ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΔΙΑΧΕΙΡΙΣΗ ΤΕΡΜΑΤΙΚΩΝ ΚΑΙ ΣΤΟΙΧΕΙΩΝ ΔΙΚΤΥΟΥ ΣΕ ΠΕΡΙΒΑΛΛΟΝΤΑ ΑΣΥΡΜΑΤΩΝ ΕΠΙΚΟΙΝΩΝΙΩΝ ΥΨΗΛΩΝ ΤΑΧΥΤΗΤΩΝ ΠΕΡΑΝ ΤΗΣ 3ΗΣ ΓΕΝΙΑΣ

ΑΠΟΣΤΟΛΟΣ Γ.ΚΑΤΙΔΙΩΤΗΣ

ΠΕΙΡΑΙΑΣ 2008

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

....

Δρ. Απόστολος Γ. Κατιδιώτης

Απόστολος Γ. Κατιδιώτης, 2008.

Με επιφύλαξη παντός δικαιώματος.

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

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

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΕΡΙΛΗΨΗ

Την τελευταία δεκαετία ο κόσμος των ασυρμάτων επικοινωνιών υπόκειται μια πληθώρα καίριων αλλαγών, οι οποίες έχουν φέρει τις ασύρματες επικοινωνίες στο προσκήνιο της διεθνούς έρευνας, τα αποτελέσματα της οποίας αντικατοπτρίζονται σε πολυάριθμες καινοτόμες τεχνολογίες / προϊόντα, όπως το WiMAX (Wordwide Interoperability for Microwave Access), τα στάνταρτς της IEEE (802.20, 802.22, κλπ), και πολλά άλλα (mesh networks, software defined radio). Στο παραπάνω πλαίσιο, ο ασύρματος κόσμος κινείται προς την κατάσταση που ονομάζεται B3G (πέραν της τρίτης γενιάς), όπου οι τηλεπικοινωνιακές υποδομές θα πρέπει να αντιμετωπίζουν συνεχώς μεταβαλλόμενες εξωτερικές συνθήκες, οι οποίες είναι ολοένα και λιγότερο προβλέψιμες ως προς την ποιότητα των υπηρεσιών που επιτρέπουν. Στον αντίποδα αυτών, οι τεχνολογίες B3G εκμεταλλεύονται την κατ' επιλογήν χρήση των πολυάριθμων διαθεσίμων τεχνολογιών (RATs), με στόχο την παροχή υψηλών επιπέδων ποιότητας υπηρεσιών σε οικιακούς και μη χρήστες. Ωστόσο, οι πολλές καινοτομίες εγείρουν το ζήτημα της επιτυχούς διαχείρισης τερματικών και στοιχείων δικτύου σε ένα τόσο ανταγωνιστικό περιβάλλον. Αυτό ακριβώς αποτελεί και το αντικείμενο της παρούσης διατριβής.

Με βάση τα ανωτέρω, το πρώτο κεφάλαιο της διατριβής ανατρέχει στις σημερινές διαθέσιμες τεχνολογίες και δη στα συστήματα B3G, καλύπτοντας έννοιες όπως η δυναμική αναδιάρθρωση τερματικών και στοιχείων δικτύων, καθώς επίσης και τα cognitive δίκτυα επικοινωνιών, ενώ αναφορά γίνεται και στην ανάγκη σχεδιασμού και ανάπτυξης των σχετικών μηχανισμών διαχείρισης, με βάση τις επιταγές των ίδιων των υποδομών, αλλά και των χρηστών.

Μετά τη δικαιολόγηση της επιτακτικής ανάγκης ανάπτυξης ευφυών μεθόδων διαχείρισης B3G συστημάτων, το δεύτερο κεφάλαιο παρουσιάζει ένα τέτοιο σύστημα, ονόματι Reconfigurable Terminal Management System (RTMS), το οποίο διαχειρίζεται ένα τερματικό B3G μέσω της μοντελοποίησης του προφίλ ενός χρήστη, της απόκτησης πληροφοριών των συνθηκών του περιβάλλοντος, αλλά και της διαπραγμάτευσης για την απόφαση σχετικά με την καλύτερη (κατ' επιλογήν) διάρθρωση του εν λόγω τερματικού.

Αναφορικά με την απόκτηση πληροφοριών για το περιβάλλον, τα επόμενα 3 κεφάλαια προτείνουν μηχανισμούς διαχείρισης του φάσματος ενός cognitive δικτύου επικοινωνιών. Για αυτό το σκοπό είναι σημαντική η ύπαρξη μηχανισμών μάθησης, ώστε ένα σύστημα να μπορεί να εκμεταλλευτεί μετρήσεις που ανακτά από το περιβάλλον του, μαζεύοντας γνώση και εμπειρία και βρίσκοντας τις δυνατότητες που του προσφέρουν οι διάφορες δυνατές αναδιαρθρώσεις του, ώστε να κατευθυνθεί στη βέλτιστη απόφαση.

Συγκεκριμένα, το τρίτο κεφάλαιο προτείνει κάποιες βελτιώσεις στη φάση χαρακτηρισμού ενός διαύλου ενός cognitive δικτύου. Η προτεινόμενη μέθοδος αποτιμά τις υποψήφιες αναδιαρθρώσεις ενός πομποδέκτη, συσχετίζοντάς τες με δυνατές ταχύτητες (bit rates) που μπορεί να επιτύχει. Για το σκοπό αυτό λαμβάνει υπόψη μετρήσεις channel-state estimation information (CSI) και αυξάνει τη βεβαιότητα σχετικά με την αποτίμηση των αναδιαρθρώσεων συγκεντρώνοντας γνώσεις και εμπειρία με τη χρήση Bayesian δικτύων. Αποτελέσματα από διάφορα σενάρια αποδεικνύουν την ορθότητα της προτεινόμενης μεθόδου.

Εν συνεχεία, στα επόμενα δύο κεφάλαια (4ο και 5ο) προτείνονται κάποια σχήματα εκμάθησης που βασίζονται σε νευρωνικά δίκτυα που χρησιμοποιούνται για την πρόβλεψη των δυνατοτήτων διαφόρων πιθανών αναδιαρθρώσεων. Συγκεκριμένα, παρουσιάζονται διάφορα ρεαλιστικά, εμπορικά πραγματοποιήσιμα σενάρια, με σκοπό το σχεδιασμό ενός, καταλλήλου δομής, νευρωνικού δικτύου, ώστε να εξαχθούν τα κατάλληλα συμπεράσματα περί της επιτυχούς χρήσης τέτοιων μεθόδων μάθησης σε cognitive δίκτυα επικοινωνιών.

Κατόπιν, το έκτο κεφάλαιο παρουσιάζει ένα συνολικό πλαίσιο διαχείρισης τερματικών και στοιχείων δικτύου συστημάτων B3G, παρουσιάζοντας τα βασικά μέρη από τα οποία αποτελείται και τον τρόπο που αλληλεπιδρούν. Αποτελέσματα από προσομοιώσεις παρατίθενται προς πιστοποίηση των ισχυρισμών περί της καταλληλότητας του προτεινομένου πλαισίου (υποδομής).

Στο τελευταίο κεφάλαιο της διατριβής παρατίθενται διάφορα συμπεράσματα, καθώς επίσης και μελλοντικές προκλήσεις που προκύπτουν εξ' αυτής.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ABSTRACT

Over the last decade the world of wireless communications has been undergoing some crucial changes, which have brought it at the forefront of international research and development interest, eventually resulting in the advent of a multitude of innovative technologies and associated products such as WiFi, WiMax, the suite of 802.20 and 802.22 standards, wireless mesh networks and software defined radio.

In this respect, the wireless world is rapidly evolving towards the “Beyond the 3rd Generation” era, where communication infrastructures need to tackle external conditions that are continuously changing, and thus become less predictable in terms of quality of service provision. On the contrary, the B3G era, through the coexistence and complementary use of a multitude of Radio Access Technologies, offers additional capabilities for providing users with advanced levels of convenience and flexibility for living and working.

As the management of terminal and network segments becomes more and more crucial, in such a multi-dimensional era, the scope of this thesis is to deal with management aspects regarding terminal and network elements in high-speed, B3G communication environments.

The first chapter includes a report in modern tendencies in the world of telecommunications, focusing in the vision of the B3G systems. The main significances that are covered in the first chapter are the dynamic reformation of networks and terminals (reconfigurability) and of cognitive networks (CN). A reference is also been made on why an advanced management functionality is rendered necessary, based on infrastructure and business level issues and requirements.

After justifying the necessity to design and develop advanced management functionality in the B3G systems, the second chapter presents such a management. The system is called Reconfigurable Terminal Management System and in general it provides the means for profile modelling, the acquisition of monitoring/ discovery/

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

context information, and the negotiation and selection of configurations, based on information deriving from policies, as well as the profiles and the context.

Regarding the monitoring/discovery/context information acquisition, the next three chapters (3 – 5) present intelligent management mechanisms for the allocation and usage of the radio spectrum, with the use of cognitive radio systems. In such a process, learning mechanisms that are capable of exploiting measurements sensed from the environment (monitoring), gathered experience and stored knowledge and discovering the capabilities of the different configurations (discovery), are judged as rather beneficial for guiding decisions and actions.

In particular, the third chapter proposes enhancements to the channel(-state) estimation phase of a cognitive radio system. The proposed method aims at evaluating the various candidate configurations that a cognitive transmitter may operate in, by associating a capability e.g., achievable bit-rate, with each of these configurations. It takes into account calculations of channel capacity provided by channel-state estimation information (CSI) and the sensed environment, and at the same time increases the certainty about the configuration evaluations by considering past experience and knowledge through the use of Bayesian networks. Results from comprehensive scenarios show the impact of the method proposed in this chapter, on the behaviour of cognitive radio systems.

Then, the next two chapters (4^{th} and 5yh one) introduce and evaluate learning schemes that are based on artificial neural networks and can be used for predicting the capabilities (e.g. data rate) that can be achieved by a specific radio configuration. Specifically, interesting scenarios, which include both commercial and simulation hardware/software products, are mobilized for the benchmarking work, conducted in order to design and use an appropriate neural network structure, while indicative results are presented and discussed in order to showcase the benefits of incorporating such learning schemes into cognitive radio systems.

The sixth chapter presents an advanced management framework, as an enabling technology for designing and developing, wireless systems in B3G environments. The chapter focuses on the main components of the proposed framework, as well as on

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

their functionality and interactions. Additionally, indicative simulation results showcase the efficiency of the proposed framework.

Finally, in the last chapter, some general conclusions that arose from the thesis are presented, as well as future challenges for the continuation of the presented work.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Στους γονείς μου, Γιώργο και Φωτεινή,

στην αδερφή μου, Ευαγγελία,

και στη Μαρία

ΠΡΟΛΟΓΟΣ

Η μεταλαμπάδευση της γνώσης από γενιά σε γενιά αποτελεί, ως πρέπει, βασικότατο στόχο κάθε ανθρώπινης δραστηριότητας. Ειδικά σε ένα τομέα τόσο ευμετάβλητο όσο οι ασύρματες επικοινωνίες, κάθε προσπάθεια προς την κατάκτηση της γνώσης είναι όχι μόνο θεμιτή αλλά και επιβεβλημένη. Αυτές και άλλες παρόμοιες σκέψεις έχουν συνοδεύσει την παρούσα ερευνητική προσπάθεια που αποκρυσταλλώνεται στη συγγραφή αυτής της διατριβής, προσπάθεια η οποία ξεκίνησε και ολοκληρώθηκε στο εργαστήριο Δικτύων Τηλεπικοινωνιών και Δικτυακών Υπηρεσιών του τμήματος Ψηφιακών Συστημάτων του Πανεπιστημίου Πειραιώς.

Ο δρόμος προς τη γνώση όμως είναι πάντα δύσκολος. Τα εμπόδια που συναντά κανείς φαντάζουν τεράστια και ανυπέρβλητα. Εμπόδια που χωρίς τη συμβολή ορισμένων ανθρώπων είναι αδύνατο να ξεπεραστούν. Έτσι, προλογίζοντας τη διατριβή θα ήθελα να απευθύνω ευχαριστίες προς αυτούς τους ανθρώπους, η συμβολή των οποίων έκανε όλα αυτά τα εμπόδια που συνάντησα στο δρόμο προς την ολοκλήρωση της διατριβής να φαντάζουν μικρότερα.

Υπεράνω όλων, η οικογένεια μου, ήταν εκεί να με στηρίζει και πολλές φορές να με ανέχεται, σε όλες τις φάσεις της ζωής μου. Ο πατέρας μου, Γιώργος και η μητέρα μου, Φωτεινή, που με πολύ κόπο κατάφεραν να μεγαλώσουν ένα παιδί, προσφέροντάς του τα πάντα, ώστε να μπορεί να επικεντρωθεί στους στόχους του και τελικά να τους πετύχει. Η αδερφή μου Ευαγγελία, που με την αγάπη της και την φροντίδα της μου έδινε πάντα κουράγιο και δύναμη να προχωρώ. Σε αυτούς οφείλω αρχικά να χαρίσω το μεγαλύτερο μερίδιο της ηθικής ικανοποίησης για το αποτέλεσμα αυτής της πολυετούς προσπάθειας και να τους πω ένα μεγάλο ευχαριστώ...

Καθ’ όλη τη διάρκεια της πορείας προς την ολοκλήρωση της διατριβής, βασικότερος αρωγός στάθηκε ο επιβλέπων καθηγητής μου, αναπληρωτής καθηγητής του Πανεπιστημίου Πειραιώς, κ. Παναγιώτης Δεμέστιχας, ο οποίος επέδειξε αμέριστη εμπιστοσύνη προς το πρόσωπό και τις δυνατότητές μου από τις αρχές της συνεργασίας μας. Η συμβολή του κ. Δεμέστιχα στην ολοκλήρωση της διατριβής μου ήταν τεράστια μιας και η συνεχής καθοδήγησή του, οι παρατηρήσεις του σε τεχνικά και μη θέματα, η υποστήριξή του, ηθική και ψυχολογική, καθώς και η πάντα

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

καλοπροαίρετη πίεση του, συνετέλεσαν ουσιαστικά στην ολοκλήρωση της παρούσας διατριβής. Ο κ. Δεμέστιχας υπήρξε για μένα ένας πραγματικός "δάσκαλος". Γι' αυτό νιώθω την ανάγκη να τον ευχαριστήσω και ελπίζω να φανώ ακόμα περισσότερο αντάξιος των προσδοκιών του.

Υπόχρεος επίσης νιώθω προς τους αναπληρωτή καθηγητή κ. Δημήτριο Σάμψων και καθηγητή Νικήτα Σγούρο, καθηγητές του Πανεπιστημίου Πειραιώς, οι οποίοι αποτέλεσαν πηγή έμπνευσης για μένα. Η συνεργασία μαζί τους, από τα προπτυχιακά μου χρόνια ακόμα, γινόταν πάντα σε άριστο και φιλικό κλίμα. Ποτέ δεν αρνήθηκαν την βοήθειά τους, οποτεδήποτε και αν τη ζήτησα και γι' αυτό θα ήθελα να τους εκφράσω την ευγνωμοσύνη μου.

Ένας ακόμη άνθρωπος στον οποίο οφείλω πραγματικά πολλά, είναι ο πολύ καλός μου φίλος, συνάδελφος και συμπαίκτης Δρ. Δημητρακόπουλος Γιώργος. Ο Γιώργος στάθηκε δίπλα μου σαν πραγματικός αδερφός. Είναι αυτός που με βοήθησε και πάντα με βοηθάει σε οτιδήποτε και αν χρειαστώ. Σε όλη αυτή τη διαδρομή μου συμπαραστάθηκε όπως λίγοι άνθρωποι θα έκαναν και γι' αυτό του οφείλω ένα μεγάλο ευχαριστώ.

Επίσης, δε μπορώ να μην ευχαριστήσω την πολύ καλή μου φίλη, συνάδελφο και σύντομα Δρα, Κρητικού Γιούλη, που μαζί ξεκινήσαμε, συνεχίσαμε και όπως όλα δείχνουν, μαζί θα τελειώσουμε αυτή την πορεία, για τη φιλία της πρώτα απ' όλα, καθώς και για την ψυχολογική υποστήριξη, αλλά και τις όμορφες στιγμές που περάσαμε μαζί στο εργαστήριο 307 της Ανδρούτσου.

Επιπλέον, θα ήθελα να ευχαριστήσω το φίλο και συνάδελφο Δρα. Κώστα Τσαγκάρη καθώς και τη Λέκτορα Βέρα-Αλεξάνδρα Σταυρουλάκη, χωρίς την βοήθεια των οποίων θα ήταν μάλλον αδύνατη η ολοκλήρωση της παρούσας εργασίας.

Ακόμα, ένα σημαντικό μερίδιο στην ολοκλήρωση αυτής της εργασίας κατέχουν τα υπόλοιπα μέλη του εργαστηρίου Δικτύων Τηλεπικοινωνιών και Δικτυακών Υπηρεσιών και πιο συγκεκριμένα οι Άγγελος Σαατσάκης, Διονύσης Πετρομανωλάκης και Μάριος Λογοθέτης. Τους εύχομαι σύντομα και αυτοί να ολοκληρώσουν το δικό τους ταξίδι προς τη γνώση.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Επίσης, θα ήθελα να ευχαριστήσω τους φίλους και συνεργάτες του «αδελφού» εργαστηρίου, Δρα Γιώργο Μπράβο, Παναγιώτη Θεοφυλάκο και Βλάση Μπαρούση και να ευχηθώ στους τελευταίους να γίνουν σύντομα διδάκτορες.

Τέλος, εξέχουσα θέση μεταξύ των ανθρώπων που με στήριξαν όλο αυτό το διάστημα κατέχει η σύντροφός μου Μαρία, η οποία ήταν πάντα δίπλα μου να με εμψυχώνει όταν έχανα το κουράγιο μου και με υπομονή και αγάπη να ανέχεται τις ιδιοτροπίες μου.

Με τιμή,

Απόστολος Γ. Κατιδιώτης

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Table of Contents

ΔΡ. ΑΠΟΣΤΟΛΟΣ Γ. ΚΑΤΙΔΙΩΤΗΣ III

ΠΕΡΙΛΗΨΗ V

ABSTRACT VII

ΠΡΟΛΟΓΟΣ XI

TABLE OF CONTENTS XV

LIST OF TABLES XVII

LIST OF FIGURES XIX

LIST OF ACRONYMS XXIII

1. WIRELESS COMMUNICATIONS, BEYOND THE 3RD GENERATION (B3G) 27

1.1. THE WIRELESS WORLD TODAY 29
1.2. RECONFIGURABILITY 31
1.3. COGNITION 33
1.4. NEED FOR ADVANCED MANAGEMENT FUNCTIONALITY 36
1.4.1. Infrastructure issues 36
1.4.2. Business level issues and requirements 37
REFERENCES 39

2. MANAGEMENT SYSTEM FOR TERMINALS IN THE WIRELESS B3G WORLD 43

2.1. INTRODUCTION 45
2.1.1. Need For Advanced Management Functionality 45
2.1.2. Relation with previous work 46
2.2. BUSINESS CASE 47
2.3. MANAGEMENT SYSTEM 49
2.4. PROFILE MODELLING 51
2.5. CONTEXT ACQUISITION 54
2.6. POLICIES 57
2.7. CONFIGURATION NEGOTIATION AND SELECTION 59
2.7.1. Operation 59
2.7.2. Cooperation with NOs 60
2.7.3. Selection of optimal configurations 62
2.8. RESULTS 66
2.8.1. Input description 66
2.8.2. Output from first scenario 70
2.8.3. Output from second scenario 74
2.8.4. Impact of cooperation and negotiation with other NOs 76
2.9. COOPERATION AND NEGOTIATION STRATEGIES 80
2.9.1. General approach 80
2.9.2. Specific schemes 80
2.10. SELECTION STRATEGY 81
2.11. CONCLUSIONS 83
REFERENCES 84

3. ENHANCING CHANNEL ESTIMATION IN COGNITIVE RADIO SYSTEMS BY MEANS OF BAYESIAN NETWORKS 87

3.1. INTRODUCTION 89
3.2. RELATED WORK AND MOTIVATION 90
3.3. FORMULATION AS A BAYESIAN NETWORK 91
3.4. LEARNING STRATEGY 93
3.4.1. Principles 93
3.4.2. Algorithm 95
3.5. RESULTS 98
3.5.1. Set-up 98
3.5.2. Presentation 98

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

3.5.3. Analysis *115*
3.6. CONCLUSIONS 116

REFERENCES 118

4. NEURAL NETWORK-BASED LEARNING SCHEMES FOR COGNITIVE RADIO SYSTEMS 121

4.1. INTRODUCTION AND PROBLEM STATEMENT 123
4.2. NEURAL NETWORKS REVIEW 124
4.3. MOTIVATION 127
4.4. BASIC NN-BASED LEARNING SCHEME 129
4.4.1. Preparation Procedure *129*
4.4.2. NN pattern selection - Results *131*
4.5. EXTENDED NN-BASED LEARNING SCHEME 137
4.5.1. Preparation Procedure *137*
4.5.2. NN pattern selection – Results *139*
4.6. DISCUSSION AND CONCLUSIONS 144

REFERENCES 146

5. PERFORMANCE EVALUATION OF ARTIFICIAL NEURAL NETWORKS BASED LEARNING SCHEMES FOR COGNITIVE RADIO SYSTEMS 147

5.1. INTRODUCTION AND PROBLEM STATEMENT 149
5.2. RELATED WORK AND MOTIVATION 150
5.3. PERFORMANCE EVALUATION 152
5.3.1. Scenario 1 .... *153*
5.3.2. Scenario 2 .... *159*
5.3.3. Scenario 3 .... *164*
5.4. CONCLUSIONS AND FUTURE WORK 175

REFERENCES 176

6. RTMS POSITIONING IN A MANAGEMENT FRAMEWORK FOR B3G ENVIRONMENTS 177

6.1. INTRODUCTION 179
6.2. ARCHITECTURE AND FUNCTIONALITY OF MANAGEMENT FRAMEWORK 180
6.2.1. High Level Description of Management Architecture *180*
6.2.2. Management Functionality *182*
6.3. CONTEXT INFORMATION ACQUISITION 183
6.4. NETWORK MANAGER (NM) - SELECTION OF OPTIMUM CONFIGURATIONS 185
6.4.1. Overview of NM *185*
6.4.2. Description of Components *185*
6.5. TERMINAL MANAGER (RTMS) 187
6.5.1. Overview of RTMS *187*
6.5.2. Description *188*
6.6. SIMULATION RESULTS 190
6.7. CONCLUSIONS 196

REFERENCES 198

7. CONCLUSIONS – FUTURE CHALLENGES 199

8. APPENDIX: PUBLICATION LIST (NOVEMBER 2008) 203

PUBLICATIONS IN INTERNATIONAL JOURNALS AND BOOKS 203
PUBLICATIONS IN INTERNATIONAL CONFERENCES 204

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

List of Tables

TABLE 1-1: BASIC CONCEPTS IN WIRELESS COMMUNICATION TECHNOLOGIES 31
TABLE 2-1: DATA STRUCTURES OF THE PROFILE MODELLING COMPONENT OF THE RTMS. 52
TABLE 2-2: DATA STRUCTURES OF THE CONTEXT ACQUISITION COMPONENT OF THE RTMS. 54
TABLE 2-3: DATA STRUCTURES OF THE POLICY COMPONENT OF THE RTMS... 58
TABLE 2-4: COST OF RECONFIGURING TRANSCEIVERS. 69
TABLE 2-5: PRIORITIES ALLOCATED TO VARIOUS TRANSCEIVER CONFIGURATIONS BY THE POLICIES. 70
TABLE 2-6: CONTRIBUTION TO OF VALUES, RESULTING FROM THE RTMS ENVIRONMENT (A) X1..X5, (B) X6..X10 71
TABLE 4-1: TEST CASES EXAMINED FOR BASIC SCHEME 133
TABLE 4-2: TEST CASES EXAMINED FOR EXTENDED SCHEME. 140
TABLE 5-1: TEST CASES EXAMINED FOR SCENARIO 1.. 155
TABLE 5-2. TEST CASES EXAMINED FOR SCENARIO 2.. 160
TABLE 5-3. CHARACTERISTICS OF THE NETWORKS USED IN SCENARIO 3. 166
TABLE 5-4. TEST CASES EXAMINED FOR SCENARIO 3.. 167
TABLE 6-1: CONTEXT INFORMATION PER RAT AND PER PRODUCER. 184

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

List of Figures

Figure 1-1: Overview of the Wireless World in the B3G Era. 30
Figure 1-2: Reconfigurable Platform 33
Figure 1-3: Simplified representation of cognitive radio cycle. 34
Figure 1-4: Cognitive radio engine. 35
Figure 1-5: Rational for the need for advanced management functionality. 36
Figure 2-1: Business case. (a) Static view: user/terminal, NOs, application/content segment. (b) Dynamic view: interactions for selecting the configurations that will lead to association with the best radio networks. 48
Figure 2-2: Functional architecture of the Reconfigurable Terminal Management System (RTMS). 50
Figure 2-3: Scope of policies with respect to an arbitrary, high-level protocol stack. 57
Figure 2-4: Overall operation of the configuration negotiation and selection component. 60
Figure 2-5: Negotiation scheme. 61
Figure 2-6: Computation of optimal configurations. 64
Figure 2-7: Utility values in the different contexts. 66
Figure 2-8: Maximum cost values in the different contexts. 67
Figure 2-9: Service area considered in the tests and respective coverage (configurations available). 68
Figure 2-10: Scenario set 1. Mean aggregate objective function values, when the service area is covered by a heterogeneous B3G infrastructure or only 3G technologies, at contexts x1 – x5 72
Figure 2-11: Scenario set 1. Mean aggregate objective function values, when the service area is covered by a heterogeneous B3G infrastructure or only 3G technologies, at contexts x6 – x10 73
Figure 2-12: Scenario set 2. Mean aggregate objective function values, when the service area is covered by a heterogeneous B3G infrastructure or only 3G technologies, at contexts x1 – x5. 75
Figure 2-13: Scenario set 2. Mean aggregate objective function values, when the service area is covered by a heterogeneous B3G infrastructure or only 3G technologies, at contexts x1 – x5 76
Figure 2-14: Results from the cooperation and negotiation with other NOs. Samples regarding the evolution of the $cst_P^{cpn}(u, s_2, q_{23}, wl_{cp}, x_{10})$ values, offered by a cooperating network, according to the reversed English model 78
Figure 2-15: Results from the cooperation and negotiation with other NOs. Samples regarding the evolution of the $cst_P^{cpn}(u, s_2, q_{23}, wl_{cp}, x_{10})$ values, offered by a cooperating network, according to the reversed Dutch model. 79
Figure 3-1: Structure of the Bayesian network. 92
Figure 3-2: Structure of the Conditional Probability Tables (CPTs). 92
Figure 3-3: Organization of CPT columns as ordered lists, for enabling fast adaptations. 93
Figure 3-4: Behaviour of the channel estimation phase: strategy for learning the configuration. 97
Figure 3-5: First set of scenarios. Behaviour of the proposed method when it has learned that the configuration can achieve: (a) 6 Mbps; (b) 12 Mbps; (c) 24 Mbps; (d) 36 Mbps; (e) 48 Mbps; (f) 54 Mbps. 101
Figure 3-6: Second set of scenarios. Behaviour of the proposed method when the calculated bit-rate changes from 6 Mbps and $|T|=2$ to: (a) 12 Mbps; (b) 24 Mbps; (c) 36 Mbps; (d) 48 Mbps; (e) 54 Mbps. 106
Figure 3-7: Third set of scenarios. Behaviour of the proposed method when the calculated bit-rate changes from 6 Mbps and $|T|=3$ to: (a) 12 Mbps; (b) 24 Mbps; (c) 36 Mbps; (d) 48 Mbps; (e) 54 Mbps. 109

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

FIGURE 3-8: FOURTH SET OF SCENARIOS. BEHAVIOUR OF THE PROPOSED METHOD WHEN THE CALCULATED BIT-RATE CHANGES FROM 24 MBPS AND $|T|$ =2 TO: (A) 6 MBPS; (B) 12 MBPS; (C) 36 MBPS; (D) 48 MBPS; (E) 54 MBPS. ... 112
FIGURE 3-9: FIFTH SET OF SCENARIOS. BEHAVIOUR OF THE PROPOSED METHOD WHEN THE CALCULATED BIT-RATE CHANGES FROM 24 MBPS AND $|T|$ =3 TO: (A) 6 MBPS; (B) 12 MBPS; (C) 36 MBPS; (D) 48 MBPS; (E) 54 MBPS. ... 115
FIGURE 4-1: TYPICAL NEURAL NETWORK STRUCTURE. ... 125
FIGURE 4-2: TRANSFER FUNCTIONS (A) LOG-SIGMOID TRANSFER FUNCTION (B) HYPERBOLIC TANGENT SIGMOID TRANSFER FUNCTION (C) LINEAR TRANSFER FUNCTION ... 126
FIGURE 4-3: COGNITIVE RADIO ENGINE ... 128
FIGURE 4-4: TIME WINDOW ... 130
FIGURE 4-5: CUMULATIVE DISTRIBUTION FUNCTIONS OF INPUT TIME-SERIES. ... 132
FIGURE 4-6: MEASURED PERFORMANCE (MSE) AT BOTH TRAINING AND VALIDATION SEQUENCES FOR ALL 84 TEST CASES – BASIC SCHEME. ... 134
FIGURE 4-7: NEURAL NETWORK FOR THE BASIC LEARNING SCHEME. ... 135
FIGURE 4-8: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN TRAINING SEQUENCE – BASIC SCHEME. ... 136
FIGURE 4-9: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE – BASIC SCHEME. ... 137
FIGURE 4-10: WEIGHT VALUES PER TIME SLOT FOR THE EXTENDED SCHEME. ... 139
FIGURE 4-11: NEURAL NETWORK FOR THE EXTENDED LEARNING SCHEME. ... 142
FIGURE 4-12: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN TRAINING SEQUENCE – EXTENDED SCHEME. ... 143
FIGURE 4-13: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE – EXTENDED SCHEME. ... 144
FIGURE 5-1: BEST NN SRUCTURE – SCENARIO 1. ... 156
FIGURE 5-2: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN TRAINING SEQUENCE – SCENARIO 1 ... 156
FIGURE 5-3: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE 1 – SCENARIO 1. ... 157
FIGURE 5-4: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE 2 – SCENARIO 1. ... 158
FIGURE 5-5: BEST NN SRUCTURE – SCENARIO 2. ... 161
FIGURE 5-6: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN TRAINING SEQUENCE – SCENARIO 2. ... 162
FIGURE 5-7: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE 1 – SCENARIO 2. ... 163
FIGURE 5-8: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE 2 – SCENARIO 2. ... 164
FIGURE 5-9: MEASURED PERFORMANCE (MSE) AT TRAINING AND BOTH VALIDATION SEQUENCES FOR ALL 20 TEST CASES – SUB-SCENARIO 1. ... 168
FIGURE 5-10: MEASURED PERFORMANCE (MSE) AT TRAINING AND BOTH VALIDATION SEQUENCES FOR ALL 20 TEST CASES – SUB-SCENARIO 2. ... 169
FIGURE 5-11: BEST NN SRUCTURE – SCENARIO 3. ... 170
FIGURE 5-12: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN TRAINING SEQUENCE – SCENARIO 3, SUB-SCENARIO 1. ... 170
FIGURE 5-13: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE 1 – SCENARIO 3, SUB-SCENARIO 1. ... 171
FIGURE 5-14: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE 2 – SCENARIO 3, SUB-SCENARIO 1. ... 172
FIGURE 5-15: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN TRAINING SEQUENCE – SCENARIO 3, SUB-SCENARIO 2. ... 173
FIGURE 5-16: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE 1 – SCENARIO 3, SUB-SCENARIO 2. ... 174
FIGURE 5-17: MEASURED PERFORMANCE (MSE) OF THE SELECTED NN IN VALIDATION SEQUENCE 2 – SCENARIO 3, SUB-SCENARIO 2. ... 174
FIGURE 6-1: CHAPTER MOTIVATION AND CONTRIBUTION. ... 180
FIGURE 6-2: ARCHITECTURE OF THE PROPOSED MANAGEMENT FRAMEWORK. ... 181

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

FIGURE 6-3: (A) SERVICE AREA CONSIDERED IN THE SCENARIO AND RESPECTIVE COVERAGE (CONFIGURATIONS AVAILABLE). (B) UTILITY VOLUME VALUES. (C) MAXIMUM PRICE VALUES. (D) COST OF RECONFIGURING TRANSCEIVER. (E) PRIORITIES ALLOCATED TO VARIOUS TRANSCEIVER CONFIGURATIONS BY THE POLICIES. 192
FIGURE 6-4: OF VALUES OF AVAILABLE SOLUTIONS (A) AT LOCATION 1, (B) AT LOCATION 2, (C) AT LOCATION 3 AND (D) AT LOCATION 4. 196

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

List of Acronyms

Acronym	*Explanation*
ANN	Artificial Neural Network
AI	Artificial Intelligence
AMPS	Advanced Mobile Phone System
AP	Access Point
B3G	Beyond the 3rd Generation
CDMA	Code Division Multiple Access
CN	Cognitive Networks
CNS	Configuration Negotiation and Selection
CSI	Channel-State Estimation Information
EDGE	Enhanced Data rates for GSM Evolution
ETSI	European Telecommunications Standards Institute
FCC	Federal Communications Commission
FDMA	Frequency Division Multiple Access
FF	Feed-Forward back-propagation
FTDNN	Focused Time-Delay Neural Network
FTP	File Transfer Protocol
GPRS	General Packet Radio Service
GSM	Global System for Mobile
HSCSD	High Speed Circuit Switched Data
HSDPA	High Speed Downlink Packet Access

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

IETF	Internet Engineering Task Force
IP	Internet Protocol
IPTV	Television over Internet Protocol
ISDN	Integrated Services Digital Network (Telephony)
ISM	Instrumentation, Scientific, and Medical band
ISO	International Organization for Standardization
ISP	Internet Service Provider
ITU	International Telecommunication Union
KPI	Key Performance Indicators
LAN	Local Area Network
MAN	Metropolitan Area Networks
MSE	Mean Squared Error
NM	Network Manager
NN	Neural Network
NO	Network Operator
OF	Objective Function
OFDM	Orthogonal Frequency Division Multiplexing
OSI	Open Systems Interconnection
P2P	Peer to Peer
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RAT	Radio Access Technology

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

RRM	Radio Resource Management
RSSI	Received Signal Strength Indication
RTMS	Reconfigurable Terminal Management System
SDMA	Space Division Multiple Access
SDR	Software Defined Radio
SLA	Service Level Agreement
TCP	Transmission Control Protocol
TM	Terminal Manager
UMTS	Universal Mobile Telecommunications System
WCDMA	Wideband Code Division Multiple Access
Wi-Fi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network
WShRNs	Wireless Short Range Network
WSN	Wireless Sensor Network
WWAN	Wireless Wide Area Network
WWW	World Wide Web

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

1. WIRELESS COMMUNICATIONS, BEYOND THE 3RD GENERATION (B3G)

Abstract:

The first chapter includes a report in modern tendencies in the world of telecommunications, focusing on the vision of the Beyond the 3rd Generation (B3G) systems. The main concepts introduced in this chapter are (i) the dynamic adaptation of networks and terminals to external requirements (reconfigurability) and (ii) the advent of cognitive networks (CN). A reference is also made on why an advanced management functionality is rendered necessary, based on infrastructure and business level issues and requirements. Parts of this chapter have been published in [1][2].

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

WIRELESS COMMUNICATIONS, BEYOND THE 3RD GENERATION (B3G)

1.1. The Wireless World Today

Wireless communications attract significant research and development effort, reflected on the progress of work performed in international projects [3], as well as on the discussions in international fora [4]. This work results in a powerful, high-speed infrastructure that offers versatile solutions to the digital information society. In this context, the technological focus is on the cooperation and coexistence of legacy Radio Access Technology (RAT) standards with currently emerging ones. The current wireless landscape is characterized by a plethora of RATs, which can be roughly classified in two major families:

- Wireless wide area networking (WWAN) technologies, which include, among others, 2G/2.5G/3G mobile communications [5], the IEEE 802.16 suite [6], WiMAX [7] and broadcasting technologies [8];

- Wireless short range networks (WShRNs), which include wireless local and personal area networks (WLANs/WPANs), as well as wireless sensor networks (WSNs) [6,9,10].

This situation is depicted on Figure 1-1.

Regarding the backbone network architecture, legacy [5] or modern paradigms [11] can be followed. Moreover, the evolution of wireless access networks is frequently referred to as B3G (Beyond the 3rd Generation) systems [3][4].

In the B3G era Network Operators (NOs) will have to address increased complexity, with respect to today. Complexity derives from two main sources:

- The inevitable heterogeneity of the network and terminal infrastructure;

- The user requirements that associate the B3G era with advanced services/applications, provided seamlessly and ubiquitously.

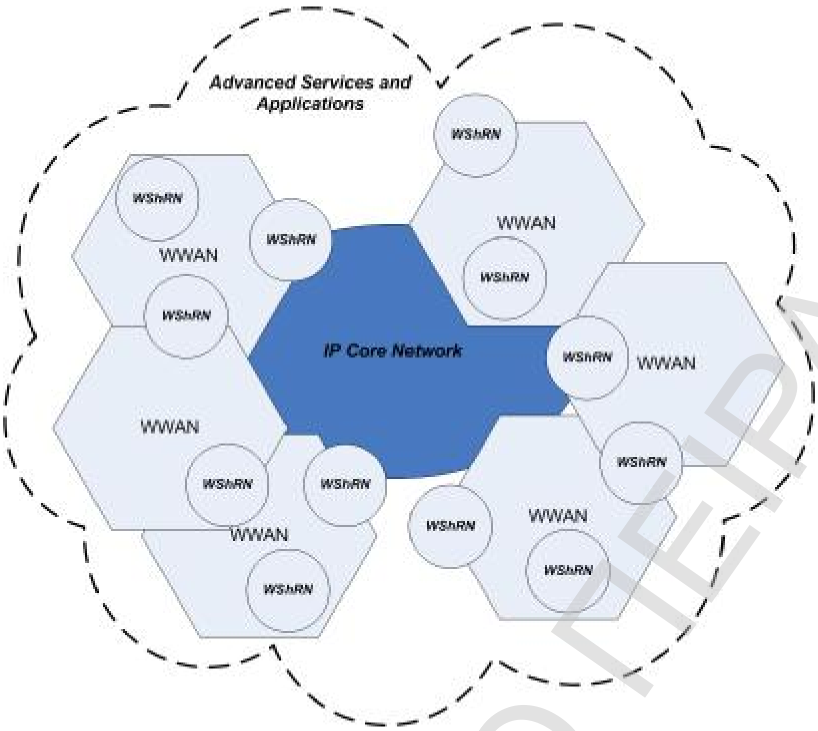


Figure 1-1: Overview of the Wireless World in the B3G Era.

To meet these objectives, NOs have to deploy complex network topologies of heterogeneous nature. The different RATs will have to co-exist, and be complementarily (and efficiently) exploited. A way to achieve this is by introducing the reconfigurability concept that is covered in the next sub-section. Each RAT has different capabilities, in terms of capacity, coverage, mobility support, cost, etc. Therefore, each RAT is best suited for handling certain situations. In this respect, a NO will have to rely on different RATs for raising the customer satisfaction, and achieving the required Quality of Service (QoS) levels, cost-effectively. QoS refers to performance (e.g., bit-rate, delay, etc.), availability (e.g., low blocking probability), reliability (e.g., low dropping or handover blocking probability), as well as security/safety (indicated also in [12]).

Another increasingly important engineering challenge in today's wireless communications domain is the proper management of the electromagnetic radio spectrum, a valuable yet limited natural resource. The current static assignment of the radio spectrum in combination with the often criticized governments' overregulation, leads to underutilization situations [13][14]. Furthermore, the disparate and highly varying contemporary radio environment calls for intelligent management, allocation and usage of the radio spectrum. One of the most prominent emerging technologies that promise to handle such situations is *cognitive radio*. This is covered in sub-section 1.3.

1.2. Reconfigurability

The concept of reconfigurability has been introduced as an evolution of Software Defined Radio (SDR) [15], often considered as a mean to facilitate interworking among versatile technologies, so as to offer alternative solutions with regards to the selection of Radio Access Technologies in certain service area regions. This is achieved by offering to terminals, as well network elements, the capability to re-adapt specific operation parameters, to the environment requirements. This capability is offered through mechanisms that enable the change of any operation parameters / software protocols of a RAT, which is called “Reconfiguration”, whenever environment conditions and requirements impose it. Reconfiguration can be realized not only by preinstalled software modules (in terminals/network elements), but also through the online (dynamic) download and installation of the necessary software, so as to enable the operation of a certain RAT.

Table 1-1: Basic concepts in wireless communication technologies

Wireless Technology	**Basic features of terminals and network elements**
Conventional wireless communication technologies	One RAT / terminal. One RAT / network element. Static (a priori) specification of network related parameters).
Cooperative B3G networks	Terminals select among numerous RATs. Alternatively, there is the possibility to simultaneously operate numerous RATs. One RAT per network element.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

	Pre-installation (offline) of the necessary software for alternative RATs operation.
Reconfigurable networks	Terminals select among numerous RATs. Alternatively, there is the possibility to simultaneously operate numerous RATs. Dynamic (online) installation and configuration of software parameters for the utilization of the selected RAT, during terminals' ad network elements' operation.

This concept constitutes a major innovation in B3G systems compared to the conventional "static" mode of operation of a today's communication network. In other words, one could easily imagine a network segment, dynamically reconfigurable, so as to utilize a completely new (just introduced) RAT, if considered more appropriate.

The Table 1-1 above summarizes the basic features of some communication-related concepts, as well as some management characteristics of theirs.

In a B3G environment, the terminal and network elements should be able to use different *configurations.* Each configuration is characterised by different parameters, e.g. software, frequency, profile, etc. Reconfiguration on the other hand provides the selection upon the appropriate configurations and the actions that needs to be taken, in order to make a change from the one configuration to the other. In general, a reconfiguration may have implications on numerous OSI layers, with more emphasis given on the PHY/MAC layers, thus it constitutes a cross-layer process). However, focusing on the lower layers, just to provide an in-depth view of this concept, we could claim that a platform that enables its transceivers or terminals to dynamically select the most appropriate RAT and frequency band for operation, could be like the one shown in Figure 1-2 below.

As can be observed, each network element of a B3G infrastructure (this includes also terminals) can readapt its mode of operation (RAT or/and frequency band used),

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

based on external conditions/requirements. This can be achieved by disposing intelligent management mechanisms, such as the ones that form part of this thesis.

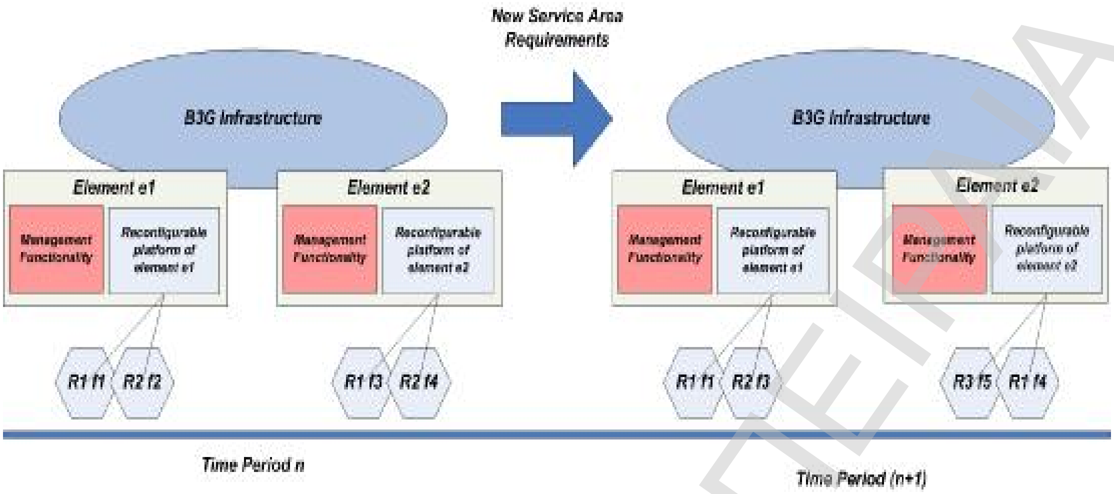


Figure 1-2: Reconfigurable Platform

1.3. Cognition

Cognitive radio systems are based on SDR technology and utilize intelligent software packages that enrich their transceivers with the highly attractive properties of self-awareness, adaptability and capability to learn. A cognitive radio system has the ability to adjust its operating parameters, observe the results and, eventually take actions, that is to say, decide to operate in a specific radio configuration (i.e., radio access technology, carrier frequency, modulation type etc.), expecting to move the radio toward some optimized operational state. In such a process, learning mechanisms that are capable of exploiting measurements sensed from the environment, gathered experience and stored knowledge, are judged as rather beneficial for guiding decisions and actions.

A direction for achieving spectrum efficiency is to equip the infrastructure with *cognitive radio* capabilities. In general, *cognitive systems dispose the capability to retain knowledge from previous interactions with the environment and determine their behaviour according to this knowledge, as well as to other goals and policies, so as to adapt to external stimuli and optimize their performance* [16, 17]. Cognition as a concept may extend to numerous parts of a communication system, i.e. to network segments, access points and even terminals. *Cognitive systems* determine their behaviour in a reactive or proactive manner, based on the external, environmental stimuli, as well as their goals, principles, capabilities, experience and knowledge. In this respect, future cognitive radio devices will have the capability, or luxury, to choose on the fly the radio configuration, by taking into account the context of operation (device status and environment aspects), goals, policies, profiles and capabilities, and machine learning (for representing and managing knowledge and experience). In a more general sense, the term *radio configuration* or simply *configuration* refers to a chosen carrier frequency and a specific radio access technology (RAT) but can be extended to include other operating parameters like transmission power, modulation type etc. This definition also allows a spectrum band to be used for operating in different RATs, in accordance with the flexible spectrum management concept [18].

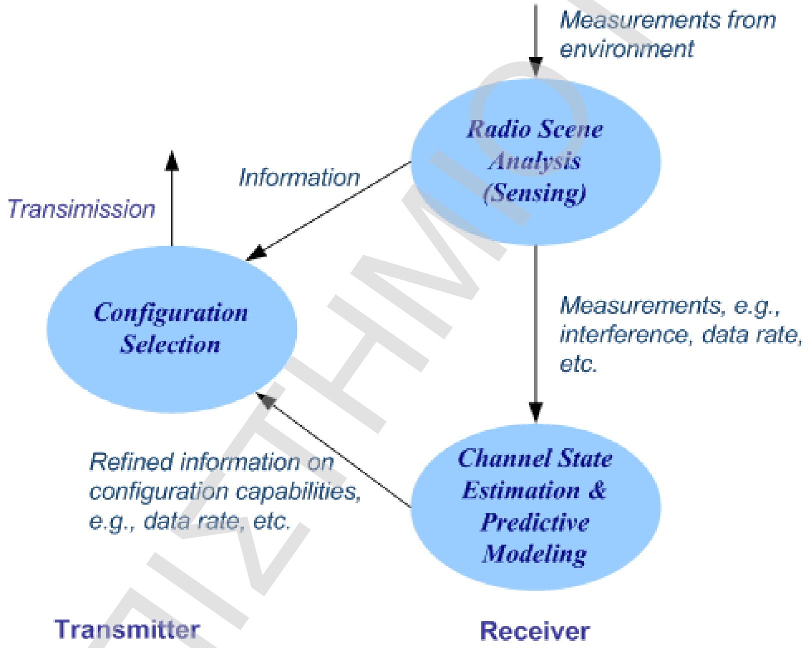


Figure 1-3: Simplified representation of cognitive radio cycle.

In [19], a typical cognitive radio operation is presented as a simplification to the "cognition cycle" initially described in [20][21] and can be divided into three, tightly interconnected phases (see Figure 1-3). The first is *radio-scene analysis*, during which different configurations are probed, and the respective environment conditions, e.g. interference related, are sensed. The second is *channel estimation and predictive modelling*, during which the capabilities of configurations are discovered (discovery process) and accordingly assessed, based on the measurements of the previous phase; moreover, experience and knowledge can be exploited in this phase. The third is *configuration selection*[1], during which the transmitter sends the desired signal by means of the "best" radio configuration (RAT, frequency, modulation, transmit power etc), as it derives from the information of the previous two phases.

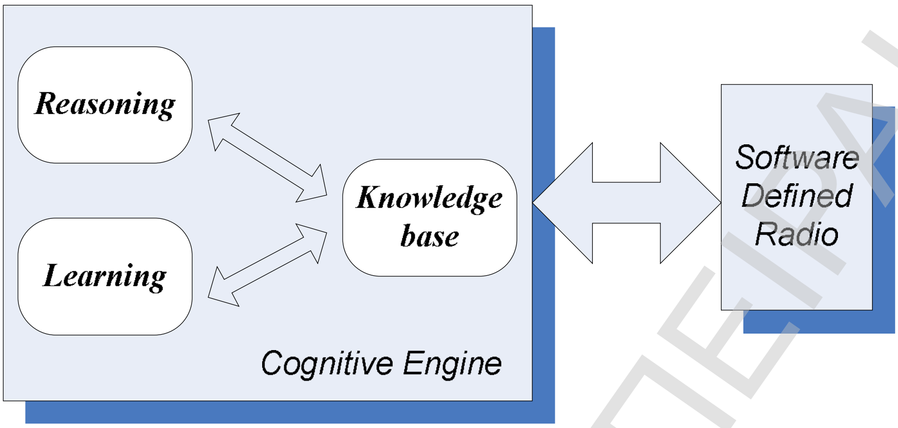


Figure 1-4: Cognitive radio engine.

The approach that is adopted here is that a cognitive radio is born by the enhancement of a software radio with cognition capabilities. Those capabilities are often provided by an intelligent software package called a *cognitive engine*, as proposed in [22] and depicted in Figure 1-4, albeit, according to FCC [23], 'neither having software nor being programmable are requirements of a cognitive radio'. Within the cognition cycle, the cognitive engine derives and enforces decisions to the software-based radio by continuously adjusting its parameters, observing and measuring the outcomes and taking actions to move the radio toward some desired operational state [24]. Meanwhile, cognitive radios are capable of learning lessons and storing them into a knowledge base, from where they may be retrieved, when needed, to assist future decisions and actions. A reasoning engine determines which actions are executable in each environment state. Considering that this could result in a computationally intensive and time-consuming process e.g. in case of a diversified radio environment with numerous state-action pairs, the need for a learning engine seems to be imperative.

[1] This is slightly different from the 3rd phase originally presented in [19] and entitled as transmit-power control and dynamic spectrum management.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

1.4. Need For Advanced Management Functionality

Advanced management functionality is required for terminals in the era of B3G wireless communications [3,4,25]. This is due to: (i) the need to exploit the heterogeneous infrastructure through evolved terminals; (ii) the demanding requirements that should be satisfied and the higher complexity of the business model. Figure 1-5 depicts the main points of the rationale for developing the management functionality.

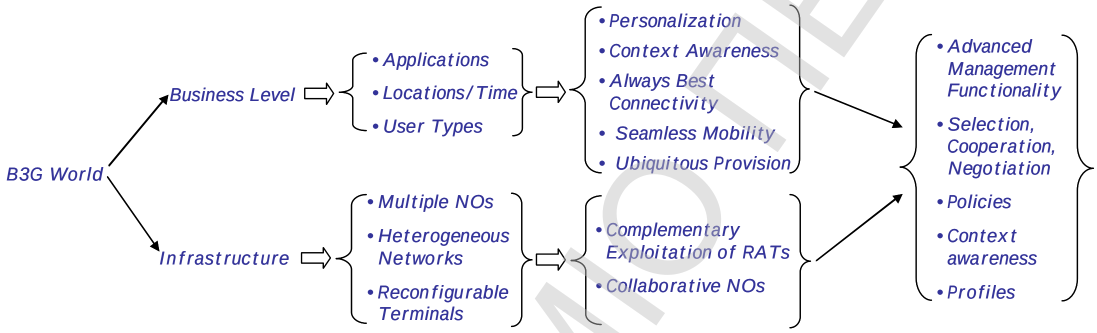


Figure 1-5: Rational for the need for advanced management functionality.

1.4.1. Infrastructure issues

At the network domain, the main B3G feature is the heterogeneity of the infrastructure. This entails that there will be co-existence and complementary exploitation [26],[27] of several RATs, which can be roughly classified into two wide sets. The first set includes the "wireless wide area networking" (WWAN) technologies. This set comprises 2G/2.5G/3G mobile communications [5], the IEEE 802.16 suite of standards [6], WiMAX [7] and broadcasting systems [8]. The second set includes technologies for "wireless short range networks" (WShRNs). This category includes wireless local/personal area networks (WLANs/WPANs) [6], as well as wireless sensor networks (WSNs) [9,10].

At the terminal domain, there are devices equipped with a set of hardware transceivers. In general, each hardware transceiver can operate a set of configurations. Each time, a specific configuration should be selected. (Re-)Configurations enable the (re-)association of the terminal with networks. A reconfiguration may involve changes in one or more of the following features: the RAT, spectrum carrier, transmission power, or the algorithms for modulation, coding and error control. More options exist if reconfigurations can be done “in software” i.e., by activating appropriate software on the hardware [3], [26],[28]; this comes in accordance with the “software-defined-radio” [28] or the “software adaptable network” [29] concepts. Terminals can be capable of “multi-homing”, when they can maintain parallel (simultaneous) connections to radio networks.

1.4.2. Business level issues and requirements

At the business level, a first feature is the existence of more NOs than today. The NOs can be classified as primary or secondary. A primary NO owns a heterogeneous infrastructure that will comprise a set of licensed WWANs, and a set of WShRNs. In principle, a secondary NO operates one or more WShRNs. The spectrum bands required by secondary NOs can be either license exempt, or rented from primary NOs (spectrum owner entities, in general), or used opportunistically [30]. In such a B3G environment, the cooperation between NOs can be envisaged, e.g., between a primary and some secondary ones. In this direction, the next chapter shows also how cooperation and negotiation [31] mechanisms can be used for materializing this cooperation, to the benefit also of the served users.

The general business level objective, especially for primary NOs, is to offer a multitude of applications to multiple types of users, locations (residential, public, or business), time zones. Applications can include voice/audio/data/video content, which is communicated in a conversational, interactive, streaming, or background manner [5]. Within the different locations and time zones a user may assume various roles, e.g., at work, private-life, shopping etc. In any case, application provision should be done at the best possible QoS and cost levels. As discussed also in [32], QoS specifies levels of performance (e.g., bit-rate, delay, etc.), availability (e.g., low blocking probability, capacity), reliability (e.g., low dropping or handover blocking probability), as well as security/safety.

The general business level objective yields the fundamental requirements that should be satisfied by the wireless B3G infrastructure of any primary NO. These are the need for *personalisation, context awareness, always best connectivity, ubiquitous provision* and *seamless mobility.*

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

- § Personalization derives from the need to support different user types and roles.
- § Context awareness derives from the need to consider the state of the served user, terminal and of their environment. The basic contextual information consists of the user and terminal identities and roles, as well as the location and time zone. From these, advanced reasoning can be deduced.
- § Always best connectivity derives from the need to optimally offer the diverse applications, in terms of QoS and cost, taking into account personalization and context information.
- § Ubiquitous provision is the requirement to offer always best connectivity, everywhere.
- § Seamless mobility renders the users agnostic of the complexity of the underlying infrastructure.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

References

[1] P.Demestichas, A.Katidiotis, D.Petromanolakis, V.Stavroulaki, "Management System for Terminals in the Wireless B3G World", accepted for publication to the Wireless Personal Communications journal.

[2] P. Demestichas, G. Dimitrakopoulos, K. Tsagkaris, V. Stavroulaki, A. Katidiotis, "Introducing Cognitive Systems to the B3G Wireless World", Cognitive Wireless Networks: Concepts, Methodologies and Visions Inspiring the Age of Enlightenment of Wireless Communications", pp. 253-269, Springer, 2007

[3] European Commission, 6th Framework Programme (FP6), Information Society Technologies (IST), Project End-to-End Reconfigurability (E2R), http://e2r2.motlabs.com, 2007

[4] Wireless World Research Forum (WWRF), www.wireless-world-research.org, 2008

[5] Third (3rd) Generation Partnership Project (3GPP), Web site, www.3gpp.org , 2007

[6] Institute of Electrical and Electronics Engineers (IEEE), 802 standards, www.ieee802.org, 2007

[7] WiMAX Forum, http://www.wimaxforum.org, 2007

[8] Digital Video Broadcasting (DVB), Web site, www.dvb.org, 2007

[9] Bluetooth, www.bluetooth.com, 2007

[10] ZigBee Alliance, www.zigbee.org, 2007

[11] "Wireless mesh networking: theories, protocols and systems", special issue in the *IEEE Wireless Commun. Mag.*, vol. 13, No. 2, April 2006

[12] W. Hasselbring, R.Reussner, "Towards trustworthy software systems", IEEE Computer, Vol. 29, No. 4 April 2006

[13] European Radiocommunications Committee (ERC), "European table of frequency allocations and utilizations frequency range 9 kHz to 275 GHz," ERC Report 25, January 2002

[14] Susan Moore, "Managing the Radio Spectrum: Hands-On or Back-Off?," *IT Professional*, vol. 6, no. 2, pp. 49-55, Mar/Apr, 2004

[15] Software Defined Radio Forum (SDRF), www.sdrforum.org, 2008

[16] P.Demestichas, D.Boscovic, V.Stavroulaki, A.Lee, J.Strassner, "m@ANGEL: autonomic management platform for seamless wireless cognitive connectivity to the mobile Internet", *IEEE Commun. Mag.*, Vol. 44, No.6, June 2006

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

[17] R. Thomas, L. DaSilva, A. MacKenzie, “Cognitive networks”, In Proc. 1st IEEE Symposium on Dynamic Spectrum Access Networks 2005 (DySPAN 2005), Baltimore, USA, pp. 352-360, Nov. 2005

[18] Leaves, P.; Moessner, K.; Tafazolli, R.; Grandblaise, D.; Bourse, D.; Tonjes, R.; Breveglieri, M., "Dynamic spectrum allocation in composite reconfigurable wireless networks," *Communications Magazine, IEEE* , vol.42, no.5, pp. 72-81, May 2004

[19] S. Haykin, “Cognitive radio: brain-empowered wireless communications”, *IEEE Journal on Selected Areas In Communications,* Vol. 23, No. 2, pp. 201-220, Feb. 2005

[20] J. Mitola *et al.*, “Cognitive radio: Making software radios more personal,” *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[21] J. Mitola, “Cognitive radio: An integrated agent architecture for software defined radio,” Doctor of Technology, Royal Inst. Technol. (KTH), Stockholm, Sweden, 2000.

[22] Clancy, C.; Hecker, J.; Stuntebeck, E.; O'Shea, T., "Applications of Machine Learning to Cognitive Radio Networks," *Wireless Communications, IEEE [see also IEEE Personal Communications]* , vol.14, no.4, pp.47-52, August 2007

[23] Federal Communication Commission: ‘Facilitating opportunities for flexible, efficient and reliable spectrum use employing cognitive radio technologies’, ET Docket 03-108, Notice of Inquiry and Proposed Rulemaking and Order (May 2003).

[24] Le, B., Rondeau, T. W., and Bostian, C. W. 2007. Cognitive radio realities. *Wirel. Commun. Mob. Comput.* 7, 9 (Nov. 2007), 1037-1048.

[25] European Commission, 7th Framework Programme (FP7), Information and Communication Technologies (ICT), Project End-to-End Efficiency (E^3), www.ict-e3.eu, 2007

[26] P.Demestichas, G.Vivier, K.El-Khazen, M.Theologou, “Evolution in wireless systems management concepts: from composite radio to reconfigurability”, *IEEE Commun. Mag.*, Vol. 42, No. 5, pp. 90-98, May 2004

[27] P. Demestichas, G. Vivier, G. Martinez, L. Papadopoulou, V.Stavroulaki, F. Galliano, M. Theologou, “Wireless beyond 3G: managing services and network resources”, *IEEE Computer*, Vol. 35, No. 8, pp. 96–98, Aug. 2002

[28] Software Defined Radio Forum (SDRF), www.sdrforum.org, 2008

[29] R. Thomas, D. Friend, L. DaSilva, A. McKenzie, “Cognitive networks: adaptation and learning to achieve end-to-end performance objectives”, *IEEE Commun. Mag.*, Vol. 44, No. 12, Dec. 2006

[30] Federal Communications Commission (FCC), “Notice of proposed rule making and order”, FCC Et docket no. 03-322, Et docket no. 03-108, December 2003

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

[31] N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, C. Sierra and M. Wooldridge, 2001 "Automated negotiation: prospects, methods and challenges", ***Int. J. of Group Decision and Negotiation***, Vol. 10, No. 2, pp. 199-215 , 2001

[32] W. Hasselbring, R.Reussner, "Towards trustworthy software systems", *IEEE Computer*, Vol. 29, No. 4 April 2006

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

2. MANAGEMENT SYSTEM FOR TERMINALS IN THE WIRELESS B3G WORLD

Abstract:

In the era of B3G wireless communications, a NO should satisfy numerous requirements, namely, personalisation, context awareness, always best connectivity, ubiquitous service provision and seamless mobility, as stated in the previous chapter. A NO can efficiently satisfy such requirements by relying on the different available radio networks of the heterogeneous infrastructure, and potentially on other cooperating networks. In this respect, the NO should possess advanced management mechanisms for driving its users to the most appropriate RATs that live up to the requirements. The presentation of such a management system is the specific contribution of this chapter. The system is called Reconfigurable Terminal Management System (RTMS). In general, it provides the means for profile modelling, the acquisition of monitoring/discovery/context information, and the negotiation and selection of configurations, based on information deriving from policies, as well as the profiles and the context. The work on this chapter focuses on the role and the information of the components of the RTMS. Concrete functionality for accomplishing the role is also presented. Nevertheless, the system is open to the integration of alternate functionality. The discussion includes a business case that presents in high level terms the role of the management system, a detailed description of the components of the this management system and results that show the efficiency of the management schemes. A summary and further research challenges, conclude this chapter. Parts of this chapter have been published in [1][2].

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

MANAGEMENT SYSTEM FOR TERMINALS IN THE WIRELESS B3G WORLD

2.1. Introduction

2.1.1. Need For Advanced Management Functionality

In the B3G world, a primary NO can *efficiently satisfy the requirements*, that derive from the business level, *by relying on the different RATs of its heterogeneous infrastructure, and potentially on other cooperating networks*. Each radio network has different capabilities, e.g., in terms of capacity, coverage, mobility support, cost, etc., and therefore, will be better suited for handling certain situations. In this respect, *the NO should possess advanced management mechanisms for driving its users to the best networks that satisfy the requirements*. The management mechanisms should be responsible for *selecting* the appropriate configurations of the user's terminal device, in order to associate it with the most appropriate radio networks.

The presentation of such a management system is the specific contribution of this chapter. The system is called ***R****econfigurable* ***T****erminal* ***M****anagement* ***S****ystem (RTMS)* [3, 4, 5]. In general, it provides the means for profile modelling, the acquisition of monitoring/discovery/context information, and the negotiation and selection of configurations, based on information deriving from policies, as well as the profiles and the context. It is used by the NO for associating its users with the most appropriate radio networks. Since some of the radio networks may belong to other NOs, part of these management mechanisms is a suite of negotiation strategies [6], which follow the reversed English and Dutch models, and allow the primary NO to interact with cooperating networks RATs. In [5] the structure of the management was introduced, together with the general role of the components. The contribution of this chapter is the formal description of the information associated with the components of the management system, and the concrete description of the functionality of the components for accomplishing their role. The system is open to the integration of alternate functionality. Therefore, this work is open for interworking with other

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

management systems of B3G networks, which can be targeted to the management of access points or network segments.

2.1.2. Relation with previous work

This work extends the legacy state-of-the art by exploiting and extending achievements in the area of the context awareness, profile modelling, policies, and configuration negotiation and selection strategies.

Firstly, it formally defines the context information that should be obtained by RTMS. This information can be obtained through various means, e.g., general context acquisition mechanisms (e.g., [7,8]), and discovery schemes (e.g., [9]), which can be sensing-based (e.g., [10,11]) and/or pilot channel based (e.g., [12]). In a similar manner, this work contributes, by providing the RTMS-relevant view, in the area of profile modelling [13] and policy-based management [14]. Compared to other recent, prominent approached in the literature [15,16,17], this work follows more classical approaches for the modelling of user preferences, since it relies on “utility” [18] values instead of distinct weights per QoS parameter, which may be more difficult to obtain. This work specifies a concise and powerful set of information, on which the RTMS role can be based.

Regarding the configuration negotiation (in other words the cooperation between NOs), this work starts from standard negotiation protocols [6], which are extended and applied to the wireless B3G world. In the area of configuration selection, work is based on an objective function (OF) that takes important aspects into account, namely, the user interest for the QoS levels offered, the affordable cost for these QoS levels, the priority of configurations and the cost of reconfiguring the terminal. The configuration selection (decision-making) scheme aggregates, in the input, important features like context, user preferences, terminal capabilities and policies from the network operator. Based on this input, decisions can be taken in an autonomic fashion [19,20], with limited collaboration [21], e.g., for the provision of policies, from the network side. In the short term, decisions of RTMS can be enforced by means of various mechanisms, for instance the media independent handover mechanisms [22]. Likewise, RTMS can benefit from emerging layer two standards defined for service

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

continuity in heterogeneous networks [22]. In the mid term, RTMS functionality can be integrated as part of these standards.

The chapter evolves as follows. The starting point, in section 2.2, is the presentation of the business case that presents in high level terms the role of the management system. Section 2.3 introduces the overall management system. Sections 2.4 – 2.7 are targeted to the components of the management system. Section 2.8 provides results, which show the behaviour and efficiency of the management schemes. This article concludes with a summary of the major concepts presented and an overview of further research challenges.

2.2. Business Case

Figure 2-1(a) depicts elements of the business case (static view). The focus is on a user that has subscriptions with a certain primary NO. The user is moving in an area, at a certain time period. The user is typically interested for applications/content. The respective terminal has some capabilities. Various RATs may serve the area, and consequently, the terminal. These belong either to the infrastructure of the primary NO, or to cooperating NOs. The RTMS system, which is distributed over the terminal and the network infrastructure, will enable the NO to maintain the user always best connected.

Figure 2-1(b) depicts interactions between entities in the business case (dynamic view). It consists of phases targeted to: (i) user/terminal authentication and localization; (ii) context acquisition; (iii) cooperation and negotiation between NOs; (iv) policy-based selection of the best configurations (RATs); (v) user/terminal operation based on the instructions received in the previous phase.

The first general phase is concerned with security and localization issues. The terminal uses an initial channel for communicating with the NO. The initial channel is necessary [22],[23] in a B3G environment, especially when spectrum is flexibly assigned to RATs. The user and the terminal device provide information that enables authentication and localization up to a certain granularity (not necessarily very fine grained). The user and terminal identities will be used for finding personalisation information. Location information will be used for context acquisition (next phase).

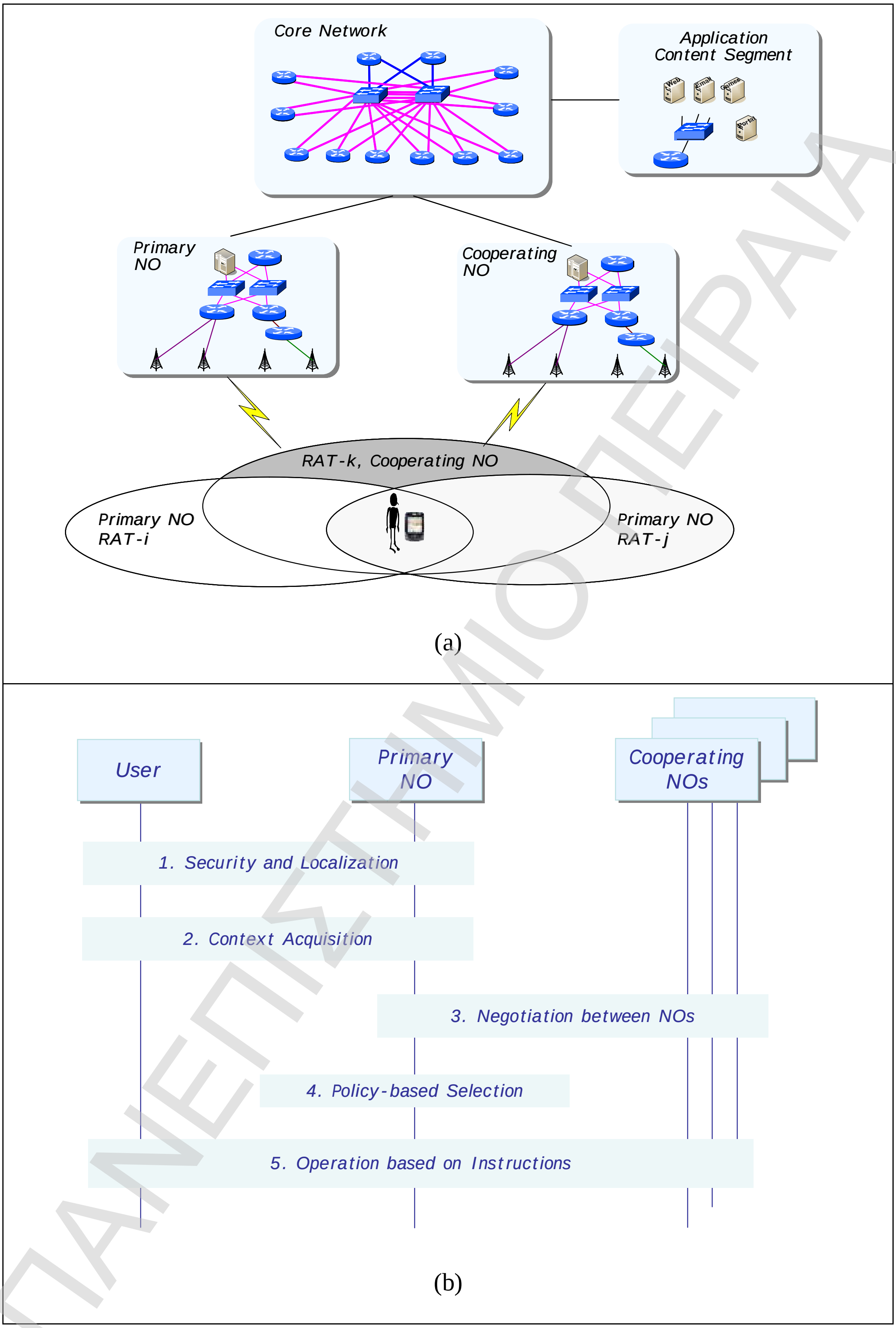


Figure 2-1: Business case. (a) Static view: user/terminal, NOs, application/content segment. (b) Dynamic view: interactions for selecting the configurations that will lead to association with the best radio networks.

The aim of the second general phase is context acquisition. The discovery of available configurations in the area of the user/terminal is of primary importance. In this respect, the NO can provide initial information to the terminal. This information can concern RATs and frequencies in the area of the terminal location. Based on this information, the terminal can tune to the different configurations and discover details on the potential performance of these RATs. Moreover, the terminal can complement the information with details, related to alternate RATs that may potentially collaborate with the primary NO.

The third general phase comprises the cooperation and negotiations between the primary NO and the alternate ones discovered during the previous phase.

In the fourth phase, the NO conducts the policy based selection of the most appropriate RATs that will maintain the user always-best connected, in a personalised, context-aware and seamless manner.

In the fifth phase, the terminal operates according to the instructions received in the previous phase.

It can be envisaged that whenever there will be a change in location the process described above will be applied again. This will contribute to the requirement for ubiquitous provision.

The next section describes the management system that supports the business case described.

2.3. Management System

Figure 2-2 depicts a high level view of the RTMS that supports the business case described in the previous section and satisfies the requirements identified in the introduction. The RTMS consists of four main components, which are called:

§ Profile modelling;

§ Context acquisition;

§ Policies;

§ Configuration negotiation and selection (CNS), based on policies, profiles and context information.

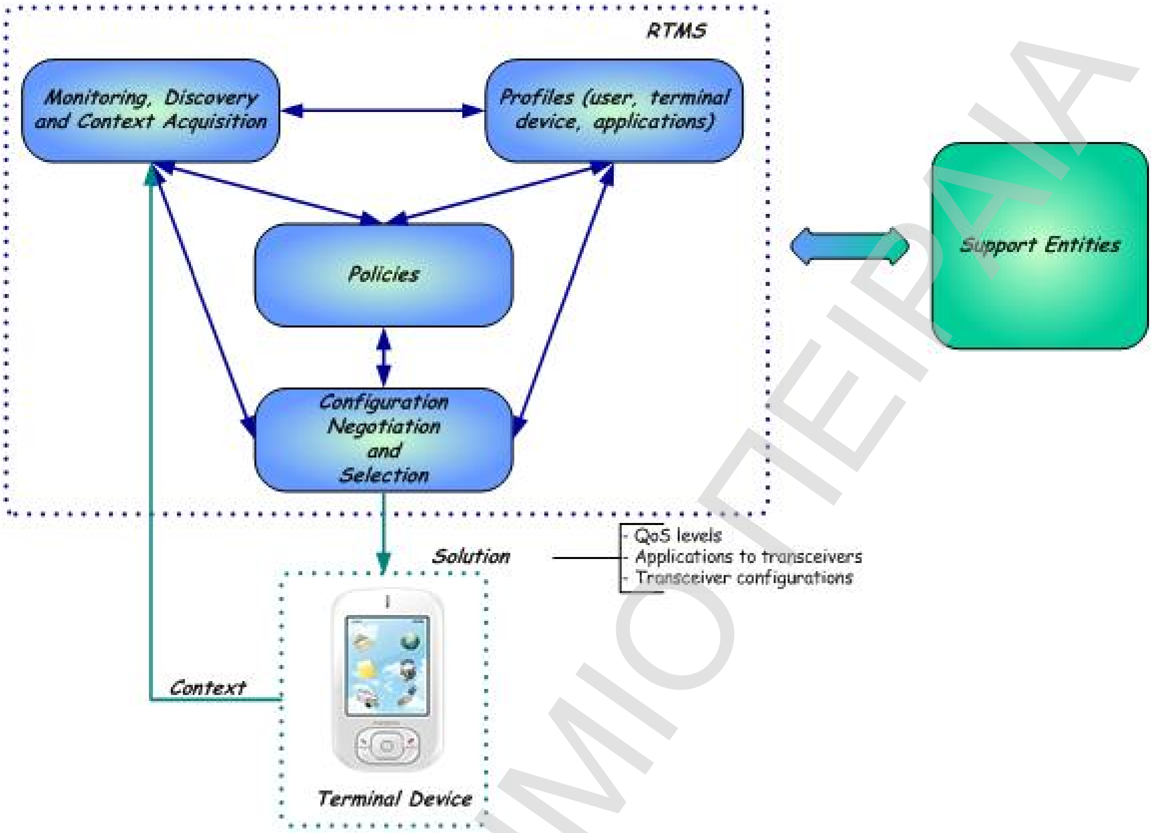


Figure 2-2: Functional architecture of the Reconfigurable Terminal Management System (RTMS).

The goal of the management system is to decide on three main configuration aspects: (i) allocation of applications to QoS levels; (ii) allocation of applications to transceivers; (iii) configurations of transceivers.

All components collaborate and contribute to the satisfaction of the requirements identified in the introduction. Nevertheless, each component has a primary requirement that is addressed.

§ The profile modelling component supports the requirement for personalisation.

§ The context acquisition component offers context awareness.

§ The policies and CNS components offer the rules and the optimisation functionality necessary for always-best connectivity, ubiquitous provision and seamless mobility.

As depicted (Figure 2-2), the RTMS can collaborate with support entities, which are management components of the network infrastructure. For instance, following policy-oriented paradigms [14], these components can offer the means for policy definitions, and act as policy decision points that also take into account other terminals in the area.

The distribution of the components and of their functionality is flexible. It is not necessary to follow one of the two extremes, which are to deploy it completely in the terminal or in the network. Following legacy, standard system engineering processes [24], precise decisions on these aspects can possibly constitute extensions of this thesis.

The rest of this chapter focuses on the role and operation of the components of the RTMS.

2.4. Profile Modelling

The role of this component is to provide information on the preferences, requirements and agreements of the user and of the terminal. Table 2-1 presents the data structures of the component. The user and the terminal device will be denoted as u and d, respectively. The user has subscription with a NO denoted as nw.

Terminal. The terminal device d has a set of wireless network interfaces (transceivers), which is denoted as $TRX(d)$. Each transceiver, $t \in TRX(d)$, can operate a set of RATs $RAT(d,t)$. For each transceiver, $t \in TRX(d)$, and RAT, $r \in RAT(d,t)$, there is a set of frequency carriers, $FRQ(d,t,r)$. Essentially, the information, up to this point, yields a set of transceiver configurations, $CFG(d.t)$. Each transceiver can operate with different configurations. One configuration is operated at a time. It is assumed that each transceiver configuration $c \in CFG(d,t)$ is described as $c = (r, f)$, where $r \in RAT(d,t)$ and $f \in FRQ(d,t,r)$. The other elements of the configuration (power, modulation, coding, error control) are part of the control domain, and therefore, are affected by the management system in a less direct way. Finally, the terminal device has a set of applications/services, $SRV(d)$, that it can support.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Table 2-1: Data structures of the profile modelling component of the RTMS.

User profile	u	Identity of user.
	nw	Identity of NO with which user u has a subscription.
	$APL(u)$	Set of applications for which user u has subscriptions with the NO nw.
	$QOS(u,s)$, K	Sets of QoS levels at which user u can use application $s \in APL(u)$. Each QoS level, $q \in QOS(u,s)$, consists of a set of K parameters, which specify the availability, performance and reliability levels required during the application provision.
	$UTL(u,s,q)$	Set of utility volume values associated with the use of application $s \in APL(u)$, at QoS level $q \in QOS(u,s)$, by user u.
	$CST(u,s,q)$	Maximum tolerable cost values, for user u, when application $s \in APL(u)$ is used at QoS level $q \in QOS(u,s)$.
Terminal device profile	d	Identity of terminal device.
	$TRX(d)$	Set of transceivers of terminal device d.
	$RAT(d,t)$	Set of Radio Access Technologies that can be operated at transceiver $t \in TRX(d)$.
	$FRQ(d,t,r)$	Set of frequencies that transceiver $t \in TRX(d)$ is capable of using, when RAT $r \in RAT(t,d)$ is operated.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

	$CFG(d, t)$	Set of transceiver configurations for $t \in TRX(d)$. Each transceiver configuration $c \in CFG(t, d)$ can be described as $c = (r, f)$, where $r \in RAT(t, d)$ and $f \in FRQ(r, t, d)$.
	$SRV(d)$	Set of applications/services that can be supported by device d.

User. The user may use (i.e. has subscriptions for) a set of applications/services $APL(u)$. For each application, $s \in APL(u)$, there is a set of QoS levels, $QOS(u, s)$, which interest the user and are agreed with the NO (through the subscription). Each QoS level, $q \in QOS(u, s)$, comprises a set of K parameters (key performance indicators). Each parameter $k = 1, \ldots K$, is related to a specific aspect that falls within the availability, performance or reliability categories [25].

For each user, u, application, s, and QoS level, q, there can be a set of utility volume values [18], $UTL(u, s, q)$, and a set of cost-related values, $CST(u, s, q)$. Essentially, each time an appropriate utility and cost value will be used, from the set, depending on the context (situation) in which the user is found. The consideration of sets of utility and the cost values, in this part of the component, entails flexibility to have different values in different contexts, as discussed in the next section.

The utility from here on expresses the user preference for a QoS level, with respect to the other permissible ones. A higher utility value means that there is a higher user interest for the particular QoS level. The utility value influences the decision regarding the best configuration, with which the user should be associated. Users should be associated with networks and RATs that can offer the most preferred QoS levels. The cost values can be seen as the maximum price that the user u has agreed for the use of application s at QoS level q.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

2.5. Context Acquisition

This component reveals the context in which the user and the terminal are found. It provides the means for understanding the status of the user, of the respective terminal device, and of their environment. Contextual information is split into categories. The first category is the basic information. The second category is the refined user profile information. Finally, the third category is the information provided by the *monitoring* and *discovery* procedures. Table 2-2 presents the data structures of the component.

Table 2-2: Data structures of the context acquisition component of the RTMS.

Basic information	$X(u)$	Set of contextual situations in which u may be found.
	$x = \{u, d, n, loc, rol\}$	An element of set $X(u)$. It refers to a user u ("who"), carrying device d, in time epoch n ("when"), moving in location loc ("where"), and assuming (conducting) a certain role (worker, private-life, etc.), rol ("what").
Refined profile information	$APL_C(u,x)$	Set of applications that interest user u in context x. This set is a subset of the overall set of applications and those supported by the device, i.e., $APL_C(u,x) \subseteq APL(u) \cap SRV(d)$.
	$QOS_C(u,s,x)$	Set of QoS levels, at which user u wants to use application $s \in APL_C(u,x)$, in context x. This set is a subset of the overall set of QoS levels, i.e., $QOS_C(u,s,x) \subseteq QOS(u,s)$
	$utl_C(u,s,q,x)$ $cst_C^{apl}(u,s,q,x)$	Utility volume and maximum cost associated with the use of application $s \in APL_C(u,x)$, at QoS level $q \in QOS_C(u,s,x)$, by user u, in context x.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

		The values belong to the sets $UTL(u,s,q)$ and $CST(u,s,q)$, respectively.
Monitoring and Discovery	$cfg_C(d,t,x)$	Configuration of transceiver, $t \in TRX(d)$ of device d, in context x.
	$CFG_C(d,t,x)$	Alternate configurations for transceiver $t \in TRX(d)$, which are available in the environment, in context x, according to the discovery procedures. These are a subset of the capabilities of the transceiver, i.e., $CFG_C(d,t,x) \subseteq CFG(d,t)$.
	$cst_C^{cfg}(d,t,c,x)$	Cost of reconfiguring transceiver t of device d, from $cfg_C(d,t,x)$, to a configuration $c \in cfg_C(d,t,x) \cup CFG_C(d,t,x)$ in context x.
	$qos_C^{cfg}(c,x)$	QoS capabilities of a configuration $c \in cfg_C(d,t,x) \cup CFG_C(d,t,x)$, in context x. Each QoS level consists of K parameters.
	$srv_C(d,t,x)$	Applications that are served by transceiver t, of device d, in context x. Inter-layer aspect.
	$qos_C^{apl}(d,t,s,x)$	QoS level offered to application $s \in srv_C(d,t,x)$ by the transceiver t, of device d, in context x.

Basic information. Each user may be found in a set of contexts, $X(u)$. Each element of the set, $x \in X(u)$, provides basic context information, which comprises the user/terminal identities ("who"), the time ("when") and location ("where"), and the user role ("what"). Therefore, x implies that, within a time epoch n, user u, carrying terminal device d, is found in a certain location loc, and assumes (conducts) a certain role (at work, at home, etc.), rol.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Refined user profile information. The following refinements can occur for the profile of user u in each context, $x \in X(u)$: (i) the set of applications, $APL_C(u,x) \subseteq APL(u) \cap SRV(d)$, that interest u; (ii) for each application $s \in APL_C(u,x)$ the set of QoS levels, $QOS_C(u,s,x) \subseteq QOS(u,s)$, that interest u; (iii) the utility, $utl_C(u,s,q,x) \in UTL(u,s,q)$, and maximum cost, $cst_C^{apl}(u,s,q,x) \in CST(u,s,q)$, when application $s \in APL_C(u,x)$ is used at QoS level $q \in QOS_C(u,s,x)$.

Set $APL_C(u,x)$ is a subset of the intersection of sets $APL(u)$ and $SRV(d)$. In other words, user u, in context x, is interested in applications for which there is subscription and can be supported by the device.

Monitoring and discovery. The monitoring procedures provide information on the current configuration. They expose the QoS level in each communication layer, as well as inter-layer aspects, namely, the assignment of applications/services to transceivers. The discovery procedures provide information on the alternate configurations available in the area.

In general, the procedures provide, in each context x, the following information: (i) the current configuration, $cfg_C(d,t,x)$, of each transceiver, $t \in TRX(d)$; (ii) the alternate transceiver configurations, $CFG_C(d,t,x) \subseteq CFG(d,t)$, that are available in the environment, and therefore are candidate for use in $t \in TRX(d)$; (iii) the cost, $cst_C^{cfg}(d,t,c,x)$, of reconfiguring $t \in TRX(d)$, from $cfg_C(d,t,x)$ to $c \in CFG_C(d,t,x)$; (iv) the QoS capabilities, $qos_C^{cfg}(c,x)$, of each configuration $c \in CFG_C(d,t,x) \cup cfg_C(d,t,x)$; (v) the set of applications, $srv_C(d,t,x) \subseteq APL_C(u,x)$, served by transceiver $t \in TRX(d)$; (vi) the QoS level, $qos_C^{apl}(d,t,s,x)$, offered to application $s \in srv_C(d,t,x)$ by transceiver t of device d.

The reconfiguration cost can express the time and the resources (battery power, memory space) required to change the configuration of the hardware. This covers also the case of software-reconfigurable transceivers [3,26]. In this case, this cost is a function of the size of the software modules that may have to be downloaded and installed, so as to deploy the configuration.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

2.6. Policies

The component can specify rules that should be followed by the RTMS. Therefore, policies may refine and complement the input designated by the profiles and the context, and moreover, they may specify strategies, algorithms and parameter values, to be followed later by the CNS. Figure 2-3 depicts the scope of policies, i.e., the relevance of rules with respect to various layers of an arbitrary protocol stack. Table 2-3 provides the data structures of the component.

Set $POL(u)$ comprises the overall set of policies that can be used for user u. In a context x, a set of policies $POL_P(u,x) \subseteq POL(u)$, is used.

The set of policies $POL_P(u,x)$ can specify the following aspects that are allowed in context x: (i) the set of transceiver configurations, $CFG_P(d,t,x) \subseteq CFG_C(d,t,x) \cup cfg_C(d,t,x)$, which are allowed for transceiver $t \in TRX(d)$; (ii) the priority, $prt_P(c,x)$, of each transceiver configuration, $c \in CFG_P(d,t,x)$; (iii) the set of applications, $SRV_P(c,x) \subseteq APL_C(u,x)$, which are allowed to be offered through a transceiver configuration c; (iv) the set of QoS levels, $QOS_P(u,s,c,x) \subseteq QOS_C(u,s,x)$, at which an application $s \in SRV_P(c,x)$ can be supported by configuration c; (v) the cost, $cst_P^{apl}(u,s,q,x)$, at which u can use s, at QoS level $q \in QOS_P(u,s,c,x)$.

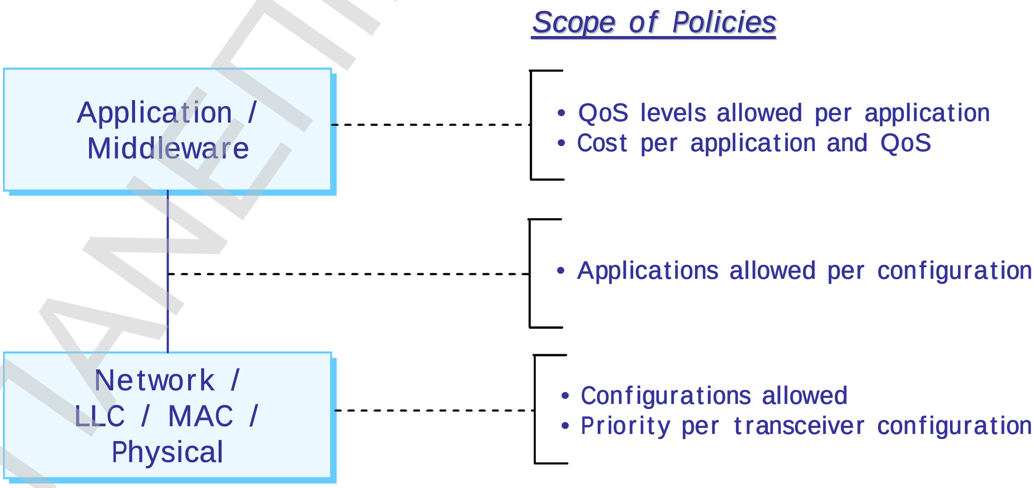


Figure 2-3: Scope of policies with respect to an arbitrary, high-level protocol stack.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

The set $CFG_P(d,t,x)$ comprises the transceiver configurations that can be used by $t \in TRX(d)$, according to the policies that should be applied in the current context x. The $prt_P(c,x)$ parameter influences (encourages or discourages) the selection of a transceiver configuration c by a transceiver $t \in TRX(d)$. The use of transceiver configurations with high priority value should be favoured. A $SRV_P(c,x)$ set also shows which applications are prohibited from being offered through configuration c. The excluded applications are elements of the $APL_C(u,x)$-$SRV_P(c,x)$ set. The sets $QOS_P(u,s,c,x)$ can be returned by the identified support entities (Figure 2-2), which can act as policy decision points that are also in position to take into account other terminals in the area.

The cost $cst_P^{apl}(u,s,q,x)$, of using an application at an offered QoS level should be less than, or equal to, the maximum tolerable cost $cst_C(u,s,q,x)$.

Table 2-3: Data structures of the policy component of the RTMS.

General	$POL(u)$	Overall set of policies that can be used for user u.
	$POL_P(u,x)$	Set of policies that can be used for user u in context x ($POL_P(u,x) \subseteq POL(u)$).
Rules per policy	$CFG_P(d,t,x)$	Set of transceiver configurations that are allowed for $t \in TRX(d)$ according to the policies in $POL_P(u,x)$. These configurations are a subset of those available in the current context, i.e., $CFG_P(d,t,x) \subseteq CFG_C(d,t,x) \cup cfg_C(d,t,x)$.
	$prt_P(c,x)$	Priority allocated to configuration $c \in CFG_P(d,t,x)$, by the policies in $POL_P(u,x)$. The value influences the selection of the configuration by a transceiver. A high priority value favours the selection of the configuration.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

	$SRV_P(c,x)$	Applications that can be offered through configuration c according to the policies in $POL_P(u,x)$. Inter-layer aspect.
	$QOS_P(u,s,c,x)$	QoS levels at which application $s \in SRV_P(c,x)$ can be supported by configuration c, according to the policies in $pol_P(u,x)$. The QoS level belongs to set $QOS_C(u,s,x)$
	$cst_P^{apl}(u,s,q,x)$	Cost associated with the use by u of $s \in APL_C(u,x)$, at a QoS level $q \in QOS_P(u,s,c,x)$, according to the policies in $POL_P(u,x)$.

2.7. Configuration negotiation and selection

This section starts from the description of the overall role and operation of the component. Then, the functionality related to the cooperation with other NOs and the selection of the optimal solution is addressed in more detail.

2.7.1. Operation

Figure 2-4 depicts the overall, general operation of the component. In the first phase, there is the acquisition of the contextual situation encountered, x, and the relevant profile information, of the user and device involved. In the second phase, there is a retrieval of the policies, $POL_P(u,x)$, that can be used. In the third phase, policy rules are evaluated, and corresponding actions are taken. Essentially, an optimisation problem is formulated appropriately, so as to take into account the conditions that derive from the policy rules.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

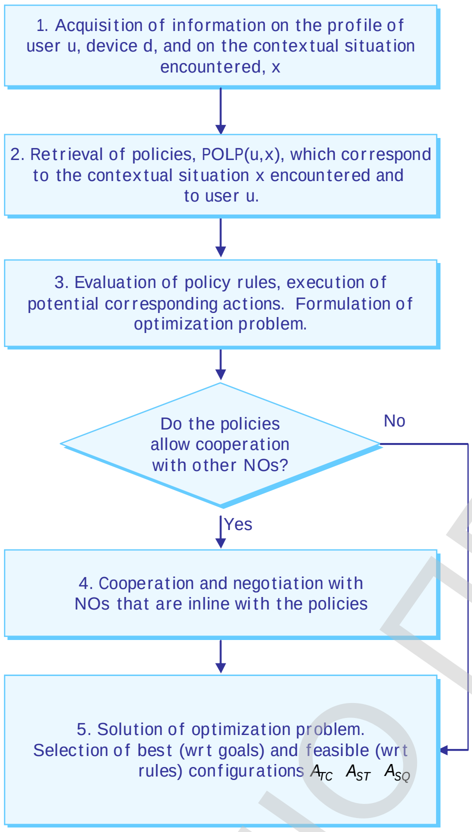


Figure 2-4: Overall operation of the configuration negotiation and selection component.

An important aspect that is investigated in the third phase is whether the policies allow the cooperation with other NOs. If this is the case, in the fourth phase, there is negotiation with the NOs, which are inline with the policies. Otherwise there is migration to the fifth phase. The information that may be collected, through the negotiation process, is appended to the data that will be used as input for the optimisation process.

The optimisation problem is solved, during the fifth general phase. As already presented, the output is the selection of the optimal configurations, i.e., the best combination of, (i) allocation of applications/services to QoS levels, A_{SQ}, (ii) allocation of applications/services to transceivers, A_{ST}, and (iii) transceiver configurations, A_{TC}.

2.7.2. Cooperation with NOs

A subset of the configurations of the $CFG_P(d,t,x)$ set may not belong to the infrastructure of nw. They may belong to a set of cooperating NOs, denoted as

$CPN_P(nw,x)$. These are the NOs that can potentially cooperate with nw, according to the policies in $POL_P(u,x)$, and have configurations in the $CFG_P(d,t,x)$ set. Therefore, there can be negotiations between nw and the NOs of the $CPN_P(nw,x)$ set.

The objective of the negotiation can be to give cost values that have the form, $cst_P^{cpn}(u,s,q,c,x)$. Each value gives the cost at which a configuration c, belonging to a cooperating NO, of set $CPN_P(nw,x)$, can support application, $s \in APL_C(u,x)$, and QoS level, $q \in QOS_P(u,s,c,x)$. These costs will be taken into account by the optimisation scheme discussed in the next sub-section. Eventually, these costs may be used for configuring the $cst_P^{apl}(u,s,q,x)$ values.

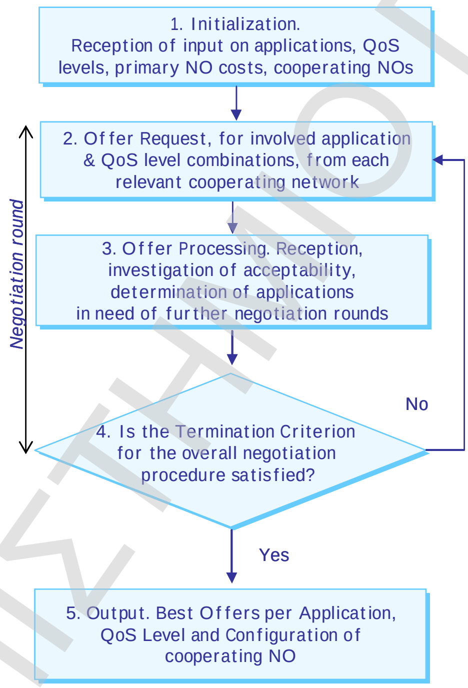


Figure 2-5: Negotiation scheme.

It should be noted that there is a relation between the cost values which should have the following form: (i) $cst_C(u,s,q,x) \geq cst_P^{apl}(u,s,q,x)$; (ii) the cost that will be offered to the user, $cst_P^{apl}(u,s,q,x)$, can be higher than the cost offered by the configuration, which will be selected for the support of the application at the QoS level. This means that according to policies there may be offered prices lower than the maximum acceptable ones in the current context. Moreover, as an outcome of the cooperation and negotiation there may be the possibility of further price reductions. It remains an issue of future research to investigate issues related to regulation activities required, in order to ensure that the NO's cooperation between will result in direct benefits (i.e., dynamic discounts) for the users.

Figure 2-5 depicts the general approach for reaching the negotiation objectives. The general approach can be customized to correspond to the specific strategies (reversed English and Dutch) [6], which have been integrated in the CNS component of the RTMS. The general and the specific approaches are discussed in more detail in section 2.9, Cooperation and Negotiation Strategies.

2.7.3. Selection of optimal configurations

In order to determine the most appropriate solution, an objective function should be optimised, under certain constraints, through appropriate algorithms.

Objective function. The objective is to offer the most preferred QoS levels, and to achieve this through the most cost-efficient configurations. In this respect, decision variables are required. These variables can be denoted as x_{tc}, y_{st} and z_{sq} where $t \in TRX(d)$, $c \in CFG_P(d,t,x)$, $s \in APL_C(u,x)$ and $q \in QOS_P(u,s,c,x)$. Variables x_{tc}, y_{st} and z_{sq} can take value 1 (0), if configuration c is (is not) assigned to transceiver t, if application s is (is not) assigned to transceiver t, and if application s is (is not) assigned to QoS level q, respectively. Based on the definitions above, the following objective function can be defined.

Maximise: $OF(A_{SQ}, A_{ST}, A_{TC}, u, d, x) =$

$$\sum_{s \in APL_C(u,x)} \sum_{q \in QOS_P(u,s,c,x)} z_{sq} \cdot utl_C(u,s,q,x) +$$

$$\sum_{t \in TRX(d)} \sum_{c \in CFG_P(d,t,x)} x_{tc} \cdot prt_P(c,x) -$$

$$\sum_{t \in TRX(d)} \sum_{c \in CFG_P(d,t,x)} \sum_{s \in APL_C(u,x)} \sum_{q \in QOS_P(u,s,c,x)} x_{tc} \cdot y_{st} \cdot z_{sq} \cdot cst_P(u,s,q,c,x) -$$

$$\sum_{t \in TRX(d)} \sum_{c \in CFG_P(d,t,x)} x_{tc} \cdot cst_C^{cfg}(d,t,c,x) \qquad (1)$$

The objective function consists of four main terms. The first term expresses the utility volume that derives from the QoS levels that will be offered to the user. The second

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

term expresses the aggregate priority of the selected configurations. The third term states the (monetary) cost at which the QoS levels can be offered to the user. The fourth terms includes the cost of reconfiguring the transceivers.

In general, the objective function aims at the following. First, to offer the QoS levels that maximise the associated utility volume (i.e., select the QoS levels that have high utility volumes). Second, to use the transceiver configurations that have high priorities according to the policies. Third, to have as small as possible monetary cost for the user. This term enables the selection of a configuration that belongs to a cooperating NO, in case the cost that will be applied for the support of an application, at a certain QoS level, is smaller compared to what is supported by nw . The fourth term aims at the minimisation of the cost of reconfiguring the transceivers.

Constraints. All allocations should comply with the information designated by the profiles, the context and the policies. All applications of the $APL_C(u,x)$ set should be served, i.e., be assigned to one transceiver. Transceivers should be assigned applications of the $SRV_P(c,x)$ set. The QoS level that will be selected for each application $s \in APL_C(u,x)$ should be in the $QOS_P(u,s,c,x)$ set. The transceiver configuration that will be selected for each $t \in TRX(d)$ should be in the $CFG_P(d,t,x)$ set. The assignment of applications to transceivers should respect their QoS capabilities. Moreover, following the proposed model, the following condition should apply regarding the cost:

$$cst_C(u,s,q,x) \geq cst_P^{apl}(u,s,q,x) \geq$$

$$\sum_{t \in TRX(d)} \sum_{c \in CFG_P(d,t,x)} \sum_{s \in APL_C(u,x)} \sum_{q \in QOS_P(u,s,c,x)} x_{tc} \cdot y_{st} \cdot z_{sq} \cdot cst_P(u,s,q,c,x) \quad (2)$$

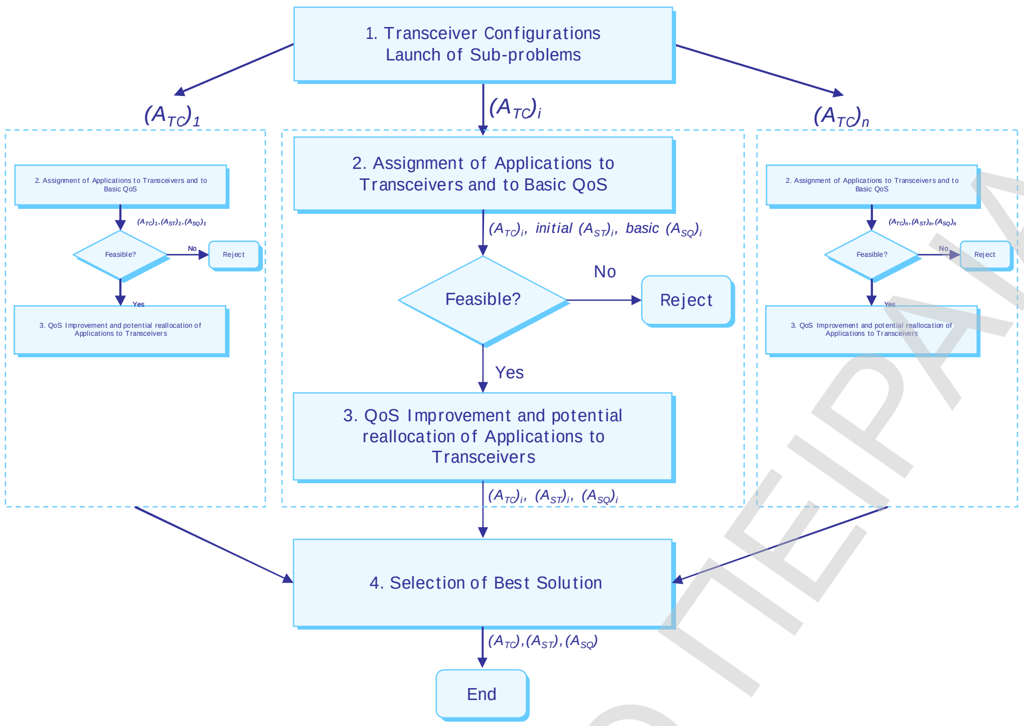


Figure 2-6: Computation of optimal configurations.

Algorithm. From a very general perspective, finding the solutions that optimize relation (1) may be a computationally demanding task [27, 28]. Nevertheless, in several instances that the RTMS will have to address, the complexity may not be very high. In general, computationally efficient algorithms are required, which may provide near-optimal solutions in reasonable time. Classical methods in this respect are simulated annealing [29], genetic algorithms [30], taboo search [31], greedy algorithms [28]. Exhaustive search can also be conducted in case the solution space is not prohibitively large.

The CNS component of the RTMS relies on a hybrid heuristic technique, originally devised for the problem described in [32]. Such an approach derives through the combination of a divide-and-conquer strategy with greedy techniques. Figure 2-6 depicts the approach for solving the overall problem. The scheme is discussed in more detail in section 2.10. The algorithm in [32] is targeted at the management of a (small) set of access points of the network infrastructure. Therefore, the problem addressed there has a larger solution space and more computing resources that can be used for the solution, compared to the RTMS case (smaller solution space and less powerful computing resources). Therefore, the selection scheme in this part of RTMS is

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

designed so as to maintain the complexity low. Primarily, this has impact in the design of the first step (explained in the last paragraph of this subsection).

The scheme consists of four main steps. The first step is targeted to the configuration of transceivers. The second phase allocates applications to transceivers, assuming a basic (nevertheless acceptable) QoS level for all applications. Step three is targeted to the potential improvement of QoS levels. The second and third phases exploit the capabilities of the selected configurations, for assigning applications/services to transceivers and offering the best possible QoS. Finally, step four is targeted to the selection of the best configuration. The fourth phase has the performance of the selected transceiver configurations, or in other words, their scores with respect to the objective function (1). Therefore, the selection of the best one (maximum score in (1)) can be done.

In the first step, a number of sub-problems is launched. In each sub-problem the configuration of transceivers is fixed. The number of sub-problems that can be examined is limited by the networks, which are available in the area, in the particular contextual situation. Intentionally, in the proposed scheme, the number of sub-problems that will be selected for examination is left open to various approaches (for enabling various policies). This choice guarantees that complexity will be at orders of magnitude similar to that of legacy mobility management schemes. This holds because one approach is to follow techniques analogous to what happens today, when terminals seek networks to connect to [37]. Following this approach, for each transceiver, there can be a very small set of the most promising configurations (each configuration can be of a different RAT) available in the area. Then each sub-problem can target the examination of the behaviour of a specific combination of transceiver configurations. Again, the number of combinations that will be examined can be limited. Nevertheless, another approach is to conduct exhaustive search of all the combinations, in case the number of available configurations is not large. Learning schemes, in accordance with the cognitive systems paradigm [33,34,35,36], are also a viable direction, for favouring the selection of the best, based on knowledge and experience, configurations.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

2.8. Results

This section shows the behaviour and efficiency of the RTMS. Comprehensive results from very indicative test cases will be used.

2.8.1. Input description

Profiles. In this part the focus is on aspects identified in section 2.4 (Table 2-1). Aspects that need to be specified are related to the user, applications, QoS levels, utility and cost values, the device and its capabilities. The focus is on a user u that carries a device d and has subscriptions with NO nw.

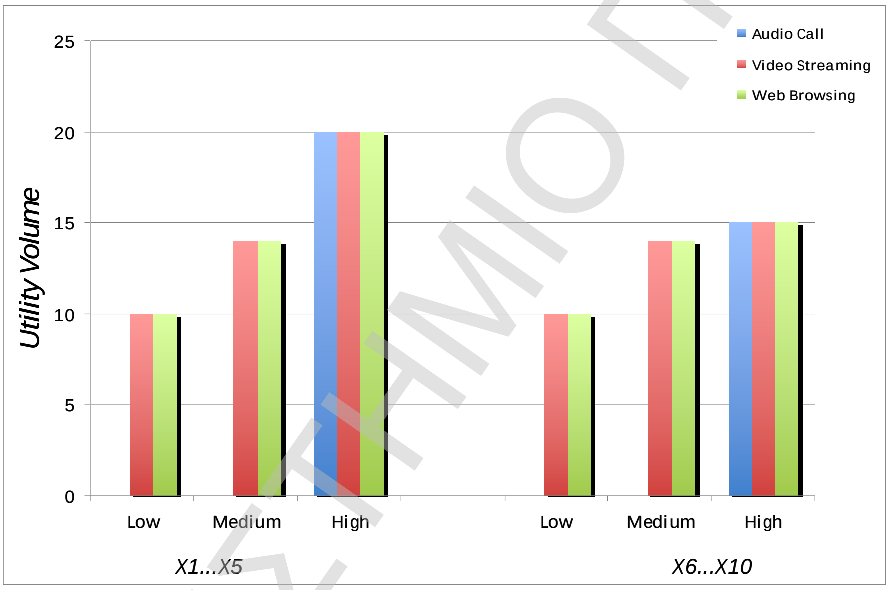


Figure 2-7: Utility values in the different contexts.

The set of subscribed applications is $APL(u) = \{s_1, s_2, s_3\}$ = {audio-call, video-streaming, Web-browsing}. Each QoS level consists of $K = 4$ parameters, namely, the blocking probability, bit-rate, bit-error-rate and dropping probability. The first parameter is related to the application availability, the second and third are related to the performance of the application, while the fourth is related to the reliability of the application provision. The sets of QoS levels are $QOS(u, s_1) = \{q_{11}\}$ = {(1%, 64Kbps, 250 msec, $10^{(-3)}$)}, $QOS(u, s_2) = \{q_{21}, q_{22}, q_{23}\}$ = {(1%, 128 Kbps, 500 msec, $10^{(-4)}$), (1%, 256 Kbps, 500 msec, $10^{(-4)}$), (1%, 512 Kbps, 500 msec, $10^{(-4)}$)}, and $QOS(u, s_3) = \{q_{31}, q_{32}, q_{33}\}$ = {(1%, 128 Kbps, 500 msec, 0.001), (1%, 256 Kbps, 500 msec, 0.001), (1%, 512 Kbps, 500 msec, 0.001)}. The work is generic with respect to the precise values of the QoS parameters. The QoS levels q_{21}, q_{22} and q_{23} can also be characterised as "low", "medium" and "high", respectively. The same characterization can be made for q_{31}, q_{32}, q_{33}. Figure 2-7 and Figure 2-8 depict the overall sets of utility values and maximum cost values, per application and QoS level, for user u, respectively. The individual values from the set are relevant to different contexts, as discussed in the following.

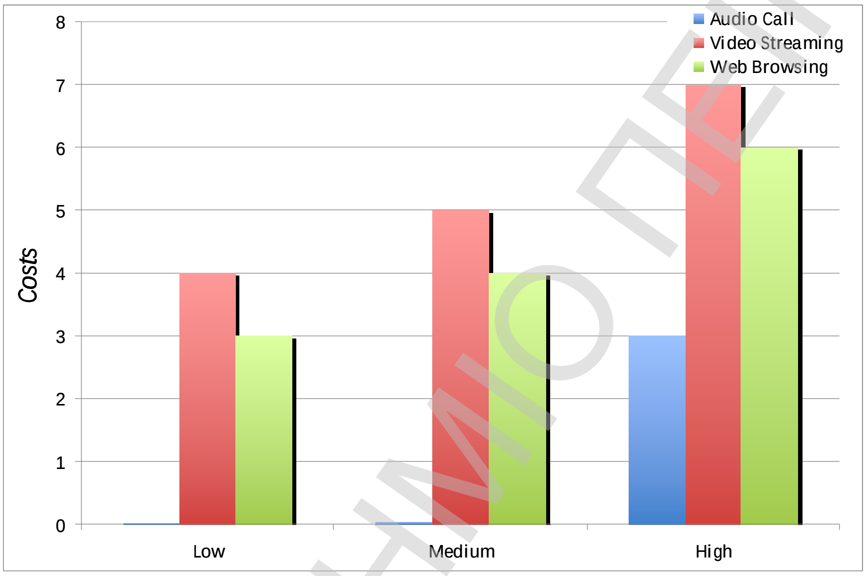


Figure 2-8: Maximum cost values in the different contexts.

Regarding the profile of d, it will be assumed that it has one or two transceivers, i.e., two main scenarios will be considered. Transceivers will be assumed capable of operating with main configurations available in the service area.

Context. In this part the focus is on aspects identified in section 2.5 (Table 2-2). Aspects that need to be specified are the basic information (location, time periods, user roles), the refined profiles (relevant applications, QoS levels and utility/cost values), and finally, the monitoring/discovery information (available configurations and their QoS capabilities).

Figure 2-9 depicts the service area and the coverage provided by the B3G infrastructure of nw, which will be considered in the tests. The tests evolve between arbitrary residential and business sites. Based on the coverage provided by the RATs of the infrastructure of nw, five main locations can be identified, loc_1, … , loc_5, as shown on the figure. Each location is characterized by the coverage that is available in the area. In general, the UMTS [37], WLAN [22] and WiMAX [38] RATs are assumed to be available. The configurations $\{um_1, um_2, um_3, um_4\}$ can associate the terminal with the UMTS RAT. The configurations that can associate the device with the WiMAX and WLAN RATs are denoted as wi and wl, respectively.

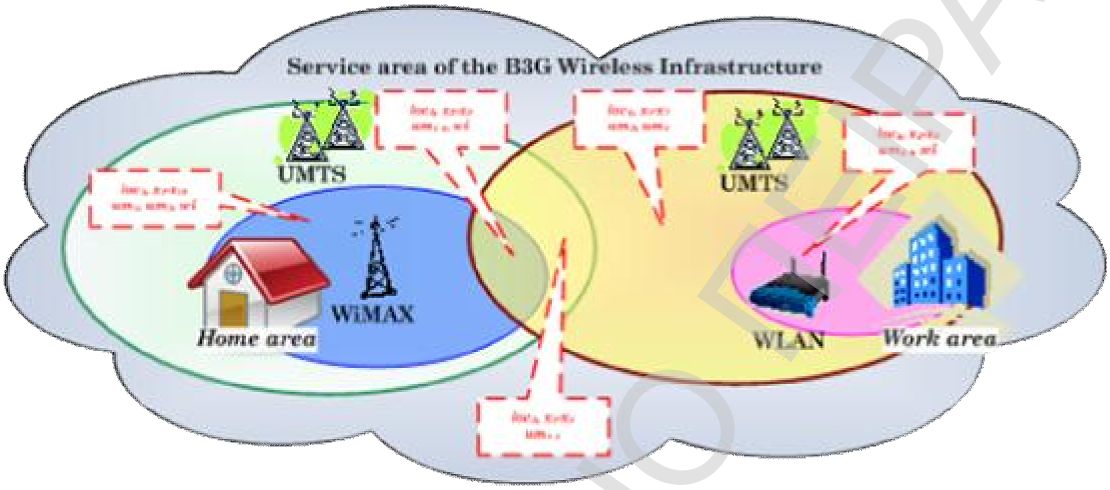


Figure 2-9: Service area considered in the tests and respective coverage (configurations available).

Ten time periods, n_1 to n_{10}, will be considered. In the first five the user role, rol_1, will be "in work". In the next five the user role, rol_2, will be "private life", respectively. These data yield that the set $X(u)$ has 10 elements, i.e., $X(u) = \{x_1, \ldots, x_5, x_6, \ldots, x_{10}\} = \{(u, d, n_1, loc_1, rol_1), \ldots, (u, d, n_5, loc_5, rol_1), (u, d, n_6, loc_5, rol_2), \ldots, (u, d, n_{10}, loc_1, rol_2)\}$.

Regarding the refinement of the profiles it will be assumed that $APL_C(u, x) = APL(u)$, $QOS_C(u, s_1, x) = QOS(u, s_1)$, $QOS_C(u, s_2, x) = QOS(u, s_2)$ and $QOS_C(u, s_3, x) = QOS(u, s_3)$ for all $x \in X(u)$. Figure 2-7 depicts the utility values that are relevant in each context. Figure 2-8 presents the maximum cost values that are relevant in each context. An important observation is that the utility values for the "high" (q_{23}, q_{33}) quality levels are lower in the case of role rol_2 ("private life"). In other words, the user is interested for the high quality level only during the working time.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Table 2-4: Cost of reconfiguring transceivers.

	Reconfiguration Cost
UMTS -> UMTS	1
UMTS -> WLan	3
UMTS -> WiMAX	3
WLan -> UMTS	3
WLan -> WLan	1
WLan -> WiMAX	2
WiMAX -> UMTS	3
WiMAX -> WLan	2
WiMAX -> WiMAX	1

Table 2-4 presents the assumed costs for reconfiguring the transceivers from one RAT to another. Different cases have been considered. The logic that has been followed is that the configurations of the WLAN and WiMAX RATs may have several common components (functions), therefore, the reconfiguration cost may be smaller, compared to that of changing between the UMTS and WiMAX/WLAN configurations, since the component similarities may be fewer.

Policies. It will be assumed that the policies do not impose constraints on the allowed configurations per transceiver. It will be also assumed that $QOS_P(u, s_2, u_i, x) = \{q_{21}, q_{22}\}$ and $QOS_P(u, s_3, u_i, x) = \{q_{31}, q_{32}\}$, where i = 1, 2, 3, 4. This means that the UMTS network cannot offer the high QoS level for the video streaming and web browsing service. In some cases in the scenarios it will be assumed that policies allow the provision of service s_1 only through UMTS. According to this model this can be formally expressed as $SRV_P(wi, x) = SRV_P(wl, x) = \{s_2, s_3\}$. Table 2-5 presents the priorities allocated by the policies to the different RATs (respective configurations).

Table 2-5: Priorities allocated to various transceiver configurations by the policies.

	Priorities
UMTS (1)	8
UMTS (2)	6
WiMAX	8
UMTS (3)	4
UMTS (4)	2
Wlan	10

2.8.2. Output from first scenario

This subsection presents the results from the first scenario (one transceiver available in the device). There is comprehensive information on how the RTMS optimises all the situations that can occur, taking into account the profile, context and policy information.

Table 2-6 illustrates results, when user u requests one particular service. Therefore, the table includes all the potential objective function values, in case u requests service s_1, s_2 or s_3, and the requested service is assigned to all the allowed QoS levels and transceiver configurations. Regarding the reconfiguration cost, which depends from the previous configuration of the terminal device, an average situation is considered (all potential cases regarding the previous configuration of the device). Table 2-6 (a) and (b) present the objective function values for the input that is valid at contexts $x_1, \dots, x_5$, and $x_6, \dots, x_{10}$, respectively. These results are exploited, in the rest of this subsection, for obtaining more general insight on the behaviour of the RTMS.

Table 2-6: Contribution to OF values, resulting from the RTMS environment (a) x1..x5, (b) x6..x10

		Objective Function (x1-x5)					
		UMTS 1	UMTS 2	WiMAX	UMTS 3	UMTS 4	WLAN
Audio Call	Low	-	-	-	-	-	-
	Medium	-	-	-	-	-	-
	High	25	23	-	21	19	-
Video Streaming	Low	14	13	11	11	7	13
	Medium	17	14	14	12	11	16
	High	-	-	18	-	-	20
Web Browsing	Low	12	10	12	8	6	14
	Medium	15	13	15	11	9	17
	High	-	-	19	-	-	21

(a)

		Objective Function (x6-x10)					
		UMTS 1	UMTS 2	WiMAX	UMTS 3	UMTS 4	WLAN
Audio Call	Low	-	-	-	-	-	-
	Medium	-	-	-	-	-	-
	High	20	18	-	16	14	-
Video Streaming	Low	13	11	14	7	5	16
	Medium	16	14	17	11	8	20
	High	-	-	16	-	-	18
Web Browsing	Low	15	12	12	10	8	17
	Medium	18	15	15	13	11	20
	High	-	-	14	-	-	19

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Figure 2-10 and Figure 2-11 illustrate the aggregate results, in case user u requests a suite of applications. All potential cases are considered. In particular, seven application demand combinations are possible, in each context. These combinations are: (i) s_1; (ii) s_2; (iii) s_3; (iv) s_1 and s_2; (v) s_1 and s_3; (vi) s_2 and s_3; (vii) s_1, s_2 and s_3. From each application demand combination there is a certain objective function value that is scored.

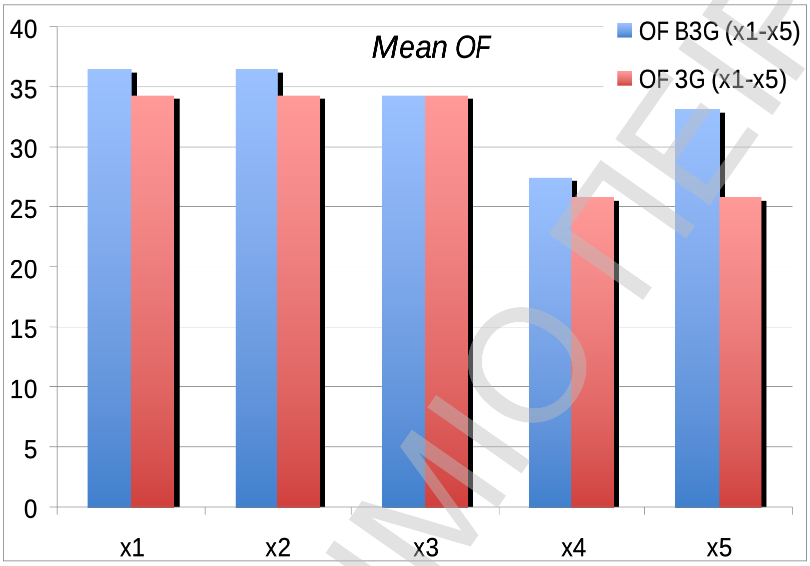


Figure 2-10: Scenario set 1. Mean aggregate objective function values, when the service area is covered by a heterogeneous B3G infrastructure or only 3G technologies, at contexts x1 – x5

Figure 2-10 shows the mean aggregate objective function value, from the seven cases, at contexts x_1, ... , x_5. Moreover, the figure also shows the mean aggregate objective function values, in case the network infrastructure of nw contained only the UMTS (3G) technologies, i.e., nw had an infrastructure that is similar to a legacy one. The improvement from the exploitation of the B3G infrastructure, through the RTMS, compared to the performance obtained from an infrastructure in which only 3G technologies exist, ranges from 8% (at contexts x_1 and x_2) to 21% (at context x_5). The complexity for obtaining these improvements, in terms of computational time, is minimal and similar to legacy situations. The RTMS schemes are based on the exploitation of a small set of configurations of UMTS, WiMAX or WLAN type. A legacy device, in each transceiver, would work with a set of configurations that has similar size. The difference is that the configurations of a legacy transceiver would be of one RAT type.

The improvements are important. They have been obtained in situations that do not always favour the RTMS. Firstly, the assumed coverage from RATs, other than UMTS, is not dense. More dense coverage would have further increased the improvement, since heterogeneous RATs offer more context-handling capabilities to the RTMS. Moreover, the improvement has been obtained in the presence of restrictive policies, i.e., to serve s_1 through the UMTS RAT only. Therefore, when s_1 is in the suite of applications requested, a configuration that associates the device with UMTS is selected. As already presented, the UMTS configurations cannot support the high QoS levels, and this stresses the performance of the RTMS.

In general, the improved behaviour is due to the following factors. First, the exploitation of the user profile and especially the user-interest for high QoS levels. Second, the exploitation of the context and policies and especially the existence of various RATs, which offer diverse QoS capabilities. Third, the existence of the RTMS that can exploit all the features.

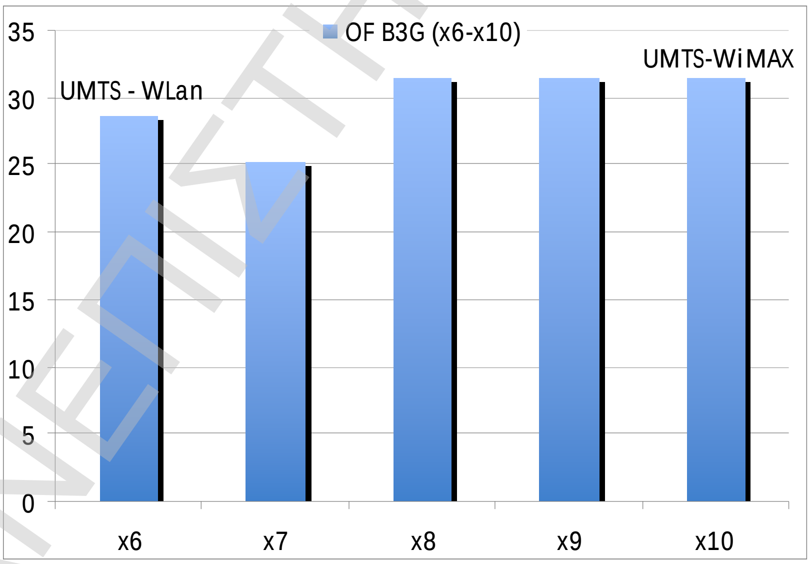


Figure 2-11: Scenario set 1. Mean aggregate objective function values, when the service area is covered by a heterogeneous B3G infrastructure or only 3G technologies, at contexts x6 – x10

Figure 2-11 shows the RTMS behaviour contexts x_6 - x_{10}. Specifically, it shows the mean aggregate objective function value, resulting from the seven application demand combinations, at contexts x_6 - x_{10}. In this case (Figure 2-11), the objective function values are optimal, but lower compared to the first case (Figure 2-10). This is anticipated. At contexts x_6 - x_{10}, when the role is rol_2, “private life”, the user interest for higher QoS levels is lower, compared to when the role is rol_1, “in work”. This yields the lower objective function values. Nevertheless, the RTMS provides optimal handling, in the sense that it offers to the users their desired QoS levels.

An important aspect, related to the second case of the first scenario (Figure 2-11), is that the RTMS prevails in an important aspect that should also be noted. By exploiting the B3G infrastructure, RTMS offers higher QoS levels, compared to the legacy situation, in terms of reliability and dependability levels. This happens because there are various options (configurations and RATs) through which context handling can be done, and the best QoS levels can be obtained.

In general, the RTMS exploits the profiles, the context and the policies, for selecting the best configurations, and therefore, yielding ubiquitous, personalised, context-aware always-best connectivity, in a seamless to the user manner.

2.8.3. Output from second scenario

The same contexts will be taken as input, as in the previous scenario. In this scenario, it is assumed that the terminal device is equipped with two transceivers, $TRX(d) = \{t_1, t_2\}$. Each transceiver is able to operate with different configuration.

Figure 2-12 and Figure 2-13 illustrate aggregate results, in case user u requests a suite of applications, which fall within the same seven application demand combinations, considered also in the first scenario.

Figure 2-12 presents the mean aggregate objective function values at $x_1, \ldots, x_5$, in case the infrastructure of nw has the technologies depicted in Figure 2-9, or only 3G components. In this case, each transceiver is able to operate a different configuration and serve a subset of the requested applications. This could lead to an overall QoS improvement, since applications can be assigned to different transceivers and different RATs. For example, at context x_1, t_1 can operate um_1 and be assigned application s_1, and t_2 can operate wi and be assigned applications s_2 and s_3. Applications s_2

and s_3 can be assigned at a "high" quality level, in contrast with the previous scenario. In general, in this scenario the improvement ranges from 13% (at contexts x_1 and x_2) to 33.3% (at x_5) when the infrastructure has configurations other than 3G.

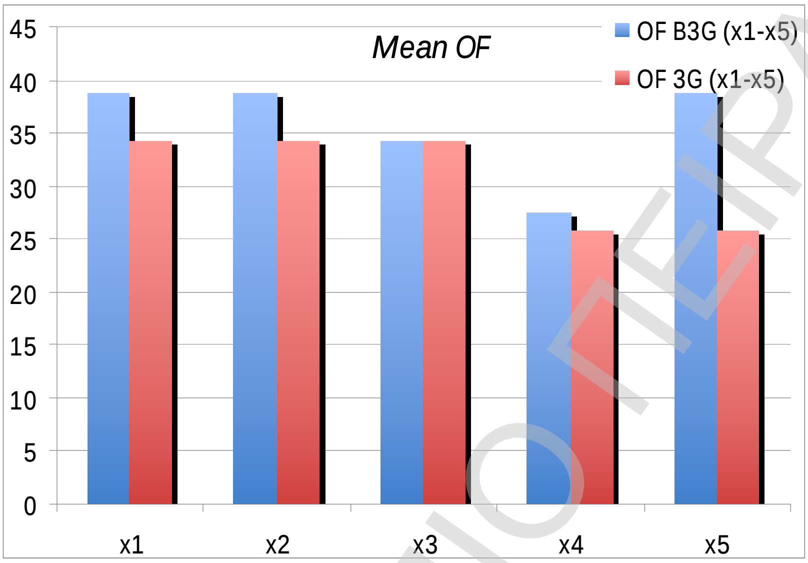


Figure 2-12: Scenario set 2. Mean aggregate objective function values, when the service area is covered by a heterogeneous B3G infrastructure or only 3G technologies, at contexts x1 – x5.

Similar deductions, as in the first scenario, can be made regarding the behaviour of the RTMS. It yields significantly improved QoS, even under certain difficult conditions, like the non-dense coverage of many RATs of the B3G environment, and the presence of restrictive policies. Again, the complexity for obtaining these improvements, in terms of computational time, is minimal and similar to legacy situations. The scheme requires that the transceivers can exploit a small set of configurations of UMTS, WiMAX or WLAN type. The set of configurations has similar size, as the one that a transceiver of a legacy device examines.

Figure 2-13 shows the RTMS behaviour in contexts x_6 - x_{10}. Specifically, it shows the mean aggregate objective function value, resulting from the seven application demand combinations. Again, in this case (Figure 2-13), the objective function values are optimal, but lower compared to the first case (Figure 2-12). This is anticipated. At contexts x_6 - x_{10}, when the role is rol_2, "private life", the user interest for higher QoS levels is lower, compared to when the role is rol_1, "in work". Nevertheless, again, the

RTMS provides optimal handling, in the sense that it offers to the users their desired QoS levels and higher reliability and dependability levels.

Compared to the first scenario the QoS provision is higher due to the two transceivers.

In general, in the second scenario as well, the RTMS exploits the profiles, the context and the policies, for selecting the best configurations, and therefore, yielding ubiquitous, personalised, context-aware always-best connectivity, in a seamless to the user manner.

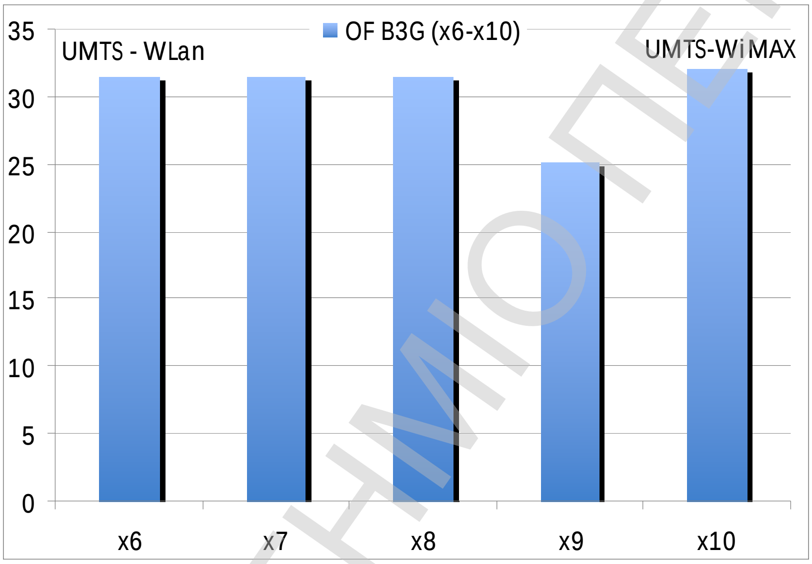


Figure 2-13: Scenario set 2. Mean aggregate objective function values, when the service area is covered by a heterogeneous B3G infrastructure or only 3G technologies, at contexts x1 – x5

2.8.4. Impact of cooperation and negotiation with other NOs

In order to show the impact of the cooperation with other NOs, it is assumed that one of the locations in the service area can be served by a cooperating network, e.g., an operating WLAN technology. The corresponding configuration is denoted as wl_{cp}. Without lack of generality, this WLAN will be assumed available in area loc_1. Therefore, the network affects contexts x_1 and x_{10}.

The WLAN network will be assumed capable of cooperating with nw, through the reversed English and Dutch negotiation schemes [6]. According to the English negotiation model, nw, through the RTMS, continuously asks for a lower cost from the cooperating NO, until an agreement is reached or a maximum number of steps is conducted. On the other hand, according to the Dutch negotiation model, nw starts by suggesting to the cooperating NOs a certain (low) cost for supporting an application at a certain QoS level. In case the cooperating networks accept, the negotiation stops. Otherwise, nw can continue to increase the cost, until a cooperating NO accepts or a maximum number of steps are conducted. In this section, it is assumed that the negotiation takes place in five steps. This is concrete evidence of the low computational time that is required by the scheme, in order to obtain the considerable improvements presented in this subsection.

The results show that negotiation can be used for obtaining a good cost from cooperating network, and enabling the user to access service s_2 at QoS level q_{23}, in context x_{10} (private life). The maximum acceptable cost is $cst_C^{apl}(u, s_2, q_{23}, x_{10}) = 7$. The cooperating network will be offering costs that, according to the formulations, are collectively denoted as $cst_P^{cpn}(u, s_2, q_{23}, wl_{cp}, x_{10})$. The offered costs, by nw in the reversed Dutch model, or the cooperating networks in the reversed English model, can be determined by pricing models [39]. This work is generic with respect to this area. Based on the maximum and the offered costs, nw, can determine the final costs that will be offered to the user, which are denoted as $cst_P^{apl}(u, s_2, q_{23}, x_{10})$. These costs will be between the cost offered by the cooperating network and the maximum tolerable (as introduced by relation (2)).

Figure 2-14 depicts the results from the reversed English negotiation model. Various cooperation cases will be considered, in order to obtain diverse insight on the potentials of the scheme. Three different cost dropping strategies have been tested, i.e., the “slow”, “medium” and “fast” strategy. Each of these strategies represents the rate that the cost is reduced by the RTMS, during the negotiation with the cooperating NO.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

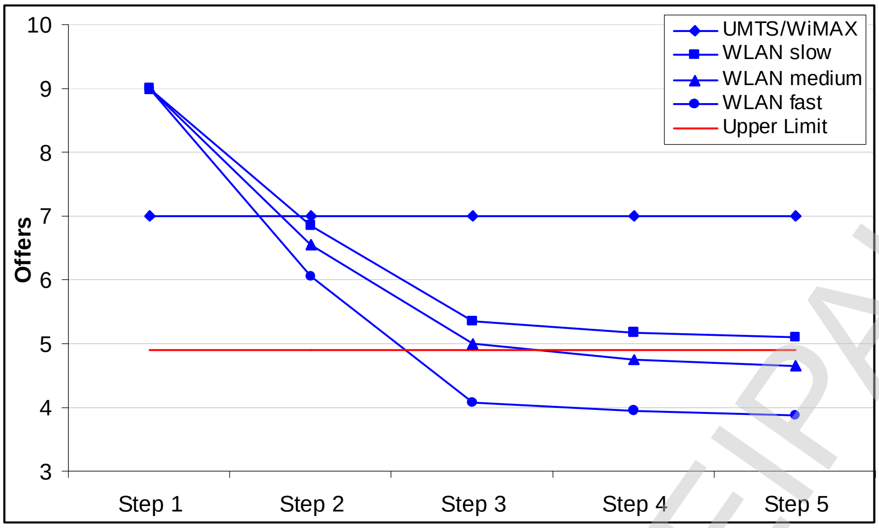


Figure 2-14: Results from the cooperation and negotiation with other NOs. Samples regarding the evolution of the $cst_P^{cpn}(u, s_2, q_{23}, wl_{cp}, x_{10})$ values, offered by a cooperating network, according to the reversed English model

In the reversed English model, the starting point cost offered by the cooperating NO for the use of the WLAN can be higher than the minimum acceptable cost. As can be observed, according to the “slow” strategy, the cost decreases slowly without reaching the minimum acceptable cost, within the range of the five negotiation steps. As a result, the negotiation did not reach to an agreement between the negotiating entities. In the second case, according to the “medium strategy, the cost degreases faster within the range of the five negotiation steps, and drops below the maximum acceptable cost. As a result, a better price is achieved through the use of the negotiation model. Therefore, the cooperating network and the WLAN technology can be used, and user u can be offered the service s_2 at QoS level q_{23}. Finally, during the “fast” strategy, an agreement is also reached but in fewer steps than in the previous case. The disadvantage in this case is that with a fast cost dropping rate, the opportunity for a better (in the case of RTMS) agreement is not exploited.

Figure 2-15 illustrates the results from the reversed Dutch negotiation model. According to this protocol, the RTMS continuously raises the cost, until the cooperative NO accepts it, or the realization of a maximum number of steps. As shown in the figure, the RTMS finds a cost area, delineated by an upper and a lower limit, to which the negotiation follows. The upper limit gives the maximum acceptable cst for the use of service s_2 at QoS level q_{23}. Above that cost, there will be no gain for the change to the WLAN, with the given role of the user. There is also

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

a lower limit, calculated also by the RTMS, which gives the starting point of the negotiation with the cooperating NO. This limit is taken around 40% lower than the upper limit. The negotiation using the Dutch model is concluded in five steps at maximum, but it may be concluded earlier if the cooperative network accepts the proposed price before reaching the fifth step. Figure 2-15 shows two different cost increase strategies ("slow" and "fast") that can occur during the negotiation. The first one ("fast") increases the cost faster and an agreement may be reached in fewer steps, while the second ("slow") may need more steps in order for an agreement to be reached, but it may also result in a smaller cost.

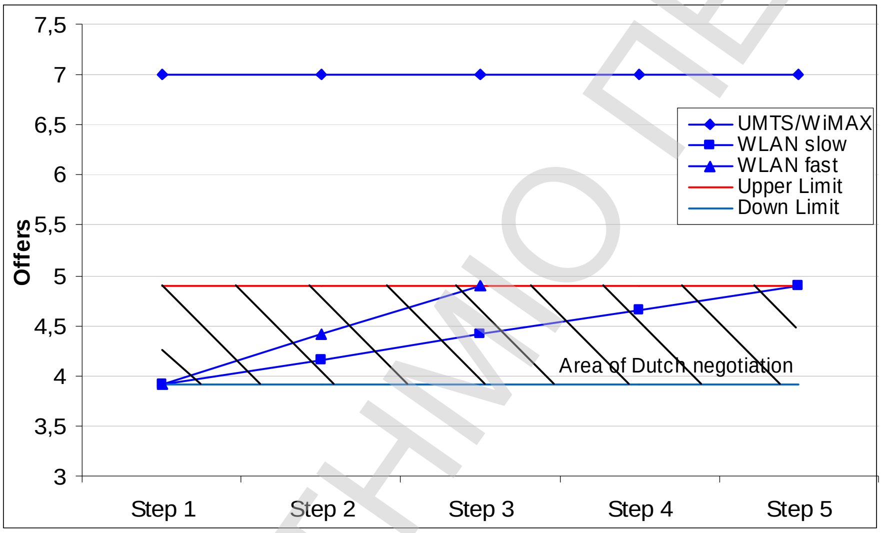


Figure 2-15: Results from the cooperation and negotiation with other NOs. Samples regarding the evolution of the $cst_{P}^{cpn}(u, s_2, q_{23}, wl_{cp}, x_{10})$ values, offered by a cooperating network, according to the reversed Dutch model.

Both negotiation cases show important potentials for establishing agreements with a cooperating NO. The agreements enable the user to access a service at a QoS level, which otherwise would not be possible. Extensions of this work will pursue further benefits for the user, through the experimentation with pricing schemes, or through the specification of regulation activities, in order to ensure that the cooperation between NOs will result to direct benefits (e.g., immediate discounts) for the users.

2.9. Cooperation And Negotiation Strategies

2.9.1. General approach

Figure 2-5 depicts the general approach for reaching the negotiation objectives. The general approach can be customized to correspond to the specific strategies that have been integrated in the CNS component of the RTMS.

The input, in the first step, conveys information on the set of applications, $A \subseteq APL_C(u,x)$, which can be served by configurations in the $CPN_P(nw,x)$ set. As introduced, the information can comprise the set of allowed QoS levels per application, and the maximum cost per application and QoS, in the particular context. The second step involves the request of offers from the networks in the $CPN_P(nw,x)$ set. Each network $i \in CPN_P(nw,x)$ obtains the information related to the applications, $A_i \subseteq A$, which can be served by configurations belonging to i. The third step includes the processing of the offers that will be received. Aspects checked can be the non-violation of acceptable limits regarding costs, or the improvement with respect to previous negotiation rounds. At the end of this step, for each network $i \in CPN_P(nw,x)$ there are cost values, $cst_P^{cpn}(u,s,q,c,x)$, for some applications $s \in A_i$ and respective QoS levels $q \in QOS_P(u,s,c,x)$, and for configurations $c \in CFG_P(d,t,x)$ that belong to network i. The fourth step examines whether a termination criterion is satisfied. If yes, the algorithm ends, otherwise, a transition to the second step is performed. In the fourth step, the termination criterion can be either the realization of a maximum of steps, or the lack of possibility to further improve the offers. In the fifth step, the $cst_P^{cpn}(u,s,q,c,x)$ values provide the best values obtained, per application, QoS and configuration.

2.9.2. Specific schemes

The general approach is flexible. It can be configured to finish in one step, e.g., in a manner similar to the first price sealed bid strategy, or to conduct multiple steps, e.g., similar to the reversed English or Dutch auction models [6].

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

In the former case all networks make their offers (bids), per application and QoS level, simultaneously. Subsequently, these offers are used in the optimization problem.

Following the English model nw may set a maximum price as a starting point for the negotiation. In each step, nw asks the cooperating networks to lower their offered costs, per application and QoS. In each step, a reduction can be either accepted or rejected by the cooperating NOs. The negotiation stops with those NOs that refuse to decrease. The termination happens after the realization of a maximum number of steps, or the refusal of further reductions, by all cooperating networks, for all application and QoS combinations (before the expiration of the maximum number of steps).

Following the Dutch model, in each step, nw makes offers to the NOs in $CPN_P(nw,x)$. The starting point is a "low" price, per application and QoS. In each step, the offered price is either accepted or rejected by the cooperating NOs. The negotiation stops with those NOs that accept the price. On the contrary, a higher price is offered, in each step, to the NOs that have not yet accepted the offer. Termination can occur in three cases. First, the realisation of a maximum number of steps. Second, the acceptance of offers, for all application and QoS combinations, by all cooperating networks (before the expiration of the maximum number of steps). Third, when the offers reach the $cst_C(u,s,q,x)$ levels.

2.10. Selection Strategy

Figure 2-6 depicts an approach for selecting the solution to the overall problem. The scheme is structured in four main steps.

The first step adresses the configuration of transceivers. The second phase allocates applications to transceivers, assuming a basic (nevertheless acceptable) QoS level for all applications. Step three is targeted to the potential improvement of QoS levels. Finally, step four is targeted at the selection of the best configuration.

In step one, a number of sub-problems is launched. In each sub-problem, the transceiver configuration is fixed. Each sub-problem, i, will consider and investigate

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

the performance of the specific transceiver configuration pattern, $(A_{TC})_i$. The sub-problems can be addressed in parallel. Each transceiver configuration pattern has certain QoS capabilities that can be offered to the applications. These QoS capabilities are exploited in the next step.

In the second step, in each sub-problem i, there will be the production of an allocation of applications/services to transceivers, $(A_{ST})_i^{init}$, and to QoS levels, $(A_{SQ})_i^{basic}$. In this step, each application $s \in APL_C(u,x)$ is assigned at the basic (lowest) acceptable QoS level of the $QOS_P(u,s,c,x)$ set, for which there is $y_{st} = 1$ and $x_{tc} = 1$, in $(A_{ST})_i^{init}$ and $(A_{TC})_i$. Specifically, after step one, each application should have a set of transceivers from which it can be served, taking into account profile, context and policy aspects. In case applications cannot be served by any transceiver, the configuration pattern is rejected, and the sub-problem is stopped. Otherwise, the algorithm evolves in a greedy manner, on an application-by-application basis. The allocation of applications that can be served by one transceiver is straightforward and is done first. For the applications that can be served by more transceivers a policy-based decision is required. For instance, in [32] an application is allocated to the transceiver that has the largest remaining QoS capabilities, and the smallest future (potential) demand (to serve applications that are unassigned).

The third step provides, for each sub-problem i, the final allocation of applications/services to transceivers, $(A_{ST})_i$, and to QoS levels, $(A_{SQ})_i$. This step improves the QoS levels that are offered to the applications. Depending on the $QOS_P(u,s,c,x)$ set, there may be the possibility for QoS improvement, perhaps by changing also the allocation to transceivers. In case higher QoS levels are associated with higher utility values, the improvement of the QoS can have a positive effect on the overall objective function value, according to (1). Therefore, a set of candidate moves (QoS improvements) is obtained. Each move consists of a QoS improvement and a potential reassignment of the application to the transceiver. Moves can be ordered with respect to their potential positive contribution on the objective function value. The algorithm can evolve in a greedy manner. In each sub-step, the algorithm can select the move that leads to the highest improvement (i.e., the one that improves most the objective function), which is also feasible from the transceiver QoS

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

capability point of view. The algorithm stops when no more QoS improvement can be made, or when the transceiver QoS capabilities are violated.

Finally, the last step is targeted to the selection of the best solution. Specifically, the triplet ($(A_{TC})_i, (A_{ST})_i, (A_{SQ})_i$), obtained in sub-problem i , which scores the highest objective function value, according to (1), is maintained as the overall solution.

2.11. Conclusions

This chapter presented a management system, called RTMS, for enabling NOs, in the era of wireless B3G communications, to drive users to the most appropriate RATs that satisfy the demanding requirements for personalisation, context awareness, always best connectivity, ubiquitous service provision and seamless mobility. The RTMS provides the means for profile management, the acquisition of monitoring/discovery/context information, and the negotiation and selection of configurations, based on information deriving from policies, profiles and context. A business case was presented, for showing the role of the management system. The work presented focused on the role and the information of the RTMS components. Concrete functionality for accomplishing the role was also presented.

Results showed the efficiency of the management schemes. The improvement from the exploitation of the B3G infrastructure, through the RTMS, compared to the performance obtained from an infrastructure in which only 3G technologies exist is significant. In general, the improved behaviour is due to the following factors. First, the exploitation of the user profile and especially the user-interest for high QoS levels. Second, the exploitation of the context and policies and especially the existence of various RATs, which offer diverse QoS capabilities. Third, the existence of the RTMS that can exploit all the features.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

References

[1] P.Demestichas, A.Katidiotis, D.Petromanolakis, V.Stavroulaki, "Management System for Terminals in the Wireless B3G World", accepted for publication to the Wireless Personal Communications journal.

[2] V. Stavroulaki, G. Dimitrakopoulos, A. Katidiotis, P. Demestichas, D. Bourse, K. El Khazen, "Negotiation of Network Services and Spectrum in B3G, Composite Radio, Environments", Innovation and the Knowledge Economy: Issues, Applications and Case Studies, pp.1103-1110, IOS Press, 2005.

[3] European Commission, 6th Framework Programme (FP6), Information Society Technologies (IST), Project End-to-End Reconfigurability (E2R), http://e2r2.motlabs.com, 2007

[4] P.Demestichas, G.Vivier, K.El-Khazen, M.Theologou, "Evolution in wireless systems management concepts: from composite radio to reconfigurability", *IEEE Commun. Mag.*, Vol. 42, No. 5, pp. 90-98, May 2004

[5] V. Stavroulaki, S. Buljore, P. Roux, E. Melin, "Equipment management issues in B3G end-to-end reconfigurable systems", *IEEE Wireless Communications Mag*, Vol. 13, No. 3, pp. 24-32, June 2006

[6] N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, C. Sierra and M. Wooldridge, 2001 "Automated negotiation: prospects, methods and challenges", *Int. J. of Group Decision and Negotiation*, Vol. 10, No. 2, pp. 199-215 , 2001

[7] M. J. Van Sinderen, A. T. Van Halteren, M. Wegdam, H. B. Meeuwissen, E. Henk Eertink, "Supporting context-aware mobile applications", *IEEE Commun. Mag.*, Vol. 44, No. 9, Sept. 2006

[8] P. Bellavista, A. Corradi, R. Montanari, A. Tononelli, "Context–aware semantic discovery for next generation mobile systems", *IEEE Commun. Mag.*, Vol. 44, No. 9, Sept. 2006

[9] K. Tsagkaris, A. Katidiotis, P. Demestichas, Neural network-based learning schemes for cognitive radio systems, Computer Communications Volume 31, Issue 14, , 5 September 2008, Pages 3394-3404.

[10] X. Liu, N. S. Shankar, "Sensing-based opportunistic channel access", *Mobile Networks and Applications* Journal, Vol. 11, No. 4, pp. 577-591, Aug. 2006

[11] H. Kim, K. G. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks", *IEEE Trans. on Mobile Computing*, Vol. 7, No. 5, pp. 533-545 May 2008

[12] J. Perez-Romero, O. Sallent, R. Agusti, L. Giupponi, "A novel on-demand cognitive pilot channel enabling dynamic spectrum allocation", In Proc. 2nd International Symposium on *New Frontiers in Dynamic Spectrum Access Networks 2007* (*DySPAN 2007*), Dublin, Ireland, April 2007

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

[13] Open Mobile Alliance (OMA), http://www.openmobilealliance.org/, 2008

[14] J. Strassner, J. Btrabsner, "*Policy-based network management: solution for the next generation*", Elsevier Science and Technology Books, Sept. 2003

[15] Q. Song, A. Jamalipour, "Network selection in integrated wireless LAN and UMTS environment using mathematical modelling and computing techniques", *IEEE Wireless Commun. Mag.*, Vol. 12, No. 3, June 2005

[16] F. Bari, V. Leung, "Automated network selection in a heterogeneous wireless network environment", *IEEE Network*, Vol. 21, No. 1, Jan/Feb 2007

[17] Q.T. Nguyen-Vuong, N. Agoulmine, Y. Ghamri-Doudane, "Terminal controlled mobility management in heterogeneous wireless networks", *IEEE Commun. Mag.*, Vol. 45, No. 4, April 2007

[18] J. Von Neumann and O. Morgenstern, "*Theory of Games and Economic Behavior*", Wiley, 1944

[19] J. Kephart, D. Chess, "The vision of autonomic computing", *IEEE Computer*, Vol. 36, No.1, pp. 41-50, January 2003

[20] P.Demestichas, D.Boscovic, V.Stavroulaki, A.Lee, J.Strassner, "m@ANGEL: autonomic management platform for seamless wireless cognitive connectivity to the mobile Internet", *IEEE Commun. Mag.*, Vol. 44, No.6, pp. 118-127,June 2006

[21] K. Nolan, L. Doyle, "Teamwork and collaboration in cognitive wireless networks", *IEEE Wireless Communications Magazine*, Vol. 14, No. 4, pp. 22-27, August 2007

[22] Institute of Electrical and Electronics Engineers (IEEE), 802 standards, www.ieee802.org, 2008

[23] Institute of Electrical and Electronics Engineers (IEEE), Standards Coordinating Committee 41 (SCC41), Dynamic Spectrum Access Networks, http://www.scc41.org, 2008

[24] T. Doran, "IEEE 1220: for practical systems engineering", *IEEE Computer*, Vol. 39, No. 5, May 2006

[25] W. Hasselbring, R.Reussner, "Towards trustworthy software systems", *IEEE Computer*, Vol. 29, No. 4 April 2006

[26] Software Defined Radio Forum (SDRF), www.sdrforum.org, 2008

[27] M.R. Carrey, D.S. Johnson, "*Computers and Intractability: A Guide to the Theory of NP-Completeness*", W.H. Freeman, San Fransisco, 1979

[28] C. Papadimitriou, K. Steiglitz, "*Combinatorial Optimization: Algorithms and Complexity*", Prentice Hall Inc., 1982

[29] E. Aarts, J. Korts, "*Simulated annealing and the Boltzmann machines*", J. Wiley & Sons, New York, 1989

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

[30] Z. Michalewicz, "*Genetic algorithms + Data structures = Evolution programs*", Springer-Verlag, Berlin, 1995

[31] F. Glover, E. Taillard, D. de Werra, "A User's Guide to Taboo Search", *Annals of Operations Research*, Vol. 41, pp. 3-28, 1993

[32] K. Tsagkaris, G. Dimitrakopoulos, A. Saatsakis, P. Demestichas, "Distributed radio access technology selection for adaptive networks in high-speed B3G infrastructures", *International Journal of Communication Systems*, Vol. 20, No. 8, pp. 969-992, Aug. 2007

[33] J. Mitola, G. Q. Maguire Jr, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999

[34] S. Haykin, "Cognitive radio: brain-empowered wireless communications", *IEEE Journal on Selected Areas In Communications,* Vol. 23, No. 2, pp. 201-220, Feb. 2005

[35] P.Demestichas, G.Dimitrakopoloulos, J.Strassner, D. Bourse, "Introducing reconfigurability and cognitive network concepts in the wireless world", *IEEE Vehicular Technology Mag.*, Vol. 1, No. 1, pp. 33-39, June 2006

[36] R. Venkatesha Prasad, P. Pawelczak, J. Hoffmeyer, S. Berger, "Cognitive functionality in next generation wireless networks: standardization efforts", *IEEE Commun. Mag.*, Vol. 46, No. 4,pp. 72–78 April 2008

[37] Third (3rd) Generation Partnership Project (3GPP), Web site, www.3gpp.org , 2006

[38] WiMAX Forum, http://www.wimaxforum.org, 2008

[39] V. Pandey, D. Ghosal, B. Mukherjee, "Pricing-based approaches in the design of next-generation wireless networks: a review and a unified proposal", *IEEE Commun. Surveys and Tutorials*, Vol. 9, No. 2, 2007.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

3. ENHANCING CHANNEL ESTIMATION IN COGNITIVE RADIO SYSTEMS BY MEANS OF BAYESIAN NETWORKS

Abstract:

This chapter proposes enhancements to the channel(-state) estimation phase of a cognitive radio system. Cognitive radio devices have the ability to dynamically select their operating configurations, based on environment aspects, goals, profiles, preferences, etc. The proposed method aims at evaluating the various candidate configurations that a cognitive transmitter may operate in, by associating a capability e.g., achievable bit-rate, with each of these configurations. It takes into account calculations of channel capacity provided by channel-state estimation information (CSI) and the sensed environment, and at the same time increases the certainty about the configuration evaluations by considering past experience and knowledge through the use of Bayesian networks. Results from comprehensive scenarios show the impact of the method proposed in this chapter, on the behaviour of cognitive radio systems. Parts of this chapter have been published in [1].

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ENHANCING CHANNEL ESTIMATION IN COGNITIVE RADIO SYSTEMS BY MEANS OF BAYESIAN NETWORKS

3.1. Introduction

As already stated in the introductory chapter, an increasingly important engineering challenge, in today's world, is the proper management of the electromagnetic radio spectrum. Thus, there is need for the development of efficient spectrum management schemes, capable of exploiting the available, underutilised frequency bands.

A direction for spectrum efficiency is to equip the infrastructure with ***cognitive radio*** capabilities [2],[3],[4],[5]. In general, cognitive radio devices dynamically select their configurations, through management functionality [6] that takes into account the context of operation (device status and environment aspects), goals and policies [7], profiles, and machine learning [8] (for representing and managing knowledge and experience). In the more general sense, the term ***configuration*** refers to a spectrum carrier and a specific RAT, but the list could also be expanded to include modulation type, transmission power etc. This definition also allows a spectrum band to be used for operating different RATs, in accordance with the flexible spectrum management concept [5].

In any manifestation, proper mechanisms for channel-state estimation are imperative for adaptive, cognitive radio systems operating in dynamically changing environments. Channel-state estimation is needed for calculating the channel capacity which, in turn, is required in order to assist the transmitter for evaluating its candidate operating configurations. More specifically, the cognitive receiver exploits the CSI in order to feed a well known theoretical formula (e.g. Shannon theorem) for the calculation of the achievable bit rate.

This is exactly where the focus of the work in this chapter is placed on. More specifically the objective of this work is to use the calculated bit rate in order to associate each of the candidate transmitter's configurations with an anticipated capability (e.g. in terms of achievable bit rate). As long as there is a clear picture about the capabilities of each configuration, the transmitter will be able, in a sequent step, to select the optimum one to use. Such decision strategies are analysed in section 2.10.

In order to increase the certainty about the configuration evaluations, a learning solution is proposed. The solution integrates knowledge and experience and relies on Bayesian Networks, which are a category of advanced machine learning schemes, suitable for reasoning probabilistic relationships [8][9][10]. Such integration in the channel-state estimation phase can be especially important for improving the robustness of the evaluation of the configuration capabilities.

The rest of the chapter is organized as follows: Related work and motivation are presented in Section 3.2. The proposed solution is presented in sections 3.3 and 3.4. Section 3.5 provides results from comprehensive scenarios that reveal the behaviour of the proposed scheme. Finally, concluding remarks are drawn in section 3.6.

3.2. Related work and Motivation

At first, this work complements the used channel-state estimation mechanisms. In general, channel estimation can be either training-based [11][12] or blind [13][14][15][16], with both cases exhibiting pros and cons in terms of bandwidth efficiency, convergence speed and estimation accuracy. When it comes to cognitive radio, the majority of channel state estimation techniques proposed in the literature, regardless of being training-based or blind, assume Orthogonal Frequency Division Multiplexing (OFDM)-based systems [17][18][19]. This can be easily justified by the fact that OFDM's inborn features, such as spectral efficiency and flexibility, render it a modulation strategy that commends itself to cognitive radio [3], albeit other proposals for the modulation scheme of cognitive radio have come since the introduction of the idea [17].

On the other hand, a cognitive radio will inherently have the ability to improve its performance through learning. Learning systems require collection of data from the environment sensed, in order to draw conclusions about the observed variables.

Machine learning techniques such back-propagation Neural Networks, Self-Organising Maps, Fuzzy Systems, Evolutionary Algorithms, Case-based Systems and of course Bayesian Networks enable such behaviour and can be used to optimize the performance by adapting the radio parameters with respect to the input variables.

Especially a Bayesian Network, [9][10], is a graphical model that encodes probabilistic relationships-dependencies among a set of variables of interest. Some of the benefits that Bayesian Networks offer when used for handling input data [20], are the abilities to handle incomplete data sets, to allow learning of causal relationships (e.g. causes and symptoms), to use prior knowledge and also, to avoid data overfitting [21] (i.e. when the network adheres to a training data set, thus being unable to perform correctly on unseen data).

Apart from medicine, bioinformatics or economics, Bayesian networks have also been used in the engineering literature e.g. for fault detection, self-management or automated diagnosis, destined for the application to wireless, cellular networks [22][23][24], and also for modelling user preferences and profiles in B3G/4G devices [25][26]. This work applies Bayesian networks for improving the performance of a cognitive radio through learning. In particular, in this chapter, a Bayesian Network is formulated in order to model the probabilistic relationship among the achievable bit rate and corresponding configuration of a cognitive transmitter.

3.3. Formulation as a Bayesian Network

Figure 3-1 depicts the approach for formulating the problem as a Bayesian network. As stated, the objective is to associate each candidate configuration with a specified capability. In the Bayesian network, random variable CFG represents the configuration that is probed, and random variable BR represents a configuration's capability, e.g., the achievable bit-rate as calculated using CSI and Shannon's theorem. CFG is the Bayesian network's predictive attribute (parent node), while BR is the target attribute (node), which can take a set of values from a reference set as will be discussed in the following. In a similar manner, although a simplified approach with one capability is assumed herewith, more capabilities, e.g., the bit error rate, etc., can be considered readily. The method relies on the constant update and maintenance of conditional probability values, of the form $Pr[BR|CFG]$, which reveal the probability that a capability (in this case the bit-rate) will be at a certain level, given that a certain configuration is used.

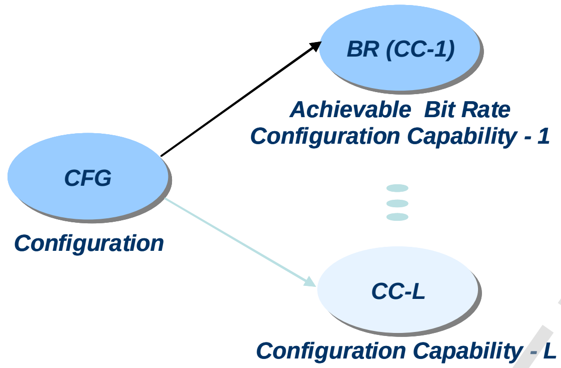


Figure 3-1: Structure of the Bayesian network.

Conditional probability tables (CPTs) can, therefore, be organized. Every node in a Bayesian Network has an associated CPT to express the probability of its state in condition to its parent states. Figure 3-2 depicts the structure of the CPTs in this case, with particular focus on the bit rate. Each column of a CPT refers to a specific configuration. If there are n possible configurations, the CPT will include n columns. Each line of the CPT corresponds to a reference, achievable bit-rate value.

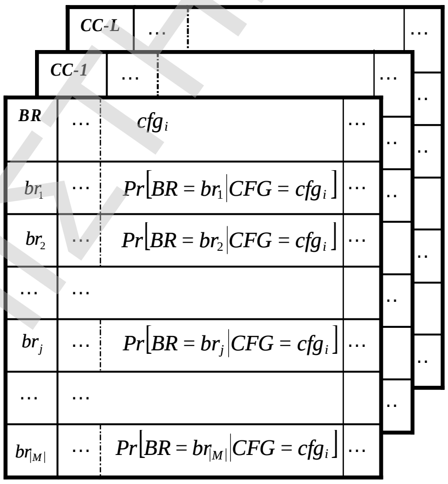


Figure 3-2: Structure of the Conditional Probability Tables (CPTs).

Those reference bit-rates comprise the set, from which random variable BR may take values. This set is selected here to be discrete [24]. Let M be that discrete set of reference, achievable bit-rate values. Without loss of generality, enumeration can be done in ascending order (i.e., $br_1 < br_2 < \ldots < br_{|M|}$). The cell at the intersection of line j $(1 \le j \le |M|)$ and column i $(1 \le i \le n)$ provides the value of the conditional probability

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

$Pr\lfloor BR = br_j | CFG = cfg_i \rfloor$, which expresses the probability that bit-rate br_j will be achieved, given that configuration cfg_i is selected.

Given a configuration, the most probable achievable bit-rate is the one that is associated with the maximum conditional probability in the respective column. In order to take into account different contexts (e.g., times in the day) there can be several CPTs. Moreover, the CPT can also be maintained as a list, sorted in descending order of the probabilities. Figure 3-3 provides an example. Configuration and bit rate pairs with high probabilities can be in the top of the list, in order to facilitate configuration selections. In the example, bit rate br_k is the most probable for configuration cfg_i.

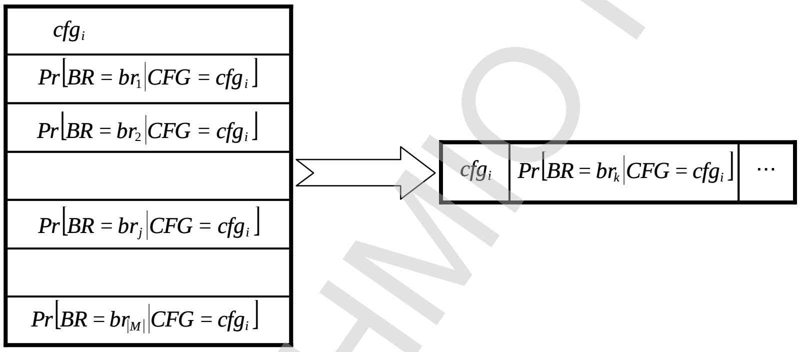


Figure 3-3: Organization of CPT columns as ordered lists, for enabling fast adaptations.

3.4. Learning Strategy

3.4.1. Principles

The capabilities of configurations are provided by the CPT. In general, there are many algorithms that are able to update the values within a CPT [1], i.e. the learning strategy of the Bayesian Network, but for brevity reasons are not addressed in this chapter.

The main formulas used for updating the CPT are:

the correction factor, cor_j,

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

$$cor_j = 1 - \frac{\left|br_j - br_{calc}\right|}{dif_{\max}} \qquad (1)$$

which is used for correcting the CPT values as follows, for each candidate value br_j:

$$Pr\left\lfloor BR = br_j \middle| CFG = cfg_i \right\rfloor_{new} = L \cdot cor_j \cdot Pr\left\lfloor BR = br_j \middle| CFG = cfg_i \right\rfloor_{old} \qquad (2)$$

the parameter L, which is a normalizing factor that guarantees that all "new" probabilities sum up to one and can be computed through the following relations:

$$L \cdot \sum_{j \in M} cor_j \cdot Pr\left\lfloor BR = br_j \middle| CFG = cfg_i \right\rfloor_{old} = 1 \qquad (3a)$$

$$L \cdot \sum_{j \in (M \quad T)} cor_j \cdot Pr\left\lfloor BR = br_j \middle| CFG = cfg_i \right\rfloor_{old} = 1\text{-}(|T| \cdot a / |M|) \qquad (3b)$$

These formulas are explained in [1].

The main purpose of this chapter is to provide a benchmarking work on the Bayesian Networks and their capability to associate each of the candidate transmitter's configurations with an achievable bit rate.

In order to address the continuously changing environment (received data), an online learning strategy is required [27][28]. The learning strategy takes into account the bit rate calculations, which are conducted using the CSI provided by the channel estimation phase, and more specifically, the "distance" (absolute difference) between those calculated values and each reference value. Assuming that, according to the calculations, a specific configuration can achieve bit-rate br_{calc}, this value can be exploited, in order to fine-tune (enhance or decrease) the values of the CPT, and therefore, increase the confidence of the capability estimations.

In general, the confidence on the capability estimations is reached when *convergence* exists. It can be defined that the proposed learning scheme *converges* when the conditional probability of the reference value, which is closest to the measured value, becomes the highest. At this point, the conditional probabilities that correspond to the other (candidate) reference values are either being reduced or reinforced less.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Convergence can also be defined differently, e.g., it can also be associated with the difference between the conditional probability of the value indicated by the calculations and all the rest. After convergence to a certain condition, there can be a set of measures that may be taken for enabling fast adaptations to future conditions. First, the number of consecutive updates, upd, which can be applied on the conditional probabilities, may not be allowed to exceed a certain maximum threshold, upd_{max}. Second, the conditional probability of a reference bit-rate value may not be allowed to fall under a certain threshold, $a/|M|$, where $0 \le a \le 1$ (recall that $|M|$ is the number of reference bit-rates). Third, the number of conditional probabilities, which fall under the minimum threshold, $a/|M|$, may not be allowed to exceed a certain maximum threshold, thr_{max}.

3.4.2. Algorithm

The following sequence of actions takes place during the channel-estimation phase of the cognitive radio process (Figure 3-4).

Step 1. Acquisition of CSI knowledge for calculating instant achievable bit rate, and inspection of whether the learning method is at a convergence stage.

The value, br_{calc}, is considered. The value derives from the calculations made, for configuration cfg_i, exploiting CSI form the previous step of the channel estimation phase of the cognitive radio process. Convergence is identified if the following two conditions hold: (i) the br_{calc} value is the same with that of the previous invocation, (ii) the probability $\Pr\left[BR = br_{calc} \middle| CFG = cfg_i\right]$ is larger than all the rest. In case there is no convergence the variable upd is set to zero, and a transition to *step 3* occurs.

Step 2. Inspection of whether further updates of the CPT are allowed, in case the channel estimation phase is at convergence stage.

Inspection of whether the number of consecutive updates that can be applied after convergence, upd, has reached the maximum threshold, upd_{max}. If the answer is positive, there is migration to *step 6*. Inspection of whether the number of conditional probabilities, that have fallen below the minimum threshold, $|T|$, has reached the maximum threshold, thr_{max}. If the answer is positive, there is migration to ***step 6***.

Step 3. Computation of the new probability values.

Computation of: (i) the correction factor, cor_j, through the set of relations (1); (ii) normalization factor, L, through relation (3a); (iii) new probability values through the set of relations (2).

Step 4. Inspection of whether the CPT should be updated.

Computation of the set, T ($T \subseteq M$), which comprises the probabilities that have fallen under the minimum allowed threshold $a/|M|$. If the number of probabilities in the T set exceeds the maximum allowed number, thr_{max}, i.e., if $|T| > thr_{max}$, there is migration to ***step 6***.

Step 5. Update of CPT.

If $|T|>0$ the following set actions are conducted: (i) The probabilities of the T set are assigned equal to the minimum threshold $a/|M|$; (ii) The new normalization factor is computed through relation (3b); (iii) the new values of the probabilities out of the T set are re-computed through the set of relations (2).

The new probability values (computed in step 3 or above) are stored in the CPT.

In case of convergence, the counter upd (consecutive updates after convergence) is increased.

Step 6. End.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

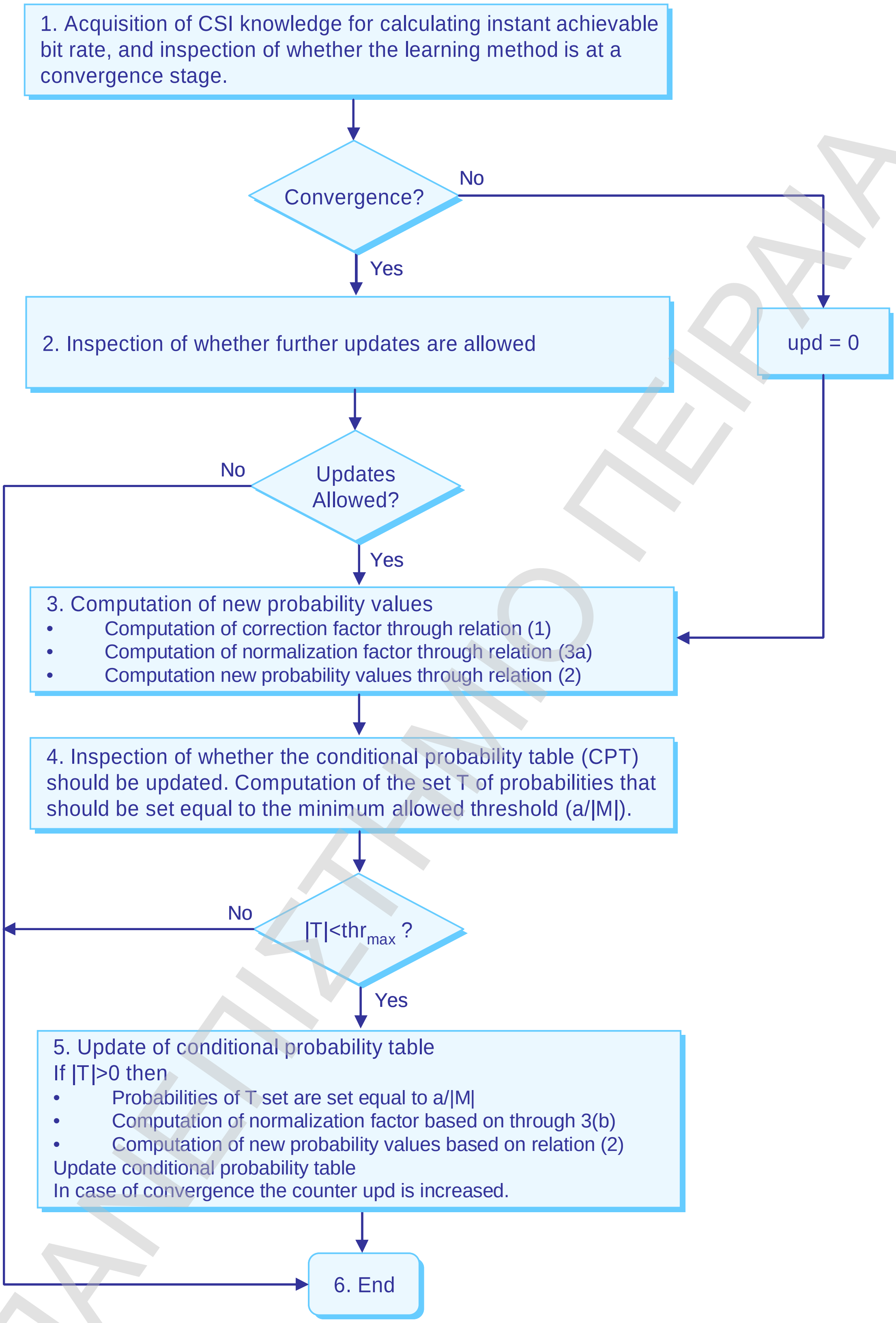


Figure 3-4: Behaviour of the channel estimation phase: strategy for learning the configuration.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

3.5. Results

3.5.1. Set-up

Various sets of scenarios are used for investigating the behaviour of the proposed method. More specifically, the focus is on how this learning method, enabled by Bayesian networks, influences and enhances the channel estimation phase of a cognitive radio process.

The scenarios concern an arbitrary configuration, denoted as c. It is assumed that there are $|M| = 6$ reference bit rate values (in Mbps). M includes the values $br_1 = 6$, $br_2 = 12$, $br_3 = 24$, $br_4 = 36$, $br_5 = 48$, $br_6 = 54$. Hence, $dif_{\max} = 48$ Mbps.

As can be seen, the capabilities of the configuration c have been chosen to be equivalent to those of legacy or emerging standards for wireless local and metropolitan area networks. In addition, parameter a, used in 3(b), has been set equal to 0.1.

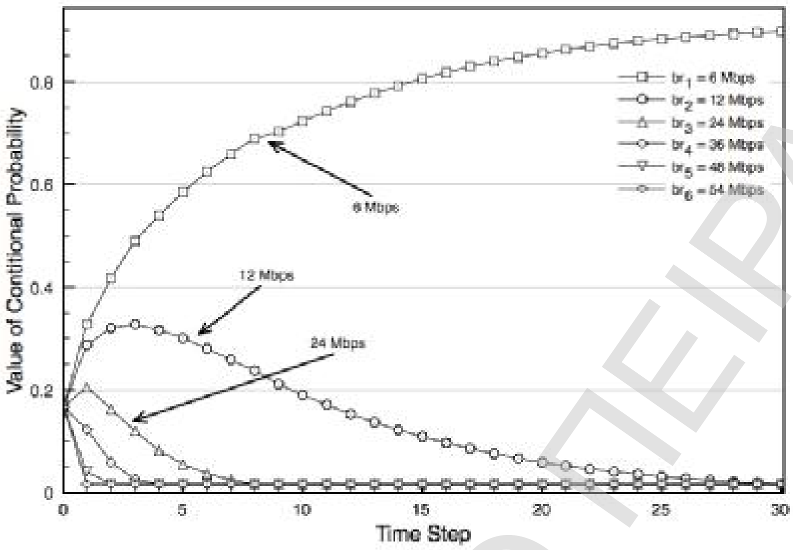
In order to ensure comprehensive testing, two categories of scenarios will be considered. In the first category the assumption is that there is no prior knowledge on the capabilities of the configuration. In the second category of scenarios (will comprise four sets), it will be assumed that the proposed method has some knowledge regarding the capabilities of the configuration. Comprehensiveness is ensured by considering, in both categories, the impact of all the potential changes from the initial conditions.

3.5.2. Presentation

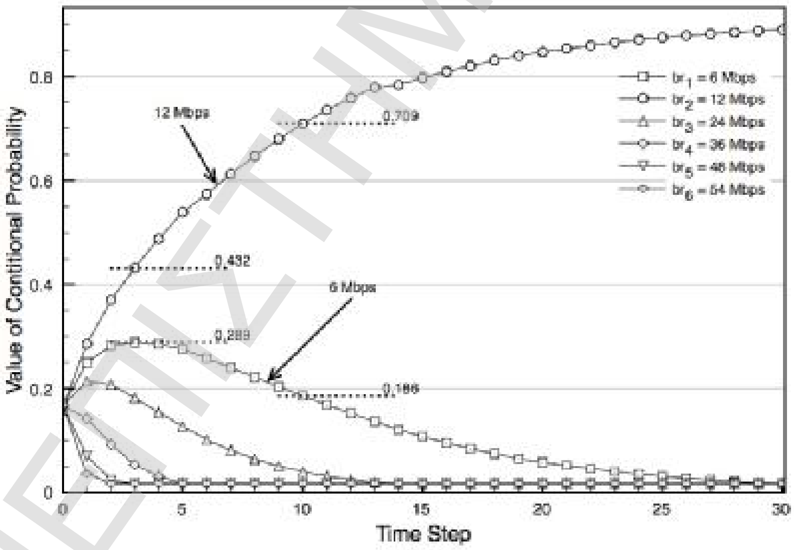
Figure 3-5 depicts the results from the first category of scenarios, in which it is assumed that there is no prior information for configuration c. The x-axis denotes the discrete time steps during which the channel estimation conducts and provides calculations for feeding the method. The y-axis shows the values of conditional probabilities of the form, $\Pr\left[BR = b \middle| CFG = c\right]$, where b can be equal to 6, 12, 24, 36, 48, 54 Mbps. Figure 3-5(a)-(f) shows the evolution of the probabilities when the

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

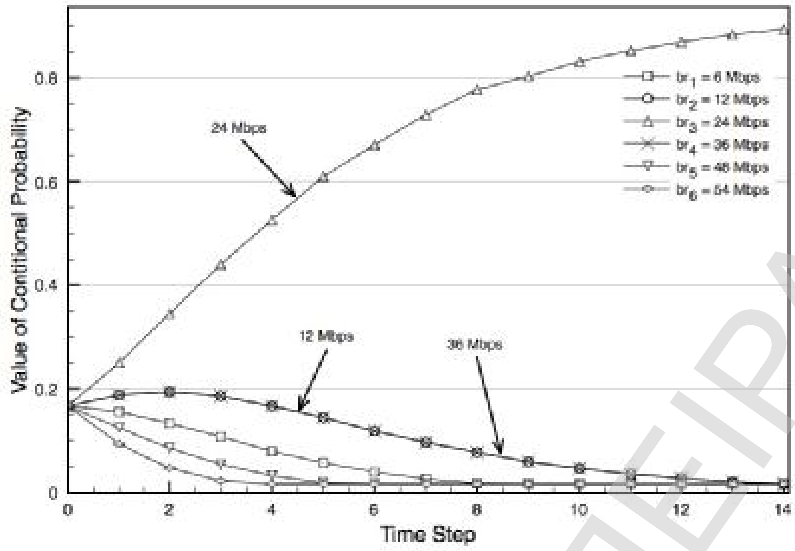
bit rate calculations indicate that the configuration can achieve 6, 12, 24, 36, 48, 54 Mbps, respectively.



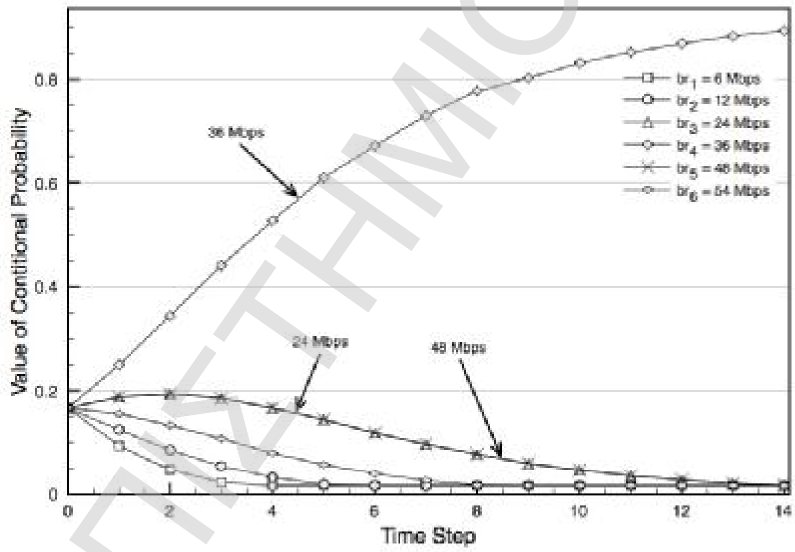
(a)



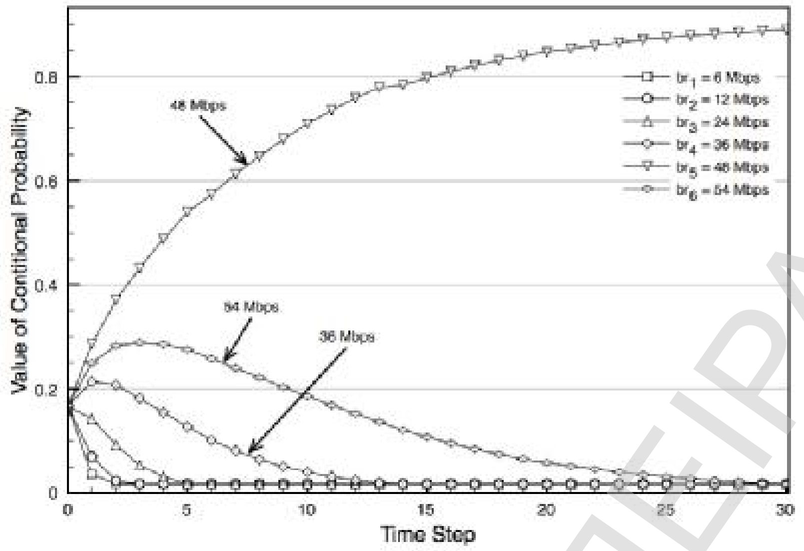
(b)


24 Mbps
12 Mbps
36 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps
Value of Contitional Probability
Time Step

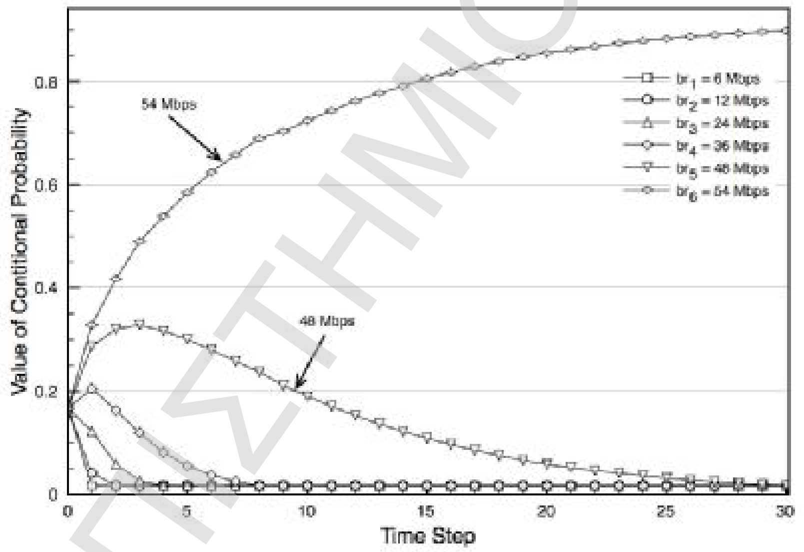
(c)


36 Mbps
24 Mbps
48 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps
Value of Contitional Probability
Time Step

(d)



(e)



(f)

Figure 3-5: First set of scenarios. Behaviour of the proposed method when it has learned that the configuration can achieve: (a) 6 Mbps; (b) 12 Mbps; (c) 24 Mbps; (d) 36 Mbps; (e) 48 Mbps; (f) 54 Mbps.

Initially, in each chart, all conditional probabilities are equal ($\Pr\left[BR = b \middle| CFG = c\right] = 0.166$), since there is no prior information for configuration c. As can be observed, the scheme readily learns the configuration capabilities, and converges to the condition indicated by the calculations. These remarks are backed up by the fact that the conditional probability, which corresponds to the calculated bit rate value, immediately becomes significant and, very soon, larger compared to all the rest.

For instance, in Figure 3-5(b) the calculations indicate that the configuration can achieve br_2 =12 Mbps. Therefore, the probability $\Pr\left[BR = br_2 \middle| CFG = c\right]$ immediately becomes significant (equals to 0.432 and 0.709 after three and ten time steps, respectively), and soon is much higher than the rest (e.g. the probability for a “neighbouring” bit-rate br_1 =6 Mbps equals to 0.289 and 0.186 after three and ten time steps, respectively). Moreover, the behaviour of the probabilities of the bit rates br_1 and br_3 =24 Mbps should be noted. Initially, they are increased, then they remain at a certain high level for an important amount of time, and after some point they start being reduced. These bit-rates are “neighbouring” to br_2. Through this behaviour, the channel estimation phase has learned and shows that these bit rates (even though less probable than br_2) are more representative of the configuration capabilities compared to br_4, br_5 and br_6. As can be observed, after three, five, thirteen and twenty-nine measurements, there are $|T|$ =2, 3, 4 and 5, respectively, probabilities that reach the minimum threshold.

Likewise, in Figure 3-5(d) the calculations indicate that the configuration can achieve br_4 =36 Mbps. Therefore, the probability $\Pr\left[BR = br_4 \middle| CFG = c\right]$ immediately reaches high levels and soon becomes larger than all the others. In this case, as well, the probabilities corresponding to values br_3 and br_5 remain at high levels for several steps. Within five, eight, thirteen steps there are $|T|$ =2, 3 and 5, respectively, probabilities that reach the minimum threshold.

Figure 3-5 (e) and (f) display the same behaviour as Figure 3-5(b) and (a), respectively. This is expected since the initial conditions are the same for all cases and also the indicated bit-rates are at the edges of the set M, for br_6 and br_1 (Figure 3-5 (f),(a)), and near the edges of the set M, for br_5 and br_2 (Figure 3-5 (e),(b)).

Next, there is the presentation of a second category of scenarios, which comprises four sets (two – five) showen in Figure 3-6 – Figure 3-9, respectively. In these sets

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

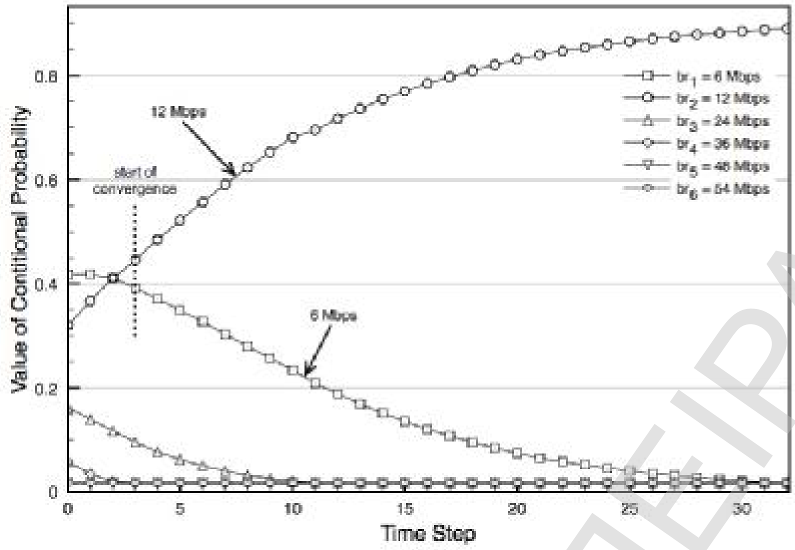
there is prior information on the capabilities of configuration c. Specifically, different situations from the first scenario will be considered as initial conditions. Then, it will be assumed that the bit rate calculations during channel estimation indicate that the capabilities of the configuration change. The objective is to see the behaviour of the proposed scheme.

Figure 3-6 presents the results from the second set of scenarios. It is assumed that the channel estimation phase has learned that configuration c can achieve 6 Mbps, and moreover that $|T|$=2 conditional probabilities have reached the minimum threshold. This is the initial condition in this scenario. In other words, the initial condition is the one of Figure 3-5(a), at time step = 2.

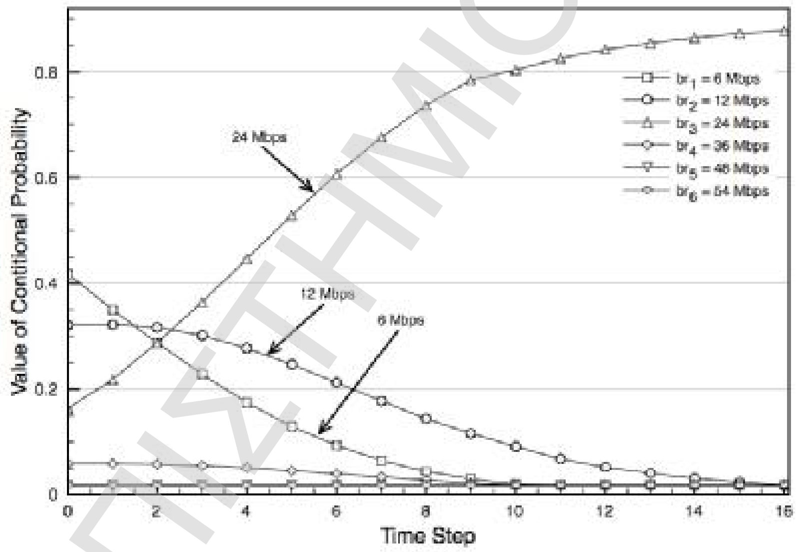
Figure 3-6(a)-(e) shows the evolution of the probabilities when the bit rate calculations indicate that the configuration can achieve 12, 24, 36, 48, 54 Mbps, respectively. Again, in each chart, the x-axis is the time domain, during which there are calculations conducted in discrete steps and provided to the proposed method. The y-axis shows the values of conditional probabilities of the form, $\Pr\left[BR = b \middle| CFG = c\right]$, where b can equal to 12, 24, 36, 48, 54 Mbps.

As can be observed in all cases, the scheme immediately starts to move towards the new situation. This is shown by the fact that immediately the conditional probability, corresponding to the value indicated by the calculations, becomes significant. The fact that there is prior knowledge on the configuration capabilities prevents the immediate convergence (which was the case in the first scenario).

This is a desirable property, for preventing oscillations regarding the estimates of the configuration capabilities, which can be due to temporarily changing environment conditions, e.g., the temporary disappearance or appearance of interferers. Nevertheless, if the change in the environment is not temporary, convergence occurs in a few steps, which range from three (Figure 3-6 (a),(b),(e)) to five (Figure 3-6(d)) (3.6 average).

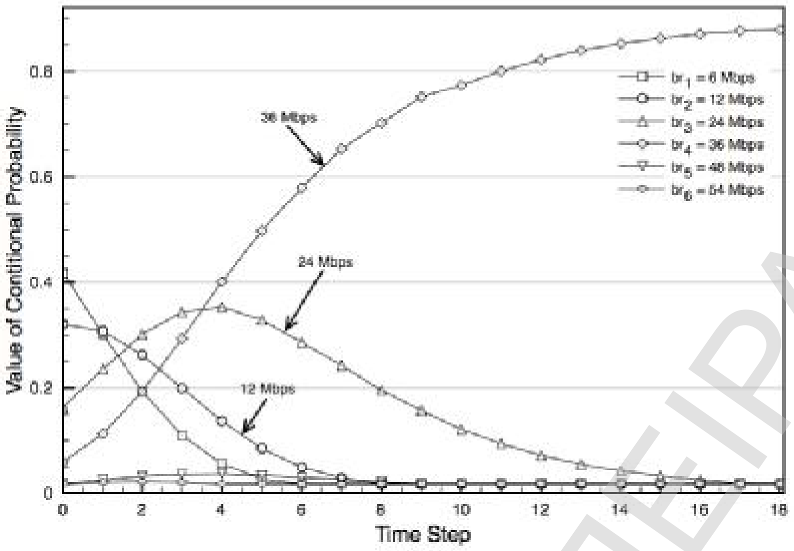


(a)

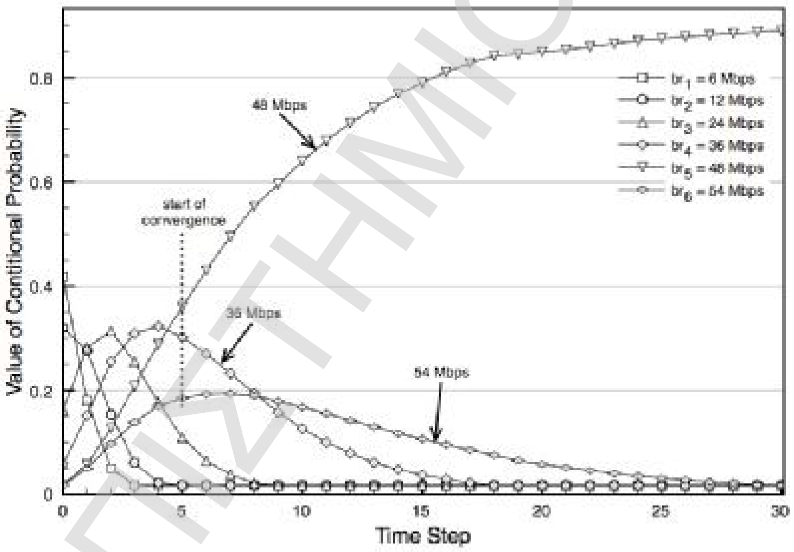


(b)

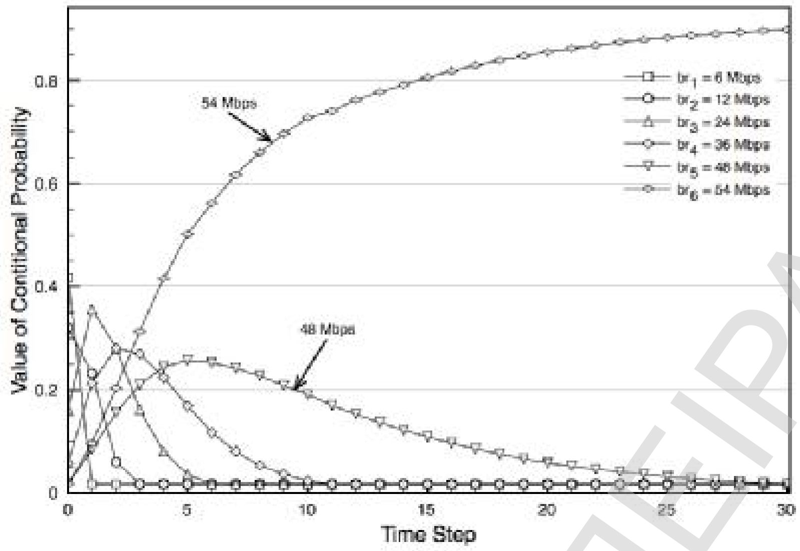
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ


36 Mbps
24 Mbps
12 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps
Value of Contitional Probability
Time Step

(c)


48 Mbps
start of convergence
36 Mbps
54 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps
Value of Contitional Probability
Time Step

(d)

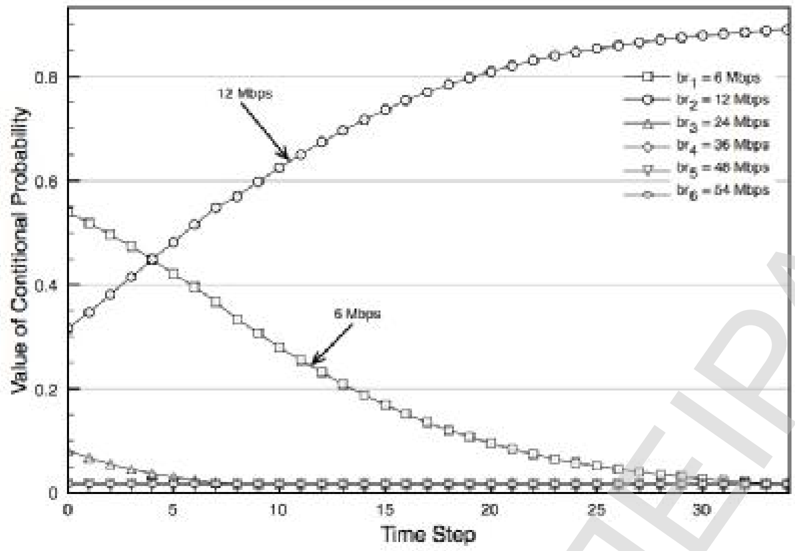


(e)

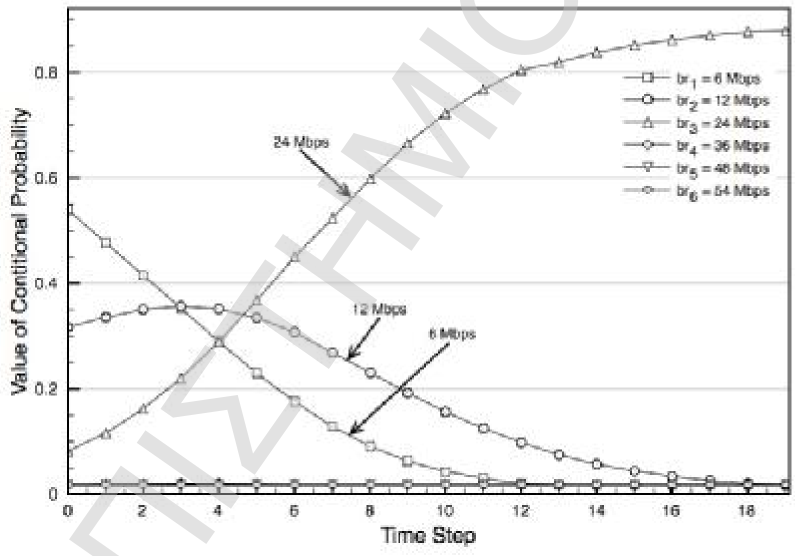
Figure 3-6: Second set of scenarios. Behaviour of the proposed method when the calculated bit-rate changes from 6 Mbps and $|T|$ =2 to: (a) 12 Mbps; (b) 24 Mbps; (c) 36 Mbps; (d) 48 Mbps; (e) 54 Mbps.

In the third set of scenarios (Figure 3-7) it is assumed that during the channel estimation phase, the proposed method has learned that configuration c can achieve 6 Mbps, and moreover that $|T|$ =3 conditional probabilities have reached the minimum threshold. In other words, the initial condition is the one of Figure 3-5(a), at time instant four. The difference of this scenario, with respect to the second one, is that there is a “higher level of convergence” to the initial condition. This means that more probabilities have fallen under the minimum threshold. The question is whether this influences the behaviour of this method, and especially, the speed of convergence to the new condition.

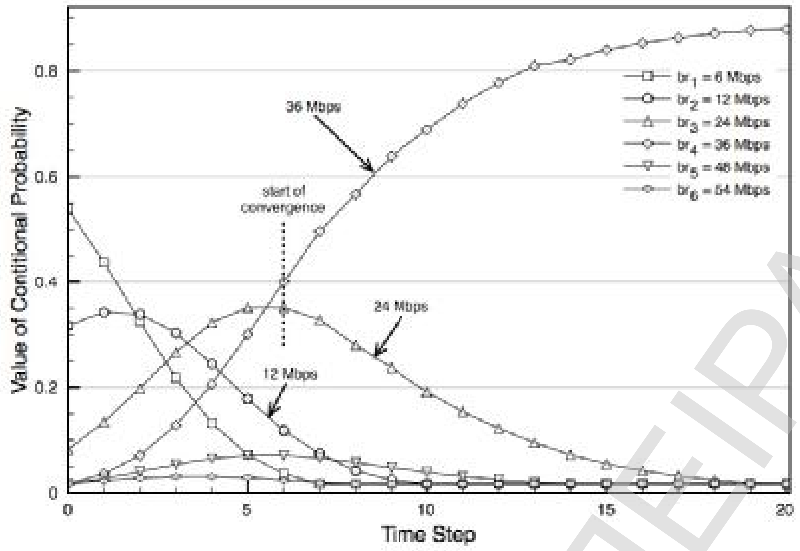
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ


Value of Contitional Probability
Time Step
12 Mbps
6 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps

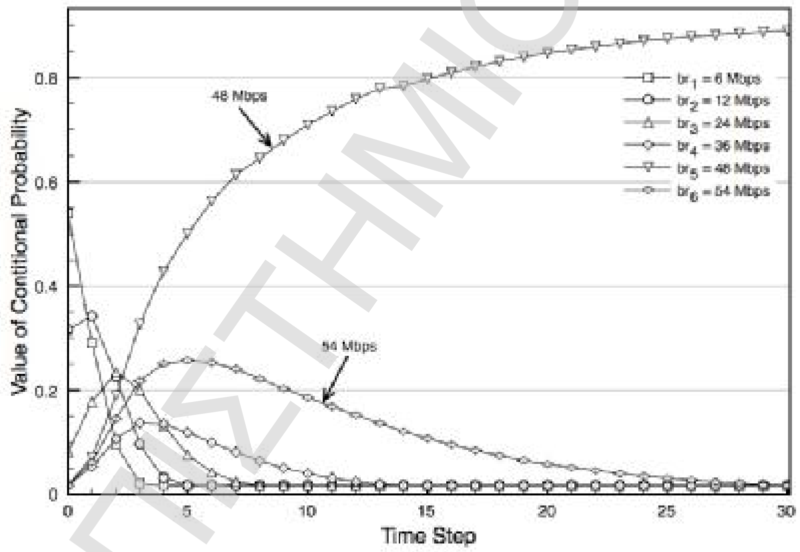
(a)


Value of Contitional Probability
Time Step
24 Mbps
12 Mbps
6 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps

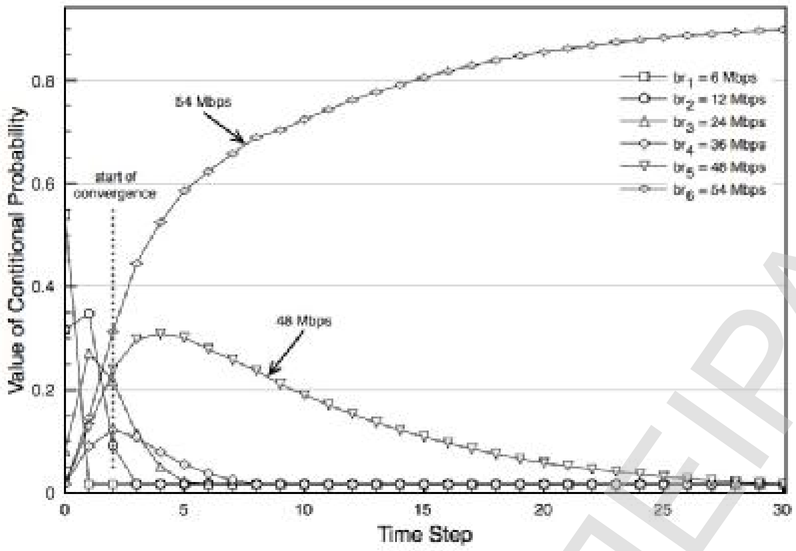
(b)


Value of Contitional Probability
Time Step
36 Mbps
start of convergence
24 Mbps
12 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps

(c)


Value of Contitional Probability
Time Step
48 Mbps
54 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps

(d)



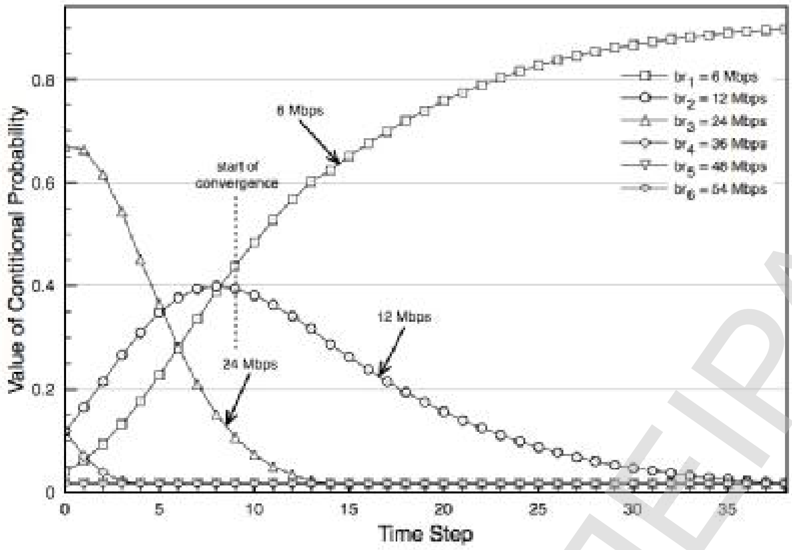
(e)

Figure 3-7: Third set of scenarios. Behaviour of the proposed method when the calculated bit-rate changes from 6 Mbps and $|T|$ =3 to: (a) 12 Mbps; (b) 24 Mbps; (c) 36 Mbps; (d) 48 Mbps; (e) 54 Mbps.

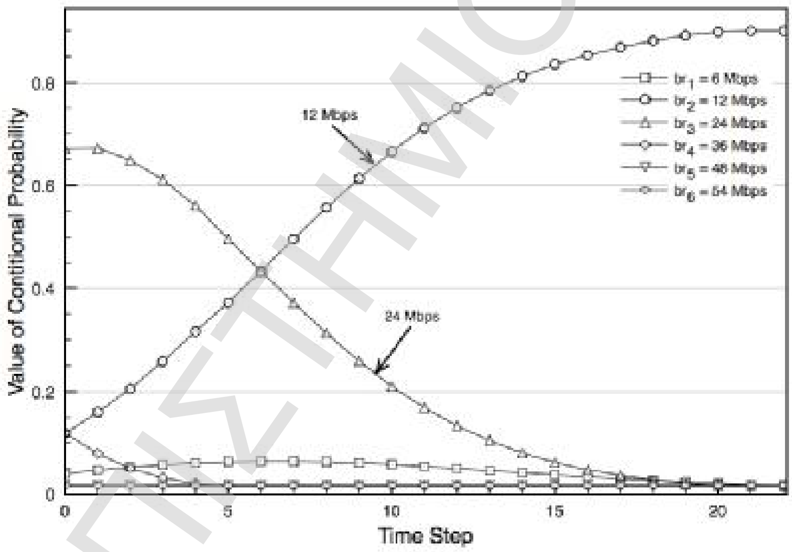
Figure 3-7(a)-(e) shows the evolution of the conditional probabilities when it is calculated that the configuration can achieve 12, 24, 36, 48, 54 Mbps, respectively. As can be observed, in all the cases of the third scenario the method converges to the new condition within few steps. The number of steps ranges again from two (Figure 3-7 (e)) to six (Figure 3-7 (c)) (4.0 average). The number of steps is slightly increased, compared to the second set of scenarios.

In the fourth and fifth set of scenarios it is assumed that the initial condition indicates that the configuration can achieve 24 Mbps. So the difference with respect to the previous two scenarios is that now a “middle” bit-rate value is taken as the initial condition.

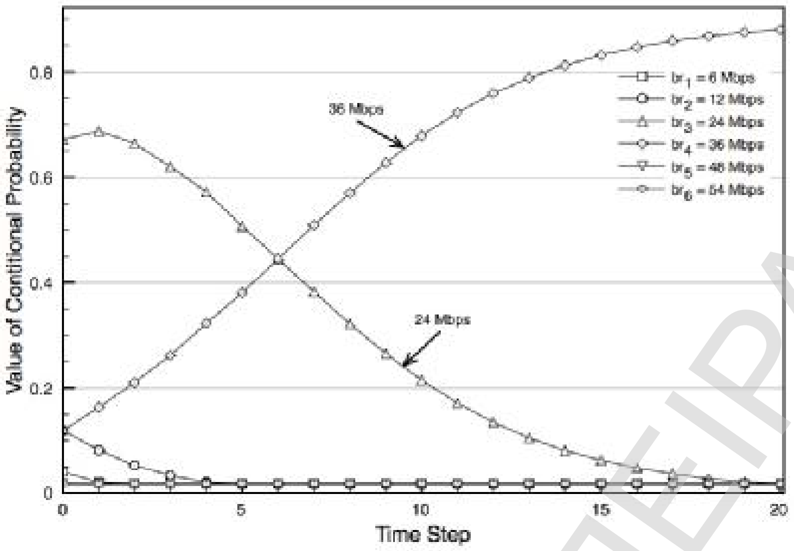
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ


0.8
0.6
0.4
0.2
0
0
5
10
15
20
25
30
35
Value of Contitional Probability
Time Step
br_1 = 6 Mbps
br_2 = 12 Mbps
br_3 = 24 Mbps
br_4 = 36 Mbps
br_5 = 48 Mbps
br_6 = 54 Mbps
6 Mbps
start of convergence
12 Mbps
24 Mbps

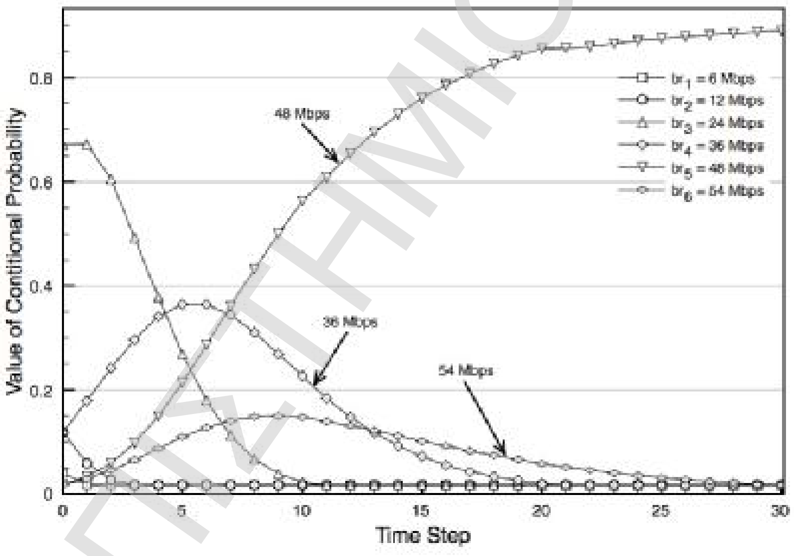
(a)


0.8
0.6
0.4
0.2
0
0
5
10
15
20
Value of Contitional Probability
Time Step
br_1 = 6 Mbps
br_2 = 12 Mbps
br_3 = 24 Mbps
br_4 = 36 Mbps
br_5 = 48 Mbps
br_6 = 54 Mbps
12 Mbps
24 Mbps

(b)

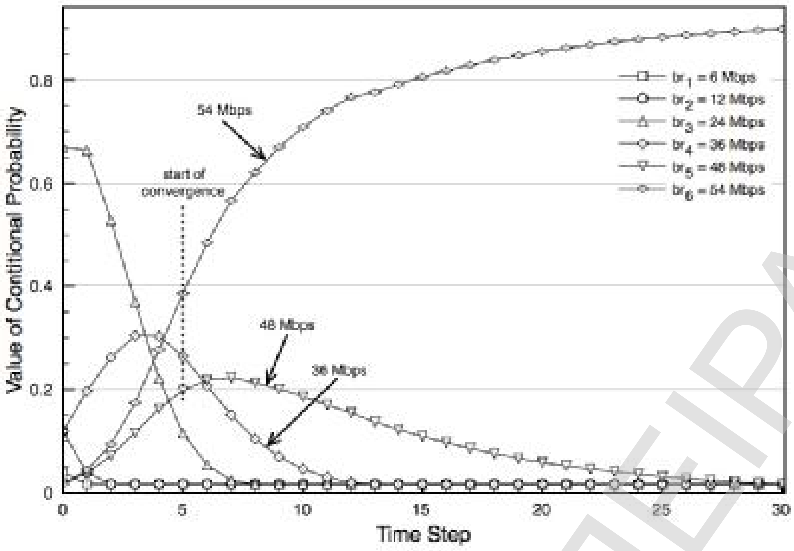


(c)



(d)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

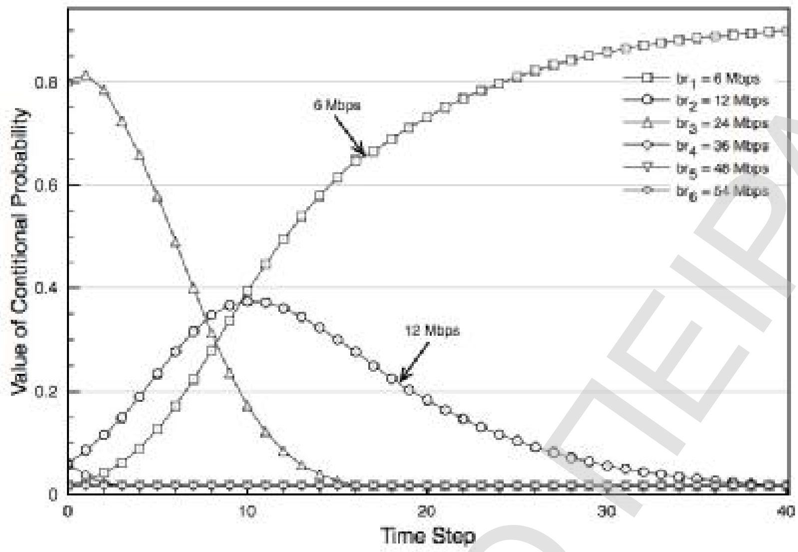


(e)

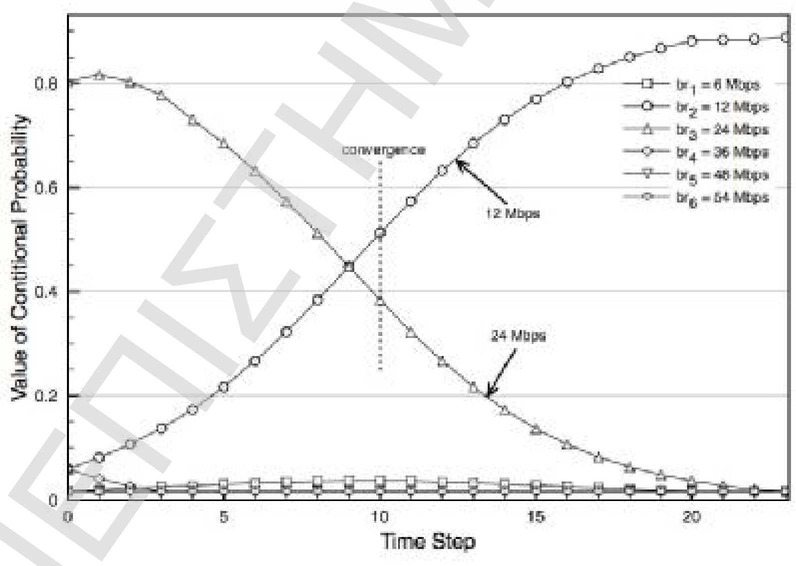
Figure 3-8: Fourth set of scenarios. Behaviour of the proposed method when the calculated bit-rate changes from 24 Mbps and $|T|$ =2 to: (a) 6 Mbps; (b) 12 Mbps; (c) 36 Mbps; (d) 48 Mbps; (e) 54 Mbps.

In the fourth scenario (Figure 3-8), the initial condition is that the configuration can achieve 24 Mbps and that $|T|$ =2 conditional probabilities have reached the minimum threshold. In other words, the initial condition is the one of Figure 3-5(c), at time step=6. Figure 3-8(a)-(e) show the evolution of the conditional probabilities in case the configuration can achieve 6, 12, 36, 48, 54 Mbps, respectively. The remarks that can be drawn from this scenario are similar to those of the second set of scenarios. Specifically, the proposed method starts immediately to move towards convergence to the new condition; convergence occurs in a few steps which ranges from five (Figure 3-8 (e)) to nine (Figure 3-8 (a)) (7.0 steps average).

In the fifth scenario (Figure 3-9), the initial condition indicates that the configuration can achieve 24 Mbps and that $|T|$ =3 conditional probabilities have reached the minimum threshold. In other words, the initial condition is the one of Figure 3-5(c), at time instant eight. Figure 3-9(a)-(e) shows the evolution of the conditional probabilities when according to calculations, the configuration can achieve 6, 12, 36, 48, 54 Mbps, respectively. The behaviour is similar to the previous scenario (8.0 steps average for convergence). Also, when the indicated bit-rate is close to the initial bit rate (i.e. br_2, br_4 on Figure 3-9 (b), (c)), the probability of the "neighbouring" bit-rate immediately raises, while all the rest probabilities drop to the minimum threshold.

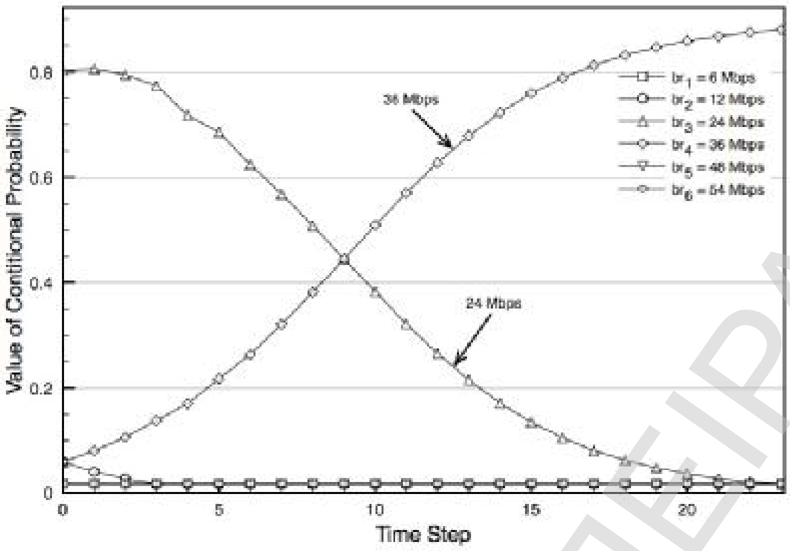


(a)

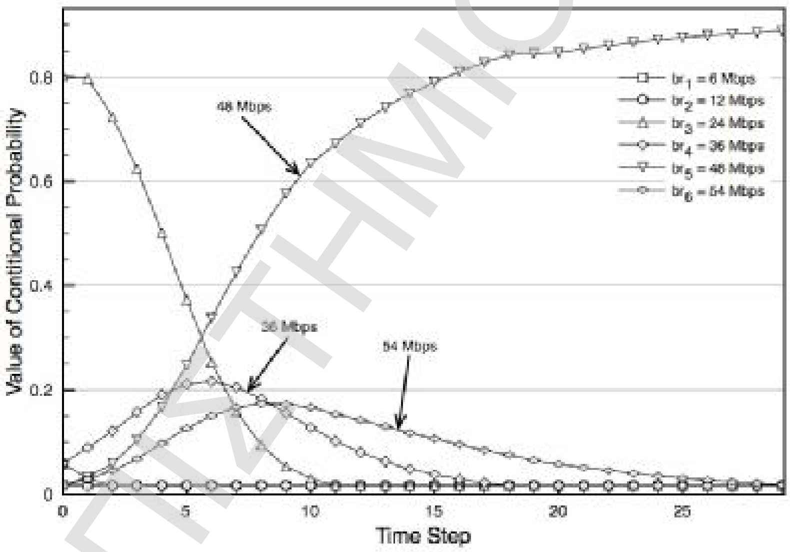


(b)

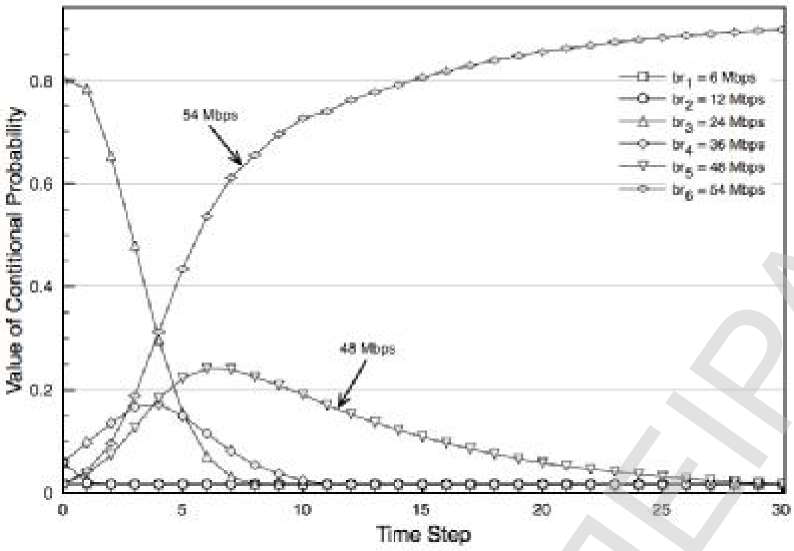
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ


Value of Contitional Probability
Time Step
36 Mbps
24 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps

(c)


Value of Contitional Probability
Time Step
48 Mbps
36 Mbps
54 Mbps
br1 = 6 Mbps
br2 = 12 Mbps
br3 = 24 Mbps
br4 = 36 Mbps
br5 = 48 Mbps
br6 = 54 Mbps

(d)



(e)

Figure 3-9: Fifth set of scenarios. Behaviour of the proposed method when the calculated bit-rate changes from 24 Mbps and $|T|$ =3 to: (a) 6 Mbps; (b) 12 Mbps; (c) 36 Mbps; (d) 48 Mbps; (e) 54 Mbps.

Contrarily, when the indicated values are not "neighbouring" (see Figure 3-9(a), (d), (e)), the "middle" values remain at a high level for a certain number of time steps, until finally reaching the minimum threshold. The results of this scenario also indicate that the "higher level of convergence" minimally impacts the overall behaviour and the speed of convergence.

3.5.3. Analysis

In summary, the behaviour of the proposed Bayesian networks' based method was tested in various scenarios, split in two categories. In the first category (Figure 3-5) the assumption was that there is no prior knowledge on the capabilities of the configuration. In the first category the method phase readily converged to all the situations that can be signalled (calculated) by exploiting the CSI from channel estimation phase. Moreover, the conditional probabilities of the bit rates, which were neighbouring to the bit rate indicated by the calculations, remained at significant levels for a certain amount of time. Therefore, these "neighbouring" bit rates appear as second-best representatives of the configuration capabilities.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

In the second category of scenarios (Figure 3-6–Figure 3-9), it was assumed that the channel estimation phase has learned the capabilities of the configuration. In the first set of scenarios of this category, the initial condition was an “extreme” value, namely, 6 Mbps (scenario sets two and three). In the second set, the initial condition taken was a “middle” value, namely, 24 Mbps (scenario sets four and five). In this category, as well, there was comprehensive investigation with respect to all the potential alterations that can be signalled, resulting from calculations of the bit rates by channel estimation phase (e.g., change from 24 Mbps to all the other values).

In the second category of scenarios it was observed that the scheme immediately starts to move towards convergence to the new condition. Convergence takes more steps compared to the first category. However, it is something positive for avoiding the impact of temporary environment changes. In any case, convergence happens in a few number of steps. Convergence is slightly faster in case the initial condition is an “extreme” value, compared to when it is a “middle” value. The “degree of convergence” to the initial condition minimally impacts the speed of convergence.

The proposed method can exploit any legacy, robust channel estimation mechanism. Assuming a mechanism is available for that purpose, it has been shown that the method can exploit the provided CSI in order to increase the level of certainty that a configuration will achieve a specific bit rate. To strengthen the importance of this statement, the results of the method can be exploited to drive the selection of one of the alternative configurations and thus, ensuring that a cognitive transmitter will always optimize its operation.

3.6. Conclusions

Cognitive radios require machine learning functionality for knowing, with high enough assurance, the capabilities of the alternative configurations in which they might operate, e.g. the achievable bit-rate. Within a cognitive radio operation, channel-state estimation provides significant information (CSI) in order to calculate achievable bit rate values and associate them with a probed configuration. In this respect, this chapter contributes to the enhancement of the channel-state estimation of a cognitive radio process, by proposing a learning method based on Bayesian

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

networks. The objective is to increase the level of certainty that a specific configuration will achieve a definite bit rate.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

REFERENCES

[1] P.Demestichas, A.Katidiotis, K.Tsagkaris, E.Adamopoulou, K.Demestichas, "Channel Estimation in Cognitive Radio Systems by means of Bayesian Networks", Wireless Personal Communications journal, to appear.

[2] J. Mitola, G. Maguire Jr., "Cognitive radio: making software radios more personal", IEEE Personal Communications Magazine, Vol. 6, No. 6, pp. 13-18, Aug. 1999

[3] S. Haykin, "Cognitive radio: brain-empowered wireless communications", IEEE Journal on Selected Areas In Communications, Vol. 23, No. 2, pp. 201-220, Feb. 2005

[4] P. Balamuralidhar, Ramjee Prasad, "A Context Driven Architecture for Cognitive Radio Nodes", Wireless Personal Communications, vol. 45, Issue 3, pp 423- 434, 2008

[5] End-to-End Efficiency (E3) project, https://ict-e3.eu/, 2008

[6] P.Demestichas, D.Boscovic, V.Stavroulaki, A.Lee, J.Strassner, "m@ANGEL: autonomic management platform for seamless wireless cognitive connectivity", IEEE Commun. Mag., Vol. 44, No. 6, June 2006

[7] J.Strassner, "Policy-based network management: solutions for the next generation", Morgan Kaufmann (series in networking), 2005

[8] Neapolitan RE. Learning Bayesian Networks. Prentice-Hall: NJ, 2003.

[9] Pearl J. Probabilistic Reasoning in Intelligent Systems. Morgan Kaufmann: San Francisco, 1988.

[10] Jensen F. Bayesian Networks and Decision Graphs. Springer-Verlag: NY, USA, 2001.

[11] Qi Jiang, Joachim Speidel, Chunming Zhao "A Joint OFDM Channel Estimation and ICI Cancellation for Double Selective Channels", Wireless Personal Communications, vol . 45, Issue 1, pp.131-143, 2008.

[12] M. Morelli and U. Mengali, "A Comparison of Pilot-Aided Channel Estimation Methods for OFDM Systems," IEEE Trans. Signal Processing, vol.49, no.12, pp.3065-3073, Dec. 2001.

[13] R. W. Heath and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," IEEE Trans. Signal Processing, vol. 47, no.3, pp.848-856, Mar.1999.

[14] S. Zhou and G. B. Giannakis, "Finite-Alphabet Based Channel Estimation for OFDM and Related Multicarrier Systems," IEEE Trans. Commun., vol.49, no.8, pp.1402-1414, Aug. 2001.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

[15] S. Zhou, B. Muquet, G. B. Giannakis, "Subspace-based (semi-) blind channel estimation for block precoded space-time OFDM," IEEE Trans. Signal Proc., vol. 50, no.5, pp. 1215- 1228, May 2002.

[16] A.Petropulu and R. Zhang, "Blind OFDM Channel Estimation through Simple Linear Precoding," IEEE Trans. Wireless Commun., vol.3, no.2, pp.647-655, March, 2004.

[17] Rashad, I.; Budiarjo, I.; Nikookar, H., "Efficient Pilot Pattern for OFDM-based Cognitive Radio Channel Estimation - Part 1," Communications and Vehicular Technology in the Benelux, 2007 14th IEEE Symposium on , vol., no., pp.1-5, 15-15 Nov. 2007.

[18] Budiarjo, I.; Rashad, I.; Nikookar, H., "Efficient Pilot Pattern for OFDM-based Cognitive Radio Channel Estimation - Part 2" Communications and Vehicular Technology in the Benelux, 2007 14th IEEE Symposium on , vol., no., pp.1-5, 15-15 Nov. 2007.

[19] A. Soysal., S. Ulukus, C. Clancy, "Channel Estimation and Adaptive M-QAM in Cognitive Radio Links", in Proc. IEEE International Conference on Communications (ICC) 08', Beijing, China, 19th -23rd May, 2008.

[20] Heckerman, D.: "A Tutorial on Learning with Bayesian Networks". In Report No. MSR-TR-95-06, Microsoft Research (1995).

[21] Tetko, I. V.; Livingstone, D. J.; Luik, A. I. Neural network studies. 1. Comparison of overfitting and overtraining, J. Chem. Inf. Comput. Sci., 1995, 35, 826-833.

[22] R. Khanafer, L. Moltsen, H. Dubreil, Z. Altman, and R. Barco, "A Bayesian Approach for Automated Troubleshooting for UMTS Networks," Proc. 17th IEEE Int'l Symp. Personal, Indoor and Mobile Radio Comm. (PIMRC '06), Aug. 2006.

[23] R. Barco, V. Wille, and L. Dı´ez, "System for Automated Diagnosis in Cellular Networks Based on Performance Indicators," European Trans. Telecommunications, vol. 16, no. 5, pp. 399-409, Oct. 2005.

[24] R. Barco, P.Lázaro, L. Díez, V.Wille, "Continuous versus discrete model in auto-diagnosis systems for wireless networks," IEEE Transactions on Mobile Computing, 06 Feb 2008, in press.

[25] A. Koutsorodi, E. Adamopoulou, K. Demestichas, M. Theologou, "Service configuration and user profiling in 4G terminals". Wireless Personal Communications, Vol 43, No 4, p.p.1303-1321, Dec. 2007.

[26] K. Demestichas, A. Koutsorodi, E. Adamopoulou, M. Theologou, "Modelling user preferences and configuring services in B3G devices", Wireless Networks, July 2007, in press.

[27] Bauer, E., Koller, D., Singer, Y.: Update Rules for Parameter Estimation in Bayesian Networks. Proceedings of the 13th. Annual Conference on Uncertainty in AI (1997) 3-13.

[28] Zhang, S. Z., Yu, H., Ding, H., Yang, N. H., Wang, X. K., "An application of online learning algorithm for Bayesian network parameter," Machine Learning and Cybernetics, 2003 International Conference on , vol.1, no., pp. 153-156 Vol.1, 2-5 Nov. 2003.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

4. NEURAL NETWORK-BASED LEARNING SCHEMES FOR COGNITIVE RADIO SYSTEMS

Abstract:

Intelligence is needed to keep up with the rapid evolution of wireless communications, especially in terms of managing and allocating the scarce, radio spectrum in the highly varying and disparate modern environments. Cognitive radio systems promise to handle this situation by utilizing intelligent software packages that enrich their transceiver with radio-awareness, adaptability and capability to learn. A cognitive radio system participates in a continuous process, the "cognition cycle", during which it adjusts its operating parameters, observes the results and, eventually takes actions, that is to say, decides to operate in a specific radio configuration (i.e., radio access technology, carrier frequency, modulation type etc.) expecting to move the radio toward some optimized operational state. In such a process, learning mechanisms that are capable of exploiting measurements sensed from the environment, gathered experience and stored knowledge, are judged as rather beneficial for guiding decisions and actions. Framed within this statement, this chapter introduces and evaluates learning schemes that are based on artificial neural networks and can be used for predicting the capabilities (e.g. data rate) that can be achieved by a specific radio configuration. In particular, the focus in this work is placed on obtaining insight on the behaviour of the presented, learning schemes, whereas useful, indicative results from the benchmarking work, conducted in order to design and use an appropriate neural network structure, are also presented and discussed. In the near future, such learning schemes are expected to assist a cognitive radio system to compare among the whole of available, candidate radio configurations and finally select the best one to operate in. Parts of this chapter have been published in [1].

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

NEURAL NETWORK-BASED LEARNING SCHEMES FOR COGNITIVE RADIO SYSTEMS

4.1. Introduction and Problem Statement

The integration of a learning engine can be important especially for the channel estimation and predictive modelling phase, for improving the stability and reliability of the discovery and evaluation of the configuration capabilities, without relying solely on the recent measurements. To this effect, many different learning techniques are available and can be used by a cognitive radio ranging from pure lookup tables to arbitrary combinations of machine learning techniques that include artificial neural networks, evolutionary/genetic algorithms, reinforcement learning, hidden Markov models, etc.

This chapter contributes towards this direction, by developing learning schemes that rely on artificial neural networks and aim at solving the problem related to the channel estimation and predictive modelling phase of cognitive radio systems and stated as follows: "*Given a candidate radio configuration, what are its anticipated capabilities (e.g., in terms of achievable data rate), taking into account recent information sensed, as well as the past experience and knowledge?*".

More specifically, the objectives of the work in this chapter can be decoupled as follows:

- First, to propose two learning schemes, a 'basic' and an 'extended' one, that are based on neural networks and are designed to enhance the learning capabilities of a cognitive terminal, in terms of assisting it to predict the data rate that a specific radio configuration could achieve if it was selected for operation, and
- Second, to perform a benchmarking work upon the proposed neural network–based (NN-based) schemes and discuss upon their applicability to future cognitive radio systems.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

The rest of the chapter is structured as follows: A review of neural networks is provided in Section 4.2 and the motivation for their application to cognitive radio systems is presented in Section 4.3. The basic NN-based learning scheme is described and evaluated in Section 4.4, while the extended NN-based scheme is presented in Section 4.5. Finally, the chapter is concluded in Section 4.6.

4.2. Neural networks review

Biological neural networks are made up of real biological neurons that are physically connected or functionally-related in the human nervous system and especially in the human brain. Artificial neural networks (ANN or simply NN) on the other hand, are made up of artificial neurons interconnected to each other to form a programming structure that mimics the behaviour and neural processing (organization and learning) of biological neurons.

Human brain can perform tasks much faster than the fastest existing computer thanks to its special ability in massive parallel data processing[2]. NNs try to mimic such a providential behaviour for solving narrowly defined problems i.e., problems with an associative or cognitive tinge [2]. To this effect, NNs have been extensively and successfully applied to pattern (speech/image) recognition, time-series prediction and modelling, function approximation, classification, adaptive control and other areas.

As stated, a neural network consists of a pool of simple processing-computing units, the 'neurons'. Within NNs three types of neurons are distinguished (upper part of Figure 4-1): input neurons which receive data from outside the NN and are organized in the so called input layer, output neurons which send data out of the NN and generally comprise the output layer, and hidden neurons whose input and output signals remain within the NN and form the so called hidden layer (or layers).

[2] Engineers of IBM's Deep Blue, chess-playing computer had a different opinion on that!

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

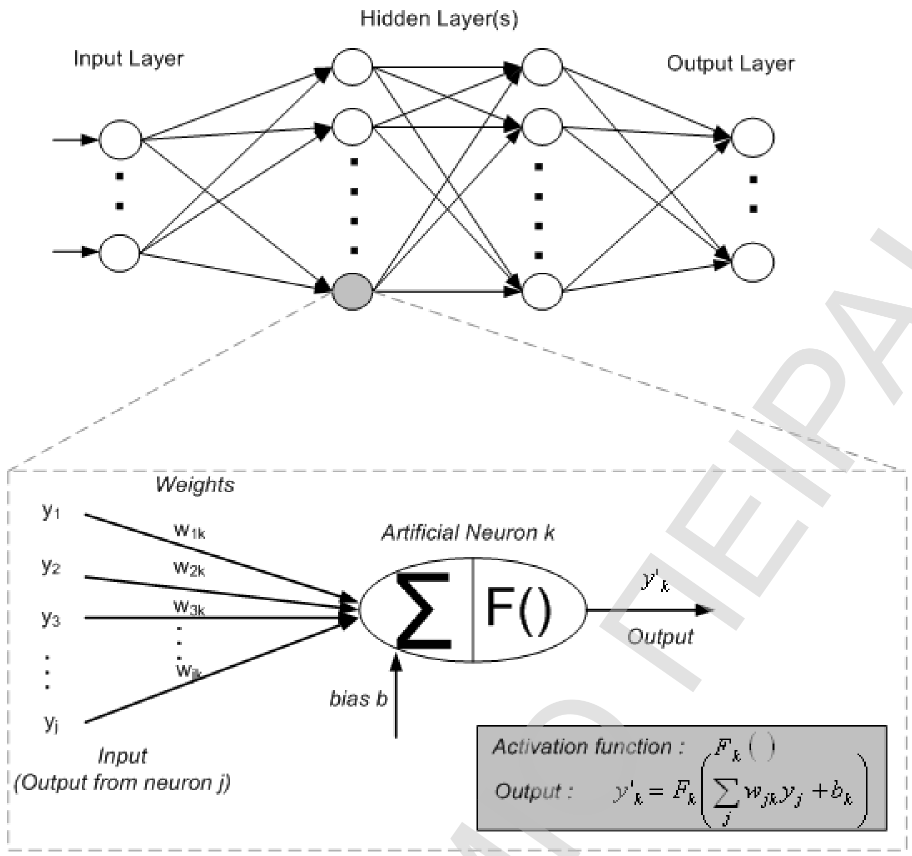


Figure 4-1: Typical Neural Network structure.

Neurons are communicating with each other by sending signals over a large number of weighted connections, thus creating a network with a high degree of interconnection. Generally each connection is defined by a weight, w_{jk}, which determines the effect that the signal of neuron j has on neuron k. Every neuron has a state of activation, be it y_k, which is equivalent to the output of the neuron. During processing, each neuron k receives input s_k from a) neighbours belonging to different layers, as well as from b) external sources a.k.a. bias offset b_k, and uses them to compute an updated level of activation $y_k^{'}$. This is done through the use of an activation function F_k as follows (see also lower part of Figure 4-1):

$$F_k\left(\sum_j w_{jk} y_j + b_k\right) \qquad (1)$$

Some sort of threshold functions can be used for that reason, such as a sgn function, a linear function, or a smoothly limiting function often being a sigmoid (S-shaped)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

function like the logistic-sigmoid transfer function a.k.a. logsig in [3] and expressed by

$$F_k(s_k) = \frac{1}{1+e^{-s_k}} \qquad (2),$$

the hyperbolic tangent sigmoid transfer function a.k.a. tansig in [3] and expressed by

$$F_k(s_k) = \frac{2}{1+e^{(-2\cdot s_k)}} - 1 \qquad (3),$$

or the linear transfer function a.k.a. purelin in [3] and expressed by

$$F_k(s_k) = s_k \qquad (4).$$

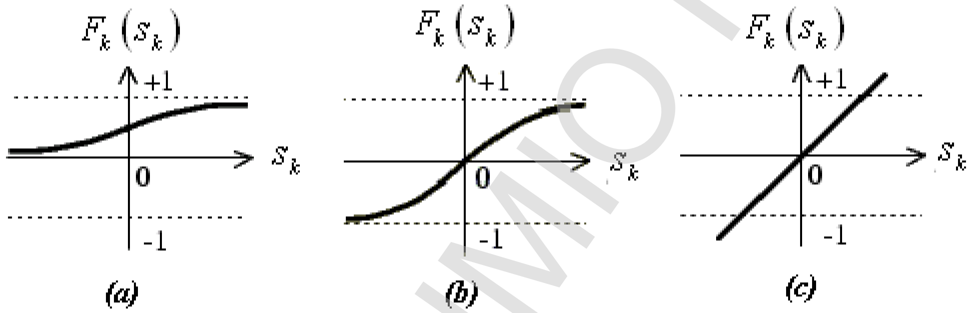


Figure 4-2: Transfer Functions (a) Log-sigmoid transfer function (b) Hyperbolic tangent sigmoid transfer function (c) Linear transfer function

Transfer functions (2), (3) and (4) are graphically depicted in Figure 4-2(a), (b) and (c), respectively. The topology of a NN plays an important role for its achievable performance. Depending on the pattern of connections that a NN uses to propagate data among the neurons, it can be classified into one over two basic (non exhaustive) categories. a) Feed-forward NNs, where data enters at the inputs and passes through the network, layer by layer, until it arrives at the outputs, with classical examples being the Perceptron [4] and Adaline [5]. b) Recurrent NNs that contain feedback connections, which are connections extending from outputs of neurons to inputs of neurons in the same layer or previous layers. In contrast with feed-forward networks, the recurrent network has a sense of history and this means that pattern presentation must be seen as it happens in time. Examples of recurrent networks have been presented by Elman [6] and Hopfield [7]. In the analysis within this chapter, both feed-forward and recurrent networks are used.

In any case, a NN has to be configured such that the application of a set of inputs produces the desired set of outputs. This can be achieved by properly adjusting the weights w_{jk} of the existing connections among all (j,k) neuron pairs. This process is called "*learning*" or "*training*". Learning can be generally distinguished between supervised and unsupervised learning (with reinforcement learning being also an option). In supervised learning, the NN is fed with teaching patterns and trained by letting it change its weights according to some learning rule, the so called back-propagation rule. The NN learns the input-output mapping by a stepwise change of the weights with the objective to minimize the difference between the actual and desired output. In the next step the actual output vector is compared with the desired output. Error values are assigned to each neuron in the output layer. The error values are back-propagated from the output layer to the hidden layers. The weights are changed so that there is a lower error for a new presentation of the same pattern. As a result of this procedure, the weights on the connections between neurons are properly adjusted so as to encode the actual knowledge of the NN. At that time, the NN can be used for the purpose that was initially set up for. On the contrary, in unsupervised learning the NN discovers remarkable features of the input data in a statistical manner by developing its own ways of classifying the input irritants. This chapter, deals only with supervised learning mode.

4.3. Motivation

The area of cognitive radios and systems is judged as rather suitable for accommodating NNs. In particular, NNs can be used in any of the phases of the cognition cycle described before, in the introductory chapter. For instance, during the radio-scene analysis, temporal statistics of a radio environment can be used to isolate distinct characteristics which in turn correspond to different modulations. These statistics can then feed a NN in order to classify a signal's modulation type as proposed in [8]. Going one step further, the extracted modulation type can be used in the sequel to characterize the whole radio configuration e.g. QPSK unveils the existence of a W-CDMA RAT.

In [9], two neural classifiers are presented and compared to each other, while used for the identification of two communication modes: a direct sequence-based WLAN

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

802.11b and a frequency hopping-based Bluetooth coexisting in the same environment and operating in the same ISM band.

In this chapter the focus is on the channel estimation and predictive modelling phases and a benchmarking work is presented that aims at evaluating the applicability of multiple types of NNs in the learning module of the cognitive engine within a cognitive terminal (see Figure 4-3).

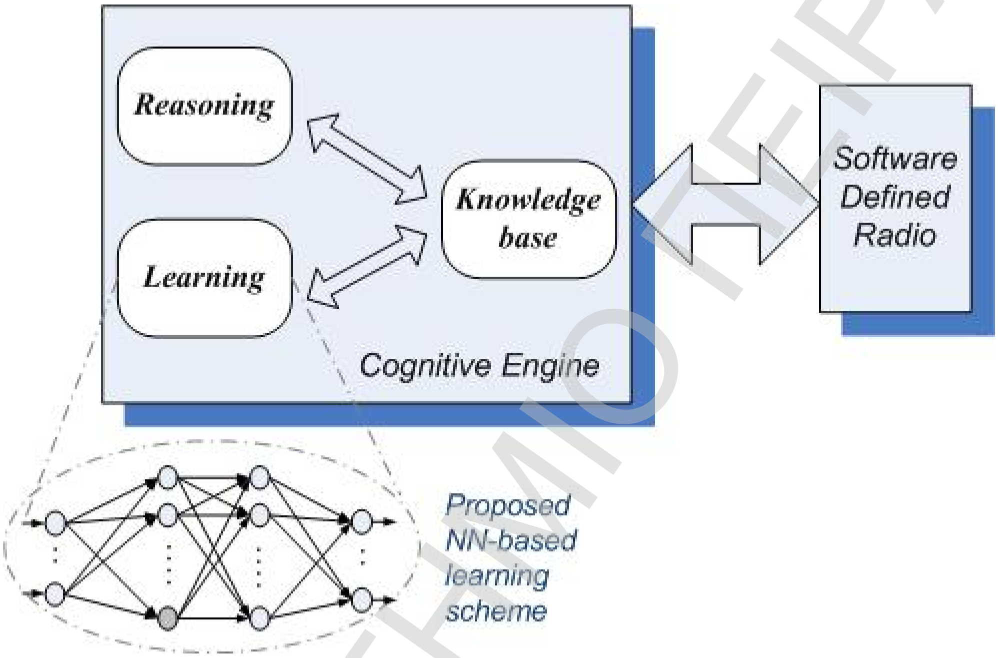


Figure 4-3: Cognitive radio engine

The proposed NN-based learning schemes should relax the reasoning process and assist in the optimum decision regarding the radio-configuration settings (mainly PHY and MAC layer) that provide the best QoS for the given problem and user/application needs. It should be noted that QoS optimization is a multi-objective problem that depends on many quality metrics with dependent relationships, including bit error rate, frame error rate, power consumption, latency, data rate etc, and as such, it should call for Pareto optimality, which balances the trade-offs among the multiple objectives. Nevertheless, the focus here is placed only on one objective: *the data rate*. This is also aligned with the initial purpose of this chapter, namely to showcase the feasibility of the proposed learning schemes, but it could be extended, as shown in the next chapter, to take into account more transmission parameters. More specifically, it is proposed a way that a NN can be used a) to learn from the information measured by the terminal during the radio-scene analysis and b) to provide in the output the data rate that is most anticipated to be obtained per radio configuration (RAT/frequency), thus behaving as a predictor of the next expected data rate. What is gained is that by associating each configuration with a predictable, achievable data rate, the proposed NN-based learning schemes may facilitate the cognitive terminal in making its decision regarding the configuration in which it should operate, selecting the best among a set of candidate ones.

Accordingly, two neural network–based learning schemes have been set up and tested: the ‘basic’ and the ‘extended’ one. In both cases, multiple types of NNs with a considerable number of adjustable parameters have been investigated. These trial and error processes, that were conducted in order to derive the best possible network in both basic and extended cases, are described and discussed more analytically in the following sections.

4.4. Basic NN-based learning scheme

It must be noted that all simulation analyses and/or results presented hereafter, assume that the NN-based scheme is tuned in an arbitrary radio configuration e.g. IEEE WLAN 802.11g. This is actually the radio configuration, the capabilities of which (i.e., data rate) need to be discovered-evaluated.

4.4.1. Preparation Procedure

In order to exhibit the applicability of such NN-based learning schemes, an algorithm is selected that will be used to train variously parameterized NNs so as to predict the data rate to be obtained by the configuration being under investigation. This algorithm aims at defining a target data rate for each of the input value(s) presented in the NN and is analysed in the sequel of this sub-section.

Let $R = \{r_k\}$, be the time-series collected by the radio scene analysis (environment sensing) phase, where each element r_k represents a data rate value at time slot k, $k \in \mathbb{N}^*$. It is also assumed that values r_k are quantized in predefined reference values from a finite set $M = \{m_1, m_2, ..., m_{|M|}\}$. A time window of n slots is used to represent past experience and knowledge collected by the NN and is depicted in Figure 4-4. At

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

any time k, the NN is fed with an input sequence $R^{in} \subseteq R$, the length of which equals the size of the time window, i.e., $R^{in} = \{r_i\}$, $i = 1,\ldots,$n. In addition, in each slot within the time window the corresponding value r_i is associated with a weight, b_i $i = 1,\ldots,$n [3]. Requiring that recently collected values should have greater weights, an exponentially weighted moving average with a smoothing factor a is used to configure b_i, i.e., $b_i = a \cdot (1-a)^i$ [10]. The objective of the above algorithm is to exponentially decrease the weighting for each older data value, giving much more importance to recent observations, while still not entirely discarding older observations.

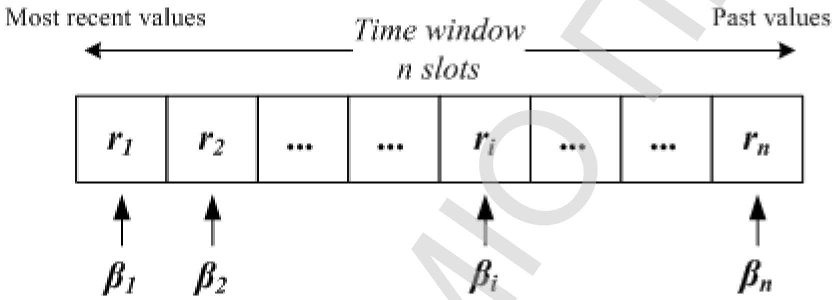


Figure 4-4: Time window.

For each r_k there will be a target data rate value, r_k^{tgt}, that will be used to train the basic NN at each time k and is derived as follows:

Consider the above specific window of n slots. In order to derive r_k^{tgt}, the input time sequence R^{in} of length n is used. Each of the reference values m_j, $j = 1,...,|M|$ is associated with a weight b_{m_j} given by $b_{m_j} = \left\{ \sum_{i=1}^{n} b_i \mid r_i = m_j \right\}$, $i = 1,\ldots,$n. Weights b_{m_j} actually represent the number of occurrences of each of the reference values m_j in M within the time window. Eventually, the target data rate value corresponding to the input time sequence within the time window will arise from the following relation:

---

[3] Weights b_i must not be confused with the weights of the connections among neurons.

$$r_k^{tgt} = \arg\max_{m_j} b_{m_j} \quad (5)$$ [4]

In other words, the target value selected is the one that has the maximum weighted sum, within the time window.

4.4.2. NN pattern selection - Results

In this sub-section the focus is given on the selection of that pattern of the NN, which gives the best performance in terms of minimizing a predefined metric such as the Mean Squared Error (MSE). All simulations and results are conducted using the Neural Network Toolbox of Matlab 7.1 [3].

Let set M contain $|M|$=6 reference bit rate values (in Mbps) as follows: m_1 = 6, m_2 = 12, m_3 = 24, m_4 = 36, m_5 = 48, m_6 = 54 i.e., the values that might correspond to the data rate obtained by a typical WLAN-equipped terminal. The time window equals to $n = 5$.

The smoothing factor a of the exponential moving average algorithm is arbitrarily set to a = 0.181, thus resulting to the respective calculated weights $\{b_i\}$ = {0.1488, 0.1217, 0.0996, 0.0814, 0,0666}. The time-series R includes values from the M set, which are randomly generated according to a selected distribution function, depicted in Figure 4-5 (normal line), that assigns bigger probability to the appearance of m_1 = 6. The target values r^{tgt} are calculated according to (5).

The NN that has been selected for the basic scheme is an Elman network [6], which is a two-layer back-propagation, recurrent network, with the addition of a feedback connection from the output of the unique hidden layer to the input layer. This

---

4 Recall that the ***argmax*** function stands for the ***argument of the maximum***. For instance, $\arg\max_x f(x)$ returns the value of the argument i.e., *x*, for which the value of the given expression i.e., *f(x)*, attains its maximum value

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

recurrent connection allows the Elman network to both detect and generate time-varying patterns. The NN uses the tansig function in equation (2) for the neurons in its hidden (recurrent) layer, and logsig function in equation (3) for the neuron in its output layer, respectively. A delay line of five slots has been inserted in the input layer, which corresponds to the time window, as mentioned previously.

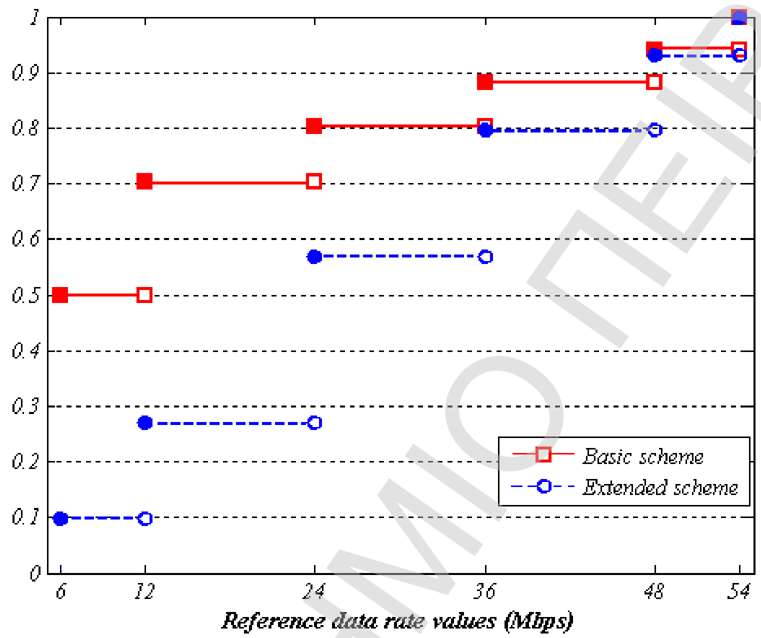


Figure 4-5: Cumulative distribution functions of input time-series.

Several test cases, each of which corresponding to an Elman type network with varying number of neurons in the hidden layer, size of training data set and also training parameters, have been examined during a trial and error procedure for finding the optimum NN design pattern.

Table 4-1 gives an overview of the parameters and their values that have been used during this investigation procedure, which produces 84 different test cases in total (all arising combinations).

For the training session, the input and target values have been properly normalized in the range of [0, 1] in a pre-processing phase [11]. During training, weights and bias values have been updated according to a gradient descent momentum and an adaptive learning rate method (a.k.a. traingdx in [3]). Finally, as already stated, the Mean Squared Error has been used as a metric for measuring the Elman neural network's performance. In this analysis, two data sets have been used, which were extracted from the whole input sequence and served as the target values for teaching the NN: a) a "training" set (seen data) which is used to build the model i.e. determine its parameters, during the so called training session, and b) a "validation" set (unseen data) which is used to measure the performance of the network by holding its parameters constant. With the term "unseen", data that have never been used to update the weights of the network are characterized.

Table 4-1: Test cases examined for basic scheme

Parameters	***Values***
Number of hidden neurons	2 / 5 / 10 / 15 / 20 / 25 / 30
Sample data points	300 / 1000 / 3000
Training epochs	200 / 500
Training learning rates	0.01 / 0.001

The importance of testing the network with both datasets, when searching for the best structure, is significant, since a small error in the training set can be misleading. If the network has not been trained well, it may not learn the basic structure of the data, but rather learn irrelevant details of the individual cases, a.k.a. *overfitting* the training data or *overtraining*. This would lead to a small error during testing with the training set, but in a large error during testing with the validation dataset. In general, performance on the training only tells us that the model learns what it's supposed to learn, but it is not a good indicator of performance on unseen data i.e. whether the NN is able to generalize well or not [12].

Moreover, the number of hidden layers and/or neurons plays a critical role in the learning process and strongly influences the performance of the network. The use of too few hidden neurons would result in a NN that is unable to learn what it is supposed to learn. On the other hand the use of too many hidden neurons would dramatically increase the time needed to learn, without yielding any significant improvement in the performance of the network. This is already captured in this chapter and is referred as overfitting. There exist some valid rules to set the number of

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

hidden nodes [12] but in general, it is better to start with a big net, train, and then carefully follow a pruning strategy for gradually reducing the size of the network [13].

As long as training is finished, the performance of the trained NN has been tested in both the "known" and "unknown" sequences comprising 100 data points each. The known sequence is actually a subset of the "training" set. Also, in order to measure the NN's degree of generalization [12], a completely unknown sequence of 100 values has been selected from the whole data set in order to constitute the so called validation set or validation sequence. During the validation, the MSE between the value produced by the NN and the expected target value has been recorded.

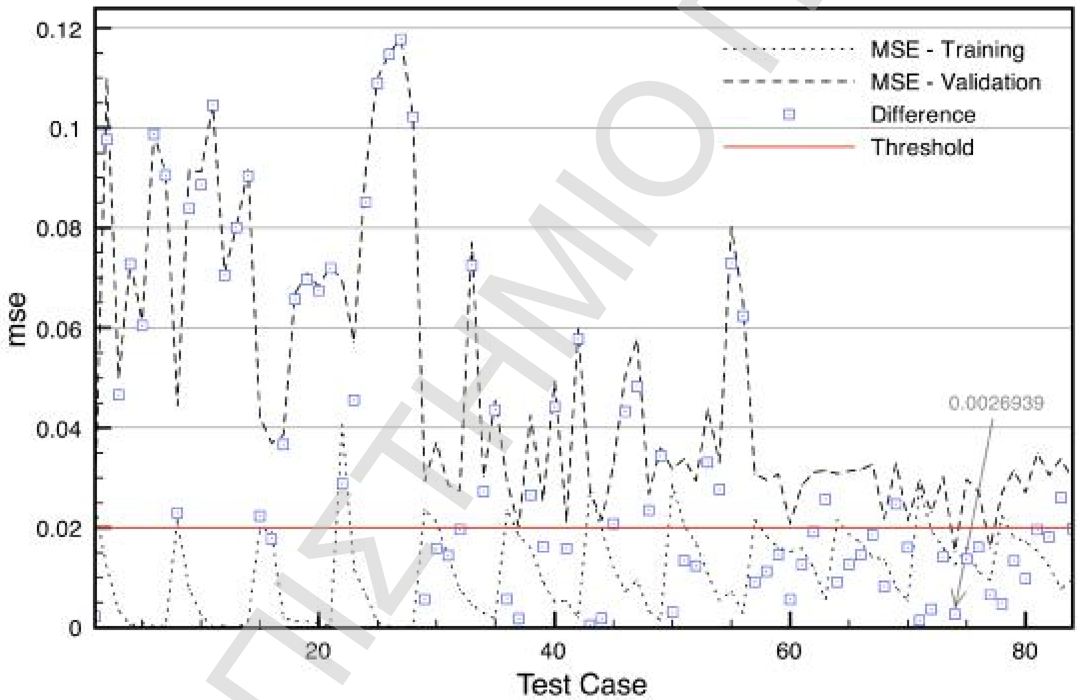


Figure 4-6: Measured performance (MSE) at both training and validation sequences for all 84 test cases – Basic scheme.

Figure 4-6 depicts the resulting MSE recorded at the end of both the training and validation sessions for all the above 84 test cases.

Accordingly, an acceptable NN design pattern should satisfy the following criteria:

- $(MSE_{trn} \leq MSE_{thres}) \wedge (MSE_{val} \leq MSE_{thres})$, where MSE_{trn} - is the final MSE produced during the training session, MSE_{val} - is the final MSE produced during the validation and MSE_{thres} - is a desirable upper threshold for the MSE which is arbitrarily set here to 0.02, and

- Minimize $\left|MSE_{trn} - MSE_{val}\right|$.

The first criterion is self-explanatory. Regarding the selection of the MSE threshold, the network has been requested to run for a number of epochs sufficient to lower the MSE to a little amount (MSE goal). Based on observations during offline trial and error efforts, which have been conducted for determining the NNs structure, but are not recorded here for brevity reasons, setting an upper threshold for the produced MSE at a value higher than that of 0.02 or 2%, resulted in lower performance in terms of how well the NN "learnt its lesson". Moreover, an upper limit of 0.02 would be enough to provide adequate performance in both seen and unseen datasets, simultaneously. This is also depicted in Figure 4-6.

As for the second criterion, it is used here in order to guarantee a certain level of generalization, meaning that the neural network must have the ability to behave efficiently when dealing with unseen input data and thus avoiding overfitting the training data.

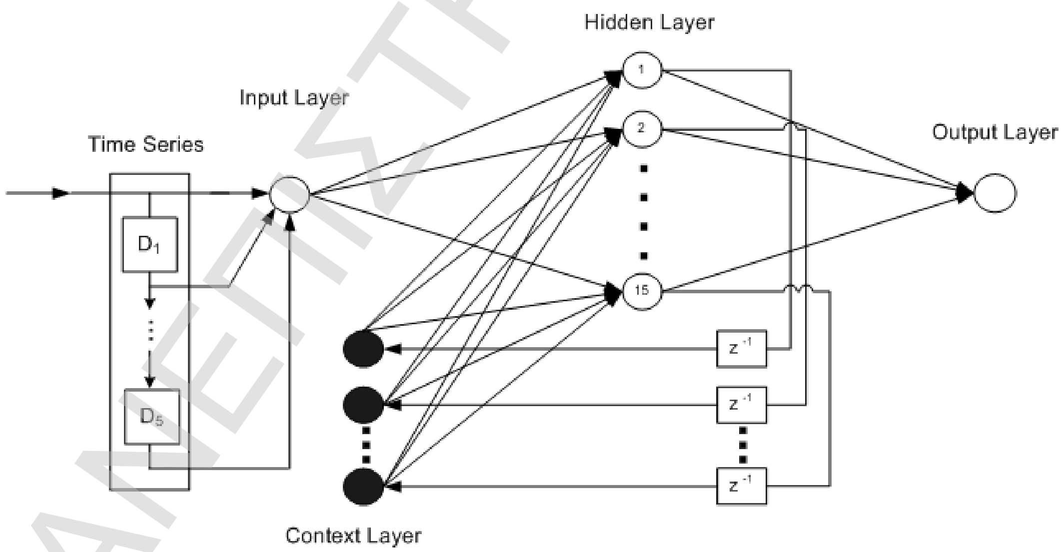


Figure 4-7: Neural network for the Basic Learning scheme.

Figure 4-6 also illustrates the difference between the MSE_{trn} and MSE_{val} produced for the above test cases. As it can be observed and also indicated by the arrow in Figure 4-6, by following the above rules, the best available network design pattern is the one corresponding to the 74th test case and depicted in Figure 4-7. This case designates an Elman NN with 15 tansig hidden nodes in the hidden layer and one logsig node in the output layer. The training session has lasted for 200 epochs and a learning rate of 0.001 has been used. Finally, a set of 3000 training data input values have been used.

Focusing now on this specific NN, the MSE that is measured during the training session, is monotonously decreasing until it reaches a constant, satisfactory goal value (training error). For brevity reasons the respective figure is not depicted here. In the case of the known sequence, the NN produces an MSE = 0.0100, while in the case of the "validation" sequence, the MSE = 0.0153. Figure 4-8 and Figure 4-9 illustrate the performance of the NN in terms of measured MSE, when applied in training and validation sequences, respectively.

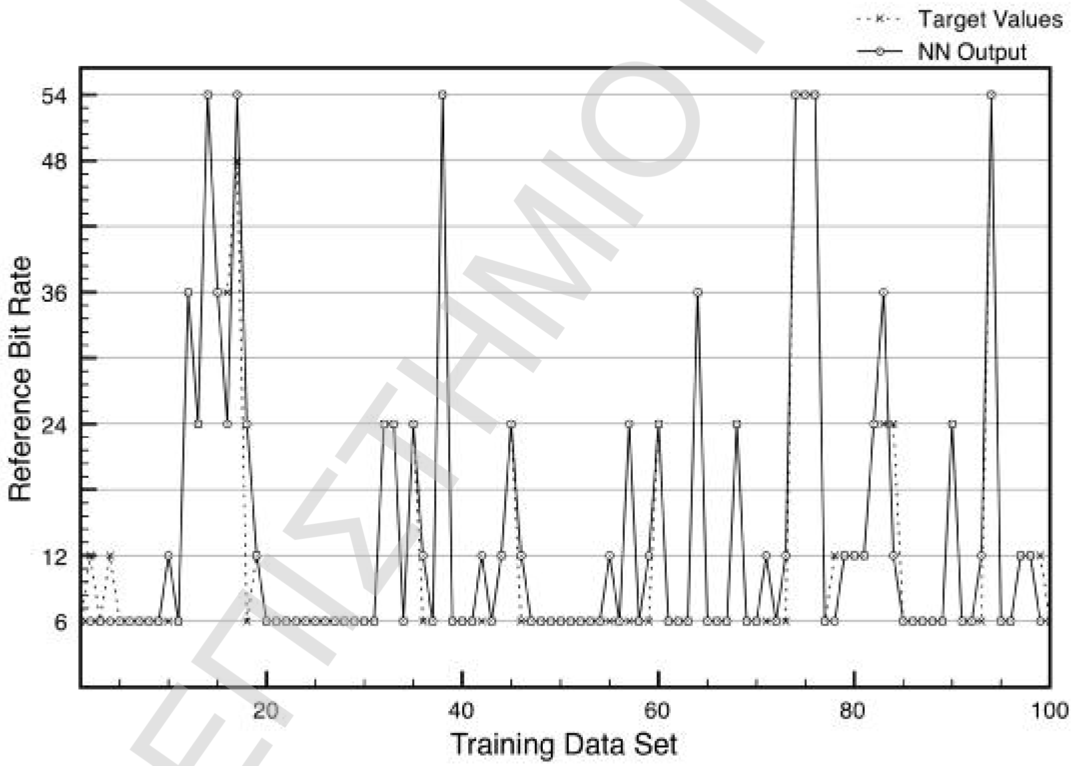


Figure 4-8: Measured performance (MSE) of the selected NN in training sequence – Basic scheme.

As it can be observed, the MSE produced during the validation naturally exceeds slightly the one produced during the training. When fed with the known sequence, the NN actual output seems to follow the target values (that are expected according to the input that feeds the NN), giving very few errors, which shows that the network has been trained well. The same applies for the unknown sequence. The NN performs well during the validation session and it can be observed that the network has learned the basic structure of the data but at the same time it is also able to generalize well.

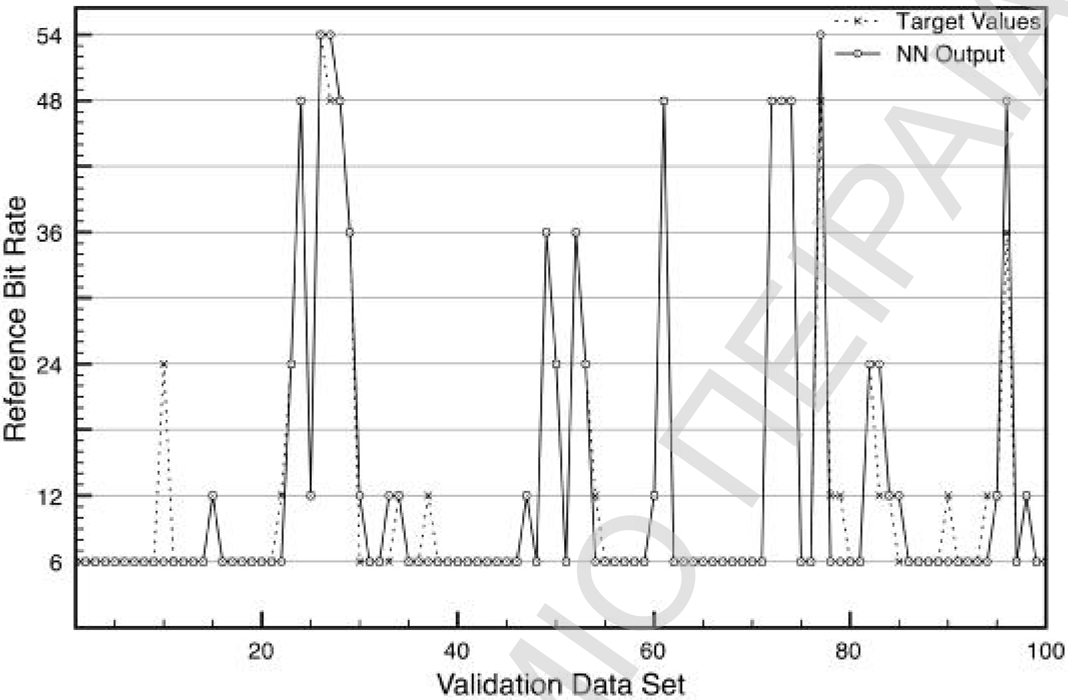


Figure 4-9: Measured performance (MSE) of the selected NN in validation sequence – Basic scheme.

4.5. Extended NN-based learning scheme

4.5.1. Preparation Procedure

For the extended NN scheme, the complexity of the problem is raised by further taking into account and co-estimating a "time zone" parameter. It is assumed that the day is divided into time zones and that during each of them, the configuration in question is associated with a mean, most usually observed data rate value, which is denoted as $\overline{m}_{tz} \in M$. This value is enhancing the proposed learning scheme with a feature of past experience. Let $R^{ext} = \left\{ r_k^{ext} \right\}$, $k \in \mathbb{N}^*$ be the new time-series collected by the radio-scene analysis. As in the basic scheme, a time window of n slots is considered and at any time k, the NN is fed with an input sequence $R^{in,ext} \subseteq R^{ext}$, the length of which equals the size of this time window , i.e., $R^{in} = \left\{ r_i^{ext} \right\}$, $i = 1, \ldots, n$.. For

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

each pair comprising an input time sequence $R^{in,ext}$ of length n and a specific time zone of the day, within which the NN is expected to operate, there corresponds a new target data rate value $r_k^{tgt,ext}$. This value will be used for supervising the training and is calculated at each time k as follows:

Step 1: Temporary target value $r_k^{tgt,tmp}$ is calculated as in (5) by applying $R^{in,ext}$ as input.

Step 2: The distance (absolute difference), dst_k, between the target value at k i.e., $r_k^{tgt,tmp}$ and mean value $\overline{m}_{tz}$, which corresponds to the considered time zone is calculated. Specifically and assuming that M is the set defined in sub-section 4.4.2, if the target value $r_k^{tgt,tmp}$ equals $m_i \in M$, $1 \le i \le |M|$ and the mean value $\overline{m}_{tz}$ equals $m_j \in M$, $1 \le j \le |M|$, then the distance dst_k, is taken equal to $|i-j|$, $1 \le i, j \le |M|$. For example, if $r_k^{tgt,tmp}$ equals $m_1 = 6$ Mbps and the mean value $\overline{m}_{tz}$ equals $m_2 = 12$ Mbps, the distance dst_k, equals $|1-2| = 1$; similarly the distance from $m_1 = 6$ Mbps to $m_3 = 24$ Mbps equals $|1-3| = 2$, and so forth.

Step 3: The new weights b'_i (where $i = 1,\ldots,$n and n is the number of slots in the time window) are recalculated with the use of an exponentially weighted moving average function, with a smoothing factor $a_k = 1 - x^{dst_k}, x \in (0 \mathbf{K} 1)$. For example, assuming that M is the set defined in sub-section 4.4.2 and that n=10, Figure 4-10 illustrates the weight values per time slot when dst_k = 1,…,5. The above rule is used so that greater distance between the calculated value and the target value should lead to a lower slope of weight decrease. A lower slope, as shown in Figure 4-10, gives high importance (high weights) to past observations so as to eliminate the abovementioned distance. The slope of weight decrease is expressed by the smoothing factor a.

Step 4: Finally, $r_k^{tgt,ext}$ is calculated from (5) by using the new weights b'_i and $R^{in,ext}$.

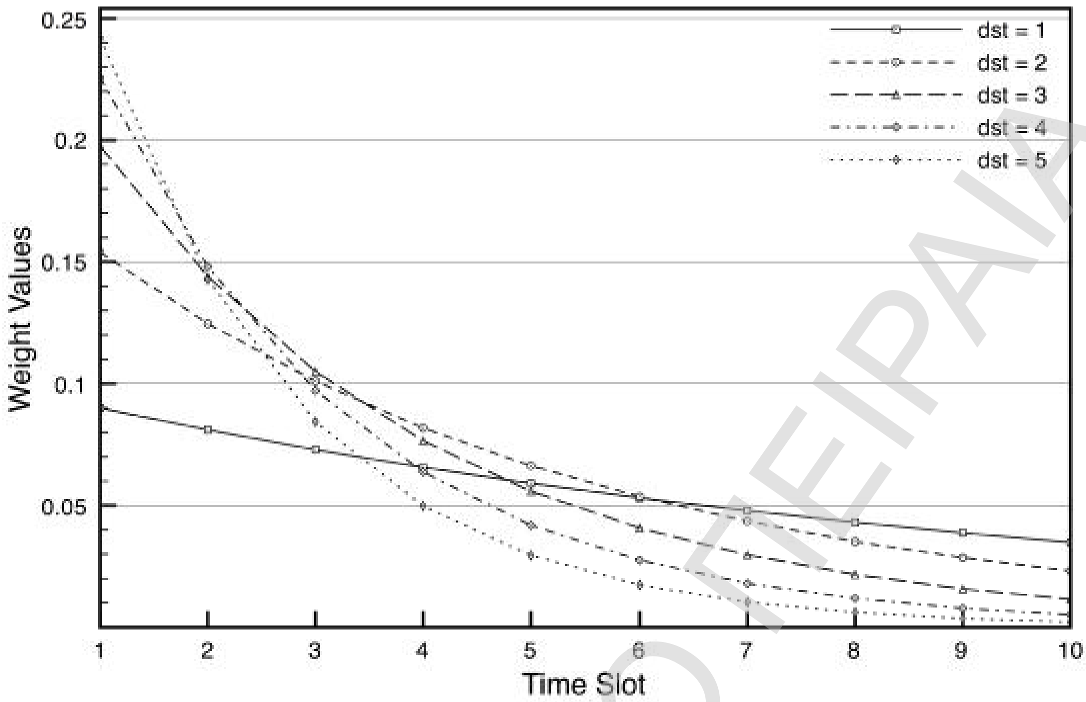


Figure 4-10: Weight values per time slot for the extended scheme.

4.5.2. NN pattern selection – Results

For the selection of the NN design pattern, the focus is, as before, on a specific, arbitrary radio configuration e.g. WLAN 802.11g. It is assumed that the number of reference bit rate values are $|M|$=6. The time window is $n=10$ and x=0.9.

It is also assumed that the day is divided in 4 equal time zones as follows: 06:00 – 12:00, 12:00 – 18:00, 18:00 – 24:00 and 00:00 – 06:00. In each of these time zones a different mean value $\overline{m}_{tz}$ is observed, let them be set equal to 24, 6, 36 and 48 Mbps for each of the four time zones, respectively; this might reveal for instance the existence of high load situation during the mid working day.

R^{ext} includes values from the M set which are randomly generated according to a selected probability distribution function, depicted in Figure 4-5 (dotted line), that assigns bigger probability to the appearance of $\overline{m}_{tz}$ depending on the time zone. The target values $r_k^{tgt,ext}$ are calculated by following the steps mentioned in the previous sub-section. The NN uses the tansig function in equation (2) for the neurons in its hidden (recurrent) layer, and the purelin function in equation (4) for the neuron in its output layer, respectively. For the training session, the input and target values have been properly normalized in the range of [-1, 1].

Table 4-2: Test cases examined for extended scheme.

	Net type	*Training Data Set Size*	*Hidden Neurons*	*Delay*	*Hidden Layers*	*Training Function*	*Epochs*	*Learning Rate*	MSE_{trn}	MSE_{val}	$\lvert MSE_{trn} - MSE_{val}\rvert$
1	FF	600	10	0	1	trainlm	300	0.001	0.1577	0.1699	0.0122
2	FF	600	10	0	1	trainlm	300	0.001	0.1622	0.1794	0.0172
3	FF	600	15	0	1	trainlm	300	0.001	0.1152	0.1400	0.0248
4	FF	600	15	0	1	trainlm	300	0.001	0.1301	0.2362	0.1061
5	Elman	2400	10	10	1	traingdx	300	0.001	0.1640	0.2125	0.0485
6	Elman	2400	10	10	1	traingdx	300	0.001	0.1841	0.2479	0.0638
7	Elman	2400	15	10	1	traingdx	300	0.001	0.1188	0.1406	0.0218
8	Elman	2400	15	10	1	traingdx	300	0.001	0.1293	0.1542	0.0249
9	FTDNN	2400	10	10	1	trainlm	300	0.001	0.1159	0.1525	0.0366
10	FTDNN	2400	10	10	1	trainlm	300	0.001	0.1323	0.2037	0.0714
11	FTDNN	2400	15	10	1	trainlm	300	0.001	0.0928	0.1285	0.0357
12	FTDNN	2400	15	10	1	trainlm	300	0.001	0.0992	0.1677	0.0685
13	Custom	2400	10	10	2	trainbr	300	0.0001	0.0549	0.0637	0.0088
14	Custom	2400	15	10	2	trainbr	300	0.0001	0.0759	0.0956	0.0197
15	Custom	2400	12	10	4	trainbr	300	0.0001	0.0722	0.0902	0.018
16	Custom	2400	20	10	4	trainbr	300	0.0001	0.0560	0.0728	0.0168

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

A number of different cases have been tested to evaluate the extended NN scheme. Table 4-2 gives an overview of the parameters used to define those test cases.

For the first set (test cases 1 – 4), a Feed-Forward back-propagation (FF) NN has been used. In the first two cases the network consisted of 10 hidden nodes in the hidden layer, while in the second two cases it consisted of 15 hidden nodes. The networks have been trained with 600 training samples, which represent a single time zone. The weights and bias values have been updated according to Levenberg-Marquardt optimization (a.k.a. trainlm in [3]) method. The training samples have been processed to provide 10 parallel inputs to the NNs, which represent the time window mentioned previously. Finally, the networks have been trained for 300 epochs at a 0.001 learning rate.

The second set (test cases 5 – 8) consisted of four Elman NNs. The first two networks consisted of 10 neurons in the hidden layer, while in the second two it consisted of 15 neurons. A delay line of ten slots has been inserted in the input layer, which reflects the time window. The traingdx function has been used for updating the weights and biases. For the training session, 2400 sample data points have been used, which per 600 samples reflect the four time zones.

The NN used for the third set (test cases 9 – 12) is the Focused Time-Delay Neural Network (FTDNN). FTDNN is a feed-forward input-delay back-propagation network, which consists of a feed-forward network with a tapped delay line at the input. FTDNN is a network well suited to time-series prediction. The delay line has been set to 10. Again, the first two networks have been configured to have 10 neurons in the hidden layer, while the second two had 15 neurons. The trainlm function has been used for the training. Also, the same 2400 sample data points as in the previous test set have been used.

Finally, the last set (test cases 13-16) uses custom feed-forward back-propagation networks. The networks have been configured to have 2 hidden layers, in the first two test cases. In the second two cases, they have been configured to have 4 hidden layers. In general, using more than one hidden layer is almost never beneficial. The only situation in which a NN with two hidden layers may be required in practice is when the network has to learn a function having discontinuities. In the first test case the 2 hidden layers have been configured to have 5 neurons each, while in the second case

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

they had 5 and 10 neurons, respectively. In the third test case the four hidden layers consisted of 3 neurons each, while in the forth case each hidden layer had 5 neurons. All networks have been configured to have a tapped delay line of 10 slots. The networks have been trained with the use of bayesian regularization back propagation (a.k.a. trainbr function [3]), which is believed to produce networks that generalize well. The training lasted for 300 epochs and the learning rate of 0.0001 has been used. The input has been the same 2400 data samples, which reflect the four different time zones, as mentioned earlier.

The best available network design pattern is the one corresponding to the 13th test case in Table 4-2, since it is the one that produces the MSE which best satisfies criteria similar to those in Section 4.4.2. This case designates a custom feed-forward back-propagation network with 2 hidden layers with 5 tansig nodes each, and a purelin node in the output layer.

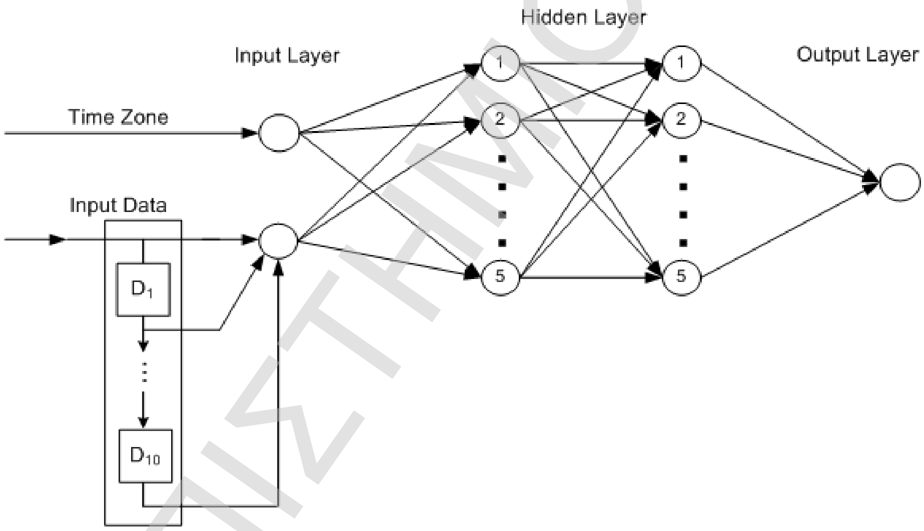


Figure 4-11: Neural network for the Extended Learning scheme.

The training session has lasted for 300 epochs and a learning rate of 0.0001 has been used. Finally, a set of 2400 training data input values have been used with a tapped delay line of 10 slots. In the sequel, the trained extended-NN has been tested in both a known (subset of training set) and an unknown (validation) sequences comprising 100 data points each. In the case of the known sequence, the NN produces an MSE=0.0549, while in the case of the validation sequence, the MSE=0.0637. Figure 4-12 and Figure 4-13 illustrate the performance of the NN in terms of measured MSE.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Again, as in the case of the basic scheme, the MSE produced during the validation naturally exceeds slightly the one produced during the training. The output of the NN during both cases is very close to the target values, which produces a very small error.

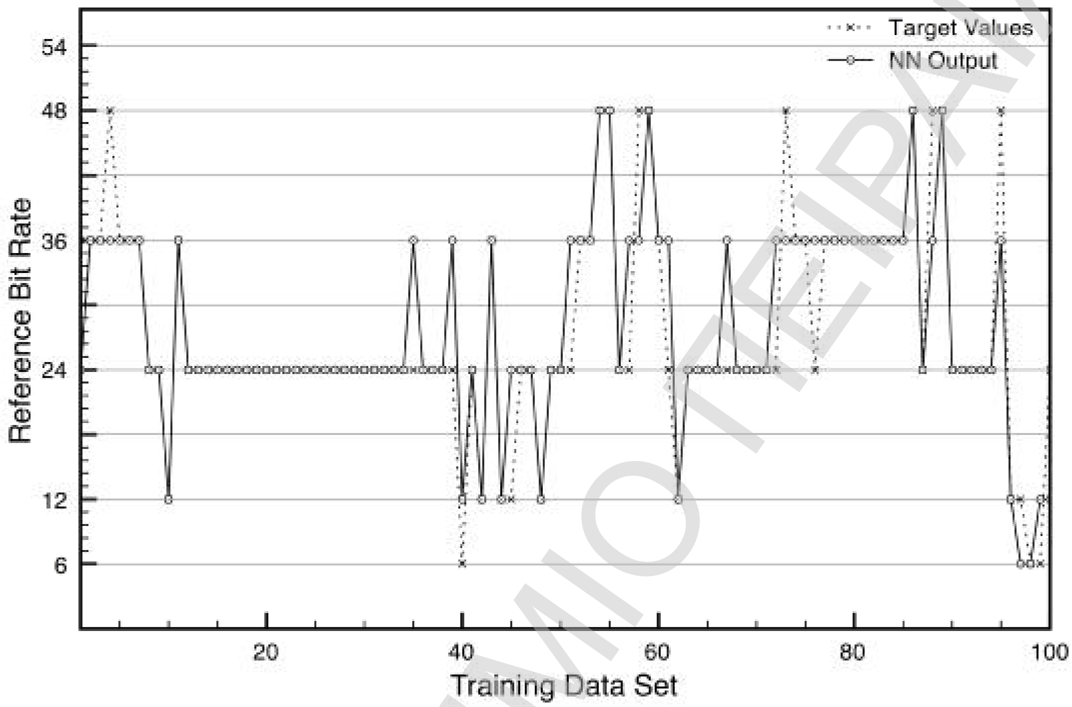


Figure 4-12: Measured performance (MSE) of the selected NN in training sequence – Extended scheme.

Due to the complexity of the problem (multiple time zones), a two hidden-layer network performs better. This observation can be generalized for all cases. The performance of the NN is dramatically increased when the number of hidden layers is increased. This seems logical, since smaller networks don't have the ability to distinguish between the time zones (separate the problem). Conversely, adding more neurons into the two hidden layer network does not raise the performance of the network. Actually, the error increases when more hidden neurons are used. This is normal since there is a theoretically best performance that cannot be exceeded by adding more neurons; the network learns irrelevant details of the individual cases. In general, the proposed NN performs well. It is able to generalize well, giving output values very close to the target values. From the above, the abovementioned NN is selected in the case of the extended scheme.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

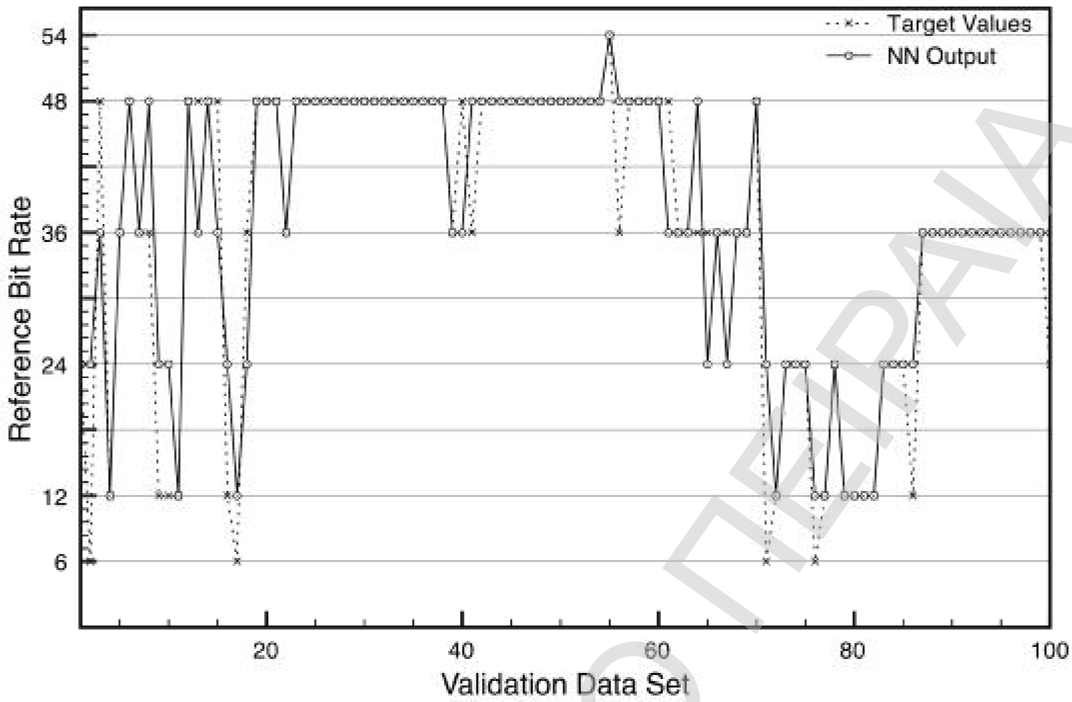


Figure 4-13: Measured performance (MSE) of the selected NN in validation sequence – Extended scheme.

4.6. Discussion and Conclusions

Cognitive radio devices need to efficiently perceive the necessity for alternating their radio configuration, to evaluate the capabilities of each of the candidate, available configurations and thereupon, to dynamically select the one, in which they must operate. To this effect, a potential solution to assist the cognitive radios in the derivation and enforcement of decisions regarding the selection of the desired radio configuration, the one that optimizes its QoS is proposed in this chapter.

The proposed solution is based on neural networks (NNs) motivated by the fact that NNs are widely different from conventional information processing as they have the ability to learn from given examples, thus being also able to perform better in cognitive tasks. Two NN–based learning schemes have been set up and tested: the ‘basic’ and the ‘extended’ one. While the former one aims at building the framework for developing such learning schemes and apply them into future cognitive radio based systems, the latter one stresses that such a learning scheme should be extensible, i.e., flexible in incorporating further information data in the learning process, given that this can bring an objective merit to the process.

The proposed NN-based schemes concern the discovery of the data rate capability of a specific radio configuration. In order for the cognitive radio to be able to select among all available candidate configurations, it should also be able to evaluate them without exception. This could be implemented either by deploying multiple parallel neural networks, (see for example the multi-processor software radio platform in [14]) or by constructing and train a network with an additional input reflecting the configuration, e.g. the access technology. Furthermore, it must be noted that apart from extending the network with the time-zone input, the NN-based learning scheme can be further fed with other information that might crucially affect the achieved data rate of a given configuration, such as location information, user preferences or even weather conditions, etc. This could be subject of future research. Learning is a continuous process during which, the NN's free parameters are adapted according to the external stimuli [12] and thus, more tests and trials, especially taking into account more realistic input time-series and environment situations, are required in order to increase the NN-based scheme's validity and robustness. Moreover, new types of NNs should be deeply explored to achieve better results, but in general the variations can be endless.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

REFERENCES

[1] K. Tsagkaris, A. Katidiotis, P. Demestichas, Neural network-based learning schemes for cognitive radio systems, Computer Communications Volume 31, Issue 14, 5 September 2008, Pages 3394-3404.

[2] Rewagad, A.P.; Soanawane, V.L., “Artificial neural network based short term load forecasting”, TENCON '98. 1998 IEEE Region 10 International Conference on Global Connectivity in Energy, Computer, Communication and Control , vol.2, no., pp.588-595 vol.2, 1998

[3] H. Demuth, M. Beale, M. Hagan, “Matlab Neural Network Toolbox User's Guide, Version 5.1”, *The MathWorks Inc., 2007*

[4] Rosenblatt, Frank (1958), The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Cornell Aeronautical Laboratory, Psychological Review, v65, No. 6, pp. 386-408

[5] Widrow B. and S. D. Sterns, Adaptive Signal Processing, New York: Prentice-Hall 1985

[6] Elman, J.L. (1990). "Finding Structure in Time". Cognitive Science 14: 179-211

[7] Hopfield, J. J.: Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. USA 79, 2554–2558 (1982)

[8] B. Le, T. W. Rondeau, D. Maldonado, C. W. Bostian, "Modulation Identification Using Neural Network for Cognitive Radios," *SDR Forum Technical Conference*, Anaheim, CA, 2005

[9] M. Gandetto, M. Guainazzo C. S. Regazzoni, “Use of Time-Frequency Analysis and Neural Networks for Mode Identification in a Wireless Software-Defined Radio Approach”, Eurasip JASP, Special Issue on Non Linear Signal Processing and Image Processing, Vol. 13, pp. 1778-1790, Oct. 2004.

[10] NIST/SEMATECH e-Handbook of Statistical Methods, available at http://www.itl.nist.gov/div898/handbook/, September, 2002.

[11] T. Masters, *Practical Neural Network Recipes in C++*, Academic Press, 1993

[12] S. Haykin, Neural Networks, *A Comprehensive Foundation*, Second Edition, Upper Saddle River, NJ: Prentice Hall, 1999

[13] B. Hassibi, D. Stork, and G. Wolff, “Optimal Brain Surgeon and general network pruning,” IEEE International Conference on Neural Networks,pp. 293–299, 1993

[14] E. Coersmeier, M. Hoffmann, H. Bothe, “Combining Cognitive Radio and Software Radio Approach for Low Complexity Receiver Architecture”, ISART 2007, Boulder, Colorado, US, 27. Feb. 2007

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

5. PERFORMANCE EVALUATION OF ARTIFICIAL NEURAL NETWORKS BASED LEARNING SCHEMES FOR COGNITIVE RADIO SYSTEMS

Abstract:

This chapter introduces and evaluates learning schemes that are based on artificial neural networks and can be used for discovering the performance (e.g. data rate) that can be achieved by a specific radio configuration in a cognitive radio system. Interesting scenarios, which include both commercial off-the-shelf and simulation hardware/software products, are mobilized for the benchmarking work, conducted in order to design and use an appropriate neural network structure, while indicative results are presented and discussed in order to showcase the benefits of incorporating such learning schemes into cognitive radio systems. Parts of this chapter have been submitted for publication in [1].

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

PERFORMANCE EVALUATION OF ARTIFICIAL NEURAL NETWORKS BASED LEARNING SCHEMES FOR COGNITIVE RADIO SYSTEMS

5.1. Introduction and Problem Statement

As already stated in the previous chapter, the integration of a learning engine can be very important especially for the channel estimation and predictive modelling phase, for improving the stability and reliability of the discovery and evaluation of the configuration capabilities. Many different learning techniques are available and can be used by a cognitive radio ranging from pure lookup tables to arbitrary combinations of Artificial Intelligence (AI) and Machine Learning techniques. Such techniques aim at answering the prominent question, which is also one of the most recognizable definitions for AI, namely "*how to make machines do things at which, at the moment, humans are better*" [2], and include among others: artificial neural networks, evolutionary/genetic algorithms, reinforcement learning, fuzzy systems, hidden Markov models, etc.

This chapter contributes in this direction and complements the previous one, by developing a learning scheme that relies on artificial neural networks and aims at solving the problem related to the channel estimation and predictive modelling phase of cognitive radio systems and can be stated as follows: "*Given a candidate configuration of a cognitive radio system, what are its anticipated performance (e.g., in terms of achievable transmission raw data rate), taking into account recent information sensed, as well as the past experience and knowledge?*".

What is gained is that by associating each configuration with a predictable transmission data rate, the proposed neural network–based (NN-based) learning schemes may efficiently characterize the achievable communication performance with respect to environmental factors and configuration parameters and accordingly, facilitate the cognitive terminal in making its decision regarding the configuration in which it should operate.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

The benchmarking work conducted in order to design and use an appropriate neural network structure is described. Additionally, the practicability and applicability of the proposed learning schemes into cognitive radio systems are supported by results obtained from long- and short- term time scale scenarios, which include both, commercial off-the-shelf and simulation hardware/software products.

The rest of the chapter is structured as follows: Related work and motivation for application of NNs to cognitive radio systems are presented in Section 5.2. Performance evaluation scenarios and results are presented and discussed in Section 5.3. Finally, the chapter is concluded in Section 5.4.

5.2. Related Work and Motivation

Several research efforts that aim at combining various artificial intelligence and machine learning techniques with cognitive radio technology, have been recently recorded in the literature.

A distributed cognitive network access scheme is presented in [3], with the objective to provide the best QoS, with respect to both radio link and core network performance and user application requirements, by using Fuzzy Logic-based techniques. Fuzzy Logic has been also used in [4] for the representation of cross-layer information targeted to the implementation of a cross-layer optimization strategy in cognitive radio networks.

Genetic algorithms is the technique that has been proposed in [5] for developing a biologically inspired model for a cognitive engine (radio architecture and algorithmic mechanisms). Even if such algorithms seem to be well-suited for handling large sets of variables, a major drawback is that they exhibit slow convergence, that prohibit them for usage in real time applications.

In this chapter, the focus is again placed on the channel estimation and predictive modelling phases of the cognition cycle, as in the previous chapter. The proposed NN-based learning scheme should assist in making the optimum decision regarding the radio-configuration settings (mainly PHY and MAC layer) that provide the best QoS for the given problem and user/application needs. It should be noted that QoS optimization is a multi-objective problem that depends on many quality metrics with

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

dependent relationships, including bit error rate, frame error rate, power consumption, latency, data rate etc. The joint optimization of such a big amount of information and parameters is non-trivial even for legacy radios with minimal programmability, while its complexity and computational load increase dramatically in a cognitive radio system where multiple radio technologies are supposed to coexist. Neural networks can outrun such hindrances by what is called black-box modelling, where input is converted to output by hiding the calculation logic from the external world.

Without loss of generality, the focus herewith is given only on one QoS metric: the transmission raw data rate. More specifically, NN-based schemes are proposed which a) are trained with measurements carried out by the cognitive terminal during the radio-scene analysis and b) provide in the output the raw data rate that is most anticipated to be obtained per radio configuration (RAT/frequency).

This work is also related to throughput estimation/prediction and rate adaptation algorithms. Such algorithms have been intensively studied in the literature and some are already deployed in commercial wireless card drivers. In [6] the authors proposed a Multilayered Feedforward Neural Network to be used by a cognitive radio as an effective technique for real-time characterization of the communication performance and eventually to comprise an alternative rate adaptation scheme. Though, that the proposed NN-based schemes may be rather seen as complementary to various rate adaptation schemes in the sense that the objective is not to reproduce the behaviour of a specific rate control algorithm, but rather to be based on its decisions as a means for training the NN to better learn its environment. Particularly in this chapter, apart from focusing only on estimating the communication performance by means of throughput prediction, NN-based schemes are proposed that are supposed to act on top of those algorithms and independently of the various, wireless card driver implementations.

Furthermore, this chapter extends the feasibility study work presented in the previous chapter ([7]), where it is demonstrated that a NN can be a good solution for incorporating learning capabilities in a cognitive cycle but, in the lack of real measurements, the work conducted was rather based on an Exponential Moving Average algorithm (utilised in several modern wireless LAN drivers) to artificially produce training data to act as input to the NNs. Nevertheless the scope remains the same. The proposed NNs are trained with the aim to increase the level of assurance that a certain operating radio configuration will lead to low or high achievable communication performance in both long- and short-term time scales.

5.3. Performance evaluation

In order to derive and evaluate the performance of the most appropriate NN structure that better fulfils the objective, several scenarios and test cases comprising both commercial and also simulated hardware and software have been set up and studied. In all scenarios, multiple, different types of NNs with a considerable number of adjustable parameters have been investigated in a trial and error procedural manner.

The data used for the test cases have been obtained from real measurements that took place in a real working environment within the university premises. Specifically, a laptop equipped with an Intel 3945ABG Wireless card has been used for measuring, among others, the maximum achievable transmission data rate, the link quality and the signal strength. The laptop has been setup with a Debian OS running on a 2.6.18 kernel and using the ipw3945 driver for the wireless card. The wireless access point (AP) used was a Linksys Wireless-G broadband router (model WRT54GS) which was able to operate in both IEEE 802.11 b/g standard modes [8][9]. This will actually comprise the radio configuration (it can be seen as one single configuration given that the operating carrier frequency is the same i.e. 2.4GHz in both modes), the capabilities of which need to be discovered-evaluated. The data collection lasted for 7 days and the services used during that period included peer-to peer (P2P) file sharing, web browsing and ftp.

Some discussion prior to continuing with the scenarios might be useful here. At first, there is no objection that the power of NNs is based on the training they received and consequently, on the availability of the set of exemplars [10], which in some cases is not easy to collect. However, the conduct of the experiments is facilitated by the nature and great availability of the needed measurements to act as input training set for the examined NNs below. On the other hand, it must be noted that there is obviously a speed versus performance trade-off while searching for the best NN structure. Though, in all the scenarios that follow, training and validation were both curried out offline. This relaxes the strict requirements of the online case for fast training and convergence and as a result, no special focus was placed on the

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

optimization of parameters that highly affect the NN’s speed, such as the training set size or the number of training epochs etc.

5.3.1. Scenario 1

For the first set of test cases of scenario 1, the focus is on the maximum achievable transmission data rate from a set of reference values that uniquely characterize each of the operating standard modes. e.g. according to IEEE 802.11g specifications [9] the achievable raw data rates are in the set {1,2,6,9,12,18,24,36,48,54} in *Mbps*. Those values are mixed with the ones from the respective IEEE 802.11b specifications [8] i.e in the set {1,2, 5.5, 11} in *Mbps*. The target is to build a NN that would be able to predict those rates in the next single step, based on past measurements.

Two types of NNs have been selected for the first set of test cases. The first network that has been used is the Focused Time-Delay Neural Network (FTDNN) which is a feed-forward input-delay back-propagation network, and consists of a feed-forward network with a tapped delay line in the input [3]. The second network type used is the Elman network [6], which is a two-layer back-propagation, recurrent network, with the addition of a feedback connection from the output of the unique hidden layer to the input layer. This recurrent connection allows the Elman network to both detect and generate time-varying patterns. Both networks are well suited to time-series prediction.

All networks used during the investigation procedure have been setup with one hidden and one output layer. The NNs use the tansig function in equation (3) for the neurons in their hidden (recurrent) layer, and purelin function in equation (4) for the neuron in their output layer, respectively. For the training session, the input and target values have been properly normalized in the range of [-1, 1], in a pre-processing phase [11]. During training, for the FTDNN networks, the weights and bias values have been updated according to Levenberg-Marquardt optimization (a.k.a. trainlm in [3]), while for the Elman networks they have been updated according to a gradient descent momentum and an adaptive learning rate method (a.k.a. traingdx in [3]). Finally, the Mean Squared Error (MSE) has been used as a metric for measuring the neural network's performance.

In this scenario, three different data sets have been used, which were extracted from the input sequence and served as the target values for teaching and validating the NN: a) a "training" set (seen data) which is used to build the model i.e. determine its parameters, during the so called training session, b) a "validation" set (unseen data) which is used to measure the performance of the network by holding the parameters found during training constant and c) a second validation set (unseen data) for increasing the robustness of the scheme. With the term "unseen" here data that have never been used to update the weights of the network are characterized. The difference between the two validation sets is that the first set includes data that have been measured during the same day as the "training" data set, while the second set includes data from a totally different day. Thereby, the performances of the networks have been tested in cases where they have been trained under different conditions (i.e. day, time of the day, interference conditions, etc.).

The importance of testing the network with both datasets, when searching for the best structure, is significant, since a small error in the training set can be misleading. If the network has not been trained well, it may not learn the basic structure of the data, but rather learn irrelevant details of the individual cases, a.k.a. *overfitting* the training data or *overtraining*. This would lead to a small error during testing with the training set, but in a large error during testing with the validation dataset. In general, performance on the training only tells us that the model learns what it's supposed to learn, but it is not a good indicator of performance on unseen data i.e. whether the NN is able to generalize well or not [12].

Moreover, the number of hidden layers and/or neurons plays a critical role in the learning process and strongly influences the performance of the network. The use of too few hidden neurons would result in a NN that is unable to learn what it is supposed to learn. On the other hand the use of too many hidden neurons would dramatically increase the time needed to learn, without yielding any significant improvement in the performance of the network. This is already captured in this chapter and is referred as overfitting. There exist some valid rules to set the number of hidden nodes [12] but in general, it is better to start with a big net, train, and then carefully follow a pruning strategy for gradually reducing the size of the network

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Table 5-1 gives an overview of the parameters and their values that have been used during the investigation procedure.

Table 5-1: Test cases examined for scenario 1.

Test Case	***Net type***	***Training Data Set Size***	***Hidden Neurons***	***Delay***	***Training Function***	***Epochs***	***Learning Rate***	MSE_{trn}	MSE_{val}	MSE_{val2}
1	FTDNN	50000	5	10	trainlm	100	0.001	0.0338	0.0181	0.0224
2	FTDNN	50000	10	10	trainlm	100	0.001	0.0322	0.0203	0.0226
3	FTDNN	50000	15	10	trainlm	100	0.001	0.0328	0.0195	0.0222
4	FTDNN	50000	30	10	trainlm	100	0.001	0.0262	0.0191	0.0221
5	FTDNN	50000	5	30	trainlm	100	0.001	0.0194	0.0081	0.0106
6	FTDNN	50000	15	30	trainlm	100	0.001	0.0304	0.0201	0.0214
7	Elman	5000	5	0	traingdx	100	0.001	0.1366	0.1254	0.1378
8	Elman	5000	5	10	traingdx	100	0.001	0.0860	0.0471	0.0504
9	Elman	5000	5	20	traingdx	100	0.001	0.0687	0.0332	0.0494
10	Elman	5000	10	0	traingdx	100	0.001	0.0373	0.0214	0.0237

As can be observed on Table 5-1, the six first test cases (1 – 6) include FTDNN networks that have been setup with different number of neurons, ranging from 5 to 30, in their hidden layer and a tapped delay line with a number of slots ranging from 10 to 30. These networks have been trained with a set of 50000 data points, for 100 epochs, with a learning rate of 0.001.

On the other hand, the next four test cases (7 – 10) present Elman networks that have been trained with a set of 5000 data points, for 100 epochs and a learning rate of 0.001. These networks have been setup with 5 and 10 neurons in their hidden layer.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Also, in some cases shown on Table 5-1, a delay line has been added to the input with varying number of slots (10 and 20).

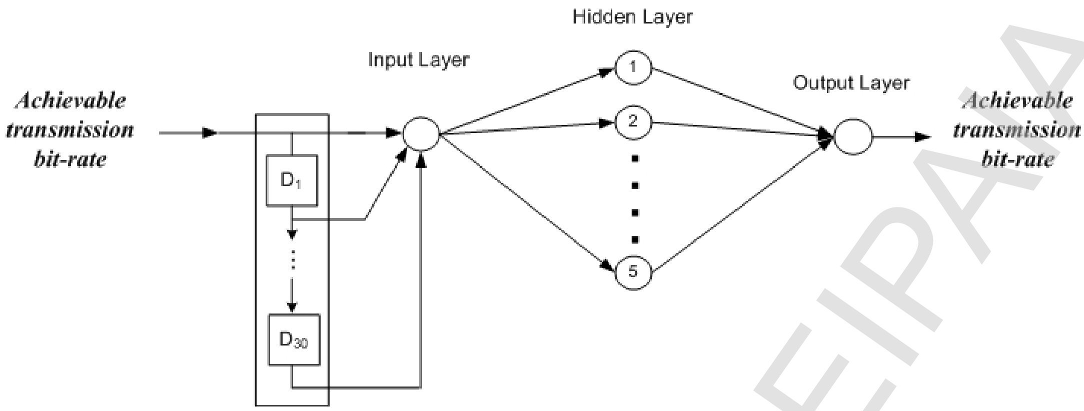


Figure 5-1: Best NN sructure – Scenario 1.

As it can be observed from the table, the best available network design pattern is the one corresponding to the 5th test case, since it is the one that produces the smallest MSE in comparison to the other cases.

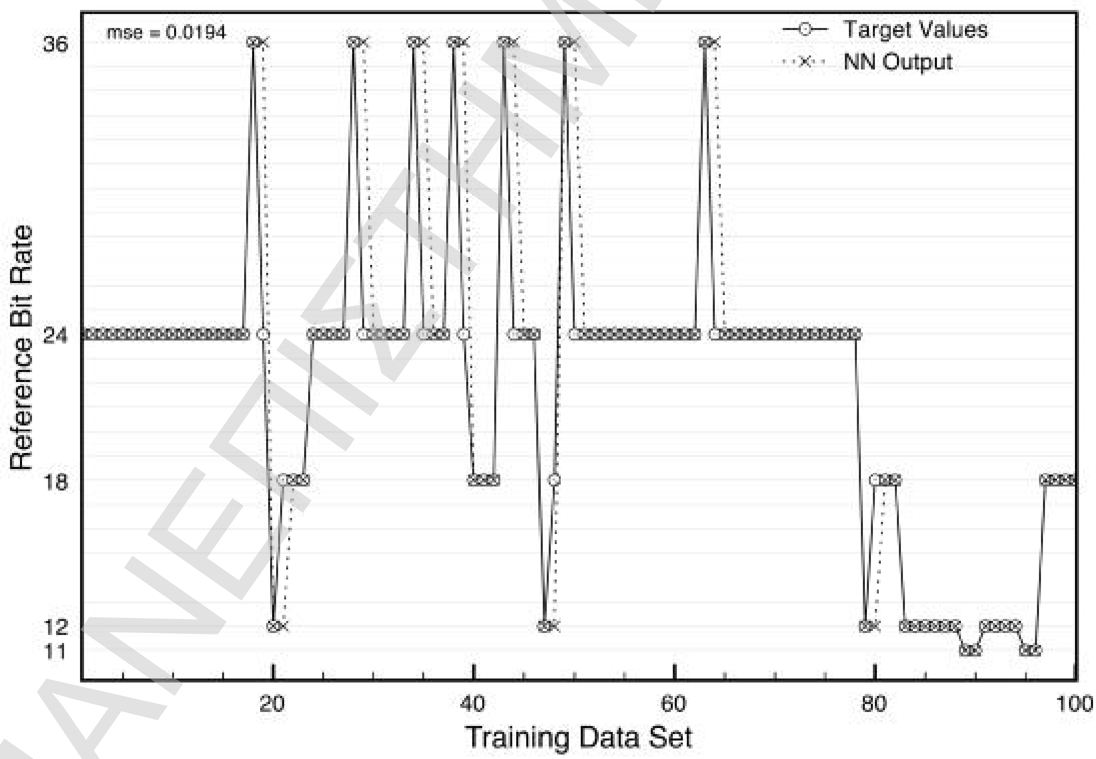


Figure 5-2: Measured performance (MSE) of the selected NN in training sequence – Scenario 1

This case designates a focused time-delay neural network with a hidden layer with 5 tansig nodes, and a purelin node in the output layer (Figure 5-1).

The training session has lasted for 100 epochs and a learning rate of 0.001 has been used. Finally, a set of 50000 training data input values have been used with a tapped delay line of 30 slots. After the training, the NN has been tested in both the known (subset of training set) and the two unknown (validation) sequences comprising 100 data points each. In the case of the known sequence, the NN produces an MSE=0.0194, while in the case of the validation sequences, the MSE=0.0081 and MSE=0.0106 for the first and the second sequence, respectively. It must be noted that the produced MSE is measured from the normalized values of the network's output. Figure 5-2 and Figure 5-3, Figure 5-4 illustrate the performance of the NN in terms of measured MSE, when applied in training and validation sequences, respectively.

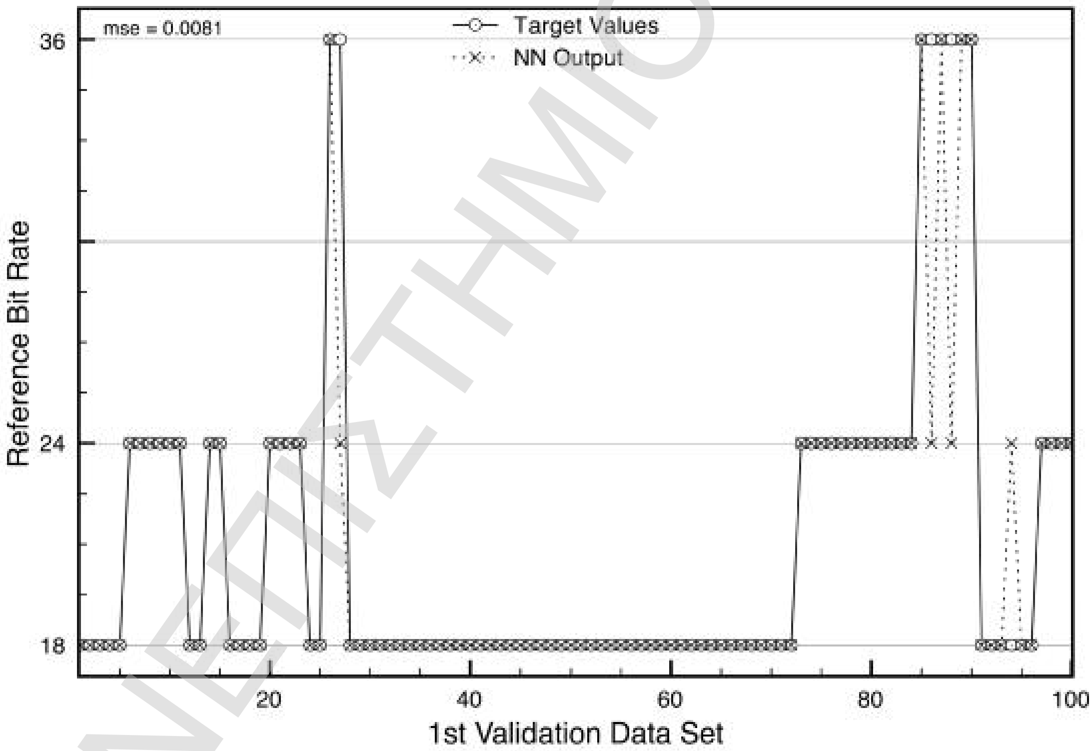


Figure 5-3: Measured performance (MSE) of the selected NN in validation sequence 1 – Scenario 1.

As it can be observed, the MSE produced from the training data set exceeds slightly the one produced during the validation. Though this might look peculiar, the reason for this is that the training set includes data with more inconsistency than in the other cases (more variations between the different reference data rates), which naturally result in a higher error.

When fed with the known sequence, the NN actual output seems to follow the target values (that are expected according to the input that feeds the NN), giving a few errors, which shows that the network has been trained well. The same applies for the unknown sequences. The NN performs well during the validation session and it can be observed that the network has learned the basic structure of the data but at the same time it is also able to generalize well.

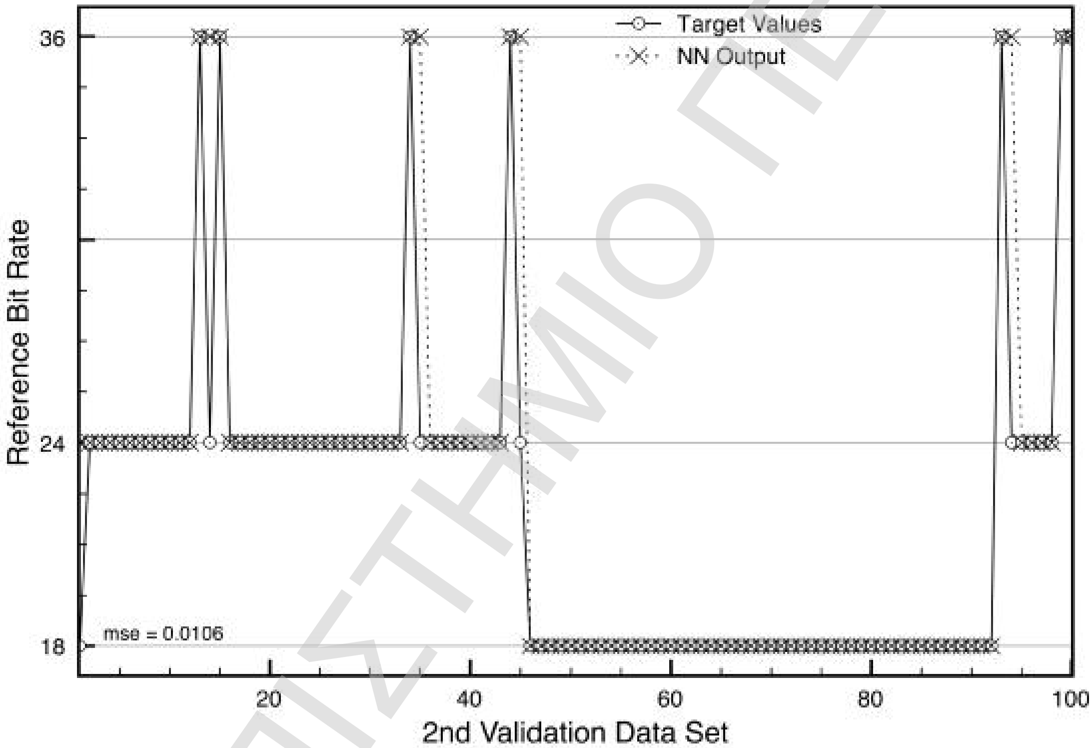


Figure 5-4: Measured performance (MSE) of the selected NN in validation sequence 2 – Scenario 1.

The above leads to the conclusion that the NN has been trained well and performs also well under the specific environment (within the university premises). In other words, the NN has obtained knowledge regarding the behaviour of the environment and it is able to make predictions at a very good level. This scenario reveals the potential of the NNs to handle time series data. The NN has learned to identify patterns and predict the achievable transmission data rate, without knowing any other details (e.g. the link quality, the signal strength (see next scenario), etc.), except for the past

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

observations. This last statement can also justify why a delay line of 30 slots gives better results, compared to the other cases.

5.3.2. Scenario 2

For the test cases of second scenario, the focus is again on the achievable transmission data rate. Though, the target in this scenario is to build a NN that would be able to predict the achievable bit rate, taken as input the quality of the link and the signal strength of the wireless transceiver. For this purpose, the same measurements collected by the wireless card have been used, as in the previous case. The link quality takes arbitrary values in the range of [1 100], while the signal strength is measured in ***dBm***. The input and target data have been properly normalized in the range of [-1 1].

A number of different test cases have been investigated. All tests that have been conducted use the FFTDNN type of neural network, given its predominance against recurrent type (Elman) in most of the situations, as revealed in Scenario 1. Again, all networks use the tansig function for the neurons in their hidden (recurrent) layer(s) and the purelin function for the single neuron in their output layer. The bias and weight values are updated according to trainlm optimization, during training sessions. Finally, once again, the MSE has been used for measuring the performance of the neural networks.

Three different data sets have been used, as in the previous scenario, which were extracted from the input sequence and served as the target values for teaching and validating the NN. These sets include a training data set and two validation data sets, as in the previous case. Table 5-2 gives an overview of the parameters and their values that have been used to define those test cases.

As it can be observed in Table 5-2, the networks used in the five first test cases (1 – 5) have been setup with one hidden layer and trained with a set of 3000 data points, for 100 epochs, with a learning rate of 0.001. The hidden layers include different number of neurons (5, 10, 15) and the tapped delay line also uses different number of slots (10, 30).

The next eight test cases (6 – 13), present networks that have been setup with two hidden layers, with a number of different neurons in each layer ranging from 6 to 20.

Table 5-2. Test cases examined for scenario 2.

	Net type	***Training Data Set Size***	***Hidden Neurons***	***Delay***	***Hidden Layers***	***Training Function***	***Epochs***	***Learning Rate***	MSE_{trn}	MSE_{val}	MSE_{val2}
1	FTDNN	3000	5	10	1	trainlm	100	0.001	0.3252	0.0756	0.0782
2	FTDNN	3000	10	10	1	trainlm	100	0.001	0.5578	0.0864	0.0896
3	FTDNN	3000	15	10	1	trainlm	100	0.001	0.3052	0.0840	0.0874
4	FTDNN	3000	5	30	1	trainlm	100	0.001	0.1273	0.0779	0.0832
5	FTDNN	3000	15	30	1	trainlm	100	0.001	0.1901	0.0903	0.0976
6	FTDNN	5000	6	10	2	trainlm	100	0.001	0.0507	0.0513	0.0714
7	FTDNN	5000	10	10	2	trainlm	100	0.001	0.0492	0.0517	0.0705
8	FTDNN	5000	15	10	2	trainlm	100	0.001	0.0506	0.0522	0.0779
9	FTDNN	5000	20	10	2	trainlm	100	0.001	0.0513	0.0541	0.0751
10	FTDNN	5000	6	30	2	trainlm	100	0.001	0.0510	0.0669	0.0743
11	FTDNN	5000	10	30	2	trainlm	100	0.001	0.0521	0.0626	0.0769
12	FTDNN	5000	15	30	2	trainlm	100	0.001	0.0495	0.0530	0.0695
13	FTDNN	5000	20	30	2	trainlm	100	0.001	0.0560	0.0783	0.0832
14	FTDNN	5000	6	10	2	trainlm	5	0.001	0.0432	0.0534	0.0623
15	FTDNN	5000	6	10	2	trainlm	10	0.001	0.0468	0.0518	0.0521
16	FTDNN	5000	10	10	2	trainlm	5	0.001	0.0489	0.0529	0.0686

The networks have been trained with a set of 5000 data points, with a learning rate of 0.001, for 100 epochs. A tapped delay line has been also used in these test cases with a number of slots ranging from 10 to 30.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Finally, the last three test cases (14 – 16), present networks which have been set up with two hidden layers, with three neurons in each layer for the first two cases (14, 15) and five neurons in each layer for the last case (16). A tapped delay line of 10 slots has been used. These three networks have also been trained with a set of 5000 data points and a learning rate of 0.001, as in the previous test cases, but they were trained for a small number of epochs (5, 10, and 5, respectively).

The best available network design pattern is the one corresponding to the 15th test case in Table 5-2, since it is the one that produces the smallest MSE in comparison with the other cases. This case designates a focused time-delay neural network with two hidden layers with 3 tansig nodes each, and a purelin node in the output layer (Figure 5-5).

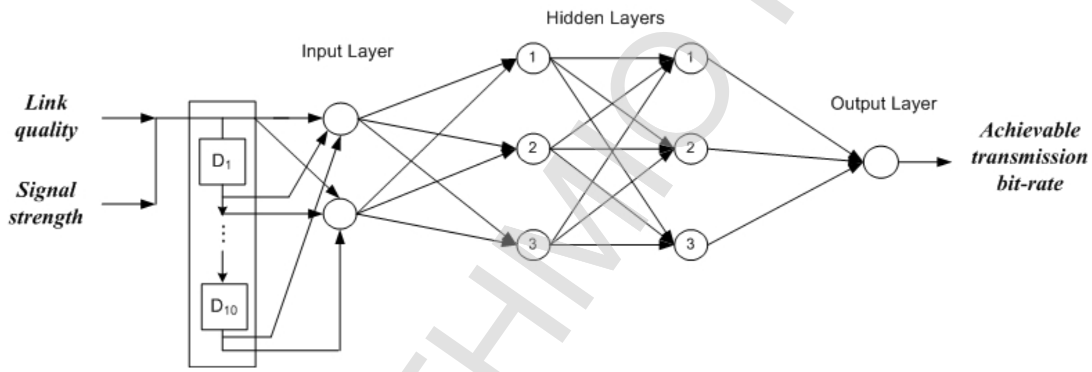


Figure 5-5: Best NN sructure – Scenario 2.

The training session has lasted for 10 epochs and a learning rate of 0.001 has been used. Finally, a set of 5000 training data input values have been used with a tapped delay line of 10 slots. As long as training is finished, the performance of the trained NN has been tested in both the “known” and the two “unknown” sequences comprising 100 data points each. In the case of the known sequence, the NN produces an MSE=0.0468, while in the case of the validation sequence, the MSE=0.0518 and MSE=0.0521 for the first and the second sequence, respectively. Figure 5-6 and Figure 5-7, Figure 5-8 illustrate the performance of the NN in terms of measured MSE.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

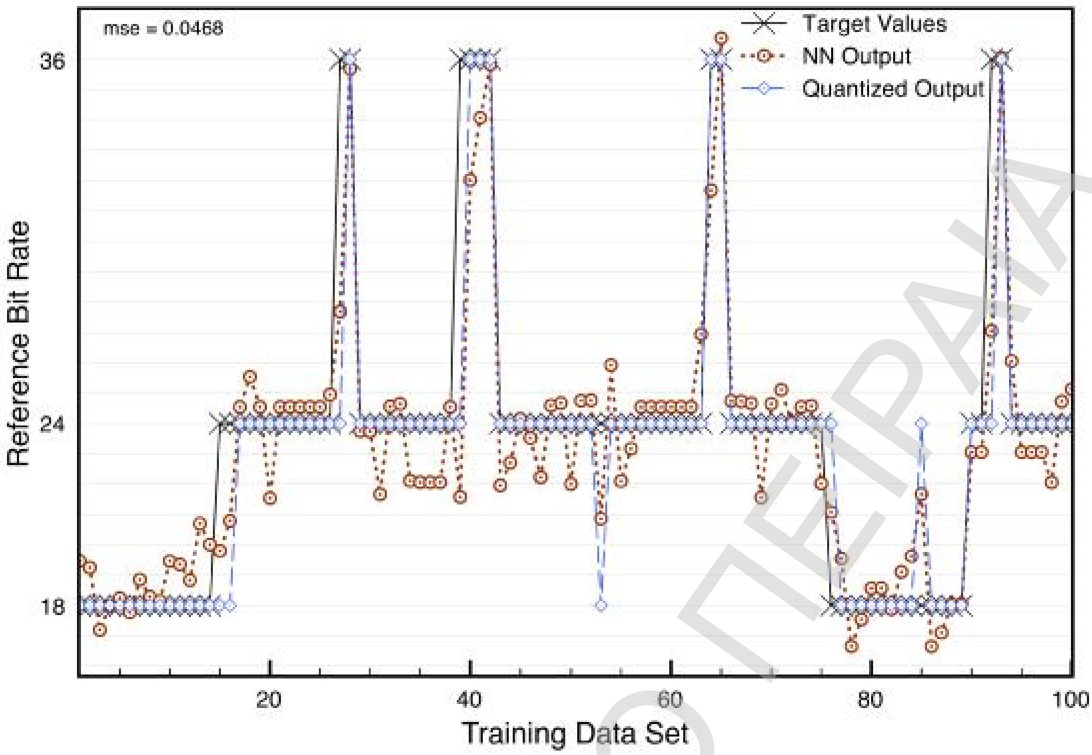


Figure 5-6: Measured performance (MSE) of the selected NN in training sequence – Scenario 2.

In this case, the MSE produced during the validation sets naturally exceeds slightly the one produced during the training, in contrast with scenario 1. The output of the NN during all test cases is very close to the target values, which produces a very small MSE. Due to the complexity of the problem (multiple loose inputs), a two hidden-layer network performs better. This observation can be generalized for all cases. In general, using more than one hidden layer is almost never beneficial. The only situation in which a NN with two hidden layers may be required in practice is when the network has to learn a function having discontinuities [11]. In this case, the performance of the NN is dramatically increased when the number of hidden layers is increased. This seems logical, since smaller networks don't have the ability to distinguish between the different types of input (separate the problem). Conversely, adding more neurons into the two hidden layer network does not raise the performance of the network. Actually, the error increases when more hidden neurons are used. This is normal since there is a theoretically best performance that cannot be exceeded by adding more neurons; the network learns irrelevant details of the individual cases. In general, the proposed NN performs well. It is able to generalize well, giving output values very close to the target values. From the above, the abovementioned NN is selected in the second scenario.

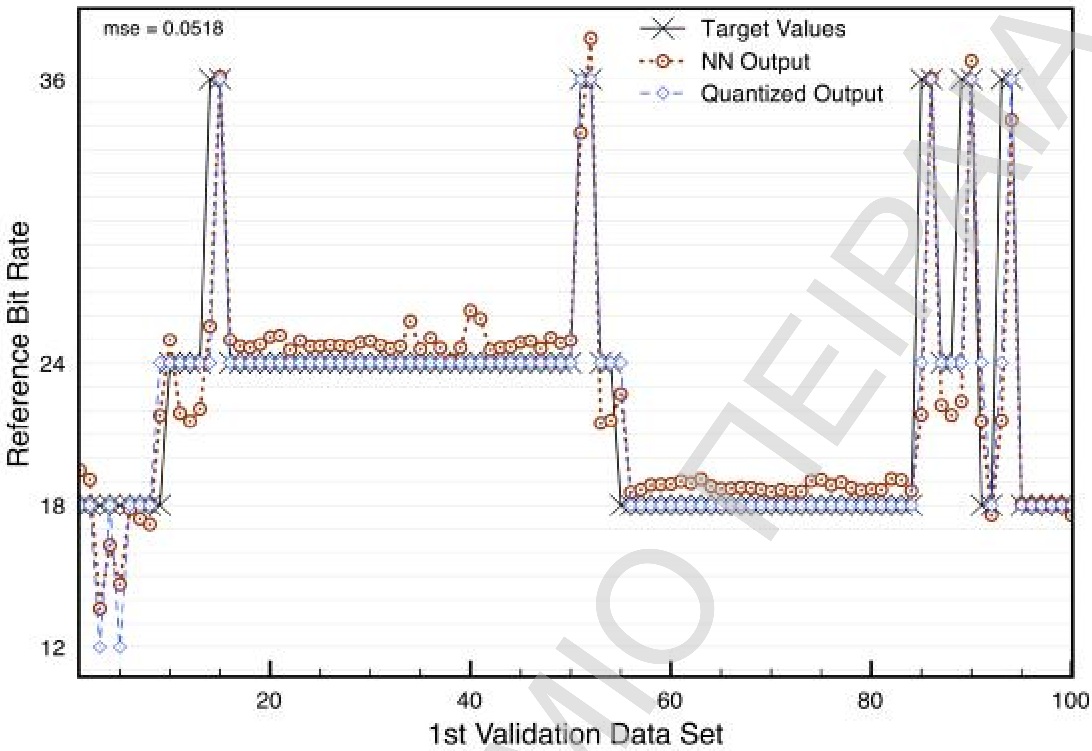


Figure 5-7: Measured performance (MSE) of the selected NN in validation sequence 1 – Scenario 2.

Comparing this scenario to the previous one, this case presents a NN that used to obtain knowledge with respect to two of the characteristics that have an impact on the achievable transmission data rate and accordingly play a major role in the prediction of the achievable transmission data rate, as stated above. The NN was able to predict at a very good level, which shows that it has learned how to associate the link quality and the signal strength with the achievable transmission data rate, in the specific environment. Moreover, this result has been achieved by using less data in the training set than in scenario 1, which can be justified by the increase in the fed inputs and can be really beneficial when a NN is destined for online training during the system operation. Finally, the use of two hidden layers resulted in the improvement of the network performance, that could not be achieved by deploying a larger delay line as in scenario 1.

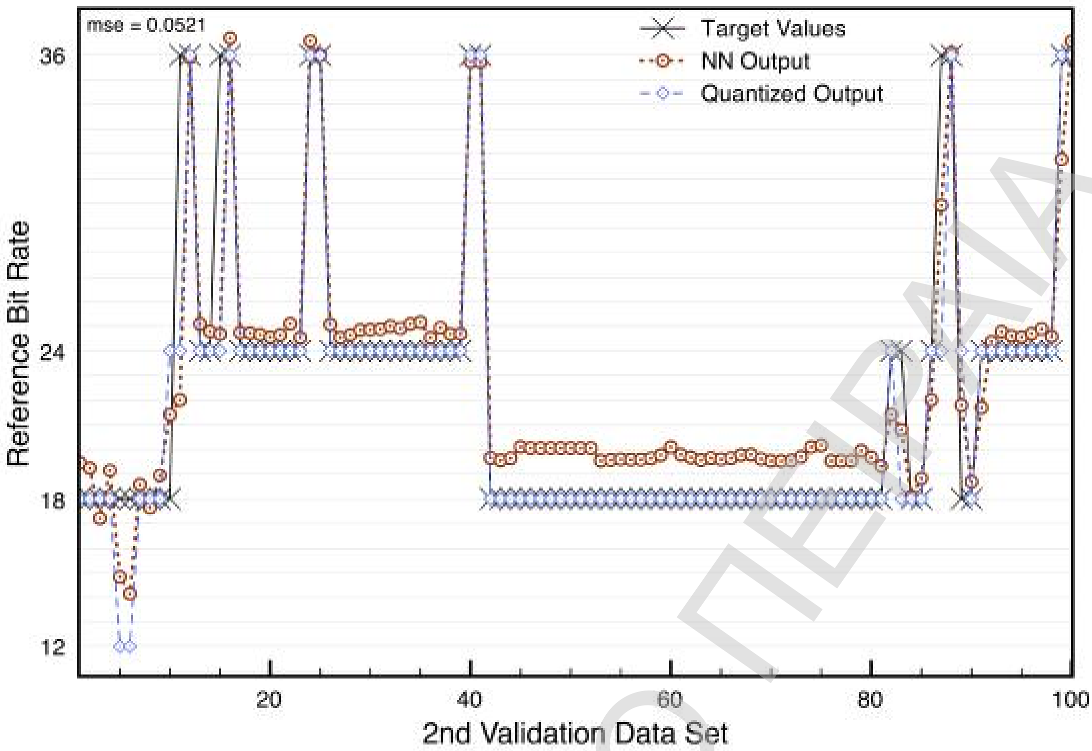


Figure 5-8: Measured performance (MSE) of the selected NN in validation sequence 2 – Scenario 2.

5.3.3. Scenario 3

The target of the previous scenarios was to build a NN that would be able a) to characterize the environment based on measurements that have been recorded for a long period of time and b) to make predictions. For that, the measurements lasted for one week, as already mentioned and a large number of data have been used to train the networks to predict following communication performance. In a real life example, such networks could be used on situations that the user is working in a specific environment, with more or less stable conditions, where the training could last longer and capture all the changes in the conditions of the environment. In such a case, the NN would be able to perform well, giving predictions close to the expected values, as attested in the previous cases.

In the third scenario, those cases that the conditions of the environment are rapidly changing are captured, and the focus is to build a NN that would be able to predict the actual achieved throughput in a short-term fashion. For that, a new set of measurements have been conducted in an environment with fast changing conditions.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

The data used for the test cases have been obtained from the same real working environment. The terminal used for the data collection was a laptop equipped with an Apple airport Wireless card running on a Mac OS X, v.10.5.4 using the airport driver for the wireless card. The wireless access point (AP) used was a Linksys Wireless-G broadband router (model WRT54GS), which has been setup to use the IEEE 802.11g protocol. Eight sets of data have been collected that represent different conditions. In the first two sets, there is no interferer in the transmission channel, in the second two there is only one interferer (a laptop operating at the same frequency band), in the third two there are two interferers and in the last two the interferers are three. In all cases the user of the equipment that collects the data is moving away from the AP and returns after a few minutes. The applications that have been used from all the equipment include ftp and web browsing. The data that have been collected through the wireless card include, among others, the maximum achievable transmission data rate measured in *Mbps*, the Received Signal Strength Indication (RSSI), measured in *dBm* and the number of bytes transmitted.

As already mentioned, in this scenario, the focus is on the throughput prediction. The target is to build a NN that would be able to predict the achieved throughput by taking as input the maximum achievable transmission data rate and the RSSI level. Two sub-scenarios have been distinguished. The first one extracts the throughput from the aggregate bytes that have been transmitted in each second of the data measurements, while in the second the throughput is extracted form the aggregate bytes transmitted in a specific time window, divided by the size of this window (the time window is set to 10secs). In other words, with the first sub-scenario the target was to study the performance of the NNs to predict the instantaneous expected throughput, while with the second sub-scenario, the performance of the NNs has been studied in the prediction of an average expected throughput, in a specific, yet short, period of time.

After a series of testing with different types of NNs (including Elman networks that have been defined in the scenario 1, linear networks and feed-forward networks), a conclusion that have been extracted is that the FTDNN type of networks performs better in all circumstances. For brevity reasons the respective tests are not depicted here. Focusing on that specific type of NN, four different NNs have been

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

distinguished to present here and analyse their performance. The characteristics of those networks are presented on Table 5-3.

The first two networks, on Table 5-3, have been setup with one hidden layer and one output layer, while the next two have been setup with two hidden layers and one output layer. All networks are using the tansig function for the neurons in their hidden layer and purelin function for the single neuron in the output layer. As can be shown on Table 5-3, the first networks have been setup with five neurons in the hidden layer, the second with 15 neurons, the third with six neurons (three in each hidden layer) and the fourth with 10 neurons (five in each layer). A delay line of ten slots has been added to the input of all networks. During training, the weights and bias values have been updated according to Levenberg-Marquardt optimization. Because of the small number of data points collected for each of the data sets, as mentioned earlier, the networks have been trained for only 20 epochs with a learning rate of 0.001.

Table 5-3. Characteristics of the Networks used in scenario 3.

	Net type	*Hidden Neurons*	*Delay*	*Hidden Layers*	*Training Function*	*Epochs*	*Learning Rate*
Net 1	FTDNN	5	10	1	trainlm	20	0.001
Net 2	FTDNN	15	10	1	trainlm	20	0.001
Net 3	FTDNN	3, 3	10	2	trainlm	20	0.001
Net 4	FTDNN	5, 5	10	2	trainlm	20	0.001

In order to test the performance of the abovementioned networks, in different conditions (i.e. move, number of interferers), a series of tests have been conducted with the use of the eight data sets that have been collected.

Table 5-4 gives an overview of those test cases and the parameters used. In test case 1 for example, the networks have been trained using the data set 1, which includes 350 data points, they have been validated using data set 2, which includes 200 data points (both data sets have been collected when there were no external interferers) and they have been also validated using data set 4, which includes 250 that have been collected

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

under different conditions (1 external interferer). The same applies for the other test cases. The training set in case 6 uses the union of data points of sets 3, 5 and 7 i.e. $3 \cup 5 \cup 7$ (700 data points). Moreover, in order to achieve a more fair comparison with the other test cases in terms of training set size, case 5 uses a subset of the union above i.e. $\subseteq \{3 \cup 5 \cup 7\}$ (235 data points) as a training set. In both cases, the networks on Table 5-3 have been validated using data sets 8 and 4, which include 200 and 250 data points, respectively.

Table 5-4. Test cases examined for scenario 3

Test Case	*Train. Set*	*Data Points*	*Valid. Set*	*Data Points*	*Valid. Set 2*	*Data Points*	*Interferers*
1	1	350	2	200	4	250	0
2	3	200	4	250	6	250	1
3	5	250	6	250	8	200	2
4	7	250	8	200	6	250	3
5	$\subseteq \{3 \cup 5 \cup 7\}$	235	8	200	4	250	mixed
6	$3 \cup 5 \cup 7$	700	8	200	4	250	mixed

The six test cases in Table 5-4 have been used for each one of the networks presented in Table 5-3 (four in total) and for each of the two sub-scenarios that were mentioned earlier. That gives a total of 48 different cases (all arising combinations).

Figure 5-9 and Figure 5-10 depicts the resulting MSE recorded at the end of both the training and validation sessions for all the above 48 cases (sub-scenarios 1 and 2, respectively).

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

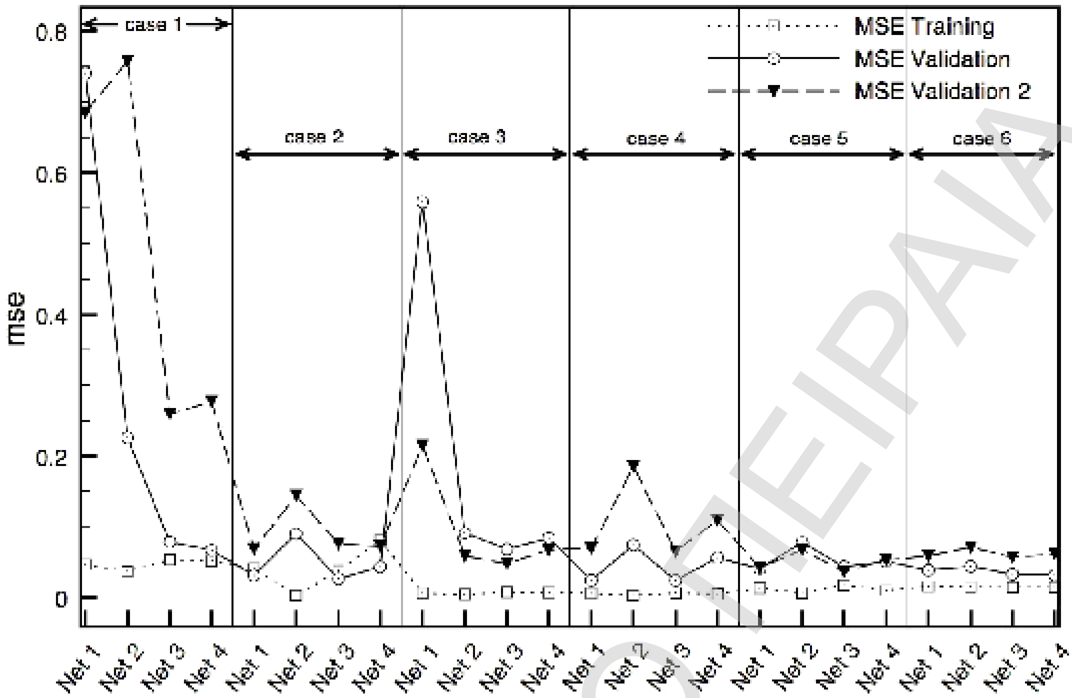


Figure 5-9: Measured performance (MSE) at training and both validation sequences for all 20 test cases – Sub-scenario 1.

Focusing on Figure 5-9, which presents the sub-scenario 2, where the expected throughput (aggregate bytes that have been transmitted in each second of the data measurements) is used as a target, various useful outcomes can be deduced. As can be observed on the figure, for the test case 1, although that the networks perform well on the training data set by producing a small MSE, which indicates that they have learned their lesson, they fail to perform well and predict the expected throughput in the cases of the validation sets. This reflects that the networks have not learned the basic structure of the data but rather irrelevant details of the individual cases.

Checking the same test case at sub-scenario 2 (Figure 5-10), the same outcome can be extracted. The networks perform well when fed with the training data set but fail with the validation sets. A possible reason for this kind of behaviour, except the one mentioned earlier, might be the data set itself. There are cases were the data that have been collected may not present the actual behaviour of a system, because of imponderable elements that might influence the data, for a short period of time. This might be the reason why all networks do not perform well when trained with data set 1, as shown on the two figures, in contrast with all the other test cases.

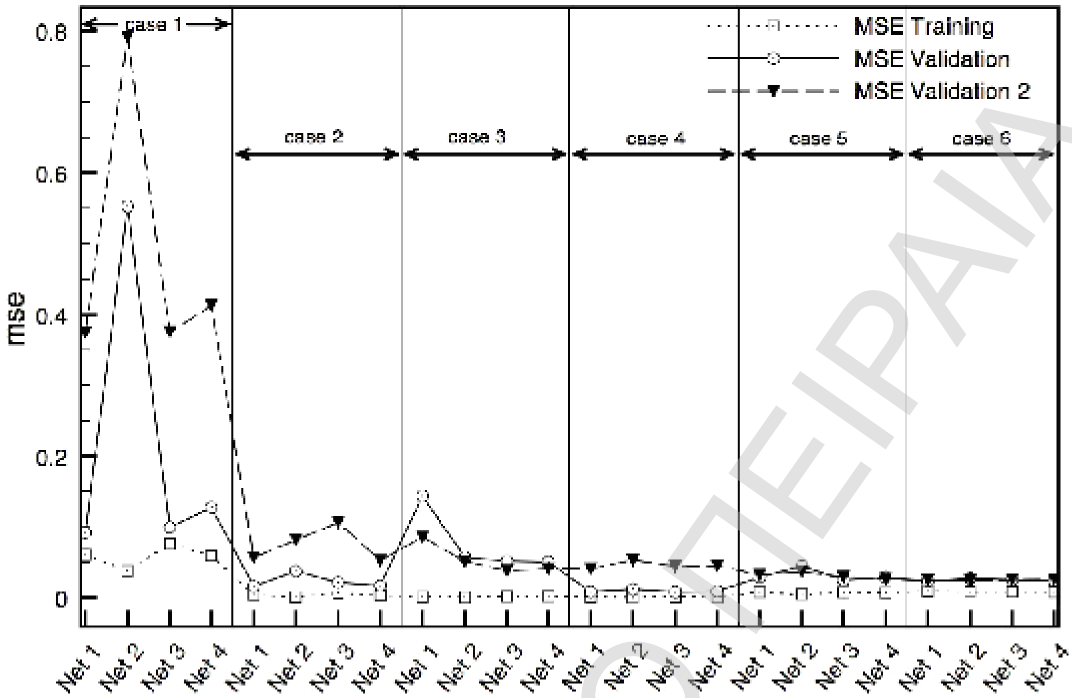


Figure 5-10: Measured performance (MSE) at training and both validation sequences for all 20 test cases – Sub-scenario 2.

Focusing again on Figure 5-9, it can be observed that the MSE produced during validation set 2, slightly exceeds the one produced during validation set 1, on most cases. This is normal since validation set 2 presents data sets that have been collected under different conditions. The same applies on Figure 5-10, as well.

By comparing the two figures, it is easy to observe that the networks trained with the data sets of sub-scenario 2 (Figure 5-10), perform better by producing a smaller MSE in all cases. This reflects that the NNs are more capable to predict the upcoming mean values of the throughput, for a specific time window, than predicting the immediate next one. In a real life scenario, this functionality could enhance a terminal device to predict the expected throughput, for example, in the next few minutes, for each of the available configurations, helping it determine which configuration best suites to the specific service demand.

Focusing again on those figures, it is difficult to determine a specific network that best suites to each sub-scenario. In general, all networks seem to perform well and close to each other. Especially, when compared to cases 5 and 6, where the networks have been trained with a mix of different data sets, all four networks produce an MSE close to each other. This is a general "expected" observation that can be verified by comparing the different six cases on both figures.

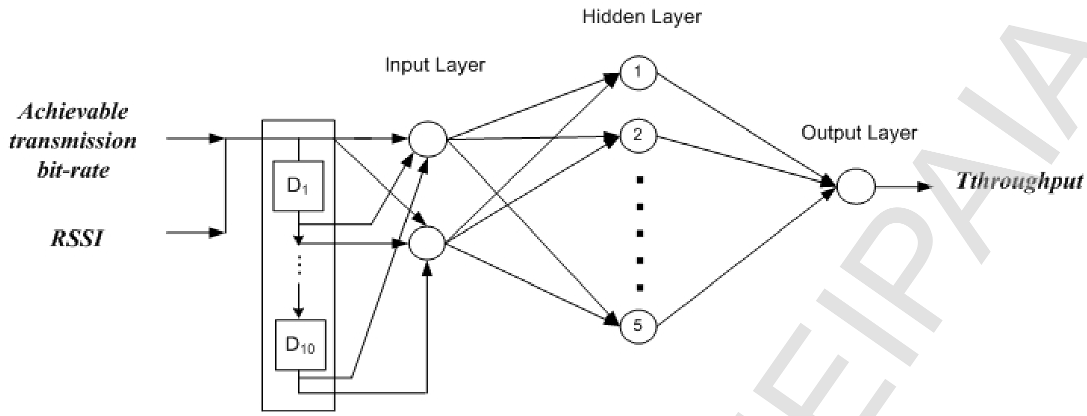


Figure 5-11: Best NN sructure – Scenario 3.

By comparing the MSE produced by all cases, the best available network design pattern is the one corresponding to Net 1 on Table 5-3, when trained with the data set presented on test case 6 of Table 5-4.

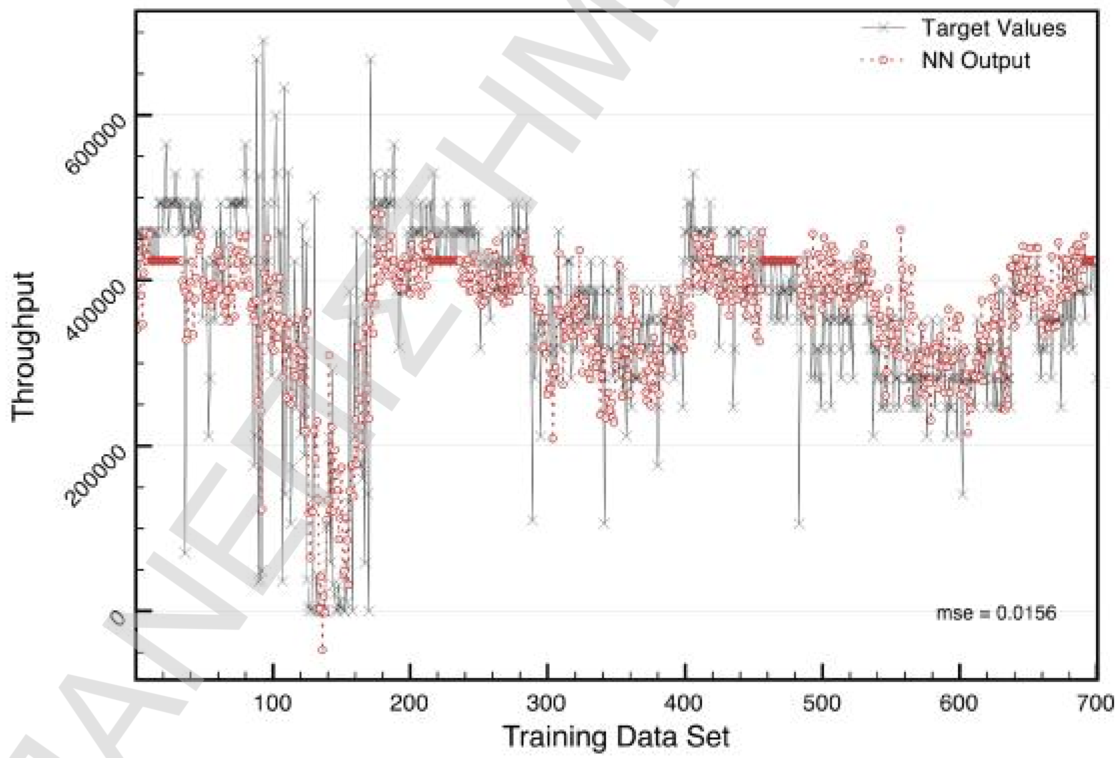


Figure 5-12: Measured performance (MSE) of the selected NN in training sequence – Scenario 3, Sub-scenario 1.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

This network design has been chosen, since it is the one that produces the smallest MSE and also uses the smallest number of neurons. This case designates a focused time-delay neural network with one hidden layers with 5 tansig neurons, and a purelin neuron in the output layer (Figure 5-11).

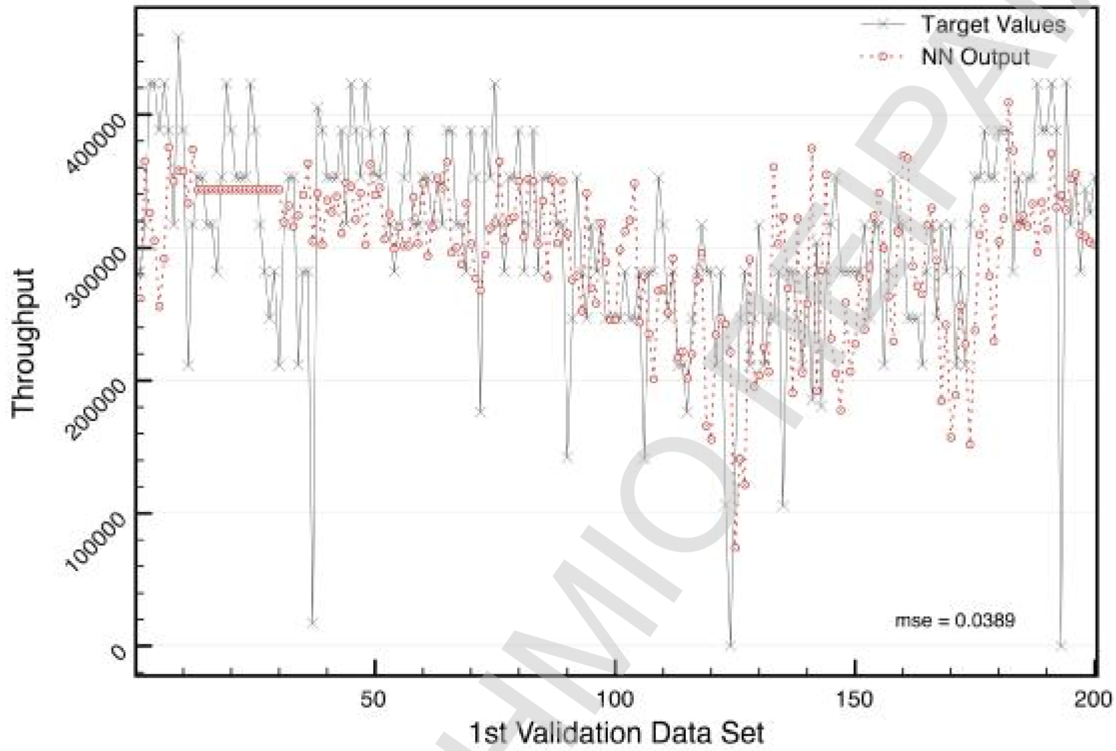


Figure 5-13: Measured performance (MSE) of the selected NN in validation sequence 1 – Scenario 3, Sub-scenario 1.

The training session has lasted for 20 epochs and a learning rate of 0.001 has been used. The set of the 700 data points which is extracted from the union of the sets 3, 5 and 7 has been used as the training data input. Finally, a tapped delay line of 10 slots has been used. As long as training is finished, the performance of the trained NN has been tested in both the “known” and the two “unknown” sequences comprising 700 data points for the “known” set and 200 and 250 data points for the “unknown” sets, respectively.

Figure 5-12 and Figure 5-13, Figure 5-14 illustrate the performance of the NN, in terms of measured MSE for sub-scenario 1, when applied in training and validation sequences, respectively. In the case of the known sequence, the NN produces an MSE=0.0156, while in the case of the validation sequence, the MSE=0.0389 and MSE=0.0588 for the first and the second sequence, respectively. As can be observed, though that the output from the NN has disparities with the expected values, in general it follows the slope of the target values.

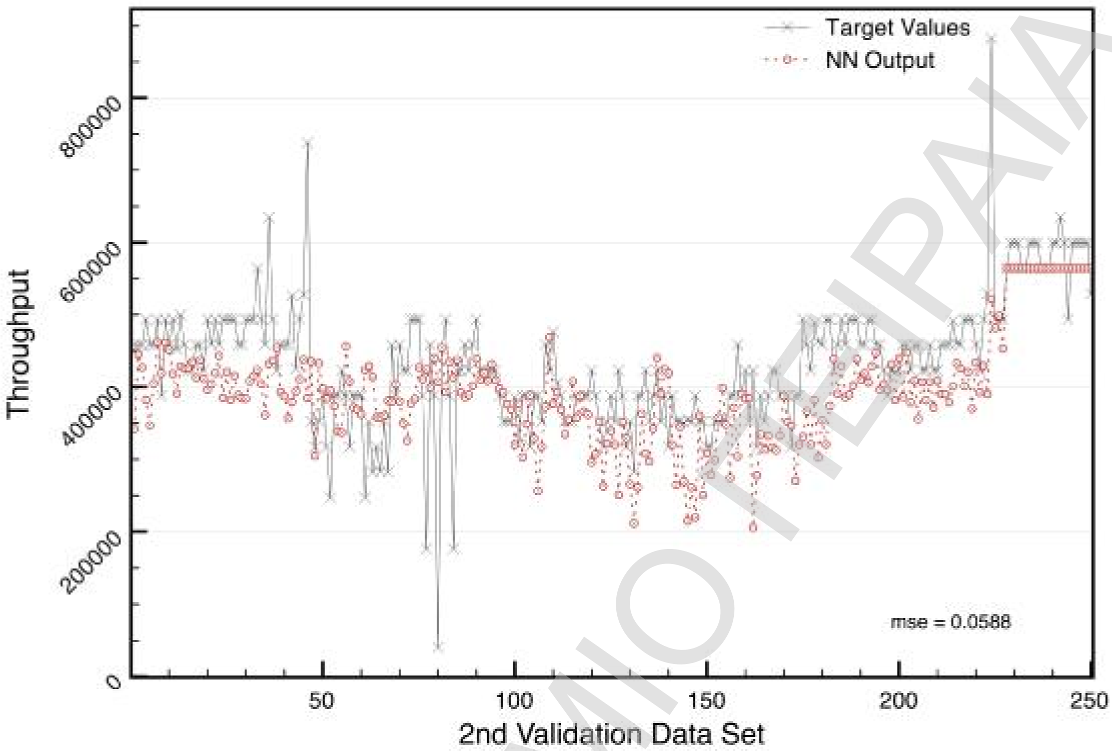


Figure 5-14: Measured performance (MSE) of the selected NN in validation sequence 2 – Scenario 3, Sub-scenario 1.

Similarly, Figure 5-15 and Figure 5-16, Figure 5-17 illustrate the performance of the same NN, in terms of measured MSE for sub-scenario 2, when applied in training and validation sequences, respectively. In the case of the known sequence, the NN produces an MSE=0.0092, while in the case of the validation sequence, the MSE=0.0232 and MSE=0.0246 for the first and the second sequence, respectively. The difference with the previous sub-scenario is huge. When the network is trained with data sets that comprises the mean expected throughput in a specific period of time, it is able to predict, with a very small error, the actual expected throughput.

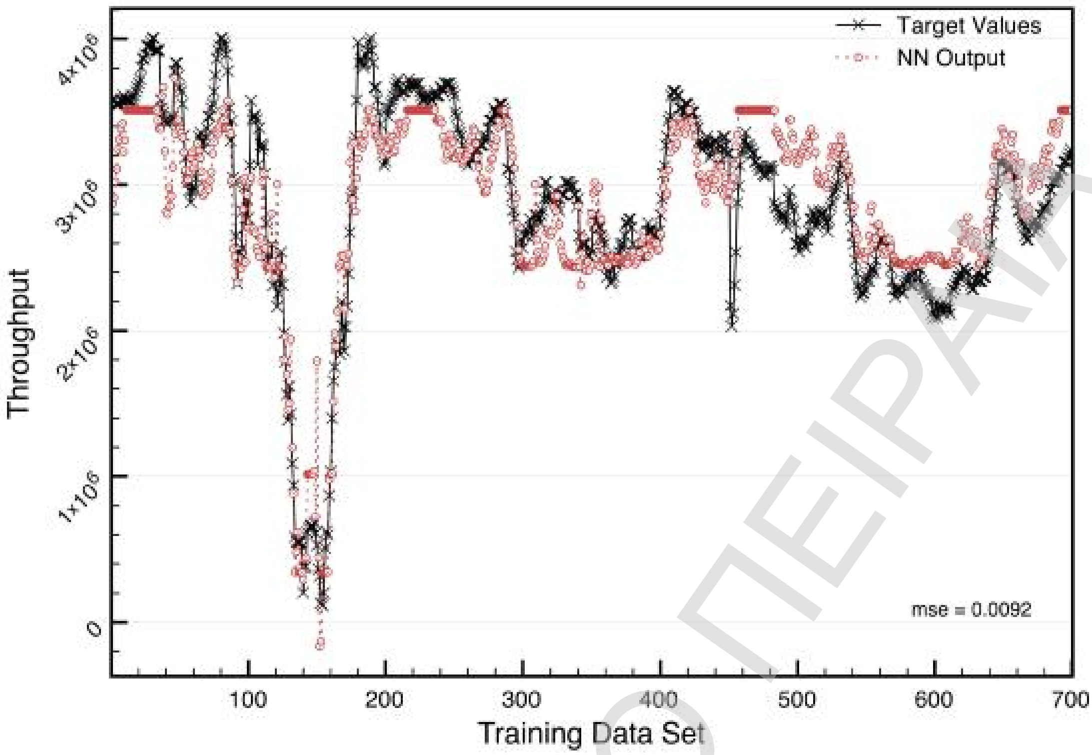


Figure 5-15: Measured performance (MSE) of the selected NN in training sequence – Scenario 3, Sub-scenario 2.

What can be educed from this scenario is that the produced NN was able to learn and perform well in the prediction of the actual expected throughput in a shorter time manner compared to the previous scenarios. Here, the NN is capable to predict with a very small error the expected throughput, based on two characteristics that influence the performance of a link. The NN gained knowledge on how to associate the maximum achievable transmission data rate and the RSSI level, with the actual expected throughput. The tests that have been conducted included cases that covered a number of different environmental conditions. The results obtained from those test cases show that the NN is able to perform well in fast changing conditions. This scenario shows that a NN could help raise the confidence on the expected performance of a link, giving more accurate and correct decisions on the selection of the appropriate configuration.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

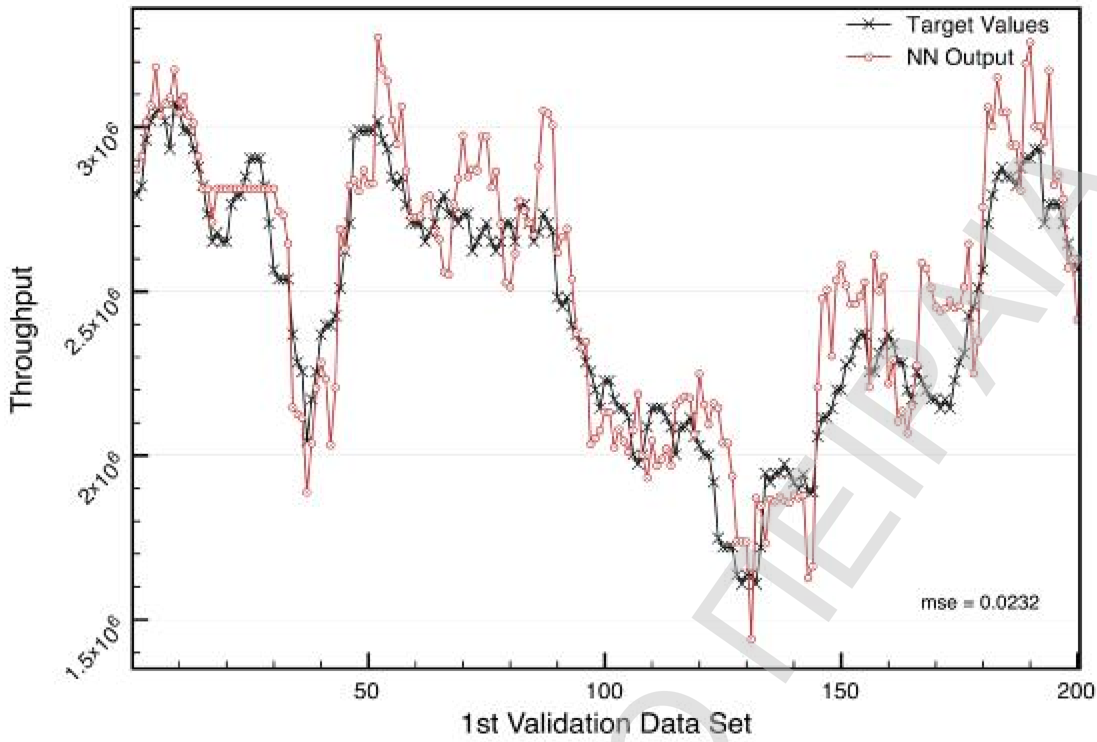


Figure 5-16: Measured performance (MSE) of the selected NN in validation sequence 1 – Scenario 3, Sub-scenario 2.

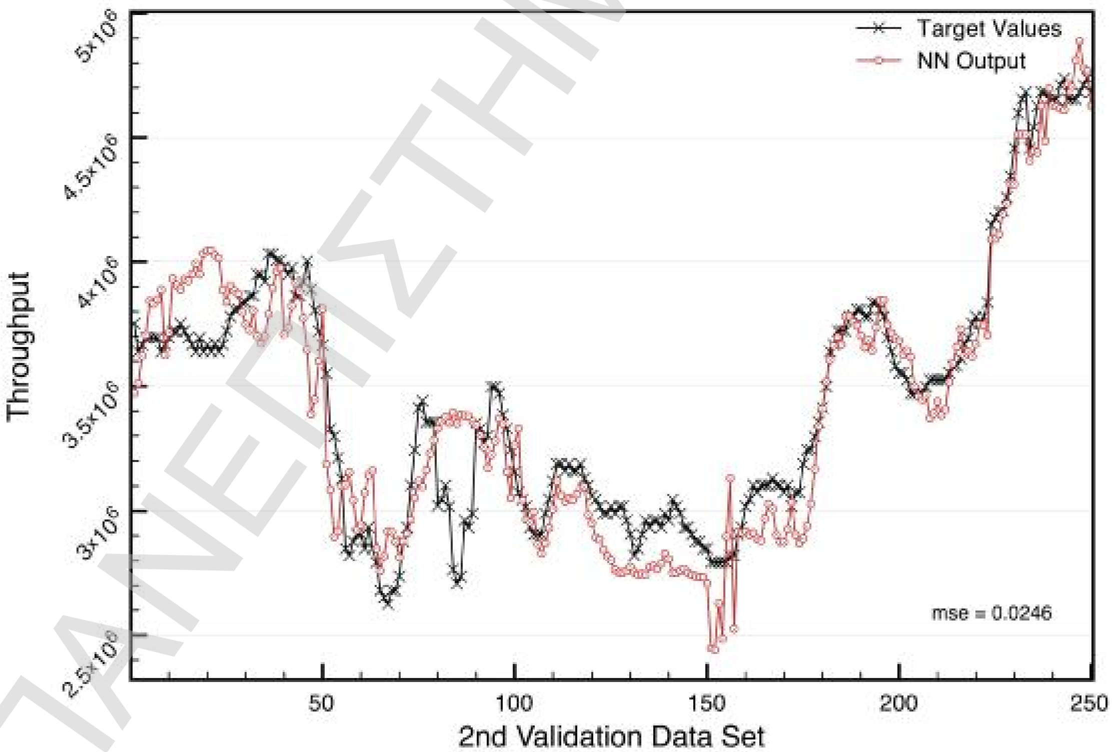


Figure 5-17: Measured performance (MSE) of the selected NN in validation sequence 2 – Scenario 3, Sub-scenario 2.

5.4. Conclusions and Future work

This chapter proposed learning schemes which are based on artificial neural networks (NNs) motivated by the fact that NNs are widely different from conventional information processing as they have the ability to learn from given examples, thus being also able to perform better in cognitive tasks. Scenarios and test cases used for the derivation of the appropriate NN structures are analytically described, while indicative results for this analysis are presented and discussed in order to showcase that the proposed schemes can be used to efficiently associate a predictable, achievable performance with a set of environmental factors and configuration parameters of a cognitive radio.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

REFERENCES

[1] A. Katidiotis, K. Tsagkaris, P. Demestichas, “Performance Evaluation of Artificial Neural Networks Based Learning Schemes for Cognitive Radio Systems”, paper is submitted to International Journal of Communication Systems, Wiley

[2] K. Knight and E. Rich, *Artificial Intelligence*. McGraw-Hill, 1994.

[3] M. Baldo, M. Zorzi, "Cognitive Network Access using Fuzzy Decision Making," *Communications, 2007. ICC '07. IEEE International Conference on* , vol., no., pp.6504-6510, 24-28 June 2007

[4] Baldo, Nicola; Zorzi, Michele, "Fuzzy Logic for Cross-layer Optimization in Cognitive Radio Networks," *Consumer Communications and Networking Conference, 2007. CCNC 2007. 4th IEEE* , vol., no., pp.1128-1133, 11-13 Jan. 2007

[5] C. Rieser, “Biologically Inspired Cognitive Radio Engine Model Utilizing Distributed Genetic Algorithms for Secure and Robust Wireless Communications and Networking,” Ph.D. dissertation, Virginia Polytechnic Institute and State University, 2004.

[6] Baldo, N.; Zorzi, M., "Learning and Adaptation in Cognitive Radios Using Neural Networks," *Consumer Communications and Networking Conference, 2008. CCNC 2008. 5th IEEE* , vol., no., pp.998-1003, 10-12 Jan. 2008

[7] K. Tsagkaris, A. Katidiotis, P. Demestichas, “Neural Network-based Learning schemes for Cognitive Radio systems” *Computer Communicagtions*, in Press

[8] IEEE Standard 802.11b, “Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: high speed physical layer extension in the 2.4 GHz band,” 1999.

[9] IEEE 802.11g, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications, Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band, http://grouper.ieee.org/groups/802/11/

[10] Laxdal, E.M.; Parra-Hernandez, R.; Dimopoulos, N.J., "Guided construction of training data set for neural networks," *Systems, Man and Cybernetics, 2004 IEEE International Conference on* , vol.6, no., pp. 5905-5910 vol.6, 10-13 Oct. 2004

[11] T. Masters, *Practical Neural Network Recipes in C++*, Academic Press, 1993

[12] S. Haykin, Neural Networks, *A Comprehensive Foundation*, Second Edition, Upper Saddle River, NJ: Prentice Hall, 1999.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

6. RTMS POSITIONING IN A MANAGEMENT FRAMEWORK FOR B3G ENVIRONMENTS

Abstract:

This chapter presents an advanced management framework, as an enabling technology for designing and developing, wireless systems in B3G environments. The chapter focuses on the main components of the proposed framework, as well as on their functionality and interactions. Additionally, indicative simulation results showcase the efficiency of the proposed framework. Parts of this chapter have been published in [1].

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

RTMS POSITIONING IN A MANAGEMENT FRAMEWORK FOR B3G ENVIRONMENTS

6.1. Introduction

This chapter presents an advanced management framework for heterogeneous wireless network infrastructures, as an enabling technology for designing and developing wireless systems in B3G environments. The need for studying and building such a framework to be based on, is mainly due to the following reasons: First, ubiquity is by definition inherent to the B3G concept where a multitude of different RATs, offered by the same or different NOs, are offered towards serving users' requests. Second, information provision and awareness are captured by what is referred throughout this chapter as context information acquisition and includes the collection of critical information e.g. load, network state, signal strength etc, as described in more detail in section 6.3. Third, intelligence lies in the overall adaptation capabilities of the infrastructure, whereas it can also be proven to exist in specialised algorithmic processes as part of the overall management functionality. Last, smart interaction, albeit it is not explicitly covered by the proposed framework, is partially covered (as will be described) by the fact that terminal operation and services are properly adapted a) according to the user preferences and b) independently of the terminal type used. Going one step further, this work places special focus on the integration of multi-standard enabled network elements and terminals into the underlying infrastructure, which is assumed to be rather beneficial, as it offers improved mobile experience and ubiquitous service provisioning, thus increasing the level of intelligence.

In the light of the above, the contribution of this chapter (shown also on Figure 6-1) is manifold: first, it describes contemporary problems that fall in the realm of the management of heterogeneous infrastructures; second, it provides a solution to those problems, through an intelligent framework for the operation and management of network segments and terminals; third, it presents comprehensive simulation results that showcase the effectiveness of the proposed mechanisms.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

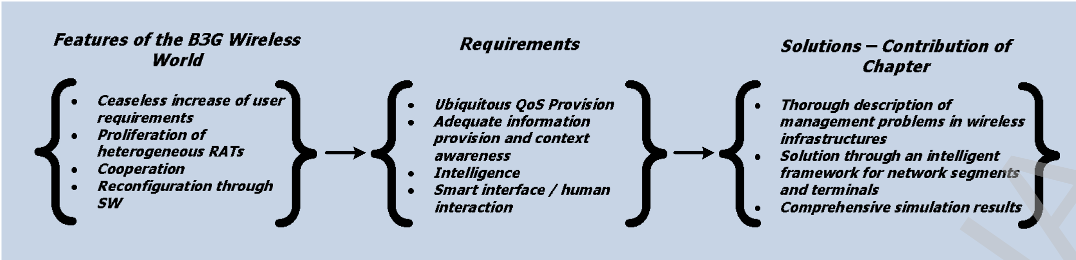


Figure 6-1: Chapter motivation and contribution.

Accordingly, the rest of this chapter is organized as follows: The next section introduces the components and the functionality of the proposed management framework. Section 6.3 focuses on the trigger for its operation, as well as the reconfiguration capabilities of the infrastructure (context awareness). Then, Sections 6.4 and 6.5 describe the decision making mechanisms of the framework from the viewpoint of the network and the terminal, respectively, showing also the way they address intelligence (adaptation). Indicative simulation results are included in Section 6.6. Finally, Section 6.7 contains some concluding remarks and potential extensions of this work.

6.2. Architecture and Functionality of Management Framework

This section emphasizes that the design and development of an efficient management architecture and functionality, acts as a necessity for leveraging the heterogeneous B3G infrastructures to high performance environments.

6.2.1. High Level Description of Management Architecture

A B3G oriented scenario naturally considers users, dispersed within the service area of a network segment (operated by one or more NOs). The segment is covered by numerous RATs provided either by legacy systems operating in a fixed-RAT fashion or by SDR-based systems disposing base stations with reconfigurable transceivers. The segment should cater for the ubiquitous provision of versatile services. At the same time, user terminals should be able to access those services from anywhere and

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

at anytime. To do so, terminals need to be able to perform any suggested changes, in order to better adapt to current conditions.

The scenario assumes that certain contextual circumstances impose some need of network segment and associated terminals' adaptation. In this respect, Figure 6-2 shows in a high level fashion the approach that will be followed in the chapter, in order to solve such a problem, through the intelligent management of the infrastructure. The proposed management framework comprises two main components, namely the Network Manager (NM) that is responsible for (re-) configuring its network elements so as to offer the NOs' services in the most efficient manner to their users/terminals, as well as the Terminal Manager (RTMS) that can be seen as the interface of the user to the service provisioning system. To this end, a fundamental concept lies in that the overall management framework acts independently of the specific network elements and terminals, thus supporting the desired intelligence.

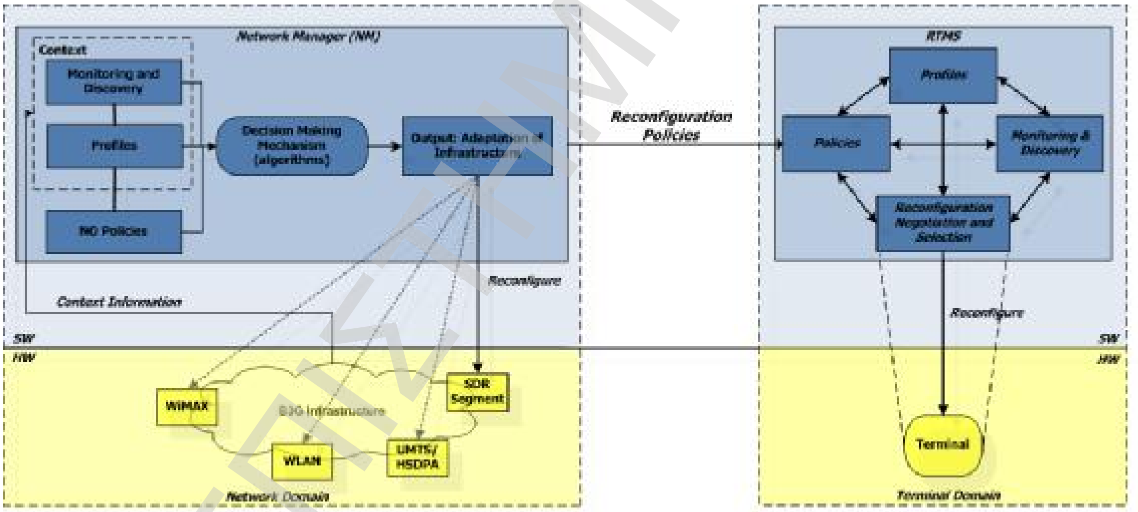


Figure 6-2: Architecture of the proposed management framework.

In addition, the system exhibits context awareness as it is capable of identifying the need for adaptation to environment requirements, e.g. whenever a degradation of network operating parameters is detected through the monitoring of specified Key Performance Indicators (KPIs). In general, this context information is related to traffic load and QoS provision, and is also enriched with information on the adaptation capabilities (for legacy and SDR-based network segments and terminals).

Intelligence is also exhibited through the use of algorithms conducted during the adaptation process. These algorithms lie within both network and terminal management systems and collaborate with each other so as to continuously provide the user with services at high QoS levels and in a ubiquitous manner. The next subsection introduces this functionality.

6.2.2. Management Functionality

It is assumed that a sudden or recurrent change in context requirements imposes some need of network segment and terminals' adaptation. The adaptation process begins with the initial trigger through context information transferred towards the NM.

In the sequel the NM follows an algorithmic process (described in [2]), targeted at deciding upon the most proper adaptation pattern, in terms of (i) allocation of RATs/ spectrum to base station transceivers, (ii) allocation of users to base station transceivers and (iii) allocation of user applications to QoS levels.

In general, the NM decisions depend on the infrastructure's capabilities. In this respect, three potential solutions are identified.

- A first possibility is the redistribution of the demand to the available RATs. This seems as the most easily deployable solution that can exploit heterogeneity, but requires cooperation among the RATs, in order to accommodate any overflow demand.
- A second potential solution would be provided by means of flexible spectrum management, which envisages dynamic selection of operating spectrum band in legacy RATs. This implies e.g. that a UMTS transceiver of a base station decides on changing its operating frequency (from 1900 MHz to 900 MHz), to cope with interference and thus increasing its serving capability.
- Third, adaptation can be guaranteed by SDR-enabled network segments, which are able to dynamically select their configuration (software activation for RAT and frequency band, modulation type, power levels, etc.), so as to respond to external stimuli.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

What can be easily observed is that the NM decisions are twofold, in the sense that they affect not only the network i.e. by deriving reconfiguration decisions for the base station transceivers, but also the users dispersed within the service area.

Specifically, regardless of the particular user terminals used, what the RTMS acquires form the NM, is a set of ***reconfiguration policies*** that serve as directives that must be taken into account during each individual terminal's adaptation process. Eventually, the RTMS, on behalf of the user-terminal, is in charge of making the final decision regarding the (re-)configuration of its operating parameters (e.g. RAT, frequency, modulation etc.). It does so by conducting an optimization process similar to the one proposed for NM [3],[4] by taking also into account an additional constraint: any decision must be fully compliant with the reconfiguration policies received from NM.

The proposed management framework is also related to the work conducted within IEEE P1900.4 Working Group [5],[6] in the sense that it accomplishes the adaptation to the conditions by distributing the decision making process among the collaborating management entities in network (NM) and terminals (RTMS) sites.

The next sections analyze (i) the acquisition of context information by the NM (Section 6.3), (ii) the NM selections and the transfer of reconfiguration policies to the terminals (Section 6.4) and (iii) the associated reconfiguration decisions taken by the RTMS (Section 6.5). Indicative simulation results that showcase the behaviour of the proposed framework are included in Section 6.6.

6.3. Context Information Acquisition

This section presents a minimal set of parameters that serve as trigger for adaptation to a new (given) context.

The acquisition of information related to the context in which the infrastructure operates, assists the NM to detect when some kind of adaptation (reconfiguration) is needed. In this respect, capacity and coverage aspects, as well as traffic demand information should be mainly provided. Therefore, the necessary input to the network resource optimization process incorporates information on the base station transceivers, such as the current load and the maximum load capability and also, information on the terminal, such as the serving and neighbouring base stations and

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

the corresponding signal quality. It also contains more static information such as base station transceivers' and terminals' identifiers, as well as relevant NO policies and profiles. As a result, the NM is provided with information from both, the terminal and network side.

Table 6-1 summarizes the KPIs collected in the case of EDGE, WIFI, HSDPA and WIMAX RATs. Through these measurements, the RAT-integrated Radio Resource Management (RRM) modules are able to support the proper QoS levels by supplying the NM with all necessary information. As observed in the table, the RTMS is aware of the quality of its connection to the currently serving network element (cell), while it is also aware of neighbouring base station transceivers by sensing the strength of their signal. This information is shared with the NM, allowing it to take efficient decisions. On the other hand, the NM is able to compute the uplink and downlink traffic for each terminal and consequently estimate the overall traffic load at the base station transceiver.

Table 6-1: Context information per RAT and per producer.

Producer	***Context Information***	***Parameters per RAT***			
		EDGE	***WIFI***	***HSDPA***	***WIMAX***
Terminal	Identification	IMEI, PLMN	MAC		
	Serving cell quality	RXQUAL	SIGNAL QUALITY	Pathloss, BLER	CINR
	Neighboring cells scanning	RXLEV	RSSI	CPICH RSCP, Ec/N0	RSSI
Access Point	Identification	BSIC	MAC, SSID	Scrambling code	
	Uplink/Downlink traffic load per MT		√	√	√
	Overall AP traffic load		Estimated	Estimated	Estimated

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

6.4. Network Manager (NM) - Selection of Optimum Configurations

6.4.1. Overview of NM

Following the identification of the need for adaptation to a given context, this section focuses on the Network Manager of the proposed management system. In particular, through the utilization of the framework presented below, the NM component is able to decide on the most efficient sets of (re)configurations and issues commands to the network elements of the segment in question. Accordingly, network elements are able to adapt to any new contextual situation providing at the same time solutions that optimize network resource usage and QoS. Likewise, the NM component issues commands (reconfiguration policies), that are targeted to the RTMS(s) and aim to assist the operating terminals to adapt to the new conditions.

The next subsections are dedicated to the description of the input, output and decision making method, for the management functionality in the network side.

6.4.2. Description of Components

6.4.2.1 Input

Monitoring and discovery. This part of the input requires interactions between the NM component and the environment. As an outcome of the monitoring process, context information is sensed for each element of the network segment and for its environment and helps the NM to extract the primary reasons for "acting". It reveals the status of the elements in the network segment and of their environment (therefore performance, fault, etc. notifications will be covered). Additionally, through this procedure, the capabilities of the candidate configurations of the transceivers of the base stations (bit rate and coverage), are identified.

Profiles. This part acquires and maintains information (data and knowledge) on the elements of the segment, such as the set of transceivers of each base station, the set of operating RATs, as well as the spectrum assigned for operation to the network segment. Moreover, this part also describes the profiles (e.g., preferences, requirements, constraints) of user classes, applications and terminals, etc.

NO Policies. NO policies designate high level rules that should be followed in context handling. Usually they are imposed by NOs/ regulators and refer to reconfiguration strategies, such as e.g. NO preferences and priorities on goals to be achieved. These are related to the maximization of the QoS levels, and the minimization of cost factors (e.g. resource consumption). Furthermore, this part provides information on NO agreements with cooperative NOs. It is noted here that NO policies should not be confused with 'reconfiguration policies' that form part of the output of the process towards terminals and are described in the following.

6.4.2.2 Output

As already stated in the introductory section, the adaptation algorithm of the NM is based on the work in [2] and is targeted at deciding upon the most proper allocations of RATs/spectrum to base station transceivers, of users to base station transceivers and of user applications to QoS levels. Accordingly, the output provides actions that will enforce such reconfiguration decisions to both the network elements and terminals. Especially regarding the latter case, the output reconfiguration decisions are first mapped into specially formatted reconfiguration policies and in the sequel, they are conveyed to the respective RTMS(s). The reconfiguration policies will be taken into account as directives during individual terminals' reconfiguration decision making and might contain, for example, suggestions for selecting specific sets of RATs/frequencies etc.

6.4.2.3 NM operation: decision making method

The overall solution approach aims (as mentioned above) at deciding for the optimum selection of RAT/spectrum, distribution of demand and assignment of QoS levels, to be applied in each of the transceivers of the base stations within the segment in question and to be properly propagated to the affected terminals. Those allocations should optimize an objective function (OF) associated with the following: (i) user classes should be allocated to their most preferred QoS levels, i.e., those that maximise the aggregate utility volume [9]; (ii) the cost of reconfiguring the network should be minimized. The last factor can be associated with the number of transceivers that change RATs or spectrum. If several solutions lead to the same aggregate utility volume the one requiring the minimum number of changes should be retained. This method is thoroughly analyzed in [2]. Additionally, given the RAT/spectrum allocation, the QoS levels assignment and the way that the demand is distributed among the various technologies, what follows is a well established problem , targeted at the assessment of a configuration scheme in a single RAT situation e.g. CDMA-based [7], OFDM-based [8] etc.

6.5. Terminal Manager (RTMS)

6.5.1. Overview of RTMS

The focus of the RTMS is on the user plane. More specifically, it comprises mechanisms that will allow a user to access services provided by the NO, at any place, any time, through any device. In other words, the RTMS addresses the issue of adapting the operation of the terminal that the user is currently interacting with, considering what best suits the user's needs at a specific moment. Thus, this adaptation should be made in accordance with the user preferences, the device capabilities, the available communication infrastructure at the present location of the user and the policies imposed by the network management side. Regarding the latter, as previously mentioned, some NM decisions may result in several reconfiguration policies, which are transferred to the terminals, operating within the same administrative domain.

In order to meet this objective of ubiquitous connectivity and service provision adaptation in a heterogeneous, B3G environment, the RTMS should be capable of:

Acquiring context information. This requirement derives from the need to consider the state of the served user, terminal and environment. The basic contextual information consists of the location of the user, the time zone, the capabilities of the currently used terminal and the corresponding available network configurations.

Maintaining and managing information on user preferences, terminal capabilities and configuration policies. The RTMS should be capable of storing accurate descriptions and representations of the aforementioned information as well as configuring and updating the respective profiles and policies.

Negotiating with the available networks and selecting the most appropriate (re-) configuration action for the terminal. The selection procedure may result into a switch from one RAT to another. This selection should be consistent with user preferences, terminal capabilities and configuration policies. The selection should also take into account the specific service area conditions and the time zone of the day. This selection is not restricted to pre-installed technologies. On the contrary, the notion of full reconfiguration supports the dynamic downloading, installation and validation of software components, required for the reconfiguration process thus enabling the operation of a RAT that was not initially included in the capabilities of the terminal.

6.5.2. Description

In order to meet the aforementioned requirements, it can be considered that the RTMS consists of the following main components: *Profiles, Policies*, *Monitoring and Discovery* and *Reconfiguration Negotiation and Selection*. This sub-section discusses on the functionality of these components.

6.5.2.1 Profiles

This component acquires, maintains and provides information on terminal capabilities and user preferences. Indicative information includes (i) the set of potential terminal configurations (the RATs that the terminal is capable of operating with, the associated spectrum and transmission power levels), (ii) the set of applications that can be used and the sets of QoS levels associated with the use of an application, (iii) the utility volume [9] that is associated with the use of an application at a certain quality level, and (iv) the maximum price that the user is willing to pay in order to use certain applications at specific QoS levels.

6.5.2.2 Policies

This component manages information related to policies. Policies reflect the NM decisions on the administrative domain in which the user is currently located. Policies define constraints related to permissible options for device configuration and service adaptation. The rules implicitly or explicitly imposed by the policies are taken into account during the RTMS decision for selecting the most appropriate configurations. Policies refine and complement the input designated by the profiles and the context.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

More specifically, policies constrain the set of permissible services and corresponding QoS levels, RATs that may be used, allowed frequency bands and power transmission levels.

6.5.2.3 Monitoring and Discovery

The role of this component is to obtain context information related to the location of the user, the present time zone, the capabilities of the device used, the available configurations in the environment, and the status of the current configuration. In more detail, the monitoring procedure, performed at regular time intervals, deals with the collection of information on the current connection and the applications used. The collected values of the monitored parameters are compared with a set of predefined thresholds. In case violations are noted and if a reconfiguration is considered appropriate, the discovery process is initiated so as to identify potential alternate configurations, e.g. RATs offering better operating conditions. It should be noted that the discovery process may not only be triggered by the monitoring procedure, but can also be performed in regular time intervals. In case the alternate RAT(s) can be deployed (considering terminal/user profiles and rules imposed by policies), the Reconfiguration Negotiation and Selection module is triggered to re-evaluate the selected RATs.

6.5.2.4 Reconfiguration Negotiation and Selection

This component decides on the most appropriate (re-) configuration for the terminal in terms of the obtained QoS levels, taking into account the current context, the user and terminal profiles, and the policies. In principle, the above factors are embedded, either through relevant decision variables or through appropriate constraints, into an objective function that has to be optimized. The OF is extracted in a similar manner to the one adopted by the NM, reflecting a maximization of the aggregate user utility volume for all services accompanied by a minimization of the related monetary cost and the reconfiguration cost. The reconfiguration cost is used to express the time and the device resources (e.g. battery, memory space) that may be required to realize a certain configuration decision. Information on the cost associated with the usage of certain services can be obtained through negotiations with networks complying with policies. In the case of the terminal participating in negotiations, three negotiation

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

schemes have been considered, namely first-price sealed bid, reversed English (Procurement) and reversed Dutch auction models [9], [10] [11]. In general, the approach is flexible. The terminal can be configured to finish the negotiation in one step, e.g., in a manner similar to the first price sealed bid strategy, or to conduct multiple steps, e.g., similar to the reversed English or Dutch auction models. The most simple negotiation scheme follows the first-price sealed bid auction. All networks make their offers (bids) simultaneously, the best offer (in terms of QoS and cost) is selected and the offered cost is taken as the agreed price. Other negotiation models investigated are a “reversed Dutch” auction model, where the Reconfiguration Negotiation and Selection continuously raises the price, until a network accepts the price and the negotiation terminates and a “reversed English” (Procurement) auction model, where the prices offered by the networks are continuously lowered until the Reconfiguration Negotiation and Selection accepts one.

6.5.2.5 RTMS Operation

A typical scenario exhibiting the RTMS’s operation and the interactions among its components is the following:

The scenario starts from a trigger that derives either from the Monitoring and Discovery (new context e.g. the user location changed), the Profiles (the user profile has been modified) or the Policies component (new policies). The trigger is then forwarded to the Reconfiguration Negotiation and Selection, which cooperates with the Profiles, Policies and Monitoring and Discovery components for acquiring information on the user and terminal profiles, the configuration policies, as well as current contextual conditions. The next step is the decision on the optimal (re-) configurations, which should lead to the best possible handling of the new context. Finally, the Reconfiguration Negotiation and Selection decisions are made known, accepted and applied.

6.6. Simulation Results

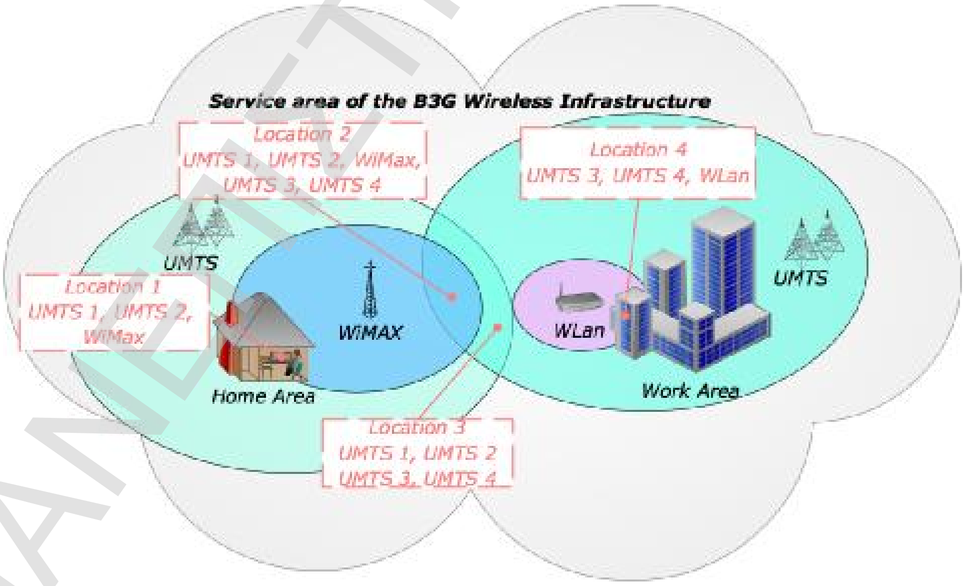
This section includes indicative evidence on the efficiency of the RTMS.

In general, a typical heterogeneous B3G infrastructure is considered that comprises a specific NO, operating three different RATs namely UMTS, WLAN and WiMAX.

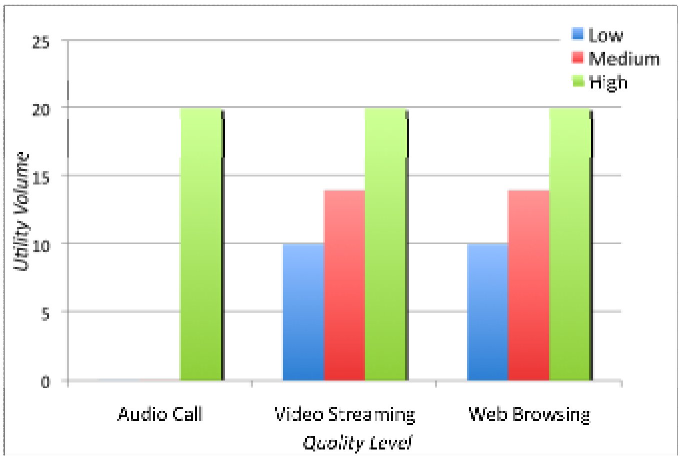
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

The OF used during the selection process will be used as a performance measure, since (implicitly or explicitly) it encompasses the QoS levels offered by the composite radio infrastructure, policies and reconfiguration cost-related aspects. It is assumed that the terminal device is capable of operating with all the configurations available in the service area.

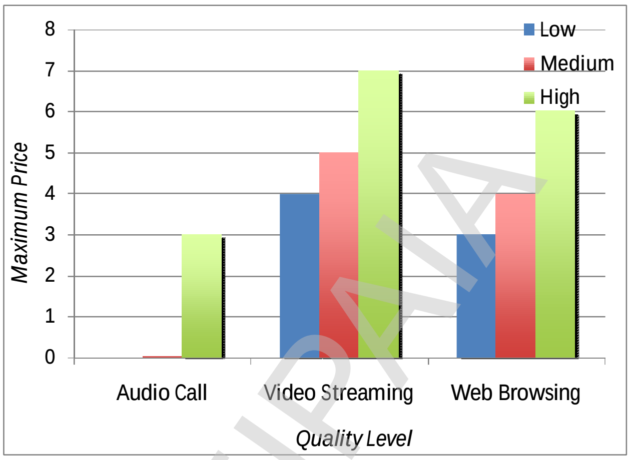
The focus is on three services, namely audio-call, video-streaming (including applications such as IPTV and mobile TV) and web-browsing, although the analysis and the results are similar in the case of other services as well. Each QoS level corresponds to a set of four reference parameters, namely blocking probability, bit-rate, bit-error-rate and dropping probability. For the audio call service only one (high) reference quality level has been defined corresponding to the following values for the above mentioned parameters {(1%, 64Kbps, 250 msec, $10^{(-3)}$) (high)}. For the video-streaming and web-browsing services, three reference quality levels (low, medium, high) have been defined, corresponding to the following associated parameters: {(1%, 128 Kbps, 500 msec, $10^{(-4)}$) (low), (1%, 256 Kbps, 500 msec, $10^{(-4)}$) (medium), (1%, 512 Kbps, 500 msec, $10^{(-4)}$) (high)}. Regarding service provisioning through specific RATs it is assumed that there is a specific policy defining that the audio call service cannot be obtained through the WiMAX RAT.



(a)



(b)



(c)

Reconfiguration Cost	
UMTS -> UMTS	1
UMTS -> WLan	3
UMTS -> WiMAX	3
WLan -> UMTS	3
WLan -> WLan	1
WLan -> WiMAX	2
WiMAX -> UMTS	3
WiMAX -> WLan	2
WiMAX -> WiMAX	1

(d)

RAT	*Priorities*
UMTS (1)	6
UMTS (2)	3
WiMAX	14
UMTS (3)	8
UMTS (4)	2
Wlan	14

(e)

Figure 6-3: (a) Service area considered in the scenario and respective coverage (configurations available). (b) Utility volume values. (c) Maximum price values. (d) Cost of reconfiguring transceiver. (e) Priorities allocated to various transceiver configurations by the policies.

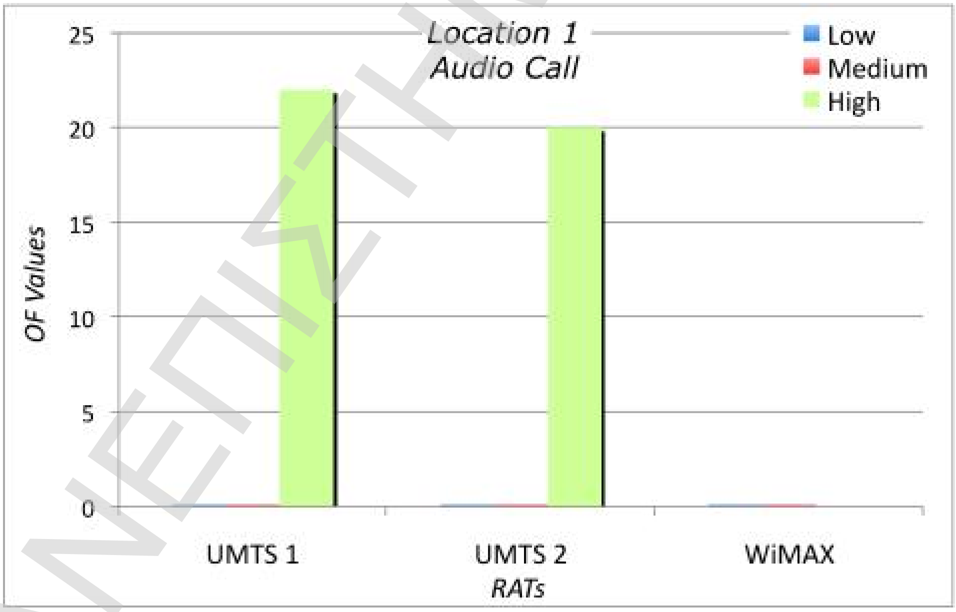
Figure 6-3(a) depicts the service area considered and the coverage provided by the B3G infrastructure. Based on the coverage provided by the RATs, four main locations can be identified, as shown on the figure.

Regarding the user profile information, Figure 6-3(b) and Figure 6-3(c) depict the utility volume values and the maximum acceptable prices that are relevant to each service and quality level, respectively. Figure 6-3(d) presents the cost value assumed for reconfiguring the device from one RAT to another. Different cases have been considered. The logic that has been followed is that the configurations of the WLAN and WiMAX RATs may have several common components (functions), therefore, the reconfiguration cost may be smaller, compared to that of changing between the UMTS and WiMAX/WLAN configurations, since the component similarities may be fewer. It should also be noted that the reconfiguration cost for each case depends on the RAT that has been selected in the previous case. Finally, regarding policies, Figure 6-3(e) presents a set of priority values which are part of the policies imposed by the NM for the use of the different RATs (respective configurations). In the test scenario, it is assumed that the terminal device is initially connected to UMTS 2. It is also assumed that the user is moving from the first location, as shown on Figure 6-3(a), towards location 4, passing from all the intermediate locations, requesting an audio call service at the first one, a video-streaming at the second and third locations and a web-browsing service at the last location.

Figure 6-4 presents the OF values for the various options investigated by the RTMS and implicitly provides insight on the efficiency of the solutions proposed for the different cases of utility volumes, prices, policies and reconfiguration cost values. At location 1, the audio call service can be obtained only through UMTS, since (as was mentioned in the assumptions outlined at the beginning of this section) policy rules set by the NM prevent the use of the WiMAX RAT for the given service. As shown in Figure 6-4(a), the optimal configuration at the given location is to acquire the service through UMTS 1 which corresponds to the maximum OF value. This is due to the fact that in the policies, a higher priority value has been set for UMTS 1 with respect to UMTS 2 (Figure 6-3(e)). The results show a 10% improvement with respect to what would happen in the absence of the RTMS in which case the terminal would remain connected to the initial RAT (UMTS 2). At location 2, the video-streaming service can be obtained through all available RATs. In this case, the solution proposed by the RTMS is to switch to the WiMAX RAT and obtain the service at the High QoS level. This is the optimal configuration as it corresponds to the highest OF value (Figure 6-4(b)). The high OF value in this case derives from the strong preference of the user for High QoS (high utility value for this level) and the fact that, in the policies, a high priority value has been set for the WiMAX network encouraging its use, even though the reconfiguration cost taken into account for switching to the WiMAX network is

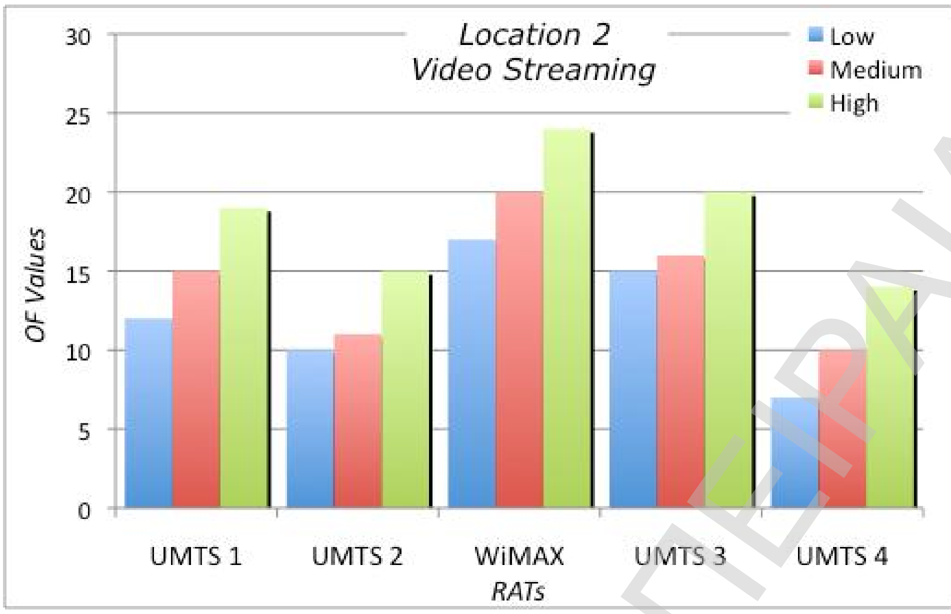
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

also high. The results show a 60% improvement with respect to the existing situation, which would be to remain connected to the UMTS 2 RAT, without exploiting the B3G infrastructure. At location 3, the video-streaming service can be obtained through the available UMTS RATs (UMTS 1-4). Again, the reconfiguration cost and the priorities set by the NM have a major impact in the reconfiguration decision. The optimal configuration proposed by the RTMS for the given case is to switch to the UMTS 3 RAT. The improvement in this case is 38% with respect to the alternative option (i.e. the lack of the RTMS), which would be to use the UMTS 2 RAT (Figure 6-4(c)). Finally, at Location 4, the user terminates the video-streaming service and initiates a web-browsing service. In the given location a WLAN RAT is available, offering the service at a high quality level. In this case, the solution proposed by the RTMS is to switch to the WLAN RAT, since it is the optimal configuration as it corresponds to the highest OF value (Figure 6-4(d)). Considering that without the RTMS the device would be connected to either UMTS 3 or UMTS 4 (i.e. it would not be able to exploit the option of the WLAN RAT) the improvement achieved is between 14% and 67%, respectively.

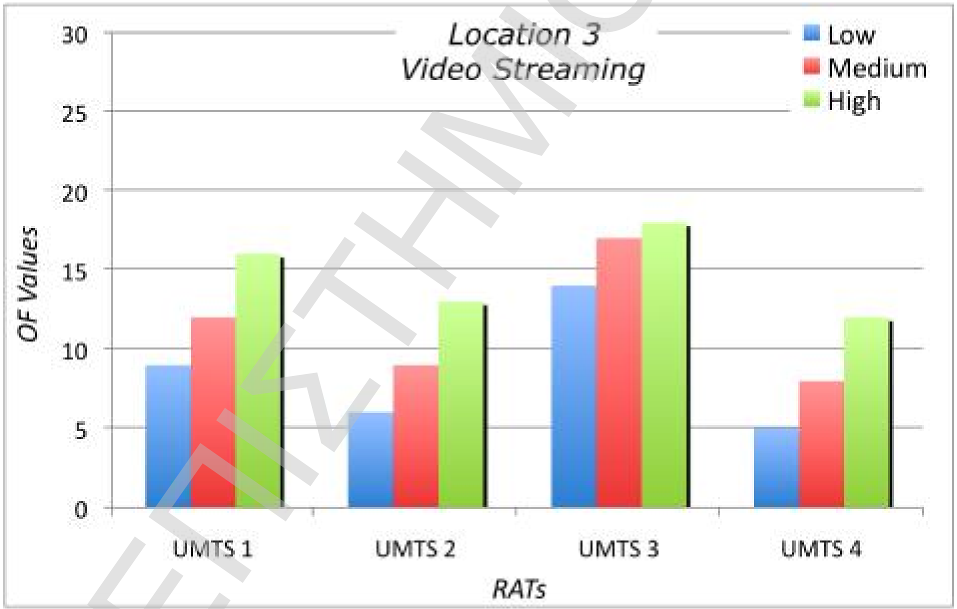


(a)

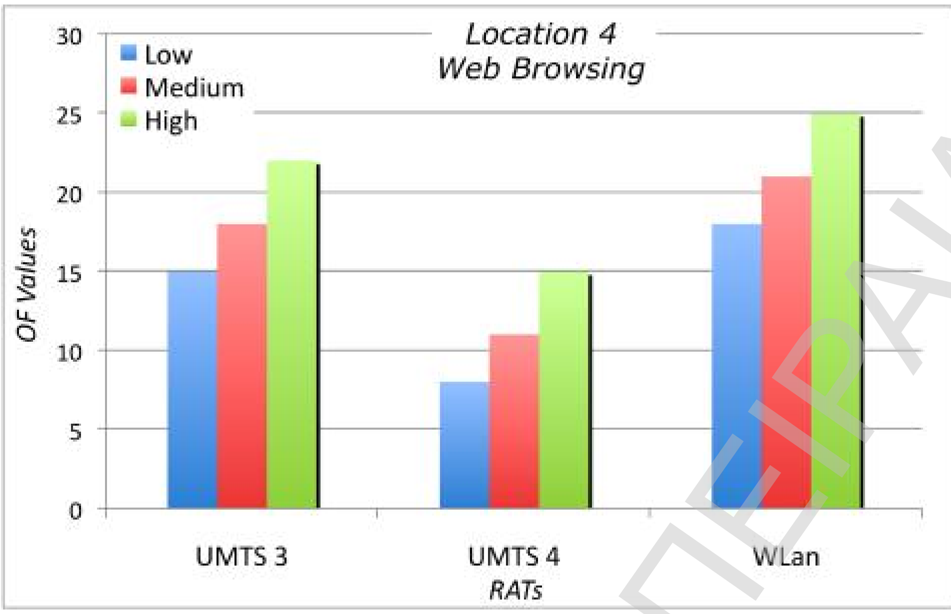
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ


Location 2
Video Streaming
Low
Medium
High
OF Values
RATs
UMTS 1
UMTS 2
WiMAX
UMTS 3
UMTS 4
0
5
10
15
20
25
30

(b)


Location 3
Video Streaming
Low
Medium
High
OF Values
RATs
UMTS 1
UMTS 2
UMTS 3
UMTS 4
0
5
10
15
20
25
30

(c)



(d)

Figure 6-4: OF values of available solutions (a) at location 1, (b) at location 2, (c) at location 3 and (d) at location 4.

In general, the improvement from the exploitation of the B3G infrastructure, through the RTMS, compared to the performance obtained from an infrastructure in which only 3G technologies exist, ranges from a 10% to 60%. The improved behaviour is due to the fact that the RTMS is capable of efficiently exploiting a B3G infrastructure. More specifically, the RTMS continuously seeks opportunities for improving the performance of the currently used device, based on user profile information, context and policies.

6.7. Conclusions

B3G wireless infrastructures have the potential to offer people an unprecedented level of convenience and flexibility for living and working. Such technologies need to be closely integrated with human interactions. To do so, the complementary use and exploitation of the numerous RATs available in the B3G world, stands as a fundamental requirement. However, this poses several issues in the management domain, regarding the intelligent management of network segments and terminals operating within heterogeneous infrastructures. The goal of this chapter, accordingly, was to present such an advanced management framework, as an enabling technology for designing and developing, wireless systems in B3G environments. Results from its application to a simulated communication infrastructure exposed its efficiency.

A natural deduction from the simulations presented is that the proposed management framework can contribute towards the realization of the concept of intelligent terminals. Nevertheless, there is still a long way to go towards ubiquitous and trustworthy service provisioning. In particular, what needs to be considered is attributing the proposed framework with learning capabilities. This will enable the recording of previous decisions, and therefore, acquisition of knowledge and experience in planning future actions fast and appropriately.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

REFERENCES

[1] G. Dimitrakopoulos, K. Tsagkaris, V. Stavroulaki, A. Katidiotis, N. Koutsouris, P. Demestichas, V. Merat, S. Walter, "A Management Framework for Ambient Systems Operating in Wireless B3G Environments", ACM/Springer Mobile Networks and Applications journal, to appear.

[2] K.Tsagkaris, G.Dimitrakopoulos, A.Saatsakis, P.Demestichas "Distributed Radio Access Technology Selection for Adaptive Networks in High-Speed, B3G Infrastructures", *International Journal of Communications Systems, Wiley,* Vol. 20, Issue 8, Aug. 2007, pp 969-992

[3] V. Stavroulaki, S. Buljore, P. Roux, E. Mélin, "Equipment Management Issues in B3G, End-to-End Reconfigurable Systems", IEEE Wireless Communications Magazine, Special issue on Composite Reconfigurable Radio Networks, Vol. 12, No. 3, June 2006, pp. 24-32.

[4] V. Stavroulaki, A. Katidiotis, D. Petromanolakis, P. Demestichas, "Equipment Management Strategies in Reconfigurable Networks", in Proc. 63rd IEEE Vehicular Technology Conference (VTC-06 Spring), Melbourne, Australia, May 2006

[5] IEEE P1900.4, http://grouper.ieee.org/groups/emc/emc/1900/4/

[6] S. Buljore et al. "Introduction to IEEE P1900.4 Activities", *IEICE Transactions on Communications, Special Section on Cognitive Radio and Spectrum Sharing Technology (Invited Paper),* Vol. E91-B, pp 2-9, Jan. 2007

[7] K. Tsagkaris, G. Dimitrakopoulos, P. Demestichas, "Policies for the Reconfiguration of Cognitive Wireless Infrastructures to 3G Radio Access Technologies", ACM/Springer Wireless Networks journal, to appear

[8] Gautam Kulkarni, Sachin Adlakha, Mani Srivastava, "Subcarrier Allocation and Bit Loading Algorithms for OFDMA-Based Wireless Networks," *IEEE Transactions on Mobile Computing* ,vol. 4, no. 6, pp. 652-662, November/December, 2005.

[9] P. Klemperer, "Auctions: Theory and Practice", Princeton University Press, 2004

[10] Auction. (2008, May 8). In Wikipedia, The Free Encyclopedia. Retrieved 10:31, May 9, 2008, from: http://en.wikipedia.org/w/index.php?title=Auction&oldid=211117120

[11] N. R. Jennings, P. Faratin, A. R. Lomuscio, S. Parsons, C. Sierra and M. Wooldridge, 2001 "Automated negotiation: prospects, methods and challenges", Int. J. of Group Decision and Negotiation, Vol. 10, No. 2, pp. 199-215 , 2001.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

7. CONCLUSIONS – FUTURE CHALLENGES

The future of wireless communications will be characterized by highly varying environments with multiple available radio access technologies exhibiting diverse features, as well as by the over-discussed utilization of radio spectrum. In such an unfamiliar landscape, cognitive radio systems are expected to play an exceptional role by juxtaposing an inherent ability to perceive, think, decide, learn and adapt to the changing environmental conditions. The investigation upon and implementation of learning schemes that will assist the cognitive radios in the derivation and enforcement of decisions regarding the selection of the best radio configuration was the subject of this thesis.

The benchmarking work conducted attests that the improvement from the exploitation of the B3G infrastructure, through the use of mechanisms presented in this thesis, is significant.

Particularly, the RTMS exploits the profiles, the context and the policies, for selecting the best configurations, and therefore, yielding ubiquitous, personalised, context-aware always-best connectivity, in a seamless to the user manner. By exploiting the B3G infrastructure, RTMS offers higher QoS levels, compared to the legacy situation, in terms of reliability and dependability levels. This happens because there are various options (configurations and RATs) through which context handling can be done, and the best QoS levels can be obtained.

For the channel estimation phase, with the use of Bayesian networks, the proposed method can exploit any legacy, robust channel estimation mechanism. Assuming a mechanism is available for that purpose, it has been shown that the method can exploit the provided CSI in order to increase the level of certainty that a configuration will achieve a specific bit rate. The results of the method can be exploited to drive the selection of one of the alternative configurations and thus, ensuring that a cognitive transmitter will always optimize its operation.

The same applies for the NN-based learning schemes presented in this thesis. Relying on the indisputable fact that cognitive radio systems have gained extremely high

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

attention from the wireless research world in the recent years, as well as they will do in the upcoming years, the main idea was to implement learning mechanisms, so as to assist cognitive radios in deciding for their operating radio configuration. The results obtained from the benchmarking work indicate that NNs could help raise the confidence on the expected performance of a link, giving more accurate and correct decisions on the selection of the appropriate configuration.

Subject for future research is to exploit the management schemes presented for enabling NOs to personalize their service offerings, instead of limiting subscribers to a fixed set of inflexible choices. Seamless mobility applications can build on frameworks like the RTMS, so as to intelligently change the services that they provide based on business policies and context.

The main directions for future work, and the respective challenges, include the following main points.

§ The realization of design analysis studies for properly deploying the RTMS support functionality in legacy and emerging wireless B3G infrastructures.

§ The enhancement of the RTMS with machine learning techniques, in order to obtain the management functionality necessary for cognitive terminals.

§ The standardization, where necessary, of parts of the information flow required for supporting the RTMS.

§ The pursuit of regulation activities, in order to ensure that the cooperation between NOs will result to direct benefits (i.e., immediate discounts) for the users.

The first point stems from the fact that the RTMS relies on support functionality, for instance on policies. This functionality has to be properly deployed in order to guarantee efficient operation. The second point stems from the fact that machine-learning techniques can offer cognition, which in turn can increase the speed with which decisions are taken, and the degree of certainty on the appropriateness of the decisions. Standardisation activities will ensure interoperability in heterogeneous infrastructures, with equipment from various manufacturers. Nevertheless, the extent of standardisation can be limited, as interoperability can rely on application layer interactions that use semantic technologies. Regulation is important for guaranteeing

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

that the benefits of inter-NO negotiations will result to benefits for the end-user. The challenges addressed above can be addressed, more or less, in parallel. The current status of the RTMS work can trigger the deployment studies, standardisation and regulation. Cognitive techniques will rely on the information specified in this work. In general, the research challenges identified are relevant to the overall field of management systems for terminals in B3G networks.

The intention for future research on the NN topic is manifold. First, the exploration upon other crucial, context information that can be used to feed the NN input layer e.g. location, user preferences or even weather conditions, etc. will be continued in order to achieve even better results. Furthermore, new types and enhanced structures of NNs that have been found to improve both short-term and long-term time series prediction capabilities could be investigated for application to the proposed schemes. Last but not least, as long as evidence on the performance capabilities of each candidate radio configuration of the cognitive terminal can be drawn, the optimization process/algorithm for selecting the optimum one also needs to be thoroughly studied as part of future work.

All in all, the major scope of any academic endeavour is to contribute, even at a minimal stage, to the ceaseless gain of knowledge. When it comes to the evolution of science and technology, the quest for supporting and directing its multi-dimensional aspects stands as a prerequisite for any further step forward. Furthermore, focusing on wireless communications, the unstoppable user desire for ubiquitous connectivity is still burning inside each and every scientific research effort, not excluding this thesis. In this respect, the goal of this thesis was to present some innovative manners to utilize some recent findings in designing and developing new ones. Its author would thus wish that this thesis has been capable of standing among numerous note-worthy relevant research efforts and that it will provide even a small hint to prospective new attempts to accompany wireless communications along the way of knowledge.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

8. APPENDIX: PUBLICATION LIST (NOVEMBER 2008)

Publications in International Journals and Books

1. K. Tsagkaris, A. Katidiotis, P. Demestichas, Neural network-based learning schemes for cognitive radio systems, Computer Communications Volume 31, Issue 14, 5 September 2008, Pages 3394-3404.

2. P.Demestichas, A.Katidiotis, D.Petromanolakis, V.Stavroulaki, “Management System for Terminals in the Wireless B3G World”, accepted for publication to the Wireless Personal Communications journal.

3. P.Demestichas, A. Katidiotis, K.Tsagkaris, E.Adamopoulou, K.Demestichas, “Channel Estimation in Cognitive Radio Systems by means of Bayesian Networks”, Wireless Personal Communications journal, to appear.

4. G. Dimitrakopoulos, K. Tsagkaris, V. Stavroulaki, A. Katidiotis, N. Koutsouris, P. Demestichas, V. Merat, S. Walter, "A Management Framework for Ambient Systems Operating in Wireless B3G Environments", ACM/Springer Mobile Networks and Applications journal, to appear.

5. P. Demestichas, G. Dimitrakopoulos, K. Tsagkaris, V. Stavroulaki, A. Katidiotis, "Introducing Cognitive Systems to the B3G Wireless World", Cognitive Wireless Networks: Concepts, Methodologies and Visions Inspiring the Age of Enlightenment of Wireless Communications", pp. 253-269, Springer, 2007.

6. V. Stavroulaki, G. Dimitrakopoulos, A. Katidiotis, P. Demestichas, D. Bourse, K. El Khazen, “Negotiation of Network Services and Spectrum in B3G, Composite Radio, Environments”, Innovation and the Knowledge Economy: Issues, Applications and Case Studies, pp.1103-1110, IOS Press, 2005.

7. A. Katidiotis, K. Tsagkaris, P. Demestichas, “Performance Evaluation of Artificial Neural Networks Based Learning Schemes for Cognitive Radio

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

Systems", paper is submitted to International Journal of Communication Systems, Wiley.

Publications in International Conferences

1. D. Petromanolakis, V. Stavroulaki, P. Demestichas, A. Katidiotis, "Management system for cognitive terminals exploiting reconfigurable connectivity", in Proc. ICT Mobile and Wireless Communications Summit, Stockholm, Sweden, June 2008.

2. V. Stavroulaki, P. Demestichas, A. Katidiotis, D. Petromanolakis, "Evolution in Equipment Management Concepts: from Reconfigurable to Cognitive Wireless Terminals", in Proc. 16th IST Mobile and Wireless Communications Summit, Budapest, Hungary, July 2007.

3. V.Stavroulaki, A. Katidiotis, D. Petromanolakis, "Management Functionality for Cognitive Wireless Terminal", in Proc. Wireless World Research Forum Meeting 17 (WWRF 17), Heidelberg, Germany, November 2006.

4. A. Katidiotis, V. Stavroulaki, P. Demestichas, M. Muck, B. Steinke, R.K. Atukula, U. Lücking, E. Patouni, P. Magdalinos, "Prototyping Environment for Equipment Reconfiguration Management and Control", in Proc. 64th Conference of the IEEE Vehicular Technology Society (VTC 2006 Fall), Montreal, Canada, September 2006.

5. V. Stavroulaki, A. Katidiotis, D. Petromanolakis, P. Demestichas, "Reconfigurable Equipment Management: Enabling Seamless Experience in Beyond 3G Wireless Networks", in Proc. 2006 IST Mobile and Wireless Communications Summit, Myconos, Greece.

6. A.Katidiotis, D. Petromanolakis, V. Stavroulaki, P. Demestichas, E. Patouni, A. Kousaridas, P. Magdalinos, N. Andriopoulos, K. El-Khazen, S. Buljore, P. Roux, C. Beaujean, T. Farnham, D. Nussbaum, A. Kountouris, F. Marx, "Prototyping for End-to-End Reconfigurable Equipment", in Proc. 2006 IST Mobile and Wireless Communications Summit, Myconos, Greece.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ

7. V. Stavroulaki, A. Katidiotis, D. Petromanolakis, P. Demestichas, "Equipment Management Strategies in Reconfigurable Networks", in Proc. 63rd IEEE Vehicular Technology Conference (VTC-06 Spring), Melbourne, Australia, May 2006.

8. A. Katidiotis, V. Stavroulaki, D. Petromanolakis, P. Demestichas, "Management Platform for Reconfigurable Equipment", in Proc. European Wireless 2006 (EW2006), Athens, Greece, April 2nd-5th, 2006.

9. V. Stavroulaki, A. Katidiotis, D. Petromanolakis, P. Demestichas, "Equipment Management Strategies in Reconfigurable Networks", in Proc. IEEE Vehicular Technology Conference (VTC2005–Spring), Stockholm, Sweden, May 30 - June 1, 2005.

10. A. Katidiotis, V. Stavroulaki, D. Petromanolakis, P. Demestichas, "Equipment Management Platform in Reconfigurable Networks" in Proc. Wireless World Research Forum Meeting 15 (WWRF15), Paris, France, December 8th-9th, 2005.

11. V. Stavroulaki, A. Katidiotis, G. Dimitrakopoulos, P. Demestichas, S. Buljore, "Negotiation and Selection of Equipment Reconfigurations in Beyond 3G Systems", in Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2005), Berlin, September 2005.

12. V. Stavroulaki, A. Katidiotis, P. Demestichas, S. Buljore, N. Alonistioti, E. Patouni, A. Glentis, F. Foukalas, J. Vogler, G. Pfeiffer, Q. Wei, S. Zhong, T. Farnham, C. Dolwin, R. K. Atukula, U. Luecking, S. Gultchev, K. Seo, K. Moessner, "A Management and Control Architecture for Enabling End-to-End Reconfigurable Equipment Operation" in Proc. Wireless World Research Forum Meeting 13 (WWRF13), Jeju Island, Korea, March 1st-2nd, 2005.

13. Stoytcho Gultchev, Kay Seo, Klaus Moessner, Vera Stavroulaki, Apostolos Katidiotis, Panagiotis Demestichas, "Control Mechanisms To Enable Equipment Reconfiguration And Facilitate Operation Of E2R Systems" SDR Forum 2005 Technical Conference.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΑ