Traffic Steering Framework for Mobile-Assisted Resource Management in Heterogeneous Networks

Dogadaev, Anton Konstantinovich; Checko, Aleksandra; Popovska Avramova, Andrijana; Zakrzewska, Anna; Yan, Ying; Ruepp, Sarah Renée; Berger, Michael Stübert; Dittmann, Lars; Christiansen, Henrik Lehrmann

Published in:
Proceedings of The Ninth International Conference on Wireless and Mobile Communications (ICWMC 2013)

Publication date:
2013

Link back to DTU Orbit

Citation (APA):
Traffic Steering Framework for Mobile-Assisted Resource Management in Heterogeneous Networks

Anton Dogadaev*, Aleksandra Checko†, Andrijana Popovska Avramova*, Anna Zakrzewska*, Ying Yan*, Sarah Ruepp*, Michael S. Berger*, Lars Dittmann* and Henrik Christiansen†
*Department of Photonics Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark
†MTI Radiocomp, Hillersd, Denmark
Email: {akdo, amch, apop, azak, yiya, srru, msbe, ladit, hlch}@fotonik.dtu.dk

Abstract—With the expected growth of mobile data traffic it is essential to manage the network resources efficiently. In order to undertake this challenge, we propose a framework for network-centric, mobile-assisted resource management, which facilitates traffic offloading from mobile network to Wi-Fi or open access small cells. A provision of desired quality of experience to the end-user is carried out by an operator-configurable monitoring application that is running on a mobile device. A potential to enhance network-centric resource management is provided by delegating traffic steering authorities to the network backbone. What is more, we give an overview of existing standardization activities on offloading the mobile traffic through Wi-Fi.

Keywords—Heterogenous Networks; Network Centric; Resource Management; Offloading

I. INTRODUCTION

Cisco forecasts that mobile data traffic will grow 13-fold from 2012 to 2017. As the 4G networks are generating 19 times more traffic than their predecessors [1], it is essential to maintain the network in a way that ensures certain Quality of Experience (QoE) level. In Long Term Evolution (LTE) networks, the load on the network has a crucial impact on achievable user data rates. Improving the network capacity by introducing additional cells challenges interference management. The exponential growth in mobile data traffic also puts additional load on the backhaul network, which becomes congested and requires considerable investments. Therefore, mobile operators would like to use, e.g., the Wi-Fi networks, or open small cells whenever possible to offload the traffic through them.

In this work, we describe a network-centric resource management framework where a mobile facilitates the networks' decisions by gathering feedback and requesting its demands. However, it is the network that makes the final decisions in order to ensure the optimum network operation, in terms of network offloading and power saving in the first place. For the mobile operators, who are interested in reducing their investments on the licensed spectrum and backhaul upgrading, our solution will enable that by facilitating the traffic offload. Users benefit from this approach, as the mechanisms for the network selection are enhanced, therefore connecting over a less loaded cellular network will improve their QoE and reduce the price of connection. Reduction of the traffic cost is especially important, because while data connections in a home country are relatively cheap, in roaming the price becomes considerable.

In the sections below, the phone, working according to our framework, will be called as a mobile device with "smarter" application on it. Such a "smarter" phone will seamlessly use, e.g., Wi-Fi or open access small cell for offload as shown in the Figure 1.

What is more, our platform will enhance possibility of ubiquitous Internet access, complementing services offering global connectivity from one account, like the iPass [2].

The remainder of this paper is organized as follows. In Section II, we state the challenges of resource management and load balancing in heterogeneous networks. Section III introduces related standardization and research work towards enhancement of handover across the standards, traffic offloading, interworking and network selection. In Section IV, we provide definition of our framework from the network-centric point of view. Section V gives detailed overview of proposed framework architecture and covers its extension possibilities by employing IEEE 802.11u and IEEE 802.21. Section VI concludes the paper.

II. RESOURCE MANAGEMENT AND TRAFFIC STEERING ISSUES IN HETEROGENEOUS NETWORKS

LTE has been deployed around the world since the end of 2009. For example in the big European cities, where coverage is assured, a lot of people are subscribing to the mobile access through LTE. It is enabled by a growing popularity of the user devices supporting the standard. In the same time the QoE is decreasing as more users share the capacity of the network.

A Wi-Fi connection is a cost efficient way to gain access to the Internet, thanks to its widely spread coverage and availability of the devices that support it. Mobile operators
could use the ubiquitous Wi-Fi coverage to ensure the mobile service accessibility, especially indoors. The opposite solution is also present in the market, where the Wi-Fi routers are connected to the Internet through LTE. It is a clear evidence of the convergence of those two standards, however taking into account the additional load on the mobile infrastructure that it imposes, it should be used carefully and only when needed.

Users are also getting used to the always-on connectivity over flat rate subscription and relying on their smartphones. When they are traveling abroad, the roaming data rates are still considerable. Seamless offloading the traffic to local Wi-Fi hotspots is a convenient solution to this problem. That is the reason for the popularity of the services like iPass, which cover a wide range of Wi-Fi hotspots. However, their coverage could be improved; they also require user to get and configure the client.

Nowadays widely-employed 3G/4G cellular networks are lacking the possibility of traffic offloading and interconnecting with other access networks that may exist in the area. In particular, Wi-Fi networks might be employed for interworking with standards for high-speed mobile communications like LTE/LTE-Advanced (LTE-A). Deployment of this functionality will sufficiently improve traffic load balancing. However, currently existing Wi-Fi networks lack the possibility of global access and authentication. In order to settle this problems ideas of Global Wi-Fi networks and their interworking with the cellular networks received increasing interest in scientific and standardization community.

In today’s mobile devices, a lot of control is still left for the user - he/she needs to know how to configure a smartphone or any other network device in order to use Wi-Fi instead of cellular connection whenever possible. What is more, users very often do not have the possibility to configure certain operations to be scheduled when they are connected through Wi-Fi, e.g., to update the installed applications, to backup data from the phone, etc. The mobile devices should require as little as possible interaction from the users, while on the other hand be able to provide (assist) the network with sufficient information which will ensure that the user’s expectations of service quality are met.

Mobile network coverage is deployed using macro cells. In densely populated areas, where high peak data rates are expected, small cells can boost the overall network capacity by introducing the additional resources. Small cells can be of micro-, pico- or femtocell size. Some of them can be for enterprise use, but also an open access small cells can be introduced. In both cases offloading the traffic from macro-to small cell will equalize the average cell load, improving the peak achievable data rates.

III. OVERVIEW OF STANDARDIZATION EFFORTS AND RELATED WORK

A lot of standardization efforts have been put to prepare the standards that ensure a seamless vertical handover between wireless technologies. In the tight coupling architecture among heterogeneous networks, a seamless handover can be achieved. The enhanced Generic Access Network (eGAN) [3], [4], a Third Generation Partnership Project (3GPP) standard, is an example of a tight coupling approach. In eGAN, the mechanisms for exchanging inter-Radio Access Technology (RAT) measurements are defined which enable well timed inter-RAT handover. Due to the fact that tight coupling approach is expensive, complex and requires modification in the terminals and access networks, a loose coupling approach is preferred. The Interworking WLAN (I-WLAN) [5] is a 3GPP standard for loose coupling interworking between WLAN and 3GPP networks. The Mobile IP is used as the mobility protocol, but there are no mechanisms defined for timely triggering of the vertical handover.

The Media Independent Handover (MIH) [6], an IEEE 802.21 standard, facilitates the vertical handover using a loose coupling approach. It provides mechanisms for the make-before-break handover by defining an initialization and preparation phases prior to the execution of a handover. In the initialization phase the current link is observed and new access networks are detected. In the preparation phase, a decision for handover is made, query regarding the available resources is performed and information, required for mobile IP execution, is retrieved. The IEEE 802.21a amendment to the MIH covers mechanisms for latency reduction during the authentication, as well as authorization and data protection for the MIH services. The main issue that is yet not defined in the standard is how to perform a reservation of the resource in order to ensure the required QoE and methods for translation of the Quality of Service (QoS) context definitions between different technologies and operators [7].

From a user’s point of view, maintaining seamless connection to Wi-Fi networks is not as straightforward as connecting to cellular networks. Often users do not recognize the available Wi-Fi networks, and, even more, they get discouraged to use Wi-Fi when credentials are necessary to connect to open or public networks. In order to provide a cellular-like experience and the Global Wi-Fi implementation, the Wi-Fi Alliance Hotspot 2.0 Specification, which references the IEEE 802.11u amendment, has been presented [8]. The IEEE 802.11u [9] aims to provide an overall end-to-end solution for interworking with external networks. It defines Layer 2 transport for a query-response protocol which can be the IEEE 802.21 protocol or the Access Network Query Protocol (ANQP). This allows users to effectively query the network for the information relevant to the network selection prior to performing the authentication procedures.

Nowadays mobile devices are capable of multihoming, allowing simultaneous connection to different networks, receiving and sending data on multiple interfaces at the same time. Seamless mobility thus becomes not just mobility per user, but more per service (flow). The IP Flow Mobility (IFOM) [10] provides the means to select and offload a single flow to a complementary access network. A shortcoming of IFOM is that it is limited to one active 3GPP and one active non-3GPP connections, while all IP flows are intended to one Packet Data Network (PDN) connection [11]. In [12], the missing interface towards the Policy and Charging Control (PCC) system is identified as a limitation for dynamic IFOM. This interface is needed in order to inform the gateways (PDN, signalling gateway) of the routing policies.

In the IEEE 802.21, a MIH Information Server (IS) is defined to maintain information of available networks. On
the other hand, 3GPP has defined the Access Network Detection and Selection Function (ANDSF) to maintain a map of coverage in a form of a static database containing the available access networks in a certain location (e.g., a cell). The mobile nodes can provide ANDSF with its capabilities (IFOM, Multiaccess PDN Connectivity or Seamless Wi-Fi) and location, so that ANDSF can limit the information that is valid for the particular mobile node. However, the standard does not include generation procedures for the policies and the dependence of the requested service, the QoS required for the service, the available signal strength levels, load balancing etc. The issue of a database design and maintenance has not been covered as well. As indicated in [13] and [14] the ANDSF does not provide sufficient tools to the operators for control of network access.

The Fixed/Mobile Convergence initiated the alignment of fixed and mobile management requirements between the 3GPP and the Broadband Forum (BBF). The Technical Report (TR) [15] defines architectural framework for interworking between Service Provider (SP) that offer 3GPP and/or fixed access. The TR considers several functionalities among which is Wi-Fi offloading and IFOM. The Broadband Policy Control Framework (BPCF), described in [16], defines a policy based activation of broadband services as well as policy management and control. The main logical entities are policy decision, policy enforcement points and admission control function. The policies are applied at an IP session, IP flow or at an aggregate level. The motivation and the architectures for interworking between 3GPP system and a fixed broadband access network using WLAN or femto access is elaborated in [17]. The authors indicate the different methods for policy control as a main problem for 3GPP-BBF interworking and discuss the procedures for initial attachment and policy and charging control session establishment.

The research community has also been focusing on suggesting and evaluating different frameworks for traffic steering. Important research directions such as: minimization of delay for the vertical handover, load balancing among networks, traffic offloading etc., have been covered to high extent. Different schemes with respect to user’s QoE and optimization of network resources have been proposed in the research literature. For example, a possible procedure for vertical handoff between cellular and Wi-Fi networks were suggested in [18]. However, this paper considers architecture for handover decision management taking into account only users and applications requirements. In [19] the decision algorithm relies on application profile handling, and manages a fleet of mobile nodes. The target of the framework is to either reduce the cost of communication from user point of view or to achieve the required QoS, while considering the energy consumption. Again the operator point of view is not taken into account, nor resource availability at the cellular or Wi-Fi networks. In [20], IEEE 802.21 is evaluated for handover of data services, where the decision is made within a time interval during which the data rates over two networks are compared. In [21], a context aware mobility framework is presented, where decisions are made in a cross-layer and interactive approach. The mobile node is responsible for determining the optimal point of attachment, while the network is responsible for optimizing the network resources.

In our previous work [22], we have introduced an approach to the network resource management, where the system functionalities are classified according to the time needed to perform particular control action. The framework is universal and applicable to a wide variety of network deployments characterized by multiple RATs, Heterogeneous Networks and different architectures of Base Stations (BSs) such as Distributed BS and Cloud Radio Access Network. In this work, we focus on traffic steering as one functional part of resource management. Here, we propose a framework where both users’ profiles per application/preferred operator and network mode are considered in the decision algorithm. More over, we argue that the network centric approach is more appropriate in achieving improved network operation and utilization, while the user is ensured with good connection and service anytime, anywhere.

**IV. Network-centric Framework Definition**

The proposed resource management framework in this paper adopts the network-centric approach, where assistance from mobile nodes is required. As the final decision is made at the network side, this framework allows the operator to be in control of the network load and resources in general. The mobile nodes are responsible to provide the network with various information such that the decision will improve users’ QoE.

As represented by Figure 2, the resource management framework consists of two parts: a network part or decision center and a mobile part or “smarter” application residing on the side of a Mobile Node (MN). The proposed decision center contains two (logical) elements: the Decision Entity and the IS. The Decision Entity is responsible for analyzing data and steering the traffic demands. It has several main operations that can indicate to the MN to perform: selection of a suitable access network from a list of hotspots/small cells, handover to a small cell or a Wi-Fi hotspot, offload traffic
flow to a small cell/Wi-Fi, etc. The goals of these operations are to ensure always best connectivity and service, optimize the network operation, balance the load among networks that belong to the operator. The IS represents a database of the (most recent) information gathered from the mobile devices, available eNodeBs and access points. The "smarter" application is responsible to send sufficient information that will aid the network to reach a decision. The users are encouraged to install this application as it will enhance the ubiquitous access to the Internet and improve the experienced QoS. In the following subsections, two approaches are presented: the basic and the enhanced framework.

A. Basic Framework

In the basic framework, the Decision Entity considers the data received from the MN, the network load and operation. Given that the mobile node is the only one able to sense alternative access networks, it is therefore responsible to detect and report a list of available networks. The "smarter" application is responsible to communicate with the decision center and send information on certain time intervals. The time intervals can be set by the decision center. The information gathered about Wi-Fi hotspot may not be just limited to the Service Set Identification (SSID). In case of support of IEEE 802.11u and ANQP, the MN can retrieve more information about the capabilities of the networks such as presented in [8]: Network Authentication Type information, Roaming consortium list, IP Address Type Availability, Hotspot Operator Friendly Name, Operating Class etc. Additionally in case of the support of IEEE 802.21, the MN can retrieve QoS parameters such as throughput, packet error rate and/or different classes of service [6].

All this information can be collected by the IS and made available to the Decision Entity for analysis. The task of the Decision Center is to provide a decision on the MNs that are best suited to be offloaded to alternative access networks and possibly forward the required security credentials for the new access network. The decision is made so that the operator's network is relieved from load, and the users that remain are ensured with the required level of quality.

B. Enhanced Framework

In the enhanced framework, the IS maintains a database (a Profile Database) of available mobile Network Profiles (NPs) and User Profiles (UPs) that indicate their operation and possibility to provide certain QoE. An example of the network profile can be presented as follows:

- **Guaranteed QoS parameters**: The available network is able to ensure a certain level of QoS parameters.
- **Non Guaranteed QoS**: The available network might be unknown to the operator, but can offer free access to the Internet without information on any QoS parameters.
- **Limited Resources**: The available network is capable to ensure a connectivity under certain limitations. For example the throughput per user is limited to a certain threshold, the number of users connected to the network is limited and/or only certain users are allowed to connect.

In order to assure its delegated functionality, Profile DB is allocated in a backbone. The Profile DB can be constructed dynamically based on the information from the MNs and/or agreements among network operators. In the second case, the networks need to update their profile in case there is a change. For example, if the network becomes overloaded and the available resources are limited. Additionally, in case of Wi-Fi hotspots, the credentials for access could be available at the Decision Center and therefore simplify the access of MNs to the available Wi-Fi.

The "smarter" application is enhanced as well. It is responsible for a profile activation depending on the type of the service that is requested by the user. The profiles can be based on internal parameters of the MN, that may include, as an example, desired throughput requirements, residual battery lifetime and activated subscription. The information on the active profiles per user is collected at the side of IS as well. The profiles can be defined according to the parameters that have higher priority, for example:

- **Required QoE parameters** profile determines the key quality parameters for the requested service. Here different classes of service can be defined depending of the type of application. For example it can be a high demand on bandwidth for streaming service, low latency for voice services, low packet error rate for file synchronization etc.
- **Minimum price** profile controls traffic consumption of the MN and its applications, minimizing the overall cost of data transmission. The application will look for Wi-Fi networks, checking their quality. Known Wi-Fi networks and security credentials could be stored and marked as the preferred ones. Application will also create UPs in open access Wi-Fi networks on behalf of the user, if needed for connecting, to facilitate the connection.
- **Maximum battery lifetime** profile controls different parameters of the MN, maximizing total time that MN can work on its battery power.

In the enhanced framework, the Decision Entity reaches decisions based on the data received from the MN, the user profiles and network conditions that are reflected through NPs. The proposed architecture and the communication among different entities in the enhanced framework is elaborated in the next section.

V. PROPOSED ARCHITECTURE FOR DATA OFFLOADING

We build our approach on the assumption that MN with "smarter" application offloads traffic to the Wi-Fi network, which is superior in terms of provided QoS parameters and most closely fits to the UP that is currently active on the MN. In the basic approach, the Decision Entity may require MN to offload all traffic each time when Wi-Fi is available for connection. However, in this paper, we propose an advanced technique that potentially can provide an optimal load balancing solution for network operators, still keeping the user of MN satisfied.
by the acquired network resources. As represented by Figure 3 a traffic offloading scenario begins on the side of MN by performing network discovery procedure in a passive low power consumption scanning mode by retrieving information about new Wi-Fi hotspots from broadcasted beacon frames. Since the MN may not recognize all received SSIDs and cannot retrieve their QoS parameters as well as an active NP, it transfers the acquired data to the network-centric Decision Center together with the request to evaluate recently discovered Wi-Fi.

Upon the retrieval of network QoSS and NPs together with UP of MN that is analyzed for offloading options, the IS executes the algorithm of decision making as depicted by Figure 4. After performing mapping procedures of received SSIDs to NPs from the Profile DB, the Decision Center analyses if the discovered networks meet the requirements and are reasonable for data offloading for a given UP, by executing the following compliance test:

- Requirement 1: NP should support connection of MN that has an active UP.
- Requirement 2: Provided QoS parameters of the network should be equal or greater than those requested by UP.

If both requirements are met, the IS puts SSIDs corresponding to current NP into the set $C$ that defines available candidates for offloading. If after searching procedures it reveals that $C = \emptyset$, the IS sends an information message to a "smarter" application that MN should stay in the serving cellular network. In this case, IS also sends notification that "smarter" application declares too high requirements for the networks that are currently available for offloading. As soon as the final list of the SSIDs that meet requirements of UP is created, the final decision may be derived. We define the task of optimal NP selection that is performed by the IS as an optimization problem that can be formulated as follows:

$$NP^* = \max_{NP_i \in C} f(NP_i, UP), \quad i = 1, \ldots, |C|,$$

where the objective function $f$ takes the set $C$ and the UP as initial parameters in order to calculate $NP^*$ that corresponds to the optimal load balancing solution and maximizes the satisfaction of the end-user by provided network resources. However, the overall solution for the introduced optimization problem highly depends on final implementation of this traffic steering framework and lies beyond the scope of the current paper.

After transmission of security credentials and the SSID that corresponds to the profile $NP^*$, offloading procedures are initiated between the MN and selected network. In this case, the MN may simply switch connection to the designated network or employ advanced protocol for traffic offloading and vertical handover, such as the MIH. If necessary, the IS also provides security parameters for authorization and indication of roaming partners, so the MN correctly identifies recently selected access network and is able to proceed with association.

Interworking solutions based on the MIH, like the IEEE 802.21, may also facilitate the task of the vertical handover realization between LTE/LTE-A and Wi-Fi networks as well as provide feasible opportunity for a possible functionality extension of the proposed framework. The MIH IS may provide information about discovered networks and indicate roaming partners for a MN. However, some procedures that are usually performed on higher layers, like pre-association and user authorization, may still require additional interaction with 3GPP and IEEE 802.11 technologies. Information exchange can be carried out by employing the ANQP from the IEEE 802.11u amendment. It can potentially simplify procedures of network identification, providing comprehensive access to information about network operators, type of the service (public, private, paid-for, etc) and connection parameters like cost of the traffic, for example. In order to have a common and up-to-date information about Wi-Fi hotspots and eNodeBs that are available in the area, their operators are supposed to negotiate a joint agreement to share this information with each other. However, we foresee that this requirement may be neglected in the nearest future after overall deployment of Wi-Fi access points with support of the Hotspot 2.0 Technical Specification that includes the IEEE 802.11u amendment. The IEEE 802.11u gives an opportunity to retrieve the information about network specifications through user generated probe requests and responses. The access network information retrieval is conducted by the ANQP, which is transported by the Generic Advertisements Service public action frames.
carrying out transportation frames of higher layer advertisements between Wi-Fi hotspots (or IS) and the MN. The information repository for this type of a protocol, in terms of this framework definition, is provided by the Decision Center. According to the fact that location of this server lies outside the scope of the IEEE 802.11u specification we bind the physical location of the IS and the Profile DB to a backbone network.

VI. CONCLUSION AND FUTURE WORK

In a scope of broad deployment of the 3G/4G access technologies, this paper introduces a novel network-centric framework for the offload decision making supported by the mobile-assisted discovery of candidate-networks.

The fast and reliable approach is provided by allocating the decision center in a backbone, which facilitates optimization of overall network resources and implementation of efficient load balancing scheme. At the same time the quality of end-user experience is controlled at an appropriate level without noticeable increase in the power consumption on its side due to the passive network discovery. In order to facilitate and accelerate adjustments according to the changes in available network resources and the Quality of Experience demands of the user, a profile-based control is presented for the user and network entities.

As another advantage, the proposed framework also fits currently available IEEE standards for the Media Independent Handover, which facilitates its implementation and retrieval of network parameters. Together with the profile approach, this resource management framework allows a network operator to determine and control the Quality of Experience of the end-user from the network-centric perspective. All the benefits listed above give potential for the presented framework to be successfully implemented and deployed in future mobile communication networks.

Performance evaluation of the proposed solution is the subject of our future work.

ACKNOWLEDGMENT

This work was partially sponsored by the Danish Advanced Technology Foundation through the research project SAIRS and by the Seventh Framework Programme for Research of the European Commission HARP project under grant number 318489.

REFERENCES