνεται όμως μία πρόβλεψη ή ένας συμπερασμός με τη χρήση ενός τέτοιου δικτύου; Ας επανέλθουμε στο δίκτυο που απεικονίζεται στο Σχήμα 8.1. Θεωρούμε ότι κάποια χρονική στιγμή θέλουμε να υπολογίσουμε την πιθανότητα η μεταβλητή D να πάρει την τιμή $d1$, δεδομένου ότι γνωρίζουμε τις τιμές των υπόλοιπων τριών μεταβλητών. Οι τιμές αυτές ονομάζονται επίσης και παρατηρήσεις (observations). Είναι λοιπόν:

$$\text{παρατηρήσεις} = \{A=a0, B=b1, C=c0\}.$$

Σύμφωνα με όσα αναφέρθηκαν προηγουμένως για τις σχέσεις ανεξαρτησίας μεταξύ των μεταβλητών ενός Δικτύου Bayes για τον υπολογισμό της εν λόγω πιθανότητας θα έχουμε:

$$P(D=d1 | A=a0, B=b1, C=c0) = \\ \alpha * P(A=a0) * P(B=b1 | A=a0) * P(C=c0 | A=a0) * P(D=d1 | B=b1, C=c0)$$

όπου το α παίζει το ρόλο του παράγοντα κανονικοποίησης. Αν αντικαταστήσουμε τις τιμές που έχουμε στις πιθανοτικές κατανομές κάθε μεταβλητής (βλ. πίνακες δίπλα σε κάθε κόμβο) τότε είμαστε σε θέση να υπολογίσουμε την ζητούμε πιθανότητα.

Τα Δίκτυα Bayes αποτελούν ένα χρήσιμο εργαλείο για τη μοντελοποίηση της γνώσης σε τομείς όπως η ιατρική, η μηχανική, η σύντηξη δεδομένων, τα συστήματα λήψης αποφάσεων, κ.α. Χρησιμοποιούνται εκτενώς σε διαδικασίες πρόβλεψης και συμπερασμού καταστάσεων.

8.2 Ασαφής Λογική

Μια λογική βασισμένη στις τιμές true και false μπορεί να αποδειχθεί ανεπαρκής για την περιγραφή του πλαισίου (context). Η Ασαφής Λογική (Fuzzy Logic) είναι η λογική που χρησιμοποιεί ολόκληρο το διάστημα μεταξύ του 0 (false) και του 1 (true) για να περιγράψει το context, παρέχοντας έτσι μεγαλύτερες δυνατότητες αναπαράστασης και χρησιμοποιείται συνήθως σε τομείς όπως η αναπαράσταση ανθρώπινου συλλογισμού (human reasoning) και η τεχνολογία αυτόματων ελεγκτών (fuzzy controllers). Οι παράμετροι movement, light, sound και at_hand του activity context αναπαρίστανται με fuzzy sets (ασαφή σύνολα). Έτσι, μια ασαφής τιμή movement δεν είναι απαραίτητα είτε walking, είτε walking_fast, είτε running, είτε halt, αφού μπορεί ο χρήστης να κινείται με ένα ρυθμό που δεν καθιστά εύκολη τη διάκριση ανάμεσα σε walking και walking_fast. Μπορεί να είναι κατά ένα ποσοστό walking, κατά ένα άλλο walking fast, κατά ένα τρίτο running και κατά ένα τέταρτο halt, με τον περιορισμό ότι το συνολικό άθροισμα των

ποσοστών αυτών να είναι ίσο με τη μονάδα. Για παράδειγμα μια τιμή movement είναι η ακόλουθη :

$$\text{movement} = \{ 0.5/\text{walking} + 0.4/\text{walking_fast} + 0.1/\text{running} + 0.0/\text{halt} \},$$

όπου το σύμβολο '+' αναπαριστά την πράξη της ασαφούς πρόσθεσης.

Ένα ασαφές σύνολο (*fuzzy set*) ορίζεται για να περιγράψει μια ανακριβή / ελλιπή / αβέβαια έννοια όπως, π.χ., ύψος, ταχύτητα, χρονική διάρκεια. Το κάθε στοιχείο x ενός διακριτού συνόλου D μπορεί να έχει διαφορετικούς βαθμούς συμμετοχής σε ένα ασαφές σύνολο A μέσω μιας συνάρτησης βαθμού συγγένειας (*membership function*), έτσι ώστε

$$\mu_A: D \rightarrow [0, 1]$$

Ένα ασαφές σύνολο A μπορεί να αναπαρασταθεί από ένα πεπερασμένο (αριθμήσιμο ή όχι) σύνολο ζευγαριών $(x, \mu(x))$, δηλαδή,

$$A = \{(x, \mu_A(x)) \mid \mu_A(x) \in [0, 1] \text{ και } x \in D\}$$

Στην δυαδική λογική η μονάδα συγγένειας λαμβάνει τιμές 0 ή 1. Ένα σύνολο C λέγεται *core* όταν $\mu_C(x) = 1, \forall x \in C$. Από ένα ασαφές σύνολο μπορούμε να ορίσουμε τις εξής ιδιότητες:

support of A (sup(A)): Είναι το crisp σύνολο όπου $\text{sup}(A) = \{x \mid \mu_A(x) > 0\}$

core of A (cor(A)): Είναι το crisp σύνολο όπου $\text{cor}(A) = \{x \mid \mu_A(x) = 1\}$

a-cut of A (A_a): Είναι το crisp σύνολο όπου $A_a = \{x \mid \mu_A(x) \geq a\}$

height of A (h_A): Είναι η μέγιστη τιμή $h_A = \max_x \{\mu_A(x)\}$

Ορίζονται διάφοροι τελεστές για τα ασαφή σύνολα. Συγκεκριμένα, οι νόρμες T και S μπορούν να θεωρηθούν ότι αντικαθιστούν του λογικούς δυαδικούς τελεστές *and* και *or* αντιστοίχως. Ένας δυαδικός τελεστής

$$t: [0, 1] \times [0, 1] \rightarrow [0, 1]$$

καλείται τριγωνική νόρμα ή εν συντομία *T-norm* όταν είναι commutative, associative, non-decreasing, έτσι ώστε για κάθε $a, b, c, d \in [0, 1]$ ισχύει ότι

$$atb = bta$$

$$at(btc) = (atb)tc$$

$$atb \leq ctd \text{ οποτεδήποτε } a \leq c \text{ και } b \leq d$$

$$at1 = a$$

Ο δυαδικός τελεστής $s: [0, 1] \times [0, 1] \rightarrow [0, 1]$ για τον οποίο ισχύουν οι ιδιότητες από 1 έως 3 ονομάζεται τριγωνική conorm ή εν συντομία *T-conorm* ή *S-norm* έτσι ώστε

$$as0 = 0$$

Κυρίως συνηθίζεται η *T-norm* να είναι το ασαφές \wedge και η *S-norm* είναι το ασαφές \vee μεταξύ ασαφών συνόλων. Αυτό σημαίνει ότι για δύο ασαφή σύνολα A και B οι ασαφείς τελεστές παράγουν ένα νέο ασαφές σύνολο ώστε:

$$A \wedge B = \{(x, \min_u \{\mu_A(u), \mu_B(u)\}), x \in D\}$$

$$A \vee B = \{(x, \max_u \{\mu_A(u), \mu_B(u)\}), x \in D\}$$

$$\neg A = \{(x, 1 - \mu_A(x)), x \in D\}$$

$$A \rightarrow B = \{(x, \max(1 - \mu_B(x), \mu_A(x))), x \in D\}$$

Εάν η μεταβλητή που θέλουμε να χαρακτηρίσουμε μπορεί να πάρει τιμές που μπορούν να αναπαρασταθούν με ασαφή σύνολα (που έχουν λογική σημασία) τότε η μεταβλητή αυτή ονομάζεται (γλωσσική μεταβλητή) *linguistic variable* και οι τιμές (γλωσσική τιμή) *linguistic values*. Για τον λόγο αυτό, κάθε μία ασαφής τιμή μιας ασαφούς μεταβλητής μπορεί να χαρακτηρίζεται από ένα σύνολο γλωσσικών όρων (*linguistic terms*), όπως π.χ., οι γλωσσικοί όροι *υψηλή*, *μεγάλο* και *τεράστια* στις εκφράσεις *υψηλή ταχύτητα*, *μεγάλο ύψος*, *τεράστια χρονική διάρκεια*.

Ένας ασαφής κανόνας αποτελείται από δύο μέρη: από το antecedent μέρος που υποδηλώνει τις συνθήκες για να ενεργοποιηθεί ο κανόνας (προγενέστερες συνθήκες) και το consequent μέρος που υποδηλώνει τα επακόλουθα του κανόνα. Κυρίως ένας κανόνας αποτελεί μια γλωσσική αναπαράσταση μιας ασαφούς συνεπαγωγής $A \rightarrow B = \neg A \vee B$ που ερμηνεύεται ως: Αν ισχύει η ασαφής παράσταση A τότε θα συνεπάγεται και η ασαφής παράσταση B .

Η πιο τυπική μορφή κανόνα είναι η σύζευξη ασαφών συνόλων που ερμηνεύονται από *linguistic variables* και *linguistic values*. Ένας ασαφής κανόνας R_j έχει την μορφή:

$$R_j: \text{If } V_{1j} \text{ is } A_{1j} \text{ and } V_{2j} \text{ is } A_{2j} \text{ and } \dots \text{ and } V_{nj} \text{ is } A_{nj} \text{ Then } C_j \text{ is } B_j$$

όπου V_{kj} είναι η k -οστή ασαφή μεταβλητή για τον κανόνα R_j που έχει ασαφή τιμή A_{kj} και τότε το επακόλουθο του κανόνα είναι C_j με ασαφή τιμή B_j . Μια βάση γνώσης με ένα σύνολο ασαφών κανόνων καλείται Ασαφής Βάση Γνώσης (ΑΒΓ). Οι ασαφείς κανόνες

χρησιμοποιούνται για ασαφείς συμπερασμούς στους οποίους παράγεται τελικά μια καθορισμένη crisp τιμή, όπως ορίζεται στην επόμενη παράγραφο.

Ο Ασαφής Συμπερασμός (ΑΣ) είναι η διαδικασία εκτίμησης των ασαφών κανόνων σε μια ABΓ, δηλαδή, ο συνδυασμός των εξόδων του κάθε κανόνα βάσει των αντίστοιχων εισόδων. Η διαδικασία του ΑΣ περιέχει τρία στάδια:

Διαδικασία Ενσωμάτωση Ασάφειας (fuzzification): Η δημιουργία συνάρτησης βαθμού συγγένειας και αναπαράσταση της εισόδου ως ασαφές σύνολο, κυρίως όταν η είσοδος είναι crisp. Παρόλα αυτά, η είσοδος μπορεί να είναι ήδη ένα ασαφές σύνολο που μπορεί να προέρχεται από κάποια έξοδο ενός άλλου ΑΣ.

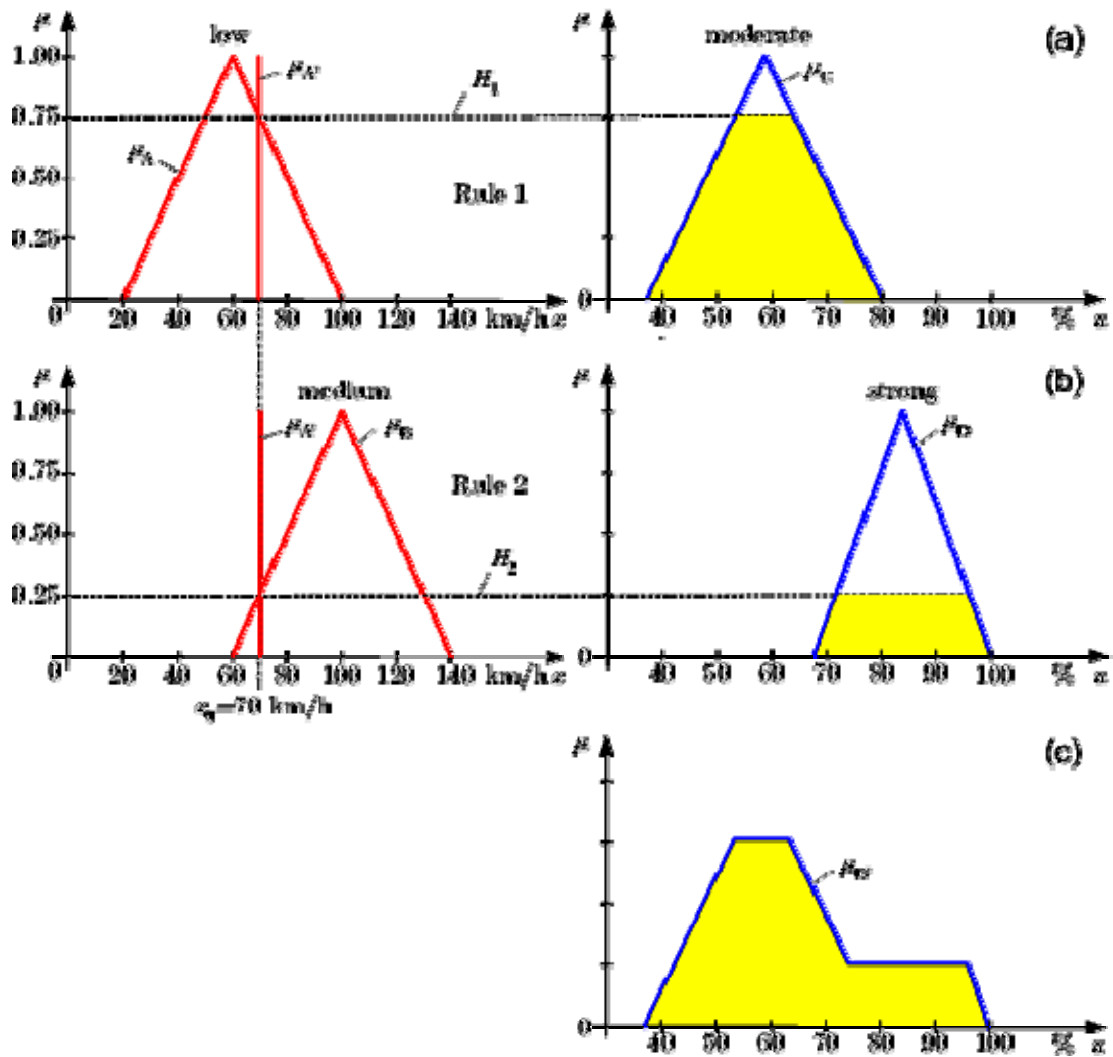
Συμπερασμός (inference): Ο υπολογισμός του ποσοστού ικανοποίησης όλων των κανόνων για την παραγωγή συμπεράσματος και η απόφαση / προσδιορισμός του ασαφούς συνόλου εξόδου. Ο υπολογισμός αυτός βασίζεται στην εφαρμογή των T-norm και S-norm ανάμεσα των συνόλων των «antecedent» και «consequent» μερών των ασαφών κανόνων από μια καθορισμένη ABΓ.

Αποσαφήνιση (defuzzification): Εάν απαιτείται η έξοδος του ΑΣ να αναπαρασταθεί σε crisp μορφή τότε εφαρμόζονται διάφοροι μέθοδοι αποσαφήνισης, όπως , που μετατρέπουν την ασαφή έξοδο σε διακριτή έξοδο

Στο Σχήμα 8.2 απεικονίζεται ο ΑΣ όταν η είσοδος στο σύστημα είναι $x_0 = 70\text{Km/h}$ με χρήση max / min συμπερασμού και οι κανόνες στην ABΓ είναι:

R_1 : **If** *velocity is low* **Then** *risk is moderate*

R_2 : **If** *velocity is medium* **Then** *risk is strong*



Σχήμα 8.2: Διαδικασία Ασαφούς Συμπερασμού

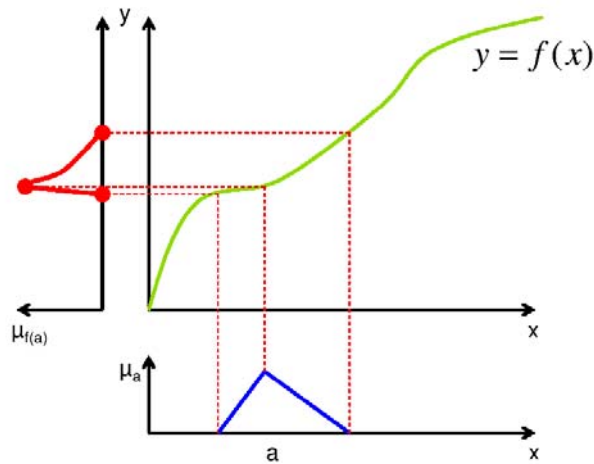
Η αρχή της επέκτασης (extension principle) περιγράφει την λειτουργική εξάρτηση μεταξύ ασαφών αριθμών και συναρτήσεων. Η αρχή αυτή επιτρέπει την εφαρμογή πραγματικών συναρτήσεων, έστω f , σε ασαφείς αριθμούς (σε ασαφείς σύνολα) έτσι ώστε για μια τιμή $x \in \mathfrak{R}$ παράγεται ένας βαθμός συγγένειας για την εικόνα του y που είναι ο μέγιστος από τους βαθμούς συγγένειας από εκείνα τα x , έτσι ώστε $y = f(x)$, δηλαδή,

$$\mu_{f(a)}(y) = \max\{\mu_a(x) \mid \forall x : f(x) = y\},$$

και στην περίπτωση που η συνάρτηση $f(x_1, \dots, x_n)$ έχει πολλές μεταβλητές τότε,

$$\mu_{f(a_1, \dots, a_n)}(y) = \max\{\min\{\mu_{a_1}(x_1), \dots, \mu_{a_n}(x_n)\} \mid \forall x : f(x_1, \dots, x_n) = y\}.$$

Στο Σχήμα 8.3 απεικονίζεται η αρχή της επέκτασης.



Σχήμα 8.3: Αρχή της Επέκτασης

8.3 Φάροι υπερύθρων

Ο φάροι υπερύθρων (Infrared beacons) εκπέμπουν συνεχώς αναγνωριστικά ή/και άλλες πληροφορίες σε μικρές αποστάσεις χρησιμοποιώντας συνήθως το πρότυπο επικοινωνίας IRDA [97]. Εγκατεστημένοι σε σταθερά και γνωστά σημεία ενός κτιρίου χρησιμοποιούνται για τον προσδιορισμό της θέσης των χρηστών.

Στη συνέχεια του παραρτήματος αυτού θα παρουσιάσουμε συνοπτικά τα χαρακτηριστικά του φάρου υπερύθρων που χρησιμοποιήσαμε στην αξιολόγηση των συστημάτων που παρουσιάστηκαν στο ΚΕΦΑΛΑΙΟ 2. Ο φάρος, που κατασκευάζεται από την εταιρεία Lesswire AG [98], απεικονίζεται παρακάτω.



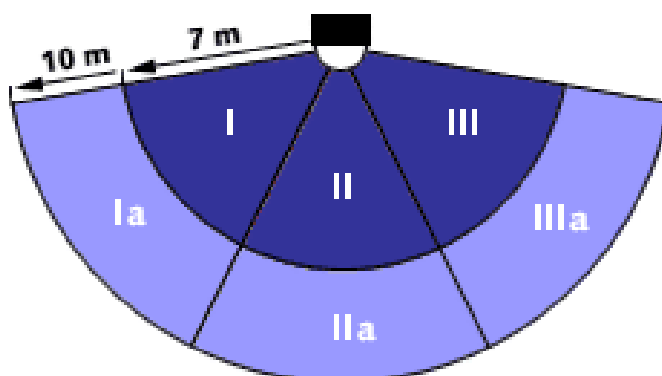
Σχήμα 8.4: Φάρος υπερύθρων της εταιρείας Lesswire AG.

Τα κύρια χαρακτηριστικά του είναι τα εξής:

- Εκπέμπει αναγνωριστικά ή/και άλλα δεδομένα σε μικρές αποστάσεις (οι ανακλάσεις είναι χρήσιμες).

- Υποστηρίζει ευρεία γωνία κάλυψης (οριζόντια και κάθετα).
- Είναι εύκολα προγραμματίσιμος μέσω υπερύθρων.
- Η ισχύς εκπομπής και η κατεύθυνσή της είναι προγραμματίσιμα χαρακτηριστικά.
- Έχει προγραμματίσιμη περίοδο εκπομπής.
- Υποστηρίζει χαμηλή κατανάλωση ισχύος.
- Τροφοδοτείται από εξωτερική πηγή ή/και εσωτερική επαναφορτιζόμενη μπαταρία
- Έχει ενσωματωμένα 2 LED που χρησιμοποιούνται για ενδείξεις σωστής λειτουργίας.
- Έχει ενσωματωμένη μνήμη (flash memory) για εύκολη αναβάθμιση του λογισμικού του.
- Η θήκη του είναι μικρή και ελαφριά.

Τα 6 προγραμματίσιμα τμήματα που καλύπτονται από τον φάρο φαίνονται στην παρακάτω εικόνα.



Σχήμα 8.5: Προγραμματίσιμα τμήματα (περιοχές) που υποστηρίζονται από το φάρο.

Τεχνικά χαρακτηριστικά

| | |
|----------------------------|---|
| Απόσταση | Προγραμματίσιμη (> 7 ή > 10 m, προαιρετικά > 3 m) |
| Γωνία, οριζόντια | >165° (3 ξεχωριστές προγραμματίσιμες περιοχές) |
| Γωνία, κάθετα | > 55° |
| Κατανάλωση ισχύος | < 3 mA (Εκπομπή 50 bytes/second) |
| Εξωτερική τάση τροφοδοσίας | 12 ... 15 V |
| Μπαταρία (προαιρετική) | 3 x AA, επαναφορτιζόμενη |
| Εξωτερική τροφοδοσία | Προσαρμογέας 12 V DC |
| Περίοδος εκπομπής | Προγραμματίσιμη (1s ... 15s) |
| Ρυθμός εκπομπής | 9.6 kbit/s |

ΟΡΟΛΟΓΙΑ

| | |
|--|---|
| Basic Probability Assignment | Βασική Πιθανότητα Καταχώρισης |
| Bayesian Network | Δίκτυο Bayes |
| Collaborative context awareness | Συνεργατική επίγνωση πληροφορίας πλαισίου |
| Context | Πλαίσιο |
| Context awareness | Επίγνωση του πλαισίου |
| Context fusion | Σύντηξη πληροφορίας πλαισίου |
| Context interpretation | Ερμηνεία πλαισίου |
| Context management | Διαχείριση πλαισίου |
| Context modelling | Μοντελοποίηση πλαισίου |
| Context reasoning | Συμπερασμού πλαισίου |
| Context representation | Αναπαράσταση πλαισίου |
| Context-aware application | Εφαρμογή επίγνωσης πλαισίου |
| Context-aware system | Σύστημα επίγνωσης πλαισίου |
| Contextual information | Πληροφορία πλαισίου |
| Coverage | Κάλυψη |
| Data fusion | Σύντηξη δεδομένων |
| Decision fusion | Σύντηξη αποφάσεων |
| Defuzzification | Διαδικασία αποσαφήνισης |
| Density | Πυκνότητα |
| Distributed systems | Κατανεμημένα συστήματα |
| Dynamic Bayesian Network | Δυναμικό Δίκτυο Bayes |
| Dynamic Data Driven Application System | Δυναμικό Σύστημα Εφαρμογής Οδηγούμενο από τα Δεδομένα |
| Epidemic model | Επιδημικό μοντέλο |
| False alarm rate | Ποσοστό ψευδών συναγερμών |
| Flooding | Πλημμύρα |
| Fuzzification | Διαδικασία ενσωμάτωση ασάφειας |
| Fuzzy controller | Ασαφής ελεγκτής |
| Fuzzy inference | Ασαφής Συμπερασμός |
| Fuzzy Logic | Ασαφής Λογική |
| Fuzzy Sets Theory | Θεωρία Ασαφών Συνόλων |
| Gossip based algorithms | Αλγόριθμοι διάδοσης φήμης |
| Hardware technology | Τεχνολογία υλικού |
| Infected | Μολυσμένος |

| | |
|---------------------------------|---|
| Inference | Συμπερασμός |
| Linguistic value | Γλωσσική τιμή |
| Linguistic variable | Γλωσσική μεταβλητή |
| Location based services | Εξαρτώμενες από τη θέση υπηρεσίες |
| Location-aware applications | Εφαρμογές υπηρεσιών θέσης |
| Maximum likelihood estimation | Εκτίμηση μέγιστης πιθανοφάνειας |
| Mobile Computing | Κινητός Υπολογισμός |
| Mobility | Κινητικότητα |
| Neighbourhood discovery schemes | Αλγόριθμοι ανακαλύψεως της γειτονιάς |
| Nodes | Κόμβοι |
| Percolation probability | Πιθανότητα διείσδυσης |
| Pervasive Computing | Διάχυτος Υπολογισμός |
| Pignistic Probability | Πιθανότητα Απόφασης |
| Proactivity | Προδραστικότητα |
| Proximity | Εγγύτητα |
| Recommendation Function | Υπηρεσία Σύστασης |
| Reputation | Φήμη |
| Reputation Manager | Διαχειριστής Φήμης |
| Self evolving trust | Αυτοεξελισσόμενη καθιέρωση εμπιστοσύνης |
| Selfish nodes | Ιδιοτελής κόμβοι |
| Situation | Κατάσταση |
| Sleep deprivation torture | Τεχνική της στέρησης ύπνου |
| Smart spaces | Ευφυείς χώροι |
| Susceptible | Εύρωτος /επιρρεπής |
| Transferable Belief Model | Μοντέλο Μεταβιβάσιμης Πεποίθησης |
| Trust | Αξιοπιστία |
| Trust Builder | Δημιουργός αξιοπιστίας |
| Trust Policy | Πολιτική εμπιστοσύνης |
| Trust Sensors | Αισθητήρες αξιοπιστίας |
| Trusted recommenders | Αξιόπιστοι συστήνοντες |
| Variability | Μεταβλητότητα |
| Wireless LANs | Ασύρματα δίκτυα |
| Membership function | Συνάρτηση βαθμού συγγένειας |
| Trust building | Οικοδόμηση εμπιστοσύνης |

ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ

| | |
|-------|--|
| ATF | Ad hoc Trust Framework |
| BPA | Basic Probability Assignment |
| CSMA | Carrier Sense Multiple Access |
| CUSUM | Cumulative Sum |
| DBN | Dynamic Bayesian Network |
| DDDAS | Dynamic Data Driven Application System |
| DE | Direct evidence |
| DoS | Denial of Service |
| DST | Dempster-Shafer Theory |
| IR | Infrared Radiation |
| IRDA | Infrared Data Association |
| LBS | Location based services |
| MAC | Media Access Control |
| MANET | Mobile Ad hoc Network |
| PDA | Personal Digital Assistant |
| RFID | Radio Frequency IDentification |
| RFTS | Recommendation Function Trust Sensor |
| RM | Reputation Manager |
| RSS | Received Signal Strength |
| RWP | Random Waypoint |
| SI | Susceptible-Infected |
| SIR | Susceptible-Infected-Recovered |
| SIS | Susceptible-Infected-Susceptible |
| TB | Trust Builder |
| TBM | Transferable Belief Model |
| TS | Trust Sensors |
| TTL | Time-To-Live |
| US | Ultrasound |
| Wi-Fi | Wireless Fidelity |
| WLAN | Wireless Local Area Network |
| ΑΒΓ | Ασαφής Βάση Γνώσης |
| ΑΣ | Ασαφής Συμπερασμός |
| BN | Bayesian Network |
| ΔΒ | Δίκτυα Bayes |

| | |
|------|---|
| ΔΔΒ | Δυναμικά Δίκτυα Bayes |
| ΔΔΒΑ | Δυναμικών Bayes Δικτύων βάσει Αξιοπιστίας Πηγών |
| ΘΑΣ | Θεωρία Ασαφών Συνόλων |
| ΣΕΠ | Συνεργατική Επίγνωση Πλαισίου |

ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

1. Weiser M. : The computer for the 21st century. Scientific American, September 1991,66-75.
2. Katz, R.H., Long, D., Satyanarayanan, M., Tripathi, S. Workspaces in the Information Age. In Report of the NSF Workshop on Workspaces in the Information Age. Leesburg, VA, October, 1996. Available at <http://www.cs.berkeley.edu/~randy/NSFWS>.
3. Weiser, M, Brown, J.S. The Coming Age of Calm Technology. In Denning, P.J., Metcalfe, R.M. (editors), Beyond Calculation: The Next Fifty Years of Computing. Copernicus, 1998.
4. D. Abowd, Towards a Better Understanding of Context and Context-Awareness, in Proc. International Conference of Human Factors in Computing Systems, (Hague, Netherlands, 2000).
5. C. Anagnostopoulos, A. Tsounis, S. Hadjiefthymiades, "Context Awareness in Mobile Computing Environments", Wireless Personal Communications, vol. 42, pp. 445-464, 2006.
6. B.Schilit and M.Theimer, Disseminating active map information to mobile hosts, IEEE Network, 8(5), pp. 22–32, Sept. 1994.
7. P.Brezillon, M.Borges, J.Pino and J.Pomerol, Context-awareness in group-work: Three case studies, Decision Support Systems, 115–124, 2004.
8. C. Anagnostopoulos and S. Hadjiefthymiades, On the Application of Epidemical Spreading in Collaborative Context Aware Computing, ACM SIGMOBILE Mobile Computing and Communications, Volume 12 , Issue 4 , pp. 43-55, 2008.
9. Mihajlovic, V., Petkovic, M.: Dynamic Bayesian Networks: A State of the Art. Technical report (2001). Report no. TR-CTIT-01-34, Centre for Telematics and Information Technology, University of Twente, Enschede, The Netherlands.
10. Jensen F.: An Introduction to Bayesian Networks. SpringerVerlag, New York, 1996.
11. Heckerman, D.: A tutorial on learning with Bayesian networks. Technical Report MSR-TR-95-06, Microsoft Research ,1995.

12. A. Dempster, N. Laird, and D. Rubin. Maximum likelihood estimation from incomplete data. *Journal of the Royal Statistical Society (B)*, 39(1), 1977.
13. Castro, P., Chiu, P., Kremenek, T., Muntz, R.: A probabilistic room location service for wireless networked environments. In: *UbiComp 2001 Conference*, Atlanta 2001.
14. D. Nauck, R. Kruse, "NEFCLASS - A Neuro-Fuzzy Approach for the Classification of Data", *ACM Symposium on Applied Computing*, pp. 461-465, 1995.
15. Want, R., Hopper, A., Falcao, V., Gibbons, J.: *The Active Badge Location System*. *ACM Transactions on Information Systems*, 10, 91-102. 1992.
16. Bahl, P., Padmanabhan, V.: *RADAR: An In-Building RF-Based User Location and Tracking System*. *Proceedings of IEEE INFOCOM*, pp. 775-784 (Tel-Aviv, Israel), 2000.
17. Ekahau, Inc : Ekahau Positioning Engine. <http://www.ekahau.com/>
18. Ladd A., Bekris K., Rudys A., Kavraki L., Wallach D., Marceau G.: *Robotics-based location sensing using wireless ethernet*. In *Proceedings MOBICOM-02*, pages 227–238, New York, September 23–28, 2002.
19. Nissanka, B. Priyantha, A., Balakrishnan, H.: *The cricket location-support system*. In *Proceedings of MOBICOM 2000*, pages 32-43, Boston, MA, August, 2000.
20. Harter, A., Hopper, A., Steggles, P., Ward A., Webster P.: *The anatomy of a context-aware application*. In *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (Mobicom '99)*, 1999.
21. Graumann, D., Lara, W., Hightower, J., Borriello, G.: *Real-world implementation of the Location Stack: The Universal Location Framework*. In *Proceedings of the 5th IEEE Workshop on Mobile Computing Systems & Applications (WMCSA 2003)*, pp. 122-128, Oct. 2003.
22. King, T., Kopf, S., Effelsberg, W.: *A Location System based on Sensor Fusion: Research Areas and Software Architecture*. *Proc. of 2nd GI/ITG KuVS Fachgespräch "Ortsbezogene Anwendungen und Dienste"*, Stuttgart, Germany, June 2005.

23. T. Roos, P. Myllymaki, H. Tirri, P. Misikangas, J. Sievanen, "A Probabilistic Approach to WLAN User Location Estimation", *Wireless Information Networks*, vol. 9, no. 3, pp. 155-164, 2002
24. C. Anagnostopoulos, Y. Ntarladimas, S. Hadjiefthymiades, "Situational computing: an innovative architecture with imprecise reasoning", *Journal of Systems and Software*, Elsevier, vol. 80, no. 12, pp. 1993-2014, December, 2007.
25. P. Korpipaa, M. Koskinen, J. Peltola, S. Makela, T. Seppanen, "Bayesian approach to sensor-based context awareness", *Pervasive and Ubiquitous Computing*, vol. 7, pp. 113–124, 2003.
26. Anagnostopoulos C., Sekkas O., Hadjiefthymiades S., "Context Fusion: Dealing with Sensor Reliability", in *Proceedings of the 2nd International Workshop on Information Fusion and Dissemination in Wireless Sensor Networks (SensorFusion07 - MASS 2007)*, Piza, Italy, October 2007.
27. Sekkas O., Anagnostopoulos C., Hadjiefthymiades S., "Context Fusion through Imprecise Reasoning", in *Proceedings of the IEEE International Conference on Pervasive Services (ICPS 2007)*, pp.88-91, Istanbul, Turkey, July 2007.
28. Sekkas O., Hadjiefthymiades S., Zervas E., "Enhancing location estimation through data fusion", in *Proceedings of the 17th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06)*, pp.1-5, Helsinki, Finland, September 2006.
29. Sekkas O., Hadjiefthymiades S., Zervas E., "Fusing sensor information for location estimation", in *Proceedings of the 10th East-European Conference on Advances in Databases and Information Systems (ADBIS 2006)*, Thessaloniki, Greece, September 2006.
30. Sekkas O., Hadjiefthymiades S., Zervas E., "Sensor Data Fusion for Location Awareness", in *Encyclopedia of Mobile Computing and Commerce vol. 2* (Ed. David Taniar), February 2007, IDEA Group Reference.
31. Papataxiarhis V., Riga V., Nomikos V., Sekkas O., Kolomvatsos K., Tsetsos V., Papageorgas P., Vourakis S., Xouris V., Hadjiefthymiades S., Kouroupetroglou G., "MNISIKLIS: Indoor Location Based Services for All", in *Proceedings of the 5th International Symposium on LBS and TeleCartography*, Salzburg, Austria, November 26th-28th, 2008.

32. Chen, S., Bao, H., Zeng, X., & Yang, Y. "A Fire Detecting Method based on multi-sensor Data Fusion," In IEEE Systems Man and Cybernetics, volume 4, 2003.
33. Rose-Pehrsson, S. L., Hart, S. J., Street, T. T., Williams, F. W., Hammond, M. H., Gottuk, D. T., Wright, M. T., & Wong, J. T. "Early warning fire detection system using a probabilistic neural network," Fire Technol. 39, no. 2, pp. 147-171, 2003.
34. Rose-Pehrsson, S. L., Shaffer, R. E., Hart, S. J., Williams, F. W., Gottuk, D. T., Strehlen, B. D., & Hill, S. A. "Multi-criteria Fire Detection Systems Using a Probabilistic Neural network," Sensor Actuat. B-Chem. 69, no. 3, pp. 325-335, 2000.
35. Doolin, D. M., & Sitar., N. "Wireless Sensors for Wildfire Monitoring," In Smart Structures and Materials, volume 5765, 2005.
36. Calle, A., Casanova, J. L., & Romo, A., "Fire Detection and Monitoring using MSG Spinning Enhanced Visible and Infrared Imager (SEVIRI) Data," J. Geophys. Res., 111, G04S06, 2006
37. Sivathanu, C., & Tseng, L. K, "Fire Detection Using Near-IR Radiation and Source Temperature Discrimination," National Institute of Standards and Technology. Annual Conference on Fire Research, 1996
38. Kosucu, B., Irgan, K., Kucuk, G. & Baydere, S. "FireSenseTB: A Wireless Sensor Networks TestBed for Forest Fire Detection", in the 5th ACM International Wireless Communications and Mobile Computing Conference (IWCMC'09), Leipzig, Germany, 2009
39. Kucuk, G., Kosucu, B., Yavas, A. & Baydere, S. "FireSense: Forest Fire Prediction and Detection System using Wireless Sensor Networks" in the 4th IEEE/ACM International Conference on Distributed Computing in Sensor Systems (DCOSS'08) , Santorini Island, Greece, 2008
40. Hefeeda, M. & Bagheri, M. "Wireless Sensor Networks for Early Detection of Forest Fires", Pisa, Italy, in the Proceedings of the 1st International Workshop on Mobile Ad hoc and Sensor Systems for Global and Homeland Security (MASS-GHS), Pisa, Italy, 2007
41. Mandel, J., Beezley, J., Bennethum, L., Chakraborty, S., Coen, J., Douglas, C., L, Hatcher, J., Kim, M., Vodacek, A. "A Dynamic Data Driven Wildland Fire Model" Computational Science - ICCS 2007, pp. 1042–1049, 2007

42. Zhanqing, L., Khananian, A., Fraser, R., & Cihlar, J. "Automatic Detection of Fire Smoke using Artificial Neural Networks and Threshold Approaches applied to AVHRR Imagery", IEEE Trans. On Geoscience and Remote Sensing, ISSN 0196-2892, 2001.
43. Page, E.S., "Continuous Inspection Schemes," Biometrika vol. 41, pp. 100-115, 1954
44. Gombay, E., & Serban, D., "An adaptation of Pages CUSUM test for change detection," Periodica Mathematica Hungarica, vol. 50, pp. 135-147, 2005.
45. D.A. Darling and P. Erdos, "A limit theorem for the maximum of normalized sums of independent random variables," Duke Math. J, vol. 23, pp. 143-155, 1956.
46. A. Walter, "Notes on the Utilization of Records from Third Order, Climatological Stations for Agricultural Purposes," Ag.Met., 4:137-143, 1967.
47. G. Shafer, "A mathematical theory of evidence," Princeton, New Jersey: Princeton Univ. Press, 1976.
48. P. Smets, R. Kennes, "The transferable belief model," Artificial Intelligence 66: 191-234, 1994
49. Yager, R.R. "On the Dempster-Shafer Framework and New Combination Rules," Information Sciences 41:93-137, 1987.
50. Sekkas O., Manatakis D., Manolakos E., Hadjiefthymiades S., "Sensor and Computing Infrastructure for Environmental Risks - The SCIER system", to appear in Advanced ICTs for Disaster Management and Threat Detection: Collaborative and Distributed Frameworks,(Eds. Dr. E. Asimakopoulou and Dr. N. Bessis), IGI Global, November 2009.
51. Zervas E., Mpimpoudis A., Anagnostopoulos C., Sekkas O., Hadjiefthymiades S., "Multisensor Data Fusion for Fire Detection", Information Fusion (2009), doi: 10.1016/j.inffus.2009.12.006.
52. Zervas E., Sekkas O., Hadjiefthymiades S., Anagnostopoulos C., "Fire Detection in the Urban Rural Interface through Fusion techniques", in Proceedings of the 1st International Workshop on Mobile Ad hoc and Sensor Systems for Global and Homeland Security (MASS-GHS 2007), Pisa, Italy, October 2007.
53. Marias G, Tsetsos V., Sekkas O., Georgiadis P., "A generic framework towards trust building in self-organized, peer, networks", in Proceedings of the 1st

International Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing (SecPerU), Santorini, Greece, July 2005.

54. Marias G., Tsetsos V., Sekkas O., Georgiadis P., "Performance evaluation of a self-evolving trust building framework", in Proceedings of the IEEE SECOVAL Workshop, pp. 132-141, Athens, Greece, September 2005.
55. Marias G.F., Papapanagiotou K., Tsetsos V., Sekkas O., Georgiadis P., "Integrating a Trust Framework with a Distributed Certificate Validation Scheme for MANETs", in EURASIP Journal on Wireless Communications and Networking (EURASIP JWCN), Article ID 78259, 18 pages, 2006, Hindawi Pub. Corp.
56. Marias G.F., Tsetsos V., Sekkas O., Georgiadis P., "Performance evaluation of a self-evolving trust building framework", to appear in International Journal of Internet Technology and Secured Transactions (IJITST), 2006, Inderscience Publishing.
57. A. Jøsang, "The right type of trust for distributed systems", in Proc. 1996 Workshop New Security Paradigms, California, USA, Sept. 1996.
58. F. Stajano, and R. Anderson, "The Resurrecting Duckling: Security Issues for Ad hoc Wireless Networks," in Proc. 7th International Workshop on Security Protocols, 1999, pp. 172-194, 1994.
59. J. Douceur, "The Sybil Attack," in Proc. 1st International Workshop on Peer-to-Peer Systems (IPTPS02), Cambridge, MA, March 2002.
60. A. Abdul-Rahman and S. Hailes, "A Distributed Trust Model", in Proc. New Security Paradigms Workshop, 1997, ACM, 1997.
61. S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehaviour in mobile ad hoc networks", in Proc. of Mobicom2000, Boston, USA, Aug. 2000.
62. S. Buchegger and J.-Y. Le Boudec. "Performance analysis of the CONFIDANT protocol", in Proc. 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing, MOBIHOC2002, Lausanne, Switzerland, June 2002.
63. P. Michiardi and R. Molva, "CORE: A collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks", In Proc. 6th IFIP Communications and Multimedia Security Conference, Portoroz, Slovenia, September 2002..

64. S. Buchegger and J-Y Le Boudec, "The Effect of Rumour Spreading in Reputation Systems for Mobile Adhoc Networks", in Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, WiOpt03, 2003.
65. V. Cahill et. al., "Using Trust for Secure Collaboration in Uncertain Environments," IEEE Pervasive Computing, July 2003.
66. C. English, S. Terzis, and P. Nixon, "Towards Self-Protecting Ubiquitous Systems: Monitoring Trust-based Interactions", Journal of Personal and Ubiquitous Computing, Sept. 2005.
67. A. Pirzada and C. McDonald, "Establishing Trust In Pure Ad-hoc Networks," in Proc. 27th Australian Computer Science Conference, 2004.
68. C. Castelfranchi and R. Falcone, "Trust is much more than subjective probability. Mental components and sources of trust", in Proc. HICSS33, Hawaii, 2000.
69. P. Michiardi, "Cooperation enforcement and network security mechanisms for mobile ad hoc networks" PhD Thesis, Ecole Nationale Supérieure de Telecommunications, Dec. 2004.
70. Y. Wang, and J. Vassileva, "Bayesian Network Trust Model in Peer-to-Peer Networks", in Proc. Agents and Peer-to-Peer Computing 2nd Intl Workshop, AP2PC 2003, July 2003, pp., 23-34.
71. S. Buchegger and J-Y Le Boudec, "A Robust Reputation System for P2P and Mobile Ad-hoc Networks", in Proc. Second Workshop on the Economics of Peer-to-Peer Systems, 2004.
72. Available at <http://www.j-sim.org/>
73. D.Demers, J.Larson, S.Shenker, H.Sturgis, D.Swinehart and D.Terry, Epidemic algorithms for replicated database maintenance, ACM Principles of Distributed Computing, pp. 1–12,1987.
74. M.Franklin and S.Zdonik, Dissemination-based information systems, Data Engineering, 19(3),1996.
75. W. O Kermack and A. G. McKendrick, A Contribution to the Mathematical Theory of Epidemics, in: Proc. Roy. Soc. Lond. A 115, 700-721, 1927.
76. Sekkas O., Piguet D., Anagnostopoulos C., Kotsakos D., Alyfantis G., Kassapoglou-Faist C., Hadjiethymiades S., "Probabilistic Information Dissemination for MANETs: the IPAC Approach", in Proceedings of the 20th

Tyrrhenian International Workshop on Digital Communications, Pula, Sardinia, Italy, September 2009.

77. Sekkas O., Anagnostopoulos C., Hadjiefthymiades S., “Enhancing Collaborative Context-Awareness through Adaptive Epidemic Dissemination”, *submitted to Pervasive and Mobile Computing Journal (PMC)*.
78. Sekkas O., Anagnostopoulos C., Hadjiefthymiades S., “Probabilistic Information Dissemination for MANETs ”, *submitted to Computer Communications Journal*.
79. P. Eugster, R. Guerraoui, S.B. Handurukande, P. Kouznetsov and A.M. Kermarrec, Lightweight probabilistic broadcast, *ACM Trans. Comput. Syst.* 21 (4), 341-374 (2003)
80. J. Luo, P. Eugster and J.P. Hubaux, Route Driven Gossip: Probabilistic Reliable Multicast in Ad Hoc Networks, in: *Proceedings of the 22nd IEEE Conference on Computer Communications*, pp 2229--2239, San Francisco, California, USA (2003)
81. H. Miranda, S. Leggio, L. Rodrigues and K. Raatikainen, A power-aware broadcasting algorithm, In: *Proc. of the 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–5, Helsinki, Finland, (2006)
82. W. Rabiner, J. Kulik and H. Balakrishnan, Adaptive Protocols for Information Dissemination in Wireless Sensor Networks, *MOBICOM 1999*: 174-185 (1999)
83. J. Kulik, W. Rabiner and H. Balakrishnan, Negotiation-Based Protocols for Disseminating Information in Wireless Sensor Networks, *Wireless Networks* 8(2-3): 169-185 (2002)
84. A. Salkham, R. Cunningham and A. Cahill, A taxonomy of collaborative context-aware systems, in: *Proc. Ubiquitous Mobile Information and Collaboration Systems*, pp. 899–911, 2006.
85. M. Boguna, R. Pastor-Satorras and A. Vespignani, Epidemic spreading in complex networks with degree correlations, *Statistical Mechanics of Complex Networks*, 127–147, 2003.
86. P. Tennert, M. Hall, B. Brown, M. Chalmers and S. Sherwood, Three applications for mobile epidemic algorithms, in: *Proc. Human Computer Interaction with Mobile Devices and Services*, pp. 223–226, 2005.

87. C. Zesheng and J. Chuanyi, Spatial-temporal modelling of malware propagation in networks, IEEE Transactions on Neural Networks, 16(5), 1291–1303, Sept. 2005.
88. A. Khelil, C.Becker, J. Tian and K. Rothermel, An epidemic model for information diffusion in Manets, ACM Modeling Analysis and Simulation of Wireless and Mobile Systems, pp. 54–60, 2002.
89. Y. Wang, D. Chakrabarti, C. Wang and C. Faloutsos, Epidemic spreading: An eigenvalue viewpoint, Symp. Reliable and Distributed Systems, 2003.
90. J. Mantyjarvi, P. Huuskonen and J. Himberg, Collaborative context determination to support mobile terminal applications, Pervasive Communications, pp. 161–168, 2003.
91. A. Lindgren, A. Doria, and O. Schelén, Probabilistic Routing in Intermittently Connected Networks, Mobile Computing and Communications Review, July 2003.
92. J.P. Tower and T. D.C. Little, A Proposed Scheme for Epidemic Routing with Active Curing for Opportunistic Networks, in: Proc. 22nd International Conference on Advanced Information Networking and Applications –Workshops,pp. 1696-1701, 2008.
93. K. A. Harras, K. C. Almeroth, and E. M. Belding-Royer, Delay Tolerant Mobile Networks (DTMNs): Controlled Flooding in Sparse Mobile Networks, in Proc. IFIP Networking, 2005.
94. Varga, A.: OMNET++ Discrete Event Simulation System User Manual. (2006)
95. E. Gombay and D. Serban, An adaptation of Page’s CUSUM test for change detection. Period. Math. Hung. 50, No. 1-2, 2005, pp. 135-147.
96. S. Ben-David, J. Gehrke and D. Kifer, Detecting Change in Data Streams, in: Proc.30th VLDB Conference, Toronto, Canada, 2004.
97. <http://www.irda.org/>
98. <http://www.lesswire.com/>