A Design Framework for Event Recommendation in Novice Low-Literacy Communities

Yimeng Deng, Klarissa T.T. Chang

Abstract—The proliferation of user-generated content (UGC) results in huge opportunities to explore event patterns. However, existing event recommendation systems primarily focus on advanced information technology users. Little work has been done to address novice and low-literacy users. The next billion users providing and consuming UGC are likely to include communities from developing countries who are ready to use affordable technologies for subsistence goals. Therefore, we propose a design framework for providing event recommendations to address the needs of such users. Grounded in information integration theory (IIT), our framework advocates that effective event recommendation is supported by systems capable of (1) reliable information gathering through structured user input, (2) accurate sense making through spatial-temporal analytics, and (3) intuitive information dissemination through interactive visualization techniques. A mobile pest management application is developed as an instantiation of the design framework. Our preliminary study suggests a set of design principles for novice and low-literacy users.

Keywords—Event recommendation, iconic interface, information integration, spatial-temporal clustering, user-generated content, visualization techniques

I. INTRODUCTION

THE fast propagation of social media has enabled the creation and distribution of user-generated content (UGC). UGC refers to media content created or produced by general public rather than paid professionals. The proliferation of UGC creates huge opportunities to explore patterns of events, such as political opinions (e.g., [1]-[2]), hotel rankings (e.g., [3]-[4]), and road traffic conditions (e.g., [5]-[6]). *Event recommendation* is the analysis and dissemination of event patterns based on information cues presented in UGC [7]. Tools supporting effective event recommendation using analyses of UGC can improve decision making capabilities [8], and address the needs of users effectively [9]-[10].

Event recommendation systems generally recommend specific items for advanced information technology users, such as sales information for loyal consumers or places of interest for travelers. Limited systems have been designed to address the needs of novice and low-literacy users in developing countries. Recommendations of high fish density areas for fishermen [11] or pest infestation trends for farmers are likely to increase productivity and earnings of these users. Novice and low-literacy users are often people who have adopted low-end information technology (e.g., mobile devices such as simple smart phones), but have not benefitted from event recommendation services.

Yimeng Deng is with the National University of Singapore, Lower Kent Ridge Road, Singapore (phone: 65-83125425; e-mail: dengyimeng@nus.edu.sg).

Klarissa T.T. Chang is with the National University of Singapore, Lower Kent Ridge Road, Singapore (e-mail: changtt@comp.nus.edu.sg).

It is necessary to close this gap by providing effective event recommendation systems that can meet the needs of novice and low-literacy users. This raises a research question: *How can we design event recommendations systems using user-generated content for novice and low-literacy communities?*

We propose a design framework for event recommendation systems with three sets of activities for process innovation. Grounded in *information integration theory* (IIT), our framework advocates that effective event recommendation is supported by systems capable of (1) reliable information gathering through structured user input, (2) accurate sense making through spatial-temporal analytics, and (3) intuitive information dissemination through interactive visualization techniques. We developed a mobile pest management application as an instantiation of the design framework for novice and low-literacy agricultural users. It incorporates iconic selection input method, spatial-temporal database and clustering algorithm, and visualization techniques derived from visual information-seeking principles [12].

Our study makes three major theoretical contributions. First, based on Walls et al.'s [13] model for *information system design theory* (ISDT), we theorize a design framework for the development of event recommendation systems. Second, we advocate three sets of activities in process innovation that can improve the effectiveness of event recommendations. Grounded in IIT, effective event recommendations can be facilitated by reliable information gathering, accurate sense making, and intuitive information dissemination. Third, we define novice and low-literacy communities as the target users. Most theoretical work focuses on advanced users, neglecting the characteristics and needs of novice and low-literacy users.

We also make three major practical contributions. First, the study provides industry insights for event recommendation systems based on the design framework. Second, practitioners who are developing information technology innovations for novice and low-literacy users can benefit from our in-depth analysis of the needs and challenges. Third, we offer guidelines for system development teams designing such event recommendation systems.

Our paper is organized as follows. We first highlight the unique characteristics of novice and low-literacy users and describe challenges associated with event recommendations for such users. We then present an overview of our design framework and elaborate on the main components. A preliminary description of the mobile pest management system is delineated. We conclude with a summary of our research contributions and potential future directions.

II. Background

The novice and low-literacy communities constitute about 3 billion people in the world population who have mobile phones and are ready to provide and consume UGC [14]. Two characteristics of such users should be taken into consideration when designing information technologies for them. First, mobile phone is the dominant information technology, thus the design framework should be built on the mobile platforms. Second, the literacy level (e.g., reading, writing and technology capabilities) of such users is usually low and functional. Friendly user interface design is needed to reduce mistakes in generating new content and facilitate understanding on information delivered to them.

Considering these challenges, three characteristics of UGC from novice and low-literacy users have proven to affect event recommendation: control on information quality, complexity of content, and enormity of data [15]. First, the quality of UGC varies drastically from excellent to abuse and spam [16]-[17], casting doubts onto the reliability of information generated. Although it may be hard to eliminate low quality information, it is possible to reduce mistakes in generating information via a well-designed user interface [18]. For instance, structured data input has been proven to improve data accuracy (e.g., [19]-[20]).

Second, mobile platforms encompass very large scale conversations across space and time [21]. The enormous information quantities make such UGC difficult to evaluate and analyze [15]. Spatial-temporal database has been applied to manage large volume electronic data [22]. Advanced analytics incorporating location and time factors can also be used to analyze large volume data sets [23].

Third, most content are tagged with contextual information (e.g., location and time) with the advent of location-based technologies [24]-[25]. It is challenging to present the multidimensional information on handheld devices [26]-[27]. Various visualization techniques are applied in different applications to solve the problem, such as text-list/map-based information visualization, adaptive filter techniques, and pop-up window techniques.

Our study aims to solve these challenges by advocating a design framework based on Walls et al.'s [13] model for ISDT (see Table I).

III. A Design Framework for Event Recommendation

Design guidelines are needed due to the lack of previous tools supporting event recommendation for novice and low-literacy users. Wall's et al. [13] model helps to guide the design product and is often used in the formulation of ISDT. It incorporates four components: (1) kernel theories, (2) meta-requirements, (3) meta-designs, and (4) testable hypotheses. The kernel theories govern meta-requirements for the design product. The meta-requirements are fulfilled by the meta-designs, and testable hypotheses are used to evaluate how the meta-designs satisfy meta-requirements. Employing IIT as our kernel theory, we propose meta-requirements, meta-designs and testable hypotheses for our design framework.

TABLE I
COMPONENTS OF THE DESIGN FRAMEWORK FOR EVENT RECOMMENDATION

1. Kernel theories Describes a set of theories from social sciences governing design requirements (for example, information integration theory)	Adoption of Information Integration Theory (IIT)
2. Meta-requirements Describes a class of goals to which the theory applies (for example, enablers of event recommendations)	Support for effective event recommendation for novice and low-literacy users by enabling: 1) Reliable information gathering 2) Accurate sense making 3) Intuitive information dissemination
3. Meta-designs Describes a class of artifacts that are hypothesized to meet meta-requirements (for example, the design of specific interface elements for novice and low-literacy users)	Incorporation of appropriate user interface, clustering algorithm, and visualization methods that are collectively capable of realizing the information gathering, sense making, and information dispensing meta-functions. Specific meta-design elements are: 1) Utilization of iconic selection as the input method 2) Utilization of spatial-temporal clustering algorithms to analyze the datasets 3) Inclusion of interactive visualization methods from visual information-seeking principles
4. Testable hypotheses Describes a set of predictions used to test whether the meta-designs satisfy the meta-requirements (for example, the empirical evaluation of specific interface elements on system outcomes)	Empirical evaluation of the meta-designs' ability to effectively recommend event by supporting the three meta-functions. Specific testable hypotheses are: 1) Ability of iconic selection to enhance the reliability of information gathering 2) Ability of spatial-temporal clustering algorithms to enhance the accuracy of sense making 3) Ability of interactive visualization techniques to enhance the intuitiveness of information dissemination 4) Ability of reliable information gathering, accurate sense making and intuitive information dissemination to enhance the effectiveness of event recommendation

IV. Kernel Theory

Information integration theory (IIT) states that each piece of information has its value (either favorable or unfavorable) and weight (perceived importance) [28]. Event recommendation systems should be able to understand the value and weight of each piece of UGC in order to identify events based on available information.

For example, Steve reports to the mobile system that he observes some pests on his farm. The value of this information is whether the content of the report, including the pest event, location and time, is trustable or not (e.g., "Is this event consistent with most people's reports about the same event?"). The weight is how much this report matters to the system for identifying the pattern of a focal event (e.g., whether it is a unique report that does not exist in the database).

If it does matter (has some weight) and the system identifies the content in the report as trustable, then this piece of information will be used by the system to analyze event patterns and disseminate a useful recommendation to the target users (i.e., Steve and other users who are probably suffering from this problem). The events identified by the system would be changed when new information comes. As IIT states, when one obtains new information, those new pieces of information will affect the result [28].

Therefore, information integration in the mobile system is an iterative process of gathering information from various sources, making sense of all the information, and dispensing the information about the latest event to target users. As such, we derive three meta-functions from IIT: *information gathering, sense making, and information dissemination*. If reliable information can be gathered, this creates opportunities for advanced sense making (e.g., about the spread of pests). If accurate sense making can be carried out, this creates opportunities for large-scale information dissemination (e.g., about event warnings and actions to deal with pests). If reliable information can be disseminated in forms that are intuitive, target users can benefit by taking timely action to deal with the event. This is an effective event recommendation process (see Figure 1).

First, the information gathering meta-function provides platforms to collect information generated by users. It helps to address the research question by answering: (1) how we can design friendly user interfaces that can facilitate user inputs via an organized manner, and (2) what interface features are needed to ensure the quality of user inputs.

Second, the sense making meta-function offers appropriate algorithms to create, compile, select, and filter the mass datasets and recognize the informative event patterns. It contributes to the system by addressing: (1) how we can organize the enormous and complex data over space and time, and (2) how we can apply clustering algorithms to analyze the spatial-temporal datasets.

Third, the information dissemination meta-function refers to the approach of sending messages to the end users. It helps to address: (1) who the user in need of receiving the recommendation is, and (2) how we can disseminate the recommendation to the user.
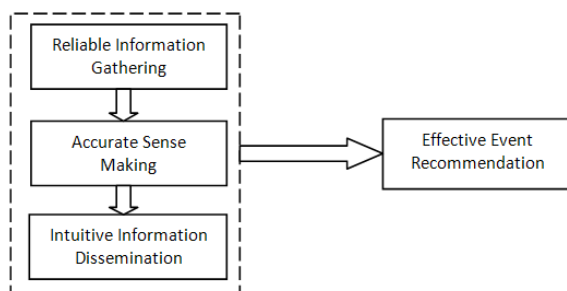

Fig. 1 Key Concepts in the Design Framework

V. META-REQUIREMENTS AND META-DESIGNS

Meta-requirements entail consideration of system features capable of representing the three meta-functions. The objective of meta-designs is to introduce a class of artifacts that are hypothesized to meet the meta-requirements [13]. This section discusses the details of the meta-requirements and meta-designs (see Table II).

TABLE II
META-REQUIREMENTS AND META-DESIGNS FOR THE SYSTEM

Meta-Function	Requirement	Proposed Design
Information Gathering	Information presentation format	Icon
	Information input method	Icon selection
Sense Making	Data storage method	Spatial-temporal database
	Data analysis method	Spatial-temporal clustering algorithm
Information Dissemination	Information design	Map-based information overview
	Interactive visualization	Zoom, filter, and pop-up windows

A. Information Gathering

The information gathering meta-function requires a well-designed user interface to guarantee the reliability of information generated by novice and low-literacy users. Many empirical studies have noted that structured data input method, such as an iconic interface and input selection, can reduce errors across different literacy levels and culture settings (e.g., [29]-[30]). As our target users are novice and low-literacy communities, we adopt icon selection as the input method to ensure reliable information gathering.

B. Sense Making

An effective data mining technique is the core of the sense making meta-function. Data mining is a process of discovering new patterns from large data sets [31]. According to Fayyad et al. [32], data mining is a knowledge discovery process that consists of applying data analysis and discovery algorithms that produce a particular enumeration of patterns (or models) over the data. In this study, we focus on two crucial components in data mining: (1) a structured data model to store large volume electronic data and (2) a data analysis algorithm to cluster the data and recognize the pattern.

Each piece of information includes the event situation, location, and time. As such, a spatial-temporal data model is needed to organize the information over space and time. Similarly, a spatial-temporal clustering algorithm, such as DBSCAN (density based spatio-clustering of application with noise), is required to identify the situation from the dynamic event patterns (e.g., [33]-[34]).

C. Information Dissemination

The information dissemination meta-function needs interactive visualization techniques to deliver intuitive information. Ben Shneiderman [12] suggests a 'mantra', an overriding principle for developing visualizations: *Overview first, zoom and filter, then details on demand*. The aim of this visual information-seeking principle is (1) to provide people with a good overview of the whole dataset, (2) to allow zooming in and focusing on details when required, and (3) to

provide dynamic queries filtering out the data that are not required [35]. Grounded in this principle, we propose that intuitive information dissemination can be achieved by a map-based information overview, zoom in/out display, content-based filter mechanism, and pop-up windows for details.

VI. TESTABLE HYPOTHESES

Testable hypotheses assess whether the meta-designs satisfy meta-requirements [13]. For the proposed design framework, this evaluates the meta-design's ability to recommend the event based on the three meta-functions of information gathering, sense making and information dissemination. In the recommendation agent literature, the effectiveness of a recommendation can be measured by the extent to which a user's decision-making processes are aligned [36]. From a decision-making perspective, an effective recommendation is able to accurately derive important patterns or phenomena of interest from reliable information sources. It should be able to present the patterns or phenomena of interest in a comprehensive format. For the proposed design framework, a suitable meta-design must incorporate features capable of recommend patterns or phenomena of interest by following the three meta-functions. In the following section, we describe a mobile pest management application developed as an instantiation of our design framework. We use this system to evaluate the effectiveness of our meta-designs.

VII. PROTOTYPE: THE PEST MANAGEMENT SYSTEM

Based on our design framework, we develop a prototype of the mobile pest management application. This system aims to gather information from individual farmers, conduct sense making to analyze pest infestation patterns and disseminate information to help farmers take timely actions. Figure 2 shows the system design. Figures 3 and 4 show the prototype .
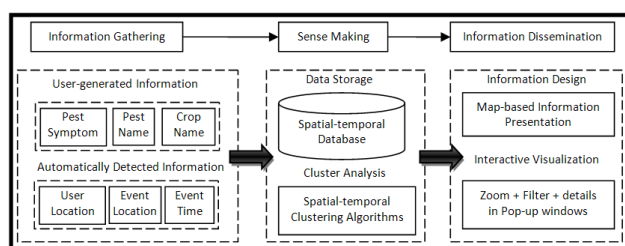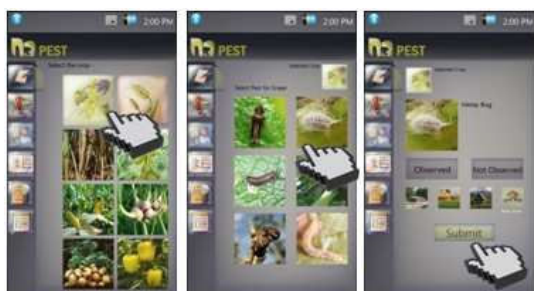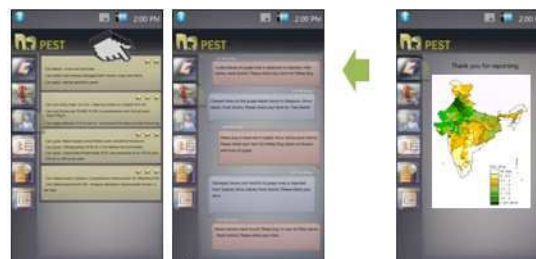

Fig. 2 System Design

Fig. 3 Information Gathering

Fig. 4 Information Dissemination

VIII. EXPERIMENT EVALUATION

Our objective is to test the hypotheses using the prototype. The meta-design's effectiveness for representing the meta-functions in the mobile pest management system will be tested in a series of laboratory experiments and field studies. In our preliminary study, we conducted a laboratory experiment with three between-subjects factors: input method (iconic vs. text selection), cluster analysis (available vs. unavailable), and visualization interactivity (high vs. medium vs. low). Thus, we employed a counterbalanced, full factorial design, which included a total of 12 combinations.

The task of reporting and retrieving pest infestation problems was chosen for the experiment mainly based on two reasons. First, a high economic impact event is an ideal context for studying the effectiveness of recommendation as the decision is vital and decision-makers will be careful for economic benefits. Second, the parties involved in the problems can represent the novice and low-literacy users focused in our design framework.

A. Independent Variables

The three independent variables are the features representing the three meta-functions. First, two methods of information input were studied: iconic and text selections. Users who generated content with iconic selection interface would select the right icons that represented their content. Users who generated content with text selection interface would select the most appropriate items from a list of textual descriptions. Second, two levels of cluster analysis were studied: available and unavailable. Participants in the available cluster analysis group would see the event distribution with visible border and severity levels of each cluster. Participants in the unavailable cluster analysis group would see the rough data point without visible distribution border. Third, three levels of interactive visualization were studied: high, medium and low. Based on visual information-seeking principles, the overview was the basic level. One more feature was added to the higher level. Participants in low level interactivity group could only see a static event overview map. Participants in medium level interactivity group could zoom in/out the map and see filtered information. Participants in high level interactivity group would see detailed information about a focal event in pop-up windows.

B. Dependent Variables

The measurements for dependent variables are adapted from previous validated scales. The effectiveness of event recommendation is operationalized as the perceived effectiveness of the seeking process and the perceived

helpfulness to decision making. Participants were asked to evaluate whether the process of seeking recommendations from the system was fast and accurate (perceived effectiveness of the seeking process). They were also asked to rate the extent to which their decision-making processes were improved by the recommendation. The questions were measured with 7-point scales anchored by the statements "strongly disagree" to "strongly agree". This determined the likelihood that the participants believed that the event recommendation was effective.

C. Participants and Experiment Procedures

180 novice and low-literacy users were recruited and randomly assigned to the 12 experimental groups. Through a role-based scenario, participants were told to report a recent pest problem on their farm and read recommendations from the system. Before they started the task, participants were asked to do a pre-study questionnaire about their demographic information and prior experience with mobile usage. Next, participants were instructed to navigate and locate information on a specific function to become familiar with the system layout and its various features.

Third, participants were instructed to generate a report about a recent pest problem on their farm. After they submitted the report, they were instructed to read recommendations from the system regarding the problem they generated. The system was deployed on Android mobile phones. To minimize the potential impact of confounding factors such as environment and mobility on user performance during the experiment, participants were asked not to pick up the mobile phone from a cradle placed on the table and were required to sit in their chair [37]. Lastly, participants were required to fill a post-study questionnaire, which included manipulation checks and measurements for dependent variables. Currently, this research is in progress and data analyses are being conducted to validate the hypotheses.

IX. CONCLUDING REMARKS

This paper aims to propose a design framework for event recommendation systems and apply it to a mobile innovation system for pest management among farmers. Our theoretical contribution is the theorizing of a design framework for mobile systems used by novice and low-literacy communities based on Walls et al.'s model for ISDT. The framework advocates the development of systems that support all three information integration meta-functions described by IIT. The practical contribution is the novel design and prototype of the mobile pest management application It incorporates iconic selection input method, spatial-temporal database and clustering algorithm, and visualization techniques derived from visual information-seeking principles. In future, using the system features and techniques, we will conduct both laboratory experiments and field experiments to test the effectiveness of the design framework.

REFERENCES

[1] L. Sarmento, P. Carvalho, M. J. Silva, and E. de Oliveira, "Automatic creation of a reference corpus for political opinion mining in user-generated content," 2009, pp. 29-36.
[2] Y. Wu, J. Wong, Y. Deng, and K. Chang, "An Exploration of Social Media in Public Opinion Convergence: Elaboration Likelihood and Semantic Networks on Political Events," presented at the Proceedings of the International Conference on Social Computing and its Applications, Sydney, Australia., 2011.
[3] A. Ghose, P. Ipeirotis, and B. Li, "Designing ranking systems for hotels on travel search engines by mining user-generated and crowd-sourced content," 2011.
[4] P. A. Johnson, R. E. Sieber, N. Magnien, and J. Ariwi, "Automated web harvesting to collect and analyse user-generated content for tourism," 2011.
[5] R. Herring, A. Hofleitner, S. Amin, T. Nasr, A. Khalek, P. Abbeel, and A. Bayen, "Using mobile phones to forecast arterial traffic through statistical learning," in 89th Annual Meeting of the Transportation Research Board, 2010.
[6] M. Haklay and P. Weber, "OpenStreetMap: User-generated street maps," Pervasive Computing, IEEE, vol. 7, pp. 12-18, 2008.
[7] M. Kayaalp, T. Ozyer, and S. Ozyer, "A collaborative and content based event recommendation system integrated with data collection scrapers and services at a social networking site," 2009, pp. 113-118.
[8] E. Kaasinen, "User needs for location-aware mobile services," Personal and ubiquitous computing, vol. 7, pp. 70-79, 2003.
[9] B. Rao and L. Minakakis, "Evolution of mobile location-based services," Communications of the ACM, vol. 46, pp. 61-65, 2003.
[10] I. A. Junglas and R. T. Watson, "Location-based services," Communications of the ACM, vol. 51, pp. 65-69, 2008.
[11] K. Chang, M. Gwee, Y. Deng, Y. Wu, X. Sha, and B. Tan, "Location-Based and Mobile Services in the Fishing Industry of the Informal Economy," presented at the Posters of 11th IFIP Working Group 9.4 Conference, Kathmandu, Nepal, 2011.
[12] B. Shneiderman, "The eyes have it: a task by data type taxonomy for information visualizations," in Visual Languages, 1996. Proceedings., IEEE Symposium on, 1996, pp. 336-343.
[13] J. G. Walls, G. R. Widmeyer, and O. A. El Sawy, "Building an information system design theory for vigilant EIS," Information systems research, vol. 3, pp. 36-59, 1992.
[14] R. Jain, V. Singh, and M. Gao, "Social Life Networks for Middle of the Pyramid," 2010.
[15] A. Abbasi and H. Chen, "CyberGate: a design framework and system for text analysis of computer-mediated communication," MIS Quarterly, vol. 32, pp. 811-837, 2008.
[16] E. Agichtein, C. Castillo, D. Donato, A. Gionis, and G. Mishne, "Finding high-quality content in social media," 2008, pp. 183-194.
[17] J. Van Dijck, "Users like you? Theorizing agency in user-generated content," Media, Culture & Society, vol. 31, pp. 41-58, 2009.
[18] D. V. Morland, "Human factors guidelines for terminal interface design," Communications of the ACM, vol. 26, pp. 484-494, 1983.
[19] N. Cheung, V. Fung, Y. Chow, and Y. Tung, "Structured data entry of clinical information for documentation and data collection," Studies in health technology and informatics, pp. 609-613, 2001.
[20] W. R. Hogan and M. M. Wagner, "Accuracy of data in computer-based patient records," Journal of the American Medical Informatics Association, vol. 4, p. 342, 1997.

[21] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," *Business horizons,* vol. 53, pp. 59-68, 2010.

[22] N. Pelekis, B. Theodoulidis, I. Kopanakis, and Y. Theodoridis, "Literature review of spatio-temporal database models," *The Knowledge Engineering Review,* vol. 19, pp. 235-274, 2004.

[23] P. Berkhin, "A survey of clustering data mining techniques," *Grouping multidimensional data,* pp. 25-71, 2006.

[24] B. Sigurbjörnsson and R. Van Zwol, "Flickr tag recommendation based on collective knowledge," in *Proceedings of the 17th international conference on World Wide Web*, 2008, pp. 327-336.

[25] T. De Pessemier, T. Deryckere, and L. Martens, "Context aware recommendations for user-generated content on a social network site," in *Proceedings of the seventh european conference on European interactive television conference*, 2009, pp. 133-136.

[26] J. Gong and P. Tarasewich, "Guidelines for handheld mobile device interface design," 2004, pp. 3751-3756.

[27] D. Zhang and B. Adipat, "Challenges, methodologies, and issues in the usability testing of mobile applications," *International Journal of Human-Computer Interaction,* vol. 18, pp. 293-308, 2005.

[28] *Information Integration Theory*. Available: http://www.cios.org/encyclopedia/persuasion/Finformation_integration_1theory.htm

[29] J. H. Kim and K. P. Lee, "Cultural difference and mobile phone interface design: Icon recognition according to level of abstraction," 2005, pp. 307-310.

[30] I. Medhi, S. Patnaik, E. Brunskill, S. Gautama, W. Thies, and K. Toyama, "Designing mobile interfaces for novice and low-literacy users," *ACM Transactions on Computer-Human Interaction (TOCHI),* vol. 18, p. 2, 2011.

[31] S. Chakrabarti, Martin Ester, Usama Fayyad, Johannes Gehrke, Jiawei Han, Shinichi Morishita, Gregory Piatetsky-Shapiro, and W. Wang, "Data mining curriculum: A proposal (Version 1.0)," in *Intensive Working Group of ACM SIGKDD Curriculum Committee*, 2006.

[32] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. (1996) From Data Mining to Knowledge Discovery in Databases. *AI Magazine*. 37-54.

[33] D. Birant and A. Kut, "ST-DBSCAN: An algorithm for clustering spatial-temporal data," *Data & Knowledge Engineering,* vol. 60, pp. 208-221, 2007.

[34] A. T. Palma, V. Bogorny, B. Kuijpers, and L. O. Alvares, "A clustering-based approach for discovering interesting places in trajectories," 2008, pp. 863-868.

[35] D. Benyon, *Designing Interactive Systems: a comprehensive guide to HCI and interaction design (Second Edition)*: Pearson Education Limited, 2010.

[36] L. Aksoy, P. N. Bloom, N. H. Lurie, and B. Cooil, "Should recommendation agents think like people?," *Journal of Service Research,* vol. 8, pp. 297-315, 2006.

[37] B. Adipat, D. Zhang, and L. Zhou, "The effects of tree-view based presentation adaptation on mobile web browsing," *MIS Quarterly,* vol. 35, pp. 99-122, 2011.