2017 31st International Conference on Advanced Information Networking and Applications Workshops

Mobile Data Usage Prediction System and Method

Yueh-Ting Lai
Telecommunication Laboratories,
Chunghwa Telecom Co., Ltd.
Taipei, Taiwan, R.O.C.
fluting103@cht.com.tw

Ya-Ping Wu
Telecommunication Laboratories,
Chunghwa Telecom Co., Ltd.
Taipei, Taiwan, R.O.C.
yaping@cht.com.tw

Chia-Hsuan Yu
Telecommunication Laboratories,
Chunghwa Telecom Co., Ltd.
Taipei, Taiwan, R.O.C.
tristan@cht.com.tw

Fang-Sun Lu
Telecommunication Laboratories,
Chunghwa Telecom Co., Ltd.
Taipei, Taiwan, R.O.C.
fslu@cht.com.tw

Chi-Hua Chen
Telecommunication Laboratories,
Chunghwa Telecom Co., Ltd.
Taoyuan, Taiwan, R.O.C.
chihua0826@cht.com.tw

***Abstract*—The popularization of Long Term Evolution (LTE) leads to the issues of mobile data usage monitoring and prediction. The analyses of these issues can be referred for the deployment of mobile network and the recommendation of customer solutions. However, a great deal of statistical computing power and time is needed for several previous mobile data usage prediction methods. Because large variances may exist among the quantities of mobile data usages from different client devices, the error of mobile data usage prediction may be large. Therefore, this study proposes a mobile data usage prediction method to analyze the historical mobile data usages of target client devices for the prediction of future mobile data usages of target client devices. Then the marketing information can be obtained to client devices in accordance with the predicted mobile data usage, and the income of operator can be increasing. In experimental results, the mobile data records from 565,606 client devices during July of 2014 were collected and analyzed for the evaluation of the proposed method. The accuracy of the proposed method was 87.97% which was higher than previous methods.**

Keywords—mobile data; data usage prediction; cellular network; k nearest neighbors

I. INTRODUCTION

With the promotion and universality of cellular networks (e.g., Universal Mobile Telecommunications System (UMTS) and Long Term Evolution (LTE)), the issues of mobile data usage prediction are more and more important for the strategies of mobile network deployment and mobile data plan recommendation. The investigation indicated that the output of industrial production of global communication devices has been increased by 0.46 million US dollars, and the annual growth rate of mobile communication industries in Taiwan is measured as 3.7% [1]. Therefore, the volumes of mobile data usages will be generated and produced, and the techniques of mobile data usage prediction are required for mobile users and telecom companies.

In previous studies, some methods based on statistical approaches and grey systems were proposed to predict mobile data usages [2]. The mean values and regression models were adopted and obtained by these approaches for mobile data usage prediction. Although lower computation complexity is required by these approaches, large variances and errors may be produced in accordance with different mobile data plans. Mobile users' behaviors may be different when they make different mobile data plans. For the improvement of mobile data usage prediction, artificial neural network (ANN) approaches were proposed to solve the non-linear problems. ANN models can be trained by using historical records, and these trained ANN models can be used to predict mobile data usages [3]. Although ANN approaches can obtain higher accuracy than statistical approaches, the ANN models may be crashed in trained stage or obtained higher errors in runtime stage with several outliers in historical records. Therefore, this study proposes mobile data usage prediction system and method based on k-nearest neighbor algorithm (kNN) to analyze the historical mobile data usages of target client devices for the prediction of future mobile data usages of target client devices. Then the marketing information can be obtained to client devices in accordance with the predicted mobile data usage, and the income of operator can be increasing.

The remainder of the paper is organized as follows. Section II proposes the mobile data usage prediction system and method. Section III illustrates the practical experimental results and gives prediction accuracy comparisons with different methods. Finally, conclusions and future work are given in Section IV.

II. THE PROPOSED MOBILE DATA USAGE PREDICTION SYSTEM

This study proposes a mobile data usage prediction system which includes (1) client devices, (2) cellular core networks, (3) an accounting computing system, (4) a database server, (5) a

978-1-5090-6231-7/17 $31.00 © 2017 IEEE
DOI 10.1109/WAINA.2017.50

mobile data usage prediction server, and (6) a decision support server (shown in Figure 1). There are c client devices and p mobile data plans for these client devices in this system. These client devices can connect the Internet through cellular core networks, and call detail records (CDRs) and communication records can be generated and obtained by the cellular core networks for an accounting computing system. The accounting computing system can calculate the mobile data usages and fees of each client device in each cycle period (e.g., m days per cycle) in accordance with mobile data plan. In this study, the value of m was adopted as 5 (i.e., 5 days per cycle) for accounting. The communication records of each client device which include the ID of client device, the ID of cycle period, the ID of mobile data plan, and mobile data usages can be recorded and stored into a database server. A mobile data usage prediction method which is performed by the mobile data usage prediction system is proposed to analyze and predict the mobile data usages of each client device in the next cycle period. Then the prediction results can be obtained to a decision support server, and the decision support server can provide strategies and updated mobile data plan for the users of client devices.

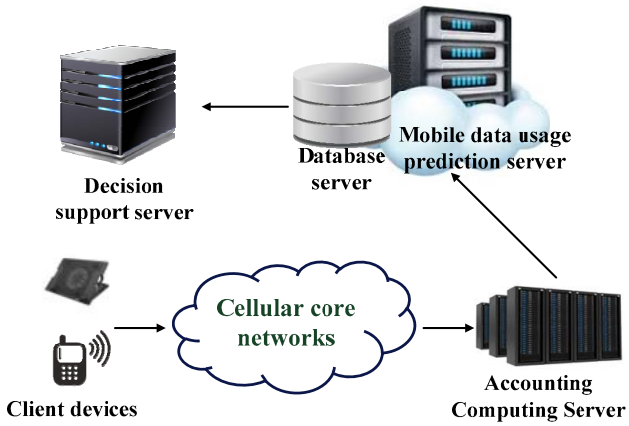

Fig. 1. The architecture of the proposed system

A. Client Devices

Client devices are telecommunication equipment which include mobile phones, tablets, and other smart handheld devices can access cellular core networks via a cellular network module. These client devices usually have a subscriber identity module (SIM) card which can be identified and authenticated by cellular core networks. User can use his client device with a SIM card to connect to the Internet through cellular core networks.

B. Cellular Core Networks

Cellular core networks which offer communication features (e.g., authentication, call control, switching, and gateways) for client devices accessing network. Cellular core networks can accept the requests from client devices in accordance with authentication and authorization processes by a SIM card. Then the connection between client devices and the Internet can be built, and the client devices can be enabled to access Internet services. The CDRs and communication records can be recorded and obtained for an accounting computing system.

C. An Accounting Computing System

An accounting computing system is a system that can obtain bill information for customers with various mobile plans and charging scenarios. The accounting computing system can receive the CDRs and communication records from cellular core networks to calculate the mobile data usages of each client device and to obtain bill information for each mobile user in accordance with user's mobile data plan. These mobile data usages can be stored into database server for mobile data usage prediction.

D. A Database Server

A database server is the storage server which performs tasks (e.g., data storage, data analyses, data manipulation, and data archive). In the proposed system, the database server is used for the storage of bill information, it can store the mobile data usages of each client device in each cycle period, and these data can be classified by mobile data plans.

E. A Mobile Data Usage Prediction Server

A mobile data usage prediction server is the critical part of the proposed system which performs the proposed mobile data usage prediction method based on kNN (shown in Figure 2). This server can retrieve the mobile data usages from database server and predict the future mobile data usages of target client devices in next cycle period for decision support server.

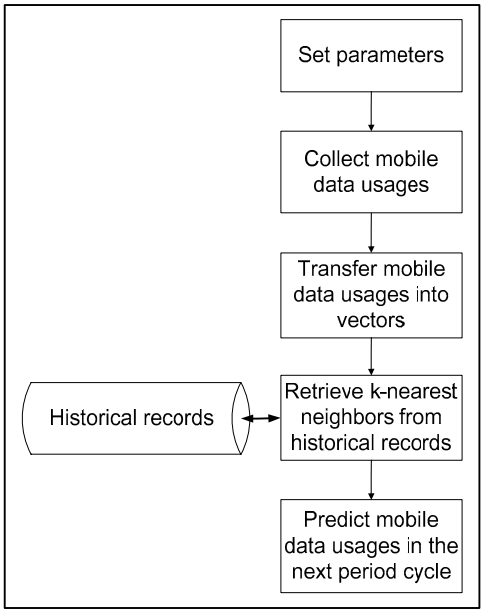

Fig. 2. The steps of the proposed method

F. A Decision Support Server

A decision support server is a system which supports the telecommunication company to improve sustainable

competitive advantages by creating suitable marketing strategies and customer services. In the proposed system, the decision support server can provide alarm mechanisms to notify customers of data usage warning and to design suitable marketing plans in accordance with the prediction results from the mobile data usage prediction server.

III. Practical Experimental Resuts and Discussions

In experimental environments, the mobile data usage records were retrieved from 565,606 client devices in the north of Taiwan during July of 2014. The mobile data plans of Chunghwa Telecom which included mPro50, mPro150, mPro450, mPro550, mPro650, and mPro750 were made by these client devices. For the proposed mobile data usage prediction method, the value of k was adopted as 10, and a period cycle included 5 days in this study. Furthermore, the previous methods (e.g., statistical mean value (SMV) and ANN) were also implemented and compared with the proposed method. The practical experimental results are showed in Table I and indicate that the accuracy of the proposed method is higher than these previous methods.

TABLE I. Accuracy Comaprisons with Different Methods

Method	SMV [2]	ANN [3]	Proposed method
Accuracy	58.33%	84.29%	87.97%

IV. Clusions and Future Work

In this study, mobile data usage prediction system and method are proposed to analyze the historical mobile data usages of target client devices for the prediction of future mobile data usages of target client devices. Furthermore, this study considers that mobile users' behaviors may be different when they make different mobile data plans. Therefore, the mobile data usages can be classified by mobile data plans, and the prediction models can be trained and used for each mobile data plan. The practical experimental results showed that the accuracy of the proposed method was 87.97% which was higher than previous methods. In the future, the alarm mechanisms and marketing plans which are generated by decision support system can be obtained to mobile users and evaluated the performance of these marketing plans.

References

[1] S.Y. Wang, "A study on the industrial trends and future prospects of global communications," *Industrial Economics and Knowledge Center*, Industrial Technology Research Institute, Taiwan, 2015.

[2] D. Naboulsi, M. Fiore, S. Ribot, R. Stanica, "Large-scale mobile traffic analysis: a survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 124-161, 2016.

[3] A.M. Kurien, G. Noel, K. Djouani, B.J. Van Wyk, A. Mellouk, "A subscriber classification approach for mobile cellular networks," *Simulation Modelling Practice and Theory*, vol. 25, pp. 17-35, 2012.