Expert Systems with Applications 41 (2014) 4996–5008

Contents lists available at ScienceDirect

Expert Systems with Applications

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

An expert system for real-time traffic management in wireless local area networks

T. Frantti [a,*], M. Majanen [b]

[a] Renesas Mobile Europe Ltd., Elektroniikkatie 10, FI-90570 Oulu, Finland
[b] VTT Technical Research Centre of Finland, Finland

ARTICLE INFO

Keywords:
Expert system
Fuzzy control
Flow control
Offloading
Real-time traffic
WLAN

ABSTRACT

The paper explores delay-based congestion and flow control and the offloading of real-time traffic from wireless local area networks (WLANs) to mobile cellular networks (MCNs) in multihomed devices. The control system developed is based on an embedded hierarchical expert system. It adjusts transceivers' traffic flow(s) for prevailing network conditions to achieve application-dependent delay and throughput limits. In wireless networks, delay and throughput depend on the packet size, packet transmission interval, and node connection density. Therefore, the controller on the destination node monitors average one-way delay and the change of one-way delay of the incoming traffic. On this basis, it adjusts the packet size and transmission interval of the source node by transmitting a control command to the source. If the prevailing level of traffic in the network exceeds its capacity despite of the control actions taken, devices prepare for developed asynchronous offloading of traffic to another access network.

The control model was validated via simulation of Voice over Internet Protocol (VoIP) traffic in the OMNeT++ network simulator. The results demonstrate that the expert system developed is able to regulate packet sizes to match the prevailing application-dependent optimum and transfer traffic to another network if the network exceed its capacity no matter the control actions taken. Although this work is motivated mainly by issues of congestion and flow control of WLAN systems and the simulations and results were prepared for the IEEE 802.11*b* system, the approach and techniques are not limited to these systems, but they are applicable for other packet switched access networks (PSANs), too.

1. Introduction

With the everyday use of cellular networks, extensive mobility during communication became self-evident reality for end users. A mobility requirement soon appeared also for Internet-based communication, and currently mobile phone networks are moving with Long Term Evolution (LTE) toward an Internet Protocol (IP) infrastructure with the ultimate aim of offering almost all the services provided by current second- and third-generation (2G/3G) cellular networks through an all-IP network (see ITU-T (2004a) and ITU-T (2004b) for more details), although Transmission Control Protocol/Internet Protocol (TCP/IP) networks were not originally designed for mobile use. For example, devices on the Internet are identified by an IP address, which has a dual role: it serves as an identifier and as a locator of the networked device. If a device changes its point of attachment, such as wireless access network, its IP address too may change, which means that other devices in the network need means to access the new address if they are to reach the device at its new location. The change in IP address causes the upper-layer protocol (at the layer above the network layer) connection to break, which is problematic for applications with more persistent connections or applications requiring registration of an IP address.

Today, mobility in IP networks relies on wireless access technologies, and the trend is toward equipping mobile devices, such as smartphones, with multiple network interfaces. The various interfaces are, in the general case, operated by different Internet service providers (ISPs), for devices' improved resilience via an opportunity to connect to the Internet through at least one of the access technologies. For end users, these technologies vary mainly in their coverage area and performance.

A networked mobile device can use the available interfaces either one at a time or several simultaneously. The device using its multiple network interfaces simultaneously is a special case of node multihoming. A multihomed device has multiple IP addresses, assigned to the same network interface or different ones.

* Corresponding author. Tel.: +358 40 547 0819.
E-mail addresses: tfrantti@ee.oulu.fi (T. Frantti), mikko.majanen@vtt.fi (M. Majanen).

http://dx.doi.org/10.1016/j.eswa.2014.01.024
0957-4174/

Node multihoming differs from site multihoming, which can be defined as an edge network configuration that has more than one service provider but does not provide transit communication between them; see the work of Baker (2011). In site multihoming, end users do not need to manage their devices in any way. When a networked device is multihomed, there are multiple paths between the source and destination devices. The paths may differ in price, data rate, latency, jitter, and packet loss. Node multihoming, with parallel connections, enables new kinds of traffic engineering techniques, whilst site multihoming is intended mainly to increase the reliability of the Internet connection.

In this paper, we explore delay-based congestion and flow control and offloading of real-time traffic from wireless local area networks (WLANs) to cellular networks in multihomed mobile devices. In WLANs, delay and throughput depend on the packet size, packet transmission interval, and connection density. Therefore, we developed and applied control systems to adjust transceivers' packet sizes for prevailing network conditions to achieve application-dependent delay and throughput limits for real-time traffic and to avoid unnecessary offloading. If the prevailing level of traffic in networks exceeds capacity regardless of the control actions, devices prepare to perform asynchronous offloading of traffic to another access network. In our work, the research assumptions can now be presented as: *"Multihomed and multiradio (WLAN and cellular) devices operating in WLAN and running real-time applications try to maximize network capacity by packet size optimization. When the performance threshold is reached, connections are asynchronously offloaded to cellular networks."* By real-time traffic we mean Voice over Internet Protocol (VoIP), video calls, and interactive games. With the assumptions above, the research question for the work can be formulated as follows: *"When should we begin to transfer traffic of real-time applications to another (cellular) network as the WLAN's traffic and number of users increase, and how can we do that to avoid unnecessary offloading?"*

The control model was validated through simulation of VoIP traffic in the OMNeT++ network simulator. Even if this work is addressed mainly to congestion and flow control of WLAN systems, the approach and the techniques are not limited to these systems. They are applicable also for other packet switched access networks (PSANs).

The rest of the paper is organized as follows. Section 2 presents a review of the literature on packet size control in WLANs, multihomed path management, and network selection. Section 3 focuses on multihome- and offloading-related standardization activities. Section 4 summarizes the hierarchical decision making system developed, which features a packet size control unit and real-time offloadig of traffic from WLANs to mobile cellular networks unit. Section 5 describes the simulation model, and Section 6 presents the simulation results. Section 7 delineates our future research. Finally conclusions are drawn in Section 8.

2. Literature review

2.1. Packet size optimization in WLANs

Korhonen and Wang (2005) have studied the effect of packet size on loss rate and delay in an IEEE 802.11-based WLAN. The analysis shows that there is a straightforward connection between packet size, bit error characteristics, and observed delay characteristics. In general, it is evident throughout the literature that the performance of wireless networking is sensitive to packet size and that significant performance improvements are obtained if a "good" packet size is used. For example, Bakshi, Krishna, Vaidya, and Pradhan (1997) show this for TCP traffic over a wireless network. Chee and David (1989), Lettieri and Srivastava (1998), and Chien et al. (1999) all have studied the relationship between frame length and throughput, but they do not propose any precise method for dynamic control of frame length to maximize throughput. Smadi and Szabados (2006) focus on optimization of packet size in error recovery but do not consider the optimal packet size for performance optimization. Sheu, Lee, Chen, Yu, and Huang (2000) present a fuzzy packet length controller (PLFC) for improving the performance of WLANs suffering from interference from a microwave oven. It was demonstrated that the PLFC improves the throughput of User Datagram Protocol (UDP) traffic from that with fixed-length packets, but the authors did not consider performance improvements when the number of users and the amount of traffic are increased. Sankarasubramaniam, Akyildiz, and McLaughlin (2003) have studied packet size optimization for energy efficiency, and Younis, Farrag, and D'Amico (2009) consider it for security and throughput, but their solutions are statistical in nature, meaning that the packet size is optimized beforehand. In the most recent of our publications (Frantti & Majanen, 2011), we presented and compared proportional-integral-derivative (PID) and fuzzy control systems to achieve maximum throughput and minimal delay by adjusting the packet sizes of UDP-based uni- or bi-directional real-time traffic in WLANs in line with prevailing channel conditions. The aim with the hierarchical expert system developed for this paper is to reach application-dependent delay and throughput limits for the maximum number of such real-time connections as VoIP calls and, if necessary, offload real-time connections to another network interface in a controlled manner.

2.2. Multihomed path management

In node multihoming, the end nodes are responsible for paths' management. Fekete (2010) notes that multihoming can be used to enable more advanced traffic engineering or load sharing techniques, such as load balancing and load spreading, by distributing traffic over multiple interfaces and addresses. In load balancing, different traffic flows are sent on different paths. In load spreading, packets belonging to the same flow are sent through different paths. Kandula, Lin, Badirkhanli, and Katab (2008), Singh, Alpcan, Agrawal, and Sharma (2010), and Yao, Kanhere, and Hassan (2009) use quality of service (QoS) parameters for balancing the traffic load over the available network interfaces. In host-centric traffic engineering, this is done by selecting the proper source IP address for outgoing packets and notifying peer nodes about the IP address for incoming packets. In network-centric traffic engineering, this is done within the network through tuning of the routing protocols (see Schiller (2005)).

2.3. Multihomed interface selection

Common approaches to network selection are based on estimated QoS parameters for each available network. Adamopoulou, Demestichas, Koutsorodi, and Theologou (2005), Bari and Leung (2007), Psaras and Mamatas (2011), Song and Jamalipour (2005), Wang, Katz, and Giese (1999), Wilson, Lenaghan, and Malyan (2005), Xing and Venkatasubramanian (2005) and Yahiya and Chaouchi (2009) consider network QoS parameters such as delay and capacity. In the literature, researchers have also proposed more user-centric criteria, among them the power consumption of devices (see Yahiya & Chaouchi (2009) and Petander (2009)) or the cost of network use (see Adamopoulou et al. (2005), Bari & Leung (2007), Wilson et al. (2005) and Alkhawlani & Ayesh (2008)). In some of the aforementioned references, including those of Adamopoulou et al. (2005), Song and Jamalipour (2005), Wang et al. (1999), and Alkhawlani and Ayesh (2008), a user is expected to supply policies and preferences. The approach of Alkhawlani and Ayesh (2008) proposes a combination of fuzzy logic and genetic