Simulation Modelling Practice and Theory 25 (2012) 17–35

Contents lists available at SciVerse ScienceDirect

Simulation Modelling Practice and Theory

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

A subscriber classification approach for mobile cellular networks

A.M. Kurien [a,*], G. Noel [a], K. Djouani [a], B.J. Van Wyk [a], A. Mellouk [b]

[a] French South African Insitute of Technology (F'SATI), Tshwane University of Technology, PBag x680, Pretoria, South Africa
[b] Network and Telecom Dept. and LiSSi Lab, University Paris-Est (UPEC), IUT Creteil/Vitry, Vitry sur Seine, France

ARTICLE INFO

Article history:
Received 2 June 2011
Received in revised form 16 February 2012
Accepted 22 February 2012
Available online 22 March 2012

Keywords:
Feature extraction
Mobile networks
Subscriber classification

ABSTRACT

The classification of subscriber types in mobile cellular networks is valuable for network service providers since it provides a mechanism to plan network services by better understanding subscriber behaviour in a network. Mobile networks contain vast repositories of data that store valuable information regarding subscriber behaviour. In this paper, a new approach for subscriber classification in mobile cellular networks is proposed. The proposed approach considers network traffic generated from a mobile cellular network operator in South Africa. The proposed approach makes use of a difference histogram approach for feature extraction and a fuzzy c-means clustering algorithm to classify traffic data into subscriber classes. To validate the proposed approach, a comparative analysis of two different multi-resolution feature extraction approaches, the empirical mode decomposition (EMD) approach and the discrete wavelet packet transform (DWPT) are compared with results obtained with the difference histogram (DH) approach. It is shown that the difference histogram provides better clustering results when compared to the two multi-resolution approaches demonstrating the potential of the difference histogram approach.

1. Introduction

Mobile cellular networks have been adopted rapidly worldwide and are considered to be amongst the most popular personal technology across the world [1]. According to the GSM association [2], it was announced that the number of global mobile subscriptions had passed 5 Billion by July, 2010. The ITU estimated that there were nearly 300 Million subscribers in Africa that rose to nearly 450 Million mobile subscribers in 2009 with an approximated growth rate of 25% [3]. As the number of subscriber bases in most mobile networks across Africa and around the world continue to increase, one of the greatest challenges faced by network operators is the efficient management of existing network resources to meet the demand in capacity as a result of the experienced growth. The management of resources needs to be viewed from two perspectives; (i) managing the existing network through efficient capacity management to meet desired quality of service (QoS) in terms of minimum grade of service which is linked to blocking probability in the network; (ii) defining an optimum threshold and determining which network infrastructure upgrades are imperative to support the rise in demand for capacity.

One way in which network operators can assess the performance of a mobile network is by evaluating the impact of subscribers on the network [4]. By understanding the influence of subscribers on the network, network providers are able to better manage the extent of capital investment on the network and manage available resources for effective QoS management. Information regarding a customer's behaviour and market segments that the customer belongs is useful to operators to make major decisions regarding business service offerings that are catered to identified market segments [5]. With the rise

* Corresponding author.
E-mail addresses: kurienam@tut.ac.za (A.M. Kurien), noelg@tut.ac.za (G. Noel), djouanik@tut.ac.za (K. Djouani), vanwykb@tut.ac.za (B.J. Van Wyk), mellouk@u-pec.fr (A. Mellouk).

1569-190X/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.simpat.2012.02.008

in demand for capacity in mobile networks across Africa and with the varying socio-economic distributions experienced in these networks, the ability to separate subscriber types into subscriber classes for effective network capacity planning and forecasting is vital. Vast repositories of various measurements are generated through counters in a typical cellular network on a daily basis.

However, the extraction of meaningful subscriber behaviour from these repositories is a challenge and requires the analysis of multiple data sources. The use of signal processing approaches have been widely used for feature extraction and classification. However, one of the problems with traditional signal processing approaches is their inefficiency in providing sufficient information in the time domain. The choice of approaches such as the wavelet transform have shown the ability to handle signals in short time intervals for high frequency components as well as long time intervals for low frequency components [6]. Wavelet based multi-resolution analysis techniques (WMAT) have been shown to be useful in denoising multi-dimensional spatial/temporal signals containing steady/unsteady noise [7]. The empirical mode decomposition (EMD) approach has been shown to be a powerful tool in analysing composite, non-linear and non-stationary signals [8]. The difference histogram approach [9], a relatively new approach for feature extraction, has been shown to have benefits in terms of computational complexity and its suitability for real-time applications.

Though the above approaches have shown their suitability and robustness in feature extraction of non-linear and non-stationary signals, their efficiency in feature extraction and classification of subscriber classes in mobile networks needs to be evaluated. In this paper, a new approach for subscriber classification in mobile networks is presented. The proposed approach makes use of the difference histogram approach proposed in [9] for feature extraction and a fuzzy c-means algorithm for clustering. To show the performance benefits of the histogram approach, the algorithm is compared with an EMD and DWPT approach. The remainder of this paper is organised as follows. The next section provide a summary of related work on network planning and optimisation strategies and feature extraction approaches. Section 3 provides a theoretical overview of the multi-resolution analysis approaches utilised in this study. Section 4 discusses the proposed subscriber classification approach for mobile cellular subscribers. Section 5 provides cluster analysis results which highlights the comparative analysis between the three feature extraction approaches. Based on the cluster analysis results, a proposed two-level hybrid channel allocation approach is presented in Section 6. The conclusion of the paper is provided in Section 7.

2. Related work

Various studies have been conducted to propose efficient network planning and optimisation strategies that make use of load balancing strategies and growth planning strategies for network evolution. With the availability of vast amounts of stored data in mobile networks, the determining of subscriber classes based on subscriber behaviour could be useful for the network planning process. The following subsections give an overview of mobile network optimisation strategies and feature extraction and classification approaches proposed in literature.

2.1. Mobile network optimisation strategies

Zang et al. propose a new efficient load balancing algorithm for channel assignment in [10]. The approach employs two thresholds (light, heavy) to classify cells into three categories according to their states. The proposed scheme runs a channel-borrowing algorithm whenever there exists a heavy cell needing free channels. Whitaker et al. [11] present the concept of marginal cost of service coverage which represents the lowest rate at which infrastructure cost must increase to facilitate higher levels of service coverage. Cell-plan infrastructure efficiency is studied from two perspectives: establishing the effect of cell density on infrastructure cost of the network, and secondly, the influence of the effects of increasing infrastructure expenditure. Network planning can be investigated from an infrastructure cost-centric point of view addressing the balance between infrastructure costs and coverage for a wide range of subscriber densities. Wireless service providers face the need to plan and rapidly evolve networks to meet subscriber demands which has an impact on the physical aspects of the network such as switching and radio equipment, available frequencies etc as described by Garcia et al. in [4].

From a planning point of view, the impact of network optimisation needs to focus on two aspects of the network: the access side and the core network side. Fig. 1 provides a basic overview of the key components of a typical cellular network highlighting the access side and the core network side. Calin et al. discus a new approach for capacity planning for growing CDMA networks. They highlight how the approach has aided in faster and more accurate capacity planning cycles while balancing QoS and capital investment constraints [12]. Mazzoni et al. investigate the characteristics of Italian cell phone users in [13]. A multi-dimensional segmentation approach is used to determine if differences exist among Italian mobile phone users leading to the identification of different market segments and secondly, to determine if it is possible to describe them using the multi-dimensional approach. In [14], Kianmehr et al. conduct cluster analysis to identify calling communities using information derived from call data records (CDR). In [15], factor analysis, clustering, and quantitative association is used to find service adoption patterns of segmented groups. From a network planning and optimisation point of view, by identifying market segments and considering that the network is not homogeneous but rather a heterogeneous network comprising of subscribers that behave differently, the network planning strategies employed for cells deployed in these segments has to be managed more effectively.