Computer Networks 54 (2010) 2117–2136

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

SUTIL – Network selection based on utility function and integer linear programming

Luci Pirmez [a], Jaime C. Carvalho Jr. [a,*], Flávia C. Delicato [b], Fábio Protti [a], Luiz F.R.C. Carmo [c], Paulo F. Pires [b], Marcos Pirmez [a]

[a] NCE/IM, Federal University of Rio de Janeiro, Postal Code 2324, Rio de Janeiro, RJ 20001-970, Brazil
[b] DIMAp, Federal University of Rio Grande do Norte, Postal Code 1524, Natal, RN 59.072-970, Brazil
[c] Inmetro, Av. Nossa Senhora das Graças, 50, Duque de Caxias, RJ 25250-020, Brazil

A R T I C L E   I N F O

Article history:
Received 1 May 2009
Received in revised form 12 March 2010
Accepted 15 March 2010
Available online 19 March 2010
Responsible Editor: T. Melodia

Keywords:
Wireless networks
Mobility management
Network selection
Seamless connectivity
Utility function
Integer linear programming

A B S T R A C T

This work presents SUTIL, a mechanism for network selection in the context of next generation networks (NGN). SUTIL selection mechanism prioritizes networks with higher relevance to the application and lower energy consumption and it enables full and seamless connectivity to mobile user devices and applications. Consequently, SUTIL contributes to realize the vision of ubiquitous computing, in which services, devices, and sensor-enriched environments interact anytime, anywhere to accomplish human designed tasks. The provided solution is based on utility function and integer linear programming and it aims at: (i) maximizing the user satisfaction while meeting application QoS and (ii) minimizing the energy consumption of devices when connecting to a target network. The solution is global since it considers for a given base station all devices that are simultaneously candidate for handoff. Simulation results showed the benefits of SUTIL usage in NGN environments.

1. Introduction

The technological advancement of wireless networks, the recent proliferation of portable devices such as laptops, palmtops and PDAs, added to the growing popularity of mobile computing, enabled realizing the scenario of ubiquitous systems, as envisioned by Weiser [1]. Ubiquitous systems are environments where devices (stations), software agents and services are integrated in a transparent and non-invasive way and work together giving support to human being activities in any time and any place. In general, these systems include sensors-instrumented environments, human–computer interfaces endowed with personalized mechanisms, and they strongly need that participant stations are connected through wireless networks. One of the major premises for ubiquitous systems is that all stations (mobiles or not) situated in any location and running user applications should be permanently connected to networks, which can be compliant to different technologies. A transparent integration of different technologies of wireless networks, particularly of WIFI, Bluetooth, GSM 3G, WIMAX and satellite networks, as well as wired networks, and a creation of a unified environment of networks and services, characterize the next generation networks (NGN) [21].

Typically, a wireless network includes at least one base station (BS) which serves a geographic area known as a cell, and one or more client stations. Each base station is the central node of a cell, and coordinates communication between all the client stations located in such cell. In general, a wireless network is composed of different cells, and has several base stations connected to each other by a backbone, providing connectivity to thousands of users, fixed or mobile. When a mobile user moves away from a base station, and gets closer to another one, he/she finishes

* Corresponding author. Tel.: +55 21 2533 6271.
E-mail address: jaimeccesarjr@gmail.com (J.C. Carvalho).

1389-1286/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2010.03.007

his/her connection with the farther BS and establishes a new connection with the closer BS. The transition of a mobile station (MS) from one BS to another one is called *handoff* or *handover* process. Ideally, the handover process must (i) be completely transparent for mobile users and (ii) guarantee that all multimedia services would not be interrupted during the process. Additionally, the process of station transition could involve only networks of the same technology (called *horizontal handoff*) or networks with different technologies, for example, from a IEEE 802.16 network to a IEEE 802.11 network (called *vertical handoff*). Choosing the best network to connect is a major challenge, due to the large number of criteria that need to be considered in the decision making. Identifying these criteria of decision is one of the main goals for achieving seamless mobility.

One of the challenges involved in NGN environments, that is "always be connected to the best network (ABC – always best connected)" [2] requires the development of solutions that consider several criteria in order to take a decision on which one would be "the best network" whenever it is needed to evaluate the transition between different cells and network operators. Examples of these criteria are: priority of service, preference and user profile, application context, signal strength, network requirements, device (station) requirements, and corporate or network provider policies. Furthermore, in order to encompass these criteria for taking decisions looking for the "best network", it is important to consider the energy consumption of mobile stations, since they are often battery-powered. Hence, a basic requirement in any process of decision for selecting the best network is to consider networks that demand a lower energy cost for stations connected to it. After a network selection, the actual process of station transition from one wireless network (cell) to the other one takes place.

The mechanism for selecting the "best network(s)" in NGN environments becomes complex due to two main factors: (i) the possibility of coexistence of different technologies for heterogeneous access; and (ii) the need of balancing users requirements with the load imposed on the selected network. In this work, we propose SUTIL, a mechanism for network selection in NGN environments that deals with such factors.

The solution provided by SUTIL is global, in the sense that, even though it considers users requirements and individual interests, it also considers the load imposed on the network, resulting from the handoff process of a set of mobile users connected to the same BS, and simultaneously candidate for handoff. SUTIL mechanism was designed to run in each BS, meaning that we have one instance of SUTIL per BS. The algorithm and protocol responsible for addressing the interaction among different instances of SUTIL are out of scope of this work. Moreover, it is important to emphasize that SUTIL addresses only the task of network selection. After the best network is selected by SUTIL, the next challenging step is to determine the right moment to trigger the handoff process. This step is out of scope of our work as well, but there are several proposals addressing this issue, such as [22].

The goals of SUTIL are threefold: (i) to maximize the satisfaction of users located in the same BS and candidate for handoff, meeting the requirements of their services in use, (ii) at the same time, to ensure the efficient use of networks resources, and (iii) to minimize the energy costs of the station when performing a handoff. Thus, SUTIL mechanism seeks to ensure that there is a fair distribution of resources (load balancing) requested by the user services among the selected networks and, at the same time, to ensure that requirements of the services would be attended by these networks.

In order to SUTIL properly operates, we assume that BSs that are candidate for a handoff process can cooperate to exchange high-level information, which describes the BSs capabilities. This hypothesis is also adopted in other works, such as in Chalmers et al. [9]. According to this assumption, whenever one of the several network providers involved in a handoff process is not willing to collaboratively exchange information, SUTIL would not work. While assuring this is a challenge, we argue that this kind of cooperation is an inevitable trend in the development of NGN. There already exist initiatives towards such cooperation. For instance, in Akyildiz et al. [17], the authors propose architecture for ubiquitous mobile communications (AMC). AMC integrates heterogeneous wireless technologies (GPRS, cdma2000, UMTS, WLAN, etc.) using two entities: a network inter-operating agent (NIA) and an interworking gateway (IG). NIA resides in the Internet and acts as a third-party eliminating the need for direct SLAs among different network operators and thus contributing to facilitate their cooperation. We argue that it is likely that the existence of mechanisms as SUTIL, that provide network selection in the context of NGNs and take into account both the networks and the users/application interests, will further motivate operators to establish agreements and upgrade their systems to incorporate them.

The selection process of SUTIL is subdivided in three phases. Considering that there is a set of MSs ready to make a handoff and a set of BSs candidate to receive such MSs, the first phase encompasses two activities: (i) verifying if all candidate networks are able to meet the requirements of all services in use in each MS of the set and, whenever a network is not able to meet some of the MSs, verifying if there is a less demanding MS in the set to assign such network (the pair user-network defined in this way is precluded from further steps of SUTIL; of course, if some network is not able to meet any user at all, it is removed from the handoff decision process); (ii) building representative services. A representative service (RS) is defined as a hypothetic service used to represent all services in use by one specific mobile user. Therefore, one RS is represented by a set of representative requirements (RRs) where each RR is calculated to attend this specific requirement for all the services in use.

In the second phase of SUTIL operation, one utility value is assigned to each network selected in the previous phase, regarding each user candidate for handoff and located in a specific BS. A utility function is defined to establish "how promising" a connection of a specific user with a target network is. It is important to emphasize that, for each user, the selected network at the end of the phase should simultaneously meet all the services in use.