A Stream-Oriented Community Generation for Integrating TV and Social Network Services

Riho Nakano
Faculty of Policy Management
Keio University
Fujisawa, Kanagawa, Japan
e-mail: s10578rn@sfc.keio.ac.jp

Shuichi Kurabayashi
Faculty of Environment and Information Studies
Keio University
Fujisawa, Kanagawa, Japan
e-mail: kurabaya@sfc.keio.ac.jp

***Abstract*— This paper proposes a stream-oriented community generation that associates real-time TV streams and Web resources by analyzing the communication among users on a social network service (SNS). The aim of this system is to provide a novel media environment for enhanced cross-media communication and discussion by dynamically creating social communities according to the real-time contexts of TV stream. The unique feature of this system is an implicit community analysis mechanism that employs the TV stream as a powerful and well-organized "context-creator" for SNS users. This system extract a group of viewers who have the same or similar interests by integrating the term co-occurrence statistics of SNS messages and their synchronicity to TV. To detect the context-dependent group of users, this system provides a dynamic feature keyword selection mechanism to create a vector space, which is specifically tailored to the TV context. The application scope of this system includes analysis of community-level sentiments in SNS messages associated with a TV program and the analysis of transitions in the sentiments of communities to develop effective advertising strategies.**

Keywords- Community Generation; Social Network Services; TV streams; Cross-media Infrastructure; Sentiment Analysis.

I. INTRODUCTION

With the popularization of Internet-enabled TVs and smartphones, the demand for integrating synchronous TV with asynchronous SNSs has increased because their relationship is complementary [1, 2]. TV motivates viewers to interact with each other by generating common interests among them, and SNS enables this interaction. Integrating TV with interactive services increases its value [3].

Although TVs and SNSs are popular information resources, studies on the implicit community analysis and community creation by integrating these media have been limited. This is because the current Internet TV (ITV) technologies focus on the integration of video streams with textual information retrieved from the Web by developing screens that are capable of displaying Web contents. Therefore, cross-media communication infrastructure is essential to get the integrated sentiment information and its context which are present in a fragmented and closed manner.

In this paper, a stream-oriented community generation is proposed to create a group of viewers, i.e., a community of users, who have common interests by integrating real-time TV streams and SNS messages. "Community Generation", which is used to widen the scope of communication by adopting a dynamic configuration, is the key concept in such communication infrastructure [4]. Community Generation enables viewers to broadcast messages about the current TV program to the appropriate audience. In order to realize community generation, this system extracts an implicit structure of viewers who have the same or similar interests by analyzing their comments about a TV stream. The association between a community of TV viewers and a community of SNS users are established according to the information appearing on both the TV and the SNS. This system automatically extracts a "community of interest" (CoI) structure from the SNS messages and TV program guides. Unlike Google TV that uses a single display for synchronous TV streams and asynchronous SNS messages, this system enables cross-media communication between the two information resources.

The aim of this system is to provide a novel media environment where a TV stream and related messages are exchanged seamlessly according to its context. As shown in Figure 1, this system uses the TV stream as a powerful and well-organized "context-creator" for SNS users. The context-creator affects a vast array of users by posting the same content at the same time over many SNSs. By introducing TV as a context-creator into SNS community analysis, the system recognizes the topic in the SNS messages by leveraging the powerful context created by the TV.

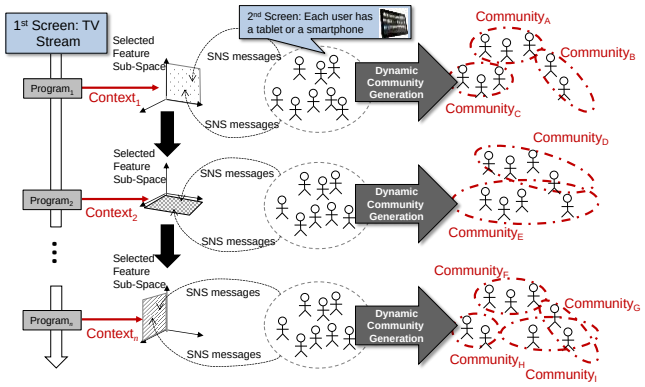

Figure 1. Fundamental Concept of Stream-Oriented Community Generation.

The most important advantages of this system are recognition of overlapping structures among communities and tracking viewers' transitions among multiple communities by detecting the sentiments expressed in SNS messages related to a TV stream. This tracking mechanism

 ISBN: 978-1-61208-200-4

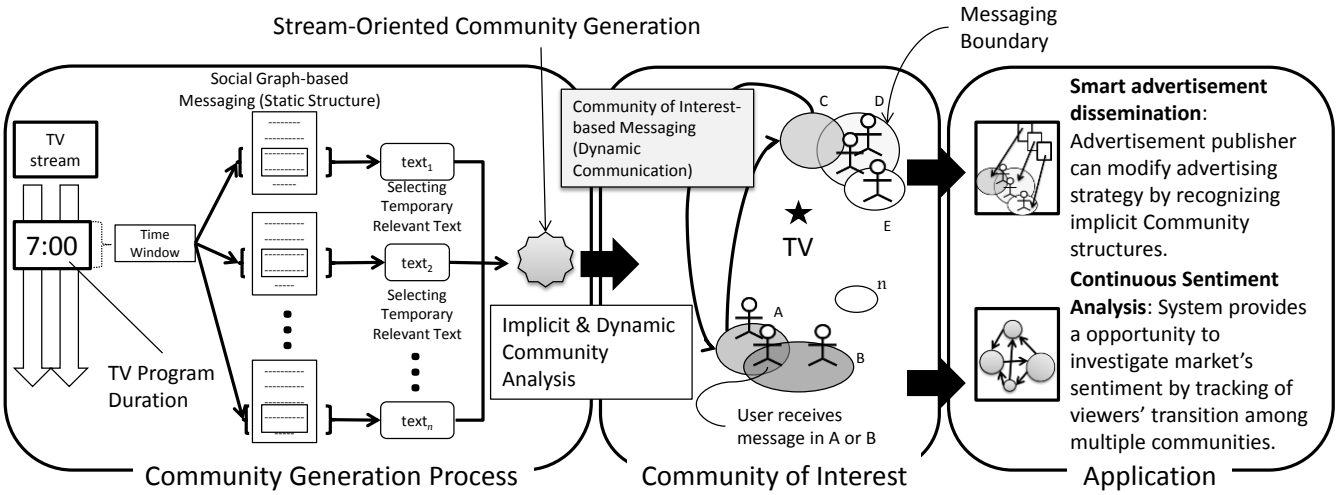

Figure 2. System Architecture of Stream-Oriented Community Generation.

focuses on the community-level viewing history, rather than a personal-level viewing history, in order to detect changes in a community's interests. The effect of a TV program can be assessed by analyzing the community-level sentiments using this tracking mechanism, and this information can be used to design TV program guides. This system reveals the transitions in the sentiments of communities, i.e., how many viewers change their opinion and how drastically they do this over the duration of a program.

The remainder of this paper is structured as follows. Section 2 presents motivating example of our community generation. Section 3 summarizes the related work briefly. Section 4 describes fundamental concept and system architecture. Finally, Section 5 concludes this paper.

II. MOTIVATING EXAMPLES

Our community-oriented community generation system's unique feature is community-of-interest-based communication mechanism that allows viewers to publish their comments about a TV program. This mechanism provides a context to the fragmented messages on the Web, which is an asynchronous medium, through TV, which is a synchronous medium. Thus, in our ITV model, TV —the 1st screen— provides contexts to the interaction on the internet-enabled second screen —smartphone or tablet. This mechanism enables the community generation system to track the sentiments of viewers about a TV program. This feature realizes the following two applications.

1) Smart advertisement dissemination: The conventional advertising strategies can be made more efficient by targeting users in a community according to their specific and common interests rather than age group or sex. Our system recognizes community structures on the basis of users' interests rather than explicitly defined community characteristics such as the follower/followee ratio. As our implicit community structures are highly dynamic and flexible, advertisement publishers can select a target community structure community structure such that the advertisement is more relevant to the targeted consumers. From users' perspective, the community structure can help the users to filter out irrelevant advertisements.

2) Continuous Sentiment Analysis: This system provides an opportunity to investigate the market's sentiment by tracking of viewers' transition among multiple communities. Our system provides "live" feedback from SNS by capturing the structure of implicit communities and their sentiment continuously. Even if the member of community is changing, the system tracks such transition and generates an appropriate community.

III. RELATED WORK

Since the invention of the Internet TV (ITV), there have been many attempts to integrate TV and new media. Tuomi [5] modeled the ITV as a substitution for a PC connected to Internet. Several methods have been developed for simultaneously displaying a TV stream and the related Web content by recommending Web resources as per the content of the TV stream [6][7]. As an advertisement recommendation method for Mobile TV, Adany et al.[8] has proposed a sequential solution procedure and several heuristic algorithms for uncertain personal advertisement allocation. TV2Web [9] is a seamless cross-media user interface that can be moved between TV screens and Web pages by linking units and displaying them smoothly using zooming metaphors. Many TV-on-Web methods have been proposed for displaying detailed information or service-related TV programs [9]-[11]. Another approach to integrate TV with the Web is the interactive approach, wherein the contents or the topic of a specific program are defined by [6] exchanging information or chatting [9] with other viewers in real time. Further, human-computer-interaction systems for TVs, PCs, and mobile devices have been proposed in [12].

IV. STREAM-ORIENTED COMMUNITY GENERATION

Figure 2 shows the system architecture of the community-oriented community generation system. This system extracts sentiments and keywords by analyzing a viewer's messages

 ISBN: 978-1-61208-200-4

about a TV stream. The extracted keywords and sentiment information are integrated with the TV stream and its details obtained using an interactive program guide (IPG). This process provides "contexts" to the Web resources. This context information is used to analyze the implicit community structure that does not involves any explicit IDs or tags such as hashtags in Twitter and group functions supported by Facebook. This section describes the fundamental architecture and functions of the stream-oriented community generation.

4.1 Data Structure

Our system uses three types of data structures.

- Text Messages Exchanged on SNS: Text messages are contents posted by SNS users. The system analyzes these text messages. A text message is a tuple consisting of timestamp information, and the user's identifier, and the text content. A set of messages is used as a primary data structure for sentiment analysis and consists of messages. Each message in a set of message is sequentially ordered according to the timestamp information in each message.
- TV Stream: The TV stream data is prefetched from an IPG and has information of the start and end time of a TV program respectively and keywords used for annotating the TV program. The annotation consists of the title, contents, performers, and word groups related to the TV program.
- Community: The system evaluates the word frequency and the sentiment information in each message in order to generate this community dynamically according to the messages or information posted by users sharing the same or similar interests. A community consists of the user ID derived from the SNS user set and keywords. Keywords include nouns and adjectives that represent users' sentiments. The system deals with the community data and text massage data as vectors in a feature vector space. Our feature vector space consists of keywords describing the TV contents and SNS messages and is a typical high-dimensional vector space. The system employs a dynamic feature selection mechanism to create the appropriate context vector space for measuring the distance between (1) a community and a message, (2) two communities, and (3) two messages.

4.2 Dynamic Feature Selection for Analyzing Communities

In order to analyze the implicit communities, our system employs a dynamic feature selection mechanism that creates a low-dimensional vector space for each community. As a fundamental data structure, this system provides a high-dimensional vector space (e.g., 3,000 dimensions) that consisting of the necessary and sufficient number of feature keywords. This full-space can be used to represent various messages and communities, but it is difficult to precisely identify the data items in the space. Each dimension in the full-space corresponds to a specific keyword corresponding to the topics in TV stream. Our concept of feature selection is that a community should be generated by considering its own context, because a different context gives a different meaning in the social network. The system does not define the relevance between a context and a community statically.

Thus, this system creates sub-space according to the community. Figure 3 shows the results of a feature selection table that defines a "context keyword" as a trigger for selecting features. This table defines the correlation between the context keyword with topic made through TV program and the feature keyword derived from the messages exchanged through an SNS. This function eliminates the ambiguity of context that is caused by summarizing an entire users' context into a feature vector.

TABLE I. FEATURE SELECTION TABLE

	Feature Keyword 1	Feature Keyword 2	…	Feature Keyword *n*
Context Keyword 1	0	1	⋯	1
Context Keyword 2	1	0	⋯	1
…				
Context Keyword *m*	0	0	⋯	0

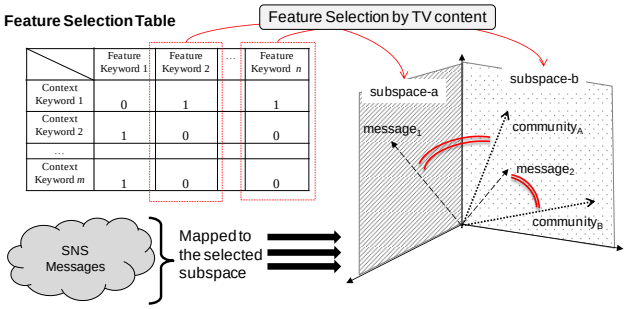

Figure 3. Feature Selection Operation for Creating Low-dimensional Vector Space to analyze Community Structure.

TABLE I shows the feature selection table. Here, the feature keyword is generated using a Web dictionary and has an enough capability to explain a user's sentiment and interest. Context keywords play a role to define the context and to recall a set of feature keywords in a specific context. The feature selection process is as carried out follows:

- Step-1: Find the context keyword from the IPG data, and generate a weighted term-frequency context vector.
- Step-2: Transmit the context vector to the feature vector space by using the feature selection table. This transmission function $f_{map}(v,x)$ that inputs a context vector v and feature selection table x is defined as follows:

$$f_{map}(v,x) := \left(\sum_{i=0}^{m} v_{[i]} \cdot x_{[0,i]}, \cdots, \sum_{i=0}^{m} v_{[i]} \cdot x_{[n,i]} \right)$$

where $v_{[i]}$ is the value of the i -th context keyword and $x_{[n,i]}$ is the value of the n-th feature keyword corresponding to the i-th context keyword. n denotes the number of context keywords in the full-space. The context feature weights are converted into feature keyword weights by using this function.

● Step-3: Select the top-k relevant feature keywords as the axis of the subspace, and compute the inner product between the vectors of community or SNS messages. This function $f_{distance}(r_1, r_2)$ that calculates the relevance between r_1 vector and r_2 vector is defined as follows:

$$f_{distance}(r_1, r_2) \coloneqq \sum_{i=0}^{n} r_{1[i]} \cdot r_{2[i]}$$

where r_1 and r_2 denotes a vector of community or a SNS message, and *n* denotes the number of features in the full-space.

● Step-4: Map every SNS message and community in the created subspace to compute the community-based relevance. The following function $f_{distance_{sub}}(r_1', r_2')$ is used to calculate the relevance by using the sub-space:

$$f_{distance\_sub}(r_1', r_2') \coloneqq \sum_{i=0}^{q} r'_{1[i]} \cdot r'_{2[i]}$$

where r_1' and r_2' denote vectors mapped into the subspace and q denotes the number of features in the subspace.

4.3 Function for Analyzing Implicit Communities

The main functions of this system can be divided into the following four categories: (1) time interval selection according to the context of the TV program, (2) community generation to create groups of users that share the same or similar interests, (3) community-of-interest(CoI)-based messaging, and (4) community transition analysis.

● Time Selection Function: Using the time interval selection function, the system extracts a set of messages from all the messages according to the time window defined for the TV program.

● Community generation Function: A set of communities is generated by morphologically analyzing the keywords for the current TV program. The community generation function generates communities according to the set of messages selected by the time section function and the candidate community to be mapped to the vector space.

● Community-of-Interest-based Messaging-Function: Using this function, the system delivers messages about a TV program to the relevant users. This messaging function involves a rather complex transfer process but provides users with an open and easy-to-use messaging platform. In concrete terms, the system selects target community from the set of all communities, and distributes the target community a SNS message.

V. CONCLUSION AND FUTURE WORKS

This paper proposes a stream-oriented community generation that associates real-time TV streams and Web resources by analyzing the communication among users on an SNS. Our approach realizes a novel communication medium that integrates messaging systems and TV streams according to viewers' interests. As a future work, we plan to implement this system on the 2nd screen (tablet devices) and perform experimental studies to demonstrate the effectiveness of our approach. We will also extend a range of association by not only focusing on the community but also handling other levels of communication scope, such as "Community & Individual Viewer," "TV-Stream & Web-Resources," "Individual Viewer & TV-Stream," and so on.

ACKNOWLEDGMENT

This research was supported by SCOPE: Strategic Information and Communications R&D Promotion Programme of Ministry of Internal Affairs and Communications, Japan: "Kansei Time-series Media Hub Mechanism for Impression Analysis/Visualization Delivery Intended for Video/Audio Media."

REFERENCES

[1] Ceasar, P. and Chorianopoulos, L. "Interactivity and user participation in the televition lifecycle:creating, sharing and controlong content," Proc.UXTV 2008, pp. 127-128, 2008.

[2] Jensen, J., Tscheligi, M., Obrist, M., and Lugmayr, A. "Changing Television Environments," LNCS5066. Springer, pp. 1-10, 2008

[3] Ursu, M.F., Sussner, J., Myrestam, U., Hall, N., Thomas, M., Kegel, I., Williams, D., Tuomola, M., Lindstedt, I., Wright, T., Leurdijk, A., and Zsombori, V "Interactive TV narratives : Opportunities, progree and challenges," ACM TOMCCAP, Vol.4, Issue4, pp. 1-39, 2008.

[4] Donnelly, V. and Merrick, P. "Community portals through community generation," SIGCAPH Comput. Phys. Handicap, pp. 9-14, 2002.

[5] Tuomi, P. "A brief history of social iTV entertainment," In Proceedings of the MindTrek '09, ACM, pp. 15-18, 2009.

[6] Ma, O. and Tanaka, K. "Webtelop: Dynamic tv-content augmentation by using web pages," In Proceedings of ICME' 2003, IEEE, Vol. 2, pp. 173-176, 2003.

[7] Nadamoto, A. and Tanaka, K. "Complementing Your TV-Viewing by Web Content Automatically-Transformed into TV-program-type Content," In Proceedings of the ACM Multimedia' 2005, pp. 41-50, 2005.

[8] Adany, R., Kraus, S., and Ordóñez, F. "Uncertain personal advertisement allocation for Mobile TV," In Proceedings of ACM MoMM '10, pp. 159-166, 2010.

[9] Sumiya, K., Munisamy, M., and Tanaka, K. "Tv2Web: generating and browsing web with multiple lod from video streams and their metadata," In Proceedings of ICKS2004, pp. 158-167, 2004.

[10] Miyamori, H. and Tanaka, K. "Webified Video: Media Conversion from TV Programs to Web Content for Cross-Media Information Integration," In DEXA2005, LNCS3588, Springer, pp. 176-185, 2005.

[11] Livingston, K., Dredze, M., Hammond, K., and Birnbaum, L. "Beyond Broadcast," In Proceedings of ACM IUI'2003, The Seventh International Conference on Intelligent User Interfaces, pp. 260-262, 2003.

[12] Martin, R. and Holtzman, H. "Newstream: A Multi-Device, Cross-Medium," and Socially Aware Approach to News Content. In Proccedings of the ACM EuroiTV 2010, pp. 83-90, 2004.

 ISBN: 978-1-61208-200-4