Friend Recommendation for Weight Loss App

Anming Li  
Department of Computing,  
The Hong Kong Polytechnic University,  
Hung Hom, Kowloon, Hong Kong SAR  
csali@comp.polyu.edu.hk

Hareton Leung  
Department of Computing,  
The Hong Kong Polytechnic University,  
Hung Hom, Kowloon, Hong Kong SAR  
cshleung@comp.polyu.edu.hk

Yvette Lui  
Department of Computing,  
The Hong Kong Polytechnic University,  
Hung Hom, Kowloon, Hong Kong SAR  
csylui@comp.polyu.edu.hk

Abstract  
Although social network had been identified as an effective way to enhance overweight and obesity intervention in literature, the specific measures for integrating social network and weight loss are very limited until now. In this study, we developed a measure for recommending friends for weight loss apps in the context of social networks. In addition to network and profile similarities that have been well documented, our measure provided methods to model weight gain related behaviors and used obtained scores to construct a "behavior network". For evaluation, we proposed two measurements and conducted an experiment on a real dataset complemented with computer generated social graph. Results validated that presented measure is able to recommend friend conducting healthier lifestyle when compared with other friend recommendation measures.

Keywords  
Social network; friend recommendation; obesity; healthcare

Introduction  
Weight loss apps typically require users to keep monitoring their food intakes and physical activities in order to calculate calorie consumption. Since weight management is a long-term process, individuals are
likely to give up in the early stage of weight management program. To address this issue, some studies [1,2] suggested involvement of partners in the process because they could provide supports and serve as emotional buffer. However, experiments involving family members, friends, neighbors and significant others reported mixed results [3-5]. It can be boiled down to the following reasons: (1) family members or friends who conduct health-risk behaviors, such as eating fast food frequently, may provide negative impact to the target person; (2) family members or friends may become more tolerant with obese persons, and thus cannot act as a rigorous supervisor [1]; and (3) family members or friends and the target person may be influenced by the same cultural norm which promotes unhealthy behaviors [2]. In this regard, a qualified weight loss companion should be: (1) close to the target person in some ways; (2) with positive influence to the weight loss process; and (3) able to resist the unhealthy cultural norm.

Numerous weight loss apps, equipped with online social networking service, manage to use social network to aid weight loss process. Taking MyFitnessPal\(^1\) as an example, members are encouraged to invite their friends to diet with them for achieving better results. Though assisting with filters such as user name, gender and age, it is labor-intensive for users to find a new friend. MyFitnessPal has about 40 million users by July 2013, and is growing continuously at a rate of 1.5 million new users per month.

In this study, we present a friend recommendation measure considering both homophily and weight gain related behaviors for weight loss apps (FRHW). Traditional friend recommenders only considered similarities between two persons but not potential influences between them. The purpose of this study is to recommend friends that are not only similar to a target person but also will benefit him/her in the process of weight loss. Two measurements are also proposed to compare different friend recommendation measures.

**Related Works**

The connection between weight management and social network had been well studied. Christakis and Fowler [1] used longitudinal statistical models to examine whether weight gain is contagious. It suggested that the chance of a person getting obese increased by 57% in case one’s friends became obese. Li et al. [2] further explored the role of social networks and social media in obesity. It identified three pathways linking obesity with social networks: social support, social integration and social capital. It advocated development of social networks specifically designed for addressing obesity. However, the obesity intervention involving spouse, friends, coworkers and neighbors reported mixed results [3,4]. Kiernan et al. [3] found that women with family support are more likely to lose weight (71.6%) while women who did not obtain support from friends are most likely to lose weight (80%). Gorin et al. [4] demonstrated that participants with one or more successful partner lost more weight compared with those with no successful partners. Shoham et al. [5] investigated social influence on adolescent’s body size, screen time and playing sports using an actor-based model. Results showed that both social influence and homophily are important to understand adolescent obesity.

\(^1\) http://www.myfitnesspal.com/
Online social media can be served as additional sources of health related information and social support. It provides accesses to people that shared same experience and 24/7 encouragement, which can be difficult for realistic friends. Maher et al. [6] reviewed papers regarding effectiveness of interventions based on social network. Those interventions aimed at changing health behaviors such as dietary intake and physical activity. 90% of the studies reported positive outcomes. Chang et al. [7] reviewed papers regarding the role of social media in online weight management. They pointed out that for those lacking social support, social media could help them establish connections and obtain supports. Hwang et al. [8] investigated peer support on an online weight loss community. As suggested in this study, 60% of the participants considered online members were more helpful than other contacts on the issue of weight loss. Convenience, anonymity and none-judgment are the three distinct features valued most by the respondents.

People tend to connect with people similar or close to them on a social graph. Based on this theory, Akcora et al. [9] proposed network and profile based measures to compute user similarity. Their network measures considered both the network similarity and strength of friendships between each other. Spertus and Sahami [10] compared 6 different measures of similarity (L1 norm, L2 norm, pointwise mutual information (PMI) with positive correlations, PMI with positive and negative correlations, Salton’s similarity measure based on inverse document frequency scaling, and log-odds) on Orkut.com data for community recommendation. Result showed that L2 norm outperformed other measures.

**Example 1.** Consider a social graph in Figure 1.

![Social Network Graph](image1)

Then we have: \( W(u) = 8, W(FG(u), E) = 7 \) and \( W(NS(u, x)) = \log(7) / \log(2 \times 8) = 0.702 \).

**Example 2.** Consider a user \( u \) with profile \( P(u) \{ \text{Gender: Female, Race: Asian} \} \) and a stranger \( x \) with profile \( P(x) \{ \text{Gender: Female, Race: Indian} \} \). Suppose there are 1000 records and distributions of gender and race are as shown in Figure 2. Weight for each item is 1.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Race</th>
<th>Ethnicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>500</td>
<td>Asian 99</td>
</tr>
<tr>
<td>Male</td>
<td>499</td>
<td>Indian 102</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Others 818</td>
</tr>
</tbody>
</table>

Then we have: \( P(u, x) = \frac{1}{2} \times \left( 1 + \log \left( \frac{1000}{500} \times \log \left( \frac{1000}{500} \right) \right) \right) = 0.74 \)

**Measures of Friend Recommendation**

The proposed measure consists of 4 components: network similarity, profile similarity, methods for scoring weight gain related behaviors, and behavior network. We combine all the components and obtain a final ranking. If a user \( u \) is connected to a target user \( x \) within a two-hop distance in the social graph, the user is called a stranger.

**Network Similarity**

For network similarity, we use the measure proposed in the literature [9]. For a user \( u \) and a stranger \( x \), the network similarity is computed as:

\[
NS(u, x) = \frac{\log |MFG(u, x)|}{\log |FG(u)|}
\]

where \( MFG(u, x) \) is the number of edges in mutual friendship graph (MFG) and \( FG(u) \) is the number of edges in friendship graph (FG). MFG is the social graph consisting of \( u \) and \( x \)'s mutual friends. FG is the social graph consisting of \( u \) and \( x \)'s friends.

**Profile Similarity**

For the profile items, we choose occurrence frequency (OF)[9]. Let \( i \) be a profile item, \( i_u \) and \( i_x \) be the values for item \( i \) in the \( u \) and \( x \) profile respectively. Profile similarity \( P(u, x) \) is computed as below:

\[
OF(i_u, i_x) = \begin{cases} 
  1, & i_u = i_x \\
  \frac{1}{1 + A B}, & i_u \neq i_x 
\end{cases}
\]

\[
P(u, x) = \sum_{k=0}^{n} w_k \times OF_k
\]

where \( A = \log (N/f(i_x)) \) and \( B = \log (N/f(i_u)) \); \( f(x) \) is the number of records with that item value and \( N \) is total number of records, \( w_k \) is the weight assigned to the \( k^{th} \)
profile item, and \( n \) is the number of profile items. Weight for the \( k \)th profile item, \( w_k \), is user-defined.

Weight Gain Related Behavior: Definition and Modeling

Weight gain related behaviors generally fall into two categories: dietary behavior and physical activity behavior. For the former, we like to examine the fruit and vegetable (referred as FV hereafter) consumption, meal eating pattern (i.e. skipping breakfast) and snacking. FV consumption is the better predictor for healthy eating. A regular meal pattern was significantly related to a low level of obesity [12]. In addition, regular meal eating patterns were associated with lower snacking level, though snacking was not associated with obesity directly. For the latter, we suggest considering aerobic activity, time spent watching TV and computer use. Exercising regularly is the key to keep fit, while more time spent on sedentary behaviors, such as watching TV and playing video games, is associated with greater risk of obesity [15].

Suppose a weight loss app user records food log and exercise log on daily basis. We define a set of weight gain related behaviors WGRB where \( B_k \) denotes the \( k \)th behavior. \( B_k(x,t) \) is the behavior score for user \( x \) over a given period of \( t \) days. In the following, we will define weight gain related behaviors one by one.

Classification of FV Intake:

We categorize FV Intake into 3 classes: low intake (\( \frac{N_{\text{FV}}(x,t)}{t} < 0.5 \)), moderate intake (\( 0.5 \leq \frac{N_{\text{FV}}(x,t)}{t} < 0.75 \)), and high intake (\( \frac{N_{\text{FV}}(x,t)}{t} \geq 0.75 \)).

**FV Intake:** Let \( B_{\text{FV}}(x,t) \) represent the behavior score of a user \( x \) for FV intake, \( N_{\text{FV}}(x,t) \) represent the number days that \( x \) has FV no less than 400g[11] during \( t \) days, then we have the function:

\[
B_{\text{FV}}(x,t) = f_{\text{FV}}\left(\frac{N_{\text{FV}}(x,t)}{t}\right) = \begin{cases} 
0, & 0 \leq \frac{N_{\text{FV}}(x,t)}{t} < 0.5 \\
0.5, & 0.5 \leq \frac{N_{\text{FV}}(x,t)}{t} < 0.75 \\
1, & 0.75 \leq \frac{N_{\text{FV}}(x,t)}{t} \leq 1
\end{cases}
\]  

(4)

**Skipping Breakfast:** Let \( B_{\text{BF}}(x,t) \) represent the behavior score for the frequency of skipping breakfast. Skipping breakfast frequently can increase the risk of gaining weight [12]. We categorize frequency of skipping breakfast using the method applied in the study [12]. Then we have:

\[
B_{\text{BF}}(x,t) = f_{\text{BF}}\left(\frac{N_{\text{BF}}(x,t)}{t}\right) = \begin{cases} 
0, & 0 \leq \frac{N_{\text{BF}}(x,t)}{t} < 0.75 \\
1, & 0.75 \leq \frac{N_{\text{BF}}(x,t)}{t} \leq 1
\end{cases}
\]

(5)

where \( N_{\text{BF}}(x,t) \) is the total number of days that user has eaten breakfast during \( t \) days.

**Snacking:** Let \( B_{\text{SN}}(x,t) \) represent behavior score of a user \( x \) for snacking. Snacking can be additional source for dietary fiber and FV. It all depends on the selections of snacks. Note that we only consider unhealthy snacks. Definition of unhealthy snacks and classification of frequency of snacking can be found in the literature [13]. Then we have:

\[
B_{\text{SN}}(x,t) = f_{\text{SN}}\left(\frac{N_{\text{SN}}(x,t)}{t}\right) = \begin{cases} 
0, & 0 \leq \frac{N_{\text{SN}}(x,t)}{t} < 1 \\
-0.25, & 1 \leq \frac{N_{\text{SN}}(x,t)}{t} < 2 \\
-0.5, & 2 \leq \frac{N_{\text{SN}}(x,t)}{t} < 3 \\
-1, & \frac{N_{\text{SN}}(x,t)}{t} \geq 3
\end{cases}
\]

(6)

where \( N_{\text{SN}}(x,t) \) is the number of snacking in \( t \) days.

**Aerobic Activity:** Let \( B_{\text{AA}}(x,t) \) represent behavior score of a user \( x \) for aerobic activity. An adult should at least spend 150 minutes on moderate-intensity aerobic activity per week (that is average 21.4 minutes per day) [14]. So we define \( B_{\text{AA}}(x,t) \) as below:

\[
B_{\text{AA}}(x,t) = f_{\text{AA}}\left(\frac{N_{\text{AA}}(x,t)}{t}\right) = \begin{cases} 
0, & 0 \leq \frac{N_{\text{AA}}(x,t)}{t} < 21.4 \\
1, & \frac{N_{\text{AA}}(x,t)}{t} \geq 21.4
\end{cases}
\]

(7)
Example 3. Consider weight gain related behaviors of a user, Tom, for a given period of time, say a week. He reported that:
- Days of eating breakfast: 4 (NBEB=4)
- Days of having 400g or more FV: 3 (NBFV=3)
- Hours spent on watching TV: 27 (NBTV=27)
- Total minutes of working out: 120 (NBAA=120)

Then the behavior score of Tom for that week is: WGRB(Tom, t) = 1/4

Example 4. Consider a social graph in Figure 3. Suppose weight for each friend is 1.

Figure 3. Social graph complemented with behavior score

Then we have the NB score for user u: NB(u) = 1/6 × ((-0.5)+ (-0.25) + 0.5+0.4+ (-0.1)) = 0.0083

where $N_{aa}(x,t)$ is total minutes of working out during the period of t days.

Time Spent Watching TV: Let $B_{TV}(x, t)$ represent behavior score of a user x for watching TV. For every two hours spent on watching TV, the chance of getting obese will increase 23%. Conversely, watching TV for less than 10 hours per week can help decrease the risk of obesity (that is average 1.43 hours per day) [15]. So we have:

$$B_{TV}(x, t) = f_{TV}\left(\frac{N_{TV}(x,t)}{t}\right) = \begin{cases} 0, & 0 \leq \frac{N_{TV}(x,t)}{t} < 1.43 \\ -0.5, & 1.43 \leq \frac{N_{TV}(x,t)}{t} < 3.43 \\ -1, & \frac{N_{TV}(x,t)}{t} \geq 3.43 \end{cases} \quad (8)$$

where $N_{TV}(x,t)$ is total hours spent on watching TV in t days.

Computer Use: Computer use referred here is the computer use in leisure time, such as online chatting, web surfing and playing computer games. We use the protocol applied to TV watching for categorizing computer use. Let $B_{CU}(x, t)$ represent behavior score of a user x for computer use in leisure time, then we have:

$$B_{CU}(x, t) = f_{CU}\left(\frac{N_{CU}(x,t)}{t}\right) = \begin{cases} 0, & 0 \leq \frac{N_{CU}(x,t)}{t} < 1.43 \\ -0.5, & 1.43 \leq \frac{N_{CU}(x,t)}{t} < 3.43 \\ -1, & \frac{N_{CU}(x,t)}{t} \geq 3.43 \end{cases} \quad (9)$$

where $N_{CU}(x,t)$ is total hours spent on using computer in t days.

Behavior Score: Let $WGRB(x,t)$ represent behavior score for a user x, we define $WGRB(x,t)$ as below:

$$WGRB(x,t) = \frac{1}{N} \sum_{k=1}^{N} w_k B_k(x,t) \quad (10)$$

where N is the number of behaviors, $w_k$ is weight of the kth behavior.

Behavior Network

In this section, we introduce a novel concept called behavior network. This concept is derived from the studies of peer effect in adolescent obesity [16-17]. Our intuition is that, if a user is able to join a network that consists of friends with healthy lifestyle, he or she is more likely to adopt the same behaviors, thus reducing the risk of gaining weight.

Definition 1 (Behavior Network) Consider a social graph $G$ of a node $x$, where each connected node is complemented with its behavior score. The score for this behavior network of $x$ is defined as:

$$NB(x) = \frac{1}{N} \sum_{k=1}^{N} w_k WGRB_k \quad (11)$$

where $w_k$ denotes the weight of the $k$th Friend, $N$ is the total number of friends of $x$, and $WGRB_k$ is the behavior score of the $k$th friend. Behavior network provides a way to evaluate the impact of social ties on a target person. A higher score for behavior network means a healthier environment.

Data Experiment

Data were drawn from Health Behavior in School-Aged Children, 2009-2010 for United States (HBSC) [18]. The investigated subjects of HBSC are high school and middle school students. In our experiment, we only consider subjects that intend to lose weight, which reduces the data set to 2320 subjects. The characteristics of those subjects are summarized in Figure 4. Profile items used in experiment are gender, age, race, grade, height and weight. The height and weight are quantized. For dietary behaviors, we
consider FV intake, skipping breakfast, and frequency of eating fast food. For aerobic activity, we consider hours spent watching TV, hours spent playing video/computer games and hours of exercise. To embed subjects in a social graph, we use NetworkX\(^2\), a python language software package, to generate a network based on Barabási-Albert model [19]. Each newly imported user connects to 4 nodes of existing network. Once all values are computed, we rank the subjects by adding up their scores for the four components.

**Comparisons of Ranking Measures**

To evaluate FRHW, we compare it with other well-known similarity measures, namely, L1 Norm, OF similarity for profile and network and profile similarity measures (NPS) proposed in literature [9]. The definitions of L1 norm for network similarity \(L_{1N}\) and profile similarity \(L_{1P}\) are given as below, respectively:

\[
L_{1N}(u, x) = \frac{|F(u) \cap F(x)|}{|F(u)| + |F(x)|} \quad (12)
\]

\[
L_{1P}(u, x) = \frac{1}{|I|} \sum_{k=1}^{|I|} I_k(u) = l_k(x) \quad (13)
\]

where \(|F(u)|\) is the friend of user \(u\), \(|F(u) \cap F(x)|\) is the mutual friends of \(u\) and \(x\), \(I_k(u)\) is the value of the \(k\)th profile item for user \(u\) and \(|I|\) is the number of items in profile.

For NPS, we combine network similarity and profile similarity as follow:

\[
NPS(u, x) = NS_{\text{norm}}(u, x) + P_{\text{norm}}(u, x) \quad (14)
\]

where \(NS_{\text{norm}}(u, x)\) is the normalized network similarity value for user \(u\) and stranger \(x\) and \(P_{\text{norm}}(u, x)\) is the normalized profile similarity value.

As suggested by the literature [1], the recommended people should not be obese because it will increase the risk of gaining weight for a target user. Surrounded by friends with healthy weight may also help establish correct perception of body image. Based on this assumption, we have our first measurement. Given a possible friend \(e\), we say that a ranking measure \(rkm_i\) wins over another ranking measure \(rkm_j\), if the number of healthy weight users in top \(k\) users recommended according to \(rkm_i\), denoted as \(NRKMi(e,k)\), is higher than the number obtained according to \(rkm_j\), denoted as \(NRKMj(e,k)\). Similarly, when \(NRKMi(e,k) < NRKMj(e,k)\) a loss for \(rkm\) occurred; when \(NRKMi(e,k) = NRKMj(e,k)\), a draw is recorded.

The average behavior score of the recommended friend is another factor we need to consider, because it indicates the lifestyle of this friend. Given a possible friend \(e\), we say that a ranking measure \(rkm\) wins over another ranking measure \(rkm\), if the average behavior score of \(e\)'s top \(k\) recommended friends is higher than the value obtained by \(rkm\), denoted as \(AVGRKMi(e,k)\) > \(AVGRKMj(e,k)\). Similarly, when \(AVGRKMi(e,k) < AVGRKMj(e,k)\), a loss for \(rkm\) occurred; when \(AVGRKMi(e,k) = AVGRKMj(e,k)\), a draw is recorded.

**Result and Analysis**

In this section, we show how the proposed measure can recommend more friends with healthier lifestyle. After computing similarities, we rank the similarity values and compare the top 10 recommendations for each user. Table 1 shows the comparison result based

---

\(^2\) https://networkx.github.io/
We also made comparisons for subjects who are not willing to lose weight. 2320 subjects are randomly selected from the same dataset while preserving the same proportion of weight status. Result is shown in Table 2. If weights for both WGRB and NB are increased, best result is achieved. One possible explanation is that in this group of people, weight gain related behaviors are more related to their weight status.

In all of the cases, the average behavior scores obtained by FRHW are higher than behavior scores obtained by other measures, which may help to construct a healthier living environment and to change a target user’s behaviors.

Limitations of this experiment are obvious. First of all, the size of data set is small. Second, it is not realistic that all the users will provide full information of their weight gain related behaviors. Third, the subjects of our experiment are all school-aged children. They cannot represent the general population. Thus we need to validate FRHW on a larger data set with a more representative population.

**Conclusion and Future Work**

Though the role of social network in overweight or obesity intervention has been well studied, measures linking them together are rarely explored. Aiming at constructing a healthy living environment for a target user, we developed a method using data obtained by weight loss app to recommend friends with healthy weight and lifestyle. It not only saves the burden of searching but also fully utilizes data gathered by weight loss app. Modeling user’s weight gain behavior is the key components of FRHW. It helps evaluate a user’s dietary behavior as well as level of physical activity. We suggested several weight gain related behaviors that can be considered in app and gave methods to score each of them. A novel concept called “Behavior Network” is built upon acquisition of behavior scores. The proposed measure is tested on a dataset with 2320 subjects. Weight status and average behavior score of recommended friends are considered as criterion of comparison. Comparing with other traditional friend recommendation measures, FRHW has ability to recommend friends with healthier lifestyle to a target user.

In the future, we need to find a way to predict the missing values of both profile items and weight related data. Secondly, we should develop a method to identify the influences of the recommended friends and their roles in helping target user as weight loss companion.

**Acknowledgments**

The authors thank the reviewers for their enlightening comments. This work was partially supported by
Internal Fund of Department of Computing, Hong Kong Polytechnic University.

References