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.

and/or exercise manually on a web site or smart phone. This 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