Technical Report

Department of Computer Science
and Engineering
University of Minnesota
4-192 Keller Hall
200 Union Street SE
Minneapolis, MN 55455-0159 USA

TR 13-034

Models for Dynamic User Preferences and Their Applications

Komal Kapoor

November 18, 2013

Models for Dynamic User Preferences and Their Applications

Komal Kapoor
Department of Computer Science
University of Minnesota
kapoo031@umn.edu

Supervised By

Jaideep Srivastava
Department of Computer Science
University of Minnesota
srivasta@umn.edu

Paul Schrater
Department of Computer Science and Psychology
University of Minnesota
schrater@umn.edu

ABSTRACT

Computational models of preferences have been applied in various domains including economics, consumer research and marketing. They are also commonly used for designing recommender agents for suggesting new content to the users based on their inferred preferences. A major challenge for such systems is to cater to the changing needs of the users over time. Although, user preferences are known to be dynamic in nature, there are few methods for predicting these dynamics in a reliable way. In this thesis, the problem of defining predictive models of dynamic user preferences is addressed. A solution to this problem is provided by formulating a framework that incorporates history and time dependent changes in user preferences for items. Two types of changes in user preferences are identified. Firstly, user's interests are modeled as either favoring familiarity or looking for exploring new content. Secondly, user's preferences for familiar items are defined to change as a function of exposure for incorporating the psychological effects of boredom from repetition. Such a framework for estimating dynamic preferences of users provides unprecedented insights to user changing needs. These insights are proposed to be incorporated in solving two important problems for content services; user retention and temporally-aware recommendations.

Categories and Subject Descriptors

H.1.2 [**Information Systems Applications**]: Models and Principles—*User/Machine Systems*; I.6 [**Simulation and Modeling**]: Applications

General Terms

Human Factors, Algorithms, Experimentation

Keywords

Dynamic Preferences, Recommender Systems, Retention, User Behavior Modeling

1. INTRODUCTION

Individual preferences have been a subject of much analysis by psychologists and sociologists. Further, preference models have been extensively used in various applications such as in modeling individual decision making in economics [7] and in predicting consumer purchase behavior [24]. This topic has recently come under the purview of the computer science community due to its applications in circumventing the information overload problem of the internet era. With the users finding themselves with more options than they can handle, the constant need for them to make choices such as which posts, blogs or articles to read, which videos or movies to watch, what music to listen, what games to play etc. can be quite overwhelming. This has led to major expansions in frontiers of user modeling primarily for the design of automated assistive agents. As a result, many businesses now incorporate recommendations as an integral part of the services offered by them.

The recommendation community has been instrumental in advancing research in representations, models and methods for extracting and applying knowledge of user preferences. While, methods have been perfected to exploit similarity structures between users and their preferences for items for recommending them new content, these models have accrued criticism for concentrating extensively or entirely on past behavior, resulting in recommendations which tend to be 'too similar' and are often disliked by the users. Furthermore, such methods have largely assumed preferences to be either static or to gradually drift with time without proposing any predictive mechanisms for such dynamics. Such an approach ignores the sequential and temporal structures in users preferences, and their future evolution. For example, let's consider the choices of a user for viewing a movie. It is easy to see that the movie that the user views today depends not only on the types of movies she generally likes, but also, the movie she saw recently allowing psychological factors such as boredom and the need for variety to emerge. Hence, there is strong temporal dependence in user choices with past choices not only informing preference inference but simultaneously modifying future preferences and choice dynamics.

Another application of user modeling has been towards directing efforts for user retention. There is tremendous competition among the rapidly increasing number of web services for survival making it vital for them to invest in growth and retention solutions. This directly results in a great deal of emphasis being placed by services on retaining and further engaging their current user base. Furthermore, with most of web services essentially acting as content delivery engines, such as StumbleUpon, Last.fm, Pandora, Spotify, YouTube etc., their ability to provide interesting content can be one of the vital factors for keeping their users engaged. As a result, the changes in user preference and the ability (inability) of the service in catering to the same, can have direct impacts on user retention metrics, which are yet to be explored in the retention community..

In this thesis, a framework for dealing with dynamics in user preferences is conceptualized. Such dynamics are defined by formalizing different states of the users and their preferences for items. Furthermore, the transitions between these state are defined as a function of time and past experiences of the user. The users are assumed to alternate between two preference states, a preference for **familiarity** and the preference for **exploration**. The familiar items are further segregated based on their level of satiation into four states namely: **Sensitization**, **Devaluation**, **Recurrence** and **Dropped**. Such dynamic preference states of the users provide insights about their changing needs, which were otherwise not available. These insights are proposed to be used to advance solutions to two major applications of user modeling described above: retention and recommendations. In order to address the retention problem, firstly, a model has been proposed for predicting the time taken by a user to return to a web service. Subsequently, the model would be refined to allow modeling the effect of particular user preference states on their return behavior. The knowledge of dynamic needs of the user is further proposed to be incorporated in the design of a new methodology for making better recommendations. An approach for clustering items based on similarities in user-item preferences and also their dynamics for recommendations, is discussed. Hence, the research agenda impacts the state-of-the-art in user modeling with application to both the areas of retention and recommendations.

The rest of the article is organized as follows. The next section reviews some relevant background for this work. Then the **Proposed Approach** section describes the framework for modeling preference dynamics in users proposed in the work. In the **Current Progress and Future Work** section the progress made in accomplishing the research agenda is summarized and directions for future work are provided. Finally, the contributions of this work are summarized in the **Conclusion** section.

2. BACKGROUND

This section reviews some theoretical and computational underpinnings of the formulation of dynamic preference states of the users.

2.1 Dynamics in Preferences

Predicting changes in preferences, based on the past behavior and experiences of the users is a non-trivial problem with no proposed solutions. Although, little studied in the user modeling community, evolution in human preferences has been a subject of much psychological research. Some key insights and findings from the state-of-the-art in behavioral psychology are discussed in this section.

Studies in psychology of preferences have been devoted towards understanding the role of familiarity and novelty in driving an individual's future interests. A correlation between familiarity and preference has been identified as the mere-exposure effect, first formalized by Zajonc [27], according to which repeated exposure to a stimuli is sufficient for an enhancement in liking. Further, Martindale has shown that prototypicality and mere-exposure are important factors for predicting aesthetic preferences [21]. Alternatively, human behavior has been described to be exploratory or sensation seeking. The inherent drive for exploration leads individuals towards desiring new and novel content. Such a preference is hypothesized to result from curiosity for new information [2]. Exploratory behavior is also linked to stimulus satiation responses arising on repeated exposure [12].

A crucial question that appears is; what determines future behavior of the organism - the desire for familiarity or a desire for exploration? Several studies have found exploration and exploitation tendencies to occur in moderation, with an excess in one leading to the other. For example, detailed experiments on the mere-exposure effect have shown that the favorable effects of exposure on preference are limited by boredom [4]. Alternatively, high uncertainty in the environment is found to result in anxiety and displeasure. Hebb [13] and Leuba [20] independently postulated this idea by suggesting that organisms are driven towards maintaining an optimal level of stimulation in their environment.

2.2 Computational Models of Dynamic User Preferences

The psychological research in preferences discussed so far has primarily been theoretical in nature. There have been efforts in the areas of consumer research and recommendations towards developing computational models for dynamic preferences in real world scenarios. Some of these models are reviewed next.

2.2.1 Consumer Choice Models for Variety Seeking Behavior

Variety seeking behavior has received a lot of attention from the consumer research community. McAlister compiled a taxonomy of factors responsible for varied behavior in consumers [22]. McAlister further modeled variety seeking behavior in soft drink preferences using a dynamic attribute satiation model [23]. The model assumed an ideal level of inventory at the attribute level and penalized departures from the ideal level. The inventory was designed to dwindle over time to incorporate the effects of forgetting.

Subsequent efforts in consumer research have advanced towards developing a general consumer choice model which allows consumers to either exhibit a short term loyalty for their last purchased brand (inertia) or devaluation for the last purchased brand (variety seeking) [11, 14]. Bawa et al [1] used a single peaked function, to model the conditional probability of repeat purchase given the number of times the brand was re-purchased since user's last switch (run length). Recent efforts have expanded these models to incorporate heterogeneities between consumers and external environment variables affecting user choices [15].

However, most of the existing efforts in consumer research have not accounted for long term changes in consumer interests. A recent work by Garcia-Torres [8] uses a utility based model of consumer choice making which incorporates the process of preference formation by allowing for integration between old habits and the acquisition of new preferences.

2.2.2 *Models for Dynamic Recommenders*

Most of the early models for recommendation assumed a static view of preferences. Such models include the popular nearest neighbors [5] and matrix factorization methods [19] for recommendation and their subsequent probabilistic renditions. However, the lack of temporal awareness in these models was obvious. Ding et al. [6] showed that it was important to include temporal changes in preferences while making recommendations. He proposed using a decay function to emphasize recency of past behavior while predicting future recommendation needs. Koren offered a better solution by modifying the matrix factorization model to incorporate changes in preferences over time [18]. However, Koren's and other similar approaches can be described as a corrective scheme for preferences changes rather than as dynamic model for preferences.

A dynamic model for preferences was proposed recently by Sahoo et.al [25] for predicting blog reading behavior in employees. They used a hidden markov model for modeling dynamic participation of users in different latent classes based on their changing preferences over time. Sahoo et. al's approach however, assumes that global preferences of a user to remain stable over time.

3. PROPOSED APPROACH

An overall framework for specifying the history and time dependent dynamics in user preferences is now described. The framework is based on formulating the notion of preference states for users and the items in their choice set. As a user exposes herself to items over time, based on her choices she naturally divides the space of items available to her (X) into two disjoint sets:

1. **Familiar Items** (X^f): Such constitute items which the user has explored in the past.

2. **New Items** (X^n): Such constitute all the items in the choice set other than the familiar items; $X^n = X - X^f$.

The preference states for users are also defined. At any point in time, a user can either prefer items from the set of familiar items, in which case she is said to be in the **familiarity** state, or the user can chose new items not explored before, in which case she is defined to be in **exploratory** state. Such a state representation embodies the inherent drives for familiarity and exploration from behavioral psychology discussed earlier.

However, the above state definition is insufficient for defining factors which produce transitions between user preferences states for familiarity and exploration. Specifically, the psychological effects of boredom for producing exploratory

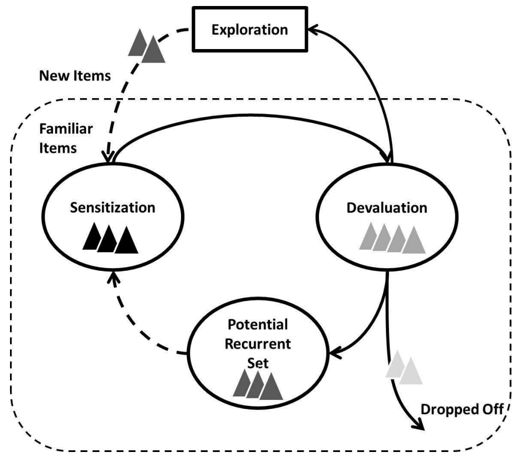

Figure 1: The user dynamic state transition model.

tendencies in individuals are considered in this work. Hence, items within the set of familiar items are further classified based on their satiation levels as belonging to one of the following preference states:

1. **Sensitization** (X^{f_s}): This state constitutes of the items which are preferable because of the user's recent exposure to them.

2. **Devaluation** (X^{f_d}): This state is defined by a decrease in interest for items which the user has already been exposed to enough times.

3. **Recurrence** (X^{f_r}): This state comprises preferable items which were temporarily devalued due to boredom. The preferences for these items have reinstated after sufficient reduction in exposure.

4. **Dropped** (X^{f_l}): This state comprises items which are devalued and beyond reinstation.

A user in the familiarity state chooses items which are sensitized or are likely for recurrence. Furthermore, the user tends to avoid items in the devalued state. Depletion in the sensitization and recurrence states and expansion in the devaluation state are identified as the factors causing the user to transition to the exploratory state. Finally, the intent of the exploratory states is to produce new additions to the familiar and sensitized state. Figure 1 shows the state transitioning structure between the various preference states for users and items outlined above.

Given, this state representation, two types of state dynamics need to be defined, namely: (a) The dynamic membership of familiar items in the Sensitization, Devaluation, Recurrence and Dropped off states with exposure and (b) A dynamic user familiarity-exploration state transition model given the preference states for the familiar items of the user.

4. CURRENT PROGRESS AND FUTURE DIRECTIONS

This section reviews the progress made in defining the dynamics for the user-item preference states outlined earlier.

Furthermore, the future directions for incorporating the dynamic model of user preferences in retention and recommendations solutions are discussed. The described work is based on modeling the consumption choices for music listeners. The music listening behavior is characterized by high level of repetition of a fairly small set of dearly loved songs with occasional exploration. It therefore allows a rich test-bed for modeling psychological factors of familiarity, boredom and exploration.

4.1 Dynamics in Preference States for Familiar Items with Exposure

In a recent work [16], a hazard based approach is defined for specifying dynamic preference states for familiar content as a function of past exposure. Hazard functions have been used in survival analysis for defining the instantaneous rate of occurrence of events. In order to use such functions for modeling user's preferences dynamics with exposure, two types of event rates are specified (a) Rate of Exit: The rate of choosing an item again given the consecutive number of times (run length) user has chosen the item in the immediate past. (b) Rate of Entry: The rate of choosing an item again given the time gap since the last time the item was chosen.

In the absence of any exposure specific dynamics, the exit and entry event rates would be constant for different lengths of exposure or absence of exposure, respectively. However, when these hazard functions computed using the music listening histories of users from the public dataset from the music service Last.fm, were compared against the average rate of return for the items, two distinctive phases corresponding to *stickiness* and *boredom* emerged. The stickiness phase was marked by a larger rate of choosing items which were chosen recently than average. Such periods of enhanced preference for recently exposed items are used to define the **Sensitization** preference state for the item. The items were also found to exist in the boredom phase wherein they were chosen with a lower rate than average after they have already been exposed to enough times in the past. The period that an item existed in the boredom phase was used to define the **Devaluation** preference state. The preference for items was found to reinstate with enough time spent away from the item. The items are then said to belong to the **Recurrence** preference state. Finally, items which are not chosen for extremely long periods of time are classified as belonging to the **Dropped** preference state.

4.2 A Familiarity-Exploratory State Transition Model

A state transition model for preferences of a user for familiarity and exploration is now proposed. The model is developed by analyzing the music listening choices of the users from the Last.fm dataset. Two types of user actions are defined: the *search* action when a user listens to song after actively searching for it and the *radio* action when a user chooses to listen to a song appearing in a radio stream. The song listened by the user can be further classified as familiar or new based on whether they belong to the familiar items or new items. Finally, the following four states are defined (a) Search Familiar: the familiarity state corresponding to a user listening to a familiar song after searching for it, (b) Radio Familiar: the familiarity state corresponding to a user listening to radio comprising of familiar songs, (c) Search New: the exploratory state corresponding to a user listening to a new song after searching for it and (d) Radio New: the exploratory state corresponding to a user listening to radio comprising of new songs. A predictive model of user state dynamics between the above four states is defined using a semi-markov model. A semi-markov model is particularly applicable for this problem as it allows defining state dependent dynamics for particular states using state-specific hazard functions.

4.3 Applications: User Retention

An approach for predicting the return time of a user was formulated for addressing the retention problem for web services, in a recent work [17]. Specifically, a Cox's proportional hazard model was used for modeling the return time data for the users. This model further allowed incorporating several user-specific covariates that affect the rate of user return. The types of covariates included in the model included those related to the typical visitation patterns of the user, their satisfaction/engagement with the service and for abstracting the effects of external factors. A few covariates were defined to capture each of the above aspects in the model. The proposed model was tested using datasets from two music services and showed better performance than other state-of-the-art data mining methods. Furthermore, the model could further improve its prediction performance using the length of absence already observed for the user. This was due to the ability of the hazard based approach to incorporate the decline in the user return rate with time spent away from the service.

Despite being a good model of user return time prediction, the model lacked an assessment of the userâĂŹss engagement and satisfaction with the service based on knowledge of the current preference states of the user. We propose to further incorporate in the model covariates related to the psychological preference states for items and users described earlier. This would allow us to analyze for the first time, how user preference dynamics impact their engagement and satisfaction with the service and in turn affect their return behavior to the domain.

4.4 Applications: Recommendation Methods

Past efforts in recommendations have largely explored similarity structures in user choices to predict their preferences for items. Some of the popular probabilistic models which learn latent similarity structures in user preference patterns to recommend them new content include variable mixture models such as probabilistic latent semantic analysis (pLSA), Latent Dirchlet Allocation etc.

These generative models are based on the principle of exchangeability, i.e. they produce the same features or clusters for different permutations of the data. However, as discussed earlier, user preference data has strong temporal dynamics resulting in dependencies between data observations from the same user. Recent advances in infinite mixture and factor models have incorporated such non-exchangeability in data observations using distance-dependent Chinese restaurant processes [3] and distance-dependent infinite latent feature models [9]. Such methods are relatively new and have been used to model temporal and spatial proximity in a few applications such as image segmentation [10] and text clustering [26]. We propose to adapt these distance-dependent formulations to simultaneously clusters users, items and their temporal dynamics resulting from exposure.

5. CONCLUSION

Figure 2 displays the proposed plan for the thesis. The progress made in completing each of the components is also specified.

Problem Description	Sub-problems	Status
A framework for modeling dynamic user-item preference states	Psychological Preference States for familiar items	Completed
	Familiarity –Exploration User State Transition Model	In Progress
Applications: User State Aware Retention Model	A Cox's Proportional Hazard Model for User Return Time Prediction	Completed
	Using Boredom and Availability Bias for Predicting Domain Devaluation	In Progress
Applications: Dynamic Recommenders	A Temporally Aware Clustering of User Item Choice Data for Recommendations	Future Work

Figure 2: A breakdown of the problems and sub-problems proposed in this thesis and their progress status.

The contributions made through this work can now be summarized. In this work, a framework for modeling changing preference of users is proposed. The framework allows modeling satiation effects in user preferences for familiar content arising due to boredom. The model also incorporates changing preferences of users for familiarity and exploration. Such dynamics are related to devaluation in preferences for the familiar content. Finally, a scheme is proposed for applying the dynamic framework for user history and time dependent preferences for providing solutions for the problems of retention and recommendations. Such an approach seeks to utilize advancements in behavioral psychology for defining predictive models of dynamic user preferences. The complexity of the web environment and presence of noise and missing information in data collected from online sources pose several bottlenecks to a potential marriage between the psychological theories and computations preference models for the real world. However, with users spending more and more time on the web, modeling properties of their behavior has become critical for advancing designs of automated agents which interact with them on a daily basis. Also, the online environment provides a humongous test bed for research on human behavior, not available before. This works seeks to provide solutions for utilizing such mutually beneficial synergies between the two domains.

6. REFERENCES

[1] K. Bawa. Modeling inertia and variety seeking tendencies in brand choice behavior. *Marketing Science*, 9(3):263–278, 1990.

[2] DeE BERLYNE. Novelty and curiosity as determinants of exploratory behaviour1. *British Journal of Psychology. General Section*, 41(1-2):68–80, 1950.

[3] David M Blei and Peter I Frazier. Distance dependent chinese restaurant processes. *The Journal of Machine Learning Research*, 12:2461–2488, 2011.

[4] Robert F Bornstein, Amy R Kale, and Karen R Cornell. Boredom as a limiting condition on the mere exposure effect. *Journal of personality and Social Psychology*, 58(5):791, 1990.

[5] Christian Desrosiers and George Karypis. A comprehensive survey of neighborhood-based recommendation methods. In *Recommender systems handbook*, pages 107–144. Springer, 2011.

[6] Y. Ding and X. Li. Time weight collaborative filtering. In *Proceedings of the 14th ACM international conference on Information and knowledge management*, pages 485–492. ACM, 2005.

[7] Ward Edwards. The theory of decision making. *Psychological bulletin*, 51(4):380, 1954.

[8] Abraham Garcia-Torres. Consumer behaviour: evolution of preferences and the search for novelty. *Cornell family papers*, 2009.

[9] Samuel J Gershman, Peter I Frazier, and David M Blei. Distance dependent infinite latent feature models. *arXiv preprint arXiv:1110.5454*, 2011.

[10] Soumya Ghosh, Erik Sudderth, Matthew Loper, and Michael Black. From deformations to parts: Motion-based segmentation of 3d objects. In *Advances in Neural Information Processing Systems 25*, pages 2006–2014, 2012.

[11] M. Givon. Variety seeking through brand switching. *Marketing Science*, 3(1):1–22, 1984.

[12] Murray Glanzer. Curiosity, exploratory drive, and stimulus satiation. *Psychological Bulletin*, 55(5):302, 1958.

[13] Donald Olding Hebb. Drives and the cns (conceptual nervous system). *Psychological review*, 62(4):243, 1955.

[14] B.E. Kahn, M.U. Kalwani, and D.G. Morrison. Measuring variety-seeking and reinforcement behaviors using panel data. *Journal of Marketing Research*, pages 89–100, 1986.

[15] Tomohito Kamai and Yuichiro Kanazawa. The latent class model of brand choice behaviors incorporating variety-seeking and state dependence. 2012.

[16] Komal Kapoor, Nisheeth Srivastava, Jaideep Srivastava, and Paul Schrater. Measuring spontaneous devaluations in user preferences. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '13, 2013.

[17] Komal Kapoor, Mingxuan Sun, Jaideep Srivastava, and Tao Ye. A hazard based approach to user return time prediction.

[18] Y. Koren. Collaborative filtering with temporal dynamics. *Communications of the ACM*, 53(4):89–97, 2010.

[19] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

[20] Clarence Leuba. Toward some integration of learning theories: The concept of optimal stimulation. *Psychological Reports*, 1(g):27–33, 1955.

[21] Colin Martindale and Kathleen Moore. Priming, prototypicality, and preference. *Journal of Experimental Psychology: Human Perception and Performance*, 14(4):661, 1988.

[22] L. McAlister and E. Pessemier. Variety seeking behavior: An interdisciplinary review. *Journal of Consumer research*, pages 311–322, 1982.

[23] Leigh McAlister. A dynamic attribute satiation model of variety-seeking behavior. *Journal of Consumer Research*, pages 141–150, 1982.

[24] Phillip Nelson. Information and consumer behavior. *The Journal of Political Economy*, 78(2):311–329, 1970.

[25] Nachiketa Sahoo, Param Vir Singh, and Tridas Mukhopadhyay. A dynamic model of employee blog reading behavior.

[26] Richard Socher, Andrew L Maas, and Christopher D Manning. Spectral chinese restaurant processes: Nonparametric clustering based on similarities. In *International Conference on Artificial Intelligence and Statistics*, pages 698–706, 2011.

[27] Robert B Zajonc. Attitudinal effects of mere exposure. *Journal of personality and social psychology*, 9(2p2):1, 1968.