AFIN 2015

The Seventh International Conference on Advances in Future Internet

ISBN: 978-1-61208-428-2

August 23 - 28, 2015

Venice, Italy

AFIN 2015 Editors

Eugen Borcoci, University "Politehnica" of Bucharest (UPB), Romania

Carlos Becker Westphall, University of Santa Catarina, Brazil

AFIN 2015

Foreword

The Seventh International Conference on Advances in Future Internet (AFIN 2015), held between August 23-28, 2015 in Venice, Italy, continued a series of events dealing with advances on future Internet mechanisms and services.

We are in the early stage of a revolution on what we call Internet now. Most of the design principles and deployments, as well as originally intended services, reached some technical limits and we can see a tremendous effort to correct this. Routing must be more intelligent, with quality of service consideration and 'on-demand' flavor, while the access control schemes should allow multiple technologies yet guarantying the privacy and integrity of the data. In a heavily distributed network resources, handling asset and resource for distributing computing (autonomic, cloud, on-demand) and addressing management in the next IPv6/IPv4 mixed networks require special effort for designers, equipment vendors, developers, and service providers.

The diversity of the Internet-based offered services requires a fair handling of transactions for financial applications, scalability for smart homes and ehealth/telemedicine, openness for web-based services, and protection of the private life. Different services have been developed and are going to grow based on future Internet mechanisms. Identifying the key issues and major challenges, as well as the potential solutions and the current results paves the way for future research.

We take here the opportunity to warmly thank all the members of the AFIN 2015 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to AFIN 2015. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the AFIN 2015 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that AFIN 2015 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the field of Future Internet.

We are convinced that the participants found the event useful and communications very open. We hope Venice provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

AFIN 2015 Chairs:
Jun Bi, Tsinghua University, China

Eugen Borcoci, University Politehnica of Bucharest, Romania
Petre Dini, Concordia University - Montreal, Canada / China Space Agency Center - Beijing,China

AFIN 2015

Committee

AFIN Advisory Chairs

Petre Dini, Concordia University - Montreal, Canada / China Space Agency Center - Beijing, China
Eugen Borcoci, University Politehnica of Bucharest, Romania
Jun Bi, Tsinghua University, China

AFIN 2015 Technical Program Committee

Rocío Abascal Mena, Universidad Autónoma Metropolitana - Cuajimalpa, México
Marie-Hélène Abel, University of Technology of Compiègne, France
Alessandro Aldini, University of Urbino "Carlo Bo", Italy
Javier A. Barria, Imperial College London, UK
Henri Basson, University of Lille North of France (Littoral), France
Khalid Benali, LORIA - Université de Lorraine, France
Jun Bi, Tsinghua University, China
Alessandro Bogliolo, University of Urbino, Italy
Eugen Borcoci, University Politehnica of Bucharest, Romania
Christos Bouras, University of Patras and Research Academic Computer Technology Institute, Greece
Tharrenos Bratitsis, University of Western Macedonia, Greece
Antonio Brogi, University of Pisa, Italy
Chin-Chen Chang, Feng Chia University, Taiwan
Grzegorz Chmaj, University of Nevada - Las Vegas, USA
Hauke Coltzau, Fernuniversität in Hagen, Germany
Javier Cubo, University of Malaga, Spain
Maurizio D'Arienzo, Seconda Università di Napoli, Italy
Jiangbo Dang, Siemens Corporation | Corporate Technology Princeton, USA
Guglielmo De Angelis, CNR - IASI, Italy
Sagarmay Deb, Central Queensland University, Australia
Gayo Diallo, University of Bordeaux Segalen, France
Daniel Díaz-Sánchez, University Carlos III - Madrid, Spain
Sudhir Dixit, HP Labs India - Bangalore, India
Jonas Etzold, Fulda University of Applied Sciences, Germany
Florian Fankhauser, TU-Wien, Austria
Wu-Chang Feng, Portland State University, USA
Alex Galis, University College London, UK
Ivan Ganchev, University of Limerick, Ireland
Rosario G. Garroppo, Università di Pisa, Italy
Christos K. Georgiadis, University of Macedonia, Greece
Apostolos Gkamas, Higher Eccelesiastic Academy Vellas of Ioannina, Greece
William I. Grosky, University of Michigan-Dearborn, USA
Vic Grout, Glyndwr University, U.K.
Adam Grzech, Wrocław University of Technology, Poland
Puneet Gupta, Infosys Labs, India

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

The City Tour Service in Mobile Ad-hoc Group

HyunKyung Yoo and Changsup Keum
Communication Service Platform Research Section
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, Republic of Korea
hkyoo@etri.re.kr, cskeum@etri.re.kr

YoungMee Shin
Energy IT Technology Research Section
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, Republic of Korea
ymshin@etri.re.kr

Abstract— **With the recent explosive growth of mobile devices, the demand of high quality contents in mobile network is increasing. To ensure the low delay contents sharing in mobile network, there is a need to deploy service platform which provides the intelligent services in the localized ad-hoc group. The service platform can provide a common structure to create and support the direct services at user's edge node. Mobile subscribers come together in same area, establish a temporal ad-hoc group and they connect to the same Wi-Fi access point. People who have common interest can communicate for sharing contents and information. For them, service platform can support the storing, caching, and management of contents, in order to share the high quality contents among proximity group in mobile network. The feature of proximity group services has instantly generated social relation. In this paper, we propose the architecture of service platform for providing a city tour service. Also, we provide the design and the development of the proximity services which is composed of a contents sharing service and a proximity chatting service. Our platform and service can be applied to the campuses and malls as well as tour group.**

Keywords-service platform; ad-hoc community; proximity service; city tour service; contents sharing

I. INTRODUCTION

The demand of mobile data services which create, store, make use of contents is increasing [1]. To deliver the high quality contents services in mobile network, it is necessary to deploy caching server in the edge node and provide the contents sharing service among the localized ad-hoc group. It is profitable to support the low delay contents through an edge node and reduce the traffic to core network [2-5].

Two major technologies that allow wireless connection between mobile devices are Bluetooth and Wi-Fi. Both are available on most modern mobile phones, tablets, laptops, and other electronic devices. Wi-Fi can achieve much higher transmission speed and wider range than Bluetooth. So Wi-Fi technology has dominated for most high-bandwidth internet-based applications [6].

For the user-centric service, the service platform is essential to provide a common structure to create and provide the direct services at user's edge node over Wi-Fi [7-9]. Also, mobile subscribers come together in same area such as city tour, establish an ad-hoc group, and get a same interest. For them, the intelligent service platform can support the storing, caching, and management of contents, in order to share the high quality contents among proximity group in mobile network.

In this paper, we propose the architecture of service platform for providing a city tour service in mobile ad-hoc group. And, we provide the design and the development of the proximity services which is composed of a contents sharing service and a proximity chatting service.

This paper is organized as follows. Section II presents the architecture of service platform for providing a city tour proximity service. We explain the functions of social sharing unit, content sharing unit, chat server, content server, service discovery unit, and mobile device. And we describe the service activation sequence flow.

In Section III, we present the city tour proximity service which is composed of a contents sharing service and a proximity chatting service. And we show the proximity chatting service flow. Finally, Section IV summarizes and concludes the paper.

II. SERVICE PLATFORM FOR MOBILE ADHOC SERVICE

The city tour services are based on the localized ad-hoc group connected to Wi-Fi. These groups have instantly generated social relation, on the basis of interest rather than intimate relations. The proximity group members can share the information of same topics at the same place at the same time. Their shared contents are photo, audio, video, text, and file. Shared contents must be guaranteed user's anonymity and considered a lifetime because of a member's temporal relation.

This proximity group service can be very useful in moving vehicle. For example, the tour guide and tourists can share the information of tour spot, restaurant, hotel, and shopping place, by making the dynamic proximity community.

A. Service Platform Architecture

For the mobile ad-hoc group services, we designed the service platform which supports the content management and the social sharing.

 ISBN: 978-1-61208-428-2

The proposed service platform architecture is shown in Figure 1. There are two parts of the mobile device and the service platform. The service platform consists of Social Sharing Unit, Content Sharing Unit, Chat Server, Content Server, and Service Discovery Unit. And Mobile Device has Client App and Discovery Agent.

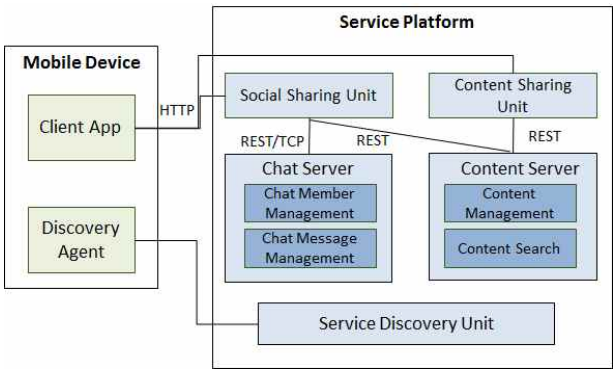


Figure 1. Service Platform Architecture

Users receive a city tour service in mobile device. The Discovery Agent of a mobile device discovers the connection information of the mobile device in the network. It gets the Access Pointer IP address and the service connection URL (Uniform Resource Locator). Client App of mobile device performs the service request and response according to user interaction.

Social Sharing Unit and Content Sharing Unit in service platform are web application for user interaction. They request the chat service or content service, and receive HTTP (Hypertext Transfer Protocol) based response from Chat Server or Content Server. They are composed of JavaScript, HTML5, and CSS (Cascading Style Sheets). JavaScript handles events, and HTML5 and CSS are the view part of web applications.

Chat Server provides the chat member management function based on REST interface and the chat message management function based on TCP socket connection. The chat member management function manages the member nicknames to guarantee an anonymity among the proximity ad-hoc community. It is the unique member identifier and it cannot be duplicated in the same Wi-Fi. The chat member management function performs the management to add and delete members. The chat message management function posts and stores the members' chat messages. If a member leaves a Wi-Fi network, the member's temporal relation disappears.

Content Server provides the contents management to cache, store, and delete contents, and the contents consumption like download or streaming. Also it supports the keyword based search.

Service Discovery Unit manages the service lists deployed to service platform and returns the service connection URL to mobile device.

B. *Service Activiation Sequence Flow*

For using the city tour service, user mobile device must be connected to Wi-Fi access point, and it gets a service connection URL. And then, mobile users are offered the city tour services from a client app deployed to mobile device.

Figure 2 shows these service activation sequence flow between mobile device and service platform. The related functions of service activation are service discovery unit, social chat unit, discovery agent, and client app.

The flows of connection to Wi-Fi network and the service activation are as follows.

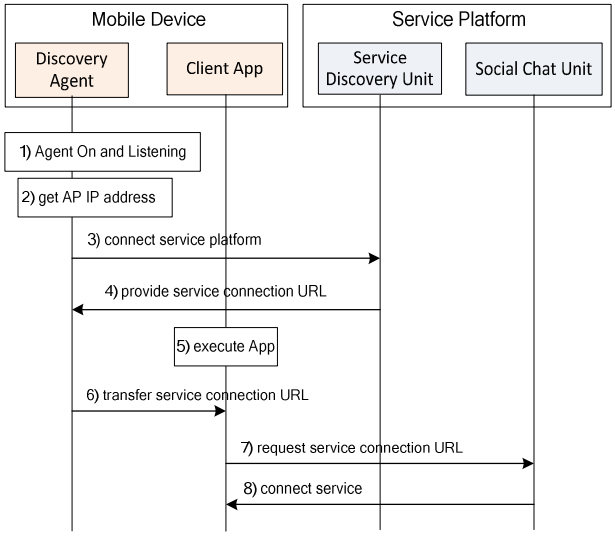


Figure 2. Service Activation Sequence Flow

- Discovery Agent of mobile device is on and waits to listen to signal.
- Discovery Agent gets the IP address of access point.
- Discovery Agent connects to Service Discovery Unit of service platform.
- Service Discovery Unit returns the service connection URL to Discovery Agent.
- Mobile user executes a Client App.
- Discovery Agent transfers the service connection URL to client app.
- Client App requests service connection URL.
- User confirms service connection and uses the service.

III. CITY TOUR SERVICE

We have implemented a city tour service which shows the temporal social relation features of service platform. Service development device was done on the Samsung Galaxy Note 8.0, running Android version 4.3 (Jelly Bean).

The demonstrated city tour service is composed of a contents sharing service and a proximity chatting service.

 ISBN: 978-1-61208-428-2

In contents sharing service, members can upload, list, and search the shared contents. The proximity chatting service provides the social sharing space to share information messages and contents among members who have same interest.

Figure 3 shows the proximity chatting service UI of city tour service. We define the intelligent agent which manages the chat member presence. When members enter the social sharing space with nickname, they can see the member lists and post the information messages. They upload contents and view shared contents and contents information.

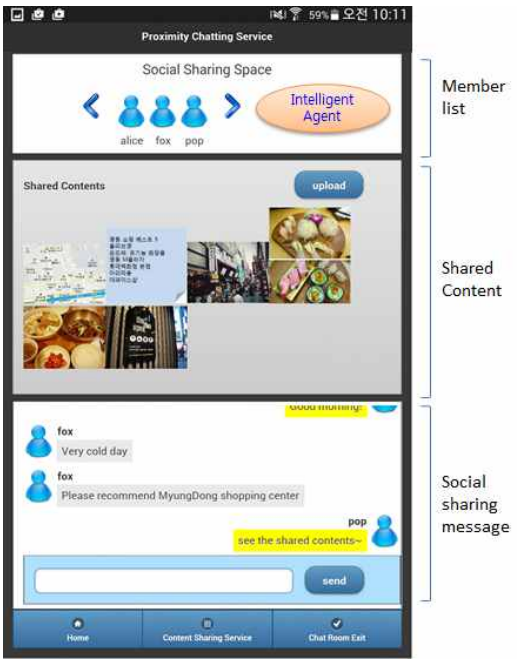


Figure 3. Proximity Chatting Service

The detail flow of proximity chatting service is shown in Figure 4. It shows the relation among components of service platform. A member creates the unique nickname and starts the proximity chatting service. Social Sharing Unit requests to get chat member lists and content lists. The chat server returns member lists and the content server returns the content lists.

When the members input the chat message for information sharing, the chat server transfers the messages to all members. While chatting, members can upload and share the related contents.

The advantage of the city tour service provides the useful information based on the temporal ad-hoc relation. People connected to the same Wi-Fi form an ad-hoc community, and people who have common interest can communicate for information sharing. Even if they are moving, service platform maintains the proximity community services.

Also, for privacy protection, when a person who uploaded content leaves a Wi-Fi network, the shared contents will be considered expired and disappear. It reflects the feature of ad-hoc services. Also, the shared contents among temporal members need to have an anonymity guaranteed.

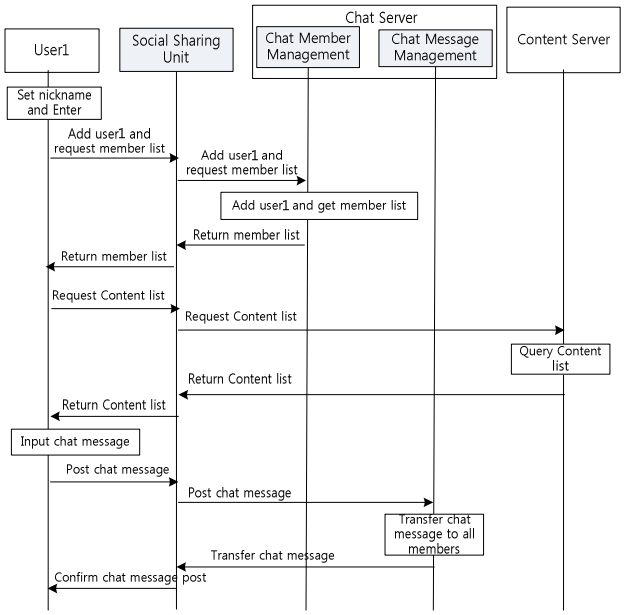


Figure 4. Proximity Chatting Service Flow

IV. CONCLUSION

In this paper, we presented the service platform for providing a city tour service in the mobile ad-hoc group. Our service platform supports the storing, caching, and management of contents, in order to share the high quality contents among the proximity ad-hoc group in mobile network.

We developed the city tour service, which is composed of a content sharing service and a proximity chatting service. The advantage of the city tour service provides the useful information based on the temporal ad-hoc community.

The lifetime of shared contents is not permanent because of a member's temporal relation. The chatting service based on user nickname can guarantee anonymity among ad-hoc members.

For further study, we have a plan to extend the service platform for the proximity intelligent advertisement and recommendation. It would be taken into account that platform provides the environment for advertisement and recommendation lifecycle management.

ACKNOWLEDGMENT

This work was supported by ETRI R&D Program funded By the Government of Korea (15ZI1410, Development of Technologies for Proximity, Real-time, and Smart Service Recommendation Platform).

 ISBN: 978-1-61208-428-2

REFERENCES

[1] Cisco white paper, “Cisco Visual Networking Index:Global Mobile Data Traffic Forecast Update, 2014–2019,” tech. rep., http://www.cisco.com, accessed Feb. 27th, 2015.

[2] Namiot, Dmitry, and Manfred Sneps-Sneppe. "Proximity as a service." Future Internet Communications (BCFIC), 2012 2nd Baltic Congress on. IEEE, 2012.

[3] Heredia, Edwin, et al. "Contextual proactivity for media sharing scenarios in proximity networks." Consumer Communications and Networking Conference (CCNC), 2014.

[4] Wang, Yufeng, et al. "Survey on mobile social networking in proximity (MSNP): approaches, challenges and architecture." Wireless networks, pp. 1295-1311, 2014.

[5] Salmon, Justin Lewis, and Rong Yang. “A Proximity-Based Framework for Mobile Services,” IEEE International Conference on Mobile Services, pp. 124-131, 2014.

[6] Kooh, Thomasgeorgescyrille, Qin Lv, and Shivakant Mishra. "Attribute Based Content Sharing in Mobile Adhoc Networks of Smartphones over Wifi." Computer Communications and Networks (ICCCN), 2012.

[7] Service Delivery Platform (SDP) definition in Wikipedia, http://en.wikipedia.org/wiki/Service_delivery_platform

[8] Pitkänen, Mikko, et al. "SCAMPI: service platform for social aware mobile and pervasive computing." ACM SIGCOMM Computer Communication Review, pp. 503-508, 2012.

[9] Rodriguez-Sanchez, M. C., et al. "GAT: Platform for automatic context-aware mobile services for m-tourism." Expert Systems with Applications, pp. 4154-4163, 2013.

 ISBN: 978-1-61208-428-2

An Analysis of the Relationship between Human Personality and Favored Location

Ha Yoon Song, Eun Byul Lee

Department of Computer Engineering,
Hongik University, Seoul, Korea
Email: hayoon@hongik.ac.kr, quftmxk@gmail.com

***Abstract*—It is a long-held belief that human personality affects the preferred locations of humans. We used statistical regression analysis in this research to examine the relationship between human personality and favored locations. The personality data is represented in accordance with the Big Five Factor (BFF) and location data was acquired from smartphone apps used by the volunteers. Instead of using raw location data, a total of 11,154 location datum was collected and then subject to further, more detailed categorization. The categorized location data were then classified into a higher-level aggregation. Finally, the relationship and the statistical results from regression analysis are presented, alongside categorized and classified location data. A summary of the relationship is provided, whereas a Coefficient of Determinant (CoD) is used for the results of the regression analysis.**

Keywords–Personality; Favorite location; Regression analysis, Personality-location relationship, Big Five Factors.

I. INTRODUCTION

The actions of humans can be predicted by time, personality factors, occupation, gender, and so on. Among these factors, human personality factors are typically fundamental in the prediction of human behavior, especially location preference. In this research, the effect that personality factors exert on a human's preferred location was examined. The five-factor model (FFM), a representative theoretical framework for understanding human personality, is comprised of the Big Five Factor (BFF), which can be obtained by using the Big Five Inventory (BFI) questionnaire set [1]. For this research, the BFFs of the 14 volunteer participants were obtained anonymously.

Nowadays, smartphones can be used to easily collect location or position data. Numerous smartphone functions including Global Positioning System (GPS), GLONASS, and indoor positioning systems are used for positioning, and humans can use smartphone apps to easily check-in wherever they are located. The most popular location apps are Foursquare and Swarm. This location data can be categorized into higher-level categories using the categories-id; for example, Vietnamese restaurant and Japanese restaurant can be categorized in the higher-level category of Restaurant, while Coffee house and Milk tea shop can be categorized in Beverage store. These categories can be identified by the categories-id. However, for the purpose of detailed research, lower-level categories are also used in this paper.

Regression analysis [2] was a helpful tool for examining the relationship between BFF and the categorized locations. Collected data sets were formed as variables for the regression analysis. In this research, an examination of the relationship between human personality and location preference was conducted. Apart from other research, regression analysis on personality data and location data were utilized. Mobile services which are strongly dependent on location can utilize the benefit of this research. For example, the result of this research will be applied for recommendation system which recommends a specific travelers attraction to a person with certain personality. A possible scenario is that a travel recommendation system based on our approach, which can match the traveler's personality with a property of travelers' attractions and can recommend a specific travelers attraction which fits the traveler's personality.

This paper is organized as follows. Section II will discuss about related works on location and personality. In Section III, personality data, location data, and the categories of locations are described as parts of the preparation for the regression analysis. Section IV will analyze the results of the regression analysis and will present examples of the way in which the BFF affects locations. Section V will present the conclusion and possible future research topics.

II. RELATED WORKS

There has been little research into how our personality relates to the categories of place that we visit. There are mainly two reasons. The first one is related with personal information protection, meaning peoples do not like to totally reveal personal information and only volunteers help to collect personality data without any legal problem. The second one is that the use of mobile device is not general at all, even though the smartphones are widely spread. This restricts the collection of location data. Therefore, only volunteers with mobile devices can provide their personality data and location data. Despite of these barriers of data collection, several related researches have found in last few years.

The relationship between social network and human personality are conducted in [3]. The effect of human personality on Facebook usage is researched. Self-report of participants are used to extract object criteria and measurements from Facebook data, and the result shows that strong relationship between human personality and actions on Facebook.

Location Based Social Networks (LBSN) are being started in this research area nowadays. One of the research [4] is mostly related to our research. Personality factors and foursquare check-ins are investigated and utilized, and the relationship is founded. Conscientiousness, Openness, and Neuroticism are found related to specific locations in this research.

 ISBN: 978-1-61208-428-2

As a result of this research, the combination of personality factors and LBSN can explain the human personality.

Regression analysis on personality data and location data can be rarely found as shown in [5] in order to predict the future human location with five volunteers' data and Back Propagation Network (BPN). An opposite of this research can be found in [6] which deduces personality factors form location data of volunteers. BFFs can be deduced from sets of location data also by BPN. Apart from the other researches, our research is a direct match between BFF (personality) and location categories. Personality data and location data collected by volunteers themselves are used with minimal treatment of the raw data.

In the next sections we will present our method to reveal the relationship between personality and location preference.

III. Personality and Location Data Preparation

Two different assortments of location data will be addressed in this section. The first method is a classification of location data. From the location data source, several information such as frequency of visit, name of location, category-id will be extracted. The second method is the categorization of location data. From the location data source, several information such as frequency of visit, name of location, name of location categories will be extracted. Note that categorization of location data is a more ramification of location data comparing to classification of location data.

Total fourteen volunteers provided their personality and location information. All volunteers are anonymous for personal information protection and called by their numbers. Most of the volunteers are university students and are in their twenties. University students tend to go out lively and visits university, restaurant, and other places frequently. BFF of each volunteer is identified by BFI. Due to the nature of volunteers that they are university students, places in university campus are taken special care and categorized as detail as possible.

TABLE I. BFF OF VOLUNTEERS

	O	*C*	*E*	*A*	*N*
Volunteer 1	3.30000	3.88889	3.25000	3.66667	2.62500
Volunteer 2	3.60000	3.33333	2.75000	3.22222	2.75000
Volunteer 3	2.70000	3.22222	3.25000	2.66667	2.75000
Volunteer 4	4.33333	3.12500	2.25000	3.20000	2.88889
Volunteer 5	4.20000	4.33333	3.50000	3.55556	2.62500
Volunteer 6	4.00000	3.66667	4.00000	3.88889	2.75000
Volunteer 7	3.50000	3.77778	3.37500	3.22222	3.00000
Volunteer 8	2.20000	3.44444	3.00000	3.11111	2.62500
Volunteer 9	2.60000	2.77778	3.37500	3.11111	2.62500
Volunteer 10	3.30000	2.88889	3.12500	3.11111	3.25000
Volunteer 11	3.40000	3.22222	3.37500	3.33333	3.12500
Volunteer 12	3.10000	3.66667	3.37500	3.22222	3.50000
Volunteer 13	3.40000	3.55556	3.62500	2.88889	2.50000
Volunteer 14	3.80000	4.00000	3.12500	3.77778	2.25000

A. *Personality data*

Research has been conducted on the various theories that seek to represent human personality. Among these theories, McCrae and Costa's [7] research on BFF is one of the foremost studies on personality representation, whereby the five factors are Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness (A), and Neuroticism (N). This distinguished Five Factor Model (FFM) is composed with BFF which can be obtained by Big Five Inventory (BFI) [8]. In order to obtain BFF of volunteers, volunteers answer the questionnaire set called BFI which is usually composed of 44 questions and the answers are selections of a degree from one to five. The BFI is designed to calculate every five factors of BFF from a combination of degrees of the 44 answers. The degree of each five factor is designed to be represented from one to five. A major merit of BFF is that the factors can be quantized and it reflects the concept of normal distribution, i.e., each quantitative factor is in a numerical form that stands for positions of normal distribution. Due to the properties of BFF, corresponding numerical data can be utilized by any algorithm and, of course, it can be applied to regression analysis as shown in this research. It is notable that each factor of BFF is mutually orthogonal, which stands that factors do not interfere each other and thus can be used as independent variables in regression analysis.

TABLE II. SAMPLE LOCATION DATA OF VOLUNTEER 1

CoV	*Location name*	*categories-id*
75	Hongik Univ.T-819	4bf58dd8d48988d19e941735
4	Starbucks	4bf58dd8d48988d1e0931735
3	CGV	4bf58dd8d48988d180941735
2	Hongik Univ.	4bf58dd8d48988d1a8941735
1	Everland	4bf58dd8d48988d182941735

Table I shows BFFs for 14 volunteers, and the values can be used to easily interpret the personality of a human. For example, in Table I, Volunteer 4 has the highest Openness rating among the 14 volunteers, implying that this volunteer is highly creative in the arts, highly intellectual, curious, and adventurous. The other BFF ratings of volunteer 4 imply planning skills, propensity for action, and diligence in regard to work. Volunteer 6 has the highest Extraversion rating, implying an energetic, verbose, and positive-minded personality.

The above BFF of Volunteers data set will be used for this research henceforth.

TABLE III. SAMPLE CLASSIFICATION OF LOCATIONS OF VOLUNTEER 1

categories-id	*Name*	*ICC*
4bf58dd8d48988d16d941735	Beverage Store	5622
4bf58dd8d48988d110941735	General Restaurant	5611
4bf58dd8d48988d19e941735	Institutions of Education	8530
4bf58dd8d48988d1e2931735	Museum and Historic Site Managing	9022
4bf58dd8d48988d1d0941735	Other Restaurant	5619
4bf58dd8d48988d182941735	Theme Park Operations	9121

B. *Classification of location data*

Location data, as different from positioning data, represent conceptual locations such as home, restaurant, school, and so on. Positioning data is typically numerical data, including latitudinal and longitudinal pairs. The location data in this research were collected using the smartphone apps named Foursquare and Swarm. Volunteers checked-in on their preferred app whenever they visited a meaningful location. The app then collated the location data based on the source data provided by the volunteers, including count of visit (CoV), the location's name, and the categories-id. Table II shows a sample of location data from volunteer 1 as an example of the data representation of this study.

In Table II, both the formal and numerical forms of the location information are presented; for example, two visits

 ISBN: 978-1-61208-428-2

TABLE IV. COUNT OF VISIT (CoV) TO THE LOCATIONS AND ICC OF VOLUNTEER 1

ICC	*8530*	*5611*	*5619*	*5291*	*5622*	*5621*	*5914*	*9900*	*9022*	*4712*	*9112*	*9121*
CoV	30	22	13	6	6	5	3	3	2	1	1	1

TABLE V. VOLUNTEERS' LOCATION DATA

Volunteer	*FI*	*RB*	*TI*	*Restaurant*	*Bar*	*BS*	*Bank*	*PAI*
1	0.15008	0.04384	0.04384	0.27487	0.02529	0.10118	0.00337	0.00000
2	0.16484	0.06154	0.01319	0.32527	0.02637	0.09890	0.00439	0.00219
3	0.03030	0.09848	0.05303	0.63636	0.02272	0.02272	0.00000	0.00000
4	0.29339	0.01222	0.02119	0.18744	0.00570	0.16299	0.00081	0.00081
5	0.30158	0.00000	0.02116	0.19047	0.01587	0.04232	0.00000	0.00000
6	0.12931	0.02586	0.02874	0.28735	0.02298	0.10344	0.00000	0.00000
7	0.13500	0.00500	0.07000	0.18500	0.01500	0.01000	0.00000	0.00000
8	0.27003	0.04748	0.01187	0.16617	0.01186	0.04154	0.00593	0.00593
9	0.24915	0.20478	0.02048	0.15017	0.00682	0.07508	0.01023	0.00341
10	0.18143	0.02109	0.08439	0.22784	0.03375	0.10126	0.00000	0.00000
11	0.26042	0.01042	0.01042	0.22916	0.03125	0.03819	0.00000	0.00000
12	0.20952	0.11429	0.04286	0.09523	0.00476	0.04285	0.00952	0.00476
13	0.18705	0.03597	0.18705	0.17266	0.01438	0.01438	0.00000	0.02877
14	0.03904	0.14571	0.28619	0.15262	0.04303	0.02951	0.00399	0.00860
Volunteer	*LS*	*IE*	*Hospital*	*CH*	*LM*	*SF*	*TP*	*BB*
1	0.00337	0.30860	0.00337	0.01349	0.01855	0.00505	0.00505	0.00000
2	0.00000	0.27033	0.00879	0.00879	0.01098	0.00219	0.00219	0.00000
3	0.00000	0.09848	0.02272	0.01515	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.28280	0.00244	0.00733	0.00570	0.00733	0.00489	0.00489
5	0.00000	0.39153	0.00000	0.01587	0.01587	0.00529	0.00000	0.00000
6	0.00287	0.37931	0.00862	0.00000	0.00862	0.00287	0.00000	0.00000
7	0.00000	0.57000	0.01000	0.00000	0.00000	0.00000	0.00000	0.00000
8	0.00000	0.41246	0.00296	0.00890	0.00890	0.00000	0.00593	0.00000
9	0.00000	0.19453	0.01706	0.00682	0.01706	0.03413	0.00000	0.01023
10	0.00000	0.23628	0.00843	0.01687	0.02109	0.06751	0.00000	0.00000
11	0.00000	0.29513	0.01736	0.05556	0.00347	0.03819	0.01041	0.00000
12	0.00000	0.42381	0.00476	0.02857	0.01428	0.00000	0.00476	0.00000
13	0.00000	0.28777	0.00719	0.01438	0.03597	0.00000	0.00719	0.00719
14	0.00292	0.07254	0.06947	0.04980	0.04857	0.02797	0.01644	0.00353

were made to a location named Hongik Univ.T-819, which has a predefined categories-id of 4bf58dd8d48988dla8941735.

According to the Standard Industrial Classification, higher categories can be created by merging subcategories using the categories-id; for example, Vietnamese restaurant and Japanese restaurant can be merged into the higher category of Restaurant. In another case, Coffee house and Milk tea shop can be merged into Beverage Store, which is a higher level category also classified in the Standard Industrial Classification. Using Standard Industrial Classification prepared by the National Tax Service and the Industry Classification Code (ICC), locations can be classified into higher-level ICC categories. Table III shows several samples of the location classifications using the data from Volunteer 1. Table III shows the categories-id, the names of the classified categories, and the ICC; for example, Hongik Univ. T-819's categories-id is 4bf58dd8d48988d19e941735, and after it was merged into Institutions of Education, a higher classification, the corresponding ICC became 8530. Similarly, categories-id 4bf58dd8d48988d182941735 was merged into Theme Park Operations with an ICC of 9121.

Table IV shows extracts of CoV values to each ICC; for example, one visit was made to the ICC location of 4712 and 22 visits were made to the ICC location of 5611. Table II, Table III, and Table IV show the sequence of the location-classification process for Volunteer 1.

It was necessary to arrange all of the data from the 14 volunteers in the same way. Table V summarizes the classification results of every volunteer. The meanings of the location-classification acronyms in Table V are:

- FI = Foreign Institutions
- RB = Retail Business
- TI = Travel Industry
- BS = Beverage Store
- PAI = Public Administration Institutions
- LS = Leasing Service
- IE = Institutions of Education
- CH = Concert Hall
- LM = Library, Museum
- SF = Sports Facilities
- TP = Theme Park
- BB = Beauty Business

Every number in Table V represents the ratio of visits according to (1), as follows:

$$VisitingRatio(loc) = \frac{count\_of\_visit\_to(loc)}{total\_count\_of\_visit} \quad (1)$$

where *loc* is one of the locations in Table V.

For example, the CoV of Volunteer 1 to Foreign Institute is 89, while the total CoV is 593; therefore, visiting ratio (FI) = 0.15008. As shown in Table V, Volunteer 1 has a visiting ratio of 0.15008 for Foreign Institute. In the case of Concert Hall, Table V shows zero values for Volunteer 6 and Volunteer 7, while other volunteers have a non-zero visiting ratio, meaning that more than one visit was made.

 ISBN: 978-1-61208-428-2

C. Categorization of location data

From the source check-in data, each location category was parsed, extracting the CoV, the location name, and the categories-name. Table VI shows several of the extracted samples, including one example where 58 visits were made to a location named Hongik Univ.T, which is categorized as University Buildings.

Instead of considering each location with the standards of location classification, the CoV was aggregated for each categories-name as shown in Table VII. Table VII shows CoV according to non-overlapped categories-name. Also, Table VII has more detailed information for Volunteer 1 than Table III; for example, bus terminals are clearly separated from train stations, i.e., categorized rather than classified. An interpretation of Table VII shows that Volunteer 1 visited Home 80 times, University Buildings 76 times, and Korean Restaurant 39 times. The sample data shown in Table VI and Table VII was collected only from Volunteer 1. Table VII shows the location data only from volunteer 1, meaning that there is a total of 14 similar tables for each volunteer.

In sum, 96 categories-names were derived from the data of the 14 volunteers with some overlapping of categories-names. After the Standard Industrial Classification and ICC, location categories with a total CoV less than 10 for all volunteers were merged into Etc. category. The final data set consists of 55 categories-names of the 14 volunteers and provided dependent variables for the regression analysis.

TABLE VI. SAMPLE OF CATEGORIZED DATA OF VOLUNTEER 1

CoV	*Location name*	*Categories-name*
58	Hongik Univ.T	University Buildings
5	CGV Ori	Theater
4	Songtan yeongbillu	Chinese Restaurant
3	Shinsegae Department Store	Department Store
3	TongtongPig	BBQ shop

TABLE VII. VOLUNTEER 1: REPRESENTATIVE CATEGORIES-NAME AND CORRESPONDING COUNT OF VISIT

Categories-name	*CoV*	*Categories-name*	*CoV*
University Lab	83	Language School	4
Home	80	Train Station	4
University Buildings	76	Art Museum	4
Cafe	60	Bank	4
Korean Restaurant	39	Fast Food Shop	4
etc.	39	Bus Stop	3
University Library	21	Cosmetics Shop	3
Dessert Shop	17	Sushi Shop	3
Outdoor Activity	15	Subway	3
Bar	12	Grocery Store	3
Chinese Restaurant	11	Shopping Centre	3
Snack Shop	11	Clothing Store	2
Noodles Restaurant	8	American Restaurant	2
Fried Chicken Shops	8	Hospital	2
BBQ Shop	8	Convenience Store	2
Japan Restaurant	8	Museum	1
Italy Restaurant	6	Gymnasium	1
Bakery	6	Salad Section	1
Theater	6	Electronics Stores	1
Airport	5	University Stores	1
Rest Area	5	Asian Restaurant	1
Department Store	5	Pizzeria	1
Bus Terminal	5	Seafood Shoppe	1

IV. Relationship Result Analysis

In this section, two sorts of location data set from Section III will be regarded as two sets of dependent variables. Corresponding results will be addressed.

A. Regression analysis on location classifications

A basic regression analysis was conducted to further develop the examination. Each volunteers' BFF values, as shown in Table I, was regarded as an independent variable. The location information, as shown in Table V, was regarded as a dependent variable. The regression analysis showed the effect of personality on location preference. The basic result is shown in Table VIII using Coefficient of Determination (CoD).

Table VIII contains the following acronyms:

- CoD = Coefficient of Determination
- IE = Institutions of Education
- PAI = Public Administration Institutions

The relationship between each variable can be represented by CoD. In Table IX, the range of CoD and the interpretations of relationships are presented. In cases where the CoD value is higher than 0.4, the relationship was considered to be high. According to this criteria, a case with a CoD value higher than 0.4 was regarded as one in which the personality and location are meaningfully related.

TABLE VIII. RESULTS OF REGRESSION ANALYSIS FOR CLASSIFIED LOCATIONS

Location Classification	*CoD*	*Location Classification*	*CoD*
Leasing Service	0.71391	Library, Museum	0.31678
IE	0.64162	Retail Business	0.26152
Restaurant	0.55895	Travel Industry	0.25720
Bank	0.52616	Foreign Institutions	0.23086
Beauty Business	0.49252	Theme Park	0.22845
PAI	0.45437	Bar	0.17817
Sports Facilities	0.39169	Beverage Store	0.15551
Hospital	0.32293	Art Viewing Facilities	0.09206

TABLE IX. RELATIONSHIP DETERMINED BY COEFFICIENT OF DETERMINATION

Range of CoD	*Interpretation*
0.9-1.0	Very High
0.7-0.9	High
0.4-0.7	Relatively High
0.2-0.4	Low
0.0-0.2	Rarely Related

Table VIII shows location classifications and the corresponding CoD. The location classifications of Restaurant, Bank, Public Administration Institutions, Leasing Service, Institutions of Education, and Beauty Business show CoD values higher than 0.4, meaning that only 6 classifications showed a meaningful relationship between personality and location.

It may be meaningful to see the effect of the subcategories of the location classifications. The next subsection will show the results for location categories.

B. Regression analysis on location categories

The regression analysis on location categories found a higher number of meaningful relationships. BFF values, as shown in Table I, were also used as an independent variable, while the total 55 location categories were used as a dependent variable. The regression analysis showed that 31 location categories had a CoD higher than 0.4, indicating a meaningful relationship between personality and location.

 ISBN: 978-1-61208-428-2

Table X shows a selection of meaningful results from the regression analysis and, in addition to the CoD, a coefficient for Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (O, C, E, A, N) was used.

The coefficient of O, C, E, A, N shows the degree of personality's effect on corresponding location categories. First, the absolute value of each coefficient stands for the strength of affection: positive coefficient means positive effect, while negative coefficient means negative effect.

The criteria of each coefficient are determined to be relative when the absolute value of the coefficient is more than 0.01, and strongly related when the absolute value coefficient is more than 0.1. Table X contains a column titled RELATION which represents the effect of personality on the location categories. The RELATION column shows the final results of this research. As a factor of meaningfulness, the plus sign (+) stands for the positive relation to the corresponding location, and the minus sign (-) stands for the negative relation to the corresponding location. For the strong positive relation, (++) is used, while the strong negative relation is indicated with the use of (- -). In addition, p-value must be considered. Once p-value is greater than 0.05 for a relation, the personality factor will be statistically disregarded and will not be presented in RELATION column.

As an example, University Buildings have an O, C, E, A, N coefficient of -0.13274, 0.07956, 0.01043, 0.00477, 0.21439, respectively. Openness and Neuroticism, which are also two factors of BFF, show an absolute value more than 0.1; therefore, - -O and ++N show that Openness strongly affects the location of University Buildings positively and Neuroticism strongly affects University Buildings negatively. Conscientiousness and Extraversion have coefficient values more than 0.01 but lower than 0.1; therefore, +C and +E are also applicable to the location category of University Building. Overall, the corresponding relationship column may contain - -O, +C, +E, and ++N. However, +C and +E must be excluded since their corresponding p-values are greater than 0.05.

Some of the RELATION columns are remained blank because the corresponding relation is negligible, whereby the coefficients of O, C, E, A, N are less than 0.01. The corresponding locations may be related to BFF since CoD is bigger than 0.4; however, the corresponding coefficient for each BFF is less than 0.01. For example, in the case of Museum, the coefficients for O, C, E, A, N are $0.00250, 0.0008, 0.00402$, -0.00828, -0.00541, respectively. Even though Agreeableness has the largest absolute coefficient value, implying a negative relationship, it is not shown.

In terms of restaurants, categories such as Korean Restaurant, Asian Restaurant, and Japanese Restaurant sometimes showed a strong relationship with BFF, while American Restaurant and Italian Restaurant showed a more minor relationship with BFF. This phenomenon might have been affected by CoV, which suggests that more data needs to be collected.

As shown in Table VII, the highest CoVs can be seen for Home and University, which is unsurprising considering that most of the volunteers were students at the time of the study. However, in Table X, University-related locations are present, while a Home location is not. The CoD for Home is less than 0.4, so maybe the location Home has a neutral relationship with human personality. Such a large CoV for University, University Buildings, University Library, University Lab, and University Stores should be categorized and more closely investigated.

Table X has the highest CoD for University Buildings of 0.87732, and O and N are RELATED for University Buildings. A similar BFF pattern could possibly be found for workers' offices. In the case of Fast Food Shop, which has a CoD of 0.75860, negative Extraversion have relations but they are not strong.

From the results in Table X, Openness, Extraversion and Neuroticism are affective factors for location visit while Conscientiousness is the negligible factor for location visit. Maybe the more data will reveal more relationship between BFF and locations since it cannot be guaranteed that our data covers all possible locations or covers all possible combinations of personality factors.

V. Conclusion

Regression analysis is usually an acceptable tool to use when deducing statistical fitness in a linear relationship between independent and dependent variables. BFF values were used as independent variables in our research, while categorized and classified location data were regarded as dependent variables. The result of the regression analysis is meaningful and, for several locations, the effect of personality was revealed. It is meaningful that we used a scientific method, instead of relying on conjecture, to identify the effect of personality on several specific locations. We also identified an aspect of the functional relationship between human personality and location preference.

Our research can be applied to many of related areas in order to increase imposed value of each area, such as enhancement of mobility model [9][10], personal travel recommendation system [11], mobile computing [12] and so on. The key is personalized future Internet, especially for location based social network related area.

An implication of this relationship emerged after the basic stage of our research, when we had obtained 2,842 location datum. Eventually, the total of 11,154 location datum was used to provide more meaningful research results. It is possible that, if more volunteer participants check-in at more locations, further insight into the personality-location relationship will be gained.

However, our research is likely to have biased results in two reasons. Even though we collected enough number of locations visited, only 14 volunteers are participating and we have a small set of BFFs comparing to the size of location data. It is an unavoidable bias and will be solved with more volunteers are involved as we continue our research. Sometimes the bias can be intentional supposing the travel recommendation system which we mentioned as an application, where travelers can be considered mostly. The second bias is that most of our volunteers are university students, which leads to lots of check-ins to university buildings. Due to this bias, however, we obtained meaningful results on the relationship between personality and university buildings. Since we continue collecting more BFFs, we expect that more distinct relationship can be found.

The more proper method of our research may be stepwise regression. Independent application of independent variable from BFF by stepwise regression will give more clear result

 ISBN: 978-1-61208-428-2

TABLE X. THE EFFECT OF PERSONALITY TO LOCATIONS WITH COEFFICOENT OF DETERMINATION

Location	*CoV*	*CoD*	O	C	E	A	N	*RELATION*
University Buildings	741	0.87732	-0.13274	0.07956	0.01043	0.00477	0.21439	- -O, ++N
Fast Food Shop	55	0.75860	0.00467	-0.01153	-0.02299	-0.00340	-0.00628	-E
Korean Restaurant	690	0.72413	0.08739	-0.03577	-0.03540	-0.02518	-0.02823	+O
University Library	104	0.66507	-0.02990	0.02783	0.00124	-0.02349	0.01764	-O
Asian Restaurant	50	0.64682	0.01263	-0.00389	-0.02293	-0.00772	0.00011	-E
Seafood Shoppe	31	0.64301	0.00038	0.00223	-0.00106	0.00057	-0.00158	
Sushi Shop	31	0.64136	0.00074	-0.00010	-0.00411	0.00259	-0.00146	
Cosmetics Shop	14	0.63816	-0.00118	-0.00579	0.00513	-0.00517	-0.00851	
Japan Restaurant	126	0.60027	0.00224	-0.01117	-0.01006	0.00342	0.00876	
Pizzeria	10	0.59626	-2.00E-05	-0.00668	-0.00069	0.00297	-0.00254	
Hotel	86	0.57299	0.00773	-0.00104	0.00437	-0.01647	-0.01546	-N
Language School	34	0.57297	-0.01789	-0.01246	-0.00626	0.01823	-0.00212	-O
Bank	42	0.56085	-0.00595	0.00112	-0.00181	0.00366	0.00048	
BBQ Shop	209	0.54467	0.02409	0.00055	0.00260	-0.02265	-0.00241	+O
Snack Shop	79	0.53491	-0.00561	-0.00752	-0.00728	0.01758	0.00432	+A
University Lab	731	0.52897	0.12886	0.08422	0.10599	-0.07828	-0.00644	
Electronics Stores	18	0.52038	0.00352	0.00499	-0.00526	-0.00484	0.00448	
Beauty Salon	24	0.51725	0.00083	-0.00273	0.00226	-0.00261	-0.00528	
Department Store	35	0.51175	-0.00407	-0.00902	-0.00134	0.00555	-0.00569	
Airport	103	0.51101	0.00615	0.00204	0.00314	-0.01302	-0.01522	-N
Rest Area	81	0.50978	0.00911	-0.00300	0.01236	-0.01191	-0.01270	
Outlet Store	11	0.48241	-0.00200	-0.00883	0.00328	0.00314	-0.00608	
Chicken Shops	90	0.48083	0.00429	-0.00713	-0.00655	0.00452	0.00644	
Public Institution	53	0.47681	-0.00358	0.00170	-0.00223	0.00301	-0.00098	
Museum	16	0.47463	0.00250	0.00081	0.00402	-0.00828	-0.00541	
University Stores	14	0.45565	-0.00135	-0.00779	0.00133	0.00766	-0.00356	
Pharmacy	53	0.44116	-9.90E-05	-0.00812	-0.00120	0.00504	-0.00706	
American Restaurant	48	0.43308	0.00320	-0.00392	-0.00216	0.00355	0.00161	
Italy Restaurant	61	0.43121	-0.00162	-0.00147	0.00241	0.00909	0.00281	
Salad Section	14	0.42998	-0.00087	0.00070	-0.00059	0.00158	0.00343	
Convenience Store	279	0.42796	-0.00051	0.00194	-0.01514	0.01295	-0.00944	

about the effect of each personality factors on location preference based on the orthogonality of each five factor.

One more possible research would include temporal information as well as location information. The check-in time and check-in duration will provide more information about the mobility pattern of each volunteer and our understanding of the personality-location relationship will become more precise. The inclusion of other personal factors in future research, including occupation, income, and gender, studies will give keener insight into this relationship.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (NRF-2012R1A2A2A03046473).

REFERENCES

[1] O. P. John and S. Srivastava, "The big five trait taxonomy: History, measurement, and theoretical perspectives," Handbook of personality: Theory and research, vol. 2, no. 1999, 1999, pp. 102–138.

[2] H. Smith, N. R. Draper, and H. Smith, Applied regression analysis. Chapman and Hall., 1981.

[3] Y. Amichai-Hamburger and G. Vinitzky, "Social network use and personality," Computers in human behavior, vol. 26, no. 6, 2010, pp. 1289–1295.

[4] M. J. Chorley, R. M. Whitaker, and S. M. Allen, "Personality and location-based social networks," Computers in Human Behavior, vol. 46, 2015, pp. 45–56.

[5] S. Y. Kim and H. Y. Song, "Predicting human location based on human personality," in Internet of Things, Smart Spaces, and Next Generation Networks and Systems. Springer, 2014, pp. 70–81.

[6] ——, "Determination coefficient analysis between personality and location using regression," in International Conference on Sciences, Engineering and Technology Innovations, 2015, pp. 265–274.

[7] P. T. Costa and R. R. McCrae, "Four ways five factors are basic," Personality and individual differences, vol. 13, no. 6, 1992, pp. 653–665.

[8] L. R. Goldberg, "The structure of phenotypic personality traits." American psychologist, vol. 48, no. 1, 1993, p. 26.

[9] D. Alberg, M. Last, and S. Elnekave, "A spatio-temporal simulation model for movement data generation," in Data Mining Workshops, 2008. ICDMW'08. IEEE International Conference on. IEEE, 2008, pp. 320–325.

[10] D. Guo and W. Cui, "Mining moving object trajectories in location-based services for spatio-temporal database update," in Sixth International Conference on Advanced Optical Materials and Devices. International Society for Optics and Photonics, 2008, pp. 71 432M–71 432M.

[11] Y. Doytsher, B. Galon, and Y. Kanza, "Storing routes in socio-spatial networks and supporting social-based route recommendation," in Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks. ACM, 2011, pp. 49–56.

[12] W.-C. Peng and M.-S. Chen, "Developing data allocation schemes by incremental mining of user moving patterns in a mobile computing system," Knowledge and Data Engineering, IEEE Transactions on, vol. 15, no. 1, 2003, pp. 70–85.

 ISBN: 978-1-61208-428-2

A Web-based System to Manage Primary Students' Homework
A Case Study for the Teaching of Portuguese

M. Candeias[1], M.G.A.D. Reis[2], E. Peres[3], J.J. Escola[1], M.J.C.S. Reis[4]
[1] University of Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal.
E-mails: {jmpcandeias, mgentil.reis}@gmail.com, {eperes, jescola, mcabral}@utad.pt
[2] Escolas João de Araújo Correia, Centro Escolar da Alameda, Peso da Régua
[3] INESC-TEC and UTAD.
[4] IEETA and UTAD.

***Abstract*—Technology has profoundly changed the way we learn and live. Internet provides a new paradigm, setting forth that education and communication approach is truly more complex than ever before. Here, we present a web-based system to support students' homework. In order to prove its usefulness, we present the results achieved with the application of a set of exercises to help improving Portuguese (mother tongue) language skills of 4th-grade primary pupils. This set of digital exercises proved to be the child's favorite, rather than their counterpart exercises in paper form, also showing a fair more positive attitude from the students' point of view. A simple survey also showed that students do prefer the use of the system here presented, when compared to other more traditional ways of practicing. We also believe that the prompt feedback about the exercises correctness, together with the training provided by the different exercises sets about the same subject, besides the exploitation of video, color, sound, etc., positively reinforce the diverse child's senses, definitely contributing to capture and motivate the child.**

Keywords–Web-based homework system; improving mother tongue teaching/learning; primary education

I. INTRODUCTION

Technology has profoundly changed the way we learn and live. This relationship appears to be quite complex, in that IT, and especially socially- and technologically-rich learning environments, seems to both require and foster skills and learning. Internet provides a new paradigm, setting forth that education and communication approach is truly more complex than ever before.

It is on the basis of such paradigm that we present a case study where a set of exercises were used in order to improve the Portuguese (mother tongue) language skills of 4th-grade primary pupils. Being part of a web-based system to support students' homework [1], a set of digital exercises proved to be the child's favorite, rather than their counterpart exercises in paper form, also showing a fair more positive attitude. To the best of our knowledge, there is no system like the one presented here available.

The paper is organized as follows. The next section is used to present some related work. Section III is used to present some of the advantages of using the web-based system presented here to help with the student homework, and related to the Portuguese mother tong teaching/learing. In Section IV, we present the theoretical model used to support our research. In Section V, we present an overview of the web-based system, and also its developing context. In Section VI, we present some of the main characteristics of the set of templates available in the system and used to create the exercises offered to the students during this study, and also the data and some of our main achievements. Finally, in Section VII, we present the major conclusions achieved with this work.

II. RELATED WORK

As far as our knowledge, there are no systems using the same approach and ideas as those used in the system presented and discussed in this paper. However, there are many web-sites and tools available for helping the teaching of virtually everything.

Concerning the class of available web-sites we want to highlight the following ones. The "Coursera" site/system [2] offers free courses online from world-class universities. The topics are varied and lectures are formatted into series of 15-minute-long clips, and students can watch videos at their convenience and in their own time. "Learnist" is the "Pinterest" of learning [3]. The interactive platform allows users collect teaching materials and educational content that are grouped into "boards". The platform is equipped with the "Learn it!" bookmarklet, which enables picking images anywhere on the web and automatically shoots them to your "Learnist" board. The "Open Culture" [4] system is a high-quality cultural and educational media. A collection of 1,100 courses, online, for free, from Ivy League universities, such as Stanford, UCLA, Columbia and Oxford University is available. Courses run the gamut from science and art to math and economics. "Brazilian" Portuguese is one of the 48 languages available to learn for free. However, the type of contents available is completely different from the ones proposed here. "Udacity" [5] is free education web-site for brainy types, founded by four Stanford roboticists, and currently offers 11 courses, all of which are in science and math-related topics. According to the web-site, plans are underway to expand the curriculum. Udacity is free of deadlines, free of prerequisites, free of quizzes and other annoying school stuff. Courses are also free of charge. TED (Technology, Entertainment, Design), the powerhouse of jaw-dropping lectures needs no introduction. Now, they're bringing their talent into education with an offshoot, Ted-Ed [6]. A treasure trove of beautifully animated and gripping videos on a wide array of subjects, such as "The Power of Simple Words" and "How Many Universes are there", is available. All videos are under 10-minutes-long. Supplemental materials, such as quizzes and discussion questions, are also available. Unfortunately, we were unable to find any lessons on Portuguese in the 108,138 lessons available. Effortless learning based on three pillars science, fun and community is the basis for the "MemRise" site [7]. Based on scientific methods for implanting new information, "MemRise"

 ISBN: 978-1-61208-428-2

proves that what may seem impossible is in fact doable. The "Guardian Wallcharts" are mapped "MemRise" style, teaching you all about cheeses, herbs and other wonders of life. The "MentorMob" site/system [8] is like the "YouTube" of learning materials: users create learning "playlists" from first-rate web-sites. MentorMob is a community, whose members share and rank each others' learning playlists.

There are many "educational web-sites" for kids/students available on the Internet. In the site "Wonderopolis" [9], a kid/student may discover a new wonder each day. Wonders are amazing facts and intriguing questions such as "Why are they called Lava Lamps?" or "What Badger has a sweet tooth?" can be answered. Surfing the web-site "Whyville" kids/students can learn through playing games and socializing [10]. J.K. Rowling created the site "Pottermore" so kids could read the books and do interactive features and games [11]. The site "National Geographic Little Kids" [12] features games, crafts and recipes, science, videos, and animal information. It is perfect for 5 years old and under kids. Math and reading video-like games like math baseball and Mad Libs Junior are available at the "Funbrain" web-site [13]. Kids/students can get into cooking on the "Spatulatta" friendly cooking web-site with lots of videos and recipes [14]. The "NGA Kids" web-site [15] gives users art adventures and activities from the United Sates National Gallery of Art. The "Yahoo Kids" web-site is the least educational of all the choices listed above, since the site includes both games and videos of all sorts. But, kids/students can find lots to learn on this interactive web-site like homework help, learning about science, and access to an encyclopedia.

Many more educational sites are available on the Internet; there are even sites with sorted lists of educational sites, such as those lists presented in [16]. However, all these sites offer exercises or sets of exercises, which can (and should) be solved by kids/students, but the teacher cannot change (customize) them to create her/his own exercises based on the existing ones, and then offer them to the kids/students. The web-based system presented here allows this.

The Learning Managements Systems (LMS), such as Moodle, Blackboard, Sakai, Desire2Learn, etc., possess their own tools for the creation and publication of exercises, and obviously should be included in the list of web-based systems and tools for helping building exercises. Also, there are other tools available on the Internet:

- "Hot Potatoes" [17] — This program is free to educators, although registration is required. There are six types of quizzes; JBC: multiple-choice exercises, JCloze: gap-fill exercises, JMatch: matching/ordering exercises, JQuiz:short-answer questions, JMix: jumbled sentences. After the creation of the quizzes, Hot Potatoes will convert the files to HTML, which the author must then (manually) upload to the server.
- "Quia! Create your own learning activities!" [18] — There are eight types of games and two types of quizzes to create at the Quia! site. The user must be registered before using Quia!. At the Quia! web-site the user enters the information and keeps track of the resulting URL. The user can create a quiz from any computer with Internet access, produce a page, and then access it again if the user remembers the URL. All pages stay on the Quia! server. The user can choose a discipline on the main page to see what others have done with Quia!
- "QuizStar" [19] — The Quiz Star is an application that allows bringing together web pages for a series of lessons. The QuizStar quizzes remain on the TrackStar server. Registration is needed.
- "Charity Kahn's JavaScript Quiz Creator" [20] — This tool creates the code for a simple multiple choice quiz. The user will need to copy the code and upload it to the server.
- "interactivetest.com" [21] — Here we can make online quizzes (up to 100 total questions) for free. For more than 100 and there is a charge. The quizzes stay on the interactivetest.com server.
- "Puzzlemaker" [22] — Is a puzzle generation tool for teachers, students and parents. The user can create and print customized word search, criss-cross, math puzzles, and more-using his/her own word lists. The user must choose a type of puzzle from the list of ten templates available and make her/his own puzzle online for classroom or home use.
- "Rubric Generator" [23] — Rubrics are intended to make grading quicker, clearer, and more objective. The site contains tools that guide the user through the process of creating the assessment tools for evaluating student performance. There are over five hundred printable rubrics on the web site.
- "Quiz your Friends" [24] — A very simple site that will guide the user through the creation of a quiz in four steps. In the final step the user chooses if the quiz will be sent via Email, Instagram, Facebook or Twitter, to the group of peoples/friends.
- A list of many more quiz builders can be found in [25].

The big differences, and main characteristics, of the system presented here and that makes it unique when compared to the systems/sites listed above are the following:

- the production/creation of exercises is based on existing templates;
- exercises can only be created based on the existing templates;
- anyone, whose access was granted by the administrator of the system, may create exercises;
- the publication of every exercise just created is done immediately and automatically after its creation, and it will be available for everyone registered in the system;
- the creation/production of a new exercise is based in three simple steps:
  1) choose a template from the list presented in the form of images containing all the exercises;
  2) change the customizable fields in the template;
  3) choose a name and save the new exercise in the exercises database;
- there is no need for any programming skills or experience.

 ISBN: 978-1-61208-428-2

As can be seen from the list above, the template dictates the “behavior” of the exercise. The exercise in “it-self” will run on the computer of the user, i.e., the exercise will not run in the “server-side” but in the “client-side”. As will be discussed in Section V (“Overview of the System”), the system was implemented using HTML 5, JavaScript, Adobe Flash, and ActionScript. The system will be used to create the exercises and store them in its internal databases. After that, the system will work as a typical web-server.

In fact, in seeking to meet the latest requirements and concepts used by leading standards in distance learning, we are making efforts to make it possible to download the exercises in a standard (SCORM and Common Cartridge) in order to make possible the execution of exercises off-line and subsequent communication of the results achieved by students at the time of reconnection to the Internet. However, at the moment, this is partially implemented.

III. Advantages of using the web-based system

In the view of several authors it seems that we are moving to a “digital learning” (e.g., [26]–[28]). According to Weller [26], there are five factors that favor de use of the Internet in education: its social acceptance, it facilitates a sense of control and therefore ownership for educators far more than previous technologies, the web browser has become a generic interface, it is both an interactive (do something with the information rather than just being a passive recipient) and personal (all learners are not the same) medium, is a sustaining (e.g., a retailer that uses it to supplement its physical shops, for instance by home delivery from supermarkets) and disruptive (e.g., it does not intend web based shopping to displace its standard retail outlets, but rather it hopes the two will complement each other, and thus altering the organization in which it is implemented) technology.

Internet can be used to supplement face-to-face teaching; for example, many campus based universities use the Internet not to replace their traditional face-to-face mode of delivery, but rather as a means of supplementing it (e.g., web pages can be used to provide additional information, or e-mail be used as a means of contacting tutors of large courses). Of course that there are topics that better suit its use (an “Introduction to Information Technology” course, and a course for teachers wanting to learn about computer conferencing, are examples of courses were using the technology effectively forms an integral part of the academic content of the course) without forget pedagogic suitability (for example, the resource based learning, where students can be presented with a wide range of resources, often external to the university, were students derive their own learning experience from these resources within the overall framework of the course). Internet’s usage also leads to institutional benefit because institutions can benefit from the use of Internet in education, both on campus and in reaching new audiences, and from the fact that it can be seen as a Computer Assisted Learning delivery mechanism—many courses are using Internet to deliver simulations, visualization aids and interactive tools.

In order to become more accurate in their work, both in reading problems and in working out solutions, primary students need more practice to reinforce what they are learning, a process which traditionally takes place on paper, and traditionally are distributed in textbooks or tutoring books. Unfortunately, these materials are usually designed for average learners, and it is often difficult to find the best-fitting content for students with differing abilities and skills. For example, in Portugal all the classes have only one textbook designed for all students, but literate students may need a higher-level tutoring, and illiterate may need a lower-level tutoring. Therefore, the depth and flexibility of ability gained from these textbooks are restricted. Some programs based on e-learning technology may provide personalized contents for learners by collecting the learning process. But primary students may become restless and unfocused when staying in front of computers during long periods of time. As putted by Warschauer, [27], “New technologies do not replace the need for strong human mentorship, but, indeed, amplify the role of such mentorship”. Obviously students must become into contact with the new teaching/studding tools progressively, in order for they to become a part of the learning environment as smoothly as possible.

Teachers, but also parents, have an important word to say in this process; as noted by Warschauer [27], “the teacher must be centrally involved, actively instructing and mentoring students, especially at the initial stages of work on a project. Unfocused instruction can leave students rudderless, and this is particularly harmful to at-risk students, such as those with learning disabilities, limited literacy, and language skills, or insufficient background knowledge”.

In general, after students have completed practice exercises or homework, very often they have to wait for their teacher to check it and provide feedback. These interruptions can reduce students’ interest in learning as well as learning efficiency.

We believe that the prompt feedback about the exercises correctness, together with the training provided by the different exercises sets about the same subject, besides the exploitation of video, color, sound, etc., positively reinforce the diverse child’s senses, definitely contributing to capture and motivate the child.

From the teacher point of view, one advantage of the system here proposed is time saving. Most of the times, the teacher writes in the exercise-book of each pupil (at least in the “blackboard”) the set of homework exercises that s/he should try to solve. This is a handy work, and consequently very time consuming. When students return with their homework, the teacher must read and correct, one-by-one, every exercise-book, in order to have and give feedback to the pupils. During this process, the teacher also analyzes each student’s work, in order to evaluate the student’s progress and potentially introduce some programme and/or pedagogical adjustments. Obviously, this analysis should also be made to the global class.

With the web-based system presented here the teacher can select different sets of exercises for each student or, alternatively, propose the same set of exercises to the whole class, and this simply by creating e-mail lists and maintaining these lists as needed. The system will then provide the teacher with statistical data concerning each student and the entire class. We believe that these data will greatly help the teacher in evaluating his/her work, and help with the mentioned adjustments. The selection of different exercises sets may not directly lead to a time reduction of the teacher’s work, but its automatic check and the statistical data do. We believe that

 ISBN: 978-1-61208-428-2

teachers can use this extra time to better understand where and why some children are having troubles.

IV. Pedagogical model

The teaching and researching communities agree that it is through the use of exercises and games that learning becomes effective, regardless of the model or philosophy used. For example, proponents of sociocultural theory claim that learning is primarily a social process mediated through interactions using tools [29]–[31]. According to Vygotsky, mediation occurs through the use of 'semiotic' and 'material' tools. The semiotic tools include symbols, signs, and spoken languages. Material tools include such items as pens, spoons, and particularly networked (Internet) computers. These tools do not simply facilitate the set of activity that might take place, but they fundamentally shape and define the type of activities that might be developed [30]. Seymour Papert proposed the use of tools, particularly the computer "a mighty education tool", in helping in the process of construction of knowledge (given rise to the "constructionist" theory), adapting the beginnings of the cognitive constructivism of Jean Piaget in order to a better use of technology, [32].

Traditionally, the dominant form of learning in schools has been teacher-directed learning or *guided learning*: "a trainer or teacher takes all the relevant decisions and the learner can and should follow him or her. He decides about the goals of learning, the learning strategies, the way to measure outcomes and he takes care of feedback, judgments, and rewards" [33]. Besides guided learning, Simons, Linden and Duffy, [33], distinguish two other ways of learning: *experiential* and *action learning*. To facilitate and support learners through the gradual and progressive acquisition of adaptive competence, novel classroom practices and cultures are needed. These practices and cultures should create the conditions for a substantial shift from guided learning toward experiential as well as action learning, resulting in a balanced and integrated use of the three ways of learning: constructive, self-regulated, and contextual or situated [34].

The constructivist view of learning has become common ground among educational psychologists (see, for example, [33], [35]). Constructivism implies that constructive learning is self-regulated. According to Zimmerman [36], self-regulation "refers to the degree that individuals are metacognitively, motivationally, and behaviorally active participants in their own learning process". According to De Corte [34], "constructive and self-regulated learning processes should preferably happen and be studied in context", and because learning is collaborative, the learning efforts are distributed over the individual student, its partners in the learning environment, and the (technological) resources and tools that are available [37].

As stated by De Corte [38], "starting as much as possible from tasks and problems that are meaningful and challenging for students, learning environments should initiate socially supported constructive learning processes that enhance students' cognitive and volitional self-regulatory skills". It is expected that students will be able to use their acquired knowledge and skills to solve situations and problems in everyday life, something that Bransford and Schwartz [39] call *preparation for future learning*.

This is the theoretical model supporting the web-based system presented here and this research.

V. Overview of the system

Before start describing the system a word about its developing context is in order.

In 2006, the "ESchool" Programme (e.escola— [40]) enabled the purchase of laptops by students, form the 5th to 10th grade, and teachers for a price of 150 euro. Later, the Portuguese government has developed a set of protocols with Intel, the leading telecommunications operators (Optimus, TMN, Vodafone, and Zon), Microsoft, "Magic Box" (Caixa Mágica), and the local members, allowing children of the primary schools access to portable computers "Magellan" (Magalhães), for a maximum cost of 50 euro [41]. Additionally, the classrooms have been equipped with interactive "blackboards".

The existence of computers in the schools is of course necessary to integrate IT effectively in the learning process, but it is by no means sufficient. Setting-up infrastructures is a relatively simple process in comparison with their actual usage, since both the necessary training and the change of habits that lead to routine usage are the result of slow and gradual processes; the need to train, motivate and involve the teachers, pupils and the community in general towards a correct use of IT tools entails a number of technical and cultural difficulties [42], [43]. Additionally, several researchers claim that it is what teachers think and believe and students' attitudes that ultimately shapes the activities in the classroom [44]–[46].

As usually happens, at least in Portugal, the lack of tools supporting the authoring and automatic checking of exercises for specific topics in primary education (e.g., check true or false sentences) drastically reduces the advantages in the use of e-learning/web-based environments on a larger scale. Some technical details of the web-based system presented here were presented elsewhere [47].

The web-based system presented here aims to contribute to the resolution of problems in two distinct classes: firstly, for the spreading, construction and sharing of educational resources, particularly those closely related to the teaching of Portuguese as mother tongue; secondly, to develop applications more "user friendly" in terms of the average user, trying to eliminate the necessary notions of programming that are currently required by the existing applications.

It is 100% visual (graphical) for the preparation of the exercises and games, which in turn are developed based on existing templates.

Obviously the need for using standards when designing and implementing tools to help in the creation of learning objects became evident. Only then the exercises and games (i.e., learning objects) will be truly available for reuse and sharing, no matter the platform the students/learners are using. In our view, SCORM (Sharable Content Object Reference Model) is the most complete standard to be used in the case presented here (Primary Students' Homework), because in its specification it includes standards developed by various third parties, integrating various standard specifications that simultaneously enable better interoperability between systems and sharing of content, and it is widely accepted that it is the best standard to use with Computer Based Training [48].

The tool has two distinct modules: one related to the development and personalization of the templates; and another dedicated to the development and personalization of

 ISBN: 978-1-61208-428-2

the exercises and games. The main blocks of the application include: user's management (registration and authentication); templates' management (creation, modification, deletion, etc.); exercise's management (creation, modification, etc.); databases' management; exercises' and templates' downloading; configuring/customization; help; and exiting application. The databases block is able to manage distinct databases, including the templates' and exercises' databases, as well as all the information related to the registered users. Also, a special database containing all the images that can be used to produce templates and exercises is implemented; obviously, the user may add new images to the database. So, in its current version, the tool implements the following data tables: users, templates, exercises, images, classes (used to classify images).

The templates management block of the application is not fully operational in its current version. In its current version, only the exercises module is fully implemented. The templates module is operational but it cannot be used to create or change an existing template (i.e., it shows the existing templates to the user and let he/she choose one from the presented list). We have created only five templates, using other tools, for demonstration proposes. When a user/teacher wants to create a new exercise he/she must follow three main steps. First, choose a template or an existing exercise, from a list presented in the form of images containing all the exercises. Second, change the customizable fields in the template. Third, choose a name and save the new exercise in the exercises database; this will generate a XML file with information corresponding to the configuration options of the exercise (exercise wording, images, correct options/answers, etc.) that will be saved along with the rest of the exercise. The number of interactions during the second step will depend on the template. Typically, it will be necessary to customize the exercise wording (i.e., the text explaining what is expected for the student to do), the images that will be used/presented to the student, as well as the correct answers/options.

In the example presented in Figure 1, besides the fields listed above, it is also possible to change the text of the "Verify" and "Try again" buttons and the text displayed when the answer is correct and when the answer is wrong. Also, it is possible to choose the degree of help/difficulty the student will have during the execution of the exercise ("easy", "medium" or "hard"). When in the "ease" mode every time a student chooses an option or enters his/her answer, he/she will have immediate feedback about its correctness. When in the "medium" mode, only when he/she presses the "verify" button will get the feedback about the correct and wrong answers. In the "hard" mode only when he/she presses the "verify" button will get the feedback about the exercise correctness, but the student will have no feedback about the correct and incorrect answers (only that there are errors, or not). All these interactions are based on choosing elements form dropdown lists, check boxes, drag and drop objects/images, among other. In the example presented in Figure 2, the student is asked to fill in the missing words ("the gaps") in order to form correct sentences; this Figure shows the exercise seen by the student.

The system was implemented using HTML 5, JavaScript, Adobe Flash, and ActionScript.

We want that in the near future this system may be used in English-speaking countries, since the only thing that needs to be changed is its interface. From the exercises point of view, they do not need adjustments, because it all depends on the templates (the templates should be designed in English, as well as everything they contain, such as video, images, text, speech, music, etc.).

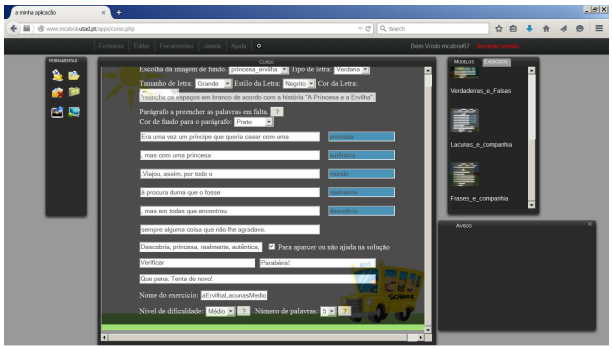


Figure 1. Creating a sample exercise a sample exercise—fill in the missing words ("the gaps") in order to form correct sentences.

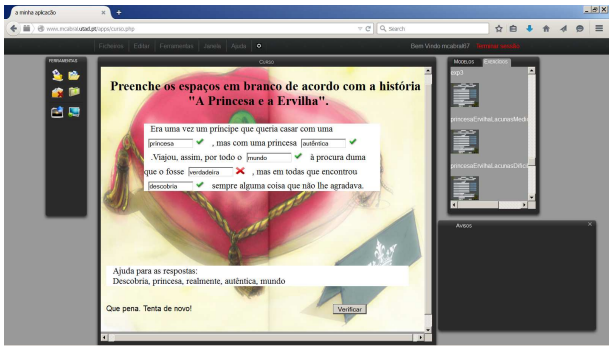


Figure 2. Solving a sample exercise a sample exercise—fill in the missing words ("the gaps") in order to form correct sentences.

VI. SOME RESULTS

Currently, there are five templates that can be used in the system to create new exercises: three relate to the teaching/learning of Portuguese mother tong, and the other two relate to the teaching of other concepts (notions of in and out, animals and fruits, count, etc.). Here we will present and discuss the results of first three templates.

The results presented in this section report to a set of exercises used to demonstrate the usefulness of the proposed web-based system. We follow the methodology proposed by Yin [49], where the researcher can conduct the research in its working context.

Bearing in mind the teaching and learning of Portuguese mother tongue, and the fact that the exercises should serve as homework for primary students, we have developed three specific templates: "fill in the gapes", "true or false", and "multiple choice". In the "fill in the gapes" template, the pupils are asked to fill the blank spaces in a sentence, or set of sentences, in order to form correct sentences. This template includes the possibility to present to the student the missing words, but not necessarily obeying the order in which they should be introduced in the sentence(s), i.e., the words are randomly presented, functioning as extra help. Also, the teacher can choose between one and five "missing words"; in

 ISBN: 978-1-61208-428-2

fact, the “gap” can be a complete sentence. In the “true or false” template, the students are asked to classify a sentence as being true or false. In the “multiple choice” template, the students are asked to select the best possible answer out of the choices from a list.

As presented above in Section V, the degree of help/difficulty the student will have during the execution of any of the exercises may be set to three different levels: “easy”, “medium” or “hard”. For our set, we have choose easy and medium levels.

We have also developed the corresponding counterpart paper format of the set of digital exercises. The exercises were presented to 4th-year (4th-grade) primary students during the months of January, February and March 2015, in the context of the project “Little Box of surprises” (*Caixinha das surpresas*). Basically there are sets of “little boxes”, each with a set of ten to twelve copies of the same story, and once/twice a month a story is chosen (by the teacher in conjunction with the librarian), and the students go to the library and read the story. At the end they are asked to respond to a set of questions, traditionally in paper form, of the types listed above (filling gapes, true/false, writing sentences/ordering words, and multiple choice). The sample exercise in Figures 1 and 2 was produced based on the fairy tale “The Princess and the pea” by Hans Christian Anderson. The full set is available in [1]. A “scanned” form of the corresponding counter-part paper format offered to the students is also available [50].

All the students have answered all the questions both in paper format and directly in the web-based system using the computers available at the library. Table I summarizes the results achieved by the 23 4th-grade students. The students have answered 12 questions of the type “True & False” (T&F), one question with 5 “gapes” in the template “fill in the gapes”, and 6 questions of the type “multiple choice”. This table presents the number of correct answers. During the application process of the web-based exercises we gave the students the possibility to answer correctly all the questions and registered the number of tries. However, due to space limitations and because in the paper form students have only one chance to answer (correctly or not) all the questions, here we present the results of the first answers (i.e., here, we consider that an answer was correct only if it was correctly answered during the first attempt). As can be seen, for the “T&F” exercises the average value is the same both for paper and web-based types. For the “gapes” and “choice” types the results were slightly better for the exercises in paper format. Note that the mode is the same for the “gapes” exercises, is better for the “choice” exercises in paper format, and is better for the “gapes” when using the web-based system.

We also presented the students a very simple survey with the following questions:

1) The web-based system was easy to use.
2) It was easy to understand.
3) I would have liked more in class use.
4) It was helpful because I had prompt feedback.
5) I would have preferred more homework questions using it (in place of other assignments).
6) I would have liked more contents of this type on it.
7) It was easier to answer the questions in it.
8) I felt better responding using the computer.

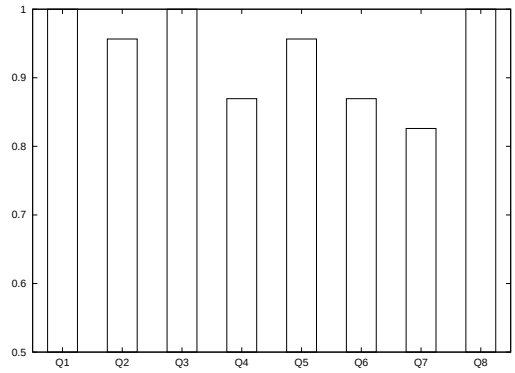


Figure 3. Survey results: average answer to each question.

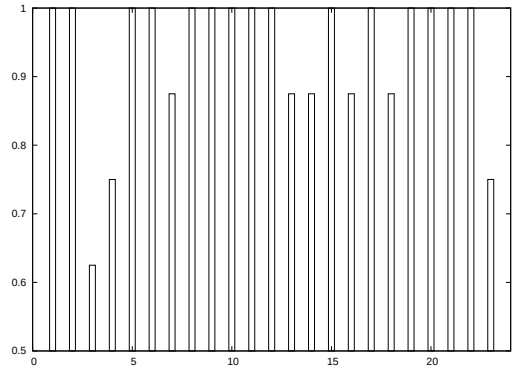


Figure 4. Survey results: average answers from each student.

Figure 3 shows the mean responses to each question of this survey. The Portuguese version of the survey (actually applied) can be downloaded [50]. As can be seen from the plots in this Figure, questions #1, 3 and 8 drew the highest responses, indicating that students found the system easy to use and feel comfortable using it. However, question #7 had the lowest response, showing that they don’t feel that it was easier to answer the questions using the computer, or at least they don’t feel that the questions were not of lower difficulty degree. This seems to be reinforced by the responses to questions #4 and 6. Average answers for each survey participant, presented in Figure 4, show that some responses seemed to be either too low or too high, i.e., showing dissatisfaction/satisfaction with the class/subject itself and not specifically with the web-based system.

Students’ comments were overall positive. The less positive comments point to a need to have more homework examples, and more time in order to gain experience in using the system and computer. Next are listed some of the comments. “I think we should use more often the computer”. “I really enjoyed doing this on the computer where everything was cool and made me learn a lot.” “I think it was easy to use the computer because almost hit all the questions at first!”. “I learned to write on the computer.” “I really enjoyed! It was great fun!”. “The computer is cool!”. “He liked to use the computer more often”. “It was fun and also spectacular”. “I wish I had more time to do so”. “I liked using the computer because it was fun to learn on the computer”.

 ISBN: 978-1-61208-428-2

TABLE I. EXERCISES' RESULTS SUMMARY FOR THE 23 STUDENTS (T&F—TRUE & FALSE; GAPES—FILL IN THE GAPES; CHOICE—MULTIPLE CHOICE; AVG—AVERAGE; STD—STANDARD DEVIATION; MOD—MODE).

		Student																						Avg	Std	Mod
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23			
Paper	T&F	10	8	10	10	8	12	11	11	11	8	9	10	10	11	10	10	10	9	7	10	11	11	11	9.9	1.2	10
	Gapes	5	5	3	5	0	5	5	5	4	5	5	5	5	5	5	5	5	5	4	5	5	5	5	4.6	1.1	5
	Choice	6	6	6	6	5	6	6	6	6	6	6	6	6	6	5	6	6	3	6	6	5	5	6	5.7	0.7	6
Internet	T&F	11	10	8	9	9	9	11	8	8	8	11	11	10	10	11	12	10	10	11	9	11	11	9	9.9	1.2	11
	Gapes	4	5	5	4	3	5	4	4	3	5	3	4	5	5	5	5	5	5	3	5	5	5	5	4.4	0.8	5
	Choice	4	4	4	3	4	5	4	5	4	5	5	4	4	5	5	6	4	2	3	5	6	6	6	4.5	1.0	4

In conclusion, we could see that there are no big differences in the results achieved by the students. However, during the answers to the exercises students showed a more positive attitude, happiness and greater openness to the use of the web-based system.

VII. CONCLUSION AND FUTURE WORK

We have presented a web-based system [1] to support students' homework. In order to prove its usefulness, we have built a set of exercises to help improving Portuguese (mother tongue) language skills of 4th-grade primary pupils. This set of digital exercises proved to be the child's favorite, rather than their counterpart exercises in paper form, also showing a fair more positive attitude from the students' point of view. A simple survey also showed that students do prefer the use of the system here presented, when compared to other more traditional ways of practicing.

We believe that the prompt feedback about the exercises correctness, together with the training provided by the different exercises sets about the same subject, besides the exploitation of video, color, sound, etc., positively reinforce the diverse child's senses, definitely contributing to capture and motivate the child.

Students with disabilities are known for possessing a set of unique characteristics that hinder their integration in school and consequently their learning. We believe that students with disabilities can be the first benefitting from using a system like the one proposed here.

As stated by Warschauer, [27], "New technologies do not replace the need for strong human mentorship, but, indeed, amplify the role of such mentorship". Obviously that the students must become into contact with the new teaching/studding tools progressively, in order for they to become a part of the learning environment as smoothly as possible.

From the examples above, we could easily concluded that the presented templates may be used to teach/learn other subjects. For example, based on the "true or false" template a teacher can build a set true or false questions for the teaching of chemistry. Obviously, the "multiple choice" template may be used to build exercises for the teaching of any subject.

REFERENCES

[1] http://www.mcabral.utad.pt/apps/. Accessed: 10-08-2015.

[2] https://www.coursera.org/. Accessed: 10-08-2015.

[3] http://learni.st/. Accessed: 10-08-2015.

[4] http://www.openculture.com/. Accessed: 10-08-2015.

[5] https://www.udacity.com/. Accessed: 11-08-2015.

[6] http://ed.ted.com/. Accessed: 11-08-2015.

[7] http://www.memrise.com/. Accessed: 11-08-2015.

[8] http://www.mentormob.com/. Accessed: 11-08-2015.

[9] http://wonderopolis.org/. Accessed: 12-08-2015.

[10] http://www.whyville.net/smmk/nice. Accessed: 12-08-2015.

[11] https://www.pottermore.com/. Accessed: 12-08-2015.

[12] http://kids.nationalgeographic.com/. Accessed: 12-08-2015.

[13] http://www.funbrain.com/. Accessed: 12-08-2015.

[14] http://www.spatulatta.com/. Accessed: 12-08-2015.

[15] http://www.nga.gov/content/ngaweb/education/kids.html. Accessed: 12-08-2015.

[16] http://www.parenting.com/. Accessed: 10-08-2015.

[17] http://web.uvic.ca/hrd/halfbaked/. Accessed: 13-08-2015.

[18] http://www.quia.com. Accessed: 13-08-2015.

[19] http://quizstar.4teachers.org/. Accessed: 13-08-2015.

[20] http://builder.i.com.com/. Accessed: 13-08-2015.

[21] http://www.interactivetest.com. Accessed: 13-08-2015.

[22] http://www.discoveryeducation.com/free-puzzlemaker/. Accessed: 13-08-2015.

[23] http://www.teach-nology.com/web_tools/rubrics/. Accessed: 13-08-2015.

[24] http://www02.quizyourfriends.com/. Accessed: 13-08-2015.

[25] http://eleaston.com/quizzes.html. Accessed: 13-08-2015.

[26] M. Weller, Delivering Learning on the Net: the why, what & how of online education. London and New York: RoutledgeFalmer, Taylor & Francis Group, 2002.

[27] M. Warschauer, "The paradoxical future of digital learning," Learning Inquiry, vol. 1, 2007, pp. 41–49.

[28] K. Moeller, U. Fischer, H.-C. Nuerk, and U. Cress, "Computers in mathematics education - Training the mental number line," Computers in Human Behavior, vol. 48, JUL 2015, pp. 597–607.

[29] L. S. Vygotsky, Mind in society: Development of Higher Psychological Processes, M. Cole, V. John-Steiner, S. Scribner, and E. Souberman, Eds. Harvard University Press, 1978.

[30] J. V. Wertsch, "The voice of rationality in a sociocultural approach to mind," in Vygostky and education: Instructional implications and applications of sociohistorical psychology, L. C. Moll, Ed. New York, NY: Cambridge University Press, 1992, pp. 111–126.

[31] C. A. Bisol, C. B. Valentini, and K. C. Rech Braun, "Teacher education for inclusion: Can a virtual learning object help?" Computers & Education, vol. 85, JUL 2015, pp. 203–210.

[32] S. M. Papert, Mindstorms: Children, Computers, and Powerful Ideas. New York: Basic Books, 1980.

[33] R. J. Simons, J. van der Linden, and T. Duffy, "New learning: Three ways to learn in a new balance," in New learning, R. J. Simons, J. van der Linden, and T. Duffy, Eds. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2000, pp. 1–20.

[34] E. D. Corte, "Learning from instruction: the case of mathematics," Learning Inquiry, vol. 1, 2007, pp. 19–30.

[35] H. Passos, "Flexible contents: Communities of practice and distributed cognitivism impact in instructional design," in 3rd International Conference of Education, Research and Innovation (ICERI2010), Chova, LG and Belenguer, DM and Torres, IC, Ed. Lauri Volpi 6, Valenica, Burjassot 46100, Spain: IATED-INT Assoc Technology Education A& Development, 2010, Proceedings Paper, pp. 329–333, 3rd International

 ISBN: 978-1-61208-428-2

Conference of Education, Research and Innovation (ICERI), Madrid, Spain, Nov 15-17, 2010.

[36] B. J. Zimmerman, “Dimensions of academic self-regulation: A conceptual framework for education,” in Self-regulation of learning and performance: Issues and educational applications, D. H. Schunk and B. J. Zimmerman, Eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 1994, pp. 3–21.

[37] G. Salomon, Ed., Distributed cognition. Psychological and educational considerations. Cambridge, UK: Cambridge University Press, 1993.

[38] E. D. Corte, “Mainstreams and perspectives in research on (mathematics) learning from instruction,” Applied Psychology: An International Review, vol. 53, 2004, pp. 279–310.

[39] J. D. Bransford and D. L. Schwartz, “Rethinking transfer: A simple proposal with multiple implications,” in Review of research in education, A. Iran-Nejad and P. Pearson, Eds. Washington, DC: American Educational Research Association, 1999, vol. 24, pp. 61–100.

[40] http://www.eescola.net/indexA.aspx. Accessed: 10-08-2015.

[41] http://eescola.pt/e-escolinha/Oferta.aspx. Accessed: 10-08-2015.

[42] M. J. C. S. Reis, G. M. M. C. Santos, and P. J. S. G. Ferreira, “Promoting the educative use of the internet in the Portuguese primary schools: a case study,” Aslib Proceedings, vol. 60, no. 2, 2008, pp. 111–129.

[43] R. Shillair, S. R. Cotten, H.-Y. S. Tsai, S. Alhabash, R. LaRose, and N. J. Rifon, “Online safety begins with you and me: Convincing Internet users to protect, themselves,” Computers in Human Behavior, vol. 48, JUL 2015, pp. 199–207.

[44] S. A. Yoon, J. Koehler-Yom, E. Anderson, J. Lin, and E. Klopfer, “Using an adaptive expertise lens to understand the quality of teachers’ classroom implementation of computer-supported complex systems curricula in high school science,” Research in Science & Technological Education, vol. 33, no. 2, May 4 2015, pp. 237–251.

[45] S. Kerckaert, R. Vanderlinde, and J. van Braak, “The role of ICT in early childhood education: Scale development and research on ICT use and influencing factors,” European Early Childhood Education Research Journal, vol. 23, no. 2, Mar 15 2015, pp. 183–199.

[46] A. Taneja, V. Fiore, and B. Fischer, “Cyber-slacking in the classroom: Potential for digital distraction in the new age,” Computers & Education, vol. 82, Mar 2015, pp. 141–151.

[47] S. Neto, A. Capela, R. Morais, E. Peres, and M. Reis, “A HTML 5-based tool to help teachers in the creation of exercises and games for the teaching of mathematical concepts in the primary schools,” in 2nd International Conference on Virtual and Networked Organizations - Emergent Technologies and Tools (ViNOrg 13), Póvoa de Varzim, Portugal, 2013, pp. 1–6.

[48] V. Gonzalez-Barbone and L. Anido-Rifon, “From SCORM to Common Cartridge: A step forward,” Computers & Education, vol. 54, no. 1, 2010, pp. 88–102.

[49] R. K. Yin, Case study research: Design and methods, 2nd ed. Newbury Park, CA: Sage, 1984.

[50] http://www.mcabral.utad.pt/HansChristianAnderson/SapatinhosVermelhos. Accessed: 10-08-2015.

 ISBN: 978-1-61208-428-2

Embodying Latent Requirements with Unexperienced Attractions through Selection of Travel Point Photographs

Momoko Kato

Graduate School of Information Science and Engineering
Ritsumeikan University
Shiga, Japan
Email: momochii@de.is.ritsumei.ac.jp

Kenjo Yasui, Hiromitsu Shimakawa

Ritsumeikan University
Shiga, Japan
Email: {kenjo, simakawa}@de.is.ritsumei.ac.jp

***Abstract*—People who want to travel search for travel points satisfying their travel requirements. They search for travel points such that they feel attractions to fulfill the requirements. They want to search for them easily and quickly. In recent years, there are many people who use the Internet to collect information on travel points. They use the Internet because they can collect the information on the travel point anytime and anywhere. On the other hand, there are various kinds of information on a lot of travel points on the Internet. Users have to embody conditions to specify the travel point they want to visit. True pleasures of travels are to experience what they have never experienced. Requirements for travel points to seem to be attractive are quite unclear. For the recommendation of travel points, people want a tool to embody easily and quickly the conditions specifying the travel point they want to visit. Current search engines for travel points on the Internet utilize the user based collaborative filtering from travel histories of travelers. However, they cannot grasp what latent requirements travelers have and what kinds of attraction they expect, because they analyze travel histories. The paper proposes a recommendation method which enables users to embody each potential requirement for their travel beforehand. The method utilizes photographs of travel points to infer what latent travel requirement users have and what kinds of attractions they expect to fulfill the latent requirements. An experiment has proved that almost two thirds of subjects are satisfied with travel points recommended using this method. Furthermore, the method allows all of the subjects to decide on travel points they want to visit in a shorter time than the method they usually use.**

Keywords–Travel; Trip; searching keyword; recommendation; preference;

I. INTRODUCTION

Some people go on a trip to relax their body and mind worn out in real life, while others aim to change their moods. They decide on vacation places based on events which incline them to trips. After they decide on vacation places, they would search for details of what they can experience there. For example, they would check foods specific to the places, famous sightseeing areas, accommodation facilities, and so on. The information collection continues until they are satisfied as long as their time allows. In these years, many people gather the vacation information with Internet, because they can search for the travel information anytime and anywhere with it. On the other hand, there is too much travel information on the Internet. A user must select appropriate travel information to decide on a vacation place. It is required for all users to find out the travel information useful to decide on vacation places, in a short time and with little effort. Specifying keywords relating to vacation places to search engines on the Internet, people planning trips browse the vacation information to decide on a vacation place with Internet. The information collection does not need much effort, if users of search engines have embodied keywords which enable them to narrow candidate vacation places meeting with their requirements. However, the information collection is not easy in many cases, because users are likely to have only unclear requirements in the beginning of planning of their trips. To make the best use of the tools to collect travel information, we should support them to embody keywords representing their vacation requirements in their minds.

The pleasure of planning trips is the process of chooseing vacation places, imaging themselves enjoying the places they search with the Internet. Some users find new vacation goals or their latent preference for trips, when they attain unknown vacation information which excites them in the planning. The new goals or latent preference for trips make it more difficult to embody keywords. Users with little knowledge on vacation cannot precisely embody keywords to represent vacation places meeting their goals and preference. They have little chance to find suitable vacation places matching their images. On the other hand, users with a lot of vacation experiences have often fixed their preference. They are likely to specify same keywords in all cases. They are disappointed with stereotyped trip images brought by search results. They fail to expand their expectation for a trip. Several travel recommendation systems have been developed with the user-based collaborative filtering from the travel histories [8] [9] [10]. But, it is the real pleasure of trips to visit places which provide entertainments users have never experienced. Users often have new goals and latent preference for trips. For new goals, the users have no knowledge to achieve them. For latent preference, the users often find them in places they visit by chance. Any existing system using travel histories cannot recommend vacation places matching with new goals and latent preference, becuase they are not attained from places they have visited in their past trips. It is likely to recommend users vacation places similar to what the users have been to. Users cannot be excited when the system recommends those vacation places.

This paper proposes a method to embody user requirements for a trip, which are ambiguous in the beginning of trip

 ISBN: 978-1-61208-428-2

planning. The method estimates goals and preference of user for trips, using photographs showing what users can experience in specific vacation places. In the process of searching for vacation places using photographs, it is easy for users to image themselves there. Suppose a user is pleased with a photograph of a specific vacation place, after imaging herself there. It is assumed that the photograph implies her goals and preference for trips, including not only apparent ones but also the latent ones. The proposed method has collected various goals and preference from many people, beforehand. Each photograph matches with some of them. The method associates terms representing them with the photograph in advance. It regards associated terms as keywords to embody her requirements for trips. The method lets the user specify the keywords in the search bar of existing Internet search engines for vacation places. It presents vacation places recommended in high ranking, which enable the user to decide on the vacation places matching with her goal and preference of trip.

The remaining of the paper is organized as follows. Section 2 describes the present state and problems of the searching for vacation places with the Internet. The method and the usability of the photograph are presented in Section 3. Section 4 illustrates experiments to verify the usefulness of the method. The paper discusses the experiment result in Section 5. Section 6 gives the conclusions and future works.

II. How users search trip places

A. Requirements in search of trip places

This paper defines a trip goal as what a user wants to experience in a vacation place. For example, a user wants to relax her body and mind with a hot-spring or refresh her in nature. In addition to that, the paper defines a trip preference as what the user wants to experience in the trip to achieve her trip goal. For example, she prefers outdoor hot springs to others, or wants to view a deep-blue sea from a height to refresh in nature. Before the trip, the user would enjoy imaging the way she would experience when she visits a specific vacation place to achieve her trip goal with her trip preference. The paper refers to a trip image as what she envisions for experience she would take in the trip. Suppose a user desires to relax her mind and body, or to refresh her viewing deep-blue sea from a height. She would search for a vacation place, imaging her experience of them. Before a trip, a user would mostly plans for the trip. She plans a trip according to conditions in the trip such as trip goals, partners, their preference, and so on. Imaging concrete situations in the trip under the conditions, the user chooses a specific vacation place. Let us consider a trip with intimate friends for a long time. The trip planner comprehends well common preference between her and the intimate friends. She would make a trip plan, choosing vacation places which make both of her and the friends enjoy. When she thinks those vacation places, she images herself and her friends enjoying scenes and their conversation at those vacation places. When she thinks that one of those vacation places is suitable in this trip, she decides to go on a trip in that place. In this way, the user enjoys the process of deciding a vacation place imaging situations in the trip, as well as the trip itself. It is due to the increase of her expectation for their trip by the imaging. She gets excited, while she is making concrete trip images along with the trip goals and trip preference. Sometimes, she finds new trip goals or trip preferences which excite her, because she attains unknown vacation information. It enhances options in the trip, which makes the trip plan more pleasant for her. On the other hand, if she fails to make her trip goals and trip preference clear, she can never proceed to the step to decide on an exciting vacation place, because she imagines no concrete situation. Travel agencies would recommend vacation places, interviewing travelers. In the interviews, they would make unclear trip goals and trip preference to clear ones to excite the travelers, in order to find good vacation places. One of roles of travel agencies is to provide unknown vacation information for customers, increasing their expectation for their trip.

B. Searching for vacation places with Internet

We have conducted a questionnaire to survey user requirement for time used for the search of vacation places, in advance. In the questionnaire, users are asked whether they prefer search in a short time if they can find suitable vacation places. About 92% people respond “yes” in this question. It reveals people want methods with which they can find information to decide on vacation places in a short time and without efforts. As one of those methods, a lot of people use the Internet. According to the investigation in 2012 [1], 45.2% people gather the vacation information using Internet, which means the Internet is more frequently used than any other major medias such as TVs, radios, and newspapers. One of reasons which make the Internet the most popular tool is that we can search for the travel information anytime and anywhere with it. For example, if users ask travel agencies to make travel plans, they must visit the agencies during their business hours. In other words, they cannot make plans for trips if they have no time in the business hours. Using the Internet, they can search for travel information even in the midnight. On the other hand, there is too much travel information on the Internet. A user must select appropriate travel information so as to decide on a vacation place she wants to visit. Keywords on vacation places must be specified in the existing search engines, such as Google [2] and major support sites for travel planning [3]. Suppose a user desires to visit hot springs she can visit within 3 hour driving from her home town. She specifies the keywords such as the province names around her home town and ”hot spring” in the search bar of Google. It provides many web pages including the province names and keyword ”hot spring” in their titles and contents. She looks for useful information in order to plan her trip from the Web pages.

C. Problems in Internet search

In the Internet search with keywords, a user must embody her trip goals and preference in advance, so that she can specify them in the search bar. It means that she can search for suitable vacation place using the Internet only if she has embodied her trip goals and trip preference. The constraint brings about harmful influences. Users with little knowledge on trips would have difficulties to embody their goals and preference for trips. The lack of embodied goals and preference prevents them from expanding trip images. Without expanded trip images, they cannot conceive good keywords to be specified in search with the Internet. Consequently, they have little chance to find suitable vacation places matching their trip images. On the other hand, users with a lot of trip experiences have often fixed their trip preference. When they search for vacation places

 ISBN: 978-1-61208-428-2

in the Internet, they are likely to specify same keywords in all cases. They are disappointed with stereotyped trip images brought by search results. They would not be excited in conventional events which they experienced in different places. They experience nothing novel. They fall into a condition that there is no places which they want to visit although they want to go on trips. There is a common issue in both of users lacking trip experiences and ones abundant in trip experiences. They cannot figure out clear trip goals and preference enough to make them exited, when they make trip plans. In order to make search of vacation places with the Internet beneficial to them, it is inevitable for them to embody their trip goals and preference.

D. Related works

There are existing studies to recommend vacation places to users using the Internet [4][5][6][7]. As it is described in the previous section, users have to embody their goals and preference of trips so as to express them as keywords. In order to solve this problem, the works in [8], [9] and [10] address to recommend suitable vacation places matching their trip goals and preference. These works use location information of pictures the users have taken in vacation places as their travel histories. The works recommend vacation places through the user-based collaborative filtering from the travel histories. However, they leave 2 issues unsolved. First, they cannot provide any vacation places matching new trip goals or preference, because they use travel histories. Users cannot be excited when they recommend the users vacation places similar to what the users have been to. Second, users get excited when they are going to experience something matching their latent trip goals and preference. However, it is not easy to embody latent trip goals and preference so that they can express them as keywords. In order to solve these issues, we need to let users get aware of their unknown trip goals and preference they have never experienced. In addition, we need to reflect their unknown trip goals or preference to the result of recommendation. Furthermore, according to Section II-A, it is important for all users to identify their trip goals and trip preference, in a short time and with little effort.

III. Method to Embody Trip Goal and Preference

A. Categories and tags

The paper proposes a method to embody trip goals and preference of users from their browsing logs of trip place photographs on a web site. When the users see photographs of vacation places, users can momentarily judge whether they want to visit or not there, because photographs give users more intuitive impressions than texts. In addition, photographs represent vacation places as they are. Users can expand easily their trip images through photographs of vacation places. Photographs express moods, which are difficult for texts to convey. Consequently, users can expand their trip image from photographs more concretely than text. The paper assumes that a photograph presented to users on a web site has one big object and background, as depicted in Figure 1. When users looking for vacation places see photographs of specific sightseeing points, they immediately judge whether vacation places in the photographs match with their trip goal or not. If they match, the method assumes they want to enjoy the mood big objects and backgrounds in the photographs bring. This

Fig. 1. Photograph using proposed method

paper refers to the mood as a category. It is expressed with a phrase such as ”hot springs” and ”landscape”. The category is used to estimate a trip goal of an user. In addition to that, this paper refers to the object shown in the photograph as its core as a tag. The tag is also expressed with a phrase such as ”outdoor bath” and ”sunset”. The tag plays an important role to estimate trip preference of the user. When a user searches for vacation places using photographs, it is easy for her to image herself in the vacation place. Imagining herself in the vacation place, she considers whether the background and the object in the photograph suit her trip preference. In this process, she embodies her unclear trip preference, seeing the object in the photograph.

Photographs used in this study have more than one category and tag. For example, the photograph in Figure 1 has the categories, ”nature” and ”Buddhist temple”, and the tags, ”tinted autumnal leaves” and ”temple”. Various categories and tags are chosen to represent trip goals and preference common to many users. We prepare so many photographs that they should cover all common trip goals and preference without loss. We make correspondence of each photograph to some of trip goals and preference. Users can embody their unknown trip goal and preference when they see the photograph in the process of imagining their trips.

B. Embodying trip goal and preference from photographs

The method makes trip goals and preference of users clear so that vacation places can be easily searched with exciting search engines on the Internet. It lets users embody their trip goals and preference, which would be used as keywords for search engines. Users are supposed to specify the keywords for search engines to find vacation places suitable for the users. Figure 2 illustrates the overview of the method. In the method, users increase points of photographs which contain her favorite combination of objects and backgrounds. The increment is referred to as “evaluation” by users. The points are provided for categories and tags associated with the photographs. In the screen where many trip photographs arranged in a random order, the method lets users evaluate her favorite ones. The trip photographs are arranged on the screen, according to the order calculated with Equation (1). p_m is the probability category m is selected. E_t is the entropy of time point t.

$$E_t = -\sum_m p_m \log_2 p_m \quad (1)$$

Through the user evaluation of photographs, the method estimates their trip goals and preference. As Equation (1)

 ISBN: 978-1-61208-428-2

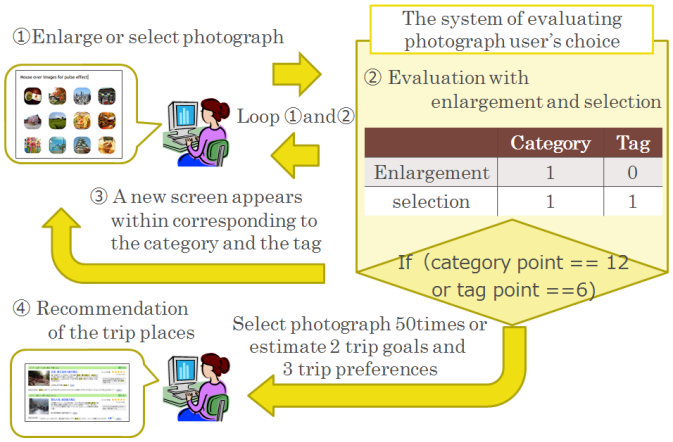


Fig. 2. Overview of method

Fig. 3. Screen of evaluation

shows, trip photographs on one screen are arranged so that the entropy in the screen is maximized in terms of trip goal and preference. When users choose specific ones from randomly arranged photographs, they are biased against the particular trip goal and preference represented with the photographs. Suppose a specific category or tag reach is evaluated more than a predefined threshold times by a user. The user is estimated to have interests in a trip goal or preference corresponding to the category or the tag. At that time, a new screen appears. The screen is filled up with new photographs corresponding to the category or the tag chosen in the previous screen. The method lets the user evaluate photographs in the new screen, again. It aims at estimating other trip goals or preference, assuming the uses have the chosen trip goals and preference. When the number of estimated trip goals and preference reach to predefined one, the method finishes evaluation of photographs by user, considering it has already gather trip goals and preference sufficiently. After that, it passes estimated trip goals and preference as keywords of search engines to recommend vacation place.

C. Evaluation with enlargement and selection

The method estimates trip goals and preference of a user by evaluating photographs. Figure 3 shows the screen for evaluation with photographs. From the screen, a user is assumed to enlarge or select photographs which suit her trip goals and preference.

A photograph is enlarged when its main object pleases her, as depicted in figure 3. Since the size of photographs for the evaluation is 64 pixels in both of the length and width, the user would desire to make their image clearer with a simple way, when they focus on them. A user can enlarge photographs when the mouse pointer hovers over them more than 1 second. When the user enlarges the photograph, the method adds a point to the categories the photograph corresponds to, because the method regards she likes the mood it presents.

The photograph is selected when the enlarged photograph is clicked. When the user judges the photograph suits her trip preference after her enlargement, she selects it to confirm its detail. The method considers the user takes the selection action, because she has strong interests in the photograph after the enlargement. The method regards she tries to confirm its detail with the selection, when she has judged its object suits her trip image well.

Suppose a user enlarges and selects the photograph having a hot-spring category and an outdoor bath tag. The method estimates that she is pleased with the mood the photograph presents. It adds a point to the hot-spring category. At the same time, the selection makes the method estimate that photograph suits her trip preference. The Outdoor bath tag gains a point.

D. Recommendation of vacation places

The trip goals and preference of the user are extracted as categories and tags represented with phrases. Using extracted phrases as keyword, the method searches for vacation places with a major travel support site [3] for travel planning. The site provides vacation places associated with the keywords in their information and reviews from other travelers. The method recommends the top 5 vacation places in the recommendation ranking. It recommends more than one vacation places because the user can choose geographically suitable places for her trip. Suppose the categories and tags finally estimated as trip goals and preference of the user are “hot spring”, “outdoor bath”, “sunset” and “sea” through the evaluation of photographs. The method searches for vacation places, presenting the estimating phrases as keywords. As a result of the search based on the keywords, some hot spring resorts are presented such as ones allowing traveler to enjoy sunset from the outdoor bath.

IV. USABILITY OF RECOMMENDATION USING PHOTOGRAPHS

A. Threshold for evaluation times

We have developed a Web based system to evaluate the usability of the method. The system presents numerous photographs using several screens. When a user evaluates a specific category or a specific tag more than predefined times, another screen appears, filled with new photographs. The thresholds for the evaluation times for categories and tags are θ_c and θ_t, respectively. We have conducted a preliminary experiment in order to decide on the thresholds. The preliminary experiment examines the combination of the 2 kinds of thresholds which brings the largest satisfaction of users for vacation places. Let us refer to the number of enlarging or selecting photographs associated with a specific category or a specific tag as the number of evaluation times. Table I

 ISBN: 978-1-61208-428-2

TABLE I. COMBINATION OF THE NUMBER OF EVALUATION

	A	B	C	D
Number of evaluation for category	8	10	12	14
Number of evaluation for tag	4	5	6	7

TABLE II. LIST OF CATEGORIES AND TAGS (1)

landscape	nature	hot springs
seaside	tinted autumnal leaves	outdoor bath
native province	cherry blossoms	indoor bath
prospect	bamboo forest	hot spring district
river	pine-covered area	public bath
star sky	valley	
night view	forest	

TABLE III. LIST OF CATEGORIES AND TAGS (2)

event	sport	Buddhist temple
carnival	camp	temple
illuminations	water sports	shrine
the Star Festival	mountain climbing	
tradition	ski slope	
	sea bathing	

TABLE IV. LIST OF CATEGORIES AND TAGS (3)

leisure spots	vehicle	history	Japanese-style hotel
amusement park	sightseeing boat	Japanese castle	seafood
aquarium	ropeway	historic spot	food of the mountain
zoo	truck	historic building	tea lunch
botanical gardens		old highway	atmosphere
shopping		garden	European
art gallery			
pasture			

shows 4 kinds of combinations of the number of evaluation times for categories and tags examined in the preliminary experiment. 100 Vietnamese university students from 19 years old to 20 years old have joined to the experiment as test subjects. The 100 test subjects are divided into 4 groups, each of them consists of 25 people. Let us refer to them such as “A”, “B”, “C” and “D”. In the preliminary experiment, subjects use the web site, imaging the place which they want to visit in Japan. They choose their favorite vacation places among recommended ones. They express their satisfaction degrees for each vacation place recommended as a final result in a questionnaire. Since combination C brings the highest satisfaction, the thresholds for the evaluation times to present a new screen are $\theta_c = 12,\ \theta_t = 6$.

B. Experiment to verify usability of photographs

We have conducted a main experiment to verify the usability of this method, under the following condition. When the number of estimated trip goals and preference reach to predefined thresholds for trip goal, θ_p, and for trip preference, θ_o, the method finishes evaluation by a specific subject. In the main experiment, $\theta_p = 3$ and $\theta_o = 2$. Furthermore, the method finishes the evaluation in case a subject evaluates photographs more than the predefined maximum number. In the main experiment, the maximum number of the evaluation is 50. Table II and Table III show the categories and tags used in this experiment. The categories and tags are determined referring to the items of trip genre in a travel support site for travel planning [3] so that any prejudice affects the arbitrariness of the experiment. To prepare trip photographs used in the experiment, the names of categories and tags in the Table II and III are specified as keywords in Google visual search [11]. From photographs the site presents, we manually select photographs suitable for trip photographs from viewpoints of comprehensibleness.

There are 4 points to be verified in this experiment. The first point is the ability of trip photographs to allow users embody their trip goals and preference. The second one is the satisfaction degree of users for vacation places recommended based on the embodiment of their trip goals and preference. Third, we have examined whether users can search for vacation places matching their trip goals and preference in a short time. Finally, we verify the method enables users to search for more easily vacation places, compared with the major support sites for travel planning. The test subjects are 18 male students and 7 female students of universities in Japan. Their ages range from 18 years old to 26 years old. The experiment compares the proposed method with the way with which people generally search for vacation places with search engines through the Internet. The paper refers it to as the conventional way. This experiment supposes 2 situations. The first one supposes a user searches a spring trip in a place where she and her friends have never been. In the second one, a user searches for vacation places which please her close friend or her boyfriend in their winter vacation. In this experiment, 25 test subjects are divided into 2 groups: A and B. The grouping aims to reduce the influence of test subjects which have variance in their ability to image the specified situation easily. The experiment lets one group apply the proposed method to the spring situation and the conventional method to the winter situation, while the other searches for vacation places for the spring situation with the conventional method and for winter situation with the proposed method. This experiment lets each of test subjects search for a single vacation place which best suits for an image the test subject has for each situation.

In the experiment, first, a user searches for one vacation place which satisfies her best for the specified situation, using the conventional method. Second, using the proposed method, the user chooses one vacation place which best suits for the other situation. Finally, in regard to the vacation places they find out in the process above, we have asked several questions to the user. After the proposed method estimates the trip goals and preference of the user, we have prepared 3 sets of keywords; the first one consists of only the trip goals, the second includes only the trip preference, and the last is composed of both of the trip goals and the trip preference. For each of 3 keyword sets, we have collected 5 vacation places at most. Since the number of collected vacation places varies with keyword sets, the total number of recommended vacation places ranges from 5 to 15. Each test subject determines a single vacation place he or she wants to visit most.

C. Embodiment of trip goals and preference

We examined whether users can embody their trip goals and preference with trip photographs. In the questionnaire

 ISBN: 978-1-61208-428-2

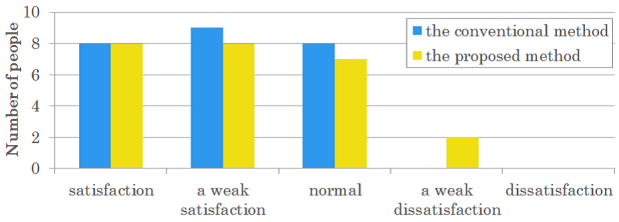


Fig. 4. User satisfaction

TABLE V. Average and dispersion of working time

	Conventional method	Proposed method
Average	35.32	7.045
Dispersion	387.25	12.39

asking “Can you embody your trip image with the trip photographs?”, about 92% users said “yes”. It showed users were able to embody their trip image. It seemed that users enlarged or selected trip photographs to see the details of what were presented in the photographs. They examined the details to see whether the photographs met their goals and preference. Operations of enlarging or selecting photographs implied trip goals and preference of users. Furthermore, in interviews after the experiment, some users uttered that the photographs made them aware trip preference they had never found. As a result, users were able to find out their latent trip preference from the photographs.

D. Satisfaction for vacation places

Let us see whether users were satisfied with vacation places recommended based on their trip goals and preference. According to the results of the experiment, about 64% of users were satisfied with recommended vacation places. Figure 4 compares the number of users satisfied with the proposed method with that of users satisfied with the conventional method for the same situations. The comparison of the 2 results indicates the proposed method recommended vacation places with the same level of satisfaction as the conventional method.

E. Time for searching

Since most users need a recommendation method which presents vacation places with few operations, it is important whether users can search for vacation places corresponding to their trips goals and preference in a short time. Table V compares the proposed method with the conventional method in terms of the working time, and its mean along with the deviation, respectively. Users took 7 minutes in average to find out favorite vacation places with the proposed one, while 35 minutes in average with the conventional one. Moreover, all of users were able to find out the vacation places in shorter time with the proposed one than the conventional one. Together with the result of the previous section, the proposed method contributes to finding as much satisfying vacation places as the conventional method in a shorter time.

F. Effort for searching

It makes no sense if users need lots of effort to search for vacation places using this method. There are the results of the questionnaire asking the following 2 points. The first question is “which method makes you go into deep thought, the conventional method or the proposed method?” The second is “which do you think impose more bothersome operations, the conventional method or the proposed method?” In the first question, about 72% users said “the conventional method”. In the second question, about 76% users said “the conventional method”. In regard to the result, we have tested any difference of this method from the conventional method with the chi-square test. The null hypothesis was that there was no difference between the conventional method and the proposed method, while the alternative hypothesis was that there was a difference between them. Table VI and VII show the test results. Regarding to the first question, there was a difference at the significant level of 5%. There was a significant difference in at the 1% level as for the second question. From both viewpoints, the proposed method allowed users to search for easily vacation places than the conventional method.

TABLE VI. RESLUT OF FIRST QUESTION

	Answer
Chi-square	4.84
Flexibility	1
p-value	0.0278069

TABLE VII. RESULT OF SECOND QUESTION

	Answer
Chi-square	6.76
Flexibility	1
p-value	0.009322376

V. DISCUSSION

A. Improving precision of estimation

Let us consider to improve the precision for the estimation of trip goals and preference of users. In the questionnaire to be answered in 5 grades, about 36 percent of users answered “normal” or “a weak dissatisfaction” with vacation places the proposed method recommends. As reasons of the dissatisfaction, some users mentioned the followings.

1) “It does not seem photographs I have chosen are not reflected to determine vacation place candidates.”
2) “Though I have chosen many food photographs, the method mainly presents photographs of ruins or red leaves of autumn in the next screen.”

The dissatisfaction may attribute to low quality in estimating trip goals and preference of users using categories and tags associated with photographs.

Reason 1 implies a problem on the way to select new photographs to be presented for users from their operations on previous photographs. New photographs are selected using categories and tags attached to previous ones. It is suspected categories and tags are determined based on the arbitrariness of the implementers of the proposed method. Suppose a user prefers a photograph because she likes the sky presented in it, while tags indicating a temple and red leaves are attached

 ISBN: 978-1-61208-428-2

to it. Though the method values tags indicating a temple and red leaves, it never picks up the trip preference of the user directing the sky. In this experiment, the implementers have determined manually categories and tags, which allows the arbitrariness of the implementers to be reflected strongly on them. Consequently, it is conceivable trip goals and preference are estimated using photographs whose categories and tags are biased. It prevents the method from estimating proper trip goals and preference users have. We need an objective way to determine relevant categories and tags.

Reason 2 notifies us of unfairness in the evaluation of photographs. The method evaluates all categories and tags with a constant weight, when a user enlarges or selects a photograph. Objects on photographs have large variety in their size. Depending on the size, they provide different impressions. For example, if a user looks at a photograph with cherry blossoms expanding, she necessarily focus on the cherry blossoms. However, if a photograph shows cherry blossoms in small size in front of big Mt.Fuji, a user may not always focus on cherry blossoms. The evaluation with constant weights is unfair for photographs containing objects of various size. One way for the improvement is to reflect the size ratio of objects to weight to evaluate photographs.

Solutions on these problems would improve the precision in the estimation of trip goals and preference of users.

B. Improving satisfaction for vacation places

Some users show dissatisfaction for recommended vacation places, because the method presents only places they have already visited, or the method recommends vacation places they cannot visit. A main reason for the dissatisfaction is lack of consideration of travel histories and hometowns of users. Without the consideration, the method may recommend vacation places users have been to, or cannot visit geographically. The method can avoid the problems, taking travel histories and hometowns of users into consideration in advance. It needs to construct a data base reflecting travel histories and hometowns of users. The data base should classify vacation places associated with categories and tags into geographical areas to calculate the accessibility of users. It should also register vacation place each user has been to in advance. Recommendation of vacation places using the data base satisfies more users.

VI. CONCLUSION

This paper has proposed a method to embody user requirements for trips. The method estimates trip goals and preference of users who enlarge and select trip photographs during they are making trip plans. After the estimation of the trip goals and preference, the method specifies them as keywords in a major support site for travel planning, to get recommends on vacation places which suit for trip goal and preference of each user. An experiment has proved about 64% users are satisfied with vacation places recommend using this method, which is as reliable as conventional methods. All of users can materially shorten the time to decide on their vacation places with the proposed method than with the conventional one. Users can decide on trip plans suitable for them in a short time with the method. We have also discussed objective categories and tags, as well as construction of a data base to enhance user satisfaction as our future works.

REFERENCES

[1] Ministry of Public Management, Home Affairs, Posts and Telecommunications (2012) White paper Information and Communications in Japan: http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h24/html/nc123340.html [retrieved: 12, 2014]

[2] Google: http://www.google.co.jp/ [retrieved: 12, 2014]

[3] Jaran saported by Recruit: http://www.jalan.net/ [retrived: 12, 2014]

[4] Vincent. W. Zheng, et al. ”Towards mobile intelligence: Learning from GPS history data for collaborative recommendation.” Artificial Intelligence 184 2012, pp. 17-37.

[5] Ye. Mao, et al. ”Exploiting geographical influence for collaborative point-of-interest recommendation.” Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. ACM, 2011, pp. 325-334.

[6] Michele. A. Burton, et al. ”Crowdsourcing subjective fashion advice using VizWiz: challenges and opportunities.” Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility. ACM, 2012, pp. 135-142.

[7] Joan. Borrs, et al. ”Intelligent tourism recommender systems: A survey.” Expert Systems with Applications 41.16 2014, pp. 7370-7389.

[8] Marco. Mamei, et al. ”Automatic Analysis of Geotagged Photos for Intelligent Tourist Services.” Intelligent Environments. 2010, pp. 146-151.

[9] Maarten. Clements, et al. ”Using flickr geotags to predict user travel behaviour.” Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. ACM, 2010, pp. 851-852.

[10] Adrian. Popescu. and Gregory. Grefenstette. ”Mining social media to create personalized recommendations for tourist visits.” Proceedings of the 2nd International Conference on Computing for Geospatial Research & Applications. ACM, 2011, pp. 37.

[11] Google visual search: https://www.google.co.jp/imghp?hl=ja&tab=wi [retrived: 12, 2014]

 ISBN: 978-1-61208-428-2

Estimating Consumer Inclination for Agricultural Products from Web Browsing History

Riki Tatsuta, Sunao Nakanishi
Graduate School of Information Science and Engineering
Ritsumeikan University, Shiga, Japan
email:{tatsuta, sunao}@de.is.ritsumei.ac.jp

Yusuke Kajiwara, Hiromitsu Shimakawa
College of Information Science and Engineering
Ritsumeikan University, Shiga, Japan
email:{kajiwara, simakawa}@de.is.ritsumei.ac.jp

***Abstract*—In recent years, more and more e-commerce sites sell agricultural products. It should be noted that users have preference of taste. They also have intention on what kinds of agricultural products they get. They consider them at the time of purchase of agricultural products in recommendation on e-commerce sites. There are a number of studies on the automatic extraction of preference of taste that users have from contexts on the web. However, there has been no study that tries automatic extraction of user intentions to purchase agricultural products on e-shopping sites. E-shopping sites without consideration of user intentions are likely to recommend agricultural products far from what users want to buy. Confused users will never visit such e-shopping site, again. We propose a model which estimates intention level of users from web contents of e-commerce sites and web browsing history of users. The proposed method refers to the degree of user interest about a topic calculated from LDA as the topic attention level. Based on the multivariate logistic regression model, the method constructs an intention level model from the topic attention level. An experiment suggests the method can calculate the safety intention level, focusing on price of agricultural products on web pages.**

Keywords–recommendation; inclination; intention; preference; agriculture; e-commerce site; browsing history.

I. INTRODUCTION

More and more e-commerce sites sell agricultural products. Sale of agricultural products on the Internet has several good effects. Farmers can increase their sales. They are also motivated through direct communications with consumers. Supply of high-quality agricultural products provides them with good brand impression [1]. E-commerce brings convenience of direct delivery of agricultural products to consumer houses. The elderly households are too weak to go out shopping food. Households who have small children have no time to go out for shopping. They are eager to buy food with Internet [2]. On the other hand, farmers need efforts to make more consumers visit their own sites to make success in e-commerce selling agricultural products. Their e-commerce sites present information which motivates consumers to purchase their products. In addition to reviews, prices, and quantity like general e-commerce sites, they have information on varieties and palatability of products. They also present cultivation methods to explain their cultivation policy [3].

Information motivating consumers to buy agricultural products varies with preference of consumers. The preference is determined by various factors. For example, preference on cooking with agricultural products ranges from foodstuff, menu, mouthfeel, taste, seasoning and flavor [4]. Apart from these, consumers have intentions on their peculiar ideas such as safety and costs [5]. Some consumers give a higher priority to safety and palatability than to prices, while others mind prices rather than appearance. A combination of preference and intentions of each consumer forms specific inclination which regulates purchase behavior of the consumer. Farmers need marketing activities to target consumers whose inclination meets their cultivation policy. However, Japanese farmers engaging in highly intensive agriculture cannot afford to consider marketing activities because they are busy with cultivating many kinds of crops. One possible solution to support attracting consumers is Search Engine Optimization (SEO), which makes the site appear more frequently in result list of Internet search engines. However, the solutions cannot support the site builders should make consumers interested in the sites so that the consumers can buy agricultural products buy according to their inclination. From the viewpoints of consumers, they want to buy agricultural products from farmers whose cultivation policy meets their own inclination. For example, consumers with safety intentions would read politely safety explanations of products in the e-commerce sites, because they want to purchase safe products as possible, as far as their price permits. They want to get foods free from pesticides. Trading on the current e-commerce never considers coincidence of farmer cultivation policy with inclination of consumers, when customers search farmers optimal to them. In fact, consumers must manually look for agricultural products fitting their needs from a lot of e-commerce sites, struggling with search engines. However, search engines disappoint them many times, introducing sites which handle mismatching agricultural products.

This paper proposes a method to estimate inclination of consumers, in order to recommend appropriate e-commerce sites to them when they search farmers. The method utilizes Web contents on the e-commerce sites and browsing history of customers. The method figures out the distribution of topics on farmer cultivation policies from a specific Web page using the Latent Dirichlet Allocation (LDA). The paper refers the product of the topic distribution and the browsing time of the web pages as the topic attention level. The method constructs an inclination level model using the topic attention level as predictor variables, based on the multivariate logistic regression. An experiment result suggests the method can calculate the safety inclination level.

This paper is organized as follows. Section II discusses the detail of inclination. The related works are also explained in Section II. Section III describes the outline of the proposed

 ISBN: 978-1-61208-428-2

considered recommendation system. The proposed method to estimate inclination is also described in Section III. Section IV presents the results of experiments and evaluation. Section V presents the discussion on improvement of the proposed method. Finally, Section VI concludes this papers.

II. PRODUCT INFORMATION MATCHING INCLINATION

Consumers have preference on agricultural products, such as taste, appearance, flavor and so on. The preference is said to differ by age, generation, gender, regions and occupations [6]. In addition, consumers would stick to their peculiar ideas such as safety and costs, when they choose products to buy. Consumers generally determine products to buy with their preference and intentions. The paper refers the combination of preference and intentions of each consumer as inclination. It usually regulates purchase behavior of the consumer.

There are studies on the estimation of consumer inclination [7][8]. For the estimation, the studies oblige consumers to fill up questionnaires and profile in free description. They analyze results with text mining technologies to find user inclination for agricultural products. However, consumer inclination changes depending on fashion and trend. For example, users minding safety, which means products are free from harmful substance to eat, have greatly increased in Japan because of the Fukushima nuclear accident. It is necessary to successively examine consumer inclination because the thinking ways of consumers vary from day to day. It is an unendurable for consumers to fill up questionnaires and profiles many times. A new method is required to extract current inclination of each consumer without bothering the consumer.

III. RECOMMENDING AGRICULTURAL PRODUCTS MATCHING INCLINATION

We aim at detecting user inclinations from Web browsing log before they decide which farming products they buy. It does not make any sense to analyze huge data obtained after they decide to buy. The log of URLs in Web browsing of users is also insufficient to achieve the aim above. We need to analyze the contents to see what topics in the pages stimulate the users. The LDA on the contents of Web pages they are interested in enables us to identify their topics.

A. Outline of recommendation with extracted inclination

To extract inclination of consumers at the time, we use their browsing history while they look for their favorite agricultural products. We extract consumer inclination from Web pages the consumer visits while he chooses agricultural products in e-commerce sites. E-commerce sites present a variety of Web pages. In the web pages, there are contents which provide stimuli to consumers, because the contents match their inclination. Inclination of each consumer is likely to appear when he selects products. It seems that Web pages vary with consumers because of their inclination. For example, it may be supposed consumers who put importance to safety purchase domestic or organic tomatoes. In this case, they are supposed to investigate description on agricultural chemicals or product places. On the other hand, consumers preferring economy to safety purchases foreign tomatoes, because of low price. In this case, consumers would browse pages of agricultural products after they sort the pages in the decreasing order of the price. It may be suspected that browsing time of Web pages gets long if

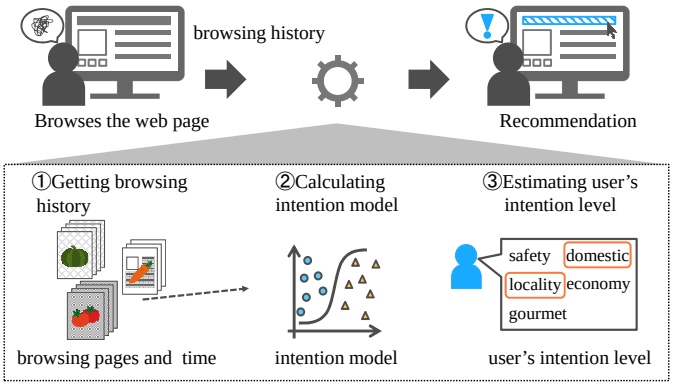


Figure 1. System configuration diagram

they have stimulated consumers. In the proposed method, we estimate consumer inclination from topics described in each Web page and the browsing time to read the Web page. In this work, we focus on inclinations: safety, domestic production, locality, economy, and gourmet from the survey results of inclinations on foods [6] and the studies of selection criterion at purchase of agricultural products [9]. As Figure 1 shows, the paper proposes a method to build a model to estimate consumer inclination from browsing history of the consumer to provide him with agricultural product information. At first, the method extracts browsing pages and browsing time from a browsing history when the consumer chooses agricultural products on e-commerce sites. Second, it calculates the degree of how much the consumer has each inclination. It refers to the degree as the inclination level. Finally, the method recommends the agricultural product information which suits his inclination.

There are various agricultural products. For example, organic tomatoes, fruit tomatoes, cherry tomatoes are classified into the category of tomatoes. These categories are referred to as product categories in the paper. Though consumers searching products visit various sites, the method values pages for a specific product category as item pages, rather than pages showing lists of item pages, which are referred to as search pages. Pages irrelevant to e-commerce sites such as Google search and Wikipedia are referred to as other pages in the paper. Item pages are classified into product categories, based on words on their titles and contents. From browsing history of each consumer, item pages are extracted for each product category. As Figure 2 shows, the proposed method builds a model to estimate the inclination of each consumer as follows.

1) The item pages are regarded as a set of words according to the Bag-of-words method [10]. The topic distribution of the item pages is figured out with the LDA. The proposed method calculates the topic attention level with the product of the topic distribution and the browsing time of each item page.
2) The method selects important variables to calculate the inclination level of the consumer.
3) Using the inclination level, the method constructs an inclination model for each consumer.

B. Weighted topic distribution

Through a morphologically analysis, the proposed method divides a set of sentences in item pages into words. It makes

 ISBN: 978-1-61208-428-2

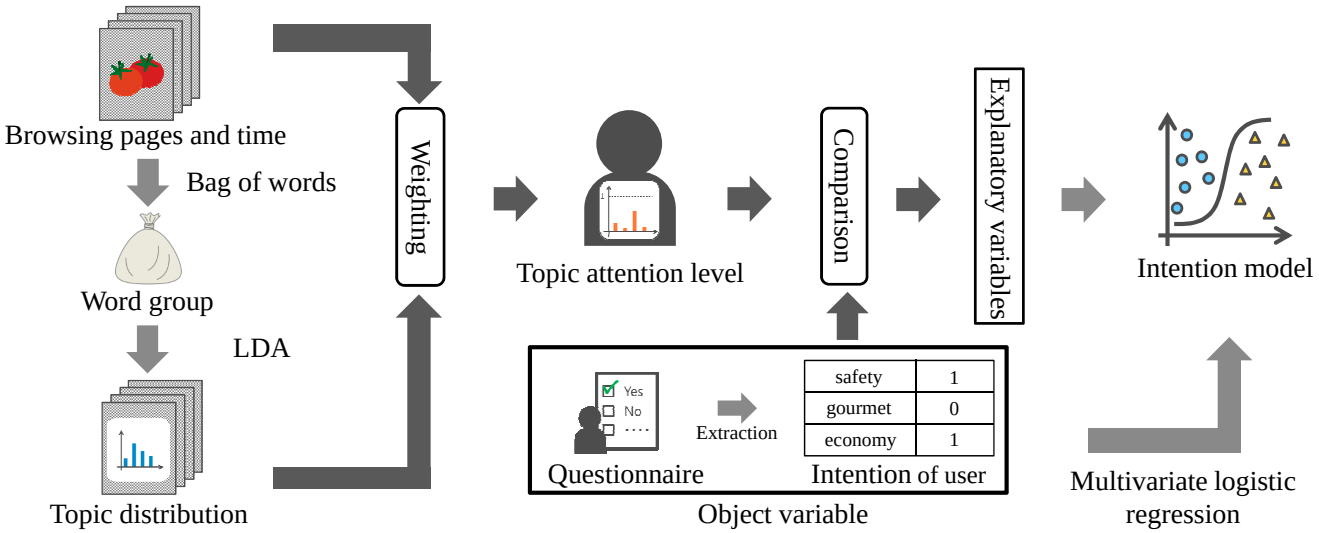


Figure 2. Proposed method

bag-of-words from nouns, adjectives and verbs which seem to be related to inclination. Using the LDA [11], the method acquires combinations of topic distribution and its word group for each page. Though words in the item page are relate to various kinds of inclination, the LDA allows to get topics classified according to each inclination.

The proposed method uses browsing time to determine whether the item page is related to a specific inclination or not. The previous work [12] has clarified a relationship of the browsing time of a Web page to the degree of user interest. LDA is a model which assumes the document is approved from two or more topics. The degree of the interest to a document is that to the topic on the document. The degree of the interest to the topic of the page are related with user's browsing time. The words which relate to several kinds of intention may be included in the e-commerce sites selling agricultural products. However, if each page is weighted by browsing time, user's focus point of words of the page is not considered.

Words which relate to the intention should be classified in one page. This technique subdivides words by a topic distribution of each item page, which enables the acquisition of the degree of each of the user's intentions. Equation (1) shows topic distribution the vectors $\boldsymbol{T_{cp}}$.

$$\boldsymbol{T_{cp}} = [T_{cp1} \quad \cdots \quad T_{cpk} \quad \cdots \quad T_{cpK}]^{\mathrm{T}} \qquad (1)$$
$$(1 \leq c \leq C), (1 \leq k \leq K), (1 \leq p \leq P_c)$$

where k, c, and p represent a specific topic, a specific product category, and a specific page, respectively. K and C are the number of topics and product categories, respectively. P_c is the number of item pages which are contained in the c-th product category. T_cpk is the k-th component of topic distribution of the p-th page of the c-th product category. Equation (2) shows vector I_c. Each component of it is a sum of T_cp weighted by the browsing time of each page in each product category.

$$\boldsymbol{I_c} = [I_{c1} \quad \cdots \quad I_{ck} \quad \cdots \quad I_{cK}]^{\mathrm{T}} = \sum_{p=1}^{P_c} (\boldsymbol{T_{cp}}) d_{cp} \qquad (2)$$

I_{ck} is the weighted vector component relating to the k-th topic of the c-th product category. d_{cp} is the browsing time of the p-th page belonging to the c-th product category. The more pages belonging to the same commodity category the consumer browses, the more I_{ck}grows. It is necessary to normalize I_{ck}, because the browsing time and the browsing count vary with consumers. Equation (3) shows $S(I_{ck})$ which is calculated with normalized I_{ck}.

$$S(I_{ck}) = \frac{I_{ck}}{|I_c|} \qquad (3)$$

In this work, we refer to $S(I_{ck})$, which is extracted from each product category as the topic attention level.

C. Selection of important variables

In Section III-C, we account for the explanatory variables of the making model as shown in Section III-D. The topic with difference of topic attention by the presence of the intention is extracted. Equation (4) shows $G\{n_{ck}\}$, the set of the topic attention level of each consumer.

$$G\{n_{ck}\} := \{S(I_{ck})_1, \cdots, S(I_{ck})_u, \cdots, S(I_{ck})_U\} \qquad (4)$$
$$(1 \leq u \leq U)$$

Whole users are divided into 2 groups: one with a specific inclination and the other without it. In Equation (5) and (6), $X\{n_{ck}\}$ and $Y\{n_{ck}\}$ show consumers with the n-th inclination and without it, respectively.

$$X\{n_{ck}\} := \{G\{n_{ck}\}|P(\{G_{ck}\})\} \qquad (5)$$

$$Y\{n_{ck}\} := \{G\{n_{ck}\}|Q(\{G_{ck}\})\} \qquad (6)$$

P is the condition that the consumer does not has the inclination. Q is the condition that the consumer has the intention. n and N corresponds to the n-th inclination and the number of it. n_{ck} is the k-th topic of the c-th product category to correspond to the n-th inclination. The method aims to build a model to estimate inclination of unknown consumers. It assumes inclination of several consumers is acquired beforehand by questionnaires or interviews. Though it obliges some efforts on

 ISBN: 978-1-61208-428-2

some consumers, there is no burden on unknown consumers. The method calculates difference of the topic attention level of the group with the inclination and that without the inclination. Let us focus on the difference of the topic attention levels of the two groups. A large difference in a topic indicates the group with the inclination pays their attention on the topic, while the other group pays no attention on it. To examine the difference of the topic attention level, the method uses median values resistant to noises. Equation (7) and (8) calculate the median values of the group with the inclination and the group without it.

$$MedianX(n_{ck}) = \begin{cases} x_{(a+1)/2}, & if\ a\ is\ odd. \\ \frac{1}{2}(x_{a/2} + x_{a/2+1}, & if\ a\ is\ even. \end{cases} \quad (7)$$

$$MedianY(n_{ck}) = \begin{cases} y_{(a+1)/2}, & if\ a\ is\ odd. \\ \frac{1}{2}(y_{a/2} + y_{a/2+1}, & if\ a\ is\ even. \end{cases} \quad (8)$$

Equation (9) shows $diff(n_{ck})$ which is the difference of $MedianX(n_{ck})$ from $MedianY(n_{ck})$.

$$diff(n_{ck}) = MedianX(n_{ck}) - MedianY(n_{ck}) \quad (9)$$

Equation (10) shows $Diff_n$ which is the set of the difference of the median for the n-th inclination.

$$Diff_n := \{diff(n_{11}), \cdots, diff(n_{CK})\} \quad (10)$$

It seems that the median with great difference is an important value to characterize a group of users who do not have intention and a group of users who have intention. This method picks out top R of large median difference in a positive direction as $High_n$ and lowest R of large median difference in a negative direction as Low_n in $Diff_n$. Topics corresponding to $High_n$ is focused by the group with the n-th inclination, but neglected by the group without it. Topics corresponding to Low_n mean the opposite case.

D. Construction of inclination model

The proposed method builds a model to estimate the consumer intention level. The explanatory variable necessary for the model is a topic attention level which belongs to $High_n$ and Low_n in III-C. The object variable is the inclination of each consumer. The method refers to it as an inclination level model. The method applies the multivariate logistic regression model, to make a good inclination level model. To improve the performance of the model, the method excludes explanatory variables belonging to $High_n$ and Low_n in a stepwise way.

The intention model distinguishes presence of a specific inclination of an unknown consumer from his topic attention level. The topics on each Web page are obtainable beforehand, applying the LDA to the contents on e-commerce sites in advance. What is required for the on-line calculation for the inclination level of the consumer is only his browsing time.

IV. EXPERIMENT

We experimented to verify the utility of the proposed method. The experiment aims at

- validation of efficiency of difference that user browses topic by presence of the intention of the user, and
- verification of validity of intention level model.

A. Experiment method

We experimented about the purchase of agricultural products using e-commerce sites. The research participants were 17. We acquired the browsing history of the research participant from beginning to the end of the experiment. We had to make the research participants unaware of the costs because that would have influenced their selection. We laid out a situation as follows “I want to celebrate grandmother’s sixty-first birthday with about 10 relatives. My mother said that I should cook for the grandmother. You don’t need to think about money, because I’ll pay. I thought that I will try to cook in grandmother’s house, but there is no supermarket nearby. So, I decided to prepare foodstuff using e-commerce sites.”. In addition, the grandmother said “I want to eat home-made dish because I dislike ready-made one. I want you to buy rice you recommend.”. Below are the proposed dishes: salad, rice cooked with matsutake mashroom, boiled dishes, and meat dish. The material necessary for dish is shown as follows: tomato, cucumber, matutake, rice, radish, taro, chicken, and meat. We presented the amount of the foodstuff necessary for the dish and the recipe of the dish to research participants. We specified e-commerce sites to buy agricultural products. We permitted them to use retrieval engines on Internet to gather information. Each of the user’s intentions was acquired in the questionnaire. We made a questionnaire, creating an association between the information that the research participants are interested in e-commerce site and each intention referring to result investigation of intention [6]. It seems that the user who has gourmet intention takes an interest in palatability. For example, information is presented on e-commerce sites such as ”sugar concentrations are higher and sweeter than another”, ”taste becomes better than another by the cultivation method of original” and ”agricultural products harvesting in the morning are very fresh”. The user who has a gourmet intention is attracted by the deliciousness or the sugar concentrations. It is seems that the user who has an economy intention is attracted by low price or large quantity. For example, information about price and quantity are presented on e-commerce sites. Therefore, the user who has an economy intention will browse a page that contains a price or quantity. However, we only asked research participants about whether or not they are attracted by price because of the quantity of agricultural products that they should buy were specified. Questionnaire items for the safety intention, the domestic intention and the locality intention are clear. After buying agricultural products, we send out questionnaires of Table I. in three grades (yes, no, unknown).

TABLE I. QUESTIONNAIRE ITEM

Intention	*Questionnaire item*
Gourmet intention	Q1. Did you interest in deliciousness of the testis?
	Q2. Did you interest in freshness?
Safety intention	Q3. Did you interest in level of safety?
Domestic intention	Q4. Did you interest in domestic production?
Locality intention	Q5. Did you interest in producing area?
Economy intention	Q6. Did you interest in value?

B. Evaluation

1) Discriminating existence or nonexistence of user’s intention by an intention level: Data to acquire topic attention level are documents of item pages which are browsed by

17 research participants. The number of commodity page is 503. We classified those pages in the commodity category. In addition, because the result that research participants bought food stuff of meat dish are divided into pork and beef, we classified them as a same commodity category. Sometimes, information not related to the retrieval intention of the user has been presented in item page. For example, there are information on recommended commodity and the commodity ranking of peculiar to sites. The documents were excluded since these documents are not relate to user's intention. The review which contain on the item page was also excluded.

2) Extraction of topic from the item page: We applied the method explained in III-B to classified item pages. Among words which appear frequently, the intention is not considered from words and symbol of the commodity category which does not relate to the intention. A bag-of-words was prepared, excluding these words. LDA is applied to the bag-of-words using collapsed Gibbs sampling [13]. The scalar value of the Dirichlet hyperparameter for topic proportions is 0.1. The beta hyperparameter for each entry of the block relations matrix is 0.1. The number of trials is 30000. Among the 5 patterns (2, 4, 6, 8, 16 topics), 4 topics presented the best performance. The top 5 words were selected from each of topics.

3) Acquisition of the user's intention: Research participants were responded to a questionnaire of Table I. If user answered YES to the questionnaire of palatability and freshness, the user is regarded to have gourmet intention.

C. Result

The method explained in paragraph III-B was applied to extract the topic. Words were labeled based on generality.

1) Extraction of topic attention level: The method explained in paragraph III-C was applied to extract the topic attention level. Figure 3 shows the median of extracted topic attention level. Table II shows the result of the questionnaire about an intention. ○ is yes. X is no. △ is unknown.

TABLE II. RESULT OF THE QUESTIONNAIRE ABOUT AN INTENTION

User	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*
Q1.	○	○	○	○	○	○	○	X	○	○	○	○	△	○	△	○	X
Q2.	○	○	X	○	○	○	○	△	X	X	△	X	○	○	○	○	△
Q3.	△	X	○	○	○	○	△	△	○	X	X	X	X	○	○	○	○
Q4.	○	○	○	○	△	○	○	○	○	X	○	○	○	○	○	○	○
Q5.	○	○	X	○	○	○	○	△	○	X	X	○	X	X	○	○	○
Q6.	○	X	○	X	○	○	○	X	△	○	○	X	○	○	X	X	○

2) Discriminating existence or nonexistence of user's intention by an intention level: We have verified whether we can discriminate user's each intention existence from user's intention level by the intention model which is explained in paragraph III-D. The Table III shows result of the multivariate logistic regression model. The significance level is less than 5%. The significance level (0.042) of a topic intention relating a value about a safety intention is extracted from result of the multivariate logistic regression model. The significance level (0.042) is a significant value. This result suggests that a safety intention is connected with topics of price. The significance level of each intention except a safety intention has been low. The significant value to estimate user's intention was not found by an intention model.

TABLE III. RESULT OF THE MULTIVARIATE LOGISTIC REGRESSION MODEL

Intention	*Label*	*Estimate*	*Std.Error*	*z value*	$Pr(>\|z\|)$
domestic	agricultural chemical cultivation method	44.99	20068.94	0.002	0.998
	brand	-494.47	168178.07	-0.003	0.998
	domestic, high quality cook	1074.23	358333.88	0.003	0.998
	deliciousness	-898.71	305721.58	-0.003	0.998
	AIC	12.773			
gourmet	cook	30.8814	21.4463	1.44	0.1499
	AIC	18.025			
economy	agricultural chemical cultivation method	-7001	741189	-0.009	0.992
	deliciousness	-23338	2467722	-0.009	0.992
	locality, value	5922	626767	0.009	0.992
	AIC	8			
safety	domestic	2.802	1.984	1.412	0.1579
	value	6.75	3.32	2.033	0.042
	AIC	20.537			
locality	locality	-3.823	2.396	-1.596	0.111
	deliciousness	8.476	6.372	1.33	0.183
	AIC	20.306			

V. DISCUSSION

A. Discussion of the intention model

The significant value to estimate user's intention has not found by an intention model. In this work, it is assumed user's intention is relate to browsing pages of time and contents. The longer users browse pages, the more they are interest in the pages. However, the time changes, depending on how to browse the page. The weighting is considered as one of the causes which makes the significance probability low. The proposed method decides weight without considering how to scan the page. It is investigated that the inspection time changes depending on how to scan the page. The examination method focuses on how to browse page using tab or click. The Table IV shows the counts of how to browse page using tab or click. If the frequency in which the page is opened with a new tab is smaller than that with the current tab, we call those users a "Click type", otherwise "Tab type". The click type are 2, 3, 4, 8, 10, 13 and 14, while the Tab type is 5, 6, 7, 9, 11, 12, 15, 16 and 17. The Table V shows the difference between the average of the browsing time from the average of the browsing frequency. It can be said that browsing time and counts of the tab type are larger than that of the click type. The difference at the browsing time appears in how to scan the page. It is necessary to consider the method of the scanning of the individual when weighting to topics using browsing time. It can be expected that the higher degree of fit model is constructed, considering the method of the scanning.

TABLE IV. THE COUNTS OF HOW TO BROWSE PAGE USING TAB OR CLICK

User	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*
Click	59	1	0	1	42	18	13	1	18	0	13	21	0	1	16	20	18
Tab	9	19	21	66	1	0	3	39	0	24	0	1	53	36	1	2	0

TABLE V. THE AVERAGE OF THE BROWSING TIME AND THE AVERAGE OF THE BROWSING FREQUENCY

	Tab type	*Click type*
Average of browsing time(second)	51.945	42.354
Average of browsing frequency(second)	1.935	1.457

 ISBN: 978-1-61208-428-2

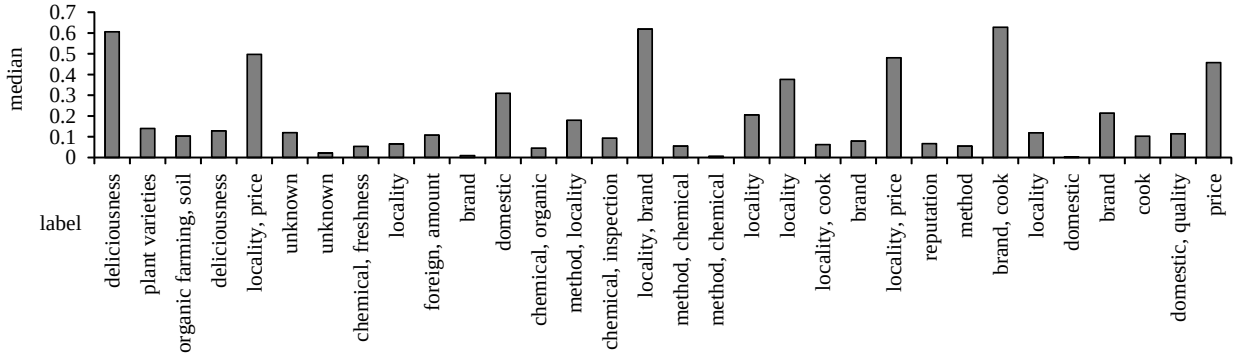


Figure 3. Median of extracted topic attention level

B. Acquisition of Web browsing history

The proposed method needs access to the web history of users. It might sound a rather demanding assumption. Users agreeing to the method are assumed to install an application to collect Web browsing histories. It is a trade-off between their advantages and disadvantages. We should consider two kinds of disadvantages: a burden to install the application, and a risk to expose their privacy. On the other hand, the users are released from stresses coming from recommendation of irrelevant information. We thought the advantages are greater than the disadvantages.

VI. CONCLUSION

The present paper estimated an intention level to present the information based on consumer intention on e-commerce sites. In this method, we extracted item pages which are contained each commodity using browsing history. We prepared a bag-of-words from documents which are contained item pages. We applied LDA to the bag-of-words to extract topic distribution. We extracted topic attention level by weighting browsing time and topic distribution. We constructed the intention model using the multivariate logistic regression model by an important topic attention level and user's intention. The user's intention is estimated from a topic attention level with the intention model in an experiment. The significance level is less than 5%. The significance level (0.042) of a topic intention relating a value about a safety intention was extracted from result of the multivariate logistic regression model. To improve fit degree, we consider transition of the pages. In addition, the questionnaire items do not necessarily correspond to separate categorize item pages. We plan to review the questionnaire.

REFERENCES

[1] Japan Finance Corporation: Survey of sixth sector industrialization in agriculture, https://www.jfc.go.jp/n/findings/pdf/topics_111202_1.pdf [retrieved: March, 2015] (in Japanese).

[2] JC general research: Survey results of the consumption behavior of vegetables and fruits, http://www.jc-so-ken.or.jp/pdf/agri/investigate/141203_01.pdf [retrieved: March, 2015] (in Japanese).

[3] Agriculture, forestry and fisheries. Information Exchange Network Business: Survey results of awareness and intention on the production and distribution information of vegetables, http://www.maff.go.jp/j/finding/mind/pdf/20040820cyosa.pdf [retrieved: March, 2015] (in Japanese).

[4] Ministry of Agriculture, Forestry and Fisheries: Japanese preference and taste, http://www.maff.go.jp/j/keikaku/syokubunka/culture/mikaku.html [retrieved: March, 2015] (in Japanese).

[5] Japan Finance Corporation: Survey of food intention: https://www.jfc.go.jp/n/findings/pdf/topics_140911a.pdf [retrieved: March, 2015] (in Japanese).

[6] Yamaguchi. K, and Takahashi. C, Studies on the Preference of Foods : Part 2 : Relationship between the attribute and the preference of food, Cookery Science 15.2, 1982, pp. 104-113.

[7] Isojima. A, Analysis of Consumers Needs for Agricultural Products by Using Text Mining, National Agriculture and Food Research Organization Vol.16, No.4, 2010, pp. 38-42.

[8] Isojima. A, Analysis of a Consumer Questionnaire Pertaining to Rice by using Text Mining. Agricultural Information Research. Vol.15, No.1, 2006, pp. 49.

[9] Isojima. A, Consumer Perceptions and Selection Standards Rice Purchases : An Investigation into Consumer Considerations, Japanese Journal of Farm Management Vol. 36, No. 3, 1999, pp. 22-32.

[10] C. D. Manning, H. Schutze, "Foundations of statistical natural language processing." MIT Press, Cambridge, MA, 1999, pp. 235-260.

[11] D. M. Blei, A. Y. Ng, M. I. Jordan, "Latent dirichlet allocation." the Journal of machine Learning research 3, 2003, pp. 993-1022.

[12] Morita. M, and Yoichi. S, "Information filtering based on user behavior analysis and best match text retrieval." Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval. Springer-Verlag New York, Inc., 1994, pp. 272-281.

[13] T. L. Griffiths, M. Steyvers, "Finding scientific topics." Proceedings of the National academy of Sciences of the United States of America 101.Suppl 1, 2004, pp. 5228-5235.

 ISBN: 978-1-61208-428-2

Secure Communication Between OpenFlow Switches and Controllers

Dominik Samociuk

Silesian University of Technology
Institute of Informatics
Email: dominik.samociuk@polsl.pl

***Abstract*—We study the applicability of different protocols related to authentication and access control for secure channel between switches and controllers in Software Defined Networks communicating with OpenFlow. We firstly show possible problems with the lack of security mechanisms in OpenFlow architecture. Then we analyze the usability, advantages, drawbacks and implementation details of Transport Layer Security, Secure Shell and IPSec protocols as the secure channel medium for OpenFlow communication between switches and controllers. Finally, we discuss their possible extensions to authentication and access control mechanisms.**

Keywords–Network security; Secure architecture; OpenFlow, Software Defined Network.

I. INTRODUCTION

OpenFlow protocol [1] has generated interest in academic and business society due to the features it offers to architects and developers of Software Defined Networks (SDN). By creating a standardized interface to connect switches with controllers, control-plane logic was moved to a centralized controller (or controller group). However, even when strictly adhering to the specification [2], OpenFlow does not enforce the use of secure communication channel between a switch and controller. Namely, the entry on the Transport Layer Security (TLS) usage was introduced in the OpenFlow specification and then modified. In the latest version, its usage is just a recommendation (not a "must-have" requirement). In the the OpenFlow switch specification ver. 1.4.0 sec. 6.3.3. it reads: "The switch and controller may communicate through a TLS connection". Moreover, due to evolving nature of the OpenFlow protocol, many vendors have not fully implemented this recommendation. Lack of the TLS adoption and problems with the implementation of the TLS infrastructure leave a clear path for attackers to infiltrate OpenFlow networks, using possible attacks described in the following sections. To worsen the security case, there are no authentication nor access control mechanisms (except for an ersatz of authentication in TLS).

Without forcing a secure communication channel, OpenFlow risks repeating the mistakes of other management protocols, designed basing on the assumptions that the link and infrastructure are secure (e.eg. Telnet, SNMPv2, TFTP). Of course, in production environment they must have been replaced with their safe versions (SSH, SNMPv3, SFTP, respectively). One of the proposed solutions for OpenFlow is controlling switches through the Internet in the architecture of branch-office network or offering switch management as "security-as-a-service". If the secure communication channel, authentication and access control are not enforced, OpenFlow will not be able to develop into the above-described roles and will be replaced by a secure protocol (like in the mentioned transition from Telnet to SSH) or by another, secured version of the protocol (like in transition from SNMPv2 to SNMPv3).

In this paper, we compare the possibility of implementing different authentication and access control mechanisms over a secure channel in OpenFlow communication with three popular protocols, namely Transport Layer Security, Secure Shell (SSH) and IPSec.

Transport Layer Security [3] and its predecessor, Secure Sockets Layer (SSL), are cryptographic network protocols for securing data communication. TLS provides confidentiality and integrity of data, as well as the authentication of the server, and sometimes of the client. It is based on asymmetric encryption and can be deployed in two modes, namely as:

- X.509 certificates and Public Key Infrastructure cryptosystem to provide authentication, encryption, integrity and non-repudiation using public and private key cryptography and digital certificates;
- Web of Trust architecture – decentralized authentication method, in which there is no hierarchical structure of the authenticating organizations and trust of each certificate is the sum of the signatures by the other members of the web, signed under this certificate.

SSH [4] is a common name for the whole family of protocols, not just terminal. It is also used for file transfer (SCP, SFTP), remote control of resources, tunneling and other applications. A common feature of all these protocols is identical to the SSH data encryption technology and user recognition. It is possible to configure SSH tunnels to transfer unencrypted traffic on the network through an encrypted channel.

IPSec [5] is a set of protocols for implementing secure connection and encryption exchange of keys between hosts. IPSec can be used for protecting the transmission in three modes:

- host-to-host – between pair of hosts;
- network-to-network – between pair of the security gateways;
- network-to-host – between the gateway and a host.

IPSec consists of at least two channels of communication between connected devices: (a) the exchange channel, through which data associated with authentication and encryption (keys) is transmitted and (b) the channel (one or more) that carries packets transmitted over the already secured line.

This paper is organized as follows. In Section 2, we discuss related works. Section 3 describes the secure channel and authentication problems with OpenFlow in the recent specification. It also presents some possible attacks and their

 ISBN: 978-1-61208-428-2

effects. In Section 4, we present technical details about the proposed solutions, including Transport Layer Security with PKI architecture, Secure Shell tunneling and IPSec protocol. In Section 5, a brief security analysis of the security of the proposed solutions is carried out. Section 6 concludes the paper and presents our planned future work on the subject, including the implementations of new solutions and experiments on PL-LAB2020 - a large testbed and experimental network being built in Poland right now.

II. Related work

So far, the following works have been focused on the secure SDN architecture using OpenFlow.

FlowVisor [6] acts as a transparent proxy between controllers and switches in order to apply limitations to the rules created by controllers. It creates slices (combinations of switch ports, MAC addresses, IP addresses, port addresses or ICMP type) of network resources and delegates the control of slices to different controllers. The role of FlowVisor is to isolate the effect of rewritten rules to the specific slice of the network i.e., one slice cannot control another's traffic.

A similar concept is FortNOX [7] – a software extension developed on NOX controller to check the flow rule contradictions in real time. It uses the role-based authorization on OpenFlow applications (in this case, something that wants to modify the network traffic using the OpenFlow protocol, e.g., firewalls, intrusion prevention systems). The difference between FortNOX and FlowVisor is that FortNOX is a single controller software, that executes parallel applications, while FlowVisor runs apart from controllers (usually on a different host). Both of these solutions restrict intrusted controllers/applications from introducing security threats. However, they rely on the assumption that the OpenFlow protocol and its communication channel are secure.

Another approach is to provide the security to Software Defined Networks using OpenFlow. NICE [8] Distributed Denial of Service protection infrastructure-as-a-service is a distributed vulnerability detection tool based on attack graph-based analytical models and reconfigurable countermeasures.

Moving Target Defense [9] in an OpenFlow environment is a mechanism to change internal hosts' IP addresses frequently and to mitigate attacks and reconnaissance from external network.

All the mentioned works are meant to provide the security to an OpenFlow-based SDN architecture, but their underlying assumption is that there are no network design vulnerabilities with the protocol. A comprehensive OpenFlow vulnerability assessment was presented in [10]. The list of vulnerabilities contain the lack of TLS adoption, flow enforcement, denial of service risk and controller vulnerabilities. The first three of them can be fully mitigated (or significantly weakened at least) by using authentication and access control mechanisms, which are examined in this paper.

III. OpenFlow secure channel problems

The OpenFlow specification in version 1.0 contains the requirement about the use of TLS [11]. However, the next version changed this requirement from 'must' to 'should'. This is also the case of the current version (1.4). There is a noticeable lack of support for TLS in current SDN switches and controllers. Table I shows TLS support offered by OpenFlow equipment vendors [10][12].

TABLE I. TLS Support in OpenFlow by vendors.

VENDOR	*TLS Support*
HP switch	No
Brocade switch	Controller port only
Dell switch	No
NEC switch	Partial
Indigo switch	No
Pica8 switch	Only new versions
Open vSwitch	Yes
NOX controller	No
Brocade Vyatta controller	Yes
POX controller	No
Beacon controller	No
Floodlight controller	No
MuL controller	No
FlowVisor	No
Big Network controller	Yes
Open Source controllers (f.e. Ryu, OpenDaylight)	Yes

The usage of Transport Layer Security has also its impact on the preparation and maintenance of Software Defined Networks based on the OpenFlow protocol. In particular, this includes generating controller and switch certificates, signing certificates with private keys, installing correct keys and certificates on devices. Assuming a topology distributed in different locations, though not connected to the Internet, technicians must prepare all components of Public Key Infrastructure and provide its security.

While the lack of TLS support is feasible in secured networks (such as data centers) where the access to physical devices is difficult, it becomes a serious security vulnerability in architectures similar to campus-style or branch-offices deployments, in which access to the network is less restricted. In the role of the management protocol in the "security-as-a-service", it can be the case when OpenFlow is transmitted by an untrusted ISP (e.g., the client is in the country interested in intercepting transmission). In such scenarios we cannot neglect the possibility of an attacker placing a device on communication path between the switch and the controller, or simply copy the flow to his/her machine (Figure 1). He/she is able then to get the configuration, insert or delete rules to modify/record sensitive data flow (configured using the OpenFlow device). Additionally, an attack can be performed without any observable differences from normal transmission, i.e., the attacker acts as a transparent proxy (compare this with the FlowVisor proxy modification of the rules through the process of forwarding messages). This type of attacks (so called 'man-in-the-middle' attacks) were very popular and successful before, [13][14]. The Software Defined Networking may decrease the difficulty of a full exploitation and may allow the attacker to automate the process. In the SDN, the man-in-the-middle attacks are arguably worse than in non-software defined networks, due to the lack of necessity to the sniff traffic to obtain plain-text credentials or the possibility to reconfigure all groups of devices in a single attack.

It is important to note that assuring a secure communication between the switch and controller is not enough – we should also authenticate all devices connected to the controller or switch. For instance, for the following two reasons the authentication and access control may be needed:

- to limit the possibility of adding a bogus device (e.g.,

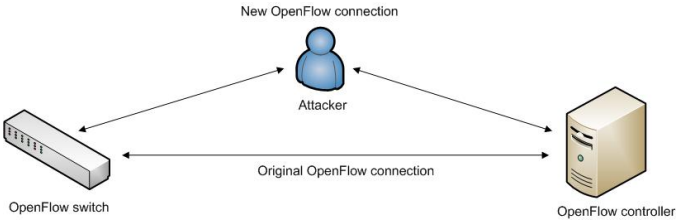


Figure 1. Man-in-the-middle attack scenario.

adding a new switch or controller to the OpenFlow domain);

- to group devices in federation (compare e.g., the Ofelia project, [15]), when different groups should be configured with different rules.

With full TLS implementation and authentication mechanism we can assure safety of the hosts and messages in transmission. We are unable, however, to detect switches that operate erroneously on rules in databases. The only solution to maintain the same view of traffic that flows through the network on the controller and switch seems to be dumping regularly and inspecting the flow tables on all hosts in the OpenFlow architecture. A potential solution described in [10] is based on generating keep-alive messages with checksums of flow tables that switches sent to controller.

IV. Secure channel creation and authentication

In this section, we describe three possibilities of creating the secure channel for communication between switches and controllers using OpenFlow. We also investigate how to further secure the OpenFlow architecture with an authentication mechanism that reduces the possibility of spoofing a device with a rogue switch or controller instead of any mitigation of eavesdropping.

A. Transport Layer Security

Deploying Transport Layer Security as an authentication and access protocol in the OpenFlow-based architectures was done by vendors using the public key infrastructure (PKI). The Web of trust architecture was neglected due to the small amount of devices in the authentication domain which could result in the vulnerabilities described in [16].

The proposed solution is an example of the peer encryption. The number of exchanges of the keys in such systems is proportional to the square of the number of users of the system (precisely, proportional to $\frac{n(n-1)}{2}$, where n is the number of users of the system). As already discussed, the reliability of the key distribution is essential for the credibility of the system. The solution to the number of the needed exchanges of keys and the need to ensure their authenticity was an application of the principle of the implied trust.

In cryptographic systems, there might be available some certification institutions, to which we have trust. This trust is supported by, for example, their protection level, their regular auditing, etc. If all users of the system have trust in the certification center, it is assumed that they also have trust in each other. In practice, this means that a user should know the address of the certification center and have its public key. In this way the user can reliably verify the authenticity of the received documents or exchange encrypted information with a reliable key from the communication partners, published by the certification center. In the PKI infrastructure, the trusted certification center is known as the Certificate Authority (CA).

To provide an opportunity to exchange keys between different systems, the cryptographic standard X.509 has been introduced also for certification. A certificate contains not only the owner of the public key signed by a certification authority, but also the information about the owner and other fields predicted by the standard X.509. Not only must CA be able to issue the certificate, but also to public and cancel the certificate, for example due to a theft or other security breach.

Equally important, as the implementation of the infrastructure, is developing and implementing the procedures for handling of the certificates. These procedures are called Certificate Practice Statement and include, among others:

- rules for issuing, verification and control the distribution of certificates;
- rules for cancellation of certificates;
- ability to recover private keys;
- methods for securing the infrastructure.

Especially in the case of PKI systems, the procedures are pillar of security.

In order to implement a TLS based secure channel and authentication mechanism, the administrator must build all the infrastructure to support it, i.e. the PKI (Figure 2). It will consist of elements such as:

- Each user of the system (in the case of an OpenFlow architecture, each switch and controller) posses key pair of private and public, stored in a secure space.
- Users of the system adopt a standard for certificate to the exchange keys.
- The existence of the certifying authority (internal or external, when developing OpenFlow in security-as-a-service manner) providing the issuance, invalidation and publication of certificates.
- Implementation of procedures to ensure the safety of the system.

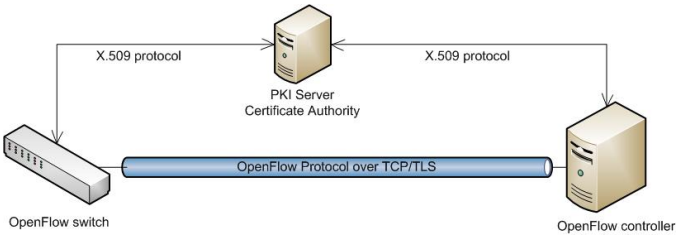


Figure 2. Architecture of the secured channel based on TLS protocol.

A simple connection between the switch and the controller illustrating the handshake with authentication by means of the Transport Layer Security mechanism consists of the following steps (by C we denote the controller; by S – the OpenFlow switch):

- S to C: sends ClientHello

The switch sends to the controller a message containing, inter alia, supported version of TLS protocol, supported methods of data encryption and compression, and session ID. This message also contains a random number which later will be used for key generation.

- C to S: sends ServerHello
  The controller responds with a similar message, in which it returns to the switch the selected parameters of the connection: TLS protocol version, supported types of encryption and compression, and a random number.
- C to S: sends Certificate
  The controller sends its certificate allowing the switch to verify its identity.
- C to S: sends ServerKeyExchange
  The controller sends its public key. The type and length of the key is determined by the type of algorithm in the previously sent message.
- C to S: sends ServerHelloDone
  The controller notifies that the switch can move to the next phase of the secure channel creation setup.
- S to C: sends ClientKeyExchange
  The switch sends to the controller an initial session key encrypted with the public key of the controller. Using the previous messages, the two random numbers (one for switch and another for controller) as well as pre-determined by the switch session key, both sides generate a session key used for the actual data exchange. The key is generated using a symmetric algorithm (typically DES). However, it is set in a safe way and known only for communicating parties.
- S to C: sends ChangeCipherSpec
  The switch informs the controller that it can switch to encrypted communication.
- S to C: sends Finished
  The switch sends this messages to report readiness to receive encrypted messages.
- C to S: sends ChangeCipherSpec
  The controller notifies that it obeyed the request - from now on the controller will only send encrypted information.
- C to S: sends Finished
  The message is sent over the secure channel to check reliability of the used mechanism.

As shown in the steps in the previous section, the default TLS mechanism provides only server authentication, resulting in authenticating the OpenFlow controller and leaving the possibility of spoofing the OpenFlow switches (grabbing configuration or modifying rules and sending them not so secured switch if possible). However, there are methods to authenticate the client switch. For this purpose, three additional messages can be used:

- C to S: sends CertificateRequest
  After submitting controller's certificate server notifies the client that it would like to receive a certificate from the OpenFlow switch
- S to C: sends Certificate
  After receiving the message ServerHelloDone the switch sends its certificate
- S to C: sends CertificateVerify
  The switch must confirm that it actually has the private key corresponding to the transmitted certificate. To prove this, the switch signs with its private key digest of all previously established connection parameters and sends it using this message.

B. Secure Shell

Another approach to encryption and authentication can be using the automatically generated keys and trust-of-first-use method as in the Secure Shell protocol. All switches are treated as SSH clients and the controller is treated as their server. The client connects to the server, authenticates with the key. In the process, it authenticates also the server key. Then the secure tunnel is created between OpenFlow's communication port on the server side (the controller) and configured port on switch. The switch sends unencrypted traffic to a local port (it is assumed that insecure architecture starts when a packet is leaving the network card and the device itself is trusted), and then the traffic is transmitted to the right port on the controller using the encrypted and secure tunnel (Figure 3). In this way, we limit the possibility of eavesdropping messages directly to the locally bugging switch and controller. We also use the authentication against the possibility of spoofing devices.

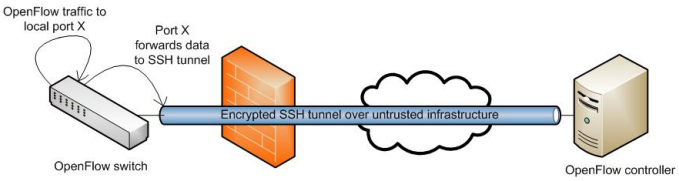


Figure 3. Secure SSH tunnel transporting the OpenFlow data.

We propose two similar authentication-related solutions, but with different level of security and configuration required. The first one is the usage of automatically generated keys and automatic acceptation of the connection by a client before the typical SSH authentication process. This will reduce the possibility of performing and attack on the transmission to the small window of the first communication between the switch and controller, but it does not require any additional configuration. Alternatively, the more secure (but requiring the involvement of the administrator) solution would be to verify the public key thumbprint of each device when connecting to the server for the first time. Additional overhead of the administrative work when adding a new switch to the OpenFlow topology will completely reduce the possibility of eavesdropping transmission of the OpenFlow messages tunneled via SSH.

C. IPSec

The last analyzed option of authentication and creation of a secure channel between the switch and controller over the OpenFlow communication is the IPSec protocol, in particular the host-to-host architecture. IPSec in the host-to-host configuration (which connects two hosts without the need of additional devices, see Figure 4), creates the secure channel between them and allows to authenticate each other, while the only needs are connections dedicated to the other side. The process can be summarized in the following five steps:

 ISBN: 978-1-61208-428-2

1) Configuring switches and controllers.
   In this phase the network administrator must prepare the information (IP addresses of both hosts, encryption key generator, pre-shared key used to initiate the connection and exchange generated keys during secure transmission) and then use them during configuration of each device. This configuration generally will be prepared once and then copied between hosts.
2) IKE phase one.
   The purpose of this phase is to authenticate hosts and set up the first secure channel for IKE exchanges. Following functions are performed: the authentication and protection of the hosts identities, the negotiation of the IKE policy to protect the exchange, the usage of Diffie-Hellman authenticated exchange to obtain matching shared keys, the setting up of the first secure channel for IKE phase two. Phase one has two operational modes: main and aggressive. The main mode consists of three exchanges (agreement of the algorithm and hashes used to secure communication, generation of secret keys components used to passing random numbers - nonces to prove identities, last exchange is the verification of the other side's identity). The aggressive mode has fewer exchanges (we obtain second IKE phase quicker) but some of the information must be exchanged before there is a secure channel available.
   Due to the security aspect of the work, we prefer the main mode in the planned implementations.
3) IKE phase two.
   In this step the IPSec tunnel is created. The following functions are performed: the negotiation of protected parameters, establishing of security associations, renegotiation IPSec SA to ensure security (additional authenticated Diffie-Hellman exchanges can be performed).
4) Data transfer.
   After the IKE phase two, the IPSec tunnel is created and OpenFlow packets can be encrypted and decrypted using the encryption mechanism specified in the configuration, resulting in authenticated secure channel transmission.
5) IPSec tunnel termination.
   After the successful transmission, the tunnel can be terminated by deletion or by timing-out.

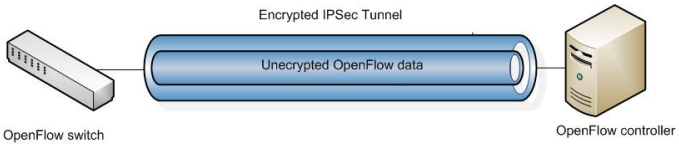


Figure 4. Secure IPSec tunnel transporting the OpenFlow data over an encrypted channel.

While TLS and SSH operate in the application layer, IPSec is a scheme operating in the network layer. Hence, only IPSec protocol protects any application traffic over an IP network. Therefore, an implementation for the OpenFlow protocol can be relatively easily migrated to another protocol or (due to evolving nature of OpenFlow) next specifications requirements.

V. Security Analysis

We analyzed benefits and drawbacks in the area of security of the proposed solutions. TLS, SSH and IPSec reduce or prevent the possibility of performing the following attacks:

- sniffing (lack of confidentiality),
- data modification,
- identity spoofing, password-based and application-layer attacks,
- man-in-the-middle attacks,
- denial-of-service attacks.

However, there are a few concerns regarding each of the solutions.

Transport Layer Security relies on secrecy of private keys and Certificate Authority trust, so assuring the security of these parts is the mission-critical aspect of maintaining this type of architecture. Some significant attacks against TLS include FREAK [17], BEAST [18] and CRIME and BREACH [19] attacks. However, assuming a proper implementation of TLS and its newest version, it is regarded as safe. Theoretically TLS can be compromised using SSL-Striping and SSL-Spitting attack, [20], but due to the newness of software-defined networking with TLS-secured OpenFlow, the attacks has not been confirmed and sufficiently researched yet.

As mentioned above, Secure Shell can be deployed using public and private keys or pre-shared key. Regardless of the option used, secrecy of keys (and CA if used) is the crucial security aspect. As regards possible attacks, using the revised version SSH-2 is assumed to be secure, however, theoretical vulnerability was discovered in [21] for the default encryption mode CBC. Therefore, we recommend the usage of CTR mode in the implementation. Another security concerns are that some institutions are able to decrypt the SSH traffic (see [22]). The details associated with such attacks were not released.

IPSec, similar to TLS and SSH relies on the secrecy of pre-shared key. However, the critical aspects are the randomness of the encryption key generator and security of the encryption and hash algorithm. As for this work, from available in IPSec implementation algoritms we assume sha1 as hash algorithm and 3DES, AES as encryption algorithms are secure. There are known attacks on IPSec when other than recommended solutions were applied [23]. Similar to SSH, there are allegations that some institutions have been working actively to insert vulnerabilities into IPSec implementations, [24].

VI. Conclusion and Future Work

In this paper, some possible solutions to the lack of authentication, access control and creation of the secure channel over the OpenFlow protocol were investigated. As argued, implementing TLS (as in the OpenFlow specification recommendations) does not address configuration problems for network operators. With the core idea of increased security of the OpenFlow transmission, a novel utilization of known Internet security systems was proposed.

By comparing TLS, SSH and IPSec, it was demonstrated that, in relation to the OpenFlow architecture usage, each of the proposed protocols has its own strengths (i.e. the ease of implementation in IPSec, or conformation with the specification in TLS) and weaknesses (possible attacks). The

 ISBN: 978-1-61208-428-2

paper showed that implementing any of the proposed solutions will result in increased security and reduction or prevention against numerous attacks on the OpenFlow protocol.

As for the future work, we are planning to utilize the PL-LAB2020 laboratory, [25], to implement each solution and analyze the security concerns mentioned in the article while verifying the performance of the protocols. The PL-LAB2020 laboratory, which is under construction now, will consist of six geographically dispersed nodes associated with leading Polish research and academic centers:

- National Institute of Telecommunication (NIT),
- Warsaw University of Technology (WUT),
- Poznan Supercomputing and Networking Center (PSNC),
- Silesian University of Technology (SUT),
- Gdansk University of Technology (GUT),
- Wroclaw University of Technology (WrUT),

connected via dedicated 2*10Gb/s fiber links. The nodes will be equipped with specialized devices for carrying out research in several directions, including the Software Defined Networking. In particular, there will be over 30 OpenFlow switches from at least 3 different vendors and a few OpenFlow controllers creating different technology domains distributed over 5 locations of PL-LAB2020 infrastructure (Figure 5). To validate the performance and analyze security of the studied solutions, PL-LAB2020 will also consist of several servers with Data Plane Development Kit and three network traffic generators and analyzers with 10Gb/s interfaces, placed in different locations.

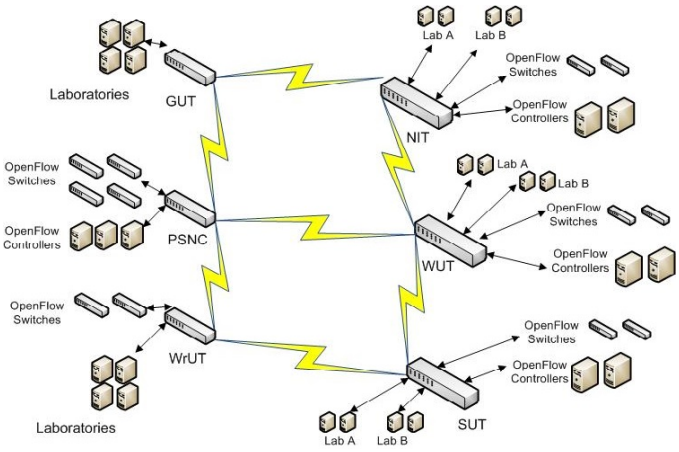


Figure 5. Architecture of PL-LAB2020.

ACKNOWLEDGMENT

This work has been supported by the National Centre for Research and Development under the European Regional Development Fund, Grant No. POIG.02.03.01-00-104/13, PL-LAB2020 project.

REFERENCES

[1] N. McKeown and et al., "Openflow: enabling innovation in campus networks," ACM SIGCOMM Computer Communication Review, vol. 38, no. 2, 2008, pp. 69–74.

[2] "Openflow switch specification: Version 1.4.0," URL: https://www.opennetworking.org/images/stories/downloads/specification/openflow-spec-v1.4.0.pdf [accessed: 2015-02-28].

[3] T. Dierks, "The transport layer security (tls) protocol version 1.2," 2008, pp. 4–68.

[4] T. Ylonen and C. Lonvick, "The secure shell (ssh) protocol architecture," 2006, pp. 4–26.

[5] N. Doraswamy and D. Harkins, IPSec: the new security standard for the Internet, intranets, and virtual private networks. Prentice Hall Professional, 2003.

[6] R. Sherwood and et al., "Flowvisor: A network virtualization layer," OpenFlow Switch Consortium, Tech. Rep, 2009, pp. 1–13.

[7] P. Porras, S. Shin, V. Yegneswaran, M. Fong, M. Tyson, and G. Gu, "A security enforcement kernel for openflow networks," in Proceedings of the first workshop on Hot topics in software defined networks. New York: ACM, 2012, pp. 121–126.

[8] C.-J. Chung, P. Khatkar, T. Xing, J. Lee, and D. Huang, "Nice: Network intrusion detection and countermeasure selection in virtual network systems," IEEE transactions on dependable and secure computing, no. 4, 2013, pp. 198–211.

[9] J. H. Jafarian, E. Al-Shaer, and Q. Duan, "Openflow random host mutation: transparent moving target defense using software defined networking," in Proceedings of the first workshop on Hot topics in software defined networks. New York: ACM, 2012, pp. 127–132.

[10] K. Benton, L. J. Camp, and C. Small, "Openflow vulnerability assessment," in Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking. New York: ACM, 2013, pp. 151–152.

[11] "Openflow switch specification: Version 1.0.0," URL: https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-spec-v1.0.0.pdf [accessed: 2015-02-28].

[12] Y. Patil, "Vulnerability analysis of openflow control channel," 2014, pp. 1–2.

[13] F. Callegati, W. Cerroni, and M. Ramilli, "Man-in-the-middle attack to the https protocol," IEEE Security and Privacy, vol. 7, no. 1, 2009, pp. 78–81.

[14] K. Ouafi, R. Overbeck, and S. Vaudenay, "On the security of hb# against a man-in-the-middle attack," in Advances in Cryptology-ASIACRYPT 2008. Springer, 2008, pp. 108–124.

[15] "OFELIA PROJECT homepage," URL: http://www.fp7-ofelia.eu/ [accessed: 2015-02-28].

[16] N. Ferguson and B. Schneier, Practical cryptography. Wiley New York, 2003, vol. 141.

[17] B. Beurdouche and et al., "A messy state of the union: Taming the composite state machines of tls," in IEEE Symposium on Security and Privacy. IEEE, San Jose, 2015, pp. 1–16.

[18] T. Duong and J. Rizzo, "Here come the ninjas," Unpublished manuscript, 2011, p. 4.

[19] D. Goodin, "Crack in internets foundation of trust allows https session hijacking," Ars Technica, 2012, pp. 1–2.

[20] M. Marlinspike, "New tricks for defeating ssl in practice," BlackHat DC, February, 2009, pp. 1–114.

[21] "SSH Vuln Note," URL: http://www.kb.cert.org/vuls/id/958563 [accessed: 2015-02-28].

[22] "Prying Eyes: Inside the NSA's War on Internet Security," URL: http://www.spiegel.de/international/germany/inside-the-nsa-s-war-on-internet-security-a-1010361.html [accessed: 2015-02-28].

[23] J. P. Degabriele and K. G. Paterson, "Attacking the ipsec standards in encryption-only configurations." in IEEE Symposium on Security and Privacy, vol. 161, Oakland, 2007, pp. 335–349.

[24] "Secret Documents Reveal N.S.A. Campaign Against Encryption," URL: http://www.nytimes.com/interactive/2013/09/05/us/documents-reveal-nsa-campaign-against-encryption.html [accessed: 2015-02-28].

[25] URL: http://http://www.pllab.pl/ [accessed: 2015-02-28].

 ISBN: 978-1-61208-428-2

Cloud Assisted Live Video Streaming over DHT Overlay Network

Pheng Un Lim and Hwangkyu Choi
Department of Computer Science and Engineering
Kangwon National University
Chuncheon, Korea
e-mail: {limphengun, hkchoi}@kangwon.ac.kr

***Abstract*—Many researches have been proposed to solve the scalability, the availability and the low-latency problems of peer-to-peer live video streaming; however, the problems are still not yet completely solved. In this paper, we propose a cloud assisted live video streaming over the distributed hash table (DHT) overlay network to tackle these problems. Our system is based on a scalable DHT network, which structures into a mesh-based overlay, to efficiently share the video stream. Moreover, cloud assist is introduced in order to maintain the availability, the low-latency and the scalability of the system. In this work, cloud server is also part of the DHT network. The role of the cloud server is to assist other nodes when the number of nodes which received video segment is less than the specific threshold. The cloud servers directly receive video segments from video server so the quality and the availability of the video segments can be guaranteed. The evaluation results show that with the help of cloud assist, the availability and the low-latency of the video segments are greatly improved if cloud has enough bandwidth.**

Keywords-video live streaming; cloud computing; peer-to-peer; DHT overlay network.

I. INTRODUCTION

Because of scalability, availability and low-latency problems, the current live video streaming still cannot meet the user demand. Many studies have been tried to solve these problems. One of the most recent researches which tried to solve these problems is peer-to-peer (P2P) live video streaming over the DHT network. Yet, the availability and the low-latency problems are still the main obstacles of this system because it fully depends on the user device's capacity. In P2P live video streaming if the user device's capacity is poor, other clients will not get the desirable quality of the video stream.

Peer-to-peer systems are distributed systems in which each node has equivalent functionality and without any centralized control or hierarchical organization [1]. The core operation of P2P system is the efficiency of locating each node in the network. The distributed hash table (DHT) is one of the most efficient methods which can be used to locate the node in mesh-based P2P network. DHT is a decentralized system that provides a lookup service similar to a hash table. The DHT node uses key to lookup for the value which associated with it. The advantages of DHT are: high scalability, reliability, and self-organizing [2]. However, only few works have been focused on using DHT with mesh-based mechanism. Thus, cloud assisted live video streaming over DHT overlay network is proposed in this paper.

In this system, the P2P network is based on a DHT using the available protocols such as Chord [1], Pastry [3], Kademlia [4], etc. A cloud server is considered as a node, the same as other users' devices. The only different is that cloud server has a big and stable capacity. The role of the cloud server is to assist other nodes when the numbers of available segments are less than the specific threshold.

The rest of this paper is organized as follow: Section 2 talks about the related works. The system architecture of the new proposed system, buffer map management, data push and pull among peers, and cloud assist scenario are discussed in Section 3. The system performance evaluation is discussed in Section 4. Finally, the conclusion is presented Section 5.

II. RELATED WORKS

There are many existing researches have been studying on P2P video live streaming; however, most of those works focused on tree-based, mesh-based with gossip protocol and others, and hybrid structure such as [5]-[7]. Only few researches focused on using mesh-based with DHT [2] and cloud assist in P2P live video streaming [8]-[10].

[2] is one of the most recent research which built a P2P live video streaming over DHT network and using chord protocol. This work was proposed to solve the problems of scalability, availability, and low-latency in P2P live video streaming systems. The performance of this system is far better than that of mesh-based and tree-based system. However, it fully depends on user devices' capacity which is small and not stable (nodes join and leave the system any time) and also the video server can become a bottleneck when the number of users is suddenly increased. In order to maintain the video quality and the scalability of the system, and also to deal with the inconsistency of the clients' devices, cloud assist has been proposed in our paper.

[8] proposed a cloud-based service called "AngelCast". AngleCast is proposed to solve the limitation of the delivered video quality. The client nodes in AngleCast are arranged in tree-based structure. [9] proposed Cloud-

 ISBN: 978-1-61208-428-2

Assisted P2P Live Streaming (Clive) to tackle the bottlenecks in the available upload bandwidth at media source and inside the overlay network that can limit the quality of the server (QoS). Cloud-Assisted Live Media Streaming (CALMS) [10] books and adjusts adaptively cloud server resources in a fine granularity to accommodate temporal and spatial dynamics of demands from live streaming users.

As most of the proposed cloud assisted P2P live video streaming researches tend to solve the problems of tree-based, mesh-based or hybrid structure, only few works are focused on hybrid structure with DHT network. Thus, in this paper, a cloud assisted live video streaming system over DHT overlay network is proposed.

III. SYSTEM ARCHITECTURE

The cloud assisted video live streaming system is built over DHT overlay network, as shown in Figure 1. In this system, cloud servers join the DHT network as a node the same as other clients and receive video segments directly from the video server.

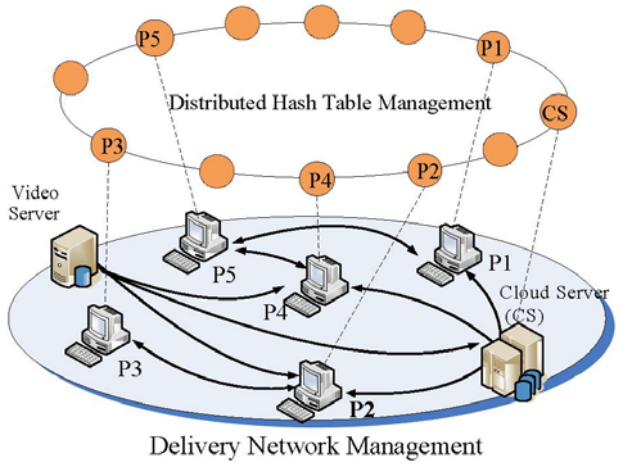


Figure 1. Cloud assisted live video streaming over DHT overlay network

A. *Buffer Map Management*

In P2P live video steaming system, the video stream is divided into segment or chunk which identified by the sequence numbers, name as segment ID [11]. Each node has a buffer to catch the video segment. A buffer map, which stores several video segments, as shown in Figure 2, is an abstract description of this buffer. Buffer map consists of a sequence of {0, 1} to indicate the buffered and empty states of the segment [12]. While playing the video stream, node can refer to its buffer map in order to know which segment is missing and which segment is already buffered. Thus, node can request the missing segment in advance. In this system, each node stores the buffer map information list of the segment which hash key's ID equals to or immediately succeeds the node's ID.

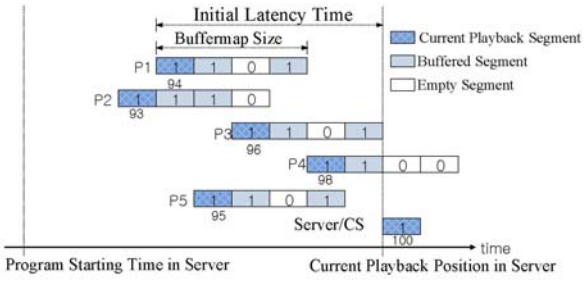


Figure 2. Live streaming scenario

One video segment can have more than one hash key because segment belongs to node in different location will generate different hash key. Hash key contains channel information, program information, segment number and location information, as shown in Figure 3.

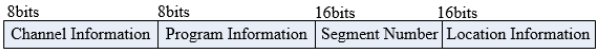


Figure 3. Hash key

The buffer map information consists of the IP address, starting segment and the current buffer map, as shown in Figure 4. The buffer map information list stores the buffer map information in the descending order of the starting segment number.

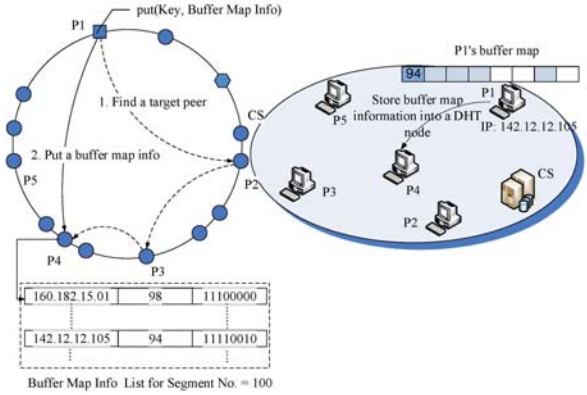


Figure 4. Storing buffer map info process

When node receives segment from server or from other node, this node has to register its buffer map information to the DHT network. The processes of registering its buffer map information are: first, hash key for this segment is created. This hash key is used to find the node to store the buffer map information. The same segment, which has the same hash key, will be stored in the same node. Then, the DHT method, *put* (key, buffer map info), is used to find the appropriate node to store the buffer map information, as shown in Figure 4.

B. *Data Delivery Management*

During playing the video stream, each node checks their buffer map information in order to know whether there is a

 ISBN: 978-1-61208-428-2

missing segment. If there is a missing segment, node uses the missing segment's ID, channel information, program information, and node's location information to create a hash key. Then, this hash key is used to find the node which stores the buffer map information list of this missing segment. After that, the requested node requests the buffer map information list from that node. The DHT method, *get* (key), is used to retrieve the buffer map information from specific node, as shown in Figure 5.

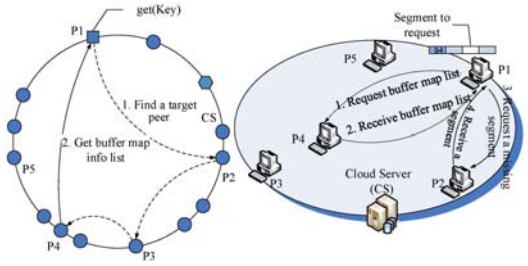


Figure 5. Request segment among peers

Finally, the requested node uses the received buffer map information list to find the node which has the missing segment and sends the request to that node.

C. *Cloud Server Assist Scenario*

The role of cloud server is to assist other nodes in the DHT network. Cloud server will assist other nodes in case the number of nodes which has the video segment is less than the threshold. The threshold is referred to the number of nodes which receive segment directly from server that can effectively share the video segment to other nodes in the network.

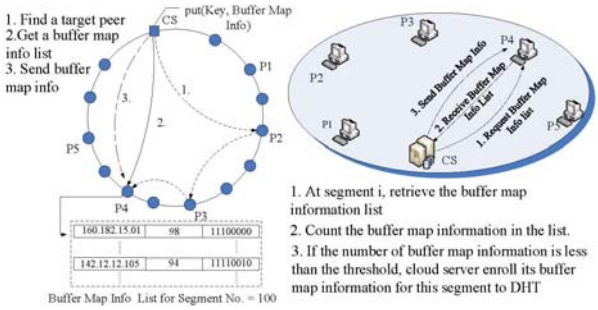


Figure 6. Cloud assist scenario

As illustrates in Figure 6, for each segment, cloud server requests the buffer map information list from the node which stores this list. Then, cloud server uses the buffer map information list to count the available segments. If the numbers of available segments is less than the threshold, cloud server generates hash key and sends its buffer map information to the node which stores the buffer map information list of this video segment. Thus, other nodes can request the video segment from cloud server. After receiving the segment, those nodes also register their buffer map to the network to share the segment to other nodes. Therefore, the cloud server will not be overloaded by many requests.

IV. PERFORMANCE EVALUATION

In order to evaluate the system performances, a simulation program has been developed. In our experiment, we set the upload bandwidth of the server and cloud to 40,000kb/s and upload and download bandwidth of nodes to 600kb/s. The video size is set to 120 and segment size is 300kb, that one segment can be played for 1 second. Thus, the video length is 2 minutes. The threshold is set to 16 percent of the total nodes. The buffer size is set to 10. After the upload bandwidth of the clouds and the nodes which have the video segment are full, the requested nodes are queued and will request the video segment again after the bandwidth is available. The number of nodes join and leave the network are set to the exponential distribution with mean value of 2 percent and 1 percent of total nodes every half second, respectively. The delay time of each segment is set to 3 seconds, means that if all nodes receive the segment during 3 seconds, the user can watch the video stream smoothly. We run each experiment five times and report the average as the final experimental results.

A. *Latency*

As illustrates in Figure 7, the average delay time of the segments without the help of cloud assist exceed 3 seconds if compare to when using cloud assist, which only exceed 3 seconds, when the number of nodes reach 3000.

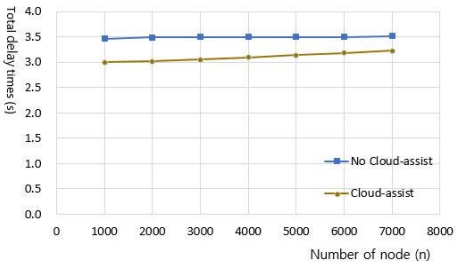


Figure 7. Delay time vs. number of nodes

As mention in the experiment setup, the threshold is set to 16 percent of total nodes. Thus, even when the number of nodes reach 8000, the video server still have enough bandwidth to send the video segment to threshold nodes (the total bandwidth which is required for sending the video segment to the threshold nodes is only 37,500KB). However, because nodes dynamically leave and join the network, the delay time of each segment cannot be maintained. For cloud-assist method, the average delay time of each segment exceed 3 seconds when number of nodes reach 3,000 because cloud server does not have enough bandwidth.

 ISBN: 978-1-61208-428-2

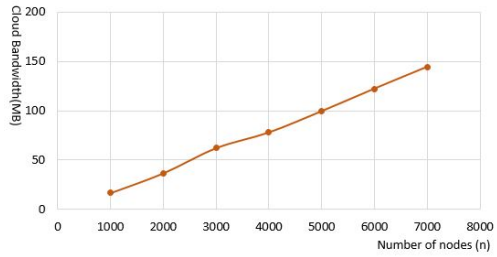


Figure 8. Cloud bandwidth vs. number of nodes

Figure 8 shows the amount of cloud bandwidth which is needed for maintaining the delay time of the network, i.e., if this amount of cloud bandwidth is used for each method, the delay time will not exceed 3 seconds.

B. Availability

Figure 9 shows the number of remained nodes when the number of nodes in the network is increased. The remained nodes refer to the nodes which do not receive the video segment after the delay time (the delay time is set to 3 seconds in this experiment) is passed. The number of remained nodes of no-cloud method is worse than cloud-assist method.

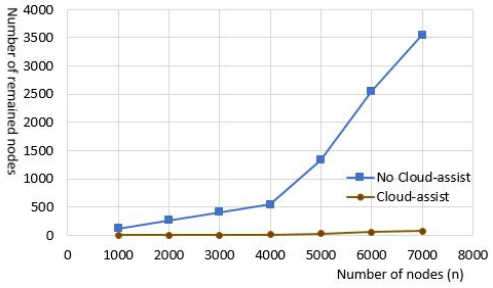


Figure 9. Number of remained nodes vs. number of nodes

Even the delay time, as shown in figure 7, of both methods has a little different, the number of remained methods of no-cloud assist is far bigger than that of cloud-assist methods. For instance, when the number of nodes reaches 3000, the delay time of no-cloud method is 3.49 and the delay time of cloud-assist method is around 3.05. However, according to figure 9, for no-cloud assist, after the delay time 3 (after 3 second) passed, the number of remained nodes is 413 compare to the remained nodes of cloud-assist methods is only around 5.

V. CONCLUSION AND FUTURE WORK

In this paper, cloud assisted live video streaming over DHT overlay network has been proposed. The DHT network can be used to solve the problem of scalability. Moreover, because cloud server is consistent, has big capacity, and receives video segments directly from video server, the availability of the video segments and the low-latency can be enhanced. The performance evaluation results show that with the present of cloud assist, the performances of the system are greatly improved in both availability and low-latency. We only build a simulation program to compare the performance of the system when there is no cloud assist and when there is a cloud assist. In the future, a full function DHT simulation process should be conducted. In addition, the threshold value and the amount of cloud capacity which is required for assisting the failure nodes also need to be studied in details in order to maximize the system performance and minimize the cost.

REFERENCES

[1] I. Stoica, et al., “Chord: A scalable peer-to-peer lookup protocol for internet applications,” IEEE/ACM Trans. Networking, vol. 11, no. 1, 2003, pp. 17-32.

[2] H. Shen, Z. Li, and J. Li, “A DHT-aided chunk-driven overlay for scalable and efficient peer-to-peer live streaming, IEEE Trans. on Parallel and Distributed Systems, vol. 24, no.11, 2013, pp. 2125-2137.

[3] A.I.T., Rowstron, P. Druschel, “Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems,” Proc. of the 18th IFIP/ACM International Conference on Distributed Systems Platforms, 2001, pp. 329-350.

[4] P. Maymounkov, D. Mazieres, “Kademlia: a peer-to-peer information system based on the XOR metric,” 1st International Workshop on Peer-to-peer Systems, 2002.

[5] F. Picconi and L. Massoulie, “Is there a future for mesh-based live video dtreaming?” Proc. 8th Int’l Conf. Peer-to-Peer Computing (P2P), 2008, pp. 289-298.

[6] J. Venkataraman and P. Francis, “Chunky spread: Multi-tree unstructured peer-to-peer multicast,” Proc. 5th Int’l Workshop Peer-to-Peer Systems (IPTPS), 2006, pp. 1-10.

[7] F. Wang, Y. Xiong, and J. Liu, “mTreebone: A hybrid tree/mesh overlay for application-layer live video multicast,” Proc. 27th Int’l Conf. Distributed Computing Systems (ICDCS), 2007.

[8] R. Sweha, V. Ishakian, and A. Bestavros, “AngelCast: Cloud-based peer-assisted live streaming using optimized multi-tree construction, Proc. 3rd Multimedia Systems Conference (MMSYS), 2012, pp. 191-202.

[9] A. H. Payberah, H. Kavalionak, V. Kumaresan, A. Montresor, and S. Haridi, “CLive: Cloud-assisted P2P live streaming,” Proc. 12th IEEE Int’l Conf. Peer-to-Peer Computing (P2P), 2012, pp. 79-90.

[10] F. Wang, J. Liu, and M. Chen, “CALMS: Cloud-assisted live media streaming for globalized demands with time/region diversities,” IEEE Infocom, 2012, pp. 199-207.

[11] Y. Zhou, D. M. Chiu, and J. C. Lui, “A simple model for analyzing P2P streaming protocols,” IEEE Int’l Conf. Network Protocols (ICNP), 2007, pp. 226-235.

[12] S. Lan, S. Zhang, X. Zhang, and Z. Guo, “Dynamic asynchronous buffer management to improve data continuity in P2P live streaming, IEEE 3rd Int’l Conf. Computer Research and Development (ICCRD), 2011, pp. 65-69.

 ISBN: 978-1-61208-428-2

An Intelligence System Based on Social Web Mining and Its Application in Health Care in Hong Kong

Kin Keung Lai[1,2]
[1] Management Sciences
College of Business, City University of Hong Kong, Hong Kong.
[2] International Business School
Shaanxi Normal University
Xian, China
mskklai@cityu.edu.hk

Juan Shi[1,2]
[1] Management Sciences
College of Business, City University of Hong Kong, Hong Kong.
[2] Management Science and Engineering
Xi'an Jiaotong University
Xian, China
juanshi4-c@my.cityu.edu.hks

Gang Chen
The 705th Research Institute
China Shipbuilding Industry Corporation
Xi'an, China
james.gang.chen@gmail.com

***Abstract*—In China, two systems—Chinese Information system for disease control and prevention (CDC) and the National Adverse Drug Reaction (ADR) Monitoring System--have been established for monitoring infectious diseases and for reporting ADR events. However, surveys have shown that nontrivial problems have affected the performance of these information systems adversely. Specifically, the absence of reasonable work and supervision systems, namely the systemic factors constitute the majority of the reasons for underperformance of these information systems. Also, not just some leaders but also quite a few doctors do not give sufficient importance to reports of infectious cases or ADR events, which is thought to be a waste of manpower and financial resources. Nowadays social media websites have attracted wide interest from both the academia and the industry, because of its very large user base, updating in real time, abundant incoming data and the fact that it offers prolific information about consumers' behavioral characteristics. In this proposal, we plan to address the aforementioned issues by implementing an Intelligence System based on Social Web Mining (ISSWM). ISSWM can help detect ADR events, assess the trends of infectious diseases, probe people's opinions towards some medical institutions or on healthcare related issues in real-time, at a cost that is extremely low compared with the existing systems.**

Keywords-big data; Web 2.0; social web mining; opinion mining; sentiment analysis; epidemics surveillance.

I. INTRODUCTION

The Chinese Information system for disease control and prevention (CDC) was established in China for control and prevention of infectious diseases in 2005. Medical and health institutions at the grass roots level are required to report cases of infectious diseases to this system, which helps the concerned department(s) to keep abreast of the situation and make timely decisions. Besides that, the National Adverse Drug Reaction (ADR) Monitoring System has also been launched to improve surveillance of ADR events. However, there are some nontrivial problems in these information reporting systems. For example, a survey, which investigates the reports of infectious cases via CDC in 17 medical institutions, has found some salient reasons that seem to be responsible for underperformance of CDC. First, reports using CDC have not been given sufficient importance by staff of these medical institutions. Second, there is a lack of effective work and monitoring mechanisms. Third, medical staff involved in reporting via CDC change frequently and lack adequate knowledge about infectious diseases [1]. As a result, missing reports rate is estimated at 11.76% and the rate of consistency between infectious diseases report cards and information input into the system is only 78.25%. Another survey about ADR has reached similar conclusions: both information reliability and timeliness leave much to be improved. From the above, it can be concluded that management factor and human factor are two main reasons of unsatisfactory performance of these information reporting systems. There is a long way to go and a lot of efforts are needed to address the problems of these systems.

Fortunately, we have entered the era of Web 2.0, which is characterized by high degree of public engagement and initiative. In China, the number of registered users of Sina Weibo, the biggest social platform, has reached 600 million and there are as many as 61.4 million daily active users [2]. It is worth noting that medical and health related data, generated on social media, constitute an important part of Big Data relevant to the medical area. Generated by users directly, they are constantly updated and reflect every aspect of people's lives [3]-[4]. More importantly, data on the social media are much easier to access via API compared with databases owned by the government. The rapid growth of online social networking sites and the public availability of data have made analysis of social media data easy and convenient [5]-[6].

This proposal envisages implementation of an intelligence system based on social web mining (ISSWM), which extracts insightful information from the unstructured or semi-structured data available on social media to answer questions related to people's main concerns in healthcare area at a certain moment. How is an epidemic spreading in

 ISBN: 978-1-61208-428-2

the country or within a certain area? What are people's specific expectations from some particular medical institutions? However, it needs to be pointed out that the purpose of ISSWM is not to substitute any existing authoritative information-gathering websites, such as CDC or the ADR monitoring system. Instead, ISSWM functions as an intelligence system complementary to the existing information systems. For example, it could serve as a plug-in for the existing systems. The advantages of ISSWM over existing information reporting systems, like CDC and ADR monitoring system, are as follows:

1. Timeliness and low cost. ISSWM can automatically collect, process and analyze data from social media websites in real time, using computer algorithms. These features of the proposed ISSWM have several implications. First, manual effort is greatly reduced and the process virtually costs nothing. Second, real-time data collection and analysis helps decision-making department(s) stay up to date with the latest situation. Third, ISSWM can effectively mitigate problems resulting from lack of reasonable regulation in CDC or the national ADR monitoring system and problems which arise from medical staff's lack of awareness regarding the importance of reporting infectious cases.

2. Tremendous user base generated by integrating users from multiple social networking websites. Since data can be collected from numerous users of multiple social media websites, ISSWM can efficiently extract information related to infectious diseases, ADR events and other related issues in healthcare sector. The very large number of users provides the much needed credibility to results of ISSWM.

3. Use of big data. During events of public interest, such as presidential debates, there are hundreds of thousands of tweets per minute [7]. Interactions on websites, such as Twitter and Sino Weibo, contain a huge amount of meaningful information and this presents some challenge to ISSWM, which is taken into account in design.

The rest of this paper is organized as follows. In Section 2, we discuss prior work and their limitations on exploiting social media data in healthcare area. Section 3 presents the framework of ISSWM. Section 4 concludes the paper.

II. LITERATURE REVIEW

Recently, a considerable amount of work has been dedicated to exploiting data on Twitter, most of which has focused on forecasting epidemics based on interactions on Twitter [8]-[12]. For example, the public sentiment with respect to H1N1 was tracked and actual disease activity was measured on the basis of information embedded in the Twitter stream [9]. Furthermore, Twitter data, such as friendship relationships [3], the proximity of users' geographical locations and users' mobility, are used to estimate the likelihood of whether a specific person is likely to get infected [10]-[11]. Influenza-related messages are identified by leveraging a document classifier and these messages achieve a correlation of 0.78 with the Centers for Disease Control and Prevention (CDC) statistics [12].

A recently emerging research area is to explore online reviews to understand patients' attitudes towards healthcare services. For example, it has been revealed that there is more than 80% agreement between patients' own quantitative ratings of care and those derived using sentiment analysis, from online free-text comments on different aspects of healthcare [13]. Reference [14] analyzed a corpus consisting of nearly 60,000 reviews with a probabilistic model of text. The output of the model was found to significantly correlate with state-level measures of healthcare.

After reviewing these literature, we find out that there are some limitations in the existing research. First, they focus on data collected from only one social media website, Twitter, as reported in [8]-[12], English National Health Service website in [13], and RageMDs in [14]. Due to a single source of data, it is doubtful whether the findings still hold in the context of the whole society; Twitter users are not necessarily representative of the wider population [15]. Second, existing research concentrates on solving one or two specific problems, such as detecting flu trend or tracking people's sentiment regarding H1N1. An integrated system is required to provide comprehensive knowledge about the current status and sentiment related to healthcare issues in the society. Third, most of prior studies have adopted a methodology called post-collection analysis [8, 10-12]. That is to say, the analysis depends on pre-retrieved tweets or a well-prepared corpus [13-14], and there is a time delay that cannot be negligible in their analysis results. Thus a real-time system is in required to ensure online analytical processing in a streaming fashion. In this way, the breakout of emergency events can be detected almost instaneously and the latest situations are reported in real time, given people's high engagement in social networking websites. Last but not the least, China has a much larger population compared with any other country, producing an unprecedented amount of social media data. However, healthcare related data on Chinese social networking websites are virtually untapped.

To bridge the abovementioned gaps, we propose an intelligence system based on social web mining (ISSWM). It collects data from multiple social networking websites, operates incessantly (24 hours*7 days) and delivers valuable results in real time. The structure of ISSWM is illustrated in Figure 1. Extensive explanation about ISSWM is presented in Section 3.

III. INTRODUCTION ABOUT ISSWM

The proposed ISSWM shall employ crawling techniques and a variety of machine learning algorithms, such as clustering, feature selection, k-nearest neighbors, regression and so on, to obtain insightful knowledge from social media websites. ISSWM mainly comprises four modules shown in Figure 1.

1. Data collecting and preprocessing module

As mentioned above, ISSWM incessantly collects data from several Chinese social networking websites, such as Sina Weibo [16], Tencent Weibo [17], Renren Net [18] and so on. Crawling technologies are also used as a complementary method to retrieve data, as there are a few limitations on obtaining tweets through API for some social networking websites [19]. Besides that, API access is not available for non-social networking websites [20]-[25]. It can

 ISBN: 978-1-61208-428-2

be expected that the raw tweets contain a lot of noisy data such as wrong spellings, punctuation, images, or tweets that have nothing to do with healthcare sector.

It is proposed to preprocess the raw data in the following way. First, regular expressions are used to delete numbers, punctuations, and non-Chinese characters. Second, stopwords, which are meaningless words, are filtered out. Third, we crawl healthcare websites, such as 120ask.com [20], mingpaohealth.com [22], etc., in order to get a corpus of healthcare related words. Inversed Term Frequency (IDF) [26] is used to evaluate how important a given word is. Words with higher IDF value are thought to be more related with healthcare sector. To further improve the accuracy of this corpus, a priori knowledge is used to filter out words which get higher IDF values but seldom appear in healthcare sector. After the healthcare corpus is established, raw tweets obtained from social networking websites are examined using this corpus. A tweet that does not share any word with this corpus is treated as non-relevant and is screened out.

Clean tweets are integrated, sorted and converted into a unified format, as illustrated in Table 1 and Table 2. Consistency needs to be checked as the tweets come from different data sources. Table 1 illustrates profile information about each user and Table 2 shows the format of information about each tweet. Relationship data such as each user's followers and friends are also kept in this step.

TABLE I. INFORMATION ABOUT USERS

Field	Description
id	ID of the user for whom to return results for table
screen_name	The screen name of the user for whom to return results for
statuses_count	The number of tweets issued by the user
description	Description about the user himself/herself
friends_count	The number of users this account is following
followers_count	The number of followers this account currently has
location	The user-defined location for this account's profile
time_zone	The time zone this user declares themselves within
…	…

TABLE II. INFORMATION ABOUT TWEET

Field	Description
id	The integer representation of the unique identifier for this Tweet
text	The actual text of the status update
created_at	Time when this Tweet was created
favorite_count	Indicates approximately how many times this Tweet has been "favorited" by Twitter users
retweet_count	Number of times this Tweet has been retweeted
place	Indicates that the tweet is associated a place
user	The user who posted this Tweet

2. Topic detection and classification module

We know that when the topic changes, keywords change correspondingly, since frequently used terms differ greatly across different areas. Latent dirichlet allocation (LDA) algorithm, which can represent each tweet as a random mixture of latent topics [27], is employed in this module to detect hot topics in social networking websites.

Suppose we would like to know what are the Top N topics attracting people's attention at a given point of time. First, we will retrieve tweets in the past month. After the preprocessing procedure (Module 1), we obtain clean tweets that focus on issues related to healthcare area, all of which are used to develop a LDA model, with the parameter "number of topics" set as N. Different weights are assigned on different topics that a given tweet can possibly be related to. The topic category on which a given tweet gets the highest weight is used to classify this tweet. In this way, all clean tweets are divided into different topic categories. Next, the N topics are sorted in terms of the number of tweets which belong to a particular topic. Finally, we get the Top N hot topics in healthcare area. The topic with a higher rank attracts more attention on social networking websites.

When a new unseen tweet arrives, it is transformed by using the established LDA model and is assigned to the topic category on which this tweet scores the highest. In this way, we identify which topic is acquiring more and more attention and, consequently, the trend of each of the Top N topics. Combing users' profile information and auxiliary information of a tweet, we can gain an even deeper understanding about the hot topics. For example, by exploiting the geographic information, we get to know the particular location where a specific topic induces a strong response among the citizens. With users' profile information, we may understand which group of people are involved in a particular topic. From time to time, we can check the content of the top N topics and get to know the change in the hot topics.

Like other hot topics, infectious diseases have their own key words, which can be obtained by a priori knowledge or learned by tweets on the internet. With these keywords, relevant tweets are extracted from multiple social networking websites and used to estimate the trend of some infectious diseases. In order to detect ADR events, a specialized corpus of medical-related words needs to be developed at the very beginning and any tweet which contains a relevant medical word or the name of a medicine in this corpus should be given special attention.

3. Module of opinion mining and sentiment analysis

After preprocessing (Module 1) and topic detection (Module 2), tweets are assigned to different topic classes. In order to grasp people's opinions or sentiments on some given topic, we use features like opinion words or phrases, negations, emoticons, syntactic dependency and so on to judge the sentiment underlying a tweet. Tweets which do not have these features are treated as neutral tweets and are discarded. Support vector machine (SVM) algorithm is employed to classify opinions into negative or positive and regression is used to identify the intensities of sentiments. As

 ISBN: 978-1-61208-428-2

both are supervised algorithm, labelled data are required to develop the model. We produce labelled data by labelling the polarities and intensities of the training tweets manually.

4. Opinion predicting module

Apart from classification of sentiments into positive or negative, we are interested in predicting people's opinions on new events. This problem has two dimensions. First, we want to forecast people's opinions on an upcoming event. Second, an event has occurred, but not everyone has publicly expressed his or her opinion toward this event. Therefore we would like to estimate the opinions of this silent group. We claim that people with similar characteristics tend to hold similar opinions toward the same event. Besides, people tend to hold similar opinions toward events of the same nature. The first hypothesis is validated in the following way.

First, we cluster people by characteristics, based on their profile information using clustering algorithms. Profile information generally includes age, occupation, gender and location and so on, as illustrated in Table 1, which can be extracted from tweets and people's own descriptions about themselves [28]. And then, we compare the similarity of people's attitudes towards a specified event, who are located in the same cluster with similarity of people's attitudes towards the same event, who are randomly chosen from a larger population. If the former similarity is significantly larger than the latter similarity in a statistical sense, the first hypothesis is validated. Following the same line, the second hypothesis could also be tested.

Based on the two validated hypotheses, we can use collaborative filtering which include memory-based collaborative filtering or model-based collaborative filtering to estimate people's opinions on new events.

IV. CONCLUSION

This paper envisages a new application of social web mining in medical and healthcare area. A system called ISSWM is proposed to exploit and analyze raw tweets from multiple social media websites. By ISSWM, we can predict the trends of infectious diseases, understand and estimate people's attitudes towards healthcare related issues, detect ADR events and stay informed about other healthcare related hot topics. ISSWM makes the best of social media websites and can deliver analysis results very efficiently at an extremely low cost.

ACKNOWLEDGMENT

The work described in the paper is financially supported by the Theme Based Research Grant, RGC of Hong Kong (TRS Project 8770001).

REFERENCES

[1] Y. He, Y. d. Liu, F. l. Ma, F. k. Xing, "Sample analysis in 17 medical institutions on the reports of infectious cases using Chinese Information system for disease control and prevention " Journal of Preventive Medicine of Chinese People's Liberation Army, no. 02, 2014, pp. 125-126.

[2] Sina (2014) Sina Weibo is the best crowd funding platform, take Dark Horse Games as an example. http://tech.sina.com.cn/i/2014-02-26/11479193540.shtml. Accessed 10 June 2015

[3] F. Griffiths, et,al, "Social networks–The future for health care delivery," Social science & medicine, vol. 75, no. 12, pp. 2233-2241, 2012.

[4] S. Brennan, A. Sadilek, H. Kautz, "Towards understanding global spread of disease from everyday interpersonal interactions," in Proceedings of the Twenty-Third international joint conference on Artificial Intelligence, 2013, pp. 2783-2789..

[5] C. McNab, "What social media offers to health professionals and citizens," Bulletin of the World Health Organization, vol. 87, no. 8, 2009, pp. 566-566.

[6] R. N. Rimal, and M. K. Lapinski, "Why health communication is important in public health," Bulletin of the World Health Organization, vol. 87, no. 4, 2009, pp. 247-247a.

[7] M. A. Russell, Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More: " O'Reilly Media, Inc.", 2013.

[8] R. Chunara, J. R. Andrews, J. S. Brownstein, "Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak," The American Journal of Tropical Medicine and Hygiene, vol. 86, no. 1, 2012, pp. 39-45.

[9] A. Signorini, A. M. Segre, P. M. Polgreen, "The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic," PloS one, vol. 6, no. 5, 2011, pp. e19467.

[10] Sadilek A, Kautz HA, Silenzio V, Modeling Spread of Disease from Social Interactions. In: ICWSM, 2012.

[11] Sadilek A, Kautz HA, Silenzio V, Predicting Disease Transmission from Geo-Tagged Micro-Blog Data. In: AAAI, 2012.

[12] A. Culotta, "Towards detecting influenza epidemics by analyzing Twitter messages," in Proceedings of the first workshop on social media analytics, 2010, pp. 115-122..

[13] F. Greaves, D. Ramirez-Cano, C. Millett, A. Darzi, L. Donaldson, "Use of sentiment analysis for capturing patient experience from free-text comments posted online," Journal of medical Internet research, vol. 15, no. 11, 2013.

[14] B. C. Wallace, M. J. Paul, U. Sarkar, T. A. Trikalinos, M. Dredze, "A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews," Journal of the American Medical Informatics Association, vol. 21, no. 6, 2014, pp. 1098-1103.

[15] D. King, D. Ramirez-Cano, F. Greaves, I. Vlaev, S. Beales, A. Darzi, "Twitter and the health reforms in the English National Health Service," Health policy, vol. 110, no. 2, 2013, pp. 291-297.

[16] Sina Weibo. "Sina Weibo API," 6th, July, 2015; http://open.weibo.com/wiki/%E9%A6%96%E9%A1%B5.

[17] Tencent. "Tencent Weibo API," 6th, July, 2015; http://dev.t.qq.com/.

[18] Renren. "Renren API," 6th, July, 2015; http://wiki.dev.renren.com/wiki/API.

[19] Sina Weibo. "Rate limiting on Sina Weibo API," 6th, July, 2015; http://open.weibo.com/wiki/Rate-limiting.

[20] 120ask. "120ask," 6th, July, 2015; http://www.120ask.com/.

 ISBN: 978-1-61208-428-2

[21] haodf. "haodf," 6th, July, 2015; http://www.haodf.com/.

[22] mingpaohealth. "mingpaohealth," 6th, July, 2015; http://www.mingpaohealth.com/cfm/Main.cfm.

[23] xywy. "xywy," 6th, July, 2015; http://www.xywy.com/.

[24] ewsos. "ewsos," 6th, July, 2015; http://www.ewsos.com/.

[25] 39net. "39net," 6th, July, 2015; http://www.39.net/.

[26] Wiki. "TF-IDF," 6th, July, 2015; https://en.wikipedia.org/wiki/Tf%E2%80%93idf.

[27] D. M. Blei, A. Y. Ng, M. I. Jordan, "Latent dirichlet allocation," the Journal of machine Learning research, vol. 3, 2003, pp. 993-1022.

[28] J. Li, A. Ritter, E. Hovy, "Weakly supervised user profile extraction from twitter," ACL, Baltimore, 2014.

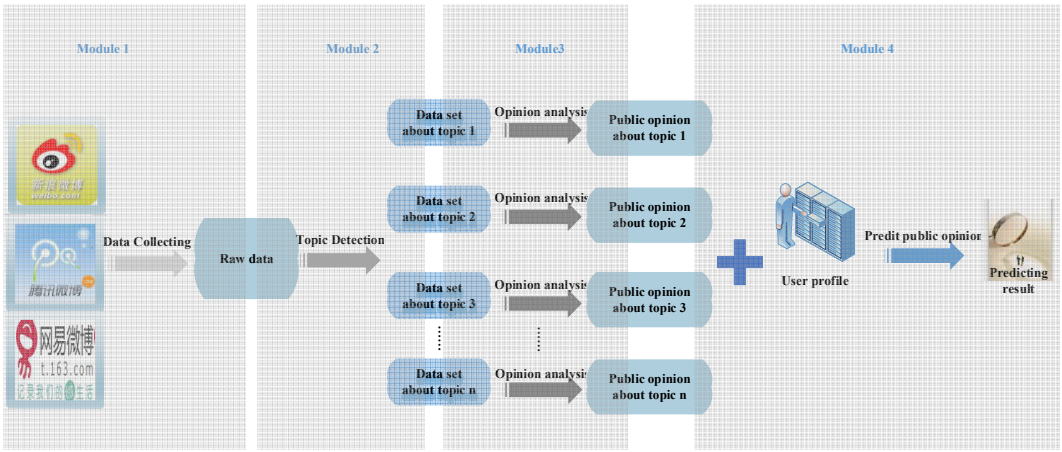


Figure 1. ISSWM for health care.

 ISBN: 978-1-61208-428-2

Energy- and Priority-Aware Traffic Engineering for Content-Centric Networking

Ling Xu, Tomohiko Yagyu

Cloud System Research Lab
NEC Corporation, Japan
Email: lingxu@gisp.nec.co.jp, yagyu@cp.jp.nec.com

***Abstract*—Content-Centric Networking (CCN) is a new network architecture aiming to solve many fundamental problems of IP networks. We are interested in the application of CCN carrier networks during natural disasters. During these times, carrier networks need to save energy until reinforcement arrives. Traffic delivery is prioritized, where high-priority traffic such as disaster news receives superior transmission quality. Multiple Tree-based Traffic Engineering (MTTE), a recently proposed traffic engineering scheme, realized the goal of energy-saving by forwarding traffic on minimal spanning trees. In addition, several priority-aware transmission schemes were proposed recently. However, no CCN communication scheme that realizes both energy-saving and prioritized routing has ever been considered. In this paper, we propose Priority-aware MTTE (PMTTE) to fill this vacuum. Given a certain energy consumption limitation, PMTTE optimizes the transmission quality of the high-priority traffic first. Then, it uses the remaining energy consumption quota to improve the low-priority traffic's quality. We compared the performance of PMTTE against priority queue-enabled MTTE (as a naive energy- and priority-aware CCN traffic engineering scheme). Simulation shows that compared with MTTE, PMTTE can improve the performance of high-priority traffic by up to 80%, and of low-priority traffic by more than 30%.**

Keywords–content-centric networking; priority-aware; energy efficiency; multiple trees

I. INTRODUCTION

A. Content-Centric Networking

Content-Centric Networking (CCN) is a network architecture that is proposed recently for solving IP architecture's problems [1]. Conventional IP networks focus on retrieving needed contents from certain hosts. Nowadays, however, users are only concerned about the contents, not from where to obtain them. In CCN networks, each piece of content is given a unique *name*. Users tell the networks the names of the needed contents. The networks independently identify hosts providing the contents and then fetch the contents to the users.

One of CCN's main features is that routers carry *caches* to reduce redundant traffic. Each host in the network may provide some contents and this host is called the *producer* of its contents. The producer splits each of its contents into a set of *chunks*. The producer also registers the name of each of its contents on each router - a process called *content publishing*. Hosts that request the content is called the content's *consumers*. When a consumer needs a piece content, it sends one request (called an *Interest*) for each chunk of the content. The Interest is then forwarded to the producer according to the information the producer registered on routers. When the producer receives an Interest, it sends the requested chunk back to the consumer. Each router along the chunk's forwarding route tries to store the chunk in its cache. The next time the router receives an Interest for the same content, if the chunk is still in its cache, the router replies directly with the chunk to the consumer.

We are interested in CCN's application in carrier networks during natural disasters and hope to make the CCN network both energy-efficient and priority-aware.

On one hand, networks consume a huge amount of energy, and improving network energy efficiency has been a popular research topic in recent years. In CCN networks, in-router caches cost extra energy. Hence, the networks generally consume more energy than conventional IP networks.

Xu et al. proposed *Multiple Tree-based Traffic Engineering* (MTTE) to reduce CCN's energy consumption [2]. Research has found that *network interfaces* (shortened as *faces* henceforth) of routers in modern networks are generally underutilized [3]. The idea of Xu et al. is to shut down as many faces of routers as possible. MTTE splits traffic on multiple tree-like networks that are generated based on the physical network. The trees are generated in such a way that the number of faces included in the trees is minimized. Since trees generally contain fewer faces than the original physical network, energy can be conserved.

On the other hand, real-world carrier networks deliver traffic of different *priorities*. During natural disasters, disaster alarms and news broadcasting are more important than entertainment TV programs. When disasters occur, the network traffic usually surges suddenly. MTTE is priority-agnostic and cannot ensure the performance of the high-priority traffic.

Several priority-aware transmission schemes have been proposed for CCN networks recently. However, to the best of our knowledge, none of the existing proposals realizes both energy- and priority-aware transmission. Through this paper, we hope to fill this vacuum.

B. Research Goal

In this paper, we assume that the network delivers two kinds of traffic: *high-priority* and *low-priority*. Examples of the high-priority traffic include live streaming such as disaster news. Such traffic is mission-critical and has to be delivered to users stably before certain deadlines. For examples, during the Great East Japan Earthquake [4], tsunami hit the seashore several minutes after the earthquake. If tsunami news can arrive at people in time, thousands of lives can be saved. The low-priority traffic can be other types of communication like binary file transmission. Compared with the high-priority traffic, low-priority traffic is less urgent and can be delayed or dropped.

The simplest idea to implement priority-aware transmission is to always deliver the high-priority traffic first using priority

 ISBN: 978-1-61208-428-2

queues. By "using priority queues," we mean that each face contains one high-priority queue and one low-priority queue. High-priority (low-priority) packets are piled on the high-priority (low-priority) queue in a first-in and first-out manner. The face always sends packets in the high-priority queue before forwarding any packets in the low-priority queue. However, this approach would result in serious quality degradation of the low-priority traffic.

Through this research, we aim to create a traffic engineering scheme that provides these properties: (1) it reduces energy consumption of the whole network system; (2) given a upper bound for the system energy consumption and a lower bound for the transmission quality of the high-priority traffic, it guarantees both the bounds without greately sacrificing the quality of the low-priority traffic.

C. Proposal Summary

When a consumer requests a piece of content, we use *session* to denote the process in which the consumer receives all chunks. Each Interest belongs to one certain session. We call sessions transmitting high- (low-) priority contents the *high- (low-) priority sessions*.

Sessions that have at least one chunk not returned before a certain deadline ϕ_{dead} are regarded as *failed sessions*. We define *Session Failure Rate* (SFR) as the ratio of failed sessions among all sessions. Session failure rate is used as the main transmission quality metric in this paper.

We propose *Priority-aware MTTE* (PMTTE), and compare the performance of PMTTE with priority queue-enabled MTTE (as a naive priority and energy-aware CCN transmission scheme). Simulation shows that, compared with MTTE, PMTTE reduces the session failure rate of high-priority traffic by 80% and of low-traffic by 30%.

The rest of this article is organized as follows: Section II highlights relevant, prior literature in this field; Section III briefly introduces MTTE since MTTE is the base of PMTTE; Section IV introduces PMTTE's design and Section V evaluates PMTTE's performance; and Section VI is the conclusion.

II. Related Work

Several priority-aware transmission schemes have been proposed for CCN networks recently. However, to the best of our knowledge, no scheme that balances the trade-off between the network energy consumption and the transmission quality of prioritized traffic has been considered yet. Kim et al. proposed reserving bandwidth for the high-priority traffic [5] where their communication scheme is similar to the IntServ bandwidth reservation scheme for IP networks [6]. During natural disasters, important messages need to be delivered to more people. Routers have limited capacity and cannot forward all packets. Psaras et al. proposed that packets should be given names that specify their priorities [7]. Routers independently decide whether to forward a packet according to the routers' remaining capacities and the packet's priority. The objective of this research is orthogonal from ours. Immediately after a natural disaster, communication traffic may dramatically increase and overload the network. Psaras et al. proposed prioritizing contents according to the importance of the contents to the contents' consumers [8]. Somaya et al. used prioritized queues for delivering packets with different deadlines [9]. Traffic is classified into multiple groups. Routers maintain queues for accommodating different groups and packets with near deadlines are delivered first. Tsilopoulos et al. realized that for streaming contents in *Information-Centric Networking* (synonym of CCN), it is inefficient for consumers to request each chunk. They suggested that routers work as proxies that request content for consumers. [10]

Research effort has been made in assessing CCN's energy efficiency using simulation [11][12]. Some research compared the energy efficiency of CCN with existing IP-based content delivery techniques such as content delivery networking and peer-to-peer networking. Song et al. [13] noticed that in modern carrier networks, a great amount of traffic is generated from the edge. [13] uses GreenTE - an existing energy-aware TE mechanism designed for IP networks [3] - for reducing energy consumption of the core network, and uses CCN for eliminating redundant traffic generated by the edge network. However, Song et al.'s approach does not reduce the energy consumption of CCN itself.

III. Background Knowledge of MTTE

In CCN, each router contains a *Forwarding Information Base* (FIB). An FIB contains a set of entries and each entry is a mapping from one prefix to one or multiple router face(s). Suppose that the FIB contains two entries $fe0$=''/Asia/'':{face2} and $fe2$=''/Asia/Tokyo/'':{face2,face5} and that an incoming packet has a name "/Asia/Tokyo/music.mp3." The router finds the face whose FIB entry's prefix matches the packet's name and forwards the packets via this face.

A. Tree-Based Congestion Control

In MTTE, the network contains a central server called the *controller*. The controller maintains a database named the *tree set*. Initially, the controller creates one spanning tree based on the physical network topology and adds the tree into the tree set. The controller sends the tree set to routers. Routers update their respective FIBs so that they can forward the packets along the trees.

When congestion is about to occur, the controller creates more trees to spread the traffic and mitigate congestion. Specifically, routers periodically report the utilization of their faces to the controller. The controller calculates the *Congestion Rate* (CR) as the maximal value of the received utilization. When $CR > \phi_{tcc}$, where ϕ_{tcc} is a system parameter, the controller creates another tree and adds it to its tree set. This mechanism is termed as *Tree-Based Congestion Control* (TCC).

B. Face Weight Calculation

We use STs to denote the tree set in MTTE. Faces included in the tree set are called *live faces*, denoted by $E(STs)$. Faces not included the tree set are called *free faces*.

When the controller creates a new tree it calculates the weight of each face. Based on the weights, the controller runs Kruskal's minimal spanning tree algorithm. The controller then asks routers to evenly deliver traffic on trees in the new tree set.

To reduce energy consumption, the new tree should not dramatically increase the number of live faces. To this end, we call the live faces with utilization higher (lower) than ϕ_{ce} the *congested faces* (*uncongested faces*). The controller assigns weights to faces so that the *uncongested faces* are chosen first, free faces are chosen later, and congested faces are chosen last.

 ISBN: 978-1-61208-428-2

Namely, underutilized live faces are more likely to be used to create the new tree, which suppresses energy consumption.

A critical design in MTTE is that the diameters of the routing trees must be suppressed. Otherwise, when the diameters are large, not only is the transmission delay increased but also the cache utilization efficiency is decreased. Congestion is more likely to occur and the transmission quality deteriorates quickly.

MTTE uses a *Betweenness*-based heuristic algorithm to suppress the diameters of the routing trees. For each face e, the controller computes its "face betweenness" ($e.be$) - a widely used metric in graph theory [14]. Informally, $e.be$ represents the number of the shortest paths in the entire network that traverse e. Imagine that routers c and p are a pair of consumer and producer, and sp is the shortest path between c and p on G. Intuitively, if more faces with high betweennesses are added to st, the probability that the data transmitted on st between c and p are delivered along the shortest path is higher. Accordingly, the diameter of st will be small. Based on this observation, in MTTE, the controller selects faces with higher betweennesses first. When the controller creates the initial tree, for each face e, it calculates $e.eb$, and sets $e.weight = 1/e.eb$. When subsequent trees are created, the controller makes the weights of uncongested faces directly proportional to the faces' utilization and makes the weights of free faces inversely proportional to the faces' betweennesses. The complete tree creation algorithm is shown in Algorithm 1.

C. Hash-Based Traffic Splitting

Each time the controller changes the tree set, it asks routers to update their FIBs. First, the controller sends the tree set to all routers. Suppose that H is a collision-proof hash function preloaded on each router, and $N(d)$ is the name of content d. Routers update their FIBs so that d is forwarded on the $H(N(d))\%K$-th tree, where K is the size of the new tree set. After the update, packets are generally evenly delivered on each tree and the congestion can be mitigated.

D. Tree Removal

When $CR < \phi_{uu}$, where ϕ_{uu} is a preloaded system parameter, the controller removes the last tree in the tree set. The controller sends the new tree set to all routers, and asks routers to update their FIBs. Routers shut down their adjacent faces that are not included in the new tree set. Accordingly, energy consumption can be reduced.

IV. PMTTE: THE DESIGN

A. Problem Analysis

PMTTE uses priority queues-enabled MTTE as the base architecture. This architecture has two problems:

First, priority queues alone are not sufficient to ensure the transmission quality of the high-priority traffic. When the network has bandwidth bottlenecks, high-priority packets will drop anyway. MTTE works well in resolving bandwidth bottlenecks by creating more trees before congestion really occurs. Hence, in MTTE, the main reason that routers drop packets can be that the traffic changes so fast that no sufficient trees are created. Certainly we can accelerate tree creation by using smaller ϕ_{tcc}. However, creating more trees consumes more energy.

1: E_{UE} = Faces in $E(STs)$ with utilization $<= \phi_{ce}$.
2: E_{CE} = Faces in $E(STs)$ with utilization $> \phi_{ce}$.
3: E_{free}: Faces not in $E(STs)$
4: U_{max} = maximum of face utilization of E_{UE}
5:
6: The controller calculates the betweenness $e.eb$ of each face e.
7:
8: **if** $|STs|$ = 0 **then**
9: // The controller is creating the initial tree
10: **for all** e in the network **do**
11: e.weight = $1/e.eb$
12: **end for**
13: **else**
14: // The controller creates a new tree for mitigating congestion
15: **for all** $e \in E_{UE}$ **do**
16: e.weight = e.utilization
17: **end for**
18: **for all** $e \in E_{free}$ **do**
19: e.weight = $U_{max} + 1/e.eb$
20: **end for**
21: **for all** $e \in E_{CE}$ **do**
22: e.weight = $U_{max} + 2$
23: **end for**
24: **end if**
25:
26: The controller generates a minimal cost spanning tree st using Kruskal's algorithm on the whole network.
27: **if** st is different from all the existing trees **then**
28: return st
29: **else**
30: return FAILED
31: **end if**

Figure 1: Based on current face utilization, the controller generates a new spanning tree (Algorithm 1 from [2]).

Second, in MTTE, routing trees are created in such a way that they heavily overlap each other for reducing energy consumption. Since priority queues aggressively drop low-priority packets, the transmission quality of the low-priority traffic will be heavily impacted when congestion occurs.

We propose *Priority-Dependent Routing* (PDR) and *Face Separation* (FS) to address these problems. Algorithm 2 lists the PMTTE algorithm's work flow and functionalities.

B. Priority-Independent Routing

We assume that each packet of the high- (low-) priority traffic contains an attribute tag that specifies content' priorities, as done in [7]. Routers can know the priorities of incoming packets by checking priority tags.

To improve the transmission quality of the high-priority traffic, we propose PDR. PDR forwards the high- and low-priority traffic on independent trees. We use *whole traffic* to denote the total of the high- and low-priority traffic. We call the trees used for forwarding the high- (low-) priority traffic the *H (L) trees*. Analogously, MTTE only has one tree set (for the whole traffic) and we call it the *W trees*. We call the three parameters required by MTTE, ϕ_{tcc}, ϕ_{ce} and ϕ_{uu}, a *threshold set*. The controller holds independent threshold sets for the high- and low-priority traffic, respectively. The ϕ_{tcc} of the H trees is generally set lower than ϕ_{tcc} of the L trees. Hence, the

 ISBN: 978-1-61208-428-2

high-priority traffic will have more sufficient routes and is less likely to get congested. Meanwhile, since we only accelerate the H trees' creation, energy consumption can be suppressed.

PDR improves the quality of the low-priority traffic as well. As traffic is separately forwarded, the congestion among traffic of different priorities in routers' forwarding queues can be mitigated. As a result, transmission quality of both the high- and low-priority traffic can be improved.

1) Implementation

To implement PDR in PMTTE, each router holds two FIBs - one for the high-priority traffic and the other for the low-priority traffic. When the system launches, the controller creates two spanning trees and adds them into the H trees and the L trees, respectively. These two trees are created using the original MTTE tree-creating algorithm. Each time the H trees (L trees) change, the controller sends the changed tree set to routers. When a router receives the H (L) trees, it updates its high- (low-) priority FIB using the similar traffic splitting algorithm of MTTE. The difference with MTTE is that, in MTTE, all prefixes (both high- and low-priority) are evenly split on the W trees. In PMTTE, the high- (low-) priority prefixes are evenly split over the H (or L) trees.

For each face e, we use $U_e(H)$, $U_e(L)$ and $U_e(W)$ to denote the utilization caused by the high-priority traffic, the low-priority traffic and the whole traffic, respectively. Similar to MTTE, routers in PMTTE periodically send the utilization of their faces to the controller. The difference between MTTE and PMTTE is that in MTTE, routers only send the $U_e(W)$. In contrast, in PMTTE, routers send both $U_e(H)$ and $U_e(W)$.

Based on the received utilization, the controller adds trees when congestion is about to occur (Algorithm 2). We use the *congestion rate of the high- (low-) priority traffic*, denoted as $\theta(H)$ and $\theta(L)$, as the maximum of the received $U_e(H)$ and $U_e(W)$. When $\theta(H)$ ($\theta(L)$) exceeds $\phi_{tcc}(H)$ ($\phi_{tcc}(L)$), the controller creates a tree for the H trees (L trees). Namely, the low-priority traffic flows in PMTTE in the same way the whole traffic flows in MTTE. Note that $\theta(L)$ is the maximum of $U_e(W)$ instead of $U_e(L)$. The latter means that when the controller creates a new tree for the L trees, the controller is more likely to pick faces used by the H trees. Later, when these faces get congested, low traffic's quality will seriously suffer.

2) Judgment of Live Faces

In the original tree creation algorithm, the controller needs to judge whether a face is live (Line 18, Algorithm 1). In MTTE, all faces used in the tree set are regarded as live faces. In contrast, in PMTTE, faces used in the H trees (rather than used in both the H trees and the L trees) are regarded as live faces when the controller creates a new tree for the high-priority tree set (Algorithm 6). The reason is as follows.

Recall that the transmission quality is closely affected by the diameters of the routing trees (Section III-B). Like in MTTE, when the controller in PMTTE creates a tree and needs to choose a free face, it chooses faces of higher betweennesses first in order to reduce the tree's diameter (Line 12, Algorithm 5). Accordingly, trees created later generally have larger diameters than trees created earlier.

The L trees are created based on the maximum of the whole traffic. Naturally, the number of trees in L tree set is likely to be larger than the number of trees in the H tree set. On average, trees in the L tree set are likely to have larger diameters than trees in the H tree set. By only using faces in the H tree set as live faces, PMTTE can suppress the diameter of the new high-priority tree.

3) Removing Underutilized Trees

If $\theta(H)$ (or $\theta(L)$) is lower than $\phi_{uu}(H)$ (or $\phi_{uu}(L)$), the controller removes the last tree from the H (or L) tree set.

C. Face Separation

Due to the overlapping nature of the tree creation algorithm of MTTE, the H trees and L trees would still heavily cover each other, which threatens the low-priority traffic's quality. To address this problem, when the controller creates a new low-priority tree, it increases the weights of faces that are used in the H trees by a factor of ϕ_{FS} (Algorithm 4). The controller can explicitly specify the extent that the L trees should disjoint from the H trees. This technique is called FS, and ϕ_{FS} is named the *Face Separation Rate*.

```
1: Creates the initial H tree;
2: Creates the initial L tree;
3: while 1 do
4:    if θ(H) > φ_tcc(H) then
5:       CreateTree(H)
6:    end if
7:    if θ(L) > φ_tcc(L) then
8:       CreateTree(L)
9:    end if
10:   Sleeps for a certain period T_update;
11: end while
```

Figure 2: PMTTE work flow

```
1: weights = {};
2: for all e in the network do
3:    weights[e] = GetFaceWeight(e, newTreeType);
4: end for
5: Calculates a Kruskal minimal spanning tree using
   weights;
6: Adds the new tree to the newTreeType trees;
7: Sends the newTreeType trees to routers for updating the
   FIBs;
```

Figure 3: The *CreateTree(newTreeType)* method. $newTreeType$ is 'H' or 'L' when the new tree to create is for the H (or L) trees.

```
1: weight = GetDefaultWeight(e, newTreeType)
2: if newTreeType is H then
3:    return weight
4: else
5:    return weight · φ_FS
6: end if
```

Figure 4: The *GetFaceWeight(e, newTreeType)* method for FS. The controller calculates the weight of a face e when creates a new tree.

D. Energy Consumption and Transmission Quality Trade-off

We use session failure rate (defined in Section I-C) as the metric for transmission quality, and use *Live Face Rate* (LER) as the metric for energy consumption. We use SFR(H) and SFR(L) to denote the SFR of the high- and low-priority traffic, respectively. LER is defined as $\gamma/|E|$, where γ is the number

```
1: if IsLiveFace(e, newTreeType) then
2:    u = GetUtilization(e, newTreeType);
3:    // e is an uncongested face
4:    if u < φce(newTreeType) then
5:       weight = u;
6:    else
7:       // e is a congested face
8:       weight = u + 2;
9:    end if
10: else
11:    // e is a free face
12:    weight = φce(newTreeType) + 1/e.betweenness;
13: end if
14: return weight;
```

Figure 5: The *GetDefaultWeight(e, newTreeType)* method.

```
1: if newTreeType is H then
2:    return e is used in H trees;
3: else
4:    return e is used in H trees or L trees;
5: end if
```

Figure 6: The *IsLiveFace(e, newTreeType)* method

of faces used in the L and H trees, and $|E|$ is the total number of faces in the physical network.

During disasters, network administrators determine a upper bound, ϕ_{energy}, for the system energy consumption and a lower bound, ϕ_{sfr}, for the transmission quality of the high-priority traffic. The administrators dynamically adjust $\phi_{tcc}(H)$ and $\phi_{tcc}(L)$ in a heuristic manner (Algorithm 8). When real system energy consumption exceeds ϕ_{energy}, $\phi_{tcc}(L)$ or $\phi_{tcc}(H)$ is increased to reduce the system energy consumption. When $LER <= \phi_{energy}$ and $SFR(H) > \phi_{sfr}$, $\phi_{tcc}(H)$ is reduced to improve the high-priority traffic's quality. When $LER <= \phi_{energy}$ and $SFR(H) < \phi_{sfr}$, $\phi_{tcc}(L)$ is reduced to improve the low-priority traffic's performance. In this way, administrators can guarantee the requestments for energy consumption and the high-priority traffic, and minimize the low-priority traffic's quality degradation.

V. Evaluation

This section evaluates PMTTE's performance by comparing PMTTE with priority queue-enabled MTTE (MTTE).

A. Topology

Our simulation is performed on ndnSIM – a simulation platform developed by UCLA for CCN-related research [15].

The simulation runs on the network topology of autonomous system 3257 (AS3257). This topology is provided in Rocketfuel network dataset [16], a dataset that has been used in network research [17][18]. Each node in AS3257 represents a router. We extract the largest connected component of AS3257 and use all the remaining nodes for creating trees. AS3257 contains three types of routers: *cores*, *gateways* and *leaves*.

```
1: if newTreeType is H then
2:    return Ue(H);
3: end if
4: return Ue(W);
```

Figure 7: The *GetUtilization(e, newTreeType)* method.

```
1: while 1 do
2:    if LER > φenergy then
3:       if φtcc(L) < 1.0 then
4:          Increases φtcc(L);
5:       else
6:          if φtcc(H) < 1.0 then
7:             Increases φtcc(H);
8:          end if
9:       end if
10:   else
11:      if SFR(H) > φsfr then
12:         if φtcc(H) > 0 then
13:            Reduces φtcc(H);
14:         end if
15:      else
16:         if φtcc(L) > φuu then
17:            Reduces φtcc(L);
18:         end if
19:      end if
20:   end if
21:   Sleeps for Tupdate;
22: end while
```

Figure 8: Energy consumption and transmission quality trade-off

According to the definition of Rocketfuel datasets, leaves are the routers with degrees equal to or less than two, gateways are the routers directly connected to the leaves, and the remaining routers are cores.

We assume that in real-world CCN networks, consumers are adjacent to leaves; low-priority producers, such as servers of entertainment applications, are adjacent to gateways and leaves; high-priority producers, such as servers of national TV stations, connect to cores. We assign one low-priority producer to each leaf and gateway, and assign one high-priority producer to each core. In total, 132 low-priority producers and 108 high-priority producers are located. Each producer provides 20 different pieces of content.

B. Simulation Parameters

We assign $ConsumersPerNode$ high- and low-priority consumers to each leaf. Each high- (low-) priority consumer randomly chooses one high- (low-) priority content from the whole high- (low-) priority prefix pool. End consumers start requesting selected content after certain delays so that network traffic volume follows a sine wave shape (for simulating the traffic spike during the disaster).

We change $ConsumersPerNode$ to control the maximal network traffic volume. Two volumes, with $ConsumersPerNode$ being 60 and 80, are evaluated. The number of high-priority consumers is 4,800 (60 x 80) and 6,400 (80 x 80), respectively. Similarly, the number of low-priority consumers is 4,800 and 6,400 as well. Each consumer issues 10 Interests per second (10 ips), meaning that maximally 10 chunks should be returned by certain producers per second. The size of each chunk is 1024 bytes. According to the simulation, the average number of hops between each pair of consumer and producer is approximately 10. Hence, maximally 983 MB (4,800 x 2 x 10 ips x 10 hops x 1024 bytes) traffic can flow within the network per second when $ConsumersPerNode = 60$. Meanwhile, each face has a capacity of 5 Mbps. The network contains 420 faces.

 ISBN: 978-1-61208-428-2

The theoretical capacity of the whole network is 260 MBps (5 Mbps x 420 / 8). Therefore, we believe that the network is sufficiently congested to evaluate the effectiveness of the proposed algorithm when $ConsumersPerNode = 60$ and 80.

The threshold set of the L trees in PMTTE is equal to the threshold set of the W trees in MTTE. Both sets have $\phi_{tcc} = 0.8$, $\phi_{ce} = 0.6$ and $\phi_{uu} = 0.5$.

The number of high- and low-priority consumers is equal. Intuitively, the high- and low-priority traffic have equal volumes. As ϕ_{tcc} of MTTEs' W trees is 0.8, ϕ_{tcc} of PMTTE's H trees should be equal or less than half of it, i.e., 0.4. For PMTTE's H trees, we set the default values of ϕ_{tcc}, ϕ_{ce} and ϕ_{uu} to 0.4, 0.3 and 0.15, respectively.

C. Performance Metrics

We evaluate the performance using SFR and LER. Session failure rate is directly decided by the session failure deadline ϕ_{dead}. ϕ_{dead} should be set according to the mean round-trip time (RTT) of the system. The RTT in different communication systems may vary greatly. In [19], the authors evaluated the Quality of Experience (QoE) of a streaming system and they assumed a RTT of 150ms. Streaming applications generally request low RTTs. In this paper, we are considering general carrier networks where both streaming and non-streaming traffic exist. Hence, the RTT in our system should be larger than the 150ms. In our simulation, we set ϕ_{dead} to 200ms.

Recall that LER is defined as $\gamma/|E|$. For MTTE, γ is the average number of faces used in the whole tree set during the simulation, For PMTTE, γ is the average number of faces used in the L and H trees during the simulation.

We expect to see that, compared with MTTE, PMTTE further reduces SFR(H) and SFR(L) at the price of a moderate increase in LER. The effectiveness of PDR and FS are evaluated, respectively.

We repeat each evaluation scenario ten times and display the mean values of each performance metrics.

Parameters used in the simulation are summarized in Table I.

TABLE I: Simulation parameters

Parameter	Value
Total number of faces	420
Total number of routers	240
Number of gateway routers	52
Number of leave routers	80
Number of core routers	108
Prefixes per producer	20
Chunk size	1024 bytes
Cache size	2000 chunks
(For PMTTE) Default H ϕ_{tcc}, ϕ_{ce}, ϕ_{uu}	0.4, 0,3, 0.15
(For PMTTE) Default L ϕ_{tcc}, ϕ_{ce}, ϕ_{uu}	0.8, 0.6, 0.5
(For MTTE) Default Whole ϕ_{tcc}, ϕ_{ce}, ϕ_{uu}	0.8, 0.6, 0.5
Duration of one simulation run	50 seconds
Update interval T_{update}	2 seconds
High-priority producer location	cores
Low-priority producer location	gateways, leaves
Session failure deadline ϕ_{dead}	0.2 seconds

D. Evaluating Priority-Independent Routing

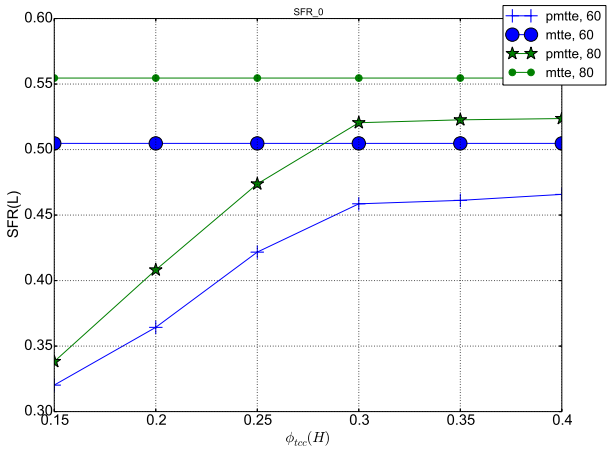


Figure 9: Session failure rate for the low priority traffic. Legends indicate the algorithm and traffic volume used for evaluation. For example, "pmtte, 60" indicates that PMTTE is evaluated when $ConsumersPerNode = 60$.

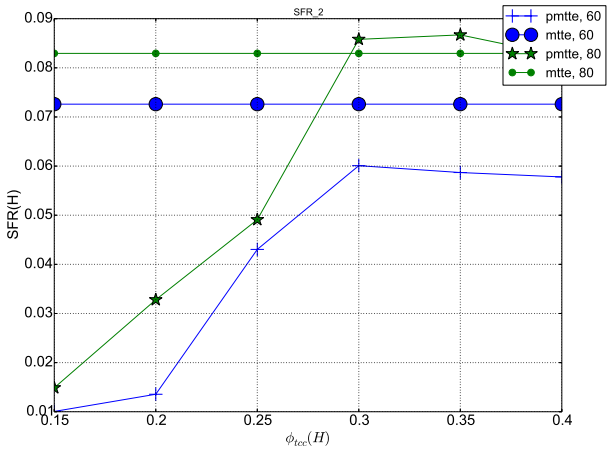


Figure 10: Session failure rate for the high priority traffic.

Figures 9 and 10 compare the session failure rates of the low-priority traffic and high-priority traffic between PMTTE and MTTE. On the Figures, $\phi_{tcc}(H)$ increases from 0.15 to 0.4 on the x-axis. For evaluation displayed in this section, ϕ_{FS} is set to 1.0 (i.e., FS is disabled).

[19] shows that in video streaming systems, the users' QoE degrades rapidly when the packet drop rate exceeds certain thresholds. Since we consider prioritized transmission, instead of limiting the absolute value of the session failure rate, it is more important to focus on the performance superiority of the high-priority traffic over the low-priority traffic. In this paper, we assume that SFR(H) should be lower than 2% to guarantee decent QoE. Figure 10 shows that when $\phi_{tcc}(H)$ is 0.15, SFR(H) can be reduced to lower than 2%.

As $\phi_{tcc}(H)$ decreases, more H trees are created and SFR(H) of PMTTE is suppressed. In Figure 10, it is notable that PDR effectively reduces SFR(H) when $\phi_{tcc}(H)$ is reduced to 0.15. SFR(H) reduces from 7.2% to 1% when

 ISBN: 978-1-61208-428-2

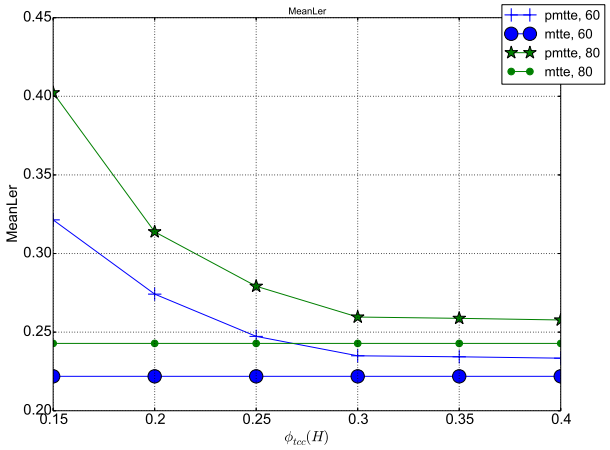


Figure 11: Live face rate.

$ConsumersPerNode = 60$, and reduces from 8.3% to 1.5% when $ConsumersPerNode = 80$. Meanwhile, when SFR(H) is less than 2%, SFR(H) can be than more than 10 times lower than SFR(L) (Figure 9 and Figure 10).

Meanwhile, as expected in Section IV-B, PDR remarkably reduces SFR(L) as well (Figure 9), as compared with MTTE. When $ConsumersPerNode = 60$, SFR(L) decreases from 51% to 33%. When $ConsumersPerNode = 80$, SFR(L) decreases from 56% to 34%.

Figure 11 plots the influence of PDR on energy consumption. Clearly, the improvement in transmission quality is at the price of higher energy consumption. Even though PMTTE still suppresses LER to be lower than 40%. In real-world network systems, the network operator can tune $\phi_{tcc}(H)$ to make the trade-off between transmission quality and energy consumption.

E. Evaluating Face Separation

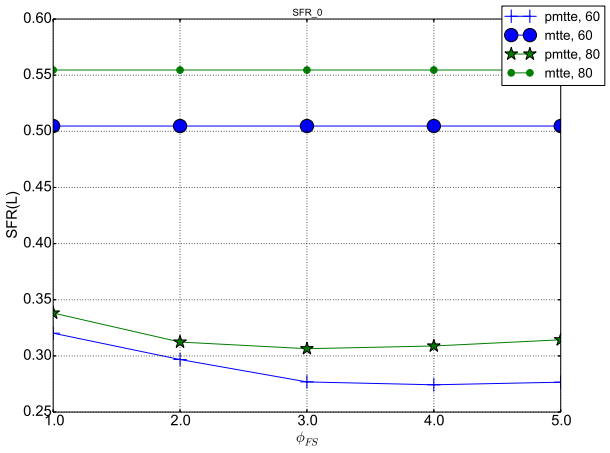


Figure 12: Session failure rate for the low priority traffic. Legends indicate the algorithm and traffic volume used for evaluation. For example, "pmtte, 60" indicates that PMTTE is evaluated when $ConsumersPerNode = 60$. (FS enabled)

Figures 12 and 13 compare the session failure rates between PMTTE and MTTE when FS is enabled. The results are

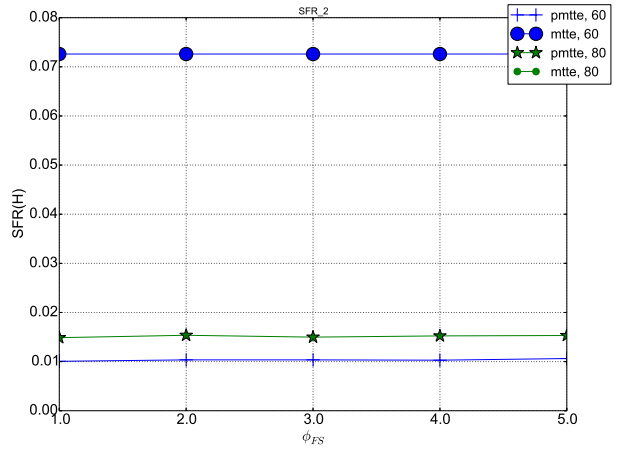


Figure 13: Session failure rate for the high priority traffic. (FS enabled)

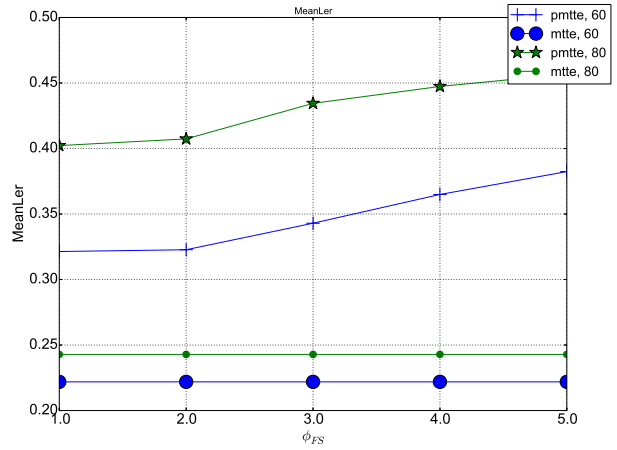


Figure 14: Live face rate. (FS enabled)

obtained with $\phi_{tcc}(H) = 0.15$ (the value obtained in Section V-D where SFR(H) of PMTTE falls below 2%). The face separating rate ϕ_{FS} increases from 1.0 to 5.0 on the x-axis. When ϕ_{FS} takes 1.0, the results are equal to those obtained when FS is disabled and PDR only is enabled.

Figures 12 and 13 clearly reveal that priority queues deteriorate the low-priority traffic. Meanwhile, FS can effectively mitigate this deterioration. As the face separation rate increases, SFR(L) decreases (33% to 27% for $ConsumersPerNode = 60$, and 34% to 32% for $ConsumersPerNode = 80$). Moreover, FS does not affect SFR(H) (Figure 13), meaning that network operators can improve the performance of the low and high traffic using PDR and FS independently.

Figure 14 shows that FS moderately increases energy consumption. However, LER is still lower than 50% in PMTTE.

VI. Conclusion and Future Work

Through this research, we aim to create a traffic engineering scheme that is both energy- and priority-aware for CCN networks. The challenge is to avoid heavily impacting the transmission quality of the low-priority traffic. Our proposal,

 ISBN: 978-1-61208-428-2

PMTTE, splits the high- and low-priority traffic on separate routing trees. This enables us to flexibly increase the tree recreation frequency for the high-priority traffic to suppress the high-priority traffic congestion. At the same time, this change mitigates the collision between the high- and low-priority traffic, which improves the quality of both. Last but not least, by aggressively separating the routes of the high- and low-priority traffic, we can actively improve the quality of the low-priority traffic. Simulation using a real-world ISP network topology shows that compared with naive priority queue-enabled MTTE, PMTTE can boost the quality of both the high- and low-priority traffic by up to 50%. As for future work, we plan to improve the transmission quality of the high-priority traffic by creating routing trees that minimize the distances between the high-priority producers and consumers. Meanwhile, MTTE is a centralized scheme and will fail when the connection between routers and the control is cut. We plan to make MTTE more fault-tolerant, especially during disasters.

ACKNOWLEDGMENT

The work described in this paper was, in part, performed in the context of the FP7/NICT EU-JAPAN GreenICN project.

REFERENCES

[1] L. Zhang et al., "Named data networking (NDN) project," Relatório Técnico NDN-0001, Xerox Palo Alto Research Center-PARC, 2010.

[2] L. Xu and T. Yagyu, "Multiple-tree based online traffic engineering for energy efficient content centric networking," in Proc. The Sixth International Conference on Advances in Future Internet (AFIN2014). IARIA, 2014, pp. 121–127.

[3] M. Zhang, C. Yi, B. Liu, and B. Zhang, "GreenTE: Power-aware traffic engineering," in Network Protocols (ICNP), 2010 18th IEEE International Conference on. IEEE, 2010, pp. 21–30.

[4] N. Mimura, K. Yasuhara, S. Kawagoe, H. Yokoki, and S. Kazama, "Damage from the great east japan earthquake and tsunami-a quick report," Mitigation and Adaptation Strategies for Global Change, vol. 16, no. 7, 2011, pp. 803–818.

[5] Y. Kim, Y. Kim, and I. Yeom, "Differentiated services in named-data networking," in Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on. IEEE, 2014, pp. 452–457.

[6] B. Braden, D. Clark, and S. Shenker, "Integrated services in the internet architecture: an overview," Internet Requests for Comments, RFC Editor, RFC 1633, June 1994, http://www.rfc-editor.org/rfc/rfc1633.txt. [Online]. Available: http://www.rfc-editor.org/rfc/rfc1633.txt

[7] I. Psaras, L. Saino, M. Arumaithurai, K. Ramakrishnan, and G. Pavlou, "Name-based replication priorities in disaster cases," in Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on. IEEE, 2014, pp. 434–439.

[8] G. Tyson, E. Bodanese, J. Bigham, and A. Mauthe, "Beyond content delivery: Can icns help emergency scenarios?" Network, IEEE, vol. 28, no. 3, 2014, pp. 44–49.

[9] S. Arianfar, P. Sarolahti, and J. Ott, "Deadline-based resource management for information-centric networks," in Proceedings of the 3rd ACM SIGCOMM workshop on Information-centric networking. ACM, 2013, pp. 49–54.

[10] C. Tsilopoulos and G. Xylomenos, "Supporting diverse traffic types in information centric networks," in Proceedings of the ACM SIGCOMM workshop on Information-centric networking. ACM, 2011, pp. 13–18.

[11] N. Choi, K. Guan, D. C. Kilper, and G. Atkinson, "In-network caching effect on optimal energy consumption in content-centric networking," in 2012 IEEE International Conference on Communications (ICC). IEEE, 2012, pp. 2889–2894.

[12] U. Lee, I. Rimac, D. Kilper, and V. Hilt, "Toward energy-efficient content dissemination," Network, IEEE, vol. 25, no. 2, 2011, pp. 14–19.

[13] Y. Song, M. Liu, and Y. Wang, "Power-aware traffic engineering with named data networking," in Seventh International Conference on Mobile Ad-hoc and Sensor Networks (MSN), 2011. IEEE, 2011, pp. 289–296.

[14] M. Girvan and M. E. Newman, "Community structure in social and biological networks," Proceedings of the National Academy of Sciences, vol. 99, no. 12, 2002, pp. 7821–7826.

[15] A. Afanasyev, I. Moiseenko, and L. Zhang, "ndnsim: Ndn simulator for ns-3," Named Data Networking (NDN) Project, Tech. Rep. NDN-0005, Rev, vol. 2, 2012.

[16] "Rocketfuel dataset," https://github.com/cawka/ndnSIM-ddos-interest-flooding/tree/master/topologies/rocketfuel_maps_cch, (retrieved: July, 2015).

[17] N. Spring, R. Mahajan, and D. Wetherall, "Measuring ISP topologies with rocketfuel," ACM SIGCOMM Computer Communication Review, vol. 32, no. 4, 2002, pp. 133–145.

[18] C. Yi et al., "A case for stateful forwarding plane," Computer Communications, 2013, pp. 779–791.

[19] T. N. Minhas, O. Gonzalez Lagunas, P. Arlos, and M. Fiedler, "Mobile video sensitivity to packet loss and packet delay variation in terms of qoe," in Packet Video Workshop (PV), 2012 19th International. IEEE, 2012, pp. 83–88.

 ISBN: 978-1-61208-428-2

On Security-Effective Mobility-QoS Management Scheme in Heterogeneous Mobile Networks

Hyeungwoo Lee
Network Development Group
SAMSUNG SDS Co. Ltd.
Suwon, Gyeonggi-do, 443-822, Republic of Korea
e-mail: hw77.lee@samsung.com (zen016@naver.com)

Jae-Young Choi and Jongpil Jeong
College of Information & Communication Engineering
Sungkyunkwan University
Suwon, Gyeonggi-do, 440-746, Republic of Korea
e-mail: {jaeychoi, jpjeong}@skku.edu

***Abstract*—To support efficient mobility, host-based mobility management protocols have been developed. The Authentication, Authorization, Accounting, and Charging (AAAC) system is used in this paper to analyze the effectiveness of the existing Proxy Mobile IPv6 (PMIPv6) and Fast Handover for PMIPv6 (FPMIPv6) network security. Furthermore, the IPv6 Mobility Management Protocol (MMP) features, performance, and seamless transfer performance in terms of packet loss probability are also analyzed. Our scheme can be efficiently used to integrate Quality of Service (QoS) and mobility to manage and control resources using a QoS Broker (QoSB). The evaluation results show a better overall performance for the fast handover structure of mobility management techniques. PMIPv6 and FPMIPv6 are, in many respects, the most efficient structures possible. Specifically, the fast handover structure of the network-based mobility management schemes shows the best results.**

Keywords—Mobility-QoS; Security-Effective; Mobility Management Protocol; PMIPv6.

I. INTRODUCTION

The wireless mobile environment is rapidly growing in the digital environment that leads by human hands. The Mobility Management Protocol (MMP) is a core protocol of the wireless mobile environment. Mobile social networking, computing, shopping, and so on will be achieved using the mobility operating system. Various MMPs have been developed for various mobility services. Particularly at the network layer, mobility support techniques have been developed by the Internet Engineering Task Force (IETF). The Mobile IPv6 (MIPv6) specification was proposed, then the Fast Handovers for MIPv6 (FMIPv6) and Hierarchical MIPv6 (HMIPv6) specifications were developed as extensions. As MIPv6 was developed, analysis of IPv6 MMP was used to improve performance [1].

When host-based MMPs operated within the wireless mobile telecom infrastructure, the telecoms companies and technical developers became aware that it was not a suitable solution for mobile services, especially for service providers, as it was necessary to equip a Mobile Node (MN) with mobility support inside the network protocol stack. Therefore, MNs had to be upgraded or developed. This increased the construction costs and complexity of the MN. Host-based MMP has led to a lack of complex control operators. A new approach to mobility services was required.

The extended protocol of Proxy MIPv6 (PMIPv6), Fast Handover for PMIPv6 (FPMIPv6), has improved the transmission rate by reducing transmission latency and packet loss. In contrast to host-based MMP, network-based MMPs (such as PMIPv6 or FPMIPv6) are in the early stages of development. Improving the security of personal authentication using the FPMIPv6 by applying the Authentication, Authorization and Accounting (AAA) mechanism has been studied. When moving between domains in management, AAA techniques for authenticating the MN are required [2]. The AAA scheme for the various wired and wireless services performs authentication, authorization, and billing. Today, many techniques in conjunction with the AAA protocol are being investigated to perform the functions of MN AAA, which is the authentication process between the MNs. For instance, Zhou, H et al. [3] proposed an FPMIPv6-based authentication technique. When the MN enters a new network, this technique protects the authenticating MN from security threats, such as Replay Attack or Key Exposure.

In this paper, we propose a security-effective mobility management scheme for IPv6-based networks using a Quality of Service Broker (QoSB). This protocol can be efficiently used to integrate Quality of Service (QoS) and mobility to manage and control resources using a QoS Broker (QoSB). The time latency is not significantly affected because of the addition of the QoSB. In PMIPv6 and FPMIPv6, the new proposal performs better with respect to the handover latency, packet loss, and handover blocking rates than the traditional MIPv6 scheme. These results are shown for PMIPv6 and FPMIPv6 on a network security system that uses Authentication, Authorization, Accounting, and Charging (AAAC). Furthermore, in this paper, we propose a unified criterion to analyze both host-based and network-based MMPs.

This paper is organized as follows. Section 2 discusses related work, and Section 3 describes the operating procedures of the proposed scheme. In Section 4, a performance evaluation of the proposed method is presented. Finally, Section 5 presents some conclusions regarding these results.

II. RELATED WORK

The PMIPv6 domain structure is composed of a Local Mobility Anchor (LMA), Mobile Access Gateway (MAG),

 ISBN: 978-1-61208-428-2

and MN. An LMA is one kind of Home Agent (HA) that serves as an MN in PMIPv6. In detail, a Home Network Prefix (HNP) is allocated to the MN that maintains the address and location information of all the MNs within the domain and also ensures connection. The MAG is responsible for network connectivity and routing functions on behalf of the MN. It also performs MN mobility signaling by tunneling through the LMA [4]. FPMIPv6 is a mobility protocol that reduces the handover latency and packet loss found in PMIPv6. It may reduce the loss of buffered packets by creating a bi-directional tunnel between the previous MAG (pMAG) and new MAG (nMAG) before making the link-layer handover. And it consists of two modes: predictive and reactive [5] [6] [7].

The QoS architecture is easily able to support end-to-end QoS in terms of the operator. When the MN is moving, it is guaranteed end-to-end connectivity and user maintenance. The architecture is designed to control the scalable deployment of resources in the access network. The core aim of the architecture is the simultaneous support of mobility and QoS. The QoS frameworks of various IETFs have considered both purposes before the final design of the QoS architecture. The advantages and disadvantages of Integration Services (Intserv) [8] and Differentiated Services (Diffserv) [9] have been discussed widely and are well known. However, they do not specifically support mobility and no hybrid solution integrates mobility and QoS. Therefore, the aim of user mobility suggests an innovative use of QoSB-related Fast HandOver (FHO) that includes a Diffserv system to efficiently control and manage the resources available [10]. This architecture is based on the concept that the user is authorized by the service provider of a contract. The QoSB, according to the user's agreement, is responsible for resource allocation in an individual subscription service. As proposed in [11], these services are generally fixed transmission services (for example, a "Guaranteed Rate of 64 KB/s," or "Target Rate of 32 KB/s") but are equipped with a potential mechanism for flexible service negotiation. The QoSB can manage the flow of resources in the core network. To reduce the signal overhead, the system is designed for a user/terminal so that it is not necessary to explicitly reserve or release resources. The services are requested by a simple Diffserv Code Point (DSCP) marking to the outgoing packet. If the MAG receives a packet from a particular user's DSCP value, it sends the required QoSB configuration. The QoSB configures the MAG to fit the appropriate QoS policy based on the information about the user. Services are implicitly suspended by an inactivity timeout. This concept is explained in more detail in [11].

III. PROPOSED SCHEME

The AAAC architecture is shown in Figure 1. It is based on the AAA architecture, which is optimized for IPv6 enhanced auditing, metering, and charging. Considering that AAAC is used for QoS in a PMIPv6 environment, the architecture is designed to offer new functionality and optimize the performance of the overall system.

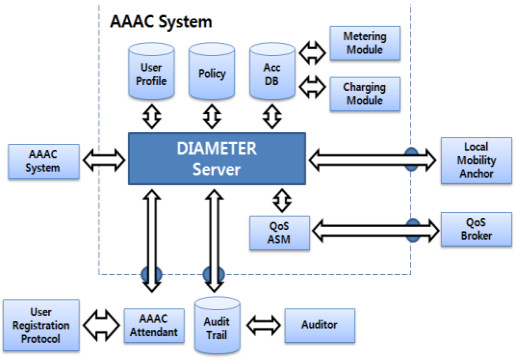


Figure 1. Enhanced Generic AAA Architecture That Supports QoS-enabled Mobility Management.

This architecture enables the subsequent auditing of AAAC using the AAAC audit trail and other factors. Hence, the policy repository is considered as a part of the policy-based AAAC system. The AAAC system supports multiple interfaces. AAAC performers can be treated with MN and interfaces. Communication is performed by the User Registration Protocol (URP). The Application Specific Module (ASM) communicates with the QoSB. The advantage of ASMs is additional flexibility, as various service equipment can be easily processed using the same method from the point of view of the AAAC system. ASM uses the AAAC protocol to communicate with the AAAC system and equipment-specified protocol to communicate with service equipment. There are clear differences in the services provided to the user in this architecture (e.g., QoS is possible and the charging system follows the AAAC requirements). For users previously provided and connected via the ASM and extended AAAC protocol, on the other hand, it is possible, if necessary, to communicate directly with the AAAC system using dedicated communication. AAAC system communication may be enhanced through an appropriate expansion by the DIAMETER default protocol. A key element of QoS service and charging is the means to measure the service used. In an IP-based measurement framework of the IETF Working Group, a variable was defined for the IP flow that depends on the needs of the network administrator. In the IPv6 network, the usage is measured according to the type of service subscribed to by the user and is sent to the QoS [2].

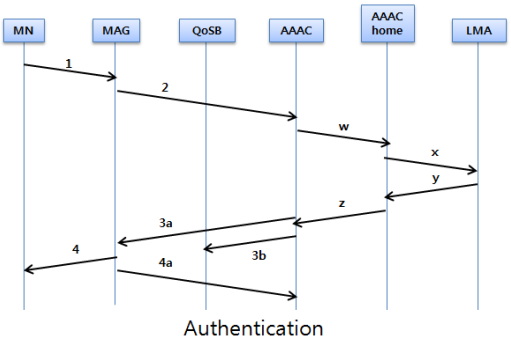


Figure 2. Registration Process.

For network operation and control, three steps may be specifically identified. 1) Registration: in this architecture, an MN/user can start using network resources after authentication and authorization, just as in today's networks. 2) Acceptance: users should be allowed to use a specific service prior to release by the network. 3) Handover: user mobility should preserve existing resources when transferring from one MAG to another.

The registration process to support end-to-end QoS is shown in Figure 2. The registration process begins after the CoA obtains the MN through automatic configuration with the two-layer identifier. When a Duplicate Address Detection (DAD) is performed, the uniqueness of the registered address is checked. When obtaining a non-authorized CoA, the user is authorized to consume only enough resources for the registration message. However, emergency calls can be made. However, as shown in Figure 2, the MN must start the authentication procedure by sending user authentication information (message 1) to the MAG for network connection. The request for this MAG is transmitted to the controlled AAAC system (message 2). For more complicated roaming, domain A (the AAAC domain) sends a registration request to the home AAAC (message w) of the MN. Domain A plays the role of a foreign domain that must contact the home AAAC of the MN. The AAAC checks first if there is a formal contractual relationship between the management domain and its own management domain (corresponding to the roaming agreement) as per the request. If the result is affirmative, the home AAAC performs authentication by verifying the provided credentials. The home AAAC sends a request to the user's LMA (message y). Finally, the home AAAC responds to the AAAC of domain A. A positive response consists of a user profile that contains the information necessary to provide the requested service in the foreign domain. The user profile contains the central management profile, including all relevant user-specific information related to the service provider. From a NVUP (network view of the user's profile), a part of the profile is sent from the AAAC server and it is necessary to provide the requested service in the foreign domain. The user profile contains the central management profile, including all relevant user-specific information related to the service provider. From a NVUP (network view of the user's profile), a part of the profile is sent from the AAAC server to the QoSB (message 3b) that also performs a DAD at this point. The other set of profiles is sent to the AAA Attendant in the MAG (message 3a). Along with the measurement and security information that is delivered to the AAA attendant, the NVUP includes all the required information related to network services. Further, the AAAC informs the MN that the registration is successful via the MAG (messages 3a and 4). After that, the MAG starts a task for the user and informs the AAAC (message 4a). Accordingly, the authentication phase is finished, and a user can access the network.

Figure 3 presents the process for granting authorization to each network service (messages 5–11). The MN sends a packet (message 5) with a DSCP code to request a specific subscription service (e.g., 256 KB/s for priority network access). One of the trailer packets, depending on the configuration of the MN, may be a dummy packet or a packet with real information. If the requested service does not comply with the policy that has already been set in the MAG, the MAG sends a request to QoSB via the QoS manager. According to the analysis, the user's NVUP, and the availability of resources for the request, the QoSB determines whether a message (message 7) is sent to the MAG. The QoS manager of the MAG sets (message 7a) the appropriate policy for the MAG, user and MN services, or notifies the user of the service denial (message 7b). After message 7a, the packet is sent to the MN that matches the configured policy rules (message 8). Packets that have different DSCP codes are subjected again to authentication. When the packet reaches the final domain with other users, it starts another QoS authorization process (message 8a). The QoS manager of the MAG sends a policy question to the QoSB (messages 9 and 10). If the QoSB has the resources, the QoSB manager receives a positive response and is configured for the MAG and its policies (message 10a). If it does not have the resources, the MAG sends a reject message regarding the service (message 10b). After message 10a, the next packet to meet the policy is able to arrive at the other terminal (message 11). In this way, two kinds of access networks can provide the QoS level of an agreement. The core network is monitored for performance (the end-to-end QoS) as expected.

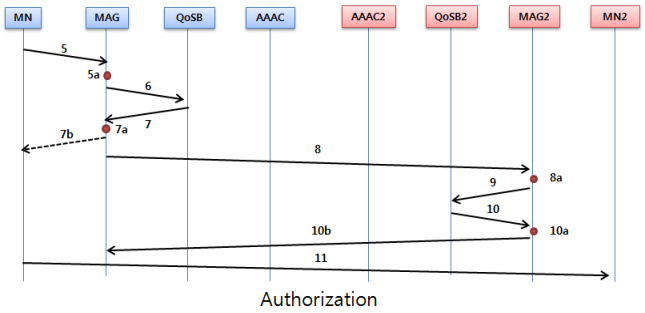


Figure 3. Authorization process.

One of the most difficult problems of IP mobility is ensuring a constant level of QoS. As shown in Figure 4, for user mobility in the network, the handover and network messages are exchanged between QoSBs using the FHO technique. When the MN begins to receive a weakened signal from the current MAG (message 1), it sends and receives the AS (Attendant Solicit), AA (Attendant Advertisement), AReq (Authentication Request) messages, and the handover procedure from the old MAG to a neighboring "new MAG" is started. The MN builds its own CoA and starts the handover process by sending an IP-handover request for the new MAG through the old MAG (message 4). The FHO module of the old MAG requests the FHO Module of the new MAG and submits it to the QoS manager module. The QoS manager immediately sends a request to the existing QoSB (message 5). The previous QoSB sends the handover request comprising the user's NVUP and a list of current user services to the new QoSB (message 6). By default, this task is transferred to a new

 ISBN: 978-1-61208-428-2

QoSB in the context of the existing QoSB. The new QoSB uses this information to check the availability of resources. The MN sends the message to determine whether to perform a handover to the QoS manager of the new MAG (message 7). This mechanism enables the QoSB to stop the handover because of QoS constraints (e.g., the loss of bandwidth resources). If the handover is possible, the QoS function manager sends this information to the FHO module (message 7a) and configures the new MAG to accommodate the moving MN. The new MAG starts the accounting process in the user's AAAC system account (messages 8 and 11). To complete the handover, the MN sends the LMA binding update (messages 9 and 10) to the LMA. In addition, the FHO module sends a handover response to the FHO module of the old MAG (message 12). The new MAG begins bi-casting. The MN handover to the new MAG is complete. If a handover is completed within a QoSB domain, the QoSB for controlling both MAGs is the same, and message 6 is not sent. The remaining messages are the same. Information-related security is exchanged between QoSBs in a similar way.

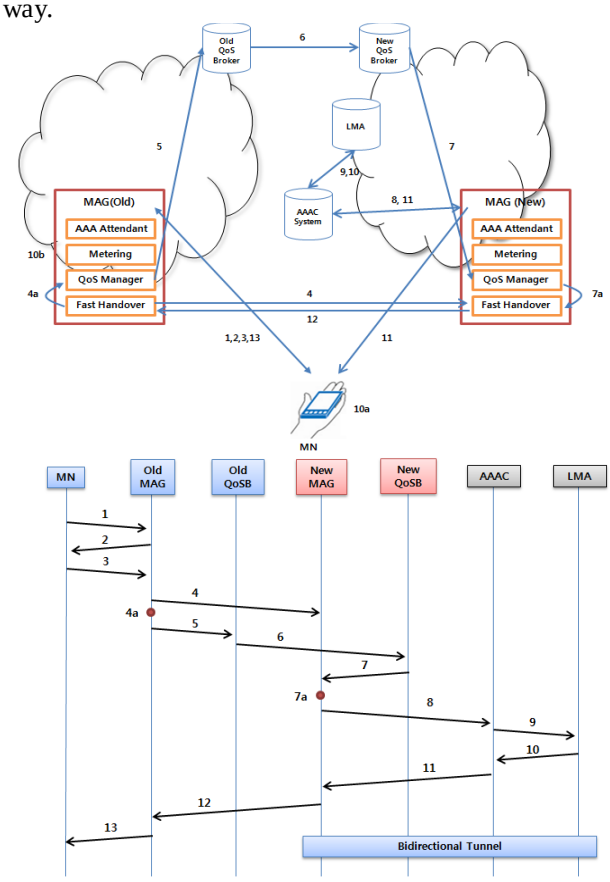


Figure 4. Handover with QoS for End-to-End QoS Support.

IV. PERFORMANCE ANALYSIS

This section evaluates the performance of the proposed method and the existing mobility management protocols.

A. Network Modeling

Figure 5 is shown a generic network topology model.

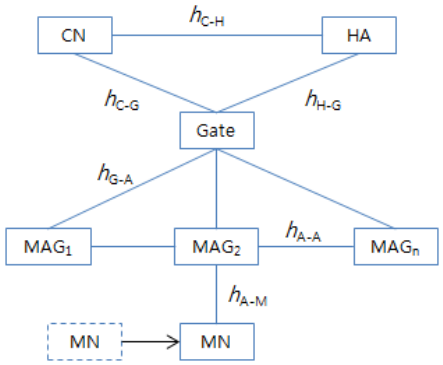


Figure 5. Network Model.

In Figure 5, the following hop count parameters are defined for describing particular paths between communication entities.

- h_{C-H}: It is the average number of hops between the correspondent node (CN) and the HA.
- h_{C-G}: It is the average number of hops between the CN and the gate.
- h_{H-G}: It is the average number of hops between the HA and the gate.
- h_{G-A}: It is the average number of hops between the gate and the MAG.
- h_{A-A}: It is the average number of hops between the neighbor MAGs.
- h_{A-M}: It is the average number of hops between the MAG and the MN.

The latency of registration lasts from when the user turns on a device until it becomes available for use. There are two types of latency. “Low-layer” latency refers to the delay when connecting to a technology after the device is ready for use. “High-layer” latency refers to the delay between sending message 1 and the arrival of message 4 (Figure 2).

$$\text{Re}\,gistration\_Delay = LowLayerDelay + HigherLayerDelay \quad (1)$$

First, this technique does not depend on a particular network protocol or architecture. Second, it is dependent on the link speed. If the user is roaming, it is the “electronic distance” between the outside and home AAACs. The link latency that occurs between the MAG and AAAC system is small enough to be negligible, as it is normally the case that the management infrastructure overprovisions the link to the resource. The processing latency in the system occurs when there is an overload in the number of requests or a database is processing more requests that its capacity. However, in the actual production of the network, this is sufficiently possible to prevent using an appropriate computing or routing tool. Thus, if a user is roaming, the limiting factor is the distance between the external and home AAACs. In this case, the latency is determined by the registration time.

Session setup latency is the time required for the user to access the network. In Figure 3, the session setup latency is

the delay between messages 5 to 11. This latency is composed of the processing time of the MAGs and QoSB, indicating the distance between the MAGs, and the link latency.

Handover latency can be an important parameter, depending on a user's sensitivity. The handover method should be quick as possible to provide seamless service to users. The handover latency is composed of transmission latency, computational latency, and two-layer handover latency.

$$Handover\_Delay = \sum Transmission\_Delays + \sum Computation\_Delays + \sum Layler2\_Handover \quad (2)$$

The transmission latency is the sum of the latency that occurs during MN-MAG, MAG-QoSB, QoSB-QoSB, and MAG-MAG communication. The global handover latency is the delay between messages 1 and 13, as shown in Figure 4. However, the time during no connectivity or when user terminal is not assigned resources is spent in the L2 handover. The handover operation is simulated to determine mobility seamlessness. After execution, measured values proving the architectural concept is obtained that showed reduced global handover latency, as well as low packet loss. The global handover latency affects the relationship between cell coverage of radio (range) and the speed at which the user can move, therefore affecting the cellular network plan.

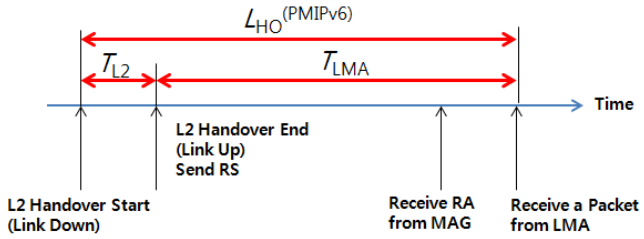


Figure 6. Timing Diagram for PMIPv6 Handover.

Figure 6 shows a timing diagram of a PMIPv6 handover. Here, $L_{HO}^{(PMIPv6)}$ is defined to be the handover latency of PMIPv6 and is expressed as follows.

$$L_{HO}^{(PMIPv6)} = T_{L2} + T_{LMA} \quad (3)$$

Furthermore, T_{LMA} refers to the time spent exchanging the PBU/PBAck messages between the MAG and LMA. During the time required to send the RS Message, the LMA receives the packet that is first sent.

$$T_{LMA} = d_{wl}(L_{RS}) + d_{wd}(L_{PBU}, h_{G-A}) + d_{lma-packet} \quad (4)$$

Variable $d_{lma-packet}$ denotes the time required for the first data packet to be sent from the LMA to the MN. Because it can be implemented by a static tunnel, a bi-directional tunnel between the LMA and MAG is not necessary. When the LMA receives a valid PBU message from the MAG, it sends a data packet and PBAck message to the MN.

$$d_{lma-packet} = d_{wl}(L_D) + d_{wd}(L_D + L_T, h_{G-A}) \quad (5)$$

Latency L_T is contained in the d_{wd} account because the data packet sent to the MN is tunneled between the LMA and MAG. This account is included in the total. The data packet is sent to the MN because of the tunnel between the LMA and MAG. This is different than in HMIPv6. Even if PMIPv6 and HMIPv6 send a similar message to the MN, the PMIPv6 reduces the packet transmission overhead of the wireless link. The FPMIPv6 has a concept similar to FMIPv6 and is composed of predictive and reactive modes.

Let $L_{HO}^{(\cdot)}$ be the handover latency in the mobility management protocol that was developed in the previous section. The $(\cdot)$ protocol is used as the indicator, $E[L_{HO}^{(\cdot)}]$ is the average value of $L_{HO}^{(\cdot)}$, TR is the residence time on the network, and its probability density function is denoted by $f_R(t)$. Here, $L_{HO}^{(\cdot)}$ is assumed to be exponentially distributed by the accumulation function $F_T^{(\cdot)}(t)$. Hence, $L_{HO}^{(\cdot)}$ is the element blocking the handover, and the handover block potential ρ_b is expressed as follows.

$$\rho_b = \Pr(L_{HO}^{(\cdot)} > T_R) = \int_0^\infty (1 - F_T^{(\cdot)}(u)) f_R(u) d_u = \frac{\mu_c E[L_{HO}^{(\cdot)}]}{1 + \mu_c E[L_{HO}^{(\cdot)}]} \quad (6)$$

where μ_c is the percentage of networks passing through a boundary of the MN. If the MAG coverage area is circular, μ_c is calculated as follows [12].

$$\mu_c = \frac{2\nu}{\pi R} \quad (7)$$

where ν is the average speed of the MN and R is the radius of the MAG coverage area.

B. Numerical Results

Performance analysis is used with the following system parameter values:

h_{C-H}=4, h_{C-H}=6, h_{H-G}=4, h_{C-A}=4, h_{A-M} =1, $E(S)$=10, τ=20ms, n=3, L_f=19bytes, D_{wl}=[10,40]ms, D_{wired}=0.5ms, BW_{wired}=100Mbps, T_{L2}=45.33ms, T_{DAD}=1000ms.

In this analysis, ρ_f ranges from 0 to 0.7 in increments of 0.05. Figures 7 shows the comparison of the handover latency. The maximum value of ρ_f increases the probability of an error in a wireless link when a packet is transmitted. The number of retransmissions of the mobility signal increases and results in an increase of the handover latency.

 ISBN: 978-1-61208-428-2

As shown in Figures 7, the handover latency of each mobility management protocol is proportional to ρ_f.

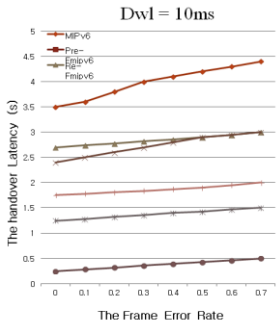


Figure 7. Handover Latency (ρ_f , D_{wl} =10ms).

Figure 8 (left) shows the transmission failure probability for v. As v increases, the MN must change rapidly. This means that the MN is required to complete the transfer of a high value in a shorter time. Therefore, as v increases, the transfer failure rate of the mobility management protocol also increases. In this analysis environment, if v is as high as 30 m, giving a handover probability of less than 0.05, only two predictive high-speed transport protocols, FMIPv6 and FPMIPv6, were able to function. Similar to the previous results, MIPv6 handover probability block performance was poor. This effect is notable as v increases. As seen in Figure 8 (right), most of the mobility management protocols are influenced by *R*. However, the performance of predictive FMIPv6 and FPMIPv6 is not affected. As with the results shown in Figures 8, the handover latency for predictive FMIPv6 and FPMIPv6 is short enough to avoid problems caused by v or *R*.

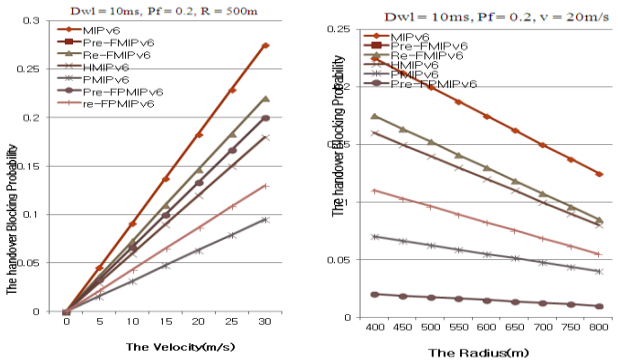


Figure 8. Handover Blocking Probability versus v (left) and *R* (right).

V. CONCLUSION

In this paper, we integrate QoS and mobility to control and manage the available resources effectively. This scheme has the advantage that the time latency is very small because of the added QoSB node. As shown in the results, the existing mobility protocols and the proposed scheme are analyzed with respect to handover latency, packet loss, and handover blocking probability in networks based on PMIPv6 and FPMIPv6. The evaluation results show a better overall performance for the FHO structure of mobility management schemes that is equally applicable in a network-based mobility management scheme. We can conclude that PMIPv6 and FPMIPv6 are the most efficient structures in many ways. In future, this approach will be applied to a variety of service platforms, such as Internet of Things (IoT) and verified in practical environments. We also plan to continue expanding its research scope.

ACKNOWLEDGMENT

This research was supported by Next-Generation Information Computing Development Program through the NRF funded by the Ministry of Science, ICT & Future Planning (NRF-2014M3C4A7030503) and Basic Science Research Program through the NRF funded by the Ministry of Education (NRF-2010-0024695).

REFERENCES

[1] J. H. Lee, J. M. Bonnin, and T. M. Chung, "Comparative Handover Performance Analysis of IPv6 Mobility Management Protocols." IEEE Tranaction on Industrical Electronics, Vol. 60, No. 3, 2013, pp. 1077-1088.

[2] V. Marques, X. P. Costa, and R. L. Aguiar, "Evaluation of a Mobile IPv6-Based Architecture Supporting User Mobility QoS and AAAC in Heterogeneous Networks," IEEE Journal on Selected Area in Communication, Vol. 23, No. 11, 2005, pp. 2138-2151.

[3] Zhou, H., Zhang, H. and Qin, Y, "An authentication method for proxy mobile IPv6 and performance analysis," Security Comm. Networks, 2009, pp. 445–454. doi: 10.1002/sec.83.

[4] H. W. Ko and J. P. Jeong, "dMMS: A Novel Distributed Dynamic Mobility Management Scheme for Minimizing Signaling Costs in Proxy Mobile IPv6 Networks," The Journal of The Institute of Webcasting, Internet and Telecommunication, Vol. 12, No. 4, 2012, pp. 65-80.

[5] S. H. Han and J. P. Jeong, "Intelligent Hierarchical Mobility Support Scheme in F-PMIPv6 Networks," The Journal of the KICS, Vol. 38, No. 4C, 2013, pp. 337-349.

[6] K. S. Go, U. S. Jung, and Y. S. Mun, "An Enhanced Fast Handover for Proxy MIPv6 Scheme for Efficient Mobile Environment of The Future Network," Journal of the institute of electronics engineerings of Korea, Vol. 48, No, 1, 2011, pp. 84-91.

[7] H. Yokota, K. Chowdhury, and R. Koodli, "Fast Handovers for Proxy Mobile IPv6," IETF RFC 5949, September 2010.

[8] R. Braden, D. Clark, and S. Shenker, "Integrated services in the Internet architecture: An overview," IETF, RFC 1633, June 1994.

[9] D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An architecture for differentiated services," IETF RFC 2475, December 1998

[10] C. Beaujean, N. Chaher, V. Marques, R. L. Aguiar, C. García, J. I. Moreno, M. Wetterwald, and T. Ziegler, "Implementation and evaluation of an end-to-end IP QoS architecture for networks beyond 3rd generation," in IST Mobile Summit, 2003, pp. 221–226.

[11] V. Marques, A. C. Casado, J. I. Moreno, and L. Rui, "A simple QoS service provision framework for beyond 3rd generation scenarios," in Proceeding of 10th International Conference on Telecommunications, 2003, pp. 1475–1481.

[12] R. Hsieh, Z.-G. Zhou, and A. Seneviratne, "S-MIP: A seamless handoff architecture for mobile IP," in Proceeding of INFOCOM, 2003, pp. 1774-1784.

 ISBN: 978-1-61208-428-2

Delay Constrained ARQ Mechanism for MPEG Media Transport Protocol Based Video Streaming over Internet

Hong-rae Lee, Tae-jun Jung, Kwang-deok Seo
Division of Computer and Telecommunications Engineering
Yonsei University
Wonju, South Korea
e-mail: kdseo@yonsei.ac.kr

Chang Ki Kim
Smart TV Media Development Team
Electronics and Telecommunications Research Institute
Daejeon, South Korea
e-mail: ckkim1@etri.re.kr

***Abstract*—MPEG Media Transport (MMT) is a new international standard aiming at addressing the emerging multimedia services over heterogeneous packet-switched networks including Internet and broadcasting networks. Due to the heterogeneous characteristics of the broadcast and broadband networks, MMT provides an efficient delivery timing model to provide inter-network synchronization, measure various kinds of transmission delays and jitter caused by the transmission delay, and re-adjust timing relationship between the MMT packets to assure synchronized playback. Exploiting the delivery timing model, it is possible to accurately estimate round-trip time (RTT) experienced during MMT packet transmission. Based on the measured RTT, we propose an efficient delay constrained ARQ (Automatic Repeat reQuest) scheme which is applicable to MMT protocol based real-time video communication over IP networks.**

Keywords-video communication over Internet; ARQ scheme; error control processing; MPEG Media Transport (MMT).

I. INTRODUCTION

In the recent years, digital broadcasting services and IP-based multimedia services over the Internet including mobile Internet have started being integrated and converged [1]. With this trend, there have been so many changes in the multimedia service environments such as media content delivery networks, diverse video signals, 4K/8K video transport systems, and various client terminals displaying multi-format signals. It has become clear that the MPEG standard has been facing several technical challenges due to the emerging changes in those multimedia service environments [2]. So, to address these technical challenges to existing and emerging MPEG standards, ISO MPEG has developed an MPEG-H standard suite (ISO/IEC 23008) for the delivery of audio-visual information compressed with high efficiency over heterogeneous environment. MPEG-H suite consists of three functional areas: High Efficiency Video Coding (HEVC) [3], 3D audio, and MPEG Media Transport (MMT) [4].

In order to deploy efficient solutions for the transport of HEVC video in an interoperable fashion, especially given the recent increased demand of multimedia delivery in the heterogeneous network environment, MPEG has launched a new standardization work item, called MMT since the middle of 2010. MMT has been working on addressing technical challenges of existing standards due to recent changes of multimedia delivery and consumption environments and new requirements from emerging use cases and application scenarios in the area of multimedia services [5]. MMT aims to overcome current limitations of available standards for media streaming by addressing streaming format that is transport and file format friendly, cross-layer optimized between video and transport layer, error resilient for MPEG streams, convertible between transport mechanisms and content adaptation to different networks [6]. The challenge to error control schemes for the real-time video communication is focused on how to endeavor to recover the packet loss and then to reduce impairment to the playback quality [7]. Many different error control techniques have been proposed to solve these issues and Automatic Repeat request (ARQ) has been known as one of the promising solutions [8].

MMT provides delivery timing model as a means to calculate jitter and the amount of delay introduced by the underlying delivery network, so that constant delay for data stream can be achieved [4]. Using the delivery timing model, an MMT receiving entity provides information in the feedback to allow the MMT sending entity to calculate the RTT. To circumvent the packet loss problem, MMT employs ARQ function as a basic error control technique in which the MMT receiving entity asks for retransmission when packet loss is detected at the receiver side [6][7]. When the RTT is so high that the retransmitted packet will not arrive in time, the retransmission will not improve the quality. The current ARQ function of MMT, however, could result in many late packets that arrive after play-out deadline because it does not consider round-trip delay caused for the retransmission. To circumvent this problem, we propose an efficient delay constrained ARQ scheme for MMT packet based real-time video communication.

The remainder of this paper is organized as follows. In Section II, we overview MMT technology and summarize the basic ARQ process supported in MMT. In Section III, we describe the proposed delay-constrained ARQ scheme. The experimental results and performance evaluation are presented in Section IV. Finally, concluding remarks are provided in Section V.

 ISBN: 978-1-61208-428-2

II. OVERVIEW OF MMT

In order to support efficient delivery and effective consumption of coded media data for multimedia services over packet-switched networks including IP networks and digital broadcasting networks, MMT defines three functional areas: encapsulation functional area, delivery functional area, and signaling functional area as illustrated in Figure 1 [4].

The encapsulation functional area defines the logical structure of media content, the Package, and the format of the data units to be processed by an MMT entity and their instantiation with ISO base media file format (ISOBMFF) as specified in ISO/IEC 14496-12. It produces Media Processing Unit (MPU) as an output.

The delivery functional area defines an application layer transport protocol including the payload format required for transferring encapsulated media data from one network entity to another. The payload format is defined to enable the carriage of encoded media data which is agnostic to media types and encoding methods.

The signaling functional area defines formats of signaling messages to manage delivery and consumption of media data.

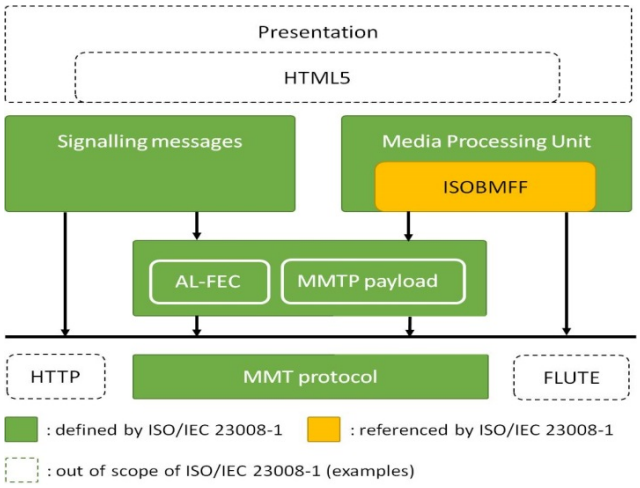


Figure 1. MMT functional areas and interface

The MMT technology adopts ARQ technique as an error control method for data transmission over error-prone networks. Like the general behavior of the ARQ scheme using a Negative Acknowledgment (NACK) [8], in case of packet loss, a NACK is sent back from the receiving entity to the sending entity and the sending entity retransmits the lost packet [6].

Figure 2 illustrates the basic operation of the MMT ARQ process [4]. The basic operation of ARQ process is as follows. The first step is the generation of ARQ configuration (AC) message by MMT sending entity, then it is delivered to the receiving entity. MMT receiving entity stores the ARQ configuration information. MMT receiving entity continues with receiving MMT media packets and checks whether there is a lost packet or not.

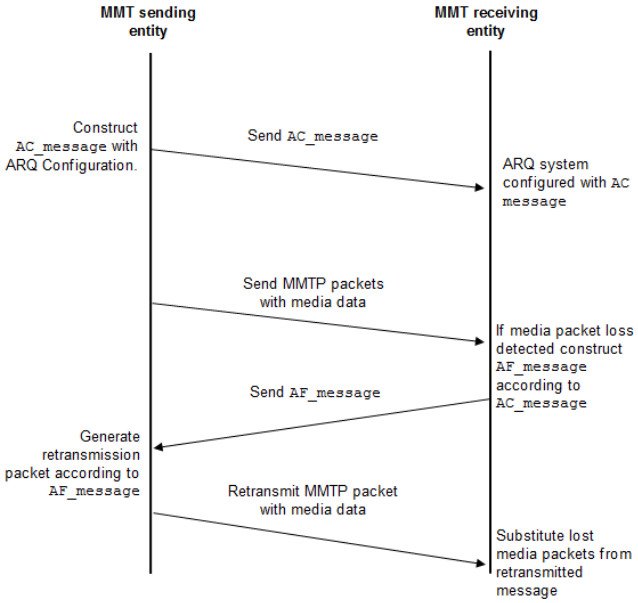


Figure 2. Operation of MMT ARQ process

Once MMT receiving entity determines that a packet has been lost, an ARQ feedback (AF) message is generated according to the configuration information defined in the AC message. The AF message is sent to MMT sending entity which will generate the retransmission packet to be sent to MMT receiving entity. MMT receiving entity is able to substitute the lost MMT packet with the retransmission packet.

III. PROPOSED DELAY CONSTRAINED ARQ MECHANISM FOR MMT PROTOCOL

One of the well-known problems in ARQ happens if networks experience severe congestion. In this situation, the retransmission packet leads more congestion and causes the further network degradation. Furthermore, in case of real-time video communication, the ARQ will be successful only if the retransmitted packet is received before its *arrival deadline*. When the RTT is so high that the retransmitted packet will not arrive in time, the retransmission will not improve the quality [7].

The objective of the proposed delay constrained ARQ is to suppress retransmitting requested lost packets that will not arrive in time for playback. By using the delay constrained retransmission, the MMT sending entity can avoid unnecessary retransmission of the out-of-date packets and therefore minimize the probability of wastefully retransmitted packets. This results in a reduced amount of data traffic wastefully injected into the network. The proposed delay constrained ARQ uses arrival deadline to

 ISBN: 978-1-61208-428-2

decide whether MMT sending entity retransmits the requested lost packet or not.

Figure 3 shows an exemplary timing diagram of the proposed delay constrained ARQ. In Figure 3, arrival deadline denotes the maximum tolerable latency for the requested retransmission packet to arrive at the receiver. If the retransmitted packet arrives later than arrival deadline, even if the packet arrives intact, it is regarded as useless and discarded so that decoding may not occur. Thus, the recovery latency for the lost packet should be taken into account for the real time video delivery wherein late repair becomes useless because of the real time nature of the data. RTT information which is crucial for the delay constrained retransmission is already available to the sender by the reception quality feedback (RQF) message supported in the MMT standard. This could be achieved by using the timestamp field existing in the MMT packet header. This timestamp field specifies the time instance of MMT packet delivery based on Coordinated Universal Time (UTC), which corresponds to the sending time at the first byte of MMT packet.

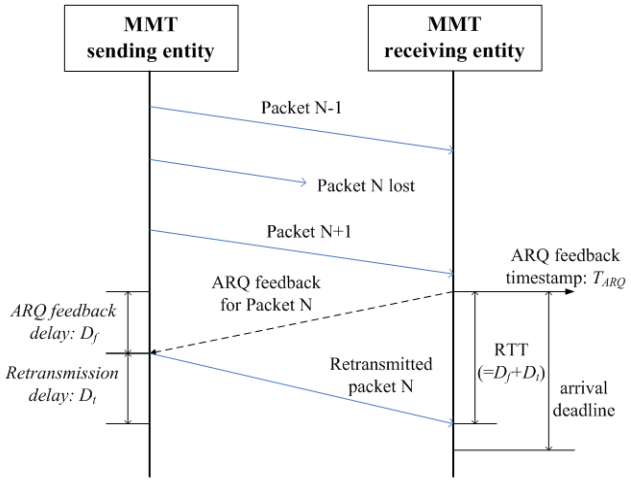


Figure 3. Timing diagram of the proposed delay-constrained ARQ

Based on the above analysis, when the MMT receiving entity detects the loss of packet, the receiver sends AF message to the MMT sending entity, and then the sender takes the decision procedure for delay constrained retransmission as shown in Figure 4 to decide whether to transmit the requested lost packet or not. The estimated RTT can be obtained *a priori* at the sending entity by RQF message provided in MMT [6]. For more accurate estimation of the RTT, up-to-date ARQ feedback delay (D_f) value can be informed to the sending entity by enclosing ARQ feedback timestamp (T_{ARQ}) into the AF message. ARQ feedback timestamp corresponds to the time instant of sending the AF message to the MMT sending entity. Using the ARQ feedback timestamp, T_{ARQ}, the sending entity can obtain up-to-date ARQ feedback delay (D_f). And this updated D_f value can be used to compute up-to-date RTT. Finally, if the RTT is estimated to be less than the arrival deadline, sender retransmits the requested lost packet to the receiver. Otherwise, the sender decides not to transmit the requested lost packet to the receiver. In order to implement the proposed delay constrained ARQ scheme, the ARQ feedback timstamp and the arrival deadline information needs to be included in the AF message of MMT. The AF signaling format supporting the proposed delay constrained ARQ scheme could be found in [9] and [10], and has been approved to be included in the standard document of MMT AMD1 (Amendment 1) in the MPEG meeting.

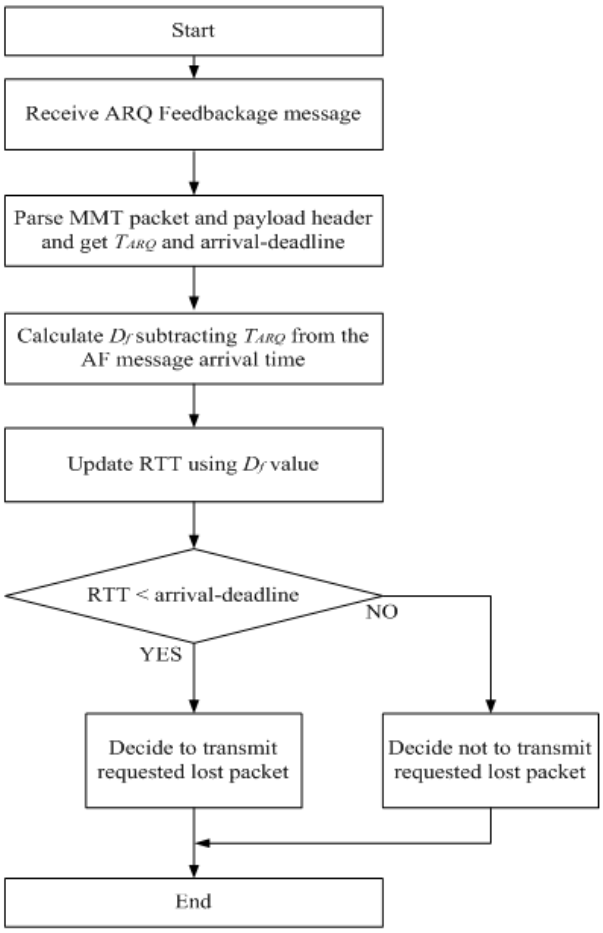


Figure 4. Procedure to decide whether to retransmit the lost packet at the MMT sending entity

The arrival deadline can be obtained at the receiver side by considering the remaining amount of safely arrived packets in the receiver buffer when the ARQ feedback message is prepared for sending. Therefore, we exploit the remaining amount of safely arrived packets in the receiver buffer and average bit rate of the MMT packet stream when the ARQ feedback message is prepared for sending.

 ISBN: 978-1-61208-428-2

IV. EXPERIMENTAL RESULTS

To evaluate the efficacy of the proposed delay constrained ARQ scheme, we performed extensive experiments over the MMT protocol-based HEVC video streaming system. The video streaming system consisted of an MMT sending entity and receiving entity. Before initiating video streaming, the AC message which includes the retransmission policy to be adopted by the MMT sending and receiving entities in the event of packet loss is sent to the receiving entity by the ARQ policy manager. The MMT sending entity retains an MMT packet in the buffer until the timeout, and it is thus available for retransmission. If MMT sending entity receives an AF message requesting retransmission of the lost packet, it checks whether the lost packets is available in the retransmit buffer and it retransmits the packet only if the arrival deadline value is greater than RTT. To verify the performance of the proposed ARQ scheme in an error-prone video transmission environment, the NIST-Net [11] Linux-based network emulation tool was used to emulate the packet-loss network environment. We used the *Stockholm* video sequence with HD 1080P resolution. The *Stockholm* video sequence was coded by an HEVC encoder, which is being increasingly used for broadcasting and mobile multimedia applications. The run-time of the generated video streaming was 4 min 39 s, and the number of generated MMT packets to be delivered was 193,042 in total. The size of each MMT packet was 1500 bytes, and the receiving buffer was set to store 300 MMT packets.

In the first experiment, the network delay was set to 50 ms while the packet loss rate has varied from 3% to 10% by the NIST-Net network emulator. Figure 5 shows the ratio of the number of renounced retransmission to the total number of retransmission request for various packet loss rates. We can observe that the ratio of the renounced retransmission increases as the packet loss rate increases. This is due to the fact that as the packet loss rate increases, the arrival deadline for the lost packet is generally decreased because of the less safely arrived MMT packets in the receiver buffer.

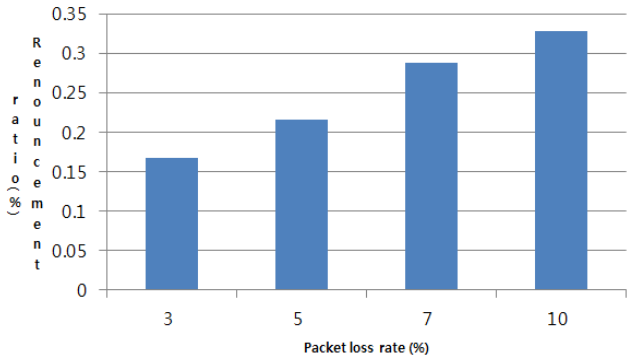


Figure 5. Ratio of renounced retransmission for various packet loss rates under fixed network delay of 50 ms

Figure 6 shows the saved network bandwidth resulted by the proposed delay constrained ARQ scheme. We can observe that the saved network bandwidth increases as the packet loss rate increases. As the packet loss rate increases, the ratio of the renounced retransmission also increases. This results in avoiding unnecessary retransmission of the out-of-date packets and therefore reduces amount of data traffic wastefully injected into the network.

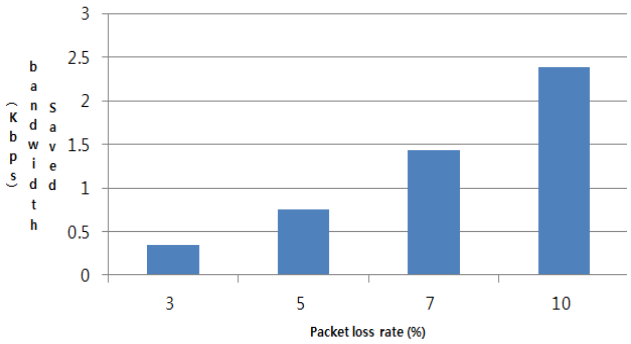


Figure 6. Saved network bandwidth for various packet loss rates under fixed network delay of 50 ms

In the second experiment, the packet loss rate was set to 10% and the network delay has varied from 100 ms to 400 ms by the NIST-Net network emulator. When compared to the simulation conditions for the first experiment, the channel condition significantly deteriorated, which showed a higher packet loss rate and much longer network delays experienced during the video streaming.

Figure 7 shows the ratio of the number of renounced retransmissions to the total number of retransmission requests for various network delays under the same test condition. It is evident that the ratio of the renounced retransmissions significantly increased as the network delays increased. This was due to the fact that, as the network delay increased, the arrival deadline for the lost packet generally decreased on account of fewer safely arrived MMT packets in the receiver buffer.

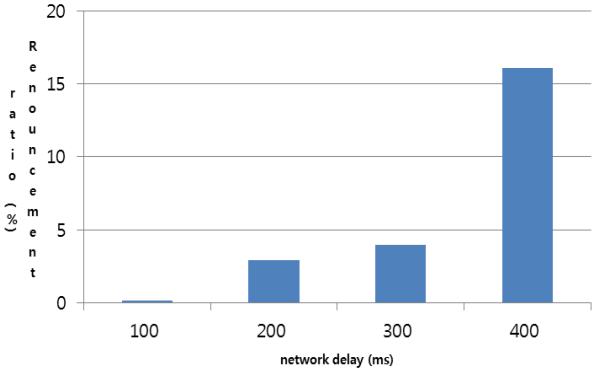


Figure 7. Ratio of renounced retransmission for various network delays under the fixed packet loss rate of 10%

 ISBN: 978-1-61208-428-2

Figure 8 shows the saved network bandwidth resulting from the proposed scheme. We can observe that the saved network bandwidth increased as the network delay increased. As the network delay increased, the ratio of the renounced retransmission also increased. This resulted in the avoidance of unnecessary retransmissions of out-of-date packets and therefore reduced a significant amount of data traffic wastefully injected into the network. As shown in Figure 8, approximately 190 Kbps were saved for the network delay case of 400 ms. This saved bit-rate amount corresponded to 13% of the total data traffic generated by the transmitted MMT packets.

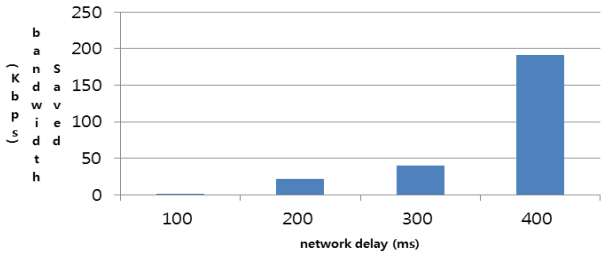
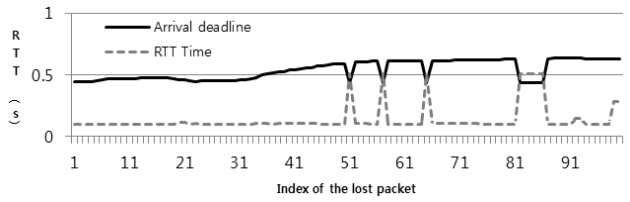
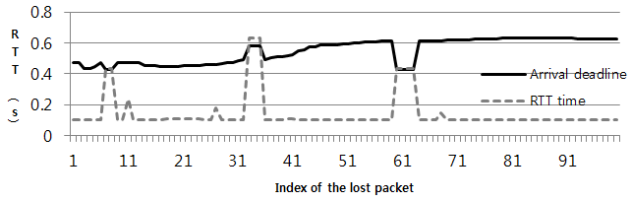


Figure 8. Saved network bandwidth for various network delays under the fixed packet loss rate of 10%



(a) packet loss rate of 5%



(b) packet loss rate of 7%

Figure 9. Comparison of RTT and arrival deadline for 100 randomly chosen lost packets during video streaming over various packet loss rates

In Figure 9, we compare RTT and arrival deadline values for one hundred randomly chosen lost packets during video streaming over packet loss rates of 5% and 7%. If the RTT is greater than the arrival deadline, the sending entity decides not to retransmit the requested lost packet to the receiving entity. Otherwise, the sending entity decides to retransmit the lost packet to the receiving entity. As shown in Figure 7, as the packet loss rate increases, the portion of the lost packet which has greater RTT value than the arrival deadline also increases.

V. CONCLUSIONS

In this paper, we proposed a delay constrained ARQ scheme to enhance the effectiveness of the basic ARQ process of MPEG MMT standard for video streaming over Internet. Using the proposed ARQ scheme, it is possible to avoid unnecessary retransmission of the out-of-date packets and therefore minimize the probability of wastefully retransmitted packets.

ACKNOWLEDGMENT

This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2015-H8501-15-1001) supervised by the IITP(Institute for Information and Communications Technology Promotion).

REFERENCES

[1] S. Aoki, K. Aoki, H. Hamada, Y. Kanatsugu, M. Yamamoto, and K. Aizawa, “A new transport scheme for hybrid delivery of content over broadcast and broadband,” Proc. Broadband Multimedia Systems and Broadcasting (BMSB), 2011, pp. 1-6.

[2] Y. Lim, K. Park, J. Lee, S. Aoki, and G. Fernando, “MMT: An emerging MPEG standard for multimedia delivery over the Internet,” IEEE Multimedia, vol. 20, no. 1, Jan. 2013, pp. 80-85.

[3] ISO/IEC 23008-2, High efficiency coding and media delivery in heterogeneous environments – MPEG-H Part 2: High Efficiency Video Coding (HEVC), 2013.

[4] ISO/IEC 23008-1, High efficiency coding and media delivery in heterogeneous environments – MPEG-H Part 1: MPEG Media Transport (MMT), 2014.

[5] Y. Lim, S. Aoki, I. Bouazizi, and J. Song, “New MPEG transport standard for next generation hybrid broadcasting system with IP,” IEEE Trans. Broadcasting, vol. 60, no. 2, Feb. 2014, pp. 160-169.

[6] ISO/IEC 23008-1 AMD 1, Additional technologies for MPEG Media Transport (MMT), 2014.

[7] Y. Wang, and Q. Zhu, “Error control and concealment for video communication: A review,” Proc. of the IEEE, vol. 86, no. 5, May 1998, pp. 974-997.

[8] M. Podolsky, S. McCanne, and M. Vetterli, “Soft ARQ for layered streaming media,” Journal of VLSI Signal Processing, vol. 27, 2001, pp. 81-97.

[9] ISO/IEC JTC1/SC29/WG11 m32363, Update of ARQ feedback message for delay-constrained retransmission in MMT, 2014.

[10] ISO/IEC JTC1/SC29/WG11 m35228, Update of Implementation guideline for delay constrained ARQ in MMT, 2014.

[11] NIST-Net, Software provided by National Institute of Standards and Technology (NIST), available on-line at http://www-x.antd.nist.gov/nistnet, [retrieved: December, 2014].

 ISBN: 978-1-61208-428-2

On Cost-Reduced Channel Changing for Mobile IPTV Services in LTE-Advanced Systems

Hyeungwoo Lee
Network Development Group
SAMSUNG SDS Co. Ltd.
Suwon, Gyeonggi-do, 443-822, Republic of Korea
e-mail: hw77.lee@samsung.com (zen016@naver.com)

Jae-Young Choi and Jongpil Jeong
College of Information & Communication Engineering
Sungkyunkwan University
Suwon, Gyeonggi-do, 440-746, Republic of Korea
e-mail: {jaeychoi, jpjeong}@skku.edu

***Abstract*— Due to the development of mobile communication technologies, multimedia services for mobile devices or other services have begun to be commercialized. In this paper, we propose a method that minimizes channel switching time by pre-computing the user's preferred channels in the LTE-Advanced System. The existing composition of the Multimedia Broadcast Multicast Service (MBMS) using the Internet Group Management Protocol (IGMP) transfers data using multicast from the Broadcast Multicast Service Center (BM-SC) to evolved Node B (eNB), and with broadcast from eNB to the User Equipment (UE). Therefore, in the case of channel switching, the IGMP Join report should be transferred to the BM-SC, where it then waits for the user's content to be retransferred from eNB. However, in the case of our proposed method, the subscriber channel technique of priorities is used to minimize the wait times between channel switching as the UE broadcasts its favorite channels in advance. Our performance is evaluated using mathematical modeling and shows that the delay time can be reduced from approximately 30 to 79% in terms of the channel switching time.**

Keywords-LTE-Advanced; MBMS; MBSFN; IGMP; Mobile IPTV; Content Database.

I. INTRODUCTION

The use of the Internet has increased greatly due to multimedia services. It is expected to increase by at least 1000-fold by 2020 [1]. In wireless environments, as the quality of the internet connection has improved, the demand for multimedia services such as Internet Protocol Television (IPTV), Video on Demand (VOD), and Voice over IP (VoIP), using a mobile IPTV is steadily increasing. IPTV is a communication and broadcasting convergence service based on IP that provides a broadcast service to the user terminal, and it can provide an interactive service according to the user's demands. In particular, mobile IPTV is mobility-granted to the existing IPTV. The Multimedia Broadcast Multicast Service (MBMS) initially appeared in Release 6. The MBMS was introduced in order to provide a service in Wideband Code Division Multiple Access (WCDMA). The most important technique of the MBMS is the Multimedia Broadcast Multicast Service Single Frequency Network (MBSFN) transmission scheme, which has been added in Release 7. The 3GPP has standardized tasks to provide Long-Term Evolution (LTE) standardization and evolved MBMS (eMBMS) in Release 8, as well as including eMBMS standards in Release 9. In addition, the eMBMS features (counting techniques to determine the number of user terminals interested in a multimedia broadcast, the broadcast service to support the mobility of the receiving user terminal technology) have been added in Release 10/11, and are progressing towards standardization in Release 12 [2].

The MBMS is an effective method for delivering multimedia content to multiple destinations by allowing for the sharing of resources more efficiently. The LTE-Advanced system has two methods for supporting the MBMS. One is a single-cell MBMS transmission method used to transmit the MBMS content to users in a single evolved Node B (eNB). Another is a multi-cell MBMS transmission method for simultaneously transmitting MBMS content to the users that belong to a group of eNBs. The MBSFN transmission scheme is greatly improved in terms of the spectral efficiency in order to improve the Signal to Interference-Plus-Noise Rate (SINR) with compared to the Universal Mobile Tele-communications System (UMTS) for MBMS. This is very useful for the cell edge, which is considered to be inter-cell interference in the UMTS, because the interference signal is decreased and the received signal is increased simultaneously. The MBSFN generally provides better performance compared to single-cell Point to Point (PTP) or Point to Multipoint (PTM) transmission [3]-[4].

As a result, the MBMS transmits content by linking one eNB to multiple UEs. If any UE requests are connected to the applicable eNB, even if the UE does not request the specific channel, along with the channel requested by the other UE, the MBMS can receive content due to the broadcasting method. When the UE changes channels, if the UE is already receiving the channel via broadcasting, there is no time delay related to the channel change. If not, it can be received after sending the desired channel to the Broadcast Multicast Service Center (BM-SC) via the Internet Group Management Protocol (IGMP) Join report. To support the improved content in LTE-Advanced/eMBMS environments, we propose a method for minimizing the channel switching time by using channel priority assignment interlinked with the Content Database.

This paper is organized as follows: Section 2 describes the related work. In Section 3, we describe the improvement of the content transmission using the channel priority scheme with an IGMP. Section 4 shows the numerical results with the analytical model, and finally we conclude with Section 5.

 ISBN: 978-1-61208-428-2

II. RELATED WORK

This section describes the LTE-Advanced System, Mobile IPTV and eMBMS.

A. LTE-Advanced System

LTE (Rel. 8/9) is based on the OFDM transmission scheme, and can be described as the evolution of 3G mobile communication technology, such as Code Division Multiple Access (CDMA), WCDMA and 3.5G High Speed Packet Access (HSPA). The LTE system operates at a rate three times higher than the 3.5G transmission system in the Downlink and the latest standard for LTE, LTE Advanced, seeks to achieve a rate more than twice as high as that found in LTE [5]-[8].

LTE and LTE-Advanced technologies increase the data transmission rate, and many technologies are being studied in order to support seamless service. Multiple Input Multiple Output (MIMO) technology in particular is being studied to make high transmission rate a core technology of LTE and LTE-Advanced systems. MIMO technology has an advantage in that there is no help from electric power or frequency, and it can obtain an antenna number proportionate to the channel capacity. Increasing the performance of mobile communication systems represents an important avenue of research. MIMO technology has rapidly evolved from Single-User MIMO to Multi-User MIMO, and recently to Mass-MIMO [14].

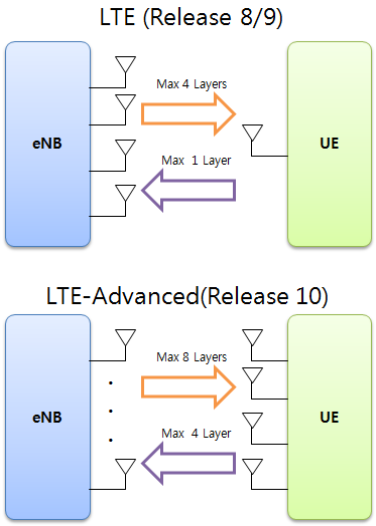


Figure 1. Antenna techniques of LTE and LTE-Advanced.

Figure 1 shows the antenna port and the spatial multiplexing layer in LTE (Rel. 8/9) and LTE-Advanced (Rel. 10) (Assuming that the UE has 4 receivers – 4 x 4). The downlink of LTE is up to support four transmit spatial layers, and the uplink (assuming that the eNB use the diversity 1 x 2) is up to support one per UE. LTE is not supported by a multi-antenna transmitter in order to simplify the baseline UE. However, multi-user spatial multiplexing (Multi-User MIMO) is supported. For MU-MIMO, the 2 UEs transmit in the same frequency and time, and eNB has to separate the UEs based on the spatial attributes. The gain of the uplink capacity, including multi-user spatial multiplexing, is generated. The maximum data transfer rate of a single user is not improved. To improve the peak data rate of a single user, in order to meet the requirements of the ITU-R (radio communication sector) for spectrum efficiency, LTE-Advanced may designate up to 8 downlink layers. This allows for 8x8 spatial multiplexing on the downlink where eight receivers are required for the UE. The UE has been specified to support up to four transmitters. Therefore, when combined with the four receivers, the eNB allows for the transmission of a 4x4 uplink.

B. Mobile IPTV

IPTV is a the TV service provider that provides services to the user via the IP network. In addition, the user has received the enhanced TV service instead of the traditional TV service through a technique called IP. Already we have been using a similar service, such as videos and web search using a computer or laptop via the Internet. The opportunity is in mobile IPTV, where users cannot watch TV services as one-way, but instead watch real-time two-way [13]. In addition, it is possible to use the enhanced TV service over the desired configuration. Moreover, once Korea had obtained popular mobile TV, such as DMB, it had a technique where you can always watch where you want. It is a concept known as Take-Out TV. However, this model of mobile TV technology extends the existing fixed-TV service, which had limitations with traditional TV services, where in particular there was still dissatisfaction with the one-way service [2]. In order to solve this problem, techniques for two-way service were introduced using mobile IPTV-based IEEE802.16, 3GPP and a variety of communication technologies (for example, 3GPP2 / BCMCS, Qualcomm / MediaFLO, IMNA / BCAST, etc.). Figure 2 shows the configuration of a mobile IPTV.

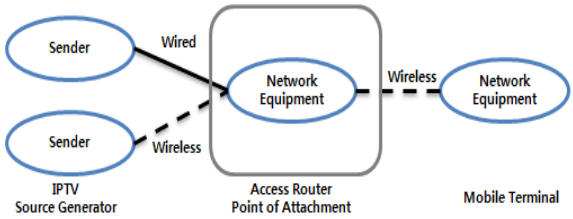


Figure 2. Configuration of mobile IPTV.

Mobile IPTV can support services anywhere, anytime and also supports seamless mobility in the wireless environment, even when moving to another radio link with a mobility supporting technology, such as the handover duration of the IPTV service. Especially if there are characteristics of a different radio link, technical issues can occur, but these technologies are expanding the research scope of the technology because they are able to expand the available range of mobile IPTV standardization that organizations are required to have [14]. In addition, the mobility studies in the Next Generation Network (NGN), ITU-T SG13 and SG19 in charge of the mobile

 ISBN: 978-1-61208-428-2

communication network standards fall jointly under the NGN-SGI (Global Standards Initiative) to form a group called Mobility Management, and the standardization progress and issues, such as Fixed Mobile Convergence (FMC), are also being studied at the same time [5]. Recently, mobile IPTV technology has been secured for the QoS and QoE of the content, security, interoperability and openness, anytime and anywhere, which aims to provide services across any device, through the combining of the next generation of web technology oriented user engagement and personalization and service convergence [16].

C. *eMBMS*

3GPP has introduced 3G / 4G user multicast information, as well as MBMS (Multimedia Broadcast / Multicast Service). MBMS may share resources efficiently, as it is an efficient method for providing multimedia content to multiple destinations. MBMS in the "LTE" context of 3G systems are evolving to become eMBMS [2]. LTE's eMBMS aims to provide broadcast and multicast service combining the high efficiency of the spectrum capacity. LTE-Advanced is designed to share services, such as digital video / audio broadcast.

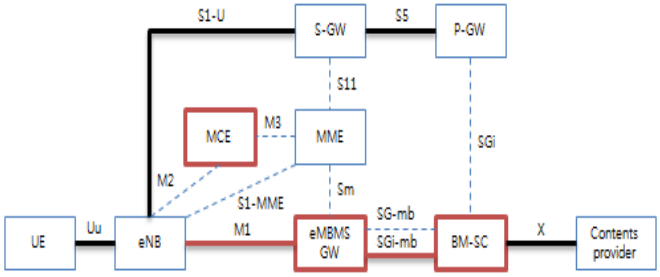


Figure 3. eMBMS architecture.

In Figure 3, eMBMS shows the UE how to perform the broadcast / multicast in the LTE-Advanced network [9]. It can send a media signal simultaneously to multiple receivers in the same area using the same eNB. In addition, the interactive features of the eMBMS system can dynamically interact with the broadcast network. To support the eMBMS service without having to change the entire structure of the existing service, the addition of a new node such as BM-SC can be supported. The eMBMS architecture is compatible with the SGSN and GGSN, such as 2G / GSM or the 3G UMTS packet core node based on the packet core domain. The eMBMS configuration of the 3GPP / LTE is shown in Figure 3. An eMBMS gateway exists between BM-SC and the eNB and performs functions such as MBMS packet transfer to the eNB by IP Multicast, user plane data header compression, and session control signaling (session start / stop). The MCE is responsible for allocating radio resources used by all eNBs belonging to one MBSFN area [15].

We first define a new interface SGi-mb and SG-mb for the BM-SC and MBMS Gateway communication. The SG-mb interface delivers authentication information and service management to the control plane. IP packets of the multimedia data are transmitted through the user plane interface's SGi-mb.

IGMP is split into two versions. It is a protocol for managing the large group membership of a multicast router and identifying a member of a multicast group on a network host. IGMPv1 is described in RFC 11/12 and IGMPv2 is described in RFC 2236. IGMPv2 is used by default in the router [10]. The operation is performed through joining, monitoring, member continuation, and leaving, as well as when a multicast group joins the router and sends a General Query message at 125-second intervals, and notifies the router to leave (IGMPv1 does not inform). Recently IGMPv3 has been under development by RFC 3376 [11]. An IGMP message is composed of 8 bytes, where the first four bits discriminate the IGMP version. Consisting of 16 bits in the checksum, there is a group consisting of a 32-bit address (IP address of Class D). The time to query is when the group address is set to zero, since that is when the report has a value that represents the address of the group.

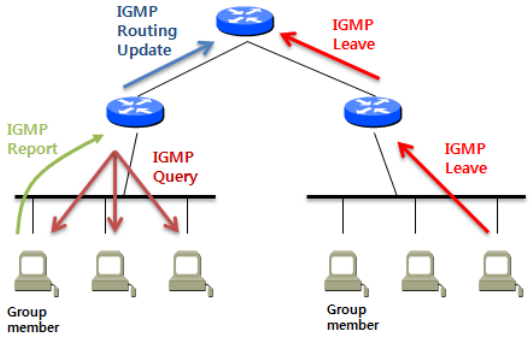


Figure 4. IGMP flow.

Figure 4 is a schematic view showing the basic operation of the IGMP. When the first process is to participate in the group, the host sends an IGMP Join message. The host does not send the report when the first process and last process leaves the group, while the router periodically sends an IGMP Query message to check whether the host is part of a group. The host then sends an IGMP Report message for each group belonging to its own process, and responds to the IGMP Query message.

Numerous standards and research papers for the Mobile IPTV service are presented and most of them address the cost by cell range. Our research, on the other hand, focuses on the improvement of the quality of user experience through minimizing the time of channel switching concerning MBMS service in LTE/LTE-A system.

III. SERVICE CONTROL OF CONTENT DATABASE

At present, several UEs are connected to one eNB, and the eMBMS transmits the content using a point-to-multipoint method. Therefore, the UE receives not only the requesting channels, but also unsolicited channels. Other UEs then request content in the same eNB. The channel change using IGMP is shown in Figure 5. After sending the IGMP JOIN report from the UE to BM-SC, the channel content packet may be transmitted from the eNB to UE.

 ISBN: 978-1-61208-428-2

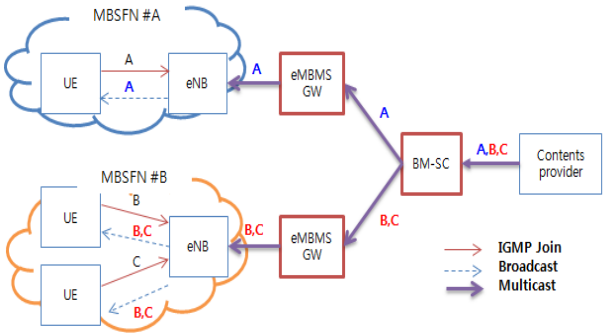


Figure 5. Channel selection procedure in eMBMS.

Figure 6 describes the process for channel change. If a UE changes from channel#1 to channel#2, the IGMP JOIN message for channel#2, the IGMP LEAVE message for channel#1, and transmits the content for channel#2, this is a source of service quality degradation due to the waiting time to perform the channel change.

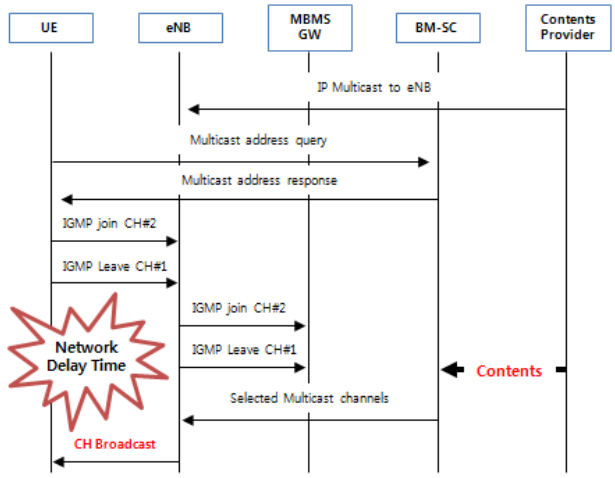


Figure 6. Flow of channel selection in mobile IPTV.

As shown in Figure 6, the channels are transmitted by multicasting from the content provider to the eNB, and then the channels requested by the UEs are transmitted by broadcasting from the eNB to UE. Therefore, when the UE changes the channel for new contents, if the UE already receives the new contents representing the requested information by broadcast message, we assume that there is no delay time for the channel change. However, there cannot be an inefficient waste of broadcast packets for all channels regardless of the subscriber. We propose a method of broadcasting with respect to the expected channel to be used.

Figure 7 shows a representation where a database node for the user to analyze the existing eMBMS system has been added. When the channel change request comes from the UE, if the channel is currently being broadcast, it sends the content of the immediately changed channel. If the channel is not being broadcast, it sends a message to the BM-SC requesting the content of the requested channel. This is the same method as the existing transmission system. Since the BM-SC has a preference pattern for UE channels, it is transmitted to the UE by the broadcast channel with selected channels that are decided by the BM-SC / content database. This method eliminates the delay time according to the channel change of the UE. The communication between the content provider / content database and BM-SC using this method is performed in the system backend, which reduces the burden of the UE and does not affect the existing channel switching time.

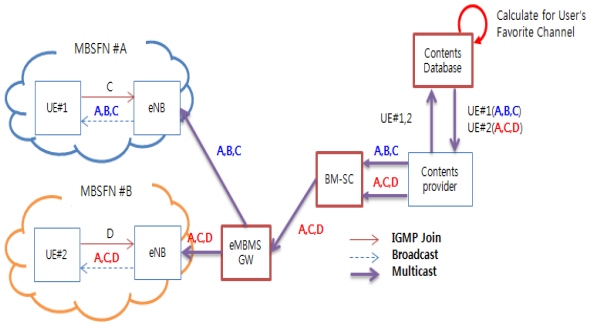


Figure 7. User rating channel database additional configuration.

IV. PERFORMANCE EVALUATION

In this section, we perform a mathematical analysis to assess the performance of MBMS channel switching time in terms of the delay time, and present some numerical results to evaluate the comparison of the channel change.

A. *Modeling Time Cost Analysis*

Table I describes the parameters which were evaluated in the case with and without a subscriber's favorite channels for the channel change time.

TABLE I. PARAMETERS USED IN NUMERICAL ANALYSIS

Parameter	Description
I_{U_u}	IGMP messages via the Uu interface
I_{U_u-BC}	Air channels transmitted by a broadcast message
I_{M1-SG}	M1/SG-mb interface using IGMP message (M1: Interface among BM-SC↔GW↔eNB)
$I_{U_u-query}$	Multicast Address Query messages via the Uu interface
I_{query}	Multicast Address Query message between eNB and BM-SC
I_x	Subscriber-specific favorite channel request / response messages
I_{x2}	Subscriber-specific favorite channel request / response messages
$I_{x2-Sort}$	Calculated per subscriber channel preferences at the content database
I_{mc}	Multicast message from the BM-SC
I_{MBMS}	Total time of MBMS

Hence, The MBMS time is calculated as follows;

$$I_{MBMS} = K_{air} \cdot \left[2(I_{U_u}) + I_{U_u-BC}\right] + K_{core} \cdot \left[2(I_{M1-SG}) + I_{mc}\right]. \quad (1)$$

Using improved functionality, the MBMS time can be classified into two. The first is when it contains a subscriber's favorite channels when broadcasting channel information:

$$I_{MBMS-with\_CH} = K_{air} \cdot \left[2(I_{U_u}) + I_{U_u-BC} \right] + K_{core} \cdot 2(I_{M1-SG}) \quad (2)$$

The second is when it does not contain a subscriber's favorite channels in broadcasting channel information:

$$I_{MBMS-with\_CH} = K_{air} \cdot \left[2(I_{U_u}) + I_{U_u-BC} \right] + K_{core} \cdot \left[2(I_{M1-SG} + I_x + I_{x2}) + I_{x2-sort} + I_{mc} \right] \quad (3)$$

K_{air} is a proportion of the air interface and K_{core} is a proportion of the core interface ($K_{air} + K_{core} = 1$) [12]. In the LTE system, we made calculations by setting a different weight for the latency time. However, the actual value of the air interface and the system performance is determined.

B. Numerical Results

Based on the analysis given so far, we compare the performance of the existing and proposed schemes. For numerical analysis, we configure the parameter values, as described in Table II, in which some of the values are taken from [12]-[16].

TABLE II. PARAMETER VALUES FOR PERFORMANCE EVALUATION

Parameter	Values
K_{air}	$0 < K_{air} < 1$, (if $K_{air} + K_{core} = 1$)
K_{core}	$0 < K_{core} < 1$, (if $K_{air} + K_{core} = 1$)
I_{U_u}	0.5
I_{U_u-BC}	0.5
I_{M1-SG}	2.5
$I_{U_u-query}$	0.5
I_{query}	2
I_x	1
I_{x2}	1
$I_{x2-Sort}$	1.5
I_{mc}	2.5

K_{core} specifies the proportion of a real-world environment to be larger than K_{air}. The content database node increases and the subscriber may encounter a delay according to the calculation for the wireless channel. In this paper, the weight of K_{air} and K_{core} was confirmed by setting three assumptions.

If there are none of the subscriber's favorite channels, as shown in Figure 8, the proportion of K_{air} and K_{core} may be increased compared to the conventional MBMS channel switching time if the subscriber's preferred channel is included (if it is broadcast). Accordingly, the proportion of K_{air} and K_{core} can confirm the degree of improvement from 22.7 to 30.7% compared to the existing procedures. Of course, the delays are reduced when the proportion of K_{core} is large.

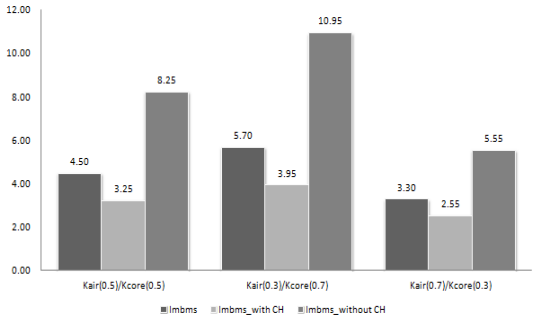


Figure 8. Comparison of the channel change.

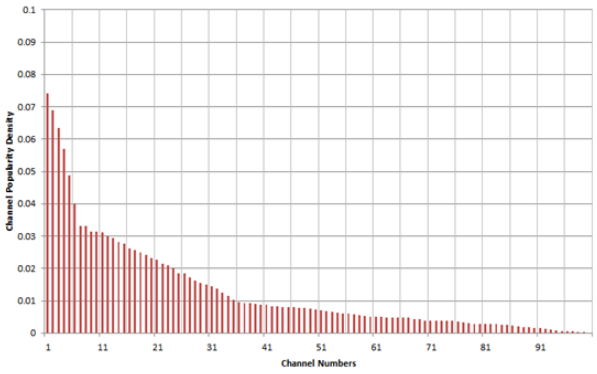


Figure 9. Channel popularity density [13].

According to [13], with respect to channel 96 for the popularity investigation result of the TV viewing conditions, $0 \le X \le 10$ is 45% as shown in Figure 9, $0 \le X \le 15$ is 65%, and $0 \le X \le 20$ indicates a probability of 70% for 5 or more channels if the shared content can be sent without channel change time.

The channel change time for the subscriber's favorite channel is analyzed. Figure 10 shows that it is possible to reduce the delay time from 57.5 to 61.8% in the case of $0 \le X \le 10$. Figure 11 shows that it is possible to reduce the delay time from 76.8 to 79.2 % in the case of $0 \le X \le 20$.

If the most popular 96 channels are shared in a service node, the probability of a viewer choosing among the channels is 0.618. Figures 10 and 11 show the cumulative channel duplication probability of 10 independent trials when total numbers of most popular shared channels are 10 and 20. Notice that the newly proposed method is able to let viewers immediately watch on the shared mobile IPTV without channel change time using the enhanced group join and leave process for LTE / LTE-A systems, when more than 5 channel requests are made for most popular channels.

 ISBN: 978-1-61208-428-2

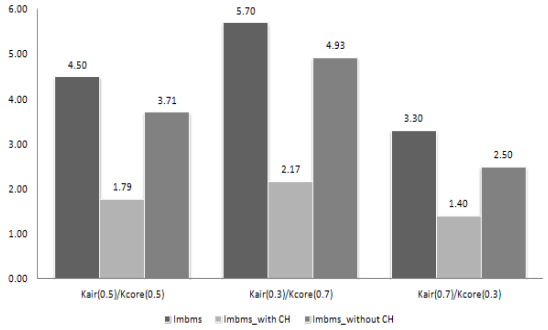


Figure 10. Comparison of channel change (*0≤ X ≤10*).

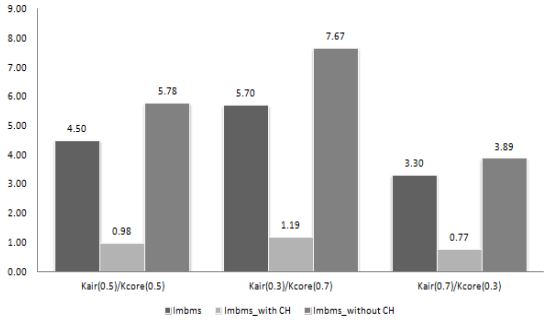


Figure 11. Comparison of Channel Change (*0≤ X ≤20).*

V. CONCLUSION

In this paper, we analyze channel change techniques and conduct a performance evaluation for mobile IPTV in LTE-Advanced environments. The proposed method, which is a curtailment of channel change time, can be effective for eliminating the delay time by broadcasting from the eNB in advance after analyzing the channel pattern of the accessing user in the content database. This method also reduces the burden of the UE and doesn't affect the existing services at all because it is performed in the back end of the BM-SC and content database system after the user initially connects. For the performance evaluation, we analyze the transmission of the content without the IGMP message using the duplication probability of the selected channel for the shared channel. If the computational accuracy of the subscriber priority channel in the content database is able to be increased, we can expect to provide a better service for MBMS and improve the performance of MBMS.

In future work, we will seek to increase the accuracy of the method for calculating the preferred channels of the subscriber in the content database.

ACKNOWLEDGMENTS

This research was supported by the Next-Generation Information Computing Development Program through the NRF funded by the Ministry of Science, ICT & Future Planning (NRF-2014M3C4A7030503) and Basic Science Research Program through the NRF funded by the Ministry of Education (NRF-2010-0024695).

REFERENCES

[1] Chih-Lin, H. Jinri, D. Ran, C. Chunfeng, J. X. Jiang, and L. Lei, "Recent Progress on C-RAN Centralization and Cloudification" IEEE Access, Vol. 2, 2014, pp. 1030-1039.

[2] J. S. Kim and S. H. Moon, "Evolved-MBMS: Mobile IPTV Technologies for 3GPP LTE," Information & Communications Magazine, Vol. 30, No. 2, 2013, pp. 66-74.

[3] R. Trestian and G. M. Muntean, "Wireless Multi-Access Environments and Quality of Service Provisioning," IGI Global, 2012.

[4] A. Alexiou, C. Bouras, V. Kokkinos, and G. Tsichritzis, "Performance evaluation of LTE for MBSFN transmissions," Wireless Networks, Vol. 18, No. 3, 2012, pp. 227-240.

[5] B. H. Shim and B. J. Lee, "Evolution of MIMO Technology," The Journal of Korea Information and Communications Society, Vol. 38A, No. 8, 2013, pp. 717-723.

[6] D. Martin-Sacristan, J. F. Monserrat, J. Cabrejas-Penuelas, D. Calabuig, S. Garrigas, and N. Cardona, "On the way towards fourth generation mobile: 3GPP LTE and LTE-Advanced," EURASIP Journal on Wireless Communications and Networking, Vol. 2009, 2009, pp. 1-10.

[7] V. Stencel, A. Muller, and P. Frank, "LTE Advanced - A Further Evolutionary Step for Next Generation Mobile Networks," in proceeding of 20th International Conference on RADIOELEKTRONIKA, 2010, pp. 1-5.

[8] B. Y. Cho and J. Y. Kim, "On the Feasibility of Dual Mode LTE-FDD/TDD and Dual Mode WiMAX-TDD/LTE-TDD," The Journal of the Institute of Internet Broadcasting and Communication, Vol 10, No 5, 2010, pp. 45-55.

[9] 3GPP, "Multimedia Broadcast/Multicast Service (MBMS); Architecture and functional description," 3GPP TS 23.246 Release 13, 2015.

[10] H. Holbrook, B. Cain, and B. Haberman, "Using Internet Group Management Protocol Version 3 (IGMPv3) and Multicast Listener Discovery Protocol Version 2 (MLDv2) for Source-Specific Multicast," IETF RFC 4604, August 2006.

[11] B. Cain, S. Deering, I. Kouvelas, B. Fenner, and A. Thyagarajan, W. Fenner, "Internet Group Management Protocol, Version 3," IETF RFC 3376, October 2002.

[12] A. Alexiou, C. Bouras, V. Kokkinos, and E. Rekkas, "Evaluation of the Multicast mode of MBMS," in Proceeding of The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07), 2007, pp. 1-5.

[13] E. J. Lee, S. K. Park, and H. S. Lee, "Cross-Layer Design of Internet Group Management Protocol for Mobile IPTV Services in WiMAX," in Procceding of 2010 Fifth International Conference on Systems and Networks Communications (ICSNC), 2010, pp. 294-299.

[14] Xiangyang Wang, Tao Jiang, Yin Xu and Dong Wang, "A downlink group pre-coding scheme for MBMS in Distributed MIMO system," Broadband Multimedia Systems and Broadcasting (BMSB), 2014 IEEE International Symposium on, 2014, pp. 1-4.

[15] Baburao Damerar, Chitti Babu, "Broadcast the User Chosen Content over LTE in a Simplified Approach Using Evolved Multimedia Broadcast Multicast Service (eMBMS) with Optimized MCE Scheduling Algorithm," Intelligent Computing Applications (ICICA), 2014 International Conference on, 2014, pp. 90-94.

[16] Junyu Lai, Guolin Sun, Gun Li, Yuankai Li, and Kaiyu Qin, " An Active Channel Changing Delay Method to Improve IPTV Channel Availability," Computing, Communication and Networking Technologies (ICCCNT), 2014 International Conference on, 2014, pp. 1-7.

 ISBN: 978-1-61208-428-2