Proceedings

First International Workshop on Recommendation Technologies for Lifestyle Change (LIFESTYLE 2012)

and

First International Workshop on Interfaces for Recommender Systems (InterfaceRS 2012)

co-located with the:

6th ACM Conference on Recommender Systems (RecSys 2012)

September 13, 2012, Dublin, Ireland

This volume is published and copyrighted by:

Bernd Ludwig
Francesco Ricci
Zerrin Yumak
Nava Tintarev
Rong Hu
Pearl Pu

ISSN 1613-0073

Table of Contents

Part I

First International Workshop on Recommendation Technologies for Lifestyle Change (LIFESTYLE 2012)

First International Workshop on Recommendation Technologies for Lifestyle Change

(LIFESTYLE 2012)

http://pc57724.uni-regensburg.de/lifestyleChange/index.php

co-located with the

ACM Recommender Systems 2012

September 13, 2012, Dublin, Ireland

Preface

The workshop on Recommendation Technologies for Lifestyle Change will be an opportunity for discussing open issues, and propose technical solutions for the designing of intelligent information systems that can support and promote lifestyle change. The objective of these systems is to provide users with up-to-date information, and help them to make choices in everyday life activities establishing a sustainable compromise between quality of life, individuality, and fun.

In today's society, particularly in the affluent society, lifestyle is influenced by technology, and the abundance of financial resources. For instance, a large variety of computer games are excessively used, and people often travels by individualized transportation means, such as car, just for fun. Moreover, the idea that technique and money can buy anything spreads also to health management: people believe that medical knowledge can be immediately applicable in case of illness, as technical knowledge can be used for repairing a broken car.

This results in lifestyles that do not care about the negative long-terms effects on the environment, but also about well-being of individual persons. The most prominent example of this is represented by various types of chronic illnesses in developed countries that result from poor lifestyle choices.

In this context, the aim of this workshop is to explore possibilities for recommender systems to support users in taking decisions related to various aspects of their lifestyle; we call them Lifestyle Change Recommender Systems (LSCRS). There are three main challenges for LSCRSs: firstly, such systems have to assess the user's context for delivering such recommendations. Secondly, in order to promote any change in user's lifestyle, they have to recommend a tailored sequence of items, mostly actions, taking into account the dependencies between the recommended items and the effects of each item recommendation. Thirdly, LSCRS have to be designed to favor the user's continuous attention, to enable the explanation of the reasons for the suggested changes in the user's future behavior, and to recall the changes already effectuated.

Hence, in order to provide an effective support to lifestyle change, recommender systems need to provide communicative capabilities, e.g, with multi-modal dialogue systems. Recommendation technologies have to initiate a feedback-change loop that could contribute to lowering the risks of severe illnesses for many individual users and improving the overall environmental situation.

In order to discuss recent developments and advances in this area, the workshop focusses on the following topics:

- Surveys of lifestyle related activities and technological approaches to monitoring them;
- Context modeling for activity recommendations;
- Formal models of sensor data for monitoring every day activities;
- User models for everyday life recommendations that provide user-tailored content;
- Motivational models for lifestyle, every day activities, and environmental responsibility;
- Recommendations of sequences of items (e.g. physical exercises for a whole week, planning meals for a month);
- Measures of the effectiveness for lifestyle change recommender systems;
- Approaches to combine sensor data and interactive user input in LSCRS;
- Strategies to cement behavioral change; Strategies for situation- and user-aware presentation of recommendations;
- Persuasive technologies for interaction with and among users on their personal situation, their habits, and their options to change their lifestyle Recommendation of activities for leisure time and lifestyle;
- Recommendation of information sources (e.g. forum entries, blogs) for LSCRS.

During the workshop, participants will present their papers and discuss contributions to the field addressing a variety of issues:

- As recommendations in this area are more dependent on the personal history of individual users rather than on the collective behavior and attitudes of many users as in more standard collaborative approaches to recommendations, the workshop participants will discuss new recommendation strategies that leverage the retrospective analysis of the user’s past actions and behavioral patterns.
- How can change in behavior be achieved by employing conversational agents? In a case study on alcohol consumption behavior, the benefits of conversational agents to persuade user to control their personal consumption of alcoholic beverages will be illustrated.
- Some contributions to the workshop discuss users’ classification, adequate user models for LSCRS, models for motivations and concerns of users, and aspects of context modeling for lifestyle change recommendations.
- A number of application domains for lifestyle change behavior will be presented ranging from recommending meals and meal plans to travel routes under ecological constraints.

August 2012

Bernd Ludwig
Francesco Ricci
Zerrin Yumak

LIFESTYLE 2012 Workshop Chairs

Organization

Bernd Ludwig
University Regensburg
Institute of Information and Media, Language and Culture
PT Building, Room 3.0.84 c, Regensburg, Germany
Phone: +941 943-3600, fax: +941 943-1954
Email: bernd.ludwig@sprachlit.uni-regensburg.de

Francesco Ricci
Free University of Bozen-Bolzano
Faculty of Computer Science
Piazza Domenicani 3, I-39100 Bozen-Bolzano, Italy
Phone: +39 0471 016 971, fax: +39 0471 016 009
Email: fricci@unibz.it

Zerrin Yumak
Swiss Federal Institute of Technology
School of Computer and Communication Sciences
EPFL / IC / IIF / LIA, INR231 (Batiment IN)
Station 14, CH - 1015 Lausanne, Switzerland
Phone: +41 21 69 36738, fax : +41 21 693 52 25
Email : zerrin.yumak@epfl.ch

Program Committee

- Christoph Bartneck, University of Canterbury (New Zealand)
- Shlomo Berkovsky, NICTA (Sydney, Australia)
- Berardina Nadja De Carolis, University of Bari (Italy)
- Pablo Castells, University of Madrid (Spain)
- Aiden Doherty, University of Oxford (UK)
- David Elsweiler, University of Regensburg (Germany)
- Jill Freyne, CSIRO ICT Centre (Sydney, Australia)
- Floriana Grasso, University of Liverpool (UK)
- Cathal Gurrin, Dublin City University (Ireland)
- Morgan Harvey, University of Erlangen-Nuremberg (Germany)
- Maurits Kaptein, Philips Research (Netherlands)
- Judith Masthoff, University of Aberdeen (Scotland, UK)
- Paul Resnick, University of Michigan (USA)
- Alan F. Smeaton, Dublin City University (Ireland)
- Ute Schmid, University of Bamberg (Germany)
- Katerzyna Wac, University of Geneva (Switzerland)

Keynote

Where Recommender Systems Can Help In Lifestyle Interventions and Where They Can't

Paul Resnick

University of Michigan School of Information

Abstract

Diet and exercise are often referred to as lifestyle "choices". And many people seem to be making poor choices. Can recommender systems help? Maybe. But we should be careful not to apply them where they can't help. I will suggest that they are potentially useful for problems of discovery and attention focusing, but not for in-depth evaluation and especially not for motivation, where the desirable action is known but not taken. I will offer examples of both promising and not-so-promising opportunities for exploration of recommender system applications in health behavior change, and suggest other conceptual frameworks that may be more useful where recommender systems are not the right lens.

Short Bio

Paul Resnick is a Professor at the University of Michigan School of Information. He previously worked as a researcher at AT&T Labs and AT&T Bell Labs, and as an Assistant Professor at the MIT Sloan School of Management. He received the master's and Ph.D. degrees in Electrical Engineering and Computer Science from MIT, and a bachelor's degree in mathematics from the University of Michigan.

Professor Resnick's research focuses on SocioTechnical Capital, productive social relations that are enabled by the ongoing use of information and communication technology. His current projects include making recommender systems resistant to manipulation through rater reputations, nudging people toward politically balanced news consumption and health behavior change, and crowdsourcing fact-correction on the Internet.

Resnick was a pioneer in the field of recommender systems (sometimes called collaborative filtering or social filtering). Recommender systems guide people to interesting materials based on recommendations from other people. The GroupLens system he helped develop was awarded the 2010 ACM Software Systems Award. His articles have appeared in Scientific American, Wired, Communications of the ACM, The American Economic Review, Management Science, and many other venues. He has a forthcoming MIT Press book (co-authored with Robert Kraut), titled “Building Successful Online Communities: Evidence-based Social Design”.

Recommending Eco-Friendly Route Plans

Efthimios Bothos
National Technical University of Athens
Athens, Greece
mpthim@mail.ntua.gr

Dimitris Apostolou
University of Piraeus
Piraeus, Greece
dapost@unipi.gr

Gregoris Mentzas
National Technical University of Athens
Athens, Greece
gmentzas@mail.ntua.gr

ABSTRACT

As personal transportation is one of the greatest contributors of CO2 emissions, means able to assist travelers in reducing their ecological impact are urgently needed. In this work we focus on travel recommenders that encourage green transportation habits among travelers who have a pre-existing interest in taking action to lessen their impact on the environment. We aim to provide urban travelers with a personalized travel recommender that will nudge them to plan routes while considering the environmentally friendliest travel modes. We present a novel, ecologically-aware approach for travel recommender systems and propose a system architecture that incorporates dimensions of recommendation information elements and profile matching methods.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Information filtering, Selection process; H.3.4 [**Systems and Software**]: User profiles and alert services

General Terms

Design, Human Factors, Algorithms

Keywords

Travel Recommenders, Choice Architecture, Nudging, Persuasive Technologies, Lifestyle Change

1. INTRODUCTION

Environmental issues are becoming increasingly pressing in our times and means to reduce the ecological impact of citizens' activities are needed urgently. A major source of environmental pollution from citizens' activities is carbon emissions due to traffic and mobility. It is estimated that urban transport in the European Union accounts for 15% of all greenhouse gas emissions [12]. As work and leisure life become progressively geographically distributed, a research issue of high importance pertains the development of methods and tools able to support and guide citizens towards pro-environmental behaviors with respect to their traveling habits and decisions.

Previous research has demonstrated that information regarding transport-related attributes such as travel time, travel costs and carbon emissions can lead to changes in citizens' travel behavior [3]. Nevertheless, although individuals base their choices on the attributes of the choice set (content), the presentation of information (context) has also a strong effect on travelers' behavior [4]. The presentation of choices, also known as "choice architecture" [17], refers to the design and incorporation of small features or nudges in the choice making process, which can assist individuals to overcome cognitive biases by highlighting the better choices for them, without restricting their freedom of choice. Tools available to choice architects can be divided into two categories: those used in structuring the choice task and those used in describing the choice options [9]. Recommender systems can act as tools for structuring the choice task and address the problem of what to present to travelers. Furthermore the use of information technologies incorporating feedback and personalization can be central to make lifestyle or behavioral changes [5] and, in our case, can nudge environmentally-responsible behavior.

In this work in progress we focus on recommender systems that encourage lifestyle changes towards green transportation habits among travelers who have a pre-existing interest in taking action to lessen their impact on the environment. We aim to provide urban travelers with a personalized travel recommender that will nudge them to plan multi-modal routes while considering the environmentally friendliest travel modes. We present a novel, ecologically-aware approach for travel recommender systems and propose a system architecture that incorporates dimensions of recommendation information elements and profile matching methods.

Our approach is detailed in Section 2. We synthesize concepts from multi-criteria decision making (MCDM) recommender systems and recommendations diversification to infuse the ecological dimension on travel recommenders. Namely, we focus on MCDM to infer user preferences and we balance the utility of routes with their carbon footprint in order to generate travel recommendations with ecological characteristics. In Section 3 we analyze the conceptual architecture of a system that implements the proposed approach. An illustrative scenario depicts the various user interactions with the proposed system in Section 4. We conclude with related

Lifestyle @RecSys'12, September 13, 2012, Dublin, Ireland

work and future directions.

2. APPROACH

Contrary to the vast majority of previous research on recommender systems that has focused on improving the accuracy of recommendations, i.e. better modeling user preferences to present individually preferred items, we focus on recommender systems as a tool for nudging users towards eco-friendly traveling decisions. Specifically, the recommender generates a list of suggested routes which reside within the limits of users' preferences and presents choices with low carbon emissions. With our approach we address the problem of a "filter bubble" [15] in its ecological dimension: users of existing navigation services may be trapped in a self-reinforcing cycle of emission-intensive travel modes while never being pushed to discover alternatives.

The problem an ecologically aware travel recommender system is asked to solve can be formulated as follows: Given a user u, find a subset $S \subseteq AvailableRoutes(u)$ such that $|S| = PresentedRoutes$ and the choice of S provides a good balance between the user perceived route utility and CO2 emissions. The research agenda of the above problem includes two main issues: First what is meant by user perceived route utility and how this is calculated and second what is the meaning of the term 'balance'. Both issues can be answered in a number of ways. Our approach is based on utility-based recommenders and involves a three-step process: users provide their preferences which are then transformed to a user perceived route utility value. In the final step, the utility and the CO2 emissions of a route are provided as input to a recommendation algorithm that selects $|S|$ results to be presented to the user.

2.1 User Preferences

Following [18] we adopt a utility based approach to elicit user preferences. In more details users provide their preferences over a set of criteria when planing a route. The revealed preferences are used to infer a user perceived utility per route.

First users are asked to assign themselves in one of six groups of drivers as identified by [2] - Hard driver, Complacent car addict, Malcontented motorist, Aspiring environmentalist, Car-less crusader, Reluctant rider (for a thorough description of these categories please see [2]). This information is asked only once and affects the level of nudging the user may be inclined to accept (i.e. an Aspiring environmentalist will be presented with more routes that involve public transportation and walking than a Hard driver).

Although most navigation applications provide the quickest routes as suggestions, in real life situations users are concerned with other aspects when deciding on a specific trip in a city. For example, the price of the ticket or the fare (e.g. for a taxi) of the transport mean might influence the user's decisions [18]. Moreover travelers interested in reducing their carbon footprint may be willing to walk a bit more or accept a longer trip. Based on the above arguments, in a second step users are asked to provide their preferences on a set of criteria which are then used to calculate a per route utility value. Indicative criteria are: preferred delay for arrival, preferred walking or bicycling time and preferred travel cost.

2.2 Routes and Utility Calculation

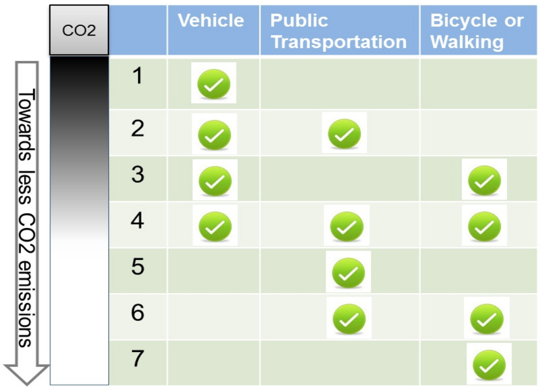

Figure 1: Travel profiles as combinations of alternative travel modes and corresponding qualitative CO2 emissions.

The alternative routes emerge from 'travel profiles' [18] which in our case are defined as the combination of one or more of the major transportation modes (personal vehicle, public transportation, walking or bicycle). In total there are $\sum_{k=1}^{3} \binom{3}{k} = 7$ travel profiles to choose from. Based on the travel mode characteristics and associated emission models of each travel profile we can infer that the use of more walking or bicycle leads to less CO2 emissions (see Figure 1), thus our aim is to nudge users into using travel profiles that include walking or bicycle.

The alternative routes are annotated with a utility value based on the submitted user preferences. To this direction Multi-Criteria Decision Making (MCDM), a set of widely studied methods in the Operations Research domain for decision making, can be employed. With MCDM a decision problem can be seen as the selection of the best alternative from a decision matrix $M \times N$ with N alternatives and M criteria. More specifically we select Multi-Attribute Utility Theory (MAUT) methods [7] which determine the utility of alternatives from user preferences on selected criteria. These methods are based on the concept that bad performing alternatives on one criterion can be compensated by good performing criteria. In our case an alternative is a route with criteria C_j. Each criterion has a weight W_j and the elements $a_{i,j}$ in the decision matrix denote the utility $U(c_{i,j})$ of criterion $c_{i,j}$. Indicative MCDM models that can be used include Weighted Sum and Weighted Product models.

In Weighted Sum Models a weighted mean over all criteria dimension for all alternatives is calculated. The result is a utility score per alternative: $U_i = \sum_{j=1}^{n} a_{ij} w_j$. Weighted Product Models multiply instead of summing up the criteria, and power instead of multiplying the weights in order to calculate the utility scores: $U_i = \prod_{j=1}^{n} a_{ij}^{w_j}$.

2.3 Recommendation Strategies

Given a set of candidate routes $AvailableRoutes(u)$ and a given threshold K of final desired number of recommendations, the optimal scenario of recommendation is finding a set of routes, that has the highest perceived utility and the

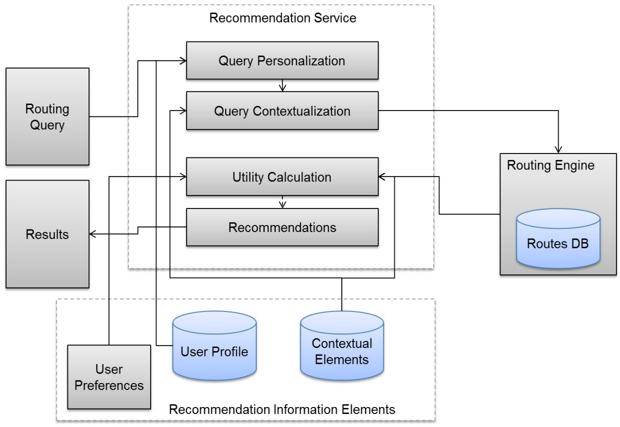

Figure 2: Proposed Architecture.

lowest CO2 emissions. However such an optimal $top-K$ answer set in general does not exist: lowering CO2 emissions typically does not correlate with the highest utility routes being selected. As a result, we have to achieve a balance between CO2 emissions and route utility. In order to generate lists of suggested eco-friendly routes, recommendation diversification algorithms can be employed following [23]. The two problems share similarities: diversification solutions attempt to identify relevant yet diversified items whereas we want to suggest relevant yet eco-friendly routes.

Two optimal algorithms are the MaxUtil which maximizes the utility of the K routes presented and the MinCO2 that minimizes the CO2 emissions of the K routes. Additional heuristic algorithms are the Swap and Greedy similarly to [20] and [22]. With algorithm Swap we begin with the K highest utility routes, and swap the route with the highest emissions with the next highest utility route among the remaining routes. A route is swapped only if the overall CO2 emissions of the displayed set is decreased. To prevent a sudden drop of the overall utility of the resulting set, a pre-defined upper-bound UB denoting how much drop in utility is tolerated has to be used. With the use of UB, swapping stops when the utility of the resulting routes becomes lower than UB. Furthermore the value of UB depends on the drivers group the user has assigned herself (see Section 2.1). With algorithm Greedy recommendation lists are formed by combining routes from different travel profiles. The list with the lowest emissions and acceptable utility is selected. Lists with acceptable utility are those whose difference with the highest utility list resides within certain limits: $HU - U_i \leq AD$ where HU is the Highest Utility, U_i is the utility of list i and AD is the Acceptable Difference which depends on the drivers group the user has assigned herself.

3. ARCHITECTURE

In this section we describe a system architecture that shows how our approach can be instantiated and extended to incorporate personal and contextual information. The proposed architecture comprises of the following components: Recommendation information elements, Recommendation service and Routing engine (see Figure 2).

3.1 Recommendation information elements

These elements incorporate the individual user profile and preferences as well as information related to the current context. In more details we identify the following information elements:

- User preferences provided by the user through a multi-criteria input interface together with the routing query before the trip planning.
- User profile configured by the user through an input interface on the first use of the system.
- Current context of the user, e.g. trip purpose (business, leisure, tourism), weather and traffic information.

3.2 Routing engine

The routing engine takes as input a set of routing options and generates a set of itineraries. It is controlled by the Recommendation service that manages the options on behalf of the user and adjusts the values based on the user's profile. Routing options to be supported include route characteristics such as travel modes. The results should include information regarding emission levels, calculated with emission models and the estimated arrival time at the destination.

3.3 Recommendation Service

This component comprises of four distinct functions responsible for personalizing and contextualizing the alternative routes to be presented to the user. The first two, query personalization and contextualization, transform the user routing query and context signals into the appropriate routing engine API parameters. Query personalization is dependent on the available transportation means the user has at her disposal i.e. car/motorcycle and bicycle and considers any disabilities the user may have. Two rules are defined for these cases:

- If the user owns a vehicle then routing results involving car/ motorcycle should be considered, similarly if the user owns a bicycle, routing results involving a bicycle should be considered.
- If the user has disabilities then bicycle and public means of transportation that do not provide amenities for persons with disabilities should be avoided.

Query contextualization considers a number of static rules to further filter the initial set of results:

- Weather data: if the day is rainy, then bike and walking time should be kept to a minimum.
- Traffic data: if there is indication of high traffic density, car time should be kept to a minimum.
- Trip purpose affects the possible delays with respect to the time of arrival. Expected delays should be minimized for business trips, can be moderately tolerable for leisure trips, and tolerable for tourism trips.

Based on the aforementioned rules, the user query is augmented and a request is sent to the routing engine for alternative itineraries.

Following query personalization and contextualization, the routing engine is triggered to generate a set of n results per travel profile given the set of personalization and contextualization parameters. Once the results are available, two

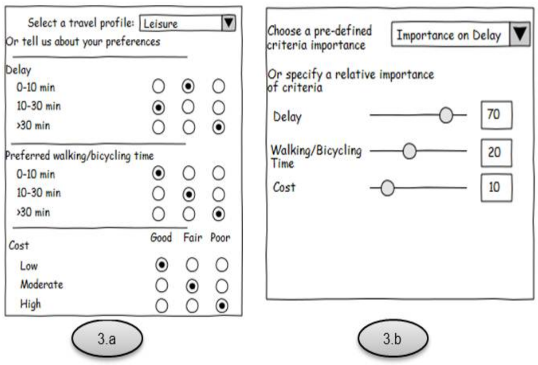

Figure 3: User input: Preferences on the criteria, and relative importance of criteria.

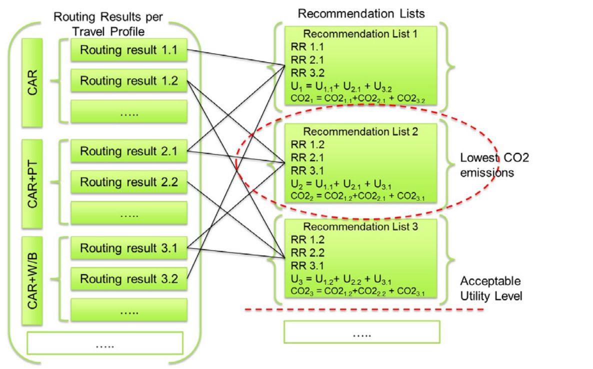

Figure 4: Recommendation lists as combinations of travel profiles. Each travel profile is a combination of one or more travel modes (CAR, PT - Public Transport, W/B - Walking or Bicycle).

more functions are triggered. The utility calculation function maps the recommendation information elements and the characteristics of the route to a perceived utility value per user and route following MCDM methods as described in Section 2.2. This step allows the projection of the user's decision strategy on the results. The final step refers to the generation of recommendations following Section 2.3.

4. ILLUSTRATIVE SCENARIO

In the following we describe an illustrative use case scenario of our approach. John is about to go out and meet his friends at a movie theater and uses his eco-friendly travel recommender to plan the route.

4.1 Query Personalization and Contextualization

The recommendation service interacts with the routing engine and retrieves a number of routes to present to John. According to the user profile, John owns a car, has no disabilities and has described himself as a 'complacent car addict'. According to the contextual information elements, the weather conditions are good, traffic is low and the trip is for leisure. A number of results are retrieved from the routing engine per travel profile.

4.2 User Preference Elicitation

John is asked to define the poor, fair and good levels of each option per criterion (Figure 3.a). Normalized scales are selected for the criteria in order to make the alternatives comparable. Similarly to [16] we employ qualitative scales which are then transformed to numerical values according to the rank order rule for further processing. The numerical mapping is 1 for poor, 2 for fair and 3 for good.

Furthermore John specifies the relative importance of criteria on a percent range, with weights summing up to a total of 100% as shown in Figure 3.b. Changes in one of the sliders in Figure 3.b adapt the values of the rest of the criteria so as to preserve the total of 100. In order to ease user input we can determine a set of predefined profiles (e.g. in the Figures we see that the 'Leisure' preferences profile has the Delay criterion set to 10-30 minutes and the 'Importance on Delay' option assigns higher weight to the 'Delay' criterion).

4.3 Utility Calculation

In this scenario we use the Ordered Weighted Averaging (OWA) MCDM method [19]. With OWA the normalized criteria values a_{ij} (numerical values of the poor, fair, good selections) are multiplied with the corresponding importance weights w_j (importance percentages). Next, rather than being aggregated, weighted criteria values $b_{ij} = a_{ij}w_j$ for each alternative i are re-ordered by descending value so that $b_{i_1} > ... > b_{i_n}$. An OWA operator is applied to the ordered criteria values that can potentially emphasize the better or the poorer values. At this preliminary phase of this work we opt for the neutral operator [19] which assigns equal weights to each criterion and the final utility scores are calculated as the weighted sum of the criterion values.

4.4 Recommendations

Using algorithm Greedy, as explained in Section 2.3, we generate lists of recommended routes by combining results from travel profiles (see Figure 4). The total utility and CO2 emissions of each list are calculated as the sum of the utilities and emissions of each element in the list. The 'Recommendation List 1' has the highest utility for John. The acceptable difference indicates that the recommendation lists one to three should be considered and from those, list 2 has the lowest emissions and is presented to John:

1. Using only his car, John can reach his destination within 30 minutes.

2. Using his car to a parking spot near his destination and then walk for 15 minutes, John can reach his destination within 40 minutes but save 20% of CO2 emissions.

3. Using his car to reach a bus stop close to his home John can reach his destination within 30 minutes and save 30% of CO2 emissions.

John decides to follow option 2 to reach his destination and save 20% of CO2 emissions.

5. RELATED WORK

Commonly, recommender systems generate prioritized lists of unseen items, e.g., music, books, by trying to predict a user's preferences based upon their profile. Travel recommender systems are designed to support travel planning decisions before travel or while on-the-move [6]. These systems capture user preferences, either explicitly or implicitly and suggest destinations to visit, points of interest (POIs), events or activities and/ or alternative routes. The main objective of a travel recommender system is to ease the information search process of the traveler and to convince her of the appropriateness of the proposed services [10].

With respect to route suggestion, certain systems consider multi-modal itineraries (i.e. routes that involve the use of more than one transportation means, for example reaching the destination with a combination of car, bus and walking). Tumas and Riccie [18] present a personalized mobile city transport advisory system that allows users to receive recommendations for personalized paths between two arbitrary points in the city of Bolzano on their mobile phone. They specify travel and user profiles which are then utilized to rank different multi-modal routes in the city and present the top ranked to users. They focus on computing suggestions according to users' travel-related preferences captured through questionnaires and based on four criteria: walking, bus changes, time of arrival at the destination and sightseeing. Zenker and Bernd [21] combine event recommendations and pedestrian navigation with (live) public transport support in order to assist passengers in finding interesting events and navigating to them.

Decision making is a central component in route planing applications [8]. In this respect, MCDM techniques have been employed to model combinations of user desires and to allow users to specify their personal decision strategies while receiving personalized alternatives adjusted to their needs. This view is similar to recent definitions of recommendation problems as MCDM problems. Multi-criteria based recommenders provide suggestions by modeling a user's utility for an item as a vector of ratings along several criteria. A comprehensive study of recommender systems based on MCDM methods was done in [11]. Other related work includes Nadi and Delavar [14] who study the use of OWA in a route planning application in order to generate personalized routes while considering user specific decision strategies in a manner that can provide multiple options regarding the user's preferred decision strategies. Additionally Rinner and Raubal [16] proposed a personalized location-based service using OWA, for finding the most preferable hotel.

Diversity in recommender systems is a research stream that tries to provide solutions beyond improving the relevance of recommendations [13]. Most approaches focus on attribute-based diversification, e.g. in [23], an order-independent intra-list similarity metric to assess the topical diversity of recommendation lists and a topic diversification approach for decreasing the intra-list similarity is introduced. Recently, explanation based approaches have been suggested [20]. They rest on the premise that for two different recommended items i and j, the closer their explanations (i.e., the sets of items that are similar to the recommended items and that are liked by the user), the more homogeneous i and j. Last but not least, in [1] a number of recommendation ranking techniques for diversity are proposed.

6. CONCLUSIONS AND FUTURE WORK

The environmental problems of our times demand new methods and applications able to provide nudges to citizens towards pro-environmental traveling choices. We suggested an approach that infuses nudges in travel recommenders. Users are presented with route alternatives that reside within the limits of their preferences and yield reduced carbon emissions. Furthermore we described a system architecture which combines multi-criteria recommendation techniques with profile matching methods.

There are various aspects that need further research. First, the field of MCDM encompasses a number of methods which could potentially fit into our problem, such as the Analytical Hierarchy Process and the Linguistic Ordered Weighted Averaging (LOWA). Our plan is to examine the applicability of these methods as well as compare them. Second, implementing and comparing the suggested algorithms for striking a good balance between routes utility and eco-friendliness will reveal which are best suited for this problem. Third we will investigate combinations of the proposed recommender with persuasive interfaces for eco-feedback. We expect that such interfaces, by informing users of their carbon footprint, can persuade them to choose the recommendations with low environmental impact. Last, we are going to evaluate our proposed approach in real life situations. To this direction, a prototype system that materializes the approach and related architecture is under development within the Peacox FP7 project. In addition two field trials in Vienna and Dublin have already been planned. In these trials a set of 65 users will use and evaluate the proposed recommendation service in their everyday life for a total duration of two weeks.

7. ACKNOWLEDGMENTS

Research reported in this paper has been partially funded by the European Commission in the Information Society Technologies (IST) project "Ecological aware navigation: usable persuasive trip advisor for reducing CO2 consumption" (Contract no.: 63761).

8. REFERENCES

[1] G. Adomavicius and Y. Kwon. Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques.

[2] J. Anable. Complacent car addicts or aspiring environmentalists? identifying travel behaviour segments using attitude theory. *Transport Policy*, 12(1):65 – 78, 2005.

[3] E. Avineri and O. Waygood. Applying goal framing to enhance the effect of information on transport-related co2 emissions. 9th Biennial Conference on Environmental Psychology, 2011.

[4] M. Ben-Akiva and S. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, 1985.

[5] P. Eslambolchilar, M. L. Wilson, and A. Komninos. Nudge: influence through mobile devices. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*, MobileHCI '10, pages 527–530, New York, NY, USA, 2010. ACM.

[6] D. R. Fesenmaier, H. Werthner, and K. W. Wober. *Travel Destination Recommendation Systems: Behavioural Foundations and Applications (Cabi Publishing)*. CAB International, 2006.

[7] J. Fülöp. Introduction to Decision Making Methods. Technical report, Hungarian Academy of Sciences, 2005.

[8] M. J. Geiger and W. Wenger. On the interactive resolution of multi-objective vehicle routing problems. In *Proceedings of the 4th international conference on Evolutionary multi-criterion optimization*, EMO'07, pages 687–699, Berlin, Heidelberg, 2007. Springer-Verlag.

[9] E. J. Johnson, S. B. Shu, B. G. C. Dellaert, C. Fox, D. G. Goldstein, G. Haubl, R. P. Larrick, J. W. Payne, E. Peters, and D. Schkade. Beyond nudges: Tools of a choice architecture. *MARKETING LETTERS*, 23(2):487 – 504, 2012.

[10] T. Mahmood, F. Ricci, A. Venturini, and W. Hopken. Adaptive recommender systems for travel planning. *Strategies*, pages 1–11, 2008.

[11] N. Manouselis and C. Costopoulou. Analysis and classification of multi-criteria recommender systems. *World Wide Web*, 10(4):415–441, Dec. 2007.

[12] A. D. May, M. Page, and A. Hull. Developing a set of decision-support tools for sustainable urban transport in the uk. *Transport Policy*, 15(6):328 – 340, 2008.

[13] S. M. McNee, J. Riedl, and J. A. Konstan. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In *CHI '06 extended abstracts on Human factors in computing systems*, CHI EA '06, pages 1097–1101, New York, NY, USA, 2006. ACM.

[14] S. Nadi and M. R. Delavar. Multi-criteria, personalized route planning using quantifier-guided ordered weighted averaging operators. *International Journal of Applied Earth Observation and Geoinformation*, 13(3):322–335, 2011.

[15] E. Pariser. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press HC, The, May 2011.

[16] C. Rinner and M. Raubal. Personalized multi-criteria decision strategies in location-based decision support. *Journal of Geographic Information Sciences*, 10:149–156, 2004.

[17] R. H. Thaler and C. R. Sunstein. *Nudge: Improving decisions about health, wealth and happiness*. Penguin, London, Mar. 2008.

[18] G. Tumas and F. Ricci. Personalized Mobile City Transport Advisory System. In W. Höpken, U. Gretzel, and R. Law, editors, *Information and Communication Technologies in Tourism 2009*, chapter 15, pages 173–183. Springer Vienna, Vienna, 2009.

[19] R. R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. Syst. Man Cybern.*, 18(1):183–190, Jan. 1988.

[20] C. Yu, L. Lakshmanan, and S. Amer-Yahia. It takes variety to make a world: diversification in recommender systems. In *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology*, EDBT '09, pages 368–378, New York, NY, USA, 2009. ACM.

[21] B. Zenker and B. Ludwig. Rose - an intelligent mobile assistant - discovering preferred events and finding comfortable transportation links. In *ICAART (1)*, pages 365–370, 2010.

[22] Y. Zheng and X. Xie. Learning travel recommendations from user-generated gps traces. *ACM Trans. Intell. Syst. Technol.*, 2(1):2:1–2:29, Jan. 2011.

[23] C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web*, WWW '05, pages 22–32, New York, NY, USA, 2005. ACM.

Learning user tastes: a first step to generating healthy meal plans?

Morgan Harvey
Computer Science (i8)
Uni of Erlangen-Nuremberg
91058 Erlangen, Germany
morgan.harvey@cs.fau.de

Bernd Ludwig
Institute for Information and Media, Language and Culture
University of Regensburg
93053 Regensburg, Germany
bernd.ludwig@ur.de

David Elsweiler
Institute for Information and Media, Language and Culture
University of Regensburg
93053 Regensburg, Germany
david@elsweiler.co.uk

ABSTRACT

Poor nutrition is fast becoming one of the major causes of ill-health and death in the western world. It is caused by a variety of factors including lack of nutritional understanding leading to poor choices being made when selecting which dishes to cook and eat. We wish to build systems which can recommend nutritious meal plans to users, however a crucial pre-requisite is to be able to recommend dishes that people will like. In this work we investigate key factors contributing to how recipes are rated by analysing the results of a long-term study (n=123 users) in order to understand how best to approach the recommendation problem. In doing so we identify a number of important contextual factors which can influence the choice of rating and suggest how these might be exploited to build more accurate recipe recommender systems. We see this as a crucial first step in a healthy meal recommender. We conclude by summarising our thoughts on how we will combine recommended recipes into meal plans based on nutritional guidelines.

1. INTRODUCTION AND MOTIVATION

In the modern developed world people have the luxury of an abundance of choice with regard to the food they eat. While huge choice offers many advantages, making the decision of what to eat is not always straightforward, is influenced by several personal and social factors [11] and can be complex to the point of being overwhelming [15].

The evidence suggests that many people are making poor dietary choices with stark consequences for their health and well-being. Societal problems such as obesity [19], diabetes [18] and hypertension [14] are all becoming more prevalent, and these conditions are strongly linked to poor dietary habits. The nutritional science literature indicates that these kinds of conditions can be prevented and sometimes even reversed through positive nutritional change [12]. Two issues, though, are that people are generally poor at judging the healthiness of their own diet [8] and even if they do recognise a problem, they lack the requisite nutritional understanding to implement positive dietary changes [4].

Paper presented at the Workshop on Recommendation Technologies for Lifestyle Change 2012, in conjunction with the 6th ACM conference on Recommender Systems.

Lifestyle @RecSys'12, September 13, 2012, Dublin, Ireland

Therefore many people could benefit from assistance that allows them to strike a balance between a diet that is healthy and will keep them well and one that is appealing and they will want to eat. After all, it is no good providing users with healthy diet plans if they do not cook and eat the dishes therein, but instead choose unhealthy meals which are more appealing to them.

We believe this is a problem for which recommender systems are ideally suited. If systems can predict dishes that the user would actually *like to eat*, this could be combined within a system modelling expert nutritional knowledge to provide meal recommendations that are both healthy and nutritious, but also appealing. Furthermore complete meal plans for individual users corresponding to nutritional guidelines given by experts could be generated algorithmically which would suit the user's personal tastes. In this paper we work towards these goals via the following main contributions:

- We collect recipe ratings data in context, in a naturalistic setting over a relatively long time period
- Users not only provide ratings data, but specify the reasons behind their rating (i.e. the content and contextual features that led them to rate in this way)
- We analyse the collected data to determine which factors might help us to better understand a user's preferences
- We discuss how these factors could be utilised to build systems which combine recipes into complete meal plans and the challenges this may present

These contributions all relate to the first aim of our work, that is, to better predict which recipes appeal to a given users and are therefore likely to prepare and eat. We conclude the paper by outlining our plans for future work, summarising some ideas on how we may combine recipe recommendations into sensible meal plans.

2. RELATED WORK

The task of understanding user preferences and suggesting appropriate recipes from a collection can be seen as a novel variant of the well-researched recommender system problem [13, 7]. Although food recommendation is not a frequently studied domain, there is a small body of appropriate related work. Early attempts to design automated systems to plan or recommend meals include CHEF [5] and JULIA [6]. Both of these systems utilise case-based planning to plan a meal to satisfy multiple, interacting constraints. [16] presented a hybrid recommender using fuzzy reasoning to recommend recipes; [9] recommended new food products to supermarket customers, and [17] proposed a system that recommends food items based on recipes recommended to groups of users, clustered by labels.

More recent efforts have tried to better understand the user's tastes and improve recipe recommendations by breaking recipes down into individual ingredients. Freyne and her colleagues [1, 2, 3] demonstrate that this approach works well, with clear improvements over standard collaborative filtering approaches. We wish to build on the success of this work to explore if other content and contextual factors influence the ratings that people assign to recommended recipes. It is our hypothesis that the process of rating a recipe is complex and several factors will combine to determine the rating assigned, beyond purely the user's tastes and that these tastes must be carefully modelled. Both negative and positive ratings could be taken into account, for example: the user may really dislike tomatoes so all recipes with this ingredient might be poorly rated.

Furthermore, not just the existence or absence of explicit ingredients in a recipe but also combination of those ingredients could be important, as could the complexity of the recipe and how long it might take to prepare. Other factors such as how well the preparation steps are described and perhaps the nutritional properties of the dish and the availability of ingredients could have a bearing on the user's opinion of the recommendation. We believe that by building recommender algorithms that incorporate or exploit these kinds of aspects we will be better able to accurately predict ratings. However we also believe that it is vitally important that such factors can be automatically ascertained from ratings data rather than replying on the users themselves. By doing so users can be left to focus on the task of rating recipes and the amount of potentially misleading bias can be minimised. Below we describe how data was collected and analysed to understand how content and contextual factors may influence the way a recipe is rated.

3. DATA COLLECTION

To collect data we developed a simple food recommender system, which selected recipes from a pool of 912 Internet-sourced recipes. This number was chosen as we believe it represents a good balance providing a sufficient variety of dishes from which we may later be able to derive plans whilst, at the same time, being small enough that the resulting ratings matrix will not be too sparse. Users were given a personalised URL and when this was accessed, they were presented with a recipe, selected at random from a list filtered to match a very basic profile. For example, users who specified being vegetarian were only recommended recipes with meta-data indicating no meat; lactose intolerant users were not suggested recipes with milk, etc. Users were not made aware of the random nature of these "recommendations" and were under the impression that the choices were tailored to them. The web page invites the user to provide a rating for the recipe in context i.e. either as a main meal or breakfast for the following day, with recipe meta-data be-

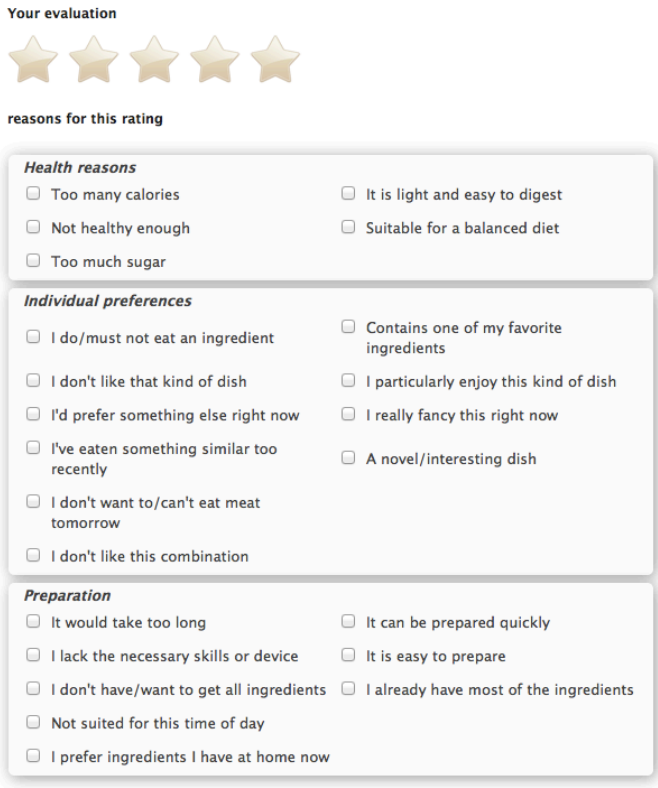

Figure 1: Screenshot of part of the user interface

ing used to determine which meals should be recommended for which time period. This is important because, in contrast to previous data collection methods, the user is not only rating the recipe with respect to how appealing it is, but also how suitable the recipe is given a specific context. Approximately 3 main meals were recommended for every recommended breakfast.

In addition to collecting ratings, the web interface offered the users the chance to explain their ratings by clicking appropriate check boxes representing different reasons. These check boxes were grouped into reasons to do with personal preferences, reasons related to the healthiness of the recipe and reasons related to the preparation of the recipe – see Figure 1. Reasons contributing positively to the ratings were shown on the right-hand side of the screen and negative reasons to the left. The listed explanations were generated through a small user study, whereby 11 users rated recipes and explained their decisions in the context of an interview. The web interface also provided a free-text box for reasons not covered by the checkboxes, however this was only very infrequently used. We did not record any information regarding whether or not the recipe was later cooked or eaten. We were concerned simply by how appealing the recipe was to the user in the occurring context.

After publicising the system on the Internet, through mailing lists and twitter, 123 users from 4 countries provided 3672 ratings over a period of 9 months. The user population grew organically over time with some users only using the system actively for a few weeks and others for longer periods - the kind of behaviour you would expect with a real system. We argue that although this is a relatively small

Rating	0	1	2	3	4	5
Count	61	818	609	822	828	534
%	1.66	22.22	16.54	22.32	22.76	14.5

Table 1: Breakdown of ratings

and sparse data set, it is an improvement on previous recipe ratings data collection methods, which have used mechanical turk (where there are no validity controls) [1, 3] and surveys where participants rate large numbers of recipes or ingredients in a single session [2]. While surveys can offer the chance to collect data on general user preferences in short time periods, they cannot account for factors, such as food availability, preparation and cooking time, previously eaten meals etc., that would influence ratings if a recipe recommender was to be used in the wild.

Our dataset also differs from previous work in terms of matrix density. The number of ratings per user follows a Zipfian distribution (median = 7, mean = 29.93 max = 395 min =1; 18 users have 1, 52 have 10+). Whereas previous food recommender papers report user - ratings densities of between 22% and 35% [1, 2, 3], our dataset exhibits a user-rating density of 3.28%, which we believe to be much more realistic and more in line with standard recommender systems collections such as movielens and netflix. In terms of ratings per recipe, our collection has a median 3 ratings per recipe (mean = 4.04, max=14, min=2). Table 3 shows the breakdown of ratings (ratings of 0 were discounted as they were marked as not being suitable as a full meal).

Our dataset is, therefore, not only realistic in terms of size, but also a suitable platform for investigation and experimentation as it is both sparse and variant in terms of ratings (sd = 1.41).

4. EXPLORATORY ANALYSIS

To learn about the decision process undertaken when users rate recipes, as well as the factors that influence this process, we analysed the reasons provided by the users when they rated. The aim here was take inspiration for the development of new and improved recommendation models. Figure 2 shows the frequency with which users indicated that particular reasons had influenced the rating they assigned. This figure demonstrates the complexity of the process with several factors - both context and content related - being indicated as being influential. Given that the focus of this work is to inform the development of recipe recommender systems, we focus primarily on factors which could be determined automatically by a system

The most common reasons for negatively rating a recipe (shaded grey in the figure) were that the recipe contained a particular disliked ingredient, the combination of ingredients did not appeal, or the recipe would take too long to prepare and cook. The most common reasons for rating a recipe positively (shaded white) had to do with ease or quickness of preparation, the type of dish or the recipe being novel or interesting. Health related reasons, such as the recipe containing too many calories, the user not perceiving the recipe as being healthy enough, or positive factors like the recipe being balanced or easily digestible were clicked less often overall. However, further analysis revealed that these were clicked very frequently for a particular subset of users. 16.3% of the recipes rated by users who clicked on health reasons at least once had a click on a health reason.

To help understand the relationships between the clicked factors and between the factors and the submitted rating we trained a number of linear models. The final model contained 23 factors in total with 17 factors which were significant i.e. the coefficient estimate is more than 2 standard errors away from 0. Highly significant factors (all p-value $\ll 0.01$) included the combination of ingredients in the recipe, whether the recipe would be suitable for vegetarians, how well the users felt the recipe fitted their own tastes and if the recipe contained a specific ingredient the user particularly likes. All of these significant indicators point to the content of the recipes (in terms of ingredients) being highly significant factors in the choice of rating and also suggest in many cases that this is dependent on the individual tastes of the users. This endorses the approach of Freyne et al., who tried to model ingredient preferences in their work. Nevertheless, the fact that ingredient factors can have both a positive and negative influence on ratings and that the combination of ingredients can be important, suggests that more complicated models may be able to better exploit ingredient information when calculating predictions.

Other important factors included whether to not the recipe would be easy to prepare and whether it suited the time of day specified (i.e. breakfast or main meal) and if the user already had the necessary ingredients at home. Interestingly, given the importance of how easy the recipe is to prepare was, the perceived time required to cook the recipe was not a significant factor. This highlight the complexity of the decision process and the number of factors - context-related and content related - which influence how a recipe is rated.

A number of factors related to how healthy the user perceived the recipe to be including if the user felt it would be light and easy to digest and if the user felt it was too unhealthy. In general these health factors did not contribute significantly to the predictive power of the linear models for all of the ratings together, however we wanted to understand if they might help predict ratings on a *per-user basis*. We looked at the correlation between calorie and fat content of recipes and the ratings provided by two groups of users, those had clicked on a health related factor once or more (Care-about-Health, $n = 53$, 2572 ratings), and those who never clicked on a health reason (Don't-Care-About Health, $n = 70$, 1110 ratings)[1]. Figures 3 and 4 show clear differences between the rating behaviour exhibited in these groups. There is a clear trend that the higher the fat content of recipes (r^2=0.88, p=0.012) or the higher the calorific content (r^2=0.87,p=0.022), the lower users in Care-about-Health group tend to rate the recipe. This trend is not present in the second group. If anything there seems to be a slight tendency toward the reverse trend whereby recipes higher in fat ($r^2 = 0.230$,$p = 0.643$) and calories ($r^2 = 0.73$, $p = 0.064$) tend to be assigned a higher rating. This observation suggests that accounting for nutritional factors will allow more accurate recommendations to be generated.

To summarise, these analyses of the collected data demonstrate the complexity of deciding how suitable a recipe will be to cooked in the near future. The results also hint that several factors could be exploited in recommendation algorithms for recipe recommendations.

[1]Nutritional content of recipes was calculated using the system as described in [10].

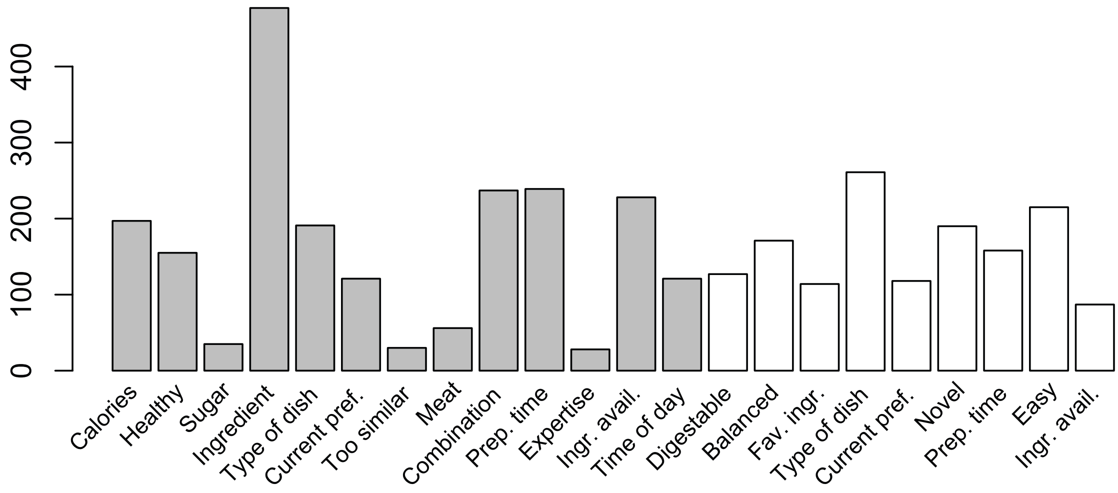

Figure 2: Reasons given for ratings

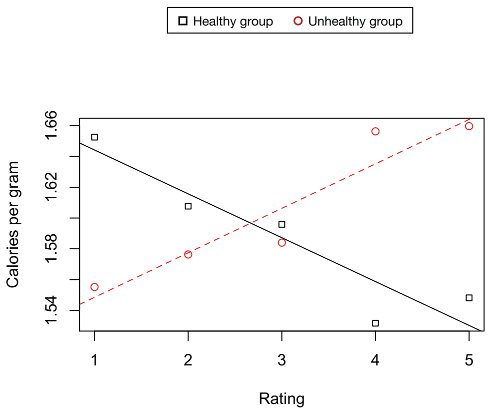

Figure 3: Influence of Calorific Content on Ratings

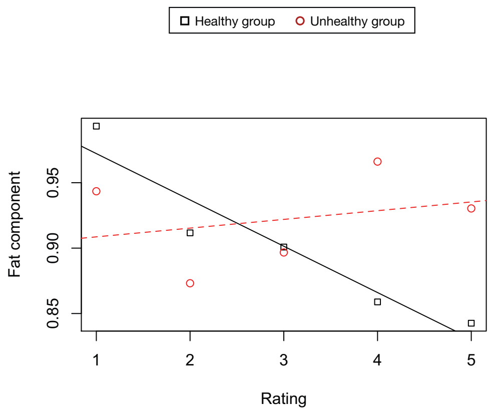

Figure 4: Influence of Fat Content on Ratings

5. BUILDING ON THESE RESULTS

In the previous section we uncovered several patterns in the data indicating that building recommendation algorithms able to account for specific content or contextual features may enable more accurate prediction of recipe ratings. Two important open questions are 1) how can we derive these contextual variables in real-life settings without asking the user to explicitly define their context? And 2) how can we best incorporate such features into recommendation models? We outline some of our thoughts on these points below:

The reasons given by users in our study and the corresponding ratings suggest the ingredients contained within a recipe are very important to the rating process. This finding endorses the approach of Freyne and her colleagues. However, it is clear from our data that ingredients can have either a positive or a negative influence on the rating. For example, if the user likes tomatoes and a recipe contains this ingredient it would be a reason for a high rating. On the other hand, however, if a user does not like tomatoes, our data shows this will negatively affect the recipe rating. Previous recommender algorithms do not account for this negative bias and we believe, based on our results, that including this would improve prediction accuracy. Future recommender models may also account for how important an ingredient is to a dish. For example, imagine a user who does not like tomatoes. For his rating of a recipe where tomato is merely a garnish, this may not have a large influence on the rating. However, if the tomato is a vital ingredient in the recipe e.g. in a tomato soup, then it is more likely to have a large influence.

Another point to consider with respect to ingredients is the coverage of particular ingredients within a collection. For example, Freyne et al.'s algorithm deals with ratings for individual ingredients. This means if egg is rated highly egg-white will be not be treated in the same way. This is exacerbated in our case by the fact that our recipes are web-sourced and may have vocabulary mis-match issues. These kinds of relationships between terms could be identified via instances of nth order co-occurrence. This could be achieved via the use of dimensionality reduction techniques such as singular value decomposition.

Reducing the dimensionality of the feature space would likely have other advantages with respect to dealing how ingredients are combined in a recipe. Our data show that the combination of ingredients can influence the rating applied to a recipe. For example, a user may rate recipes with tomato highly and recipes with pineapple similarly highly on average. However, recipes which combine these ingredients may be given a very low rating. On the other hand, tomato and basil are a combination that work well together and this may have an extra positive influence on the data. Dimensionality reduction techniques, such as SVD or Bayesian Latent Variable models, should implicitly deal with these kinds of patterns.

Our analyses further suggest that including nutritional information in recommendation models should allow more accurate prediction of ratings. We identified two groups of users who behaved very differently based on whether or not they at some point checked that the healthiness of a recipe as an explanation for a rating. The "healthy group" tended to assign a lower rating to recipes higher in calorie and fat content, while the "unhealthy group" displayed, if anything, the opposite predisposition. The group to which a user should be assigned could be obtained explicitly from the user or, preferably, could be learned from ratings data. For example, recipes could be assigned a healthiness score based on nutritional guidelines from health experts and learn which group a user belongs to based on the way they rate recipes with high or low health scores. We acknowledge that the nutrition-aware models may improve performance by offering unhealthy dishes to the users that prefer such dishes and this could be against our long-term goals. We would, however, deal with this issue when combining recipes into meal plans as explained below.

6. CONCLUSIONS AND FUTURE WORK

In this paper we have investigated the decisional process involved in rating recommended recipes. We collected ratings data for recipes and context and statistically analysed the reasons behind assigned ratings. Our future goals in the short term include building on this work to design models that better predict user food preferences using the ideas suggested above. We are continuing to collect data and hope to investigate how performance of models change as the collection size increases.

The presented work represents a single component in a much larger project aimed at building recommender systems that promote healthier dietary choices. In the longer term we plan to move beyond the recommendation of recipes in isolation to recommending dietary plans (7 - 30 days). This involves recommending sequences of recipes under constraints. These constraints will include user preferences of combining recipes and nutritional knowledge, such as the daily recommended intake suggested by the WHO, and user activity patterns. The WHO guidelines provide a means to calculate recommended calorie intake based on a user's profile, as well as a breakdown of the percentage of energy that should come from different types of sources (proteins, fats, carbs, fibre etc.)

One way of modelling this situation is to view it as a graph problem, where the shortest pathes should be computed in a graph where nodes correspond to meals. A week with three meals per day would be represented by a graph with 7 * 3 nodes where edges correspond to dishes (e.g. spaghetti carbonara is an edge from breakfast today to lunch today). A possible cost function could be the distance from the intake estimated from the ingredients and the portion size compared to the recommended daily value. Evaluating the output of such algorithms will be a challenge beyond algorithmics and will involve collaboration with nutritional scientists working on on the project.

Acknowledgements

The authors would like to thank Mario Amrehn and Stefanie Mika for their hard work with the data collection.

7. REFERENCES

[1] J. Freyne and S. Berkovsky. Intelligent food planning: personalized recipe recommendation. In *IUI '10*, pages 321–324, New York, NY, USA, 2010. ACM.

[2] J. Freyne and S. Berkovsky. Intelligent food planning: personalized recipe recommendation. In *IUI '10*, pages 321–324, New York, NY, USA, 2010. ACM.

[3] J. Freyne, S. Berkovsky, and G. Smith. Recipe recommendation: accuracy and reasoning. In *Proc. UMAP*, UMAP'11, pages 99–110, Berlin, Heidelberg, 2011. Springer-Verlag.

[4] J. F. Guthrie, B. M. Derby, and A. S. Levy. *America's Eating Habits: Changes and ConsequencesAgriculture Information Bulletin No. (AIB750)*, pages 243–280. US Department for Agriculture, 1999.

[5] K. Hammond. Chef: A model of case-based planning. In *Proceedings of the National Conference on AI*, 1986.

[6] T. Hinrichs. Strategies for adaptation and recovery in a design problem solver. In *Proceedings of the Workshop on Case-Based Reasoning*, 1989.

[7] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich. *Recommender systems – An introduction.* Cambridge Univ. Press, 2011.

[8] G. Johansson, A. Wikman, A. M. Ahrn, G. Hallmans, and I. Johansson. Underreporting of energy intake in repeated 24-hour recalls related to gender, age, weight, day of interview, educational level, reported food intake, smoking habits and area of living. *Public Health Nutrition*, 4(4):919–27, 2001.

[9] R. D. Lawrence, G. S. Almasi, V. Kotlyar, M. S. Viveros, and S. S. Duri. Personalization of supermarket product recommendations. *Data Min. Knowl. Discov.*, 5(1-2):11–32, January 2001.

[10] M. Müller, M. Harvey, D. Elsweiler, and S. Mika. Ingredient matching to determine the nutritional properties of internet-sourced recipes. In *Pervasive Health 2012*, 2012.

[11] M. Nestle, R. Wing, L. Birch, L. DiSogra, A. Drewnowski, S. Middleton, M. Sigman-Grant, J. Sobal, M. Winston, and C. Economos. Behavioral

and social influences on food choice. *Nutrition Reviews*, 56(5):50–64, 1998.

[12] D. Ornish, S.E. Brown, J.H. Billings, L.W. Scherwitz, W.T. Armstrong, T.A. Ports, S.M. McLanahan, R.L. Kirkeeide, K.L. Gould, and R.J. Brand. Can lifestyle changes reverse coronary heart disease?: The lifestyle heart trial. *The Lancet*, 336(8708):129 – 133, 1990.

[13] F. Ricci, L. Rokach, B. Shapira, and P.B. Kantor, editors. *Rec. Systems Handbook*. Springer, 2011.

[14] P. Scarborough, P. Bhatnagar, K. Wickramasinghe, K. Smolina, C. Mitchell, and M. Rayner. Coronary heart disease statistics. *British Heart Foundation Health Promotion Research Group Department of Public Health, University of Oxford*, 2010.

[15] B. Scheibehenne, R. Greifeneder, and P. M. Todd. Can there ever be too many options? a meta-analytic review of choice overload. *Journal of Consumer Research*, 37:409–425, 2010.

[16] J. Sobecki, E. Babiak, and M. Słanina. Application of hybrid recommendation in web-based cooking assistant. In *Knowledge-Based Intelligent Info. and Engineering Sys.*, pages 797–804. Springer, 2006.

[17] M. Svensson, J. Laaksolahti, K. Höök, and A. Waern. A recipe based on-line food store. In *5th Int. Conf. on Intelligent User Interfaces*, IUI '00, pages 260–263, New York, NY, USA, 2000. ACM.

[18] Diabetes UK. Reports and statistics on diabetes prevalence.

[19] WHO. World health organization: Chronic disease information sheet http://www.who.int/mediacentre/factsheets/fs311/ (accessed feb 14th 2012).

Intrapersonal Retrospective Recommendation: Lifestyle Change Recommendations Using Stable Patterns of Personal Behavior

Robert G. Farrell, Catalina M. Danis, Sreeram Ramakrishnan, Wendy A. Kellogg
IBM, T J Watson Research Center
19 Skyline Drive
Hawthorne, NY 10532
{robfarr, danis, sramakr, wkellogg}@us.ibm.com

ABSTRACT
Leading a healthy lifestyle can prevent or delay medical conditions, elevate mood, improve energy, stabilize sleep, and have other positive effects. Recommender systems are one possible technology to support making lifestyle changes. Recommender systems often use ratings of other users to make recommendations, but this approach may be problematic for making lifestyle change recommendations because of the large variations in human behavior . This paper proposes Intrapersonal Retrospective Recommendation as a new method for generating lifestyle change recommendations that uses only personal history. We explain the benefits and drawbacks of this approach and suggest some future directions.

Categories and Subject Descriptors
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Design, Experimentation, Human Factors.

Keywords
Lifestyle change, habit, retrospective, recommendation, recommender system

1. INTRODUCTION
Changing long-term behavior can be a challenging task for anyone. Bad habits can be entrenched and good habits difficult to establish, transition and sustain. In this work, we focus on lifestyle changes motivated by maintaining or improving personal health. For example, changing a diet in order to achieve weight loss, or supporting the individual in evolving a sustainable exercise regimen comprised of flexibility, strengthening, and difficult to maintain over time as motivation and commitment wax and wane. Thus, there is an opportunity for recommender systems (Ricci, Rokach, and Shapira, 2010.) to suggesting incremental changes to one's routines that collectively can bring about a lifestyle change.

This paper introduces *Intrapersonal Retrospective Recommendation (IRR)* as a promising method of generating lifestyle change recommendations. The key idea behind this approach is that recommendations can be based on what behaviors worked and did not work for the individual in the past. Stable patterns of behavior within a prior time period may be more predictive of an individuals' future behavior than the common behavior patterns of other users. Hence, behavioral patterns in periods of success at lifestyle change or maintenance that are not being followed can be recommended when the user is facing a similar goal but not succeeding. Similarly, the system can recommend cessation of behavior patterns that are found in prior periods of failure at lifestyle change or maintenance as long as a similar goal is being pursued.

The remainder of the paper is structured as follows: the next section introduces the problem of lifestyle change recommendation; section three explains some of the problems with using collaborative filtering in this domain; section four introduces the IRR method; section five reviews some related work; section six provides a real-world example; and section seven provides a generalized algorithm. We wrap up with a discussion and some conclusions.

2. THE PROBLEM
We define *lifestyle* as the pattern of behavior choices an individual makes during a period of time. Mobile phones, sensors, and other devices are making it increasingly possible to collect fine-grained behavioral data about individuals, often with little work on the part of the user. In addition, tools[1] such as Lose It!, DailyBurn™ FitDay™, and MyNetDairy™ allow their users to track their food

Lifestyle@RecSys*'12*, September 13, 2012, Dublin, Ireland

[1] LoseIt! is an unregistered trademark of FitNow, Inc., DailyBurn™ is a registered trademark of Daily Burn, FitDay is a registered service mark of Internet Brands, Inc., MyNetDiary is a registered trademark of 4Technologies Corporation.

tracking makes it possible for users to maintain a long-term fine-grained history of their lifestyle-related choices. In addition, meters, scales, and so on allow individuals to track *measurements* such as weight, waist size, and number of calories above or below some target budget or level to determine progress against their goals over time.

We have been developing a system to analyze a person's tracking data and display information useful for making and maintaining a healthy lifestyle. Alone, these data may be too detailed for people to distinguish meaningful patterns. Recommendations have a significant role to play in helping users dynamically make intelligent choices that help achieve their goals. In our system, each individual has a set of lifestyle goals in the form of constraints to satisfy over an interval of time. For example, Bob wants to do aerobic exercise three times per week. Aerobic exercise is an *activity,* each time that Bob performs that activity is a *behavior* and repeated behaviors are a *behavioral pattern* (i.e., Bob aerobic exercise three times per week). Given this formulation, the recommendation problem is to suggest one or more activities, either individually or in a sequence, for a user, given their history of behaviors by finding stable behavioral patterns.

In the examples in this paper, we use the term *item* to refer to a food (an eating activity) or exercise (a physical activity). This could be at various level of specificity. For example, the item could be "coffee" or "coffee with crème and sugar" or the item could be "running" or "running two miles in 10 minutes". While it is possible for users to explicitly rate items, tracking data indicating that a user consumed a food or performed an exercise is itself an implicit rating. We consider the count of the number of times the user performed the behavior, total amount of the food or time exercising, and total calories burned or consumed to be part of the implicit rating of an item.

3. CHALLENGES WITH USER-USER SIMILARITY APPROACHES

Collaborative filtering (Resnick et. al., 1994) is one of the most successful approaches to generating recommendations. It uses the known ratings of a group of users to make predictions about the unknown ratings of other users. The prediction accuracy of collaborative filtering depends on the similarity of the ratings from the group of users. There are several reasons to believe that LCRSes based on collaborative filtering may have relatively low prediction accuracy:

1. **Sparse Item Space:** The items in LCRSes may be selected by users from a very large item space or even constructed by users, thus reducing the probability of two individuals rating the same item. For example, one user may like Vietnamese frog's legs but it might not ever appear on the menus of other users. While comparing items at a higher level of abstraction may increase the item overlap, recommendations of highly abstract items may be less useful. Even if two users rate single items, impact on measurements may vary. For example, how foods are prepared makes a large difference in their calorie count.
2. **Diverse User Characteristics and Goals:** There may be large individual differences in lifestyles due to differences in individual characteristics (age, height, gender, weight, etc.) and goals (lose 100 pounds vs. lose 5 pounds.) For example, males and females at different weights typically have quite different calorie targets and this impacts food choices and exercise regimens.
3. **Varied Contexts:** Ratings across individuals in real-world situations, whether explicit or implicit, may diverge because of the varied contexts in which ratings were collected. For example, one user may rate pizza high and another low because of the quality of the different pizza parlors they frequent.
4. **Distinctiveness:** The diversity of ratings may be exacerbated by the fact that, for many people, lifestyle is, by definition, aimed at being distinctive. People may seek out items that are unique or significantly different than others. For example, one person might search out exotic foods and another unusual places to exercise, thus reducing their similarity.

While many of these issues can be addressed by having a larger data set, a technique is needed that is not sensitive to the sparse item space, diverse user characteristics and goals, varied contexts, and distinctiveness.

4. INTRAPERSONAL RETROSPECTIVE RECOMMENDATIONS

An alternate technique for generating recommendations is to use the user's own history of choices. We call this *Intrapersonal Retrospective Recommendation (IRR)*. IRR is based on the observation that most people are highly *stable* (i.e., change slowly over time) and highly *distinctive* (i.e., different that others) in their lifestyle-related choices. When a set of outcome measures are tracked (e.g., weight), IRR can return users to stable patterns that worked in the past (i.e., where measures were consistent with their goals) or avoid stable patterns that did not work in the past (i.e., where measures were short of goals). We call a stable pattern of lifestyle-related choices over a time period a *prior self.* In the lifestyle change domain, for example, an individual may have the prior self who ran at least twice a week and who consistently ate high calorie desserts and the prior self who did not exercise. A key insight is that prior selves can serve the role of "similar users" for the purposes of recommendation.

To illustrate this idea, suppose Rachel has toast and coffee every morning and is maintaining her weight. One day she starts buttering her toast every day. If Rachel's peers have coffee and orange juice for breakfast, then collaborative filtering would recommend orange juice to Rachel. IRR would instead recommend she use less butter or butter her toast less often, since during her stable period she did not butter her toast.

The time dimension plays an important role in retrospective recommendations. For example, the "stage" of the individual with respect to his or her goal can influence what is recommended. Early on, simply limiting quantity might lead to successful adherence and weight loss. Later, once the individual has lost significant weight but not yet achieved his goal, it might be necessary to also change the types of foods that are recommended.

We anticipate that IRR will not be as sensitive to some of the problems we outlined with user-user similarity approaches. First, the item space may be smaller because individuals may explore only a small number of items; trying a new item is often risky or breaks an established habit. Second, given that individuals are relatively stable in their preferences over time, user diversity may be lower with prior selves because individual characteristics and goals will have not changed significantly. The context of item ratings within individuals will tend to be consistent over time, improving the stability of IRR. Finally, distinctiveness is not an issue since all comparisons are intrapersonal.

Retrospective recommendation has a number of other advantages. First, the lifestyle change data needed can be made private, thus circumventing the need for data from other users that would be required for other approaches, such as social recommending. Second, IRR may be more transparent (Herlocker, 2004) since the behaviors being recommended will be familiar. Finally, recommendations from one's own history are easier to trust than recommendations from others.

5. RELATED WORK

Other researchers have investigated how to make lifestyle change recommendations. Van Pinxteren, Geleijnse, and Kamsteeg (2011) created a recipe recommender system that suggested healthy alternatives to commonly selected recipes by using a recipe similarity measure. Luo, Tang, and Thomas (2010) created a system to recommend home nursing activities and home medical products. Hammer, Kim, and André (2010) describe a rule-based recommender system for diabetes patients that balance short-term user preferences with long-term medical prescriptions. Sami, Nagatomi, Terabe, and Hashimoto (2008) designed a system to recommend leisure-time physical activities and identified the problem of varied contexts across individuals, primarily the issue that people prefer different places to exercise. Wiesner and Pfeifer (2010) developed a semantic distance metric for health concepts and used it to make personalized health recommendations from an electronic health record. None of these systems make use of the user's history.

There has been recent interest in temporal variation, such as adjusting recommendations to situations where the end users interests "drift" over time (Cao, Chen, Xiong, 2009). Koren and Bell (2011) show how the predictive accuracy of matrix factorization models can be improved using temporal information.

Tanaka, Hori, and Yamamoto (2010) developed LifeLog, a recommender system that captures a history of offline and on-line Web activities and recommends information on Web sites to help users "enjoy waves of information again". This system uses prior stable patterns from personal history, but does not make recommendations based on success or failure relative to personal goals and it does not make lifestyle change recommendations.

6. AN EXAMPLE

We have collected log data from several individuals who are trying to lose weight. Here is a typical log for food and exercise:

Date	**Item**	**Time**	**Number**	**Units**	**Calories**
3/19/12	*Coffee*	*Breakfast*	*32*	*oz*	*9*
3/19/12	*Melon*	*Lunch*	*1*	*cup*	*61*
3/19/12	*Running*	*Exercise*	*50*	*minutes*	*535*

Each line indicates a particular instance of an item or type of behavior that happened at a date and time. The item can be a type of food and an amount (number and units) or a type of exercise with an amount of time. The calories for each item are listed.

On June 1st, we collected log data from three users (A, B, and C) who logged their food and exercise daily. User A lost 20 pounds and logged for 6 months, User B lost 9 pounds over 7 months, and User C lost 55 pounds over 36 months. While all three users were undergoing significant weight loss, their trajectories included plateaus and periods of weight gain. While we do not have enough data to compare with other approaches, we have used IRR to generate recommendations at various points in time for each of these users. User A provided feedback on the suitability of the recommendations.

After three months of steadily losing weight at a rate of 3.3 pounds per month, User A hit a plateau i.e., his weight leveled off. During the 4th month, User A's net calories reached 13,500 calories, whereas during the previous 3 months, User A's average net calories had leveled off at 10,000 calories per month. Thus, we can differentiate the three months of making steady progress toward the weight loss goal and the month of making little progress toward the goal.

We analyzed User A's logs over the 4 month period to find each food and exercise frequency and cumulative calories. During the fourth month, (period two) User A had many of the same patterns of behavior as in the first three months (period one). He ate Coffee, Juice and Milk at breakfast, chocolate squares and nuts as a snack, and field greens and salad dressing at lunch. However, in period two User A started drinking beer and had increased his consumption of chocolate, wine, kung pao chicken, quiche, pizza, bananas, brown rice, and jelly. Together, these differences accounted for the majority of the calorie change. User A also at tortilla chips in month 3 but then stopped in month 4, probably a positive development. However, if tortilla chips were substituting for French fries, for example, then it might be better to go back to tortilla chips. While we cannot be sure that these patterns had become new habits, it might still be useful to recommend changes early.

Given the patterns, we generated recommendations by suggesting decreasing frequency or portion size for food and increasing frequency, intensity, or time for exercise. User A then sorted the recommendations into three categories of suitability: "follow" (likely to take the suggestion), "consider" (like the suggestion but unlikely to follow it), and ignore (don't like the suggestion and won't follow it). Here is how he sorted the recommendations:

- Follow – reduce quiche and chocolate;
- Consider – stop beer, reduce wine and pizza;
- Ignore – reduce kung pao chicken, bananas, and brown rice.

User A reported that eating chocolate squares was a behavior acquired during the weight loss, so it was relatively easy to moderate. The wine and pizza were more entrenched habits. In future work we would like to address the problem of entrenched habits and strong preferences against recommended items.

7. AN ALGORITHM FOR GENERATING RETROSPECTIVE RECOMMENDATIONS

Given our analysis of log data, we designed the following general algorithm for IRR:

1. **Find Periods of Success and Failure:** Establish the individual's goals as a Boolean combination of measures to maintain, increase, or decrease (e.g., maintain weight while reducing fat intake by 10%)Calculate periods of consistent goal achievement (e.g., maintaining weight) or failure over time using historical data (e.g., weight 190 and fat intake of only 25mg/day average). A period of 1 week is used as a minimum length period since physical changes are difficult to measure in smaller periods.
2. **Find Stable Patterns:** Identify stable patterns as repeated items within each period contributing most and least to the goal (i.e., net calories). For example, eating a diet snack an average of 4 times per month while reducing fat intake or running on the treadmill 12 times a month while losing

weight. Any item in the log more than once is used, but the top N=20 items are selected according to their contribution toward the goal (i.e., highest net calories.)

3. **Find Potential Changes in Stable Patterns**: Compute 3 categories of changes across the "current period" and prior periods:
    a. *Cessation* – Patterns that existed but no longer appear in the current period (e.g., no more toast and skim milk at breakfast). For this we use the immediately prior period.
    b. *Formation* – Patterns that emerged in the current period (e.g., chocolate croissant and bacon start appearing at breakfast)
    c. *Substitution* – Patterns that existed but have been modified (e.g., toast has butter, a cup of skim milk instead of 4 oz)
4. **Determine If Potential Changes in Stable Patterns Will Contribute or Detract from Goals:** Label stable patterns of change as *contributing to* or *detracting from* the user's goals. For example, toast and skim milk at breakfast add 180 calories but with only 5mg of fat. On average, other breakfast choices of 180 calories had 10 mg of fat. Therefore, toast and skim milk would be labeled as contributing to the goal. For formation, the proportion of contribution to goals is used (e.g., chocolate croissant had a disproportionate contribution to both fat and calories.) For substitution, the difference can be evaluated as a formation (e.g., adding butter had a disproportionate contribution to fat). The results are ordered by net calories added or subtracted.
5. **Recommend Changes With The Largest Impact**: Recommendations are generated to a) decrease the frequency and/or portion size of high calorie foods; b) increase the frequency of low calorie foods; or c) increase the frequency, intensity, or time for exercise. These changes are based on the net calories saved. For example, if the pattern of eating hamburgers started in period 4 and it contributes significantly to the net calories then recommend reducing the number and/or size of hamburgers and offer.
6. **Tie Changes to Particular Times and Places** Recommendations can be associated with the appropriate time of day (i.e., meal.) or place (i.e., a restaurant). Recommendations can be offered repeatedly to establish new habits.

8. DISCUSSION
IRR appears to generate useful recommendations in some cases, but we have not done a formal evaluation. There are significant challenges with using IRR, many of which were identified in (Herlocker, 2004). First, it can suffer from the cold start problem. There may be no tracking data at the start, particularly since users may start using the service to make a lifestyle change without having already performed tracking. The system can wait for enough tracking data to be available before making a recommendation, but presumably users may simply be tracking an initial state that is far from their goals. Second, in its most basic form IRR guarantees that items will never be novel, since they are selected from the user's own history. Users may be dissatisfied with a system that repeatedly recommends known items. Third, the coverage of an intrapersonal retrospective recommender system will be low, since it only covers the part of the item space that the user has already explored.

Some of these problems may be solved with a hybrid solution. Ideal profiles of foods and exercise could be stored for various weight ranges and weight loss targets and used in lieu of historical data. In this case, retrospective recommendation could recommend the foods and exercises from the ideal profile needed to establish new patterns, transition from old patterns to new patterns, or .maintain existing patterns.

9. CONCLUSION
Lifestyle change is a challenging domain for recommender systems. People are often purposefully distinctive in their lifestyles. The complexity of the real-world activities of eating and exercising makes it difficult to find similarities between users. We introduced Intrapersonal Retrospective Recommendation as an alternative recommendation method that uses an individual's own history of goal achievement to identify behavior patterns to re-establish, transition, or sustain. Next, we hope to evaluate this approach on a larger data set and integrate the system into an overall personal health solution.

10. ACKNOWLEDGMENTS
Thanks to Jim Christensen for his early involvement in the medical adherence project and Sen Hirano for his feedback on this paper.

11. REFERENCES
[1] Cao, H., Chen, E., Yang,J., and Xiong., H. 2009. Enhancing recommender systems under volatile user interest drifts. In *Proceedings of the 18th ACM conference on Information and knowledge management* (CIKM '09). ACM, New York, NY, USA, 1257-1266.

[2] Hammer, S., Kim, J. and André, E. 2010. MED-StyleR: METABO diabetes-lifestyle recommender. In *Proceedings of the fourth ACM conference on Recommender systems* (RecSys '10). ACM, New York, NY, USA, 285-288.

[3] Herlocker, J.L., Konstan, J., Terveen, L.G. and Riedl, J.T. 2004. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.* 22, 1 (January 2004), 5-53.

[4] Koren, Y. and Bell, R.M. 2011. Advances in collaborative filtering. Recommender Systems Handbook, pages 145-186.

[5] Luo, G., Tang, C. and Thomas, S.B. 2010. Intelligent personal health record: experience and open issues. In *Proceedings of the 1st ACM International Health Informatics Symposium* (IHI '10), Tiffany Veinot (Ed.). ACM, New York, NY, USA, 326-335.

[6] Resnick, P., Iacovou,, N., Suchak, M., Bergstrom, P., and Riedl., J. 1994. GroupLens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (CSCW '94). ACM, New York, NY, USA, 175-186.

[7] Ricci, F.. Rokach, F. and Shapira, B. Introduction to Recommender Systems Handbook, Recommender Systems Handbook, Springer, 2011, pp. 1-35

[8] Sami, A., Nagatomi, R., Terabe, M., and Hashimoto, K. 2008 Design of A Physical Activity Recommendation System. In *Proceedings of the IADIS European Conference on Data Mining* (Amsterdam, The Netherlands 22-27 July 2008). DM '08.

[9] Tanaja, K., Hori, K., and Yamamoto, M. 2010. Development of a Recommender System based on Extending Contexts of

Content and Personal History. *Journal of Emerging Technologies in Web Intelligence*, Vol 2, No 3 (2010), 197-203, Aug 2010.

[10] Wiesner, M. and Pfeifer, D. 2010. Adapting recommender systems to the requirements of personal health record systems. In Proceedings of the 1st ACM International Health Informatics Symposium (IHI '10), Tiffany Veinot (Ed.). ACM, New York, NY, USA, 410-414.

[11] van Pinxteren, Y., Geleijnse, G. and Kamsteeg, P. 2011. Deriving a recipe similarity measure for recommending healthful meals. In *Proceedings of the 16th international conference on Intelligent user interfaces* (IUI '11). ACM, New York, NY, USA, 105-11

User Modeling for Pervasive Alcohol Intervention Systems

Ugan Yasavur
School of Computing and Information Sciences
Florida International University
Miami, FL
uyasa001@fiu.edu

Reza Amini
School of Computing and Information Sciences
Florida International University
Miami, FL
ramin001@fiu.edu

Christine Lisetti
School of Computing and Information Sciences
Florida International University
Miami, FL
lisetti@cs.fiu.edu

ABSTRACT

In this paper, we have proposed a user model for computer based drinking behavior change intervention and recommender systems. We discuss specific requirements of user modeling in health promotion and specifically alcohol interventions. We believe that making behavior change systems available pervasively may lead to better and sustainable results. Therefore, our proposed user model takes advantage of the target-behavior related features such as contextual features (e.g., social interactions, location, and time). The proposed user model uses well-validated questionnaires to capture target-behavior specific aspects. We also introduced approaches for enhancing users' experience in the model creation stage by using Embodied Conversational Agents(ECAs) and users' affective states.

Keywords

User modeling, tailoring, alcohol intervention, behavior change, lifestyle change recommender systems (LSCRS).

1. INTRODUCTION

The positive effect of tailoring and personalization on lifestyle change systems is evidenced by several studies [20] [33] [34]. For effective tailoring in lifestyle change systems, comprehensive user characteristics and personal profile/model related to the target behavior need to be acquired and maintained.

Explicit and implicit modeling is needed in healthy behavior promotion systems. In addition, the user model for health behavior change systems must be specialized according to a target behavior (e.g excessive drinking, lack of exercise, obesity). Explicit ways to create a user model or user profile may include conducting assessments with the use of validated questionnaires, psychometric instruments and screening instruments. Implicit ways to build user-profile may include tracking motivation, stage of change, affective features, spatio-temporal events and some data interpretation and mining.

Paper presented at the Workshop on Recommendation Technologies for Lifestyle Change 2012, in conjunction with the 6th ACM conference on Recommender Systems.

Lifestyle@RecSys'12, September 13, 2012, Dublin, Ireland.

Explicit modeling is generally used in the initial user profile creation stage and does not require continuous updates. Implicit modeling facilitates the maintenance of context-related variables in order to increase the context-awareness (e.g. users' physical and social environments) of the system.

After initial creation of a user profile, context-related and affective features need to be kept up-to-date and other profile features must be updated less frequently.

We focus on one target behavior, namely alcohol consumption related behavior change. Therefore, our proposed user model targets lifestyle change systems which aim to promote decreasing or stopping alcohol consumption.

In the following sections, first we study the state of the art in user modeling in life style change recommender systems and behavior change intervention systems in Section 2. Then, we extend the explicit(target-behavior specific) and implicit(target-behavior related) features to build and maintain a user model in Section 3.

2. RELATED RESEARCH

Personalization and tailoring are used in variety of different domains including e-commerce[24], social networks [35], entertainment [18] [7] and health [26] [27]. Whereas collaborative and content-based recommender systems provide a good level of personalization in e-commerce, social networks and entertainment domains, the behavior change domain requires a different approach. The demographic information, user interests, goals, background information and individual traits are the most commonly used user profile features in recommender systems. While these features are still useful in health behavior change systems, different target behaviors requires different modeling features (e.g. consequences of drinking and dependence on alcohol for drinking behavior change; and family history and Body Mass Index (BMI) for obesity).

In addition to personal information, it is useful to benefit from research on context-aware systems [2]. By the increase in usage of smart mobile phones and mobile social network applications, it has recently become possible to track context-related information about users. The most widely used features in context-aware systems are location [7] and time [18] [7]. It is also useful for health promotion recommender systems to use findings of context-aware systems which focus on inferring users' states and activities including social interactions [36] [32]. From continuously posted data on social networks, it is possible to detect social interaction [6].

Recently, there has been an increasing interest in user modeling based on affective features [4]. The user's affective states can be an indicator for the relevance of the recommended item to the user's interest.

In behavior change systems, personalization according to affective state plays a particularly important role because delivering appropriate messages according to current emotions of the user can increase the effectiveness of health promotion interventions [25].

In the health intervention systems which use Embodied Conversational Agents(ECAs)[13] as a user interface, additional personalization can increase the efficacy of the intervention system. Several studies show that concordance of patient and physician increases patient satisfaction [15] [23]. Also, related research on race concordance of the virtual character and the user implies that racial adaption of ECA and user has positive impact on user's satisfaction [25].

In the context of the computer-based alcohol interventions, although there exists some effort in web-based alcohol interventions for personalization and tailoring, they mainly focus on personalization of feedback for conducted assessments [9] [27], [19].

While all mentioned interventions provide personalized feedback, few of them [27], [19] provide feedback based on theoretical constructs (e.g., Transtheoretical Model of Behavior Change). Drinker's Check Up (DCU) [19] provides personalized feedback based on available normative data and uses elements of behavior change models. Responsible Drinking Program[27] makes further personalization by dynamically tailoring feedback across multiple interactions of the client. Although the explicit information acquired from the users is only used for tailoring the feedback, these brief interventions provide good sources for target-behavior specific user modeling. They do not focus on user modeling and personalization in the course of long term behavior change period.

It has been concluded by several extensive surveys on alcohol interventions [8] [43] that computer based interventions have positive effect on reducing or stopping drinking. To maintain motivation and make the behavior change sustainable, we can use behavior change support systems in the form of social networks, mobile applications, lifestyle change recommender systems, and motivational systems.

In the next section we discuss our proposed comprehensive user model which can be used as a reference for alcohol intervention systems and behavior change support systems.

3. THE PROPOSED USER MODEL

Our proposed user model is shown in Figure 1. The model is updated after each assessment and after perception of new affective and contextual features of the user. Assessments provide information about different aspects of the client's drinking. We use some well-validated [19] assessment instruments to gain understanding of the user's drinking psychometric aspects. In addition to assessment results, it is beneficial to monitor the user's affective states via a camera to be able to adapt the recommendations and messages with the user's affective states.

The proposed user model is composed of features grouped under two categories, *target-behavior specific features* (explicit features) and *target-behavior related features* (implicit features). In the following sections, we explain the importance of each feature and the aspects of the problematic drinking behavior that each feature captures.

3.1 Target-Behavior Specific Features

Our target behavior in this paper is *alcohol drinking*. So, in this section we focus on the assessment instruments which can capture specifically the user's alcohol consumption behavior features. The assessments used in this paper are standardized assessment measures proved to be effective in alcohol consumption behavior change [40].

3.1.1 Consequences of Drinking

"Drinking Consequences" feature set assesses the *negative consequences* of the user's drinking. Drinker's Inventory of Consequences (DrInC) [28] is a reliable, valid, clinically useful, and self-administered instrument to assess the negative consequences of drinking. DrInC includes a set of questions in *five different* areas: physical, inter-personal, intra-personal, impulse control, and social responsibility.

The user answers each question in a 4-point Likert scale. Then, by adding up the responses in each area, we calculate his/her score in that area. These scores show the severity of an individual's problems.

The recommender system can use these scores in order to prepare the best personalized feedbacks and recommendations based on the consequences that alcohol has had on the user's life. According to the [28] this feature set should be updated on weeks 1, 8, 16, 26, 52, and 68 of intervention.

Intra-Personal: This feature is assessed using 8 questions which reflect the *subjective perceptions* of the user about her/his drinking. These questions query the user's feeling experienced because of drinking (bad, unhappy, or guilty), personality change experiences (e.g. aggressive, depressive), interference with personal growth, moral life, interests and activities, and interested lifestyle.

Inter-Personal: The focus of this feature is to find out the impact of drinking on the *user's relationships*. So, we query the user's experiences of damage/loss of friendship/love, impairment of parenting and causing harm to the family, concern about drinking from family or friends, damage to reputation, and embarrassing actions while drinking. The assessment of this feature is performed using 10 questions.

Social Responsibility: We use this feature to describe the role-fulfillment of the user from the *other people's point of view*. We use 7 questions to query the user's work/school problems (missing days, poor quality, fired or suspended), financial problems, and failings to meet expectations.

Physical: This feature is assessed using 8 questions that reflect the *negative physical states* resulting from user's drinking. These questions query the user's hangovers, sleeping problems, sickness, harm to health, appearance, eating habits, sexuality, and injury while drinking.

Impulse Control: This feature includes 12 questions about other *unhealthy lifestyles* exacerbated by drinking (e.g., smoking, drugs, and overeating), risk taking and impulsive actions of the user, troubles with law, and damages to people and property.

3.1.2 Motivation to Change

To assess the stage of user's *readiness* and *motivation to change*, we use an instrument called SOCRATES [31]. This instrument involves 19 questions categorized in three domains: *ambivalence*, *recognition*, and *taking steps*. Questions are answered in a 5-point Likert scale. A behavior change recommender system can use these scores to capture the readiness of the user to change before providing recom-

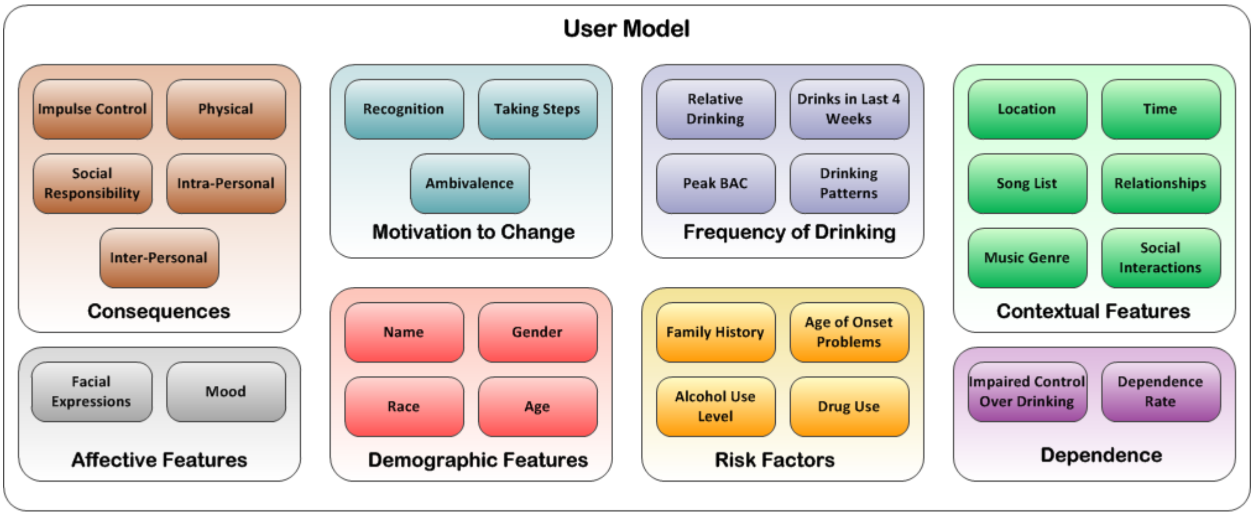

Figure 1: User Model

mendations to change the user's behavior change.

Recognition: The recognition score shows the degree of the user's *awareness* about his/her drinking problems, and the degree of his/her *desire* to change. Therefore, higher degrees of this feature show more desire and motivation to change from the user.

Ambivalence: Ambivalence score shows the degree of *uncertainty* of the user about whether s/he drinks too much, is in control, is hurting others, or is alcoholic. A high ambivalence score shows openness of the user to change. A low ambivalence score has two possible reasons: (1) user knows that his drinking is causing problems (high Recognition); or (2) user knows that s/he does not have drinking problems (low Recognition).

Therefore, we can use this feature to decide whether the user is open to reflections and recommendations or is not ready yet.

Taking Steps: This feature shows the degree of the user's *successful experience in changing* drinking behavior. So, high "Taking Steps" score can be interpreted as (1) need help to persist on the change behavior, and (2) need help to prevent backsliding to the previous drinking behaviors. On the other hand, low scores in this feature show no recent behavior changes in user.

3.1.3 Dependence to Alcohol

We assess the user's degree of dependence to the alcohol using a self-administered 20-item questionnaire called Severity of Alcohol Dependence Questionnaire (SADQ-C) [41]. This feature can be used to predict the likelihood of achieving control-drinking goals, and likelihood of withdrawal.

Questions are answered in a 4-point Likert scale, so the range of the score will be from 0 to 60. Scores higher than 30 for males and 25 for females show severe alcohol dependence and probable need of medical intervention. Scores in 16-30 range show moderate dependence. Otherwise, the user has mild physical dependency.

3.1.4 Risk Factors

We use the Brief Drinker Profile (BDP) [30] to assess some information about the family drinking history, other drug use, additional life problems, motivation for treatment, and history of problem development. Information derived from this feature set can be used in selecting the treatment approaches for user [29] in the behavior change recommender systems. According to the BDP manual [30], the non-static features of this group should be updated every three months.

Age of Onset Problems: This feature involves the user's age in which s/he first took a drink, the age in which s/he first became drunk, and the age in which drinking started affecting his/her life. This feature is static and does not need updates later.

Family History: This feature includes the alcohol problem history of the person's family. User can place his/her family drinking in different categories of abstainer, light drinker, moderate drinker, heavy drinker, problem drinker, or alcoholic. If the user's family does not have any drinking history, it means that his/her drinking patterns were acquired, not inherited. To assess genetic risk factors, the alcohol problems of his/her other biological relatives are queried too.

Drug Use: Since using other drugs can increase the risk of alcohol problems, the type and frequency of the possible used drugs in the last 3 months is queried.

AUDIT Score: Alcohol Use Disorders Identification Test (AUDIT) [5] is a 10-item questionnaire that we use to identify people whose alcohol consumption has become hazardous or harmful to their health. The amount and frequency of drinking, alcohol dependence, and problems caused by alcohol are queried using this instrument. Questions are scored using a 5-point Likert scale. The total score is the summation of all the answers. Table 1 shows the way AUDIT scores are interpreted.

The cut-off numbers may be different based on average body weight, gender, race, and cultural standards.

3.1.5 Frequency of Drinking

This category of features describes the user's drinking patterns and amount of alcohol consumption. So, the alcohol behavior change recommender systems can use them as indicators of the user's drinking pattern and provide more personalized recommendations for the user.

Table 1: AUDIT score interpretation.

AUDIT Score	Interpretation
score < 4	No drinking problems
$4 \leq$ score ≤ 8	Harmful for ages under 18 and females
score > 8	Alcohol dependence
$8 <$ score ≤ 15	Should be advised to reduce drinking
$16 \leq$ score ≤ 19	Should be suggested counseling
score ≥ 20	Should be warranted further diagnose

Drinking Pattern: A drinker may have one of the two drinking patterns: *steady* or *periodic*. A drinker with steady drinking pattern drinks at least once a week and about the same amount every week. A drinker with periodic drinking pattern drinks less often than once a week and is abstinent between drinking episodes.

Drinks in Last 4 Weeks: This feature includes the number of standard drinks that a user had per week in the last four weeks. A standard drink is a 12 oz beer (5% alcohol), a 5 oz wine (12.5% alcohol), or a 1.5 oz liquor (40% alcohol).

Relative Drinking: This feature shows the user's statistical standing relative to the other U.S. people with the same gender.

Peak BAC: Blood Alcohol Concentration (BAC) is the amount of alcohol contained in a person's blood and is measured as weight per unit of volume. Widmark's [44] basic formula for calculating BAC is as follows:

$$\%BAC = (A \times \frac{5.14}{W} \times r) - 0.015 \times H \tag{1}$$

Where, "A" is the total number of liquid ounces of alcohol that the person has drunk since the commencement of drinking. It is calculated by multiplying the number of liquid ounces of drink by its percentage of alcohol. "W" is the person's weight in pounds. "r" is the alcohol distribution ratio which is 0.73 for men and 0.66 or women. "H" is the number of hours between commencement of drinking and the time of BAC calculation.

3.2 Target-Behavior Related Features

These features are not specific to the target behavior but they are implicitly related with the target behavior. For example, demographic information of the user have significant role in personalizing the recommendations and using the normative data to interpret the *target-behavior specific features*. As a concrete example, the normative data used for rating the dependence to alcohol and consequences of drinking depend on the user's gender, race, and age. In addition to the demographic information, we studied affective and contextual features which provide important target-behavior related information.

3.2.1 Demographic Features

Demographic features can be used to improve interpretation of the other feature scores and to improve interaction with the user. Studies [28], [22] show that people of different **genders**, **ages**, and **ethnicities** experience different types of negative consequences after drinking. For example, women have more sleeping problems after drinking while men have more sexual and money problems after drinking.

Therefore, taking the demographic data into account in the user model enables recommending more accurate feedback and exercises to the user.

We can build rapport with the user by calling the user with his/her **name** during the intervention and personalize his/her experience.

For the systems that use ECAs as the interface, they can adapt the ECA's **race** and **gender** to the user's. Research shows that patient-physician race concordance can lead to better health outcomes [15] and that people respond to the ethnicity of ECAs in the same ways of that of humans.

3.2.2 Affective Features

The problem drinkers, who experience intense feeling of depression, discontent and indifference to the world around them, report that they drink to relax or reduce anxiety symptoms [39]. Another research found that emotions and affective states of a person, depending on personality types, predict motives for problem drinking [16]. Therefore emotions and affective states of a problem drinker is crucial for the user model. They can help to fine-tune appropriateness of recommendations and interventions and improve context awareness.

The emotions and affective states can be also used to improve user's experience in the systems which use ECAs as the user interface. The user's experience may affect implicitly the amount and accuracy of the disclosed information. Building a close relationship with the user facilitates his/her behavior change and affects the accuracy of the information disclosed [38], [42].

While the instruments demonstrated can be used as self-administered via form-based interface, the suggested style to administer them is to be delivered via a face-to-face interview [28]. The face to face interviews can be conducted by ECAs [25] which can build a close relationship with the user and have positive effects on the interview process.

Monitoring the facial expressions and mood helps to determine the user's emotions and affective states. In the next section, we described each of these non-verbal signals in more details.

Facial Expressions: According to [3], the facial expressions are the most important modalities in human behavioral judgment. Thus, including facial expressions in human affect analysis can increase the accuracy [12] of the analysis.

Using facial expressions, the behavior change recommender system can recognize the effect of the recommended message/feedback on the user, and his/her affective state.

The user's emotional facial expressions can be recognized through a camera using a real-time facial expression recognition system and categorized into the universal emotion categories [17]: happy, sad, angry, surprised, and neutral.

Mood: Mood is the user's background state of well-being which is often modeled on a bipolar scale of positive-negative valence. Mood changes much slower than emotion and lasts longer time (e.g, minutes to days). Therefore, unlike facial expressions that are updated in real-time, mood can be updated less frequently (e.g., every 5 minutes) in the user model.

To capture the user's mood, we suggest to get the average of the user's categorized emotional facial expressions in a time window and to classify the user's emotions to positive and negative emotions.

3.2.3 Contextual Features

The advancement of the technology on mobile devices, increasing usage of mobile applications, and location-based social networking systems such as Facebook Location[1] and FourSquare[2] introduced new possibilities in development of the context-aware systems. Other than location and time information, social networking and micro-blogging services (Twitter[3]) also offer possibilities to track mood [11], social interactions, relationships, and social ties of the user.

Recently, increased popularity of the music-based social networks[4] and their tight integration to the general purpose social networks introduced new possibilities to improve context awareness of the systems. Research [14] shows that listening some music genres is positively associated with alcohol use. It is also possible to identify personal song lists which lead to alcohol use by tracking multiple context related parameters. For example variation of mood depending on the listened songs and music genres might give important insight about the factors which prepare appropriate psychological conditions for alcohol use.

The location, time of the day, social interactions and mood tracking [11] can help to understand specific conditions which result in alcohol use such as physical environment, psychological conditions, and social conditions.

Several studies show the relationship between reasons and motivations for drinking [1], [22], [21]. Their results imply that contextual awareness will have positive effect on intervention and support systems.

These results implies that personalization and tailoring, based on the contextual factors, are crucial for the alcohol intervention and behavior support recommender systems. Thus, in our proposed user model, we propose to use available information from social networking services and mobile applications to monitor drinking related contextual features.

4. CONCLUSION

In this paper we proposed a user model for alcohol related lifestyle change recommender systems. We proposed target-behavior specific features and target-behavior related features for the user model. We identified the importance of each feature group for the alcohol related intervention and recommender systems. We proposed a user model composing of *eight different groups of features*, *consequences of drinking*, *motivation to change*, *dependence to alcohol*, *risk factors*, *frequency of drinking*, *demographic* features, *affective* features, and *contextual features*.

5. REFERENCES

[1] A. Abbey, M. J. Smith, and R. O. Scott. The relationship between reasons for drinking alcohol and alcohol consumption: An interactional approach. *Addictive Behaviors*, 18(6):659 – 670, 1993.

[2] G. Adomavicius and A. Tuzhilin. *Recommender Systems Handbook*. Springer US, Boston, MA, 2011.

[3] N. Ambady and R. Rosenthal. Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, 111(2):256–274, 1992.

[4] I. Arapakis, I. Konstas, and J. M. Jose. Using facial expressions and peripheral physiological signals as implicit indicators of topical relevance. In *Proceedings of the 17th ACM international conference on Multimedia*, MM '09, pages 461–470, New York, NY, USA, 2009. ACM.

[5] T. F. Babor, J. C. Higgins-Biddle, J. B. Saunders, and M. G. Monteiro. *AUDIT: The Alcohol Use Disorders Identification Test. Guidelines for use in primary health care*. World Health Organization, Department of Mental Health and Substance Dependence, 2 edition, 2001.

[6] N. Banerjee, S. Agarwal, V. Bahl, R. Chandra, A. Wolman, and M. D. Corner. Virtual Compass: relative positioning to sense mobile social interactions. In *Proc. Eighth International Conference on Pervasive Computing*, number LNCS 6030, pages 1–21, Helsinki, Finland, May 2010.

[7] V. Bellotti, B. Price, P. Rasmussen, M. Roberts, D. J. Schiano, A. Walendowski, B. Begole, E. H. Chi, N. Ducheneaut, J. Fang, E. Isaacs, T. King, M. W. Newman, and K. Partridge. Activity-based serendipitous recommendations with the Magitti mobile leisure guide. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*, page 1157, 2008.

[8] B. M. Bewick, K. Trusler, M. Barkham, A. J. Hill, J. Cahill, and B. Mulhern. The effectiveness of web-based interventions designed to decrease alcohol consumption–a systematic review. *Preventive medicine*, 47(1):17–26, July 2008.

[9] B. M. Bewick, K. Trusler, B. Mulhern, M. Barkham, and A. J. Hill. The feasibility and effectiveness of a web-based personalised feedback and social norms alcohol intervention in UK university students: a randomised control trial. *Addictive behaviors*, 33(9):1192–8, Sept. 2008.

[10] T. Bickmore. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human*, 12(2):293–327, 2005.

[11] J. Bollen, H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1):1–8, Mar. 2011.

[12] G. Caridakis, L. Malatesta, L. Kessous, N. Amir, A. Raouzaiou, and K. Karpouzis. Modeling naturalistic affective states via facial and vocal expressions recognition. In *Proceedings of the 8th international conference on Multimodal interfaces*, pages 146–154. ACM, 2006.

[13] J. Cassell and E. J. Sullivan, S. Prevost, & E. Churchill. Embodied Conversational Agents. *Social Psychology*, 40(1):26–36, 2000.

[14] M.-J. Chen, B. a. Miller, J. W. Grube, and E. D. Waiters. Music, substance use, and aggression. *Journal of studies on alcohol*, 67(3):373–81, May 2006.

[15] L. a. Cooper, D. L. Roter, R. L. Johnson, D. E. Ford, D. M. Steinwachs, and N. R. Powe. Patient-centered communication, ratings of care, and concordance of patient and physician race. *Annals of internal medicine*, 139(11):907–15, Dec. 2003.

[16] M. L. Cooper, V. B. Agocha, and M. S. Sheldon. A motivational perspective on risky behaviors: the role of personality and affect regulatory processes. *Journal of personality*, 68(6):1059–88, Dec. 2000.

[1] http://www.facebook.com/about/location
[2] https://foursquare.com
[3] http://www.twitter.com
[4] http://www.spotify.com

[17] P. Ekman, W. V. Friesen, and P. Ellsworth. What emotion categories or dimensions can observers judge from facial behavior? In P. Ekman, editor, *Emotion in the human face*, pages 39–55. Cambridge University Press, New York, 1982.

[18] D. Gavalas and M. Kenteris. A web-based pervasive recommendation system for mobile tourist guides. *Personal and Ubiquitous Computing*, 15(7):759–770, May 2011.

[19] R. K. Hester, D. D. Squires, and H. D. Delaney. The Drinker's Check-up: 12-month outcomes of a controlled clinical trial of a stand-alone software program for problem drinkers. *Journal of substance abuse treatment*, 28(2):159–69, Mar. 2005.

[20] P. Krebs, J. O. Prochaska, and J. S. Rossi. A meta-analysis of computer-tailored interventions for health behavior change. *Preventive medicine*, 51(3-4):214–21, 2010.

[21] E. Kuntsche, R. Knibbe, G. Gmel, and R. Engels. Who drinks and why? A review of socio-demographic, personality, and contextual issues behind the drinking motives in young people. *Addictive behaviors*, 31(10):1844–57, Oct. 2006.

[22] J. W. LaBrie, J. F. Hummer, and E. R. Pedersen. Reasons for drinking in the college student context: the differential role and risk of the social motivator. *Journal of studies on alcohol and drugs*, 68(3):393–8, May 2007.

[23] T. a. Laveist and A. Nuru-Jeter. Is doctor-patient race concordance associated with greater satisfaction with care? *Journal of health and social behavior*, 43(3):296–306, Sept. 2002.

[24] G. Linden and B. Smith. Amazon. com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE*, 7(February):76–80, 2003.

[25] C. Lisetti, U. Yasavur, C. D. Leon, and R. Amini. Building an On-Demand Avatar-Based Health Intervention for Behavior Change. *Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference*, 2012.

[26] M. L. a. Lustria, J. Cortese, S. M. Noar, and R. L. Glueckauf. Computer-tailored health interventions delivered over the Web: review and analysis of key components. *Patient education and counseling*, 74(2):156–73, Feb. 2009.

[27] L. Mauriello. An Internet-Based Computer-Tailored Intervention to Promote Responsible Drinking: Findings from a Pilot Test with Employed Adults. *Alcoholism Treatment*, 30(1):1–15, 2011.

[28] W. Miller and J. Tonigan. *The drinker inventory of consequences (DrInC)*. 1995.

[29] W. R. Miller and R. K. Hester. Matching problem drinkers with optimal treatments. In W. R. Miller and N. Heather, editors, *Treating addictive behaviors Processes of change*, pages 175–203. Plenum Press, 1986.

[30] W. R. Miller and G. A. Marlatt. Brief Drinker Profile. *Psychological Assessment Resources, Odessa, FL*, 1984.

[31] W. R. Miller and J. S. Tonigan. Assessing drinkers' motivation for change: The Stages of Change Readiness and Treatment Eagerness Scale (SOCRATES). *Psychology of Addictive Behaviors*, 10(2):81–89, 1996.

[32] E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell. Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, SenSys '08, pages 337–350, New York, NY, USA, 2008. ACM.

[33] S. Noar, C. Benac, and M. Harris. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychological bulletin*, 133(4):673, 2007.

[34] S. M. Noar, N. Grant Harrington, S. K. Van Stee, and R. Shemanski Aldrich. Tailored Health Communication to Change Lifestyle Behaviors. *American Journal of Lifestyle Medicine*, 5(2):112–122, Nov. 2010.

[35] D. Quercia and J. Ellis. Using mobile phones to nurture social networks. *IEEE Pervasive Computing*, 09:12–20, 2010.

[36] K. K. Rachuri, M. Musolesi, C. Mascolo, P. J. Rentfrow, C. Longworth, and A. Aucinas. Emotionsense: a mobile phones based adaptive platform for experimental social psychology research. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, Ubicomp '10, pages 281–290, New York, NY, USA, 2010. ACM.

[37] C. R. Rogers. The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, 21(2):95–103, 1957.

[38] B. G. Silverman, J. Holmes, S. Kimmel, C. Branas, D. Ivins, R. Weaver, and Y. Chen. Modeling emotion and behavior in animated personas to facilitate human behavior change: the case of the HEART-SENSE game. *Health care management science*, 4(3):213–28, Sept. 2001.

[39] M. M. Spada and A. Wells. Metacognitions, emotion and alcohol use. *Clinical Psychology & Psychotherapy*, 12(2):150–155, Mar. 2005.

[40] D. D. Squires and R. K. Hester. Using technical innovations in clinical practice: the Drinker's Check-Up software program. *Journal of clinical psychology*, 60(2):159–69, Feb. 2004.

[41] T. Stockwell, D. Murphy, and R. Hodgson. The severity of alcohol dependence questionnaire: its use, reliability and validity. *British journal of addiction*, 78(2):145–155, 1983.

[42] a. L. Suchman, K. Markakis, H. B. Beckman, and R. Frankel. A model of empathic communication in the medical interview. *JAMA : the journal of the American Medical Association*, 277(8):678–82, Feb. 1997.

[43] A. White, D. Kavanagh, H. Stallman, B. Klein, F. Kay-Lambkin, J. Proudfoot, J. Drennan, J. Connor, A. Baker, E. Hines, and R. Young. Online alcohol interventions: A systematic review. *J Med Internet Res*, 12(5):e62, Dec 2010.

[44] E. M. P. Widmark. *Principles and Applications of Medicolegal Alcohol Determination*. Biomedical Publications, Davis, California, 1981.

Part II

First International Workshop on Interfaces for Recommender Systems (InterfaceRS 2012)

First Workshop on Interfaces for Recommender Systems (InterfaceRS 2012)

http://www.abdn.ac.uk/~csc284/InterfaceRS/

co-located with the
6th ACM Conference on Recommender Systems (RecSys 2012)

September 13, 2012, Dublin, Ireland

Preface

Since the emergence of recommender systems, a large majority of research focuses on objective accuracy criteria and less attention has been paid to how users interact with the system and the efficacy of interface designs from users' perspective. Well-designed user interfaces have the capability of enhancing user interaction experience and overall satisfaction. For example, explanation interfaces can increase user confidence in their decision choices and inspire user trust and loyalty to the used system. Nowadays, a variety of novel recommendation technologies have been developed to meet different needs (e.g., group and social recommenders). Recommender systems have also extended to new application platforms (e.g., mobile devices). In addition, heterogeneous information resources have been incorporated into recommender systems (e.g., psychological factors, social media). This brings forward new challenges in designing effective and efficient interfaces for these new recommender applications.

This half-day workshop brought together researchers and practitioners around the topics of designing and evaluating novel intelligent interfaces for recommender systems in order to: (1) share research and techniques, including new design technologies and evaluation methodologies (2) identify next key challenges in the area, and (3) identify emerging topics. This workshop aimed to create an interdisciplinary community with a focus on the interface design issues for recommender systems and promoting the collaboration opportunities between researchers and practitioners.

The papers in this volume discuss the following three key issues:

- Visualization and exploration in large and multi-dimensional datasets
- Emotional transferal in recommendations
- Social aspects of interfaces

The paper "*TopicLens: An Interactive Recommender System based on Topical and Social Connections*" looks at visualization of data sets. The authors use a river interface metaphor for navigating topics, items as well as people in three case studies: Twitter, New York Times articles and movies via the Facebook API. The paper "*CoFeel: Using Emotions for social interaction in Group Recommender Systems*" surveys the social interaction between users, and considers how they interact with each other using emotions in a mobile based feedback interface. The paper "*Graph Embeddings* *for Movie Visualization and Recommendation*” proposes a novel way of navigating and exploring a large number of recommendations visually, using a dendogram representation. A demo of this interface is available at: http://graph.bunchwars.com/.

We would like to thank all the authors for their submissions, our Program Committee and sub-reviewers for their precious work.

InterfaceRS 2012 Workshop Organizing Committee

August 2012

Workshop Co-chairs

Nava Tintarev
Dept. of Computing Science,
University Of Aberdeen,
MacRobert Bld., Room 821, Aberdeen, AB24 5UA, United Kingdom
Tel.: +44 1224 27-2839
Fax: +44 1224 27-3422
E-mail: n.tintare@abdn.ac.uk

Rong Hu
Human Computer Interaction Group
School of Computer and Communication Sciences
École Polytechnique Fédérale de Lausanne (EPFL)
EPFL SCI IC PFP, BC 145 (Bâtiment BC), Station 14, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 13 27
E-mail: rong.hu@epfl.ch

Pearl Pu
Human Computer Interaction Group
School of Computer and Communication Sciences
École Polytechnique Fédérale de Lausanne (EPFL)
EPFL SCI IC PFP, BC 107 (Bâtiment BC), Station 14, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 60 81
E-mail: pearl.pu@epfl.ch

Program Committee

- Li Chen, Hong Kong Baptist University, China
- Henrik Eneroth, Antrop, Sweden
- Alexander Felfernig, Graz University of Technology, Austria
- Franca Garzotto, Politecnico di Milano, Italy
- Helen Hastie, Heriot Watt University, UK
- Eelco Herder, L3S Research Center, Germany
- Alejandro Jaimes, Yahoo! Research, Spain
- Joseph Konstan, University of Minnesota, USA
- Jesse Vig, Palo Alto Research Center, USA

TopicLens: An Interactive Recommender System based on Topical and Social Connections

Laura Devendorf
University of California
Berkeley
CA 94720
ldevendorf@berkeley.edu

John O'Donovan
University of California
Santa Barbara
CA 93106, USA
jod@cs.ucsb.edu

Tobias Höllerer
University of California
Santa Barbara
CA 93106, USA
holl@cs.ucsb.edu

ABSTRACT

This paper describes TopicLens, an interactive tool for exploring and recommending items within large corpora, based on both social metadata and topical associations. The system uses a hybrid visualization model that represents topics and content items side by side, allowing the user to actively explore recommendations rather than passively viewing them. The approach provides insight into the composition of relevant topics as they relate to the meta-data of underlying texts. We describe a novel approach to sorting and filtering, which can be topic or document-driven, and two novel interaction styles termed "view inversion" and "human-review", each of which enable novel perspectives on topic modeled sets of documents. To evaluate the system, three use cases are presented to highlight interesting insights across three different data sets using our novel recommendation interface.

1. INTRODUCTION

Recommender systems attempt to ease the information overload problem by providing the right information to the right person at the right time [31, 19, 33]. However, presentation *mechanisms* for these systems are becoming increasingly important, as they are applied to increasingly more diverse data on the social web. For example, Herlocker's early experiments on the value of explaining recommendations [19] have informed and influenced many of today's recommender system designs. Tintarev and Masthoff [38] survey the role of explanation as an integral part of the recommendation process and outline seven distinct advantages of providing explanation. More recent efforts to analyse the effect of "inspectability and control" [21], interactive visual feedback [6], and dynamic critiquing [30, 10] clearly show that the interface components play an important role in a user's acceptance and overall trust in a recommendation.

In this paper we focus on one specific interface design (Figure 1 for exploration of recommendations which have been derived from a topic modeling algorithm. Topic modeling is a statistical method for extracting relevant topics from a large corpus of text. Visualization of connections formed through topic modeling can enable users to quickly identify trends and other insightful details from a large data set. Successful visualizations are especially effective at highlighting patterns within high dimensional data. Such visualizations may also allow the user to navigate and dynamically filter information in order to extract specific and relevant items. Example use cases are:

- To augment the users ability, beyond keyword based search and navigation, to discover topical composition and interrelationships in texts (i.e. recommendation via topic associations).
- To highlight popular trends and conversations within social networks.
- To compare bodies of text, visually exploring similarities, differences and patterns in the underlying texts for better personalized result sets.

The focus of this paper is largely on the UI design and on novel interaction techniques to represent connections formed over large text datasets using topic modeling or other automated text analysis algorithms. The key elements in our visual representations include:

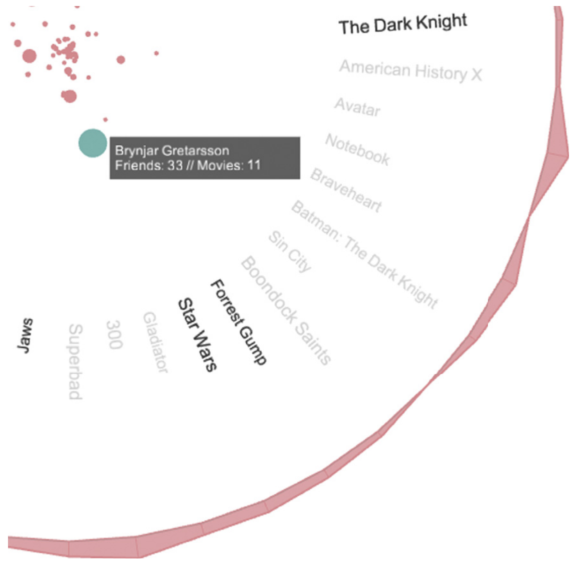

Figure 1: A snapshot of TopicLens interactively recommending movies from Facebook API. The segment shows the popularity of each item among friend groups on the outer ring, and highlights recommendations in bold on the inner ring. This view is highly dynamic and changes based on mouseover interactions

Interface@RecSys'12, September 13, 2012, Dublin, Ireland. Paper presented at the Workshop on Interfaces for Recommender Systems 2012, in conjunction with the 6th ACM conference on Recommender Systems.

- *Recommendable Item*: An abstract entity which can translate to either a text document or a user within a social network. These are conceptually grouped because they are both represented by collections of terms. For example, in the Twitter data set, a user is represented as a collection of Tweets.

- *Topic*: Multinomial distributions over a set of terms, which can be associated with content items.

While established representations, such as word clouds and tree maps [35] can be useful for visualizing frequency in topic-item relationships, we describe a model that also preserves and represents relationships at the meta-data level. This allows users not only to see which topics arise, but also how they arose and under what conditions. The approach enables more informed reasoning about documents a user wishes to investigate, while highlighting trends over a number of different types of networks with respect to a particular investigation.

Microsoft's "Twahpic" [29] approach to visualizing topics in conjunction with meta-data leverages a composite view that optimizes its visualization strategy for each different facet of the data. This strategy is effective for illustrating and highlighting the multifaceted nature of the data, but is difficult to navigate due to the separation of each frame and the segregation of the data networks. In short, the interaction model helps a user form impressions of the data rather than supporting investigations into the data.

Work by Cao et al. in [8] shows a benefit of using multiple approaches to visualizing the different facets of the data, and in this paper, we will present a model that takes a hybrid approach rather than a segregated approach in order to facilitate navigation and interaction with the data. The key features of the proposed technique are as follows:

- Presents a choice of view modes, sorting parameters and controls for navigation and dynamic filtering.

- Enables a user to filter topics in relation to the pre-existing networks in the data.

- Allows for human oversight of algorithmically generated results.

- Enables exploration of dataset as a map, traversing and isolating regions of particular interest in order to extract relevant items.

- Caters to diverse topic modeling scenarios, including additional data such as social and information networks.

In the remaining sections, we will discuss the related research and provide a brief background of topic modeling before describing in detail the design decisions made when developing the TopicLens interface. The design decisions include those related to overall structure and the mapping of formal elements to relational information. Novel aspects of the interface are also discussed, particularly new techniques that we have termed ï£¡view inversionï£¡ and ï£¡human review.ï£¡ We will then present three applications of the system, one of which uses data that does not contain topic-based relations, thus highlighting a more generalized application of the design.

2. RELATED WORK

Due to the proliferation of data available on the web, there is an increasing need for better techniques for exploration of large amounts of text data. This is commonly known as addressing an information overload problem [20]. Ongoing research has produced proactive, query-based solutions in the fields of search [12] and reactive or filter-based approaches in the field of recommendation [20, 7]. In the context of this work, we are especially interested in approaches that employ visual and interactive methods to tailor an information space to a user's individual needs. The novel approach presented in this paper employs a statistical method known as Latent Dirichlet Analysis (LDA) or "topic modeling" [4, 3] to discover useful linkages between documents upon which visualizations are built.

Mathematical Theory	theorem lemma proof follow constant bound exist definition
Software Engineering	software process tool project development design system developer
Gene Expression	protein genes expression network motif interaction pathway genome
Politics and Society	political social policy economic china law government national
Business and IT	business firm services customer technology management market product
Fluid Dynamics	flow velocity wall fluid turbulence reynold pressure channel

Table 1: Examples of LDA topics learned on a corpus of research papers

While there has been a significant amount of research in this domain from a variety of perspectives, from early approaches such as [27, 40, 18] to more recent work in [36, 37, 39, 22, 25], visual techniques for exploring large sets of documents have not yet been widely adopted.

2.1 Topic Modeling

LDA or "topic modeling" is a statistical technique introduced by Blei et al. [4] that computes focused probability distributions over the words in a set of documents. The algorithm functions by mapping documents onto a smaller number of "topics". In this sense, a topic consists of a multinomial distribution over words or stemmed terms in a document set. For example, as $p(w|t)$, for $t \in 1 \ldots T$, where T is the number of topics [4, 16]. In many cases, topics are displayed as a list of the top n words with the highest probability in the set. Table 1 from [15] shows some example topics produced by an LDA algorithm. In this case, the words "theorem, lemma, proof, follow, constant..." seem to relate to the topic "Mathematical Theory". Recent research in [9, 26] has shown that although LDA topics can be misinterpreted, they are generally well understood by users. Techniques for the automatic labeling of topics have been presented in [23].

In TopicLens, topics are leveraged to form associations among items in a large corpus, and these associations are used to produce informative and highly flexible representations of the broader content item space, using novel layout and interaction techniques. Before describing our approach to visualizing a topic space, we now present a discussion of existing approaches to visualization of large document sets.

Many approaches in the literature dealing with the representation of large text collections, ranging from traditional static representations, e.g. [18], to more recent and highly interactive representations which use advanced methods to relate documents together, e.g. [**?**]. They can rely on pre-existing meta data, or can compute relations on the fly. In this paper, we present a novel interactive design and layout for exploring topic based and social network relations in large document sets. Before presenting the prototype system in detail, the following section provides a brief account of the design choices for using a combination of river and graph-like visual representations in the system.

2.2 The Need for a Hybrid Model

As shown in Figure 2, we are supporting exploration of multi-faceted data in a variety of ways. Specifically, examples are demonstrated on three different network types: social network data with

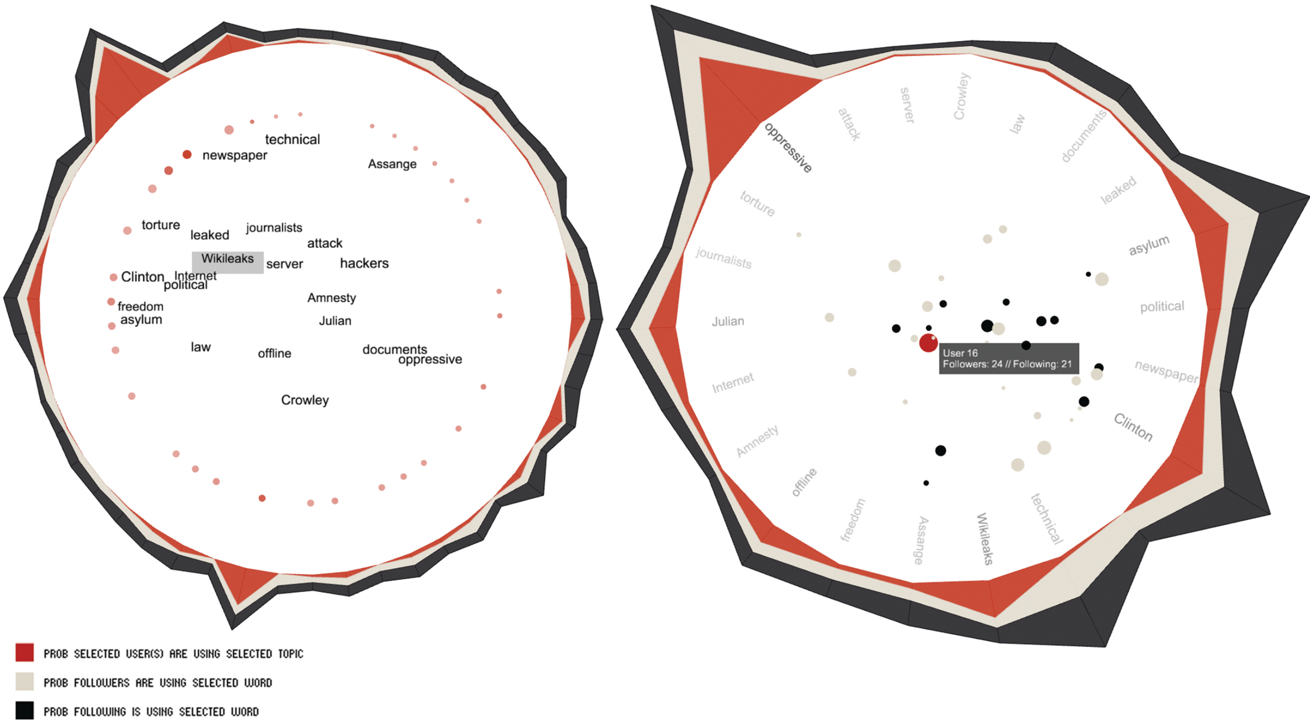

Figure 2: Two detail views of the TopicLens visualization, each showing connections between items and related topics. In this case the data is from the Twitter social network, so our generic "items" represent Twitter users. Frequency measures are shown on the outer river-like component. The two views are of the same data, with items and topics inverted.

unidirectional edges (followers and followees) from Twitter; augmented with topic relations, and a topic modeled network of news articles from the New York Times; and social network data with bi-directional connections from Facebook. Across all examples, the goal is to use simple interaction and novel layouts to facilitate user comprehension of complex data, particularly to communicate the "credibility' factor of peers in a network with respect to particular topics of interest. This complexity would be inherently difficult to communicate with a single visualization technique such as a river or graph visualization. Accordingly, we have opted for a hybrid approach which uses a graph-like mechanism similar to TopicNets [15] for highlighting relations between document and topic nodes, and a river-like view similar to ThemeRiver [17] overlaid to communicate frequency or "credibility" of different sets of peers within the context of a topic selection. This approach has been successful in applications such as Freire's ManyNets [13].

3. DESIGN CONSIDERATIONS

At the core, the TopicLens interface seeks to empower the user to explore a large datasets based on a number of factors. We designed the interface with the idea that potential users would benefit most from learning and engaging in the system rather than making sense of the data at a glance. We see applications of our system being beneficial for any researcher who is looking to glean insights into a large body of text. This includes analysts of social networks as well as scholars in the humanities who may want to use TopicLens to explore trends in the bodies of work by a single author or works belonging to a single or set of genres. We provide functionality with the goal of avoiding a crowded interface and we took great measures to ensure clarity and consistency across multiple view modes. In our informal tests and observations of interactions with the system, we have found it easy to learn and that users take quickly to the dynamic filtering and sorting tools we provide.

4. VISUALIZATION DESIGN

The most prominent feature of the visualization, shown in Figure 2, is its use of the wheel to structure information. Using a circular structure allows us accommodate variability in the size of the datasets. The wheel dynamically expands to fit the data and contracts upon filtering. Zooming and font size are adjusted in order to keep information present within the visualization space, regardless of how much there is to display.

The visualization is designed to fit within a rectangular window with width larger than height. The exact dimensions can vary and in our examples, we found it most effective to use a full screen view on a high-resolution display (1280x1024 and higher), especially when dealing with large sets. The left side of the screen contains the controls and legends and the wheel rotates on an axis in the center of the screen. A static camera is also positioned at the center of the screen, allowing the user to zoom in towards and away from the center. The river is positioned along the outer edge of the wheel and protrudes in different directions depending on the current data selection.

4.1 Organization

In order to support the user in exploring the data at varying levels of detail, the organization of the visualization needs to clearly distinguish the different relationships that are represented. We classify those relationships into three types: primary, secondary and ternary.

The data we collect has pre-existing relationships as formed through meta-data (primary relations), the topic modeling algorithm provides information about relationships between items and topics (secondary relationships), and we found it helpful to further analyze the topics in relation to items and item meta-data (ternary relationships). By dividing the wheel into three concentric regions, we were able to map each type relationship to its own location on the wheel. As you travel from the center out, the information represented reflects a increasing number of factors. The wheel, combined with zooming, was intended to give the user the idea that zooming out will provide them with a big picture, birds-eye overview of the data and zooming in closer will focus on the finer detailed relationships. The following paragraphs provide a detailed explanation of the relationship types and the regions they map to.

4.1.1 Primary Relations: Center

Primary relations are formed though associations in item meta-data. In the analysis of Twitter networks, a single item represents a Twitter user. Item meta-data includes, but is not limited to, a list of followers of this user and a list of other Twitter users that this user is following. In the case of topic modeling run over New York Times articles, primary relationships would be formed between two or more articles that share the same author formed by two articles. Primary relationships are mapped to the center so these relationships can be viewed in a local space. Figure 2 shows primary relations through coloring in the view on the right. In the view on the left, topics are featured in the center. Since primary relations don't exist within topics, no explicit color mapping is represented.

4.1.2 Secondary Relations: Center & Inner Ring

Secondary relationships occur as a result of the topic modeling and define the relationships between topic and item nodes. Each of these relationships occurs with a given probability as defined by the LDA algorithm. These relationships as well as their respective probabilities are represented by interactions between the center and inner ring. While the nodes in the center are not bound to any axis or predetermined path, the nodes in the inner ring are equidistantly laid out in a circle. This is primarily because the inner ring also functions as the axis points for the river visualization but also reinforces simplicity by defining only one type of data to be related spatially. On the left side of Figure 2, highlighting Wikileaks changed the opacity of the nodes on the inner ring in order to indicate how related each item is to this topic. On the right, highlighting User 16 changed the opacity of the topics in the inner ring, similarly showing the strength of the connection.

4.1.3 Ternary Relations: Outer Ring / River

Ternary relationships are formed between the topic modeled results and the meta-information of the items related to those results. Using the river visualization to graph these relationships allows us to see an overall frequency of the node in addition to the meta-information frequencies within the same space. Depending on the data and filtering, the river model can be customized to show any particular facet of the meta-information. Figure 2 is showing average probabilities over each facet of item meta-data in relation to the selected item. The colors in the river match the colors of the meta-data in the center, reinforcing this relationship.

4.2 Visual Mappings

Because the TopicLens visualization needs to encode a rich variety of data, we took care to make the visual encoding of different relationships and concepts distinct. In order to maintain simplicity we map objects and relationships to specific formal elements. Depending on the underlying dataset, visual features may be turned on and off in order to keep the visual complexity to a minimum.

At the root, our information display consists of two basic entities: topics and content items. Items are mapped to circles and topics are mapped to rectangles with the text label of the topic in the center. We made these entities distinct in order to visually and conceptually separate them. The topic text is always visible but the item text is only present on demand. Similarly, the circular shape of the item is always visible but the rectangular shape of the topic is only visible on selection.

Color is used to visually group items based on meta-data. For instance, if there is meta-information about item categories, each category type would map to a unique color. This mapping was chosen partly because it enables a quick visual grouping of items and extends to a large number of categorizations. Another reason for choosing color, was its ability to support a visual connection between the meta-data of the individual item and the corresponding meta-data represented in the river. This offers the user two levels of understanding by illustrating how the meta information is connected to the item as well as the topic.

Opacity is used to illustrate secondary relationships, relationships between topics and content items. These relationships occur with a probability specified by the LDA algorithm. Opacity is an effective means of illustrating these connections as it indicates relative strength. Darker nodes have strong probabilities of relation, lighter have weaker ones. If a node is unrelated, it is removed from the space. Secondary relations are highlighted upon interaction as the user must specify a single item or topic in order to view its connections. If multiple items or topics are selected, then the opacity value is determined by the average probability from all nodes in the selected set.

Position and order are used in conjunction to highlight patterns in the data. Patterns are exposed by using the ordering of the items or topics on the inner ring to position the items or topics in the center. Each value begins in the center of the circle and is pulled towards all of its related nodes in the inner ring. The strength of attraction depends on the probability of the connection between the item and the topic. The result is a spatial grouping of items or topics that share similar relationships. A number of interaction techniques for positioning items on the inner ring will be discussed in the following sections.

Size is used to illustrate measures of numerical magnitude such as frequency or number of relations. Similar to position and ordering, some mappings of attributes to size can be more informative than others. For this reason, we allow the user to indentify the node attribute that determines node size.

5. IMPLEMENTATION

This visualization evolved through a number of design iterations. Using Processing to program the design and interaction allowed us to easily explore changes in the design and instantly see the results. The Processing framework also made it simple to program animations and transitions between states. A number of libraries were used to extend the scope and flexibility ofProcessing. The PeasyCam library provided the basic virtual viewpoint control, the ControlP5 library was used to implement text boxes, range sliders and list boxes and an OpenGL library was used to add custom functionality into the system such as smoothing and alpha blending.

The TopicLens application creates node and edge objects by parsing configuration and data files on load. During the execution of the program, nodes and edge objects are referenced in order to create dynamic links. Links are the elements that are drawn to the canvas and much of the code is devoted to maintaining those links and

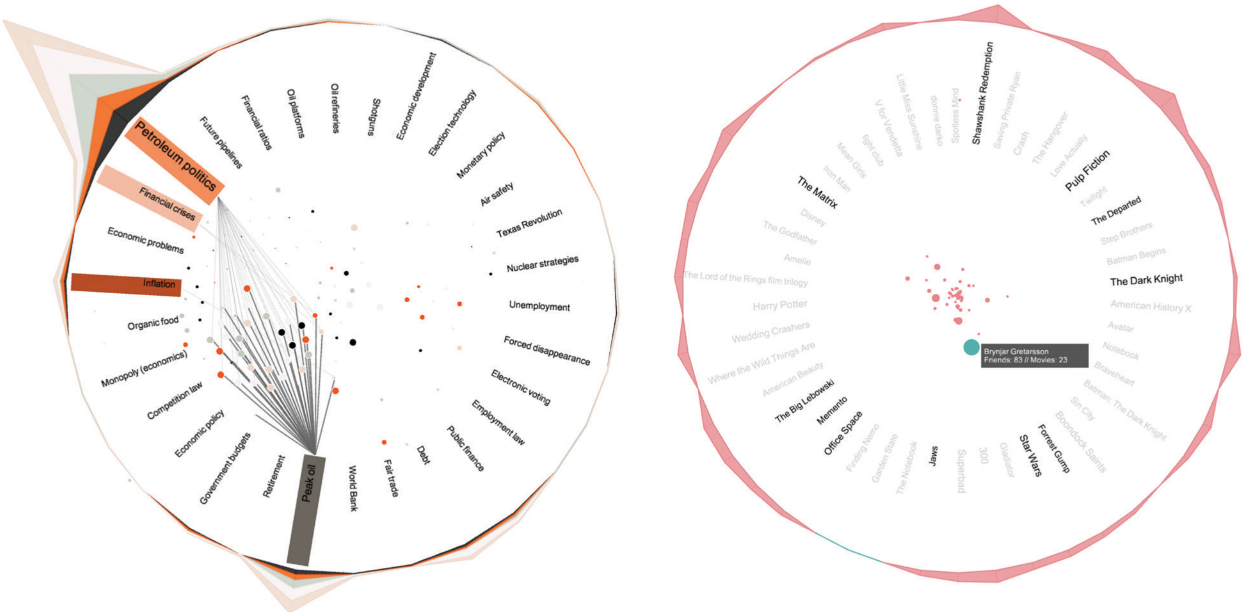

Figure 3: TopicLens view details for varying data sets. Left: TopicLens view of news articles from the New York Times showing topics on the inner ring and articles/items in the center. The view shows selection of an individual topic (peak oil) and edges linking to related articles. Right: TopicLens view of a Facebook social network showing an individual's friend network and associated item preferences.

dynamically updating their values to indicate relationships. The smooth transitions were created using an integrator class that allows the user to specify characteristics such as mass, position, damping and attraction. When a link targets a given position, the integrator dynamically updates its position depending on its physical characteristics.

6. USE CASES AND DISCUSSION

In order to showcase the flexible applicability of our visual model, we present three use cases that explore different dynamic datasets. Each use-case will discuss the design decision made to cater to the specific data domain as well as a usage scenario to illustrate its potential for a variety of applications.

6.1 Recommending Credible Information In Twitter

Preserving social network relations in topic modeled systems allows us to glean insights into the networks and salient topics therein. This example is catered specifically as an attempt to visualize credibility in Twitter networks. Our definition of "credibility" relates to the probability by which a user is connected with a particular topic, based on LDA analysis over a bag of words representation of all of that user's tweets. In analyzing credibility, we also examine that user's followers and followees and their respective associations with the given topic.

In this visualization, which is represented earlier in Figure 2, each topic node contains a label that represents the list of words in a mined topic. Primary relationships are formed between a user and their followers and followees. Secondary relationships occur between users and extracted keywords and the ternary relationships represent probabilities over the meta data. In this example, the user meta data contains a list of other users following this user and a list of the users this user is following. One type of ternary relationship in this example is the relationship between a topic and the average probability that the users friends are discussing that topic.

As noted in Section 4.1, ternary relations are mapped to the river and the nodes in the inner ring form the axis points. The river displays specific information depending on the organization of the nodes within the space. This approach affords the user an opportunity to uncover potentially interesting relations in the following 6 view configurations.

With topic nodes in the center, and user nodes on the outer ring:

- *Upon selection of an individual user*, the river view shows that user, their friends and their followers' probabilistic association with each topic on the outer ring.
- *When no user is selected*, the river shows the average probability for each topic across all users.
- *When a topic is selected*, the river shows each user's association with that topic.

With user/item nodes in the center, and topic nodes on the outer ring:

3
- *When a topic is selected*, the user's friends and follower's opacity is varied to represent association with that topic.
- *When a user is selected*, their association with each topic on the outer ring is shown in the river view.
- *When no user is selected*, the probability of each topic in the global space is shown on the outer ring

For this scenario, the river represents three probabilities for each node, the average probability of the user using the topic, the average probability of the user's followers using the topic and the average probability that the people following this user are using the topic. Since topics are represented along the inner ring, this information is available for every topic. Each of the probabilities is represented on the river, using color matching to indicate the group or single user it applies to. To further explain what the river is visualizing, a legend on the bottom left of the interface dynamically updates, explaining the current model. In this case mode visibility is of particular importance as the river maps different values through the life of the visualization.

When a Twitter user is highlighted in the space, interactions take place at each of the three levels. In the center, the primary relationships are presented through colors. All users who don't belong to this user's network are removed and the remaining users are color coordinated to indicate whether they are a follower of the selected user, or someone the selected user is following. Spatially, each user is attracted to the topic nodes in the inner ring by the positioning algorithm mentioned above. Topics related to the selected node vary by opacity in order to indicate the strength of connection.

The probability mappings were specifically designed to investigate credibility or trust. The top left of Figure 2 shows a network with two people selected. All of the nodes in the set represent both of the selected people's networks. On the outer river, one can see the probability distributes for this network over each topic in the network. From the river you can conclude that these two users are using the topic "Crowley" quite a bit, however their friends and followers are not. For this reason, they may not be a trusted source for this topic since their followers do not appear to be interested in similar topics. On the other side of the visualization is the topic "asylum" which is being used largely by the network and not so much by these users.

Drawing firm conclusions at this level is not necessarily reliable but better information can be introduced by selecting the right network. For instance, if you know the terms "Assange", "Julian" and "Wikileaks" are all terms related to Wikileaks, then you could select those terms from the visualization and view the results over the given network of users associated with those terms. By investigating the probability of these three words occurring together across the social network you may be able to visualize trends about who is followed, by whom and for what reasons.

6.2 Recommending New York Times Articles

In the example shown on the left of Figure 3, topic modeling was performed on a set of New York Times articles and is used for investigation and discovery of related articles that may not have been discovered through traditional search models. Each document node represents an article and topic nodes represent the topics extracted from those articles. Each article contains information about the section of the paper which it belonged to, such as opinion, world or national news.

Two unique design features were included in this interface to improve the functionality in relation to the underlying data. The first one is colored rectangles on topics. These colors are used to reinforce ternary connections though the use of color averaging. The color of the rectangle is determined by the category of each of the articles associated with it. Should a color tend heavily towards a single category's color, on could deduce that the topic tends to appear most frequently within that category. The actual distribution of the categories is explicitly represented in the river.

The second unique feature is the use of lines. When hovering over a topic, darkened lines extend form the topic itself to all related documents. Lighter lines then extend from each of those related documents to all of the other topics they are related to. This conveys information to the user about other topics related to their selection. The user is able to specifically locate the documents that contributed to this relationship by following the lines or selecting multiple topics and browsing the filtered document space.

The lines are particularly useful for illustrating how two topics are related to each other and upon what criteria. This is helpful when browsing for articles associated with a given theme. Let's say a researcher is looking for references on "peak oil." Searching for and selecting "peak oil" from the space would show the researcher other related topics as well as articles specifically contain the relation. If one of the related articles contains a topic that is also of interest to him or presents a particularly interesting comparison, he or she can easily isolate and obtain information about the articles containing both topics by filtering the space and hovering over the document node, revealing information about the article such as title, date and author.

TopicLens could also be used visualize trends associated with a subset of articles. Say a user read a few articles in the Times Opinion section and they would like to find other articles about similar and related subjects. This could be accomplished by typing each article name into the search field. This would in turn select each of the corresponding articles in the space and illustrate the topics associated with them. In order to remove outliers, they would adjust the slider to specify the amount of documents that need to be associate a topic in order for it to appear in the space. After this filtering step, they are presented with a number of related topics, the most popular being the largest and darkest. By selecting that topic, the space is reorganized to show all the the articles related to that topic. The user can now visually browse these articles and quickly identify which one appeared in the opinion section, based on color. Hovering the mouse over a document node would reveal its specific information and provide access to the full text..

6.3 Recommending Movies via Facebook API

In this example, shown in the right of Figure 3, we use the proposed framework to visualize data that is not topic modeled in order to show how the interface also operates on similarly structured datasets. Reinterpreting the definitions of recommendable item and topic allows us to use the existing visual model for this dataset. In this example, a single Facebook user takes the place of a recommendable item and a movie takes the place of a topic. Since movies can be related to any number of Facebook users and Facebook users can be associated with any number of movies, this dataset can function similarly to the topic model examples. Each item-topic, or rather user-movie, combination is assigned the probability of 1 since the user has specified explicitly that they like the given movie. This visualization is able to provide exploratory views of the most popular movies within a Facebook friend network as well as the least popular movies. It can also isolate pockets of users that are fans of these most or least popular movies. Essentially this view is a visual representation of a social collaborative filtering process, since items which are popular among Facebook friends are promoted for a single target user receiving the recommendation.

7. CONCLUSION

In summary, this paper has presented a novel interactive interface for recommending interesting topics and documents from within a large corpus. The design is a hybrid which combines river and graph-like representations of recommended items and can be easily adapted and customized by the end user for different use cases. We have also introduced novel interaction methods that support human skills in the exploration of topic modeled data sets. In doing so, we have extended the efficacy of both the system and the algorithm, allowing the user to navigate large datasets and uncover patterns. Details of our design choices and methodology have been discussed, and demonstrated over three example applications, including social network data from Twitter augmented with topic modeling over users' tweets, a topic modeled set of New York Times news articles, and social network data from Facebook, including item preferences. In each example case, we have discussed ways in which the approach facilitates discovery of relevant information which may go undiscovered in traditional analysis tools. We have also demonstrated TopicLens' ability to act as a flexible interaction layer, supporting exploration of multiple application domains.

8. ACKNOWLEDGMENT

The authors would like to thank Peter Pirolli and Bongwon Suh for use of their Twitter data set. This work was partially supported by the U.S. Army Research Laboratory under Cooperative Agreement No. W911NF-09-2-0053; by NSF grant IIS-1058132; and by the U.S. Army Research Laboratory under MURI grant No. W911NF-09-1-0553; The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of ARL, NSF, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation here on.

9. REFERENCES

[1] M. Analyzer. Matheo analyzer, database analysis and information mapping. http://www.matheo-analyzer.com/, 2010.

[2] K. Andrews, W. Kienreich, V. Sabol, J. Becker, G. Droschl, F. Kappe, M. Granitzer, P. Auer, and K. Tochtermann. The infosky visual explorer: exploiting hierarchical structure and document similarities. *Information Visualization*, 1:166–181, December 2002.

[3] D. Blei and J. Lafferty. Correlated topic models. In *Advances in NIPS 18*, pages 147–154. MIT Press, Cambridge, MA, 2006.

[4] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

[5] I. Borg and P. Groenen. *Modern Multidimensional Scaling: Theory and Applications*. Springer, 2005.

[6] S. Bostandjiev, J. O'Donovan, , and T. Hollerer. Tasteweights: An interactive hybrid recommender system. In *6th ACM Conference on Recommender Systems, Dublin, Ireland, September 9 to 13th*, 2012.

[7] R. Burke. Knowledge-based recommender systems. In *Encyclopedia of Library and Information Systems*, volume 69, 2000.

[8] N. Cao, J. Sun, Y.-R. Lin, D. Gotz, S. Liu, and H. Qu. Facetatlas: Multifaceted visualization for rich text corpora. *IEEE Transactions on Visualization and Computer Graphics*, 16:1172–1181, November 2010.

[9] J. Chang. Reading Tea Leaves: How Humans Interpret Topic Models. 2009.

[10] L. Chen and P. Pu. Critiquing-based recommenders: survey and emerging trends. *User Model. User-Adapt. Interact.*, 22(1-2):125–150, 2012.

[11] D. R. Cutting, D. R. Karger, and J. O. Pedersen. Constant interaction-time scatter/gather browsing of very large document collections. In *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval*, SIGIR '93, pages 126–134, New York, NY, USA, 1993. ACM.

[12] B. D. Davison, T. Suel, N. Craswell, and B. Liu, editors. *Proceedings of the Third International Conference on Web Search and Web Data Mining, WSDM 2010, New York, NY, USA, February 4-6, 2010*. ACM, 2010.

[13] M. Freire, C. Plaisant, B. Shneiderman, and J. Golbeck. Manynets: an interface for multiple network analysis and visualization. In *CHI '10: Proceedings of the 28th international conference on Human factors in computing systems*, pages 213–222, New York, NY, USA, 2010. ACM.

[14] B. Fry. *Visualizing data - exploring and explaining data with the processing environment*. O'Reilly, 2008.

[15] B. Gretarsson, J. O'Donovan, A. Asuncion, D. Newman, S. Bostandjiev, T. Hï£¡llerer, and P. Smyth. Topicnets: Visual analysis of large text corpora with topic modeling. *ACM Transactions on the Web: Special Issue on Intelligent Text Visualization*, 16:1172–1181, November 2011.

[16] T. L. Griffiths and M. Steyvers. Finding scientific topics. *PNAS*, 101(Suppl 1):5228–5235, 2004.

[17] S. Havre, E. Hetzler, P. Whitney, and L. Nowell. Themeriver: Visualizing thematic changes in large document collections. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):9–20, 2002.

[18] M. A. Hearst. Tilebars: visualization of term distribution information in full text information access. In *CHI '95: Proc. of the SIGCHI Conf.*, pages 59–66, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

[19] J. L. Herlocker, J. A. Konstan, and J. Riedl. Explaining collaborative filtering recommendations. In *CSCW*, pages 241–250, 2000.

[20] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22:5–53, January 2004.

[21] B. Knijnenburg, J. O'Donovan, S. Bostandjiev, and A. Kobsa. Inspectability and control in social recommenders. In *6th ACM Conference on Recommender Systems, Dublin, Ireland, September 9 to 13th*, 2012.

[22] S. Koch, H. Bosch, and T. Ertl. Towards content-oriented patent document processing. *IEEE Symposium on Visual Analytics, Science and Technology*, pages 203–210, 2009.

[23] Q. Mei, X. Shen, and C. Zhai. Automatic labeling of multinomial topic models. In *KDD*, 2007.

[24] N. E. Miller, P. Chung Wong, M. Brewster, and H. Foote. Topic islands - a wavelet-based text visualization system. In *Proceedings of the conference on Visualization '98*, VIS '98, pages 189–196, Los Alamitos, CA, USA, 1998. IEEE Computer Society Press.

[25] D. Newman, T. Baldwin, L. Cavedon, E. Huang, S. Karimi, D. Martinez, F. Scholer, and J. Zobel. Visualizing search results and document collections using topic maps. *Web Semantics: Science, Services and Agents on the World Wide Web*, 8(2-3):169–175, July 2010.

[26] D. Newman, Y. Noh, E. Talley, S. Karimi, and T. Baldwin. Evaluating topic models for digital libraries. In *JCDL '10: Proceedings of the 10th annual joint conference on Digital libraries*, pages 215–224, New York, NY, USA, 2010. ACM.

[27] K. A. Olsen, R. R. Korfhage, K. M. Sochats, M. B. Spring, and J. G. Williams. Visualization of a document collection: the vibe system. *Inf. Process. Manage.*, 29(1):69–81, 1993.

[28] Questel. Qpat, intellectual property patent and trademark searching. http://www.qpat.com/, 2010.

[29] D. Ramage, S. Dumais, and D. Liebling. Characterizing microblogs with topic models. In A. Cohn, editor, *Proceedings of the Fourth International on Weblogs and Social Media*, pages 130–137. AAAI Press, 23–26 May 2010.

[30] J. Reilly, K. McCarthy, L. McGinty, and B. Smyth. Dynamic critiquing. In *ECCBR*, pages 763–777, 2004.

[31] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: An open architecture for collaborative filtering of netnews. In *CSCW*, pages 175–186, 1994.

[32] D. A. Rushall and M. R. Ilgen. Depict: Documents evaluated as pictures. visualizing information using context vectors and self-organizing maps. In *Proceedings of the 1996 IEEE Symposium on Information Visualization (INFOVIS '96)*, INFOVIS '96, pages 100–, Washington, DC, USA, 1996. IEEE Computer Society.

[33] U. Shardanand and P. Maes. Social information filtering: Algorithms for automating "word of mouth". In *CHI*, pages 210–217, 1995.

[34] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations, 1996.

[35] B. Shneiderman and M. Wattenberg. Ordered treemap layouts. *Information Visualization, IEEE Symposium on*, 0:73, 2001.

[36] W. Spangler, J. Kreulen, and J. Lessler. Mindmap: Utilizing multiple taxonomies and visualization to understand a document collection. *Hawaii International Conference on System Sciences*, 4:102, 2002.

[37] J. T. Stasko, C. Görg, and Z. Liu. Jigsaw: supporting investigative analysis through interactive visualization. *Information Visualization*, 7(2):118–132, 2008.

[38] N. Tintarev and J. Masthoff. A survey of explanations in recommender systems. In *Data Engineering Workshop, 2007 IEEE 23rd International Conference on*, pages 801–810. IEEE, 2007.

[39] Wanner. Towards content-oriented patent document processing. *World Patent Information*, 30(1):21–33, 2008.

[40] J. A. Wise, J. J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow. Visualizing the non-visual: Spatial analysis and interaction with information from text documents. In N. D. Gershon and S. Eick, editors, *IEEE Information Visualization '95*, pages 51–58. IEEE Computer Soc. Press, 30–31 Oct. 1995.

[41] P. C. Wong, B. Hetzler, C. Posse, M. Whiting, S. Havre, N. Cramer, A. Shah, M. Singhal, A. Turner, and J. Thomas. In-spire infovis 2004 contest entry. *Information Visualization, IEEE Symposium on*, 0:r2, 2004.

CoFeel: Using Emotions for Social Interaction in Group Recommender Systems

Yu Chen
Human Computer Interaction Group
Swiss Federal Institute of Technology
CH-1015, Lausanne, Switzerland
yu.chen@epfl.ch

Pearl Pu
Human Computer Interaction Group
Swiss Federal Institute of Technology
CH-1015, Lausanne, Switzerland
pearl.pu@epfl.ch

ABSTRACT
Group and social recommender systems aim to suggest items of interest to a group or a community of people. One important issue in such environment is to understand each individual's preference and attitude within the group. Social and behavioral scientist have evidenced the role of emotions in group work and social communication. This paper aims to examine the role of emotion for social interaction in group recommenders. We implemented CoFeel, an interface that aims to provide emotional input in group recommenders. We further apply CoFeel in a GroupFun, a mobile group music recommender system. Results of an in-depth field study show that by exchanging feelings with other users, CoFeel motivates users to provide feedback on recommended items in a natural and enjoyable way. Results also show that emotions do serve as an effective and promising element to elicitate users' attitudes, and that they do have the potential to increase user engagement in a group. Based on suggestions collected from users, we propose other potential recommendation domains of CoFeel.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces –*Graphical user interfaces (GUI), User-centered design.* H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces - *Organizational design, Web-based interaction*

General Terms
Design, Human Factors

Keywords
Group and Social Recommender Systems, Interface Design, Mobile Interface, Affective Interface, Emotional Feedback

1. INTRODUCTION
Nowadays, sharing, coordination, cooperation and communication among group members are becoming indispensible in online environment. Such groups can be constituted by families selecting a recipe together, colleagues working on same projects, and social club members planning a culture event. These are examples of small groups, normally less than hundreds of people. In group environment, group decision-making becomes a problem due to information overload. Group recommender systems (GRSs) aim to alleviate information overload by suggesting items to a group of people.

Group recommendation problem is not only "the sum of members" (Jameson, 2004). As the audiences move from individuals to groups of people, challenges arise such as aggregating preferences and arriving at equilibrium point of expectations. Picture yourself sitting together with your friends and selecting a music playlist for a birthday party. The selection process does not only depend on the verbal indication on preferences and choices, but also on various non-verbal channels such as individuals' emotion within the group. Social and behavioral scientists have long been studying the social role of emotion in group environment. Our goal is to set a basic understanding of using emotion for social interaction in group recommenders with a particular focus on the following two questions.

1) What are the roles of emotional information in group recommender systems?
2) How to design such interface that is useful, easy to use and playful?

To answer these questions, we introduce CoFeel, an affective interface that allows users to provide emotional input in recommender systems. We further implemented CoFeel in GroupFun, a mobile group music recommender system. The rest of the paper is organized as followed. Section 2 discusses existing work and how they related with our work, and particularly, why emotions play an important role in group and social environment. This is followed by design and usage of CoFeel interface in Section 3 and how to apply CoFeel in GroupFun in Section 4. After reporting the results of a small-scale qualitative user study in Section 5, this paper discusses further application scenarios in Section 6 and concludes with limitations and future work.

2. RELATED WORK

2.1 Social Interaction in Group Recommender Systems
Jameson studied some of the key user issues for group recommender systems (Jameson, 2004) and investigated several measures for promoting collaborating and coordination. These measures mainly aim at designing user interfaces to enhance mutual awareness. Mutual awareness in group recommender systems includes membership awareness, preference awareness and decision awareness.

Membership awareness allows users to check which users are in the group. Being aware of members in a group facilitates users to decide how to behave and thus enhances trust in a group

Interface@RecSys*'12*, September 13, 2012, Dublin, Ireland

recommender (Yu, Zhou, Hao, & Gu, 2006).
Preference awareness enables users to be aware of the preferences of other members. A user study on PolyLens reveals that users would like to see each other's preference information, even at the expense of some degree of privacy loss. Preference awareness in group recommender systems are categorized into three levels: zero awareness, partial awareness and full awareness. Zero preference awareness means that users only know their own preferences, as shown in MusicFX (J. F. Mccarthy, 1998a). Zero preference awareness systems are simple but do not inspire user trust. Partial awareness in group recommenders allows users to apply preference information from other group members (Kudenko, Bauer, & Dengler, 2003). However, it is prone to social loafing, a phenomenon when people contribute less in a social environment than when they work individually. In full preference awareness, users are aware of other members' preferences. One typical technique for is Collaborative Preference Specification (CPS) (Jameson, 2004), as presented in CATS, PocketRestaurantFinder (J. F. Mccarthy, 1998b) and Travel Decision Forum. CPS in group recommender systems enables persuasion, supports preference explanation and justification and reduces conflict. Decision awareness is important in helping users arrive at a final decision. Decision awareness is a status in which users are aware of the decision making process of other members. Existing group recommender systems include the following decision making styles: (1) zero awareness - simply translating the most highly rated solution into action without the consent of any user (e.g. in MusicFX), (2) partial awareness - one or a selected set of representatives of the group are responsible for making the final decisions (e.g. INTRIGUE and PolyLens), and (3) full awareness - arriving at final decision through face-to-face discussions (e.g., CATS) or mediated discussions (e.g., MIAU (Kudenko et al., 2003) and Travel Decision Forum). However, none of the work addresses the role of emotion in decision-making or group interaction.

2.2 Interface in Group Recommender Systems

"Group interfaces differ from single-user interfaces in that they depict group activity and are controlled by multiple users rather than single user" (Ellis, J. Gibbs, & Rein, 1991). Therefore, interface adequacy has more requirements in group recommenders compared with individual recommenders. Flytrap (Crossen, Budzik, & Hammond, 2002) visualizes recommended items by using colors and locations. Songs personalized for different users are displayed with different colors, and the closer the songs are to the center, the more likely they will be played. PolyLens (Connor, Cosley, Konstan, & Riedl, 2001) supports three models of visualizing recommendation UI. Group-only interface only displays movies from group recommendation. Composite interface displays a list of recommended movies with both group and individual member predictions. Individual-focused interface shows the items for other individual users' preferences. CATS (K. Mccarthy et al., 2006) offers users personal space and group space. In group space, each user has a snowflake with a different color and the size of snowflake indicates preferences of individual users. This allows users to check the interest of other users for a particular resort. Additionally, each icon presents a resort, and its size grows or shrinks in accordance with the preference of the whole group.
Travel Decision Forum (Taylor, Ardissono, Goy, & Petrone, 2003) introduces an animated character for each group member currently not available for communication. By responding with speech, facial expressions, and gesture to proposed solutions; a representative conveys to the current online users some key aspects of its corresponding offline user's responses to a proposed solution. This is one of the few work that employs non-verbal channels in group environment.

2.3 Emotion in Recommender Systems

Musicovery[1] and *Stereomood*[2] have developed an interactive interface for users to select music category based on their mood. *Musicovery* classifies mood by two dimensions: dark-positive and energetic calm. It uses highly interactive interface for users to experience different emotion categories and their corresponding music. However, such recommender does not support interaction in social group environment. The main goal of studying recommender systems is to improve user satisfaction. However, satisfaction is a highly subjective metric. Masthoff and Gatt (Masthoff, 2005) have considered satisfaction as an affective state or mood based on the following aspects in socio- and psycho-theories: 1) mood impacts judgement; 2) retrospective feelings can differ from feelings experienced; 3) expectation can influence emotion and 4) emotions wear off over time. However, they did not propose any feasible methods to apply the above psychological theories. They also proved that in group recommender systems, members' emotion can be influenced by each other, and this phenomenon is called emotional contagion.

2.4 Emotions and Decision Making

Our everyday experiences leave little doubt that our emotions can influence decisions we make. For instance, experiment results (Raghunathan & Pham, 1999) showed that in gambling decisions, as well as job-selection decisions, sad individuals are biased in favor of high-risk and high-reward options, whereas anxious individuals are biased in favor of the opposite. On the other hand, (Isen, 2001) reveals evidence that in most circumstances, positive affect enhances problem solving and decision making, leading to cognitive processing that is not only flexible, innovative, but also thorough and efficient. (Schwarz, 2000) has addressed the influence of moods and emotions experienced at the time of decision making, affective consequences of decisions and the role of anticipated and remembered affect in decision making. (Bechara, 2004) further proves the influence of emotions on decision-making from neurology. (Velásquez, 1997) and (Gratch & Rey, 2000) also modeled emotion-based decision making.

2.5 Social Role of Emotions

(Keltner, 1999) integrate claims and findings concerning the social functions of emotions at the individual, dyadic, group, and cultural levels of analysis. On dyadic level (a group of two), emotional expressions help individuals know others' emotions, beliefs and intensions, and thus rapidly coordinating social interactions. Emotional communication also evokes complementary and reciprocal emotions in others that help individuals respond to significant social events. Emotions serve as incentives or deterrents for other individuals' social behavior. On group level, emotions have claimed to help individuals solve the problem of identifying group members. Displaying emotions may help individuals define and negotiate group-related roles and status. Collective emotional behavior may also help group members negotiate group-related problems. Study results from

[1] Musicovery. http://musicovery.com/

[2] Stereomood. http://www.stereomood.com/

(Ketelaar & Tung Au, 2003) are discussed in terms of an "affect-as-information" model, which suggests that non-cooperating individuals who experience the negative state associated with guilt in a social bargaining game may be using this feeling state as "information" about the future costs of pursuing an uncooperative strategy.(Bowles & Gintis, 2002) suggest that prosocial emotions, such as shame, guilt, (K. Mccarthy et al., 2006)pride, regret, and joy, play a central role in sustaining cooperative relations, including successful transactions in the absence of complete contracting. (Hareli & Rafaeli, 2008) propose that organizational dyads and groups inhabit emotion cycles: emotions of an individual influence the emotions, thoughts and behaviors of others; others' reactions can then influence their future interactions with the individual expressing the original emotion, as well as that individual's future emotions and behaviors. (Barsade, 2001) proved that the leaders transfer their moods to group members and that leaders' moods impact the effort and the coordination of groups. (Hancock et al., 2008) have investigated emotion contagion and proved that emotions can be sensed in text-based computer mediated communications.

3. CoFeel: Providing Emotional Input

3.1 Design Goals

As the first step to investigate the social role of emotions, we design an interface that helps users to provide emotional input. Since this input is also users' feedback, we cross-use "emotional input" in this paper. We refer to the guidelines for designing recommender systems, proposed by (Pu, Chen, & Hu, 2011). Designing CoFeel should meet the following design principles.

1. **Usefulness**. Users are able to provide emotional feedback using CoFeel.
2. **Ease to use.** Users find CoFeel easy to use and easy to learn.
3. **Playfulness**. Users find it fun, playful and entertaining to use CoFeel.

3.2 What is it?

CoFeel aims to enhance group experience by enhancing self-presence and mutual awareness within a group. By exchanging feelings with other users, CoFeel aims to motivate users to provide feedback on recommended items in a natural and easy way. It is implemented as an infrastructure, which can be easily extended to various group recommendation domains.

We choose Geneva Emotion Wheel (GEW) introduced by Scherer (Scherer, 2005) for users to label emotions, i.e., attitude to recommended items, see Figure 1. Using GEW to label emotion has two advantages: natural tagging of discrete categorical words and the possible mapping of these labels to a two-dimensional space (valance-arousal). In each emotion, users can choose different sized circle. As such, users can assign different intensity values to the emotion they choose.

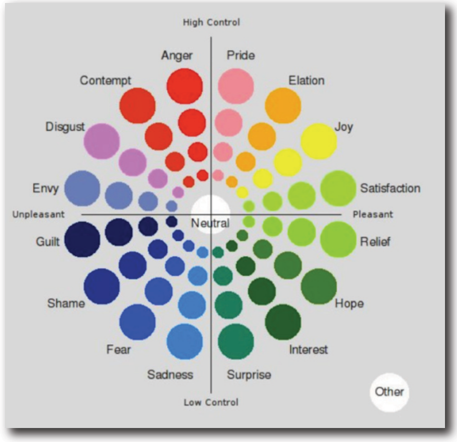

Figure 1. Geneva Emotion Wheel (Scherer, 2005)

We adopt Scherer's color wheel style and choose 8 emotions for CoFeel Emotion Plate: excited, joyful, surprised, calm, sad, fear, distressed, aroused, as is shown in Figure 2. Each emotion class provides a scale from 1 to 5 indicating the intensity of the emotion. In order to enhance user engagement in interacting with the CoFeel, we design each emotional position as a hole and a ball is rolling on the surface of emotion plate. Users interact with the plate by placing the ball in the hole that corresponding to the emotional state. The aim of using the plate-hole-ball metaphor is to enhance user affordance to interact with the interface.

Figure 2. Interface of CoFeel Emotion Plate

3.3 How to use it?

We implement CoFeel emotion plate on mobile phones. Since we have chosen the metaphor of a plate, it is natural that a ball can roll around the surface. Users can select the emotion, i.e., place the ball, by tilting the plate surface. Once users confirm an emotion, they can simply click a 'track' button, which is around the emotional plate, see Figure 3. The phone detects user movement and direction of surface plate using sensors on mobile phones, i.e., accelerometer and gyroscope. We designed this way in order to make the proces more fun and engaging. We have also filtered out constant accelormeter data when users are walking, travelling and etc. In this way, users can input their emotion in a stable way.

Figure 3. Interacting with CoFeel Emotion Plate

4. PROTOTYPE

4.1 GroupFun: a music recommender system

In order to test the applicability of CoFeel, we implemented GroupFun, a mobile group music recommender system. Its function is to come up with common playlists for user created groups. Users can create groups and share their music taste with their group members by rating songs in GroupFun. When GroupFun generates a common playlist for a group, the criterion is to take into account the music taste of all of its contributing members. Figure 4 shows the group function of GroupFun. Users can use CoFeel for two purposes: 1) providing emotional feedback to a song and 2) leaving mood traces on the timeline of a song.

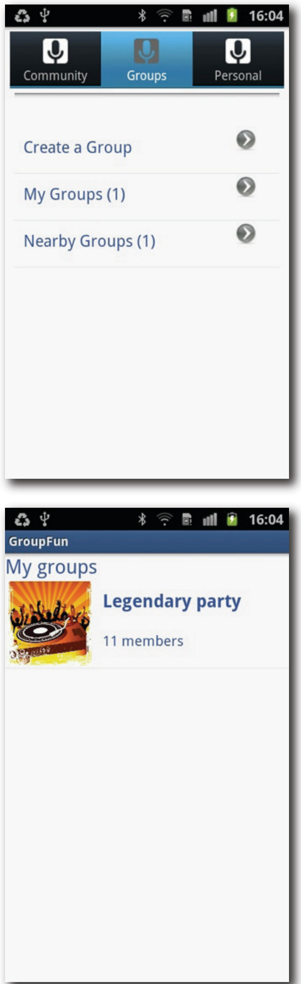

Figure 4. Group function of GroupFun

4.2 Providing Emotional Feedback to a Song

Emotional feedback can be used as an explanation interface for rating. Users can choose the emotion category and its intensity using CoFeel, see Figure 5. As we introduced in Section 3.3, users hold the phone and roll around the indicator ball around the surface of emotion plate, as is shown in Figure 6.

Figure 5. Providing emotional feeback to a whole song

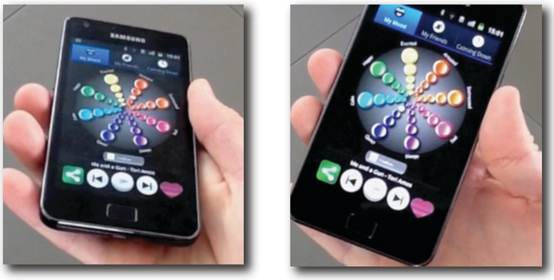

Figure 6. Interacting with CoFeel in GroupFun

After selecting, emotional feedback is recorded with the song, as is shown in Figure 7. The color dots right to the title of a song indicates the type and intensity that users have chosen, which correspond to the colors in CoFeel. The intensity of emotions is visulized with transparency of circles. For example, the song 'We will rock you', is rated as an 'exciting' song, with the level of 3 out 5.

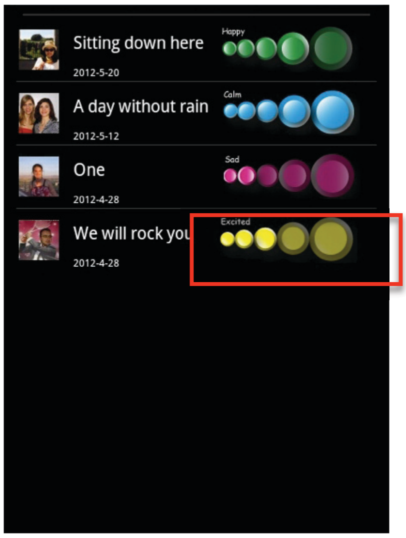

Figure 7. Visualzing friend's emotional feedback in GroupFun

4.3 Emotional Traces in Timeline of a Song

Users can also leave emotional traces throughout the timeline of a song. Figure 8 is an example way to visualize the traces as music score. User emotions are distinguished by different colors, corresponding with colors in CoFeel. Intensities of emotions correspond to the line. The position of dots in the lines represents the relative position of the moment when user leaves emotional comments. For example, a user is listening to "Paradise" from Coldplay. The last two red dots represent users' emotion towards the end of the song: aroused with different levels of intensities.

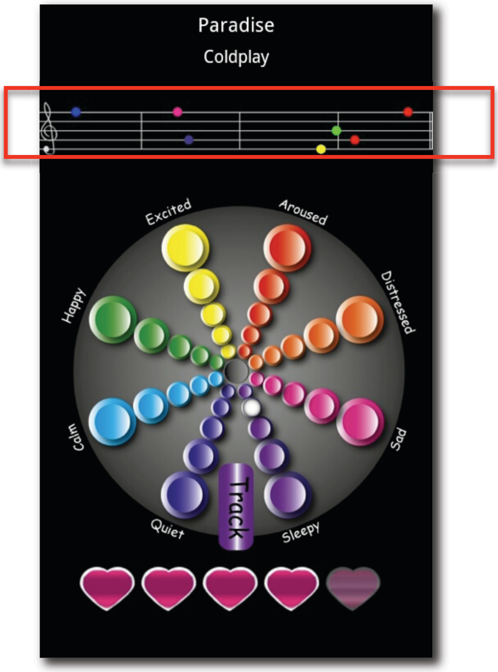

Figure 8. Leaving emotional comments in a timeline for a song

5. Experiment

5.1 Goals

To the best of our knowledge, our work is the first to propose providing emotional feedback in group recommender systems. Therefore, the main purpose of evaluation is not to prove its superiority to other means of feedback or replace them. Rather, we aim to understand users' opinions towards emotional feedback and the design of CoFeel interface, including their degree of acceptance and suggestions. To be more specific, we aim to investigate two research questions:

1) Is emotional feedback useful for social interaction in group recommender systems?
2) Has CoFeel successfully been designed as an effective and playful interface to provide emotional feedback?

5.2 Design and Procedure

In order to answer the above questions, we carried out a small-scale qualitative user experiment, with emphasis on learning from users through active listening, inspection and observation. In addition to normal users, we also showed GroupFun to domain experts. Based on the above two types of interviewed users, we divide the experiment to two steps.

Step 1: Evaluate with normal users

The goal of evaluate with normal users is to observe how they interact with CoFeel, particularly whether they have encountered any usability problems. However, we evaluate CoFeel interface using GroupFun, *without* explicitly telling users what we were evaluating and observing.

Four users participated in the experiment. Each user is distributed with an Android phone installed with GroupFun. Before experiment, we assigned each participant with a specific group with 11 members. The 11 members come from his/her Facebook friends. Each group is recommended with a music playlist. Since the accuracy of recommendation is out of scope of this paper, we use choose most popular songs, i.e., top 40 songs in the experiment week.

Before exposing users with application and systems, we ask the following questions to warm them up.

1) How often do you listen to music?
2) In which context do you listen to music?
3) Which kind of device do you use to listen to music?
4) What do you think about the relation between music and emotion?
5) Do you share music among friends?

During the experiment, the participants explore and experience GroupFun freely, with particular focus on CoFeel interface. We observe how they interact with GroupFun and CoFeel, the whole process of which is recorded. In the meantime, they can ask any questions and raise their concerns. After the experiment, we ask for users' comments.

Step 2: Interview domain expert

Different from experiment with normal users, the goal of interviewing domain experts is to understand the role of emotions in social and group environment and whether CoFeel contributes to this purpose. Additionally, the focus shifts from observation to listening for their feedback and suggestions. We invited a doctor in the field of social psychiatry and interviewed them for feedback in emotional design. They first briefly play around with GroupFun and CoFeel then commented on the design from the theoretical function point of view.

5.3 Results

Step 1: Evaluate with normal users

We summarize the demographic information as below in Table 1.

ID	User 1	User 2	User 3	User 4
Occupation	Student	Student	Student	Consultant
Gender	Male	Female	Male	Female
Age	22	26	25	32
Music exp. (App.)	>12 h/day	>8h/day	2 h/day	2-4 h/day
Devices for listening	Mobile phones and laptop	Computer, car, MP3 player	Computer	Mobile phones
Music context	Studying, designing, walking	Working, cooking, driving, before sleep	Relaxing	Travelling, meditation, music lessons
Sharing music with friends	Spotify, Facebook	Facebook, Twitter, Google +	Google+, Facebook, Email	CDs, DVDs

Table 2. Demographic information of interviewed users

From the interview process, we discovered some interesting phenomenon.

1. They hardly notice that music is more frequent in their life than their perception. When asked how often they listen to music, 3 out of the 4 interviewed users answered: not very often. However, when we ask them to recall the last song they listened to recently, they finally discover much more scenarios and time that they listen to music. This implies that users tend to use listening to music as background tasks.

2. The methods they listening to music tend to be mobile and pervasive. From user evaluation, we found that 3 out of 4 users listen to music on the go. Such mobile devices can be smart phones, mp3 players, laptops, in-car entertaining system and etc.

3. They choose music based on different context. When asked what types of music they listen to. Their answers usually start with "er", "well, depends…". Then they elaborate how they choose music in different contexts, e.g., studying, driving, cooking etc.

4. They are intrinsically willing to share music among friends. Surprisingly, all interviewed users share and discuss about songs among their friends. As one user mentioned, "I share a song with friends, either because I like it, or I think my friend may like it, or it include our shared memory, or it suits the current context."

We further observe users when they are playing around CoFeel emotion plate in GroupFun. Not surprisingly, we observed some common phenomenon during their interaction with system.

1. During the whole process they interact with GroupFun, they spend the majority of time exploring CoFeel, out of curiosity and fun.

2. The first time when they saw the interface, their mental model of choosing emotion is by clicking. After few seconds, they realized how the ball is moving.

3. They learnt to use CoFeel to keep track of their mood in very short time.

This implies that given the fact that CoFeel is a novel interface, users enjoy playing with it and can learn how to use it in short time.

After using interacting with the system, we further interviewed them for feedback on the design of CoFeel and its usage in GroupFun. We received both many encouraging and promising comments as well as suggestions.

Overall, users were excited to talk about CoFeel emotion plate. As users commented: "The plate reminded me of a game I played when I was young, very intuitive and entertaining." "It is simply artistic and charming." "I like the visual effect. It is beautiful". From the received comments, users are generally impressed by the visual effect of CoFeel.

When asked whether CoFeel, i.e., emotional feedback, is useful in GroupFun, all of them agree it is useful. "It is interesting way to comment on a song." "In this way, my friend understand why I like this song and I also know their styles and favorite songs better." "I used the emotional re-tweeting function in one micro-blogging system, which is a fast and convenient way to express multi-dimensional meanings." "Sometimes I don't know how to express my feeling and comments for a song. They are abstract and I'm a person of few words. Emotional feedback looks like I'm choosing my comments from a set of words. It is a take-away style. Everything is predefined and very quick."

At the mean time, they suggest further application scenarios for using CoFeel in social interaction. "It will be interesting to see a music messaging system where people communicate emotions via music." "What about an interface for mixed emotions?" "Re-tweeting a song attached with emotions would be cool!"

From the qualitative analysis above, we conclude that CoFeel has fulfilled the goals we have set in Section 3.1: *usefulness*, *ease of use* and *playfulness*.

Step 2: Interview domain expert

Furthermore, we interviewed a doctor in children and adolescent psychiatry. From mental health perspective, he pointed out that discussing with friends with/using music is also used to enhance people's mental health. This process is called music therapy. Music and mood is by nature connected. Meanwhile, encouraging discussion about mood among a social group also brings benefit to enhance users' mental state, under the condition that the process should be fun. This method is also known as social therapy. He also commented on GroupFun with CoFeel as followed. "Your software, I find it very interesting, especially the idea of self-regulation by the music and the group's involvement even if it is a virtual interaction. In short, fun and social group, they are two very important elements, not just for people with depression, but also for everyone who is interested in this type of language. Every day, we all have moments of frustration and we all seek for self-solutions and be content with a group that gives us support and sense of belonging."

From the interview results, we find that theoretically providing emotional feedback has a positive effect on encouraging group interaction and engagement. A further discovery is that social interaction that takes place within a group also enhances user mood and mental state.

5.4 Implications

We summarize the findings from the above user study about providing emotional feedback in group recommender systems.

1. Providing emotional feedback enhances **mutual awareness** of user preferences within a group. Users know the reasons their friends like a song.

2. A well-designed interface for emotional feedback offers social affordance and invites **users engagement** in the system. When users know the items their friends like and the reasons of liking, they are more likely to experience the recommended items, i.e., music. This encourages users to be more engaged in the system.

3. Social interaction in turn strengthens users' sense of **social belonging** and enhances their emotional state.

6. LIMITATIONS AND DISCUSSIONS

This work has some limitations that we would like to continue in the future. First of all, CoFeel collects explicit emotions reported by users. Sometimes, users are not aware of their emotional attitude. Thus we also aim to consider users' implicit emotional feedback. Additionally, the study is limited within individuals with manipulated friend groups instead of users within a group. Furthermore, as an in-depth qualitative user study, we only invited a few users and domain experts. In order to further validate our hypotheses, we need more groups and users and conduct larger scale user studies for quantitative analysis. It would also be interesting to let users use GroupFun with their friends in real life and observe their behavior and attitude.

Despite of the limitations, using emotions for social interaction implies a much broader usage context. CoFeel not only applies in music recommender systems but also various other domains. Based on feedback received from interviewed users, we propose the following example domains where emotional feedback can be useful: movies, tourists, product, hotels, food and etc. One thing in common in the above domain is the capability for the items to elicit emotions. This has been cross validated by social and behavioral scientists.

7. CONCLUSIONS

We hypnotize that using emotion to enhance social interaction in group recommenders. We have implemented CoFeel, with the goal of designing an interface that is easy to use and enjoyable for users to leave emotional attitude. We further applied CoFeel in GroupFun, a group music recommender system in mobile phones. CoFeel can be used in two modes in GroupFun: elicitation of emotional attitude towards a whole song or emotional traces in the timeline of a song. We then conducted an in-depth qualitative experiment with users, observing their interaction with GroupFun and CoFeel, followed by interviews with them. Besides normal users, we also showed our prototype to domain experts and received positive feedback from them, both theoretically and practically. Results show that providing emotional feedback not only enhances mutual awareness of user preferences, but also encourages social interaction. In essence, providing such social affordance using emotions in group environment in turn promotes users' enthusiasm in interacting with system. Based on discussion with users, we are more convinced that emotional feedback, i.e., CoFeel, applies not only in music domain, but also in many others, such as travel, movie and product recommendations.

8. ACKNOWLEDGEMENT

We thank Swiss National Science Foundation for supporting this work and project. We also thank all participants for their interest in our project, their valuable time and suggestions.

9. REFERENCES

[1] Barsade, S. G. (2001). Working Paper Series OB Organizational Behavior " The Ripple Effect : Emotional Contagion In Groups ", (August).

[2] Bechara, A. (2004). The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage. Brain and cognition, 55(1), 30-40. doi:10.1016/j.bandc.2003.04.001

[3] Bowles, S., & Gintis, H. (2002). Prosocial emotions.

[4] Connor, M. O., Cosley, D., Konstan, J. A., & Riedl, J. (2001). PolyLens : A Recommender System for Groups of Users. ECSCW 2001 (pp. 199-218).

[5] Crossen, A., Budzik, J., & Hammond, K. J. (2002). Flytrap: intelligent group music recommendation. the 7th international conference on Intelligent user interfaces (pp. 184-185).

[6] Ellis, C., J. Gibbs, S., & Rein, G. . (1991). Groupware: some issues and experiences. Commun. ACM, 34(l).

[7] Gratch, J., & Rey, M. (2000). Modeling the Interplay Between Emotion and Decision-Making.

[8] Hancock, J., Gee, K., & Ciaccio, K. (2008). I'm sad you're sad: emotional contagion in CMC. Proceedings of the 2008 ACM, 295-298. Retrieved from http://dl.acm.org/citation.cfm?id=1460611

[9] Hareli, S., & Rafaeli, A. (2008). Emotion cycles: On the social influence of emotion in organizations. Research in Organizational Behavior, 28, 35-59. doi:10.1016/j.riob.2008.04.007

[10] Isen, A. M. (2001). An Influence of Positive Affect on Decision Making in Complex Situations: Theoretical Issues With Practical Implications. Journal of Consumer Psychology, 11(2), 75-85. doi:10.1207/S15327663JCP1102_01

[11] Jameson, A. (2004). More Than the Sum of Its Members : Challenges for Group Recommender Systems, 48-54.

[12] Keltner, D. (1999). Social functions of emotions at four levels of analysis. Cognition & Emotion, (July 2012), 37-41. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/026999399379168

[13] Ketelaar, T., & Tung Au, W. (2003). The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. Cognition & Emotion, 17(3), 429-453. doi:10.1080/02699930143000662

[14] Kudenko, D., Bauer, M., & Dengler, D. (2003). Group Decision Making through Mediated Discussions, 238-247.

[15] Masthoff, J. (2005). The Pursuit of Satisfaction : Affective State in Group Recommender Systems. User Modelling.

[16] Mccarthy, J. F. (1998a). MUSICFX : An Arbiter of Group Preferences for Computer Supported Collaborative Workouts. ACM conference on Computer supported cooperative work.

[17] Mccarthy, J. F. (1998b). PocketRestaurantFinder: A Situated Recommender System for Groups (pp. 1-10).

[18] Mccarthy, K., Salam, M., Coyle, L., Mcginty, L., Smyth, B., & Nixon, P. (2006). CATS : A Synchronous Approach to Collaborative Group Recommendation. American Association for Artificial Intelligence (pp. 86-91).

[19] Pu, P., Chen, L., & Hu, R. (2011). A user-centric evaluation framework for recommender systems. Proceedings of the fifth ACM conference on Recommender systems - RecSys '11, 157. New York, New York, USA: ACM Press. doi:10.1145/2043932.2043962

[20] Raghunathan, R., & Pham, M. (1999). All Negative Moods Are Not Equal: Motivational Influences of Anxiety and Sadness on Decision Making. Organizational behavior and human decision processes, 79(1), 56-77. doi:10.1006/obhd.1999.2838

[21] Scherer, K. R. (2005). What are emotions? And how can they be measured? Social Science Information, 44(4), 695-729. Retrieved from http://ssi.sagepub.com/cgi/content/abstract/44/4/695

[22] Schwarz, N. (2000). Emotion, cognition, and decision making. Cognition & Emotion, 14(4), 433-440. doi:10.1080/026999300402745

[23] Taylor, P., Ardissono, L., Goy, A., & Petrone, G. (2003). INTRIGUE: Personalized recommendation of tourist attractions for desktop and handset devices. Applied Artificial Intelligence, 37-41. doi:10.1080/08839510390225050

[24] Velásquez, J. D. (1997). Modeling Emotion-Based Decision-Making.

[25] Yu, Z., Zhou, X., Hao, Y., & Gu, J. (2006). TV Program Recommendation for Multiple Viewers Based on user Profile Merging. User Modeling and User-Adapted Interaction, 16(1), 63-82. doi:10.1007/s11257-006-9005-6

Graph Embeddings for Movie Visualization and Recommendation

Michail Vlachos
IBM Research - Zurich, Switzerland

Daniel Svonava
Slovak University of Technology, Slovakia

ABSTRACT

In this work we showcase how graph-embeddings can be used as a movie visualization and recommendation interface. The proposed low-dimensional embedding carefully preserves both local and global graph connectivity structure. The approach additionally offers: a) recommendations based on a pivot movie, b) interactive deep graph exploration of the movie connectivity graph, c) automatic movie trailer retrieval.

1. INTRODUCTION

As we are moving gradually from the era of information to the era of recommendation, interactive interfaces that engage the users and let them easily discover the data of interest will be of increasing importance. In this work we explore how graph-embedding techniques can be used as the basis of an interactive recommendation engine for movies.

Several recommendation systems have appeared in the literature in the recent years for recommending videos [5, 9] or movies [2, 6, 8, 10]. These, however, rarely focus on visually driven interfaces. In our scenario, we use a movie-actor database as the underlying graph structure. We use textual features to describe the movie objects. Given a selected pivot movie the system can retrieve a set of similar movies which are portrayed and clustered on two dimensions. Our method presents a novel way of capturing both neighborhood and cluster structure. Neighborhood information is preserved by retaining the Minimum Spanning Tree (MST) structure on two-dimensions. This also partially preserves the global graph structure, as the MST represents the dataset 'backbone'. In addition to the neighborhood structure, the method also retains the cluster structure which can be visualized at different granularities. Finally, the proposed mapping can accommodate both metric and non-metric distance functions.

Data-embedding techniques have been extensively used for visualizing high-dimensional data. Examples include the Bourgain embedding [3], FastMap [7], BoostMap [1] and ISOMAP [12]. The goal of those projection methods is to preserve all distances approximately, while our approach preserves a subset of distances (spanning-tree distances) *exactly*.

We use the proposed graph embedding as the entry point for a movie recommendation interface. Our methodology allows the exploratory visualization of the movie graph space and incorporates additional capabilities, such a filtering of the graph based on various criteria, real-time graph exploration and automatic retrieval of the movie trailer from the Internet.

Paper presented at the Workshop on Interfaces for Recommender Systems 2012, in conjunction with the 6th ACM conference on Recommender Systems. Copyright @ 2012 for the individual papers is held by the papers' authors. This volume is published and copyrighted by its editors. Interface@RecSys'12, September 13, 2012, Dublin, Ireland
.

2. DESCRIPTION OF OUR APPROACH

The proposed method combines the visual simplicity and comprehensibility of the minimum spanning tree (MST) with the grouping properties of hierarchical clustering approaches (dendrograms). Given pairwise distances describing the relationship of high-dimensional objects, the objective is to preserve a subset of important distances as well as possible on 2D, while at the same time accurately conveying the cluster information.

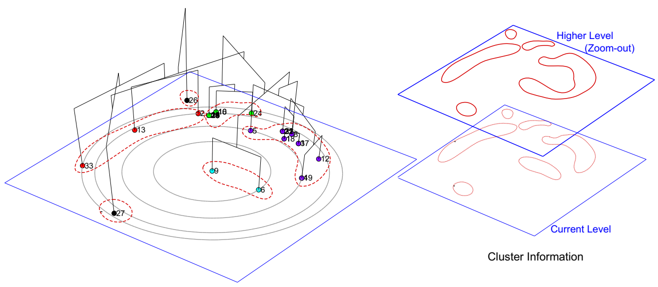

Figure 1. Conceptual illustration of our approach. The spanning tree structure is preserved and projected onto a 2D plane, while cluster structure is overlayed in the third dimension. 'Cutting' the dendrogram at a certain level projects the cluster structure onto 2D.

Our approach works as follows. First, we construct a Minimum Spanning Tree (MST) layout on the 2D plane in such a way that all distances to a user-selected pivot point and the neighborhood distances on the MST are exactly preserved. This construction carefully considers how to best portray the original object relationships for either metric or non-metric distance functions. Secondly, the dendrogram cluster hierarchy is constructed so that it can be positioned exactly *on top* of the MST mapping of objects. The cluster hierarchy can be frozen at any resolution level (tomographic

	Visualization Method	Trailer	Filters	Special Features
netflix.com	Tables	Yes	Limited	Very large user-base
jinni.com	Linear Treemap	No	Yes	Semantic search
IMDB.com	Icon Tiles	Yes	No	Comprehensive database
MovieLens.org	Tables	No	Yes	Recommendations through initial test
Our method	Minimum Spanning Dendrograms	Yes	Rating/Genre/Year	Clustering/Deep Graph Exploration

Table 1. Overview of features for several extant movie recommender systems

view) to convey the multi-granular clusters that are formed. This concept is elucidated in Figure 1: objects are properly mapped on the 2D plane whereas on the third dimension the hierarchy of the clustering structure is portrayed. Naturally, this is only a conceptual illustration of our approach. In practice, cluster information is also projected onto two dimensions, e.g., by properly coloring the nodes belonging to the same cluster. Therefore, by 'cutting' the dendrogram derived on a user-defined level, clusters on 2D can be formed, expanded and contracted appropriately, as the user drills up or down on the cluster hierarchy.

The work presented here constitutes an demo prototype of hhe visualization technique presented in [11]. Here we explore in more detail how the proposed high-dimensional data embedding methodology can be used as the interface of a movie search engine. In Table 1 we present briefly the differences of our approach with respect to prevalent movie search and recommendation systems.

2.1 Neighborhood Preservation

We will first explain how to capture on two dimensions the relationship between a set of high-dimensional objects. As not all pairwise distances can be retained on two dimensions we choose to maintain, as well as possible, the spanning tree distances which partially capture the local relationships and also record information about the general global structure [11].

We begin by constructing the spanning tree on the original high-dimensional objects. One object is selected as pivot and mapped in the center of the 2D plane coordinate system. By traversing the spanning tree, objects are positioned on the 2D plane by triangulating the distances to two objects: the pivot object and the neighboring point previously mapped on the spanning tree.

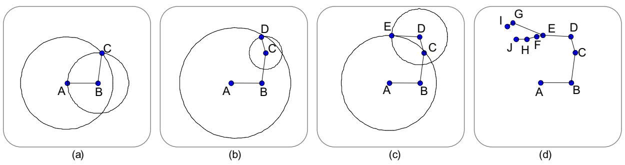

Figure 2. Two dimensional mapping of MST of objects

Example: Suppose the first two points (A and B) of the MST have already been mapped, as shown in Fig. 2(a). Let's assume that the second distance preserved per object is the distance to a reference point, which in our case is the first point. The third point is mapped at the intersection of circles centered at the reference points. The circles are centered at A and B with radii of $d(A, C)$ – the distance between points A and C– and $d(B, C)$, respectively. Owing to the triangle inequality, the circles either intersect in two positions or are tangent. Any position on the intersection of the circles will retain the original distances towards the two reference points. The position of point C is shown in Fig. 2(a). The fourth point is mapped at the intersection of the circles centered at A and C (Fig. 2(b)) and the fifth point is mapped similarly (Fig. 2(c)). The process continues until all the points of the ST are positioned on the 2D plane and the final result is shown in Fig. 2 (d).

The mapping technique presented will retain *exactly* the distances between all points and the pivot sequence, and also between the nodes that lie at the edges of the spanning tree. This creates a powerful visualization technique that not only allows to preserve nearest neighbor distances (local structure), but in addition retains distances with respect to a single reference point, providing the option of a global data view using that object as a pivot.

Layout optimization using simulated annealing: For reaching maximum visual clarity we try to minimize the number of intersected graph edges. Recall that when triangulating the position of a third point to its neighbor and the pivot, the algorithm proceeds by identifying the intersection between two circles. One can readily see this in Figure 3; there are two positions in which a newly mapped point can be placed.

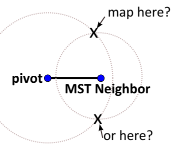

Figure 3. Selecting which of the two positions a new point is mapped to

We employ a probabilistic global optimization technique based on *simulated annealing (SA)* [4] that intelligently selects which of the two mapping positions to use, so as to minimize the number of crossed edges. SA is an effective optimization method when the search space is discrete, which is exactly the situation we face. Our experiments on the movie graph database suggest that the simulated annealing process is very effective in reaching an improved layout.

Using Non-metric Distances: When the underlying distance measure obeys the triangle inequality the circles around the reference points are guaranteed to intersect. However, many widely used distance functions (e.g., dynamic warping, longest common subsequence) violate the triangle inequality, and thus the corresponding reference circles may not necessarily intersect. In such cases, one needs to identify the position where to place an object with respect to the two circles, in such a way that the object is mapped as close as possible to the circumference of both circles. We need to identify the locus of points that minimize the sum of distances to the perimeters of two circles. One can show that the desired locus always lies on the line connecting the centers of the two circles. An example is shown in Fig. 4.

2.2 Cluster Preservation

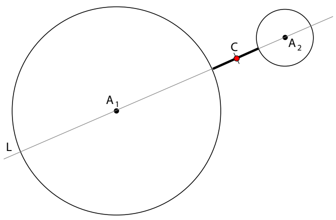

Figure 4. Discovering the mapping point for non-metric distances if the circles do not intersect.

Now we turn our attention to capturing and conveying the cluster information on 2D. Recall that the input for the algorithm is a matrix of pairwise distances. Based on the pairwise distances given, one can build a hierarchical dendrogram. The dendrogram construction is based on a single linkage approach that merges *closest* singleton objects and clusters.

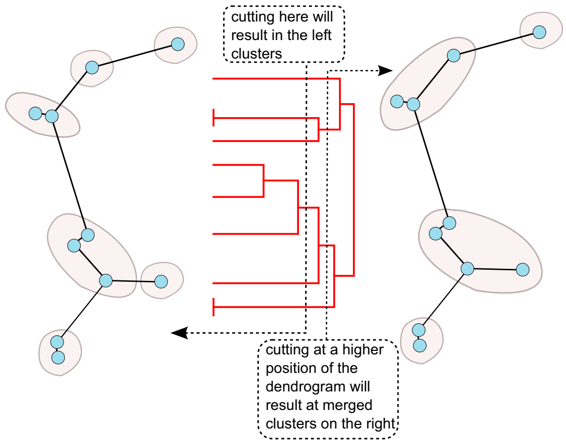

Figure 5. Cluster information conveyed by variable thresholds on a dendrogram information, thus imposing 'zoom-in' and 'zoom-out' in the cluster structure.

Given the minimum spanning tree, the clustering process can be sped up, because the merging order of the MST algorithm is the same as the merging order of the single linkage hierarchical clustering approach. In addition, this reflects the fact that one can achieve an *exact registration* of the constructed hierarchy on top of the MST-mapped points, because the clustering order is the same as the order crystallized in the spanning-tree mapping. This can be verified in Fig. 1 where the single linkage dendrogram is positioned exactly on top of the spanning-tree mapping.

The above observation combined with the hierarchical cluster information provides the capacity for multi-granular cluster views on 2D by interactively setting variable threshold levels on the resulting dendrogram. In this way, a 'tomographic view' of the clusters with formation of variable size clusters is possible. The concept is illustrated in Fig. 5: by cutting the dendrogram at a lower threshold, six clusters are created on the left. Imposing a higher threshold, clusters are merged progressively as shown on the right side. Our prototype implementation conveys cluster information by coloring the node perimeters and the connected edges.

3. GRAPHICAL INTERFACE

On top of the proposed visualization methodology we have built a movie recommendation engine that allows the interactive exploration of a large movie graph. Our sample database consists of 125,000 movies and 955,000 actors. We augment the information on each movie by attaching additional unstructured information from the web. So, each movie is described by a *bag-of-words* pertinent to the genre, actors, director(s), language, and a set of keywords relevant to the plot. To evaluate movie similarities we follow an IR-driven approach by considering the cosine similarity between the bag-of-words. More complex functions can be also accommodated, however this approach already gave very satisfactory and intuitive results. So, similarity between movies is based solely on the content of the movie rather than on any collaborative features (or ratings). We follow this path in an effort to introduce a factor of *serendipity* in the recommendation process. Like this, more obscure movies can appear in the recommendation if they share similar content (e.g. plot or mood) with the one that the user selected. In general, we have noticed that movie recommender systems that consider collaborative features (e.g. users that have rated highly movie A have also rated highly movie B) tend to have a strong bias toward blockbuster movies that most users (typically) have already watched.

GUI and Functionalities: The data visualization is accommodated through a graphical web interface, shown in Figure 6. The interface allows the user to search for a movie, and subsequently displays the proposed visualization graph, placing the movie selected as the center (pivot) object. The user can then easily identify other relevant movies, with the option to retrieve detailed information about the movie, such as participating actors or the movie plot. The user can modify the number of displayed movie clusters or even watch the movie trailer. The application allows the exploration of both sides of the movie-actor bipartite graph: either by deeper exploring of the movie graph (by clicking on a movie) or by searching/filtering for movies of a particular actor (by clicking on an actor's image). Additional filtering functionalities include: a) Filtering by rating, e.g., when interested in retrieving movies with a rating $> Y$. b) Filtering by year, when the user is interested only in recent movie releases.

Below we provide specific examples of our mapping, highlighting its ability to be used as an effective and interactive movie recommendation system. Number of clusters can be interactively modified and clusters are conveyed using varying border and edge colors.

Examples: Selecting 'The Titanic' as pivot movie produces the 2D mapping shown on the left-hand side of Fig. 7. The user can navigate through the graph and identify similar movies. Clustered with Titanic are the movies: 'Titanic (1953)', 'Poseidon' and 'Shakespeare in Love'. Another cluster displayed, shown in light green, includes movies like: 'The Notebook', 'Atonement', 'Purple Rain'; all romantic drama movies. Similarly, selecting 'Star Wars' as pivot movie correctly packs closely the remaining Star Wars movies (Fig. 7, right-hand side). An adjacent cluster includes parodies of Star Wars, like 'Spaceballs' or 'Thumb Wars'. Other related movies include adventure films like the 'Lord of the Rings' trilogy, or sci-fi action thrillers like 'Aliens' and 'Star Trek'. Additional illustrative examples can be found in the provided video.

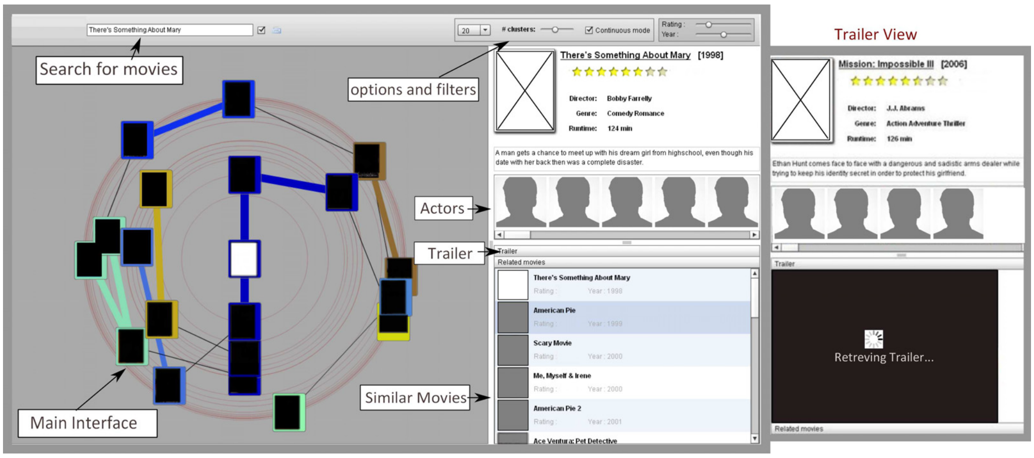

Figure 6. The interface of the Movie Recommendation web application

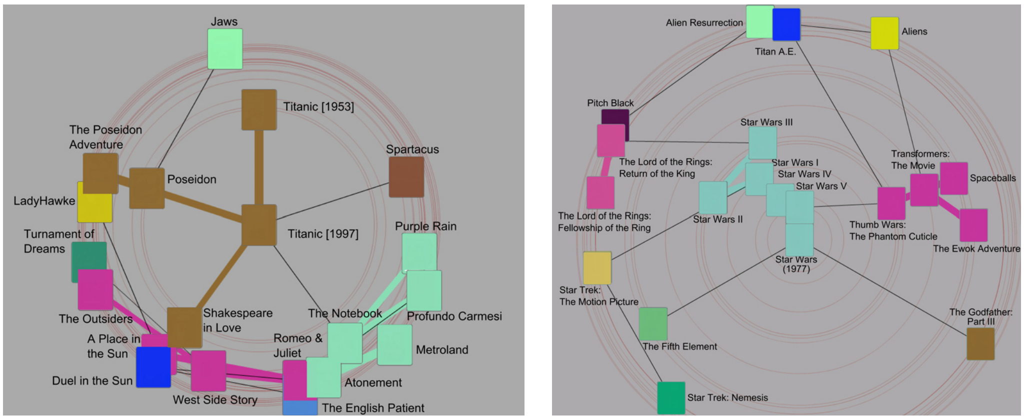

Figure 7. Using the proposed mapping on the movie graph. *Left:* Pivot movie is 'Titanic'. *Right:* 'Star Wars' selected as pivot

Benefits of simulated annealing (SA): The proposed SA component probes multiple node placement configurations, and picks the one that minimizes the number of intersected edges. To measure its effectiveness we select 1000 movies at random and retrieve the k nearest neighbors. Table 2 reports the median number of edge intersections with and without the SA component. We observe that the SA implementation significantly reduces the edge intersections and hence the screen clutter, providing a more intelligent node placement. Note, that both variations provide the same distance preservation with respect to the pivot and the MST distances.

Table 2. Node placement using simulated annealing (SA) significantly reduces the number of edge intersections

	Graph edge intersections	
	Without SA	With SA
k=20	36	12
k=30	108	48
k=40	222	112
k=50	346	212

References

[1] V. Athitsos, J. Alon, S. Sclaroff, and G. Kollios. BoostMap: An Embedding Method for Efficient Nearest Neighbor Retrieval. In *IEEE Trans. Pattern Anal. Mach. Intell. 30(1)*, pages 89–104, 2008.

[2] J. Bennett and S. Lanning. The Netflix Prize. In *Proc. of KDD Cup and Workshop*, 2007.

[3] J. Bourgain. On Lipschitz embeddings of finite metric spaces in Hilbert space. In *Israel Journal of Mathematics, 52*, pages 46–52, 1985.

[4] V. Cerny. A thermodynamical approach to the travelling salesman problem: an efficient simulation algorithm. In *J. of Opt. Theory and Applications 45*, pages 41–51, 1985.

[5] J. Davidson et al. The YouTube Video Recommendation System. In *Proc. of ACM Recommender Systems*, pages 293–296, 2010.

[6] E. A. Eyjolfsdottir, G. Tilak, and N. Li. MovieGEN: A Movie Recommendation System. In *UC Santa Barbara: Technical Report*, 2010.

[7] C. Faloutsos and K. I. Lin. FastMap: A fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets. In *Proc. of SIGMOD*, 1995.

[8] F. Gruvstad, N. Gupta, and S. Agrawal. Shiniphy - Visual Data Mining of movie recommendations. In *Stanford University: Technical Report*, 2009.

[9] T. Mei, B. Yang, X.-S. Hua, L. Yang, S.-Q. Yang, and S. Li. VideoReach: an online video recommendation system. In *Proc. of SIGIR*, pages 767–768, 2007.

[10] B. N. Miller, I. Albert, S. K. Lam, J. A. Konstan, and J. Riedl. MovieLens Unplugged: Experiences with an Occasionally Connected Recommender System. In *Proc. of IUI*, pages 263–266, 2003.

[11] D. Svonava and M. Vlachos. Graph Visualization using Minimum Spanning Dendrograms. In *IEEE International Conference on Data Mining*, 2010.

[12] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science 290*, pages 2319–2323, 2000.