Management Strategies for Distributed Cross-Layer Reconfigurations in the Context of Cognitive, B3G Infrastructures

G. Dimitrakopoulos, K. Tsagkaris, E. Adamopoulou, K. Demestichas, P. Demestichas

University of Piraeus

Department of Digital Systems

80 Karaoli Dimitriou str., 18534, Piraeus, GREECE, e-mail: gdimitra@unipi.gr

***Abstract*—B3G (Beyond the 3rd Generation) wireless infrastructures are increasingly aligned with cognitive networking principles. Cognitive networks dispose mechanisms for dynamically selecting their configuration (algorithms and parameter values, at different layers of the protocol stack), through appropriate management functionality that takes into account the context of operation, profiles, goals and policies. This paper focuses on such management functionality by addressing a pertinent problem, dealing with "Distributed, Cross-Layer Reconfigurations" (DCLR). Our work contributes in four main areas. First, we formally define and solve a fully distributed problem version, which is very important for the management of a particular reconfigurable element, in a cognitive context. Second, we propose robust (stable, reliable), learning and adaptation, strategies for estimating (discovering) the performance potentials of alternate reconfigurations. Third, we give a computationally efficient solution to the problem of exploiting the performance potentials of reconfigurations, rating reconfigurations, and finally, selecting the best ones. Finally, we present results that expose the behavior of our schemes.**

***Index Terms*— Cognitive networks, Cross-layer optimization, Utility, Learning and Adaptation, Bayesian networks**

I. Introduction

IMMENSE research and development effort is being dedicated to the development of new wireless networking technologies, in order to deliver powerful and affordable, high-speed, wireless access solutions. Currently, the wireless access landscape includes a multitude of technologies available to the mean user. Moreover, the wireless world is migrating towards the era of B3G (Beyond the 3rd Generation) [1] wireless access communications. The motivation is to increase the exploitation of the available technologies. The main idea is that a network operator (NO) can rely on different radio access technologies (RATs), for achieving the required capacity and QoS (Quality of Service) levels, in a cost efficient manner. The B3G concept can be realized through cognitive (adaptive, reconfigurable) networking potentials [2][3][4][5][6]. Cognitive networks, reactively or proactively, adapt to the environment requisitions, by means of self-configuration. Self-configuration is applied, for tackling complexity and scalability. Reconfiguration may affect all layers of the protocol stack. Specifically, 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 this respect, it is believed that cognitive networks enable the realization of B3G infrastructures with reduced capital expenditures (CAPEX).

The realization of cognitive, wireless access, networks requires advanced management functionality, which will be in charge of finding the best reconfigurations. This paper provides such management functionality, by addressing an important problem for the management of a reconfigurable network element, which operates and is managed in parallel with other elements. The problem is called "Distributed, Cross Layer Reconfigurations" (DCLR). Capabilities are exploited in the provision of the highest possible QoS levels, at the appropriate capacity levels. This exploitation yields a rating of the candidate reconfigurations, and leads to the selection of the best one.

The organization of the paper is as follows. Section II presents the motivation for this work and describes the overall context of the DCLR problem. Section III defines the DCLR problem, and outlines the two main parts of its solution. Sections IV and V describe the two solution parts in detail. Section VI provides results that show the behavior and efficiency of our schemes, and section VII includes concluding remarks.

II. Motivation and High Level Problem Description

In general, cognitive systems determine their behavior, in a reactive or proactive manner, based on external stimuli, goals, principles, capabilities and experience. In the case of cognitive networks, this definition can be translated as the capability to dynamically select the network's configuration, through appropriate management functionality that takes into account the context of operation, profiles, goals and policies, and machine learning.

As can be deduced from the definitions above, cognitive networks consist of reconfigurable platforms and management functionality. The role of reconfigurable platforms is to enable the dynamic selection of the appropriate configurations. As can be shown in Fig. 1, a cognitive element, according to requirements, is based on a reconfigurable platform and thus may (i) change the RAT it operates with and maintain the spectrum, (ii) maintain the RAT and change the spectrum and (iii) change both, RAT and spectrum.

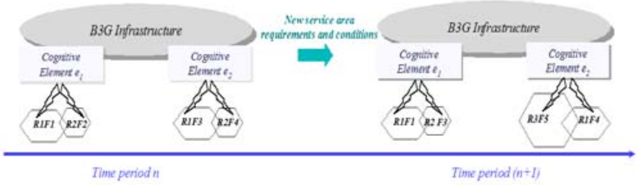

Fig. 1: Role of reconfigurable platforms

In the light of the above, each element may be multi-standard. Only a subset of technologies is used, namely those that are most appropriate for the context of operation. Specifically, each element is controlled by management functionality that has to solve a problem of cross-layer flavor. The functionality helps the network to continuously observe the environment, looking for potential changes that can affect its operation. Observations form the basis for initiating machine-based analysis (reasoning) to see if the reconfiguration process should be invoked. Once the decision is taken, the network acts accordingly. This loop [6] can be augmented by a machine learning process [7], which leads to cognition. The loop is guided by a set of policies and goals. Fig. 2 provides the overall description of the management functionality (DCLR) proposed in this paper.

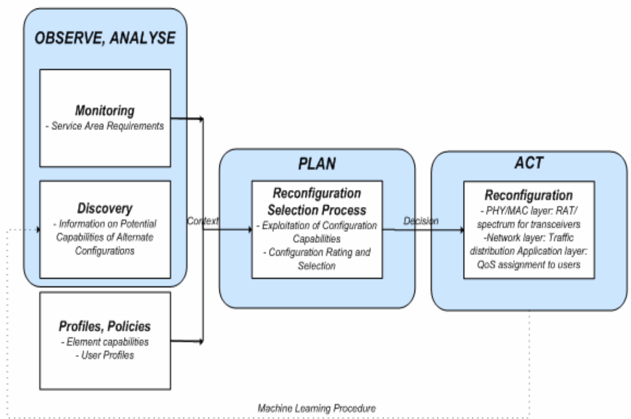

Fig. 2: The DCLR problem

The proposed management mechanisms undertake decisions that affect the protocol stack in a cross-layer fashion. The next section, accordingly, describes the DCLR problem in detail with the input and the output (decisions for element reconfiguration); sections 4 and 5 present the solution.

III. DCLR Problem Statement

A. DCLR input

The input is classified in three main categories: (i) monitoring, (ii) discovery, and (iii) profiles.

Monitoring. This part gathers monitoring information for estimating traffic and mobility characteristics in the service area. This includes the users and the services requested. Regarding mobility, assuming a semi-stationary state has been reached, we can associate each user with a location.

Discovery. This part exploits basic monitoring information for estimating the capabilities (achievable bit rate) of the candidate configurations. These capabilities can change over time, as they are influenced by the changing conditions in the environment, especially the behavior of "near-by" elements. This poses a significant engineering challenge: how to increase the degree of assurance that, by assigning a certain configuration c (combination of RAT r and frequency f) to a transceiver of an arbitrary element, the resulting maximum achievable bit rate will be, e.g., x Mbps and y km, respectively. Section IV solves this problem through a strategy based on Bayesian networks [8].

Profiles. This part provides information on the candidate configurations of the element (element profiles), such as the set of transceivers of the element, the set of operating RATs, as well as the set of spectrum carriers. Moreover, this part also describes the profiles (e.g., preferences, requirements, constraints) of user classes, applications and terminals, as well as the policies and agreements of the NO.

B. DCLR output

Reconfiguration decisions comprise: (i) transceiver reconfigurations; (ii) QoS assignment; (iii) traffic distribution.

Transceiver reconfigurations. These denote the allocation of a certain configuration to each transceiver.

QoS assignment. This is associated with the allocation of applications to QoS levels.

Traffic distribution. Finally, users are allocated to transceivers. This is aligned with the provision of applications through permissible RATs, in accordance with the (terminal) profiles, NO policies and agreements.

Objective function. The reconfiguration decisions should optimize an objective function that consists of two main parts. (i) The first part is targeted to the maximization of the aggregate utility volume. 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. The rationale is the preferred reconfigurations are those that require fewer changes should be preferred.

IV. Robust Discovery of Reconfiguration Capabilities

This section presents the learning and adaptation method for estimating the likelihood that reconfiguration c is associated with achievable bit rate $br_e(c)$.

A. Formulation through Bayesian networks

Let CFG be a random variable, representing the configuration that is probed, and BR be another random variable representing a configuration's capability, namely, the achievable bit-rate. It can be assumed that CFG is the Bayesian network's predictive attribute (node), while BR is the target attribute. The method relies on the constant update and maintenance of conditional probability values of the form $\Pr\left[BR \middle| CFG\right]$.

A conditional probability table (CPT) can, therefore, be organized. Fig. 3 depicts the structure of the CPT. Each column of the 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 an achievable bit-rate value. Note that a discrete set of m reference bit-rate values can be defined (also for reducing complexity 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 $\Pr\left[BR{=}br_j \middle| CFG{=}cfg_i\right]$. It expresses the probability that bit-rate br_j will be achieved, given that configuration cfg_i is selected.

	…	cfg_i	…
br_1	…	$\Pr\left[BR{=}br_1 \middle| CFG{=}cfg_i\right]$	…
br_2	…	$\Pr\left[BR{=}br_2 \middle| CFG{=}cfg_i\right]$	…
…		…	
br_j	…	$\Pr\left[BR{=}br_j \middle| CFG{=}cfg_i\right]$	…
…		…	
br_m	…	$\Pr\left[BR{=}br_m \middle| CFG{=}cfg_i\right]$	…

Fig. 3: General CPT structure

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. Configuration and bit rate pairs with high probabilities can be in the top of the list, in order to enable fast configuration selections.

B. Solution: Learning and Adaptation Strategy

The capabilities of configurations are provided by the CPT. The next step is to describe how to update the CPT, taking into account the measurements (environment sensing) of the cognitive radio system and, more specifically, the "distance" (absolute difference) between each reference value and the measured value.

Let us assume that an environment sensing shows that a specific configuration can achieve bit-rate br_{meas}. This measurement 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. Let $dif_{\max}$ be the maximum difference between the reference bit-rate values, i.e. $dif_{\max} = br_m - br_1$.

Then, the following correction factor, cor_j, can be computed for each reference achievable bit-rate value br_j:

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

It holds that $0 \le cor_j \le 1$. A value close to one reflects that the corresponding reference value br_j is close to the measured value br_{meas}, thus the corresponding conditional probability value should be reinforced accordingly. The opposite stands for a value that is close to zero. Given a candidate configuration cfg_i, the correction of the CPT values can then be done as follows, for each candidate value br_j:

$$\Pr\left[BR = br_j \middle| CFG = cfg_i\right]_{new} = L \cdot cor_j \cdot \Pr\left[BR = br_j \middle| CFG = cfg_i\right]_{old} \quad (2)$$

Parameter L is a normalizing factor that guarantees that all the "new" probabilities sum up to one. The system *converges* when the conditional probability of the reference value, which is closest to the measured value, becomes the highest. At this point, the probabilities of the other candidate, reference values are either being reduced or reinforced less. After convergence, there can be a limit on the number of consecutive updates that can be applied on the conditional probabilities. This will assist fast adaptation to change in conditions. For the same reason, the minimum probability of a reference bit-rate value may not fall under a certain threshold, a/m, where $0 \le a \le 1$ (m is the number of reference bit-rates). In such cases, the normalization factor, L, 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 minimum threshold.

V. SELECTION OF RECONFIGURATIONS

This section exploits the potential capabilities of candidate reconfigurations. This yields a rating of reconfigurations, and eventually, leads to the best reconfigurations. In general, this part of the solution of the DCLR problem consists of four phases (Fig. 4).

The first phase finds different valid transceiver reconfigurations, each one constituting a sub-problem, subject to parallel processing.

In the second phase, in each of the sub-problems, the demand distribution, is computed. At this phase, the QoS levels offered to users are kept to their lowest acceptable values. If a solution cannot be provided under these conditions, the reconfiguration is rejected.

Then, in the third phase, the QoS level offered to users is gradually improved, until either no further increase is possible (users are at the maximum QoS level) or there is no more capacity available.

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.

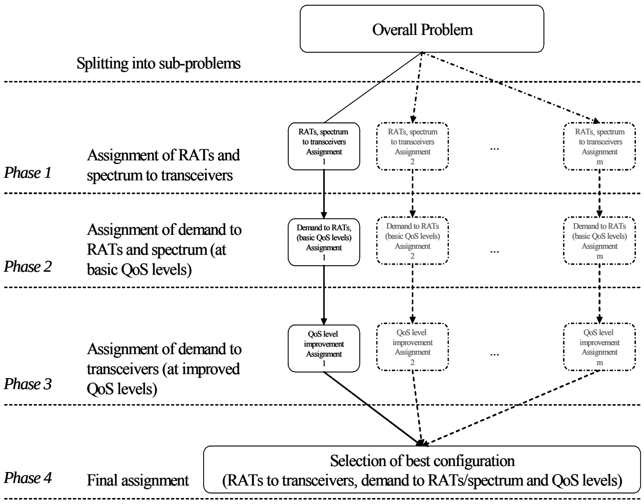

Fig. 4: Strategy for the solution of the problem

VI. RESULTS

A. Robust Discovery

An indicative scenario is realized, driven by the estimation of the conditional probability values. In all scenarios, parameter a has been set equal to 0.1. The selected 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 *BR* values (in Mbps): br_1 =0.5, br_2 =1.0, br_3 =1.5, br_4 =2.0, br_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 favored). 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. 5 depicts the distribution of conditional probabilities in four cases. In each case br_{meas} is (in Mbps) 1.2, 1.5, 2.1 and 0.75, respectively. The algorithm is applied in five steps. Initially, the conditional probabilities are uniformly distributed, i.e. equal to 0.2 (step 1). We calculate the correction factors and then we compute the new (adjusted) conditional probabilities. The results are further analyzed in the following. Fig. 5 shows that our model correctly and quickly adapts to the situation, by selecting br_2 as the most probable value, in the second step. As should be done, from the beginning there are high values for br_2 and br_3, a slight diminishment for br_1 and br_4, and a severe degradation for br_5. As the scheme is further applied, and since br_{meas} does not change, the most probable value is actually further reinforced.

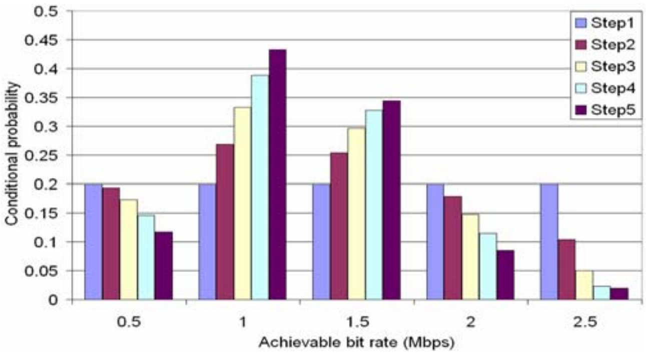

Fig. 5: Robust learning and adaptation - results

B. Reconfigurations Selection

A simple service area is covered by a network segment. The segment consists of a number of reconfigurable elements. Elements operate in parallel. The behavior of these elements and the service area requirements cause reconfiguration triggers, which will be in the focus of this subsection. The demand in the element's service area includes nine different cases studied, each one corresponding 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.

Each element is equipped with 3 reconfigurable transceivers. Each transceiver may select between two configurations. 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 bit rate 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.

Finally, two services are available, 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 configuration c_1.

All in all, we are able to make scenarios, combining the capabilities of the configurations, in order to see which configuration fits better the traffic mixes. Such a scenario would assume that $br_e(c_1)$=1 Mbps and $br_e(c_2)$=4 Mbps. The coverage pattern for c_1 is about 1000m, and for c_2 about 500m. Fig. 6 shows indicative results.

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.

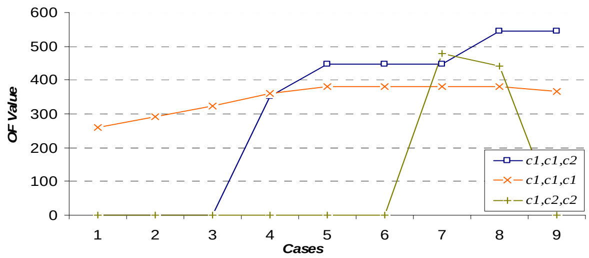

Fig. 6: Reconfigurations Selection - Results

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. 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 the objective function value decreases.

The behaviour 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.

VII. Conclusions

B3G infrastructures can be efficiently realized by exploiting cognitive networking potentials. Cognitive, wireless access, infrastructures dynamically reconfigure to the appropriate RATs and spectrum, in order to adapt to the environment conditions and requirements. This is achieved by disposing reconfigurable platforms, controlled and supported by advanced management functionality. This paper provided such management functionality by addressing a problem, defined and solved by proposing robust learning and adaptation strategies for estimating the performance potentials of alternate reconfigurations. We gave an efficient solution to the problem of exploiting those potentials, and presented results that expose the 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 DCLR 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. Another issue for future study is to complement the distributed DCLR scheme with a second tier of, more centralized, management functionality, 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 2nd tier will restore the performance to the desired levels.

Acknowledgment

This work has been performed in the framework of the European funded project E^{2}R II. The project has received research funding from the Community's Sixth Framework programme. The authors would like to acknowledge the contributions of their colleagues from the E^{2}R II consortium. The views expressed in this document do not necessarily represent the views from the complete E^{2}R II consortium. The Community is not liable for any use that may be made of the information contained therein.

References

[1] Wireless World Research Forum (WWRF), www.wireless-world-research.org, 2006.

[2] FP6/IST project E2R (End-to-End Reconfigurability), www.e2r.motlabs.com, 2005

[3] 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

[4] P. Demestichas, G. Dimitrakopoulos, K. Moessner, T. Dodgson, D. Bourse (editors – ch.8: Reconfigurability), "Technologies for the Wireless Future: Wireless World Research Forum (WWRF), Vol. 2", J. Wiley and Sons, 2006, ISBN: 0-470-02905-6.

[5] J. Mitola III, G. Maguire Jr. "Cognitive radio: making software radios more personal", IEEE Personal Commun., Vol. 6, No 4 , pp13 -18, August 1999

[6] R.W.Thomas, L.A. DaSilva, A.B. MacKenzie, "Cognitive networks", IEEE DySPAN 2005, First Symposium on Dynamic Spectrum Access Networks, USA, Nov. 2005.

[7] M. A. L. Thathachar, P. S. Sastry, "Networks of Learning Automata", Kluwer Acad. Publishers, 2004.

[8] R.E.Neapolitan, "Learning Bayesian Networks – Series in Artificial Intelligence" , Prentice Hall, 2002.