Reconfiguration Discovery and Selection in the Context of Autonomic Management of Cognitive Wireless Infrastructures

{P.Demestichas, G.Dimitrakopoulos, K.Tsagkaris, K.Demestichas, J.Adamopoulou}[1], J.Strassner[2]

[1] University of Piraeus, Piraeus, GREECE, gdimitra@unipi.gr
[2] Motorola Labs, Corporate Technology, Schaumburg, USA,

Abstract. B3G (Beyond the 3rd Generation) wireless infrastructures shall be flexible enough, so as to adapt to environment requisitions. Flexibility can be efficiently realized by exploiting cognitive networking concepts. Cognitive, wireless access, infrastructures can dynamically select their configuration, in principle through self-management, realized in a distributed manner, at a maximum level of autonomy. This paper presents the fundamental components of the necessary management architecture in support of (re)-configuration decisions. Additionally, two essential aspects are addressed, i.e. (i) the provision of robust (stable, reliable), learning and adaptation, strategies for estimating (discovering) the performance potentials of alternate reconfigurations, and (ii) the description of a computationally efficient solution to the problem of exploiting the performance potentials of reconfigurations, rating reconfigurations, and in the end, selecting the best ones. Finally, results that showcase the behavior of the presented schemes, are indicated.

Keywords Autonomic computing, Cognitive networks, Greedy strategies, Learning and Adaptation, Utility

1. Introduction

As globalization necessitates worldwide capabilities in communication technology, the world of telecommunications is currently undergoing some radical changes. Utmost research interest is attracted by wireless communications, bringing them at the forefront of technological evolution. The vision of future wireless networks is reflected upon a harmonized coexistence and cooperation of disparate Radio Access Technologies (RATs), basically consisting of cellular technologies [1] such as 3G/4G and different kinds of Broadband Wireless Access (BWA) technologies that have emerged [2] and are mainly represented by Wireless Local Area Networks (WLANs) and Wireless Metropolitan Area Networks (WMANs). The ultimate goal is to achieve ubiquitous wireless access and provision of pioneer services comparable to those traditionally offered through wired systems, by combining the benefits from all RATs that exist in such heterogeneous environments [3],[4].

Framed within this statement stands the migration of the wireless world towards the era of B3G (Beyond the 3rd Generation) wireless access communications. The main idea is that a Network Operator (NO) can rely on different RATs, for achieving the required capacity and quality of service (QoS) levels, in a cost efficient manner. Each distinct RAT has capabilities and characteristics, in terms of capacity, coverage, mobility support, cost, which make it more suitable for certain environment conditions / requirements. A NO can select those that are best suited for delivering capacity and QoS levels, given the current context.

The B3G concept can be realized through cognitive (adaptive, reconfigurable) network concepts [5][6], in conjunction with network cooperation [7][8][9]. Cognitive networks, reactively or proactively, adapt to the environment requisitions, in principle, by means of self-configuration (self-management). Self-configuration is applied, for tackling complexity and scalability. Reconfiguration may affect all layers of the protocol stack, namely, the physical, MAC (Medium Access Control) and LLC (Logical Link Control), network, transport, middleware and application layers.

The general definition of cognitive networks implies some very advanced capabilities in the specific case of cognitive, wireless access, networks, operating according to the B3G paradigm. As part of the reconfiguration, at the physical and MAC layers, there can be elements (hardware components, such as transceivers) that dynamically change the RATs they operate and the spectrum they use, in order to improve capacity and QoS levels. In other words, a hardware component (transceiver) will be changing RATs (and spectrum), in space and time, in order to adapt to new conditions. In this respect, it is believed that cognitive networks enable the realization of B3G infrastructures with reduced capital expenditures (CAPEX).

The realization of cognitive networks calls for a thorough analysis of the management architecture in support of reconfiguration decisions, as well as its requisite functionality. This paper deals with both issues, by presenting the basic components of this architecture, as well as by addressing a problem targeted at the management of a reconfigurable network element, which operates and is managed in parallel with other elements. The problem is called "RAT and Spectrum selection, QoS assignment and Traffic distribution" (RSQT). Our focus is on a highly distributed problem version. Our scheme requires minimal interactions with other management/managed elements. Eventually, our goal is to employ autonomic computing concepts [10][11][12][13] in B3G infrastructures, in order to offer users with seamless mobility and experience.

In summary, the organization of the paper is as follows. Section 2 presents the basic features of cognitive networks and describes their associated management architecture that operates in accordance with autonomics. Section 3 defines the RSQT problem outlining the two main parts of its solution. Section 4 describes the first main part of the RSQT solution, i.e., the robust methods for estimating the performance of reconfigurations. Section 5 is the second part of the RSQT solution, which is a computationally efficient algorithm to the problem of exploiting the capabilities of reconfigurations, as well as rating and selecting the best reconfigurations. Section 6 provides results that show the behavior of our schemes, and section 7 includes concluding remarks.

2. Cognitive, Wireless Access Networks

Cognitive wireless networks have been proposed for the realization of the B3G vision, with reduced CAPEX. This is achieved through their inherent ability to adapt to varying requirements (e.g., change RATs and spectrum at the PHY/MAC layers).

A. Motivation and Features

A complementary idea, to the cognitive network concept, is to merely have cooperating networks, operating different RATs. Then, the NO can select the best, among a set of alternate, cooperative networks, in order to offer the best possible services to its customers. However simple this may sound, the overall goal of providing seamless mobility and connectivity may still be difficult to achieve. The main reason is that mere cooperation, as described, implies that the entire set of RATs should be a priori deployed by the NO. This is not the most efficient way to reduce CAPEX.

Cognitive wireless networks do not presume the fixed deployment of technologies in terminals and network segments; rather, they have embedded intelligence that enables them to learn, from previous interactions with the environment, and, based on those interactions, adapt their functionality according to different external stimuli. This is depicted in **Fig. 1**. Each segment is made up of cognitive elements. Each element and terminal is reconfigurable (can operate with alternate configurations) and has the intelligence to select the best configuration, in order to adapt to the environment conditions. In this context, reconfiguration at the PHY/MAC layers provides the ability to dynamically select the most appropriate RATs and spectrum, in order to better handle business-, service-, resource-, location-, and time-variant requirements.

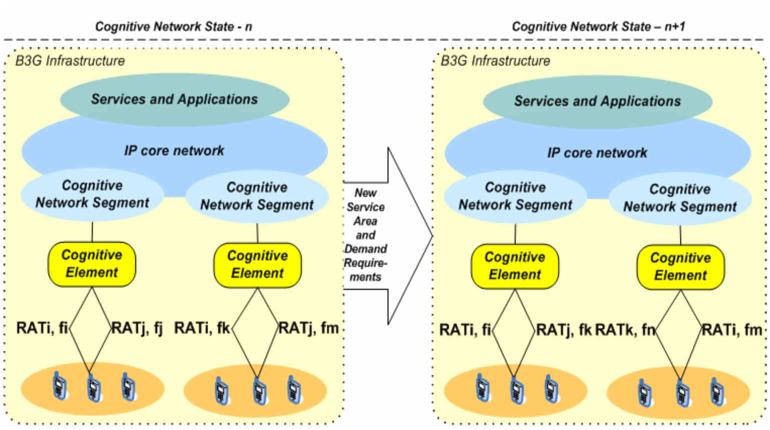
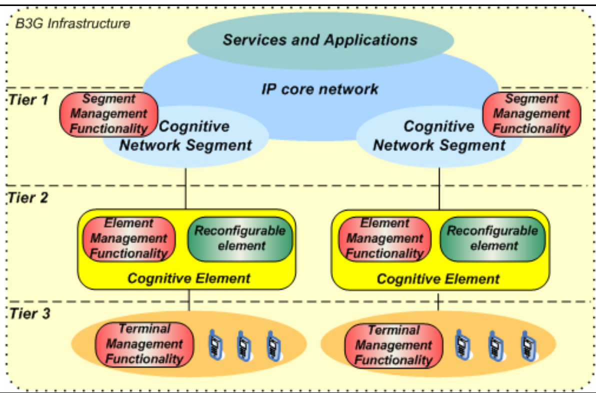


Fig. 1. Example of cognitive wireless network: elements may change RAT, frequency, or both, when new conditions are identified

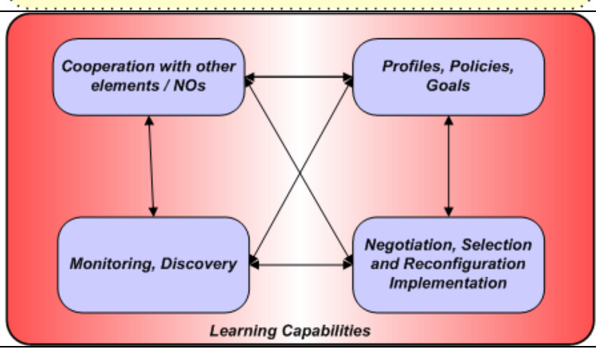
The alternative configurations that should be utilized are known by the cognitive elements, enabling context-aware selection. Configurations change in time and space. Reconfigurations are software-defined. Therefore, a reconfiguration is done by activating the appropriate software, which implements the selected RAT.

B. End-to-End Management Architecture

Since a cognitive network consists of numerous elements and terminals of highly heterogeneous natures, located in different places, a centralized management approach becomes prohibitively complex. Hence, distributed management approaches, relying on pertinent technologies, e.g. autonomic computing, are currently in the focus [11]. This approach can offer scalability and modularity (providing low complexity). **Fig. 2** (a) depicts the overall management architecture of a B3G infrastructure.



(a)



(b)

Fig. 2. (a) Overall Management Architecture, (b) Management functionality for individual element / reconfigurable terminal

Its entities are organized in a hierarchical manner that consists of three-tiers. At tier-1, each (red) entity controls a whole network segment (subset of the network). Tier-1 is made up of mechanisms whose primary purpose is to coordinate with the backbone network, as well as the decisions of tier-2 management entities. Accordingly, the entities at the second tier manage a particular reconfigurable network element (access point). The entities at the third tier are targeted to terminals. Tier-2 and tier-3 entities have the internal structure shown in **Fig. 2** (b).

- The "Monitoring - Discovery" component is set to continuously sense the environment, so as to discover the capabilities of alternate configurations.
- The "Cooperation with other elements/NOs" component can communicate with other elements / NOs so as to acquire their requests, offers, etc.
- The "Profiles, Policies, Goals" component provides user, application and element requirements and characteristics, as well as policies and business goals of the NO.
- The "Negotiation, Selection and Reconfiguration Implementation" component decides upon and implements the changes to be made on the reconfigurable element, based on policies, profiles and the integrated learning capabilities.

Cognitive radio architectural approaches should provide for the maximum possible individual as well as group operation, so as to decrease the system's complexity and support its scalability. As already introduced, it is anticipated that such a fully-distributed approach can be provided by the use of autonomic computing [10][13]. Autonomic Computing is derived from the human nervous system – just as the human nervous system performs involuntary actions (such as pumping blood) to free the human brain to address other tasks, autonomic computing systems perform tasks that previously required intensive manual operation (such as (re)configuring a device) to enable the autonomic system to perform more strategic tasks (such as optimization and planning).

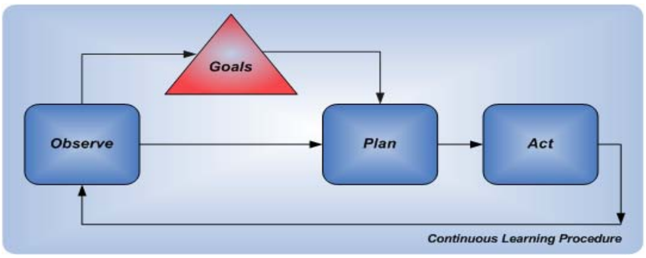


Fig. 3. Autonomic operation of cognitive element

Local optimization is achieved by sending to entities the appropriate policies, which direct self-management towards a global operational goal. Ideally, the distributed decision entities have full knowledge of the context, thanks to cognitive support

functionality (enhanced with learning capabilities). This is shown in **Fig. 3**, where cognition is reflected on a feedback loop (Observe- Plan- Act - OPA).
In this context, the network continuously observes the environment, looking for potential changes that can affect its operation. Observations form the basis for initiating machine-based reasoning to see if the reconfiguration process should be invoked. Once the decision is taken, the network acts accordingly. This loop is repeated inside a machine learning process [14], which leads to cognition. The loop is guided by a set of goals, which take observations into account when planning actions.
This section presented management architecture for cognitive wireless networks. The components of **Fig. 2** should be further enhanced, primarily through autonomic-computing and learning functionality, fine-tuned and validated. A basic model for discovering and selecting reconfigurations is presented in the next section.

3. RSQT Problem Statement

RSQT is seen as a main part of the management functionality required for taking reconfiguration decisions in the context of cognitive, infrastructures. **Fig. 4** provides the overall problem description and the relevant data structures.

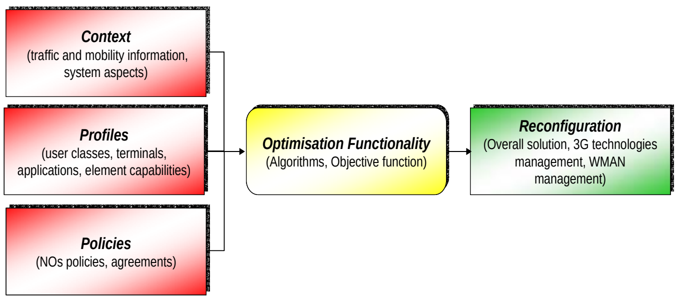


Fig. 4. RSQT problem overview

The input to the problem is classified into three main categories: (i) context; (ii) profiles; (iii) policies.
The context part provides information on the candidate configurations of the element. In general, an arbitrary reconfigurable element, e, will have a set of transceivers, T_e. Each transceiver, $t \in T_e$, will be capable of operating a set of RATs, $R_e(t)$. Moreover, there is a set of spectrum carriers, $F_e(t,r)$, with which t can operate RAT r. In general, for all $t \in T_e$, $F_e(t,r) \subseteq F_r$, where F_r is the set of spectrum carriers, with which RAT r can be operated, e.g., due to regulation or technological reasons. Each transceiver, t, has a set of candidate configurations, $C_e(t)$. Each

configuration, $c \in C_e(t)$, is a set (r, f), where $r \in R_e(t)$ and $f \in F_e(t,r)$. Finally, the union of the sets above readily leads to the aggregate sets of RATs R_e, carriers F_e, and configurations C_e, which can be used in e.

Additionally, this part exploits basic monitoring information for estimating traffic requirements and mobility characteristics in the service area. Set U_e denotes the set of users in the service area of e. The set of services (applications) requested is S_e. It is assumed that each user $u \in U_e$ requests a service, $s(u) \in S_e$. This model covers physical users that require more than one service. Regarding mobility, assuming a semi-stationary state has been reached, we can associate with each user u a location $l(u)$. This is the subset of the service area, of reconfigurable element e, in which u is found.

Furthermore, basic monitoring information for estimating the capabilities (capacity and coverage) of the candidate configurations are exploited. These capabilities can change over time, as they are influenced by the changing conditions in the environment, especially the behaviour of "near-by" elements. All elements act in a completely distributed (autonomic) manner. This poses a significant engineering challenge: how to increase the degree of assurance that, by assigning a certain configuration, c=(r, f), to transceiver t of e, the resulting capacity, $cp_e(c)$, and coverage, $cv_e(c)$, will be, e.g., x Mbps and y km, respectively. A probabilistic model, as well as a learning and adaptation strategy, should be adopted. The resulting problem is: *Given a specific candidate configuration c =(r, f), how can we predict the most probable values of the random variables capacity and coverage?*

Section 0 solves this problem through a robust, learning and adaptation, strategy, based on Bayesian networks. Our solution does not violate the autonomy of network elements. In fact, no cooperation (e.g., no message exchange) between the different network elements is needed. Our schemes complement legacy schemes for the sensing (initial discovery) of the performance potentials of reconfigurations.

Profiles and policies. This part describes the profiles (e.g., preferences, requirements, constraints) of user classes, applications and terminals, as well as the policies and agreements of the NO. Set $Q(s,u)$ provides the target QoS levels at which service s ($s \in S_e$) should be offered to user u ($u \in U_e$). Set $R(s,u)$ specifies the set of RATs, through which service s can be offered to user u. The provision of service s, at QoS level q, to user u, is associated with a utility volume (importance), $uv(s,q,u)$.

The output of the problem is classified into the following categories: (i) *transceivers reconfiguration*; (ii) *QoS assignment*; (iii) *traffic distribution*.

Transceiver reconfigurations. These are denoted as A_{TC}= $\{c_e(t) \mid t \in T_e\}$. Each element of the set, $c_e(t)$, is the overall reconfiguration of transceiver t, and corresponds to a pair [$r_e(t)$, $f_e(t)$]. Functions $r_e(t)$ and $f_e(t)$ are the RAT and spectrum, respectively, allocated to transceiver t. Each element of the set should respect the relation $c_e(t) \in C_e(t)$, or equivalently, $r_e(t) \in R_e(t)$ and $f_e(t) \in F_e[t, r_e(t)]$. This guarantees compliance with the element capabilities (i.e., permissible allocations of RATs and spectrum to transceivers).

QoS assignment. This is expressed through set A_{UQ}= { $q_e(u)$| $u \in U_e$ }. Function $q_e(u)$ is the QoS level that will be offered to user u, by reconfigurable element e. Each function should preserve the relation $q_e(u) \in Q(s(u), u)$, for ensuring compliance with the profiles and agreements, i.e., the provision of applications at the appropriate, acceptable QoS levels.

Traffic distribution. Set A_{TU} ={ $u_e(t)$ | $t \in T_e$ } expresses the new traffic distribution, due to the reconfiguration. Each element of A_{TU}, $u_e(t) \subseteq U_e$, includes the users that will be served by transceiver t of element e. For each $u \in u_e(t)$, the relation $r_e(t) \in R(s(u), u)$ should hold. This guarantees the provision of applications through permissible RATs, in accordance with the (terminal) profiles, NO policies and agreements. In addition, the relation $nr(c_e(t), u_e(t), A_{UQ}) \leq cp_e(c_e(t))$ guarantees that the capacity constraint of each transceiver is preserved. Function $nr(...)$ is kept flexible, in order to ensure applicability with various RATs.

Objective function. The reconfiguration decisions should optimise an objective function that consists of two main parts. The first part is targeted to the maximisation of the aggregate utility volume deriving from the A_{UQ} allocation 00. This is the quantity $\sum_{u \in U_e} uv[s(u), q_e(u), u]$. The rationale is that users should be assigned to their most preferred QoS levels, to the largest extent possible. (ii) The second part of the objective function is targeted to the minimization of the number of required changes. These changes are seen as the cost of reconfiguring the element, according to allocation A_{TC}. The rationale is that among reconfigurations that exhibit the same performance, those that require fewer changes should be preferred. The number of changes can be linked with the RAT and spectrum changes in each transceiver (e.g., see **Fig. 1**).

4. Robust Discovery of Reconfiguration Capabilities

This section presents the learning and adaptation method for robustly estimating the likelihood that the selection of a reconfiguration pattern is associated with a specific capacity and coverage.

A. Formulation through Bayesian networks

Fig. 5 depicts a Bayesian network that is proposed for modeling the specified problem. CAP and COV are random variables representing capacity and coverage, respectively. CFG is another random variable representing a configuration, i.e. a combination of RAT and spectrum allocation.

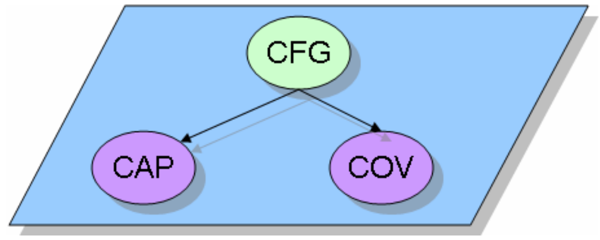


Fig. 5. Bayesian Network for RSQT

CFG is the Bayesian network's predictive attribute (node), while CAP and COV are the target attributes. The goal is the computation of the maximum value of the joint conditional probability $\Pr[CAP, COV | CFG]$. In this subsection, we show that the desired probability $\Pr[CAP, COV | CFG]$ is equivalent to the computation of the product of the conditional probabilities $\Pr[CAP | CFG]$ and $\Pr[COV | CFG]$, i.e.,

$$\Pr[CAP, COV | CFG] = \Pr[CAP | CFG] \cdot \Pr[COV | CFG] \quad (1)$$

Relation (1) means that capacity and coverage can be decoupled. Two independent conditional probability tables (CPTs) can be organized, i.e. one for capacity and one for coverage. Each CPT refers to a particular RAT. Each column of the CPT refers to a specific configuration (i.e., RAT and carrier frequency). Each line of the CPT corresponds to a capacity (or coverage) value. In this sense, a discrete set of potential capacity values is defined. Each cell (intersection of line and column) provides the probability (likelihood) that the configuration (corresponding to the column) will achieve the potential capacity value (corresponding to the line). Given a configuration, the most probable value of the capacity is the value that corresponds to the maximum conditional probability.

The proof of relation (1) starts from the joint probability formula, which suggests that:

$$\Pr[CAP, COV, CFG] = \Pr[CAP, COV | CFG] \cdot \Pr[CFG] \quad (2)$$

According to the Bayesian chain rule formula:

$$\Pr[CAP, COV, CFG] = \Pr[CFG] \cdot \Pr[COV | CFG] \cdot \Pr[CAP | COV, CFG] \quad (3)$$

Hence, from (2) and (3):

$$\Pr[CAP, COV|CFG] = \Pr[CAP|COV, CFG] \cdot \Pr[COV|CFG] \quad (4)$$

Assuming that capacity's probable values are not dependent on coverage's probable values, it can be deduced that:

$$\Pr[CAP|COV, CFG] = \Pr[CAP|CFG] \quad (5)$$

The combination of relations (5) and (4) leads to relation (1).

B. Solution: learning and adaptation

In the previous subsection, we defined that the capabilities of configurations are modeled through the CPTs. The next step is to describe how to update the CPTs. **Fig. 6** is the general representation of the process.

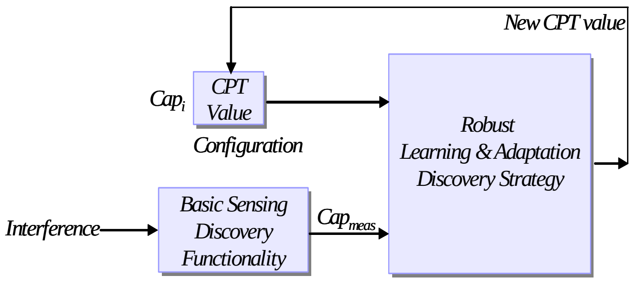


Fig. 6. Process for updating the CPT values

This learning and adaptation process yields the robust methods for discovering the performance capabilities of candidate configurations. The update process takes into account the system's measurements and, more specifically, the "distance" (absolute difference) between each probable value and the measured value.

Let us assume that measurements (obtained through basic discovery-sensing functionality) show that a specific configuration can achieve capacity cap_{meas}. This measurement can be exploited, in order to fine-tune (enhance or decrease) the values of the CPTs, so as to increase the degree of assurance of future predictions. Let $dif_{\max}$ be the maximum difference between the probable capacity values, i.e., $dif_{\max} = cap_m - cap_1$.

Then, the following correction factor, cor_i, can be computed for each candidate capacity value cap_i:

$$cor_i = 1 - \frac{|cap_i - cap_{meas}|}{dif_{\max}} \quad (6)$$

It holds that 0≤cor_i≤1. A value close to 1 reflects that the corresponding candidate value cap_i is close to the measured value cap_{meas}, thus it should be reinforced accordingly. The opposite stands for a value that is close to 0.

The correction of the $\Pr[CAP = cap_i | CFG]$ values can then be done as follows for each candidate capacity value cap_i:

$$\Pr[CAP = cap_i | CFG]_{new} = nf \cdot cor_i \cdot \Pr[CAP = cap_i | CFG]_{old}$$

(7)

The parameter nf is a normalizing constant whose value can be computed by requiring all the "new" probabilities to sum up to 1.

The system *converges* when the most probable candidate capacity value (i.e. the one with the maximum probability) is reinforced, while the probabilities of the other candidate capacity values are either reduced or reinforced less. After convergence, we limit the number of consecutive updates that can be done on the probability values associated with each capacity value. This is done for assisting fast adaptation to new conditions. For the same reason, we do not allow that a probability falls under a certain threshold, a/m, where 0<a<1 (m is the number of potential capacity values). In such cases, the normalization factor, nf, is computed by requiring all the other "new" probabilities to sum up to 1-($k \cdot a/m$), where k is the number of probabilities that are assigned equal to the threshold.

5. Selection of Reconfigurations

This section presents the second part of the solution of the RSQT problem. It exploits the potential capabilities of candidate reconfigurations. This yields a rating of reconfigurations, and eventually, leads to the selection of the best reconfigurations.

In general, this part of the solution of the RSQT problem consists of four phases, for reduction of the associated complexity.

The first phase finds different valid transceiver reconfigurations, which constitute sub-problems, to be launched in parallel. In each sub-problem, there is a certain transceiver configuration, i.e., an allocation of RATs and spectrum to transceivers. The capabilities of this configuration will be explored in the next two phases, and finally, the fourth phase will consist in the selection of the best solution

In the second phase, in each sub-problem (i.e., for each transceiver configuration found in the first phase), there is an allocation of demand to transceivers, at the basic QoS levels. The configuration explored in the sub-problem has certain capabilities in terms of capacity and coverage. Therefore, each user (demand portion) can be served by a set of transceivers. The user is allocated to the best transceiver according to the policy (e.g., selection of the one with the largest available capacity). The phase stops when all users are allocated to transceivers, or when the demand that cannot be served by the configuration exceeds a certain threshold.

In the third phase, there are attempts to improve QoS (which was assigned at the basic level in the previous phase). QoS levels are gradually augmented, in a greedy manner, starting from those that lead to a larger increase in the utility volume (and hence, users' satisfaction). The algorithm stops when no more QoS increase is possible, due to either user profiles, or available capacity.
Finally, the fourth phase includes the selection of the best configuration. The configurations have scored a certain performance in the previous two phases. This performance derives from the provision of desired QoS levels. This is reflected in utility volumes. A policy based decision is required for selecting the best configuration. An approach can be to select the configuration that requires the least changes on the already established configuration.
Essentially, the previous phases provide a rating of reconfigurations, with respect to the objective function. So the last, or fourth, phase selects the best reconfiguration, i.e., the reconfiguration with the highest objective function value.

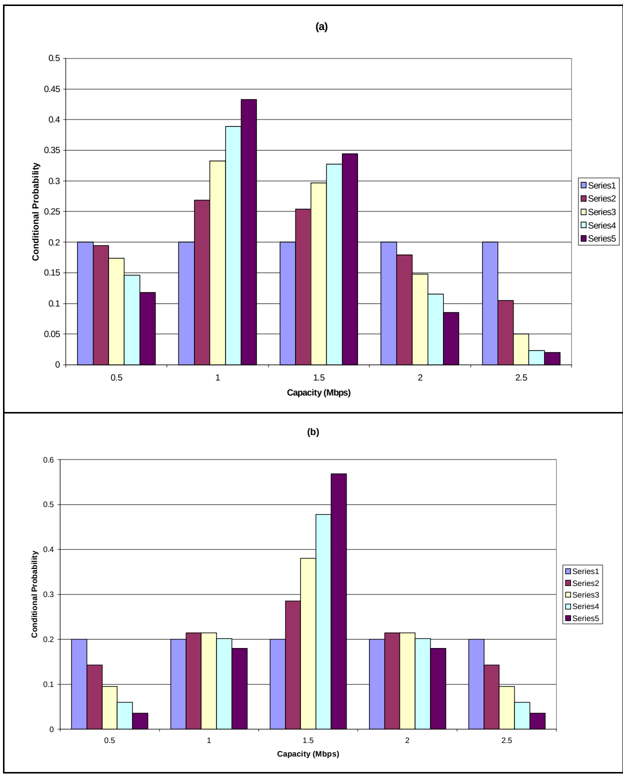
6. Results

This section presents results on the efficiency of the robust discovery and reconfiguration selection methods.

A. Robust discovery

Four scenarios are realized. The scenarios are driven by the estimation of capacity's conditional probability values. The coverage values can be determined in the same way. In all the scenarios, a has been set equal to 0.1.
Scenario 1. This scenario helps the in-depth understanding of our proposed technique. Our focus is on an arbitrary configuration $c_1=(r_1, f_1)$. It is assumed that there are $m=5$ candidate capacity values (in Mbps): $cap_1=0.5$, $cap_2=1.0$, $cap_3=1.5$, $cap_4=2.0$, $cap_5=2.5$. Hence, $dif_{\max}=2$ Mbps. It should be noted that a denser grid of candidate values could be selected (actually, in this case our results would have been favoured). What is more, the distance between two subsequent candidate values needs not be the same. Also, only three consecutive reinforcements are allowed, after convergence.

Fig. 7 (a)-(b) depicts the distribution of conditional probabilities in two cases. In each case cap_{meas} is (in Mbps) 1.2 and 1.5, respectively.



(a)

(b)

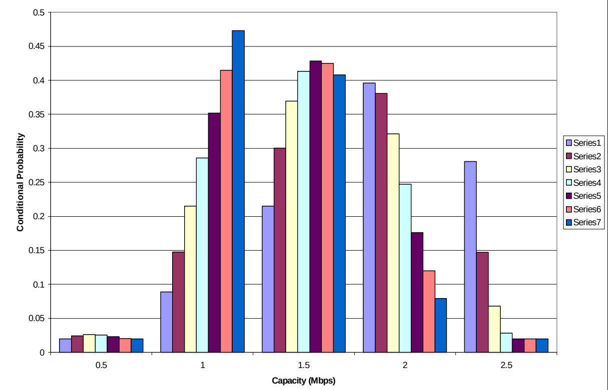
Fig. 7. Learning and adaptation of potential capabilities of arbitrary reconfiguration. Distribution of conditional probabilities corresponding to candidate capacity values, when is sensed: (a) =1.2 Mbps; (b) 1.5 Mbps.

The algorithm is applied in five series of runs. Initially, the conditional probabilities are uniformly distributed, i.e., equal to 0.2, in all scenarios (Series 1). By using (6), we calculate the correction factors. Then, by using (7), we compute the new (adjusted) conditional probabilities. The results for each case are further analyzed in the following.

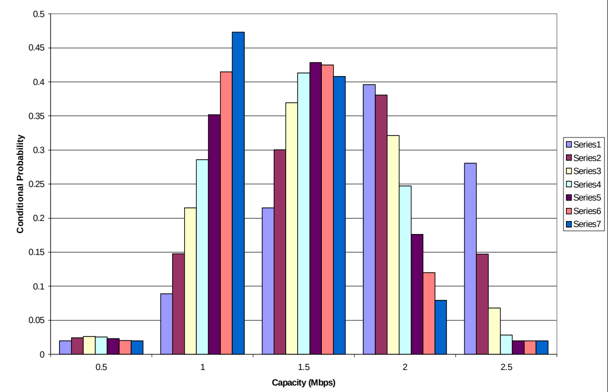
Fig. 7 shows that our model correctly and quickly adapts to the situation, by selecting cap_2 as the most probable value, in the second series. As should be done, from the beginning there are high values for cap_2 and cap_3, a slight diminishment for cap_1 and cap_4, and a severe degradation for cap_5. As the scheme is further applied, and since cap_{meas} does not change, the most probable value is actually further reinforced. Our learning and adaptation model accurately adapts to the second case also, in which cap_{meas} =1.5. Quickly, there is peak at cap_3, whereas cap_2 and cap_4 remain practically the same, and finally, cap_1 and cap_5 suffer significant diminishment.

Scenario 2. The goal of this scenario is mainly to examine how many steps it takes for the scheme to adapt to a new situation, in other words, to a sudden big change in the environment conditions. **Fig. 8** (a)-(b) shows the speed of the adaptation when there is a sudden degradation of the measured capacity. In the first figure, we allow only three consecutive reinforcements, of the most probable value, after convergence. In the second figure, only one reinforcement is allowed.

Fig. 8 (a) shows what happens when cap_{meas} suddenly becomes 1.1 Mbps. The goal is to examine how quickly the system can adapt and converge to cap_2, which is the most proper value, based on the measurements. Our starting point is when cap_{meas} is 1.1 Mbps. As can be observed in **Fig. 8** (a), in 4 steps (Series 2-5) the most probable value drops from cap_4 to cap_3. In another six steps (Series 6-11), candidate value cap_2 and cap_3 are suggested as the most likely ones. Finally, in the next step (Series 12) cap_2 becomes the most probable one.



(a)



(b)

Fig. 8. Learning and adaptation when the capacity of arbitrary reconfiguration suddenly degrades from 2.1 to 1.1 Mbps. Speed of adaptation when the number of consecutive updates after convergence is: (a) three; (b) one.

Fig. 8.(b) shows what happens if only one reinforcement is allowed after convergence. Again, our starting point is when cap_{meas} is 1.1 Mbps. In just 2 steps (Series 2-3) the most probable value drops from cap_4 to cap_3. In another 3 steps (Series 4-6), candidate value cap_2 almost reaches cap_3, and in the next step (Series 7) cap_2 becomes the most probable one.
Another conclusion can be deduced from the aforementioned simulations. The number of consecutive reinforcements after convergence clearly affects the model's adaptation speed. High number of consecutive reinforcements reduces the adaptation speed.

B. Reconfiguration selection - exploitation of reconfiguration capabilities

Three scenarios are realized. Initially, we present the input to this phase. Then, we analyze the results from the scenarios.
Context. We consider a simple service area that is covered by a network segment. The segment consists of a number of reconfigurable elements that operate in parallel. The behavior of these elements and the context information cause reconfiguration triggers to a random element, which will be in the focus of this subsection. **Fig. 9** refers to the demand in the element's service area.

	Data		***Voice***	
	%	***sessions***	***%***	***sessions***
Case1	0	0	100	260
Case2	6.5	16	93.5	228
Case3	14	32	86	196
Case4	23	48	77	164
Case5	33	64	67	132
Case6	45	80	55	100
Case7	59	96	41	68
Case8	76	112	24	36

Case9	97	128	3	4

Fig. 9. Demand in the element

Nine different cases are studied. Each case corresponds to a different traffic mix (combination of voice and data sessions). Initially, the demand for voice dominates. Gradually, the demand for the data service dominates. The demand is taken uniformly distributed within the service area.

Moreover, each element is equipped with 3 reconfigurable transceivers. Each transceiver may select between the two configurations studied in the previous sub-section. In doing so, the resulting overall configurations for each element can be denoted as e.g. (c_1, c_1, c_2), implying that two transceivers are assigned configuration c_1, while the third one is assigned configuration c_2 and so forth. Additionally, the assignment of configuration c_2 to all transceivers is not considered, since it would lead to coverage holes.

As aforementioned, the configurations c_1 and c_2 have different capacity capabilities, 1 or 2Mbps for c_1, and 4, 7 or 10Mbps for c_2. It is also assumed that c_1 can achieve larger coverage than c_2, i.e., the larger the capacity is, the smaller the coverage becomes.

Profiles and Policies. Set S_e consists of two services, a voice service (s1) and a data service (s2). Whereas the voice service is associated with a fixed quality level, for the data service, a set of quality levels is provided. Moreover, s1 can only be offered through r_1 (therefore, configuration c_1). **Fig. 10** contains the acceptable QoS levels, the utility volume, whenever a service is offered at a certain QoS level, as well as the bandwidth requirements per service.

	Voice (c1)		***Data (c1 or c2)***	
QoS levels	***Bit Rate (kbps)***	***Utility Volume***	***Bit Rate (kbps)***	***Utility Volume***
0	16	1	32	2
1			64	4
2			128	8
3			256	16

Fig. 10. Profiles and policies: services, QoS levels, utility volumes

All in all, we are able to consider various scenarios, combining the capabilities (capacity and coverage) of the configurations, in order to see which configuration fits better the traffic mixes.

Scenario 1. This scenario assumes that $cp_e(c_1)$=1 Mbps and $cp_e(c_2)$=4 Mbps. The coverage pattern for c_1 is about 1000m, and for c_2 about 500m. **Fig. 11** shows indicative results.

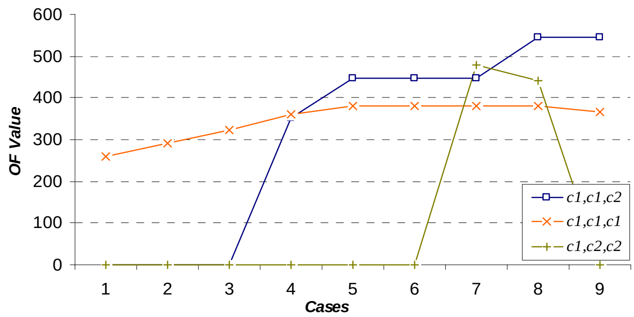


Fig. 11. Scenario 1; objective function value achieved by configurations

Configuration (c_1, c_1, c_1) increases the objective function value as the data load increases. This happens because, at the same time, voice load decreases, and therefore, there is spare capacity that can be exploited in offering higher QoS to more data sessions. At some point, the objective function value remains the same (cases 5-8), since the increase deriving from new data sessions is compensated by the decrease in the voice sessions. Finally, in case 9, the data sessions have become so many, that for some users the QoS levels offered need to be degraded, compared to case 8, and thus, the objective function value decreases.

The behavior of the (c_1, c_1, c_2) configuration is similar. Initially (cases 1-3), the configuration cannot handle the demand, because the voice load dominates and exceeds the capacity of the two c_1 transceivers. Starting from case 4, the voice sessions have decreased and can be accommodated by the two transceivers, configured with c_1. Consequently, (c_1, c_1, c_2) yields the highest objective function value. This occurs since the spare capacity is exploited for providing higher QoS to data services. Higher objective function values are achieved, compared to (c_1, c_1, c_1), because $cp_e(c_2)$ is higher. At some point the improvement stops, because the overall load is heavy, and therefore, some of the QoS levels have to be degraded again.

Finally, configuration (c_1, c_2, c_2) exhibits an acceptable performance only at certain traffic mixes. Specifically, its objective function value is initially zero, until the voice sessions can be accommodated by a single c_1 transceiver. This occurs in case 7. Then, (c_1, c_2, c_2) proves itself to be appropriate, but only until the data sessions have become far too many and cannot be catered for by c_2's limited coverage (the distribution of users within the element is uniform).
Comparing now the alternatives, we find that at the very initial demand patterns, the (c_1, c_1, c_1) configuration performs better. However, as data sessions increase, the (c_1, c_1, c_2) configuration becomes superior, due to the spare capacity that can upgrade QoS levels offered to continuously coming data sessions. This excellence of (c_1, c_1, c_2) is though sometimes marginal compared to (c_1, c_1, c_1). Additionally, at certain traffic mixes with few voice and many data sessions, (c_1, c_2, c_2) exhibits the best performance, due to its large overall capacity.
Scenario 2. This scenario assumes that $cp_e(c_1)$=1 Mbps and $cp_e(c_2)$=7 Mbps. The coverage pattern for c_1 is about 1000m, and for c_2 about 250m. **Fig. 12** shows indicative results.

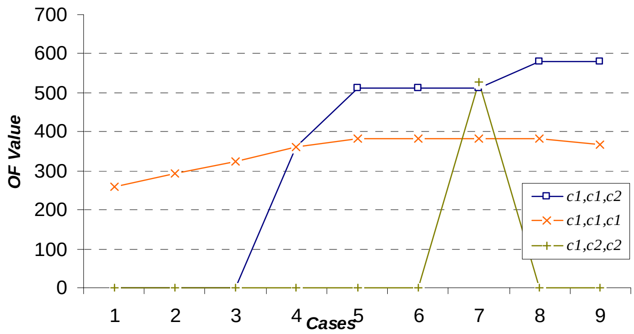


Fig. 12. Scenario 2; objective function value achieved by configurations

The study of the curve, leads to conclusions that are similar to the ones obtained in scenario 1. Regarding configuration (c_1, c_1, c_1), exactly the same results are reached, as expected, since $cp_e(c_1)$=1. So, configuration (c_1, c_1, c_1), the results are the same since $cp_e(c_1)$=1. Regarding the (c_1, c_1, c_2) configuration, initially, it is inappropriate. Starting from case 4, its performance increases, along with the increase in the data sessions, up to the point where the system reaches its "pole" capacity. However, it should be noted that this configuration gives now higher objective function values than scenario 1, due to the fact that $cp_e(c_2)$= 7 Mbps. Again,

configuration (c_1, c_2, c_2) exhibits no performance at the initial traffic loads, and is, in general, appropriate only when voice sessions have decreased enough and can be served by a single c_1 transceiver. Moreover, (c_1, c_2, c_2) provides now a higher value in case 7 than in scenario 1 (due to the current higher $cp_e(c_2)$ value), but proves itself inappropriate right after that, due to its restricted coverage capabilities that cannot cater for the uniformly distributed coming data sessions.
In general, comparing again the available configurations, we find that at the very initial demand patterns, (c_1, c_1, c_1) outperforms the rest configurations. As data sessions increase, there is superiority of the (c_1, c_1, c_2) configuration, which is important compared to (c_1, c_1, c_1). Additionally, (c_1, c_2, c_2) performs better than the rest configurations at certain traffic mixes (case 7).
Scenario 3. This scenario assumes that $cp_e(c_1)$=2 Mbps and $cp_e(c_2)$=4 Mbps. The coverage pattern for c_1 is about twice the coverage of c_2. **Fig. 13** shows indicative results.

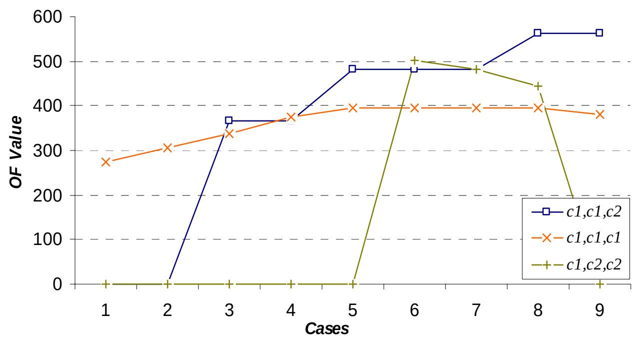


Fig. 13. Scenario 3; objective function value achieved by configurations

Examining (c_1, c_1, c_1) at a first stage, the results obtained are naturally same in rationale as in the previous scenarios. However, the fact that $cp_e(c_1)$=2 Mbps leads in general to a higher objective function value compared to the previous scenarios. Regarding (c_1, c_1, c_2), it can be put into effect also from case 3, since the larger capacity of the 2 c_1 transceivers is enough for the existing voice sessions. Finally, (c_1, c_2, c_2) still remains appropriate only in certain traffic patterns. Specifically, it could be taken into consideration only in cases 6, 7 and 8, where, on one hand, voice sessions can be efficiently served by one c_1 transceiver, whilst data sessions fall within c_2's coverage area. Much to our anticipation, this curve moves to the left in

comparison to scenarios 1 and 2 (where $cp_e(c_1)$=1 Mbps), justified by the fact that now a single c_1 transceiver can accommodate more voice sessions than before.
The comparison among the configurations examined, shows again that at the very initial demand patterns, (c_1, c_1, c_1) outperforms the rest configurations. As data sessions increase, there is a considerable superiority of the (c_1, c_1, c_2) configuration, even more important than in scenarios 1 and 2, due to the large capacity of c_1. The (c_1, c_2, c_2) configuration performs better than the rest configurations at certain traffic mixes (cases 6 and 7), since it can more efficiently guarantee the desired utility.

7. Conclusion

B3G wireless infrastructures can be efficiently realized by exploiting cognitive network concepts. Cognitive, wireless access, infrastructures dynamically reconfigure to the appropriate RATs and spectrum, in order to adapt to the environment requirements and conditions. Reconfiguration decisions call for advanced management functionality. This paper provided such management functionality by addressing a pertinent problem, called "RAT and Spectrum selection, QoS assignment and Traffic distribution" (RSQT). We formally defined and solved a fully distributed problem version. We proposed robust (stable, reliable), learning and adaptation, strategies for estimating (discovering) the performance potentials of alternate reconfigurations. We gave a computationally efficient solution to the problem of exploiting the performance potentials of reconfigurations, and presented results that expose the behaviour and efficiency of our schemes.
One of our future plans is to further employ autonomic computing principles in the direction of realizing cognitive, wireless access, infrastructures. Our goal is to develop an autonomic manager, which will encompass the RSQT scheme. The manager will consist of policies, context perception capabilities, reasoning algorithms, learning functionality and knowledge engineering, technologies for the representation of ontologies and semantics. All these will yield a system that hypothesises on causes to a problem, and subsequently, validates or falsifies the hypothesis.
Another issue for future study is to complement the distributed RSQT scheme with a second tier of, more centralised, management functionality. The centralised functionality will be invoked when the distributed components cannot converge to acceptable solutions. The synergy of the two tiers will guarantee that whenever the distributed components diverge from the near-optimal performance levels, the application of the second tier will restore the performance to the desired levels.
Another issue for further study is to exploit the RSQT scheme for enabling NOs to personalise their service offerings, instead of limiting subscribers to a fixed set of inflexible choices. Seamless mobility applications can build on schemes like RSQT to

intelligently change the services that they provide based on business policies and context.

References

1. Third (3rd) Generation Partnership Project (3GPP), Web site, www.3gpp.org , 2006.
2. Institute of Electrical and Electronics Engineers (IEEE), 802 standards, Web site, www.ieee802.org, 2006.
3. Wireless World Research Forum (WWRF), www.wireless-world-research.org, 2006.
4. Internet Engineering Task Force (IETF), www.ietf.org, 2005.
5. FP6/IST project E2R (End-to-End Reconfigurability), www.e2r.motlabs.com, 2006.
6. 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.
7. FP5/IST project MONASIDRE (Management of Networks and Services in a Diversified Radio Environment), www.monasidre.com, 2003.
8. 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, 2002
9. P.Demestichas, V.Stavroulaki, L.Papadopoulou, A.Vasilakos, M.Theologou, "Service configuration and traffic distribution in composite radio environments", IEEE Transactions on Systems, Man and Cybernetics Journal, vol. 33, No. 4, pp. 69-81, Nov. 2003
10. J. Kephart, D. Chess, "The vision of autonomic computing", IEEE Computer, Vol. 36, No.1, pp. 41-50, January 2003.
11. P.Demestichas, D.Boscovic, V.Stavroulaki, Al Lee, J.Strassner, "ATLAS: Autonomic management platform for seamless cognitive wireless connectivity", *IEEE Commun. Mag*, June 2006.
12. J. Strassner, "Policy-based network management", Morgan Kaufmann Publishers. U.S.A., 2005.
13. J.Strassner, "Autonomics – A critical and innovative component of seamless mobility", Technical Report, http://www.motorola.com/mot/doc/5/5978_MotDoc.pdf, Motorola, Dec. 2005.
14. M. A. L. Thathachar, P. S. Sastry, "Networks of Learning Automata", Kluwer Acad. Publishers, 2004.
15. A.Mas-Colell, "Microeconomics", Oxford University Press, U.K., 1995
16. J.Tirole, "The Theory of industrial organization", MIT Press, Cambridge, Mass., U.S.A., 1998 G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.