Industrial Management & Data Systems

Friend recommendation for healthy weight in social networks: A novel approach to weight loss

Anming Li Eric W.T. Ngai Junyi Chai

Article information:

To cite this document:
Anming Li Eric W.T. Ngai Junyi Chai , (2015),"Friend recommendation for healthy weight in social networks", Industrial Management & Data Systems, Vol. 115 Iss 7 pp. 1251 - 1268
Permanent link to this document:
http://dx.doi.org/10.1108/IMDS-04-2015-0130

Friend recommendation for healthy weight in social networks

A novel approach to weight loss

Anming Li and Eric W.T. Ngai
Department of Management and Marketing, The Hong Kong Polytechnic University, Kowloon, Hong Kong, People's Republic of China, and
Junyi Chai
Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands

Received 14 April 2015
Revised 31 May 2015
Accepted 2 June 2015

Abstract

Purpose – The purpose of this paper is to propose a new approach recommending friends to social networking users who are also using weight loss app in the context of social networks.

Design/methodology/approach – Social network has been recognized as an effective way to enhance overweight and obesity interventions in past studies. However, effective measures integrating social network with weight loss are very limited in the healthcare area. To bridge this gap, this study develops a measure for friend recommendation using the data obtained by weight loss apps; designs methods to model weight-gain-related behaviors (WGRB); constructs a novel "behavior network;" and develops two measurements in experiments to examine the proposed approach.

Findings – The approach for friend recommendation is based on Friend Recommendation for Health Weight (FRHW) algorithm. By running this algorithm on a real data set, the experiment results show that the algorithm can recommend a friend who has a healthy lifestyle to a target user. The advantages of the proposed mechanism have been well justified via comparisons with popular friend recommenders in past studies.

Originality/value – The conventional methods for friend recommenders in social networks are only concerned with similarities of pairs rather than interactions between people. The system cannot account for the potential influences among people. The method pioneers to model a WGRB as recommendation mechanism that allow recommended friends to simultaneously fulfill two criteria. They are: first, similarity to the target person; and second, ensuring the positive influence toward weight loss. The second criterion is obviously important in practice and thus the approach is valuable to the literature.

Keywords Healthcare, Obesity, Social network, Recommendation

Paper type Research paper

1. Introduction

1.1 Background

Recommendation system (RS) processes the information by suggesting to users the objects that are possibly compatible with their interests (Adomavicius and Tuzhlin, 2005; Kim, 2011). RS has been comprehensively investigated in different perspectives such as preference (Koutrika and Ioannidis, 2010), trust (Bedi *et al.*, 2007), decision making (Chai *et al.*, 2014), and position localization (Bao *et al.*, 2015). Online social

The authors are grateful for the constructive comments of the referees on an earlier version of this paper. The third author thanks LEI Chun-Fong for helpful comments.

Industrial Management & Data Systems
Vol. 115 No. 7, 2015
pp. 1251-1268

0263-5577
DOI 10.1108/IMDS-04-2015-0130

networks offer users opportunities to expand their circle of friends through a RS (Curras-Perez *et al.*, 2014). The system breaks regional restrictions and unites like-minded people who have similar interests or goals (Yang *et al.*, 2012). For example, a soccer fan may associate with other fans to enjoy FIFA World Cup.

Online social networks are great sources of health-related information and social support. They provide access to people who share the same experience and 24/7 encouragement, which can be difficult for realistic friends. Goel and Goldstein (2014) reported the application of social networks to predict the behavior of individuals with respect to the large social data in marketing environments. Centola (2010) investigated the influence of social network structure on diffusion by studying the spread of health behavior through artificially structured online communities. He found that the proliferation of behavior across clustered-lattice networks would be farther and faster than that across corresponding random networks.

A weight management program could be a long-term process and requires constant self-monitoring. Individuals are likely to concede in the early stage and eventually fail the weight management program. An obese user who participates in a weight management program may want to know other participants in the program in social networks, thus promoting each other for better achievement. Weight loss apps respond to such requests and typically require users to monitor their food intake and physical activities to calculate calorie consumption.

1.2 Literature review

Existing literature provided some valuable solutions for better weight management. Maher *et al.* (2014) reviewed the effectiveness of interventions based on social network. Those interventions aimed to change health behaviors such as dietary intake and physical activity. Among the studies, 90 percent reported positive outcomes. Chang *et al.* (2013) reviewed the role of social media in online weight management. They pointed out that social media could help those who lack social support establish connections and obtain support. Hwang *et al.* (2010) investigated peer support in an online weight loss community, and suggested that 60 percent 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.

Li *et al.* (2013) explored the role of social networks and social media in obesity intervention. They identified three pathways that link obesity with social networks, namely, social support, social integration, and social capital. They advocated for the development of social networks specifically designed to address obesity. However, Kiernan *et al.* (2012) and Gorin *et al.* (2005) found that obesity intervention involving spouse, friends, co-workers, and neighbors had mixed results. Kiernan *et al.* (2012) found that women with family support were more likely to lose weight (71.6 percent), whereas women who did not obtain support from friends were most likely to lose weight (80 percent). Gorin *et al.* (2005) demonstrated that participants with one or more successful partners lost more weight compared with those without successful partners.

A landmark study by Christakis and Fowler (2007) used longitudinal statistical models to examine whether weight gain is contagious. They showed that the risk of a person becoming obese increased by 57 percent if a friend became obese. They further revealed that the involvement of partners in weight management process could provide support and emotional buffer to people who face a similar problem. In other words, obesity is "socially contagious," implying that people tend to be obese if their family and friends are obese. Pagoto (2012) observed that, "not only do we share habits, but we subtly (and sometimes not so subtly) reinforce each other's habits. To the extent that you are locked in a social circle that strongly reinforces unhealthy habits, you will have great difficulty consistently living a healthy lifestyle." The experiments of Kiernan *et al.* (2012), Gorin *et al.* (2005), and Shoham *et al.* (2012) involving multiple roles of concerned persons (e.g. family members, friends, and neighbors) revealed mixed results. Shoham *et al.* (2012) investigated the social influence on body size, screen time, and playing sports of adolescents using an actor-based model. The results showed that both social influence and homophily are important to understand adolescent obesity.

Equipped with online social networking service, some weight loss apps have been developed using social network to aid the weight loss process. For example, MyFitnessPal (www.myfitnesspal.com/) encourages members to invite their friends to go on a diet with them to achieve better results. We also observed that users of weight loss forum also proactively attempted to seek encouragement, incentive, and support by gathering people who share the same goals, starting weight, and so on. This effort is mostly achieved by sending out a post on a forum. Finding a friend can be filtered using user features, such as name, gender, location, and age. Either way, the process is labor-intensive for users. MyFitnessPal has approximately 40 million users as of July 2013 and continues to grow at a rate of 1.5 million new users per month.

1.3 Motivation and our approach

Based on past works, weight loss intervention involving friends and family may fail because the concerned persons with health-risk behaviors, like frequently eating fast food, could have a negative effect on the target people; the concerned persons may be more tolerant with obese persons, and therefore cannot act as rigorous and helpful overseers (Christakis and Fowler, 2007); the concerned persons and the target person may be influenced by a similar cultural norm, which could promote or strengthen unhealthy behavior (Li *et al.*, 2013). The criteria that qualify a companion for weight loss should include: familiarity to the target person in some ways, positive influence on the weight loss process, and ability to resist the unhealthy cultural norm, if any.

In our literature review, we observed that the concept of "weight loss with friends" is widely accepted in social community, such as self-organized weight-loss activities in social communities. Despite the benefits offered by social network for body weight management, the effective measure is surprisingly very limited. Existing mobile app or weight loss software possesses the ability to collect user information on daily diet, activities, and so on, but they have not been properly considered in weight loss app. On the measures of social network, most past studies only considered the similarities among persons, but they cannot ensure the presence of both similarity and positive influences among persons in friend recommendation. Motivated by these issues, this study proposes an approach to use friend recommendation in social network in body weight management.

This friend recommendation approach is referred to as Friend Recommendation for Health Weight (FRHW). This work systematically examines the measure, which consists of four components, namely, network similarity, profile similarity, methods for scoring weight-gain-related behavior (WGRB), and behavior network. Our aim is to recommend friends who both possess similar features of targeted users and bring positive influence on the entire weight-reducing process. The former two components are common measures frequently adopted in past studies. The latter two components were developed in this paper by incorporating domain knowledge. In Section 2, we elaborate the first two components for general friend recommendation. We will elaborate the latter two measures in Section 3. We subsequently provide a mechanism that uses all four measures to rate the WGRB. We construct a novel behavior network using all output scores in Section 4. Finally, we combine all the components and obtain a final ranking for friend recommendation in Section 5. In data experiments, we use two measurements as the criteria to examine the performance of our solution. Our results are compared with the results of other representative measures. Section 6 concludes this paper.

2. General measures for social network-based recommendations

2.1 Related works

Intuitively, social networking users tend to connect with people who are close or similar to them or who share similarities with them (Ngai *et al.*, 2015). Therefore, many friend recommenders build their measures on computing user similarity from different perspectives. Chen *et al.* (2009) performed two experiments on Beehive, an enterprise social networking site, to assess four friend recommendation algorithms. They concluded that known contacts could be effectively discovered by using the network-based algorithm whereas the algorithm that was based on similarity of user-created content, such as profile entries, tags, and statuses could be used to recommend new friends. Akcora *et al.* (2011) proposed the network and profile-based measures to compute user similarity. The proposed measure considers network similarity and strength of friendships. Spertus and Sahami (2005) compared six similarity measures, such as, first, $L1$ norm; second, $L2$ norm; third, pointwise mutual information (PMI) with positive correlations; fourth, PMI with positive and negative correlations; fifth, Salton's similarity measure based on inverse document frequency scaling; and sixth, log-odds on Orkut.com data for community recommendation. They argued that one of the similarity measures, $L2$ norm, outperformed the other five measures. Boriah *et al.* (2008) compared existing 14 methods for similarity computation of categorical data. They applied these methods for outlier detection. The results showed that none of them can be totally superior to the other methods, and each method should be applied according to different needs. Vosecky *et al.* (2009) identified users across different social networks, and defined profile vector, which represents a user profile in the form of vector. This similarity can be computed and stored in a vector known as similarity vector. In Sections 2.2 and 2.3, we briefly introduce two kinds of similarity, namely, network and profile similarities, which have been frequently referred in past studies.

2.2 Network similarity

Network similarity measures the degree of overlap between two users in social network. Specifically, the measure attempts to define closeness between the two users. Suppose a user x is connected to a target user u within a two-hop distance in the social graph, the user is called a stranger. For network similarity, we use the measure from Akcora *et al.* (2011). For user u and stranger x, network similarity can be computed via the equation below:

$$NS(u,x) = \frac{\mathrm{Log}\left(\left|MFG(u,x).E\right|\right)}{\mathrm{Log}\left(2\left|FG(u).E\right|\right)} \tag{1}$$

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

Example. Suppose we consider a social graph as shown in Figure 1. The corresponding results are as follows: $FG(u).E = 8$, $MFG(u,x).E = 7$, and $NS(u,x) = \log(7)/\log(2 \times 8) = 0.702$, wherein 0.702 indicates the proximity between two persons in a network. A larger number indicates that the two persons are closer to each other.

2.3 Profile similarity

Profile similarity is used to represent the similarity between two users in the form of profile items. We use the method provide by Akcora *et al.* (2011) to calculate such profile similarity, which is based on occurrence frequency (hereafter OF; Jones, 1993). Suppose i be a profile item, i_u and i_x be the value for item i in u and x profiles. Profile similarity $P(u,x)$ can be computed via the following equations:

$$OF(i_u, i_x) = \begin{cases} 1, & i_u = i_x \\ \frac{1}{(1+A\times B)}, & i_u \neq i_x \end{cases} \tag{2}$$

$$P(u,x) = \frac{1}{n}\sum_{k=1}^{n} w_k \times \ OF_k \tag{3}$$

Therein, $A = \log(N/f(i_u))$ and $B = \log(N/f(i_x))$.

Function $f(x)$ represents the number of records with that item value. N is the total number of records; w_k is the weight assigned to kth profile item; n is the number of profile items. The weight for kth profile item, denoted as w_k, is user-defined.

Example. We consider user u with profile P_u{Gender: Female, Race: Asian} and stranger x with profile P_x{Gender: Female, Race: Indian}. Suppose there are 1,000 records in total and distributions of gender and race are as shown in Table I. The weight for each item is one. Then, the corresponding results are as the follows:

$$P(u,x) = \frac{1}{2} \times \left\{ 1 + \left(1 + \log\frac{1,000}{80} \times \log\frac{1,000}{102} \right)^{-1} \right\} = 0.74$$

The value of $P(u,x)$ represents the similarity of two users in terms of profile items. Two users are more similar if the number is larger.

3. WGRB: definition and modeling

In this section, we define and model WGRB. WGRB comprises dietary behavior and physical activity. They are easy to quantify and record in mobile apps. Some of them are actually used in the mobile apps available in the market. The connection

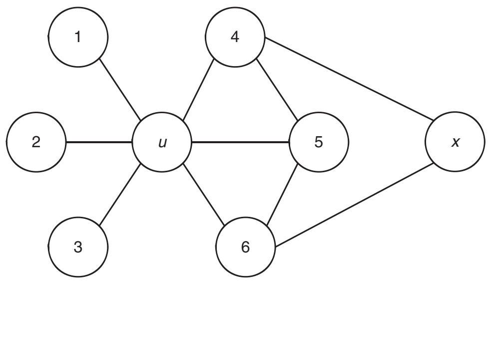

Figure 1.
Social network graph

between these kinds of behavior and obesity or weight gain have been investigated and supported by numerous studies (World Health Organization (WHO), 2003; 2010; Ma *et al.*, 2003; Hu *et al.*, 2003; Forslund *et al.*, 2005).

For dietary behavior, we examine the consumption of fruit and vegetable (FV hereafter), meal eating pattern (i.e. skipping breakfast), and snacking, because these behaviors are studied in plenty of research works. First, FV consumption is the better predictor for healthy eating. High intake of energy-dense food is considered as one of the convincing factors that contribute to weight gain while high intake of NSP (dietary fiber) will decrease the risk, as pointed out by a report of WHO (see the reference Trogdon *et al.*, 2008). FV are considered low-energy-dense food and rich in dietary fiber and water. Intake of FV is negatively associated with weight gain (Kim, 2011). Second, Ma *et al.* (2003) indicated that a regular meal pattern was significantly related to a low level of obesity. In the WHO report, breakfast feeding is categorized as a probable factor. Third, regular meal eating patterns were also associated with low snacking levels, although snacking was not associated with obesity in a direct manner.

For physical activity behavior, we consider three factors that are commonly regarded as most important weight-gain-associated factors, that is, aerobic activity, time spent watching TV, and computer use. Exercise is generally the key for maintaining fitness. According to Hu *et al.* (2003), considerable time spent on sedentary behavior, such as watching TV and playing video games, is associated with greater risk of obesity.

Suppose a weight loss app user records food log and exercise log on daily basis. We define a set of WGRB, where B_k denotes kth behavior. $B_k(x, t)$ is the behavior score for user x over a given period of t days. We define WGRB separately as shown in Sections 3.1-3.7. We derive the following equations from obesity and nutrition researchers. To the best of our knowledge, this paper is the first to formalize and combine these findings by considering the features of weight loss apps.

3.1 FV intake

Let $BFV(x,t)$ represent the behavior score of user x for FV intake and $N_{FV}(x,t)$ for the number of days that x has FV no less than 400 g during t days. The value of 400 g is supported by the report of the WHO (2003) and Agudo (2005).

The subsequent equations are valid:

$$B_{FV}(x,t) = f_{FV}\left(\frac{N_{FV}(x,t)}{t}\right) = \begin{cases} 0, & 0 \leqslant \frac{N_{FV}(x,t)}{t} < 0.5 \\ 0.5, & 0.5 \leqslant \frac{N_{FV}(x,t)}{t} < 0.75 \\ 1, & 0.75 \leqslant \frac{N_{FV}(x,t)}{t} \leqslant 1 \end{cases} \tag{4}$$

	Frequency
Gender	
Female	502
Male	498
Ethnicity	
Asian	80
Indian	102
Others	818

Table I. Distribution of gender and ethnicity

In the study, FV Intake is divided into three categories:

(1) low intake if $N_{FV}(x,t)/t < 0.5$;

(2) moderate intake if $0.5 \leqslant N_{FV}(x,t)/t < 0.75$; and

(3) high intake if $N_{FV}(x,t)/t \geqslant 0.75$.

3.2 Skipping breakfast

Let *BBE*(*x,t*) represent the behavior score on the frequency of skipping breakfast. According to Ma *et al.* (2003), skipping breakfast frequently could increase the risk of gaining weight. We categorize the frequency of skipping breakfast using this method. The following equations are then valid:

$$B_{BE}(x,t) = f_{BE}\left(\frac{N_{BE}(x,t)}{t}\right) = \begin{cases} 0, & 0 \leqslant \frac{N_{BE}(x,t)}{t} < 0.75 \\ 1, & 0.75 \leqslant \frac{N_{BE}(x,t)}{t} \leqslant 1 \end{cases}, \tag{5}$$

where $N_{BE}(x,t)$ represents the total number of days that the user has eaten breakfast during t days.

3.3 Snacking

Let $B_{SN}(x,t)$ represent the behavior score of user x for snacking. Increasing snacking frequency will cause increasing in energy intake, it also easily results in bad eating habit such as overeating. But association between snacking and weight gain is inconclusive. Snacking can be an additional source of dietary fiber and FV. Generally, the argument depends on the selection of snacks. Given little research explores relationship between snack selection and obesity, we only consider the aspect of the unhealthy snacks. Unhealthy snack means junk food, sugary food, and beverage. Classification of frequency of snacking is illustrated according to Forslund *et al.* (2005). Given this information, the following equation is valid:

$$B_{SN}(x,t) = f_{SN}\left(\frac{N_{SN}(x,t)}{t}\right) = \begin{cases} 0, & 0 \leqslant \frac{N_{SN}(x,t)}{t} < 1 \\ -0.25, & 1 \leqslant \frac{N_{SN}(x,t)}{t} < 2 \\ -0.5, & 2 \leqslant \frac{N_{SN}(x,t)}{t} < 3 \\ -1, & \frac{N_{SN}(x,t)}{t} \geqslant 3 \end{cases}, \tag{6}$$

where $N_{SN}(x,t)$ represents the number of snacking in t days.

3.4 Aerobic activity

Let $B_{AA}(x, t)$ represent the behavior score of user x for aerobic activity. Through simple calculation of time, an adult should at least spend 150 min on moderately intense aerobic activity per week or 21.4 min per day on average. This calculation is based on the report from WHO (2010). Then, we define $B_{AA}(x, t)$ via the following equations:

$$B_{AA}(x,t) = f_{AA}\left(\frac{N_{AA}(x,t)}{t}\right) = \begin{cases} 0, & 0 \leqslant \frac{N_{AA}(x,t)}{t} < 21.4 \\ 1, & \frac{N_{AA}(x,t)}{t} \geqslant 21.4 \end{cases} \tag{7}$$

where $N_{AA}(x,t)$ represents the total number of minutes of exercise during the period of t days.

3.5 Time spent watching TV

Let $B_{TV}(x, t)$ represent the behavior score of user x for watching TV. According to Hu *et al.* (2003), watching TV for less than 10 hr per week decreases the risk of obesity or 1.43 hr per day on average. Nevertheless, the risk of becoming obese will increase 23 percent for each 2 hr per day increment in watching TV (i.e. 3.43 hr per day). Given this information, we define $B_{TV}(x, t)$ using the following equation:

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

where $N_{TV}(x,t)$ represents the total number of hours spent watching TV in t days.

3.6 Computer use

Computer use in this study refers to its use during leisure time for online chatting, web surfing, and playing games. We employ the protocol applied to watching TV to categorize computer use. Let $B_{CU}(x, t)$ represent the behavior score of user x for computer use in leisure time. Then, the following equation is valid:

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

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

3.7 Behavior score

Let $WGRB(x,t)$ denote the behavior score for user x. We define $WGRB(x,t)$ via the following equation by using the previously obtained scores:

$$WGRB(x,t) = \frac{1}{N}\sum\nolimits_{k=1}^{N} w_k B_k(x,t) \tag{10}$$

where N is the number of behavior of users as listed above (i.e. WGRBs) and w_k is weight of kth behavior.

Example. Consider WGRB of a user, Tom, for a given period, say a week. The weight assigned to each behavior is 1. Suppose the information provided by Tom is as follows:

- days of eating breakfast: 4 ($N_{BE} = 4$);
- days of having 400 g or more FV: 3 ($N_{FV} = 3$);
- hours spent watching TV: 27 ($N_{TV} = 27$); and
- total minutes of exercise: 120 ($N_{AA} = 120$).

Then, the behavior score of Tom for that week can be calculated as:

$$WGRB(Tom, t) = 1/4 \times \{0+0+(-1)+0\} = -0.25$$

Given the assumption that a user uses a mobile app to track his/her WGRB, we attempt to consider all factors that may contribute to weight gain and use them to model the user behavior. This measure is provided to quantitatively measure the behavior of a user. Usually, a negative number indicates that the user has an unhealthy lifestyle, which may contribute to weight gain. A positive number represents the health degree of the user lifestyle. Therefore, the lifestyle is healthier as the number is larger.

4. Behavior network

We develop a behavior network using the output scores obtained from Section 3. We were inspired by the studies of peer effect in adolescent obesity (Trogdon *et al.*, 2008; Ali *et al.*, 2011). Suppose a user joins a social network that consists of friends with healthy lifestyle. This user tends to adopt similar behavior, and thereby reduces the risk of weight gain.

4.1 Definition: behavior network

Consider social graph G of node x, in which each connected node is complemented with its behavior score. The score for this behavior network of the node x can be defined via the following equation:

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

where w_k represents the weight of the kth friend, N is the total number of friends of x, and $WGRB_k$ is the behavior score of the kth friend. The defined behavior network provides a method to evaluate and verify the effect of social ties with respect to the target person. Obviously, a higher score in behavior network implies a healthier environment.

Example. Consider a social graph as shown in Figure 2. Suppose the weight for each friend is one. Then, NB score for user u is calculated as $NB(u) = 1/6\times \{(-0.5)+(-0.25)+0.5+0.4+(-0.1)\} = 0.0083$.

In this study, we assume that the behavior of friends will affect the weight loss process of the user. We use $NB(u)$ to measure one's social ties in terms of whether they promote a healthy lifestyle or whether they will help in the weight loss process. If one gains high $NB(u)$,

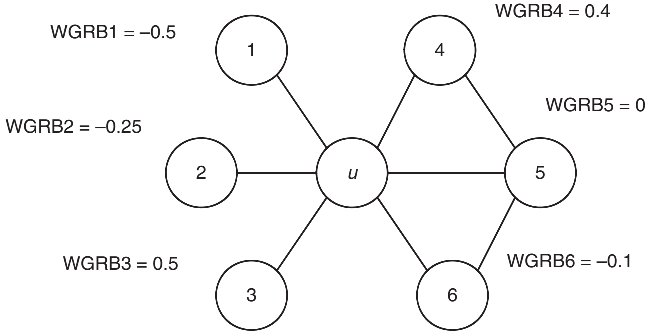

Figure 2. Social graph complemented with behavior score

then his/her social network can positively influence his/her weight loss. By contrast, if one's $NB(u)$ is low, then he/she is in a social network that may hinder his/her weight loss process.

4.2 FRHW algorithm

We include all parts of our method into FRHW algorithm that can be run in the computer. The pseudo codes are shown below. As a preliminary setting, the number of friends of a user is known. This algorithm calculates profile similarity, the number of profile items, the possible values for each profile item, and the number of users who have specific values for a profile item. Then, the profile and network similarities of each user are computed with respect to other users in the database. After generating the WGRB score of users, this algorithm is run to calculate NB value for each user. Finally, the four values obtained above are combined to get final score of FRHW.

N denotes the number of records in database, N_p is the number of profile items in a profile, N_{WGRB} is the number of WGRBs, and $w(\cdot)$ corresponds to weight function:

Algorithm: FRHW
Description:
for $u \leftarrow 1$ to N, do
    $FG(u).E \leftarrow$ get number of friends of u
end for
//Get Statistical Data for Each Profile Item
for each profile item i in profile
    for each possible value i_x of i
        $f(i_x) \leftarrow$ get number of users whose i is equal to i_x
    end for
end for
for u←1 to N, do
    for v←1 to N, do,
        if u = v
        continue
        end if
        //Compute Profile Similarity
        for each profile item i in profile
            if $f(i_u) = f(i_v)$
                $OF_i(u,v) = 1$
            else
                $OF_i(u,v) = \frac{1}{\log\left(1+\frac{N}{f(i_u)}+\frac{N}{f(i_v)}\right)}$
            end if
            $P(u,v) = P(u,v) + w(i)OF_i$
        end for
        $P(u,v) = P(u,v)/N_p$
        //Compute Network Similarity
        $|MFG(u,v).E| \leftarrow$ get number of common friends of u and v
        if $|MFG(u,v).E| = 0$
            $NS(u,v) = 0$
        else
            $NS(u,v) = \frac{Log(|MFG(u,v).E|)}{Log(2*|FG(u).E|)}$
        end if

```
        end for
        //Compute Scores for Weight Gain Related Behaviors (WGRBs)
        for each weight gain related behavior, Bi:
            Bi(u)←get behavior score of user u
            WGRB(u) = WGRB(u)+w(Bi) Bi(u)
        end for
        WGRB(u) = WGRB(u)/N_WGRB
    end for
    for u←1 to N, do
        if |FG(u).E| = 0
            NB(u) = 0
        else
            for each friend of user u, v
                NB(u) = NB(u)+w(u,v)*WGRB(v)
            end for
                NB(u) = NB(u)/|FG(u).E|
        end if
    end for
    for u←1 to N, do
        for v←1 to N, do,
            if u = v
                continue
            end if
        FRHW(u,v) = w(P)* P(u,v)+w(NS)* NS(u,v)+w(WGRB)* WGRB(v)+w(NB)* NB(v)
        end for
        rank FRHW scores for u
    end for
```

5. Experiments and analyses

5.1 Data preparation

Our experiments adopted the data from a Health Behavior in School-Aged Children survey that was conducted in the USA during the 2009-2010 cycle. The survey focussed on the experiences and attitudes of school-aged children toward various health-related behaviors, including dietary habits and physical activities. The data were collected through questionnaires. A total of 314 schools participated in the survey and 12,642 valid samples were collected. Additional details on this survey could be found in Iannotti (2013). By considering only those subjects who aimed to lose weight, we reduced the data set to 2,320 subjects. Table II summarizes the features of our subjects.

The profile items used in the experiment include gender, age, race, grade, height, and weight. Height and weight were quantized while other items are categorized by following codebook of the survey. For dietary behavior, we considered FV intake, skipping breakfast, and frequency of eating fast food. For aerobic activity, we employ the hours of spent watching TV, the hours of spent playing video/computer games, and the hours spent for exercise.

Given that the survey was in the form of a questionnaire, all aforementioned behaviors and their corresponding option values were coded in the original data set. As the first step of this experiment, we ran a MATLAB program to match these codes to our behavior models. To embed the subjects in a social graph, we used NetworkX, a Python language software package, to generate a network according to the

	Frequency
1. Gender	
Female	1,245
Male	1,074
NA	1
2. Weight status	
Underweight	32
Healthy weight	751
At risk of being overweight	514
Overweight	562
NA	461
3. Ethnicity	
Black or African American	367
White	1,065
Asian	79
Hispanic	491
Two or more races	152
Others	65
NA	101
4. Age	
10 yrs old	191
11 yrs old	304
12 yrs old	420
13 yrs old	446
14 yrs old	405
15 yrs old	372
16 yrs old	156
17 yrs old	24
NA	2
5. Grade	
Grade 5	281
Grade 6	353
Grade 7	453
Grade 8	438
Grade 9	407
Grade 10	388

Table II.
The five features of the data set

Barabási-Albert model (Albert and Barabasi, 2002). NetworkX is a scale-free network that follows power laws. More details on the software can be found in https://networkx.github.io/. Each newly imported user is connected to four nodes of the existing network. Subsequently, we ran a MATLAB program that followed the persudo-code and obtained a 2320×2320 matrix. Each row represented the information and the sorted FRHW score of a user.

5.2 Competitors of ranking measures

To compare the results, we examined representative similarity measures in past studies, such as *L*1 norm (Akcora *et al.*, 2011); OF similarity for profile, and network profile similarity (NPS) measures (Spertus and Sahami, 2005). This section describes methods to calculate the competitors.

The definitions of $L1$ norm for network similarity ($L1_N$) and profile similarity ($L1_P$) are represented below:

$$L1_N(u,x) = \frac{|F(u) \cap F(x)|}{|F(u)||F(x)|} \tag{12}$$

$$L1_P(u,x) = \frac{1}{|I|}\sum_{k=1}^{n}\begin{cases}1, & I_k(u) = I_k(x)\\ 0, & \text{otherwise}\end{cases} \tag{13}$$

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 kth profile item for user u, and $|I|$ is the number of items in profile.

For NPS, we combine and calculate network and profile similarities using the following equation:

$$NPS(u,x) = NS_{norm}(u,x) + P_{norm}(u,x), \tag{14}$$

where $NS_{norm}(u,x)$ is the normalized network similarity value with respect to user u and stranger x and $P_{norm}(u,x)$ represents the normalized profile similarity value.

5.3 Measurements for evaluations

Considering the suggestion of Christakis and Fowler (2007), the recommended people should not be obese because it will increase the risk of weight gain for the target user. Given that the user is surrounded by friends with healthy weight, the establishment of the correct perception on body image is also helpful. Based on this assumption, the first measurement used to evaluate measures can hold.

5.3.1 The first measurement. Given a possible friend e, 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 $N_{RKMi}(e,k)$, is higher than the number obtained according to rkm_j, denoted as $N_{RKMi}(e,k) > N_{RKMj}(e,k)$. Similarly, when $N_{RKMi}(e,k) < N_{RKMj}(e,k)$, a loss for rkm_i occurs; when $N_{RKMi}(e,k) = N_{RKMj}(e,k)$, a draw is recorded.

The average behavior score of a recommended friend reflects the lifestyle of this friend. The second measurement evaluates and verifies the different measures.

5.3.2 The second measurement. Given a possible friend e, a ranking measure rkm_i wins over another ranking measure rkm_j, if the average behavior score of e top k recommended friends is higher than the value obtained by rkm_j, denoted as $AVG_{RKMi}(e,k) > AVG_{RKMj}(e,k)$. Similarly, when $AVG_{RKMi}(e,k) < AVG_{RKMj}(e,k)$, a loss for rkm_i occurs; when $AVG_{RKMi}(e,k) = AVG_{RKMj}(e,k)$, a draw is recorded:

Example. Suppose two friend recommendation solutions, A and B, exist. Then, we obtain the top five recommended friends by A and B for a specific user Tom, separately. The behavior scores for these candidates are retrieved (see Table III). The average behavior score of A is AVGRKMA (Tom, 10) = 1/5×(1−1+0+1+1) = 2/5; the average score of candidate B is AVGRKMB (Tom, 10) = 0. Then, we conclude that A is better than B in this test case.

5.4 Results and discussions

The proposed measures were more suitable for recommending friends with healthier lifestyles. Our experiments computed and ranked the similarity values, and then compared the Top k ($k = 10$ in our experiments) recommendations for each user. The comparison result was based on weight status of the recommended friends as shown in Table IV.

Weight status was considered by FRHW and its competitors, but FRHW performed better because it recommended more friends with healthier weight, with the max in the quantity of win (i.e. 1,743) and the min in the quantity of loss (i.e. 522). This result showed using FRHW performed better than other competitors. $L1_N$ did not account for the profile items in the calculation. Profile items including age, height, and weight were significantly associated with the weight status. Compared with the other three measures of $L1_p$, OF, and NPS, the result had a positive effect on the performance of FRHW.

On the weight loss cases, several users with unhealthy weight were recommended repeatedly because of the high behavior scores. For example, 91 overweight or at risk of being overweight subjects were recommended for ten or more times in loss cases for $L1_P$, accounting for 92.6 perecnt of the recommendations. The median behavior score of these subjects was 0.82, within a range (0, 1). We observed that the same situation occurred on the equal cases. A possible reason is that people who intend to lose weight purposely choose healthy behavior. Meanwhile, those changes cannot be reflected on their weight status in a short time.

Furthermore, we compared the subjects who were not willing to lose weight. A total of 2,320 subjects were randomly selected from the identical data set while preserving the same proportion of weight status. Table V shows the results of the comparisons.

Table III. Behavior scores of top 5 recommended friends

Solution	Behavior scores				
A	1	−1	0	1	1
B	1	−1	−1	0	1

Table IV. The comparison results of Top k ($k = 10$) recommended friends with healthy weight

	Win	Equal	Loss
FRHW vs $L1_N$	1,743	55	522
FRHW vs $L1_P$	1,107	313	900
FRHW vs *OF*	1,087	373	860
FRHW vs *NPS*	1,090	372	858

Table V. The comparison results for people who do not want to lose weight

	Win	Equal	Loss
FRHW vs $L1_N$	1,773	149	398
FRHW vs $L1_P$	1,367	275	678
FRHW vs *OF*	1,288	387	645
FRHW vs *NPS*	1,289	386	645

The best result can be achieved when the weights for both WGRB and NB increased because WGRB was more associated with weight status in this group.

The average behavior scores obtained using FRHW were higher than that obtained from competitors. The result showed that our approach can reasonably help construct a healthier living environment and to optimize the behavior of target users. This experiment also has some limitations. First, the size of data set was relatively small. Second, this experiment assumed that users will provide full information of their WGRB, but in the real world, this assumption is not true all the time. Third, the subjects in our experiment were school-aged children rather than the general population. Validation of FRHW on larger data sets would be appealing.

6. Conclusions

Social network is the grouping of individuals into specific groups by basically communicating via the internet and/or mobile technology. Body weight management through social networking is more effective than conventional means because it is highly conducive to connect people based on shared interests and goals, which could be free of geographical limitations. Despite well-documented role and value of using social network in overweight or obesity intervention in past studies, the measures that link them together were rarely explored. To develop a healthy living environment for people in need, we developed an approach which utilized data obtained from weight loss apps to recommend friends who have healthy weight and lifestyle. The benefit of our approach is to relieve the burden of searching, as well as fully utilizing the data gathered by weight loss apps.

This study systematically elaborated four key components of FRHW, namely, network and profile similarities (Section 2), methods for scoring WGRB (Section 3), and behavior network (Section 4). We validated that our method can properly examine the dietary behavior and the level of physical activity of users. We suggested a series of WGRB that can be considered in apps. We developed methods to score each of them. A novel "behavior network" was constructed upon acquisition of behavior scores. The proposed measure was tested on a real data set with 2,320 subjects. We used two measurements as criteria to evaluate our method and conducted the comparisons. Our experiment examined two aspects of recommended friends, namely, the weight status as the first measurement and the average behavior score as the second measurement. The comparison results showed that FRHW had the better ability to recommend friends with healthier lifestyle to the target users.

The theoretical implications of this study include the extension of conventional friend recommender that only considered the similarities between pairs of people to the solid user-oriented considerations. For the practical implications, the present approach in terms of developed measures modeled WGRB and pioneered to construct a new form of social networking called "behavior network," which could be valuable for the development of software apps. In the future, we plan to explore how each WGRB affects the outcome of a weight loss program. Given that weight loss is affected by various factors, we cannot easily determine the weight for each behavior in the WGRB score calculation. We hope to obtain data from real weight loss applications and analyze the correlation between a specific WGRB and outcome. The findings of this study can be applied to weight loss app, particularly to help their users achieve better results. We also want to know how these behaviors spread through an online social network. To achieve this goal, a method for identifying the influences of recommended friends and their roles as weight loss companions must be developed.

References

Adomavicius, G. and Tuzhlin, A. (2005), "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17 No. 6, pp. 734-749.

Agudo, A. (2005), *Measuring Intake of Fruit and Vegetables*, *World Health Organization*, Geneva.

Akcora, C.G., Carminati, B. and Ferrari, E. (2011), "Network and profile based measures for user similarities on social networks", *Proceedings of the IEEE International Conference on Information Reuse and Integration (IRI 2011)*, pp. 292-298.

Albert, R. and Barabási, A.L. (2002), "Statistical mechanics of complex networks", *Reviews of Modern Physics*, Vol. 74 No. 1, pp. 47-97.

Ali, M.M., Amialchuk, A. and Heiland, F.W. (2011), "Weight-related behavior among adolescents: the role of peer effects", *PLoS One*, Vol. 6 No. 6, pp. 1-9.

Bao, J., Zheng, Y., Wilkie, D. and Mokbel, M. (2015), "Recommendations in location-based social networks: a survey", *Geoinformatica*, Vol. 19 No. 3, pp. 525-565.

Bedi, P., Kaur, H. and Marwaha, S. (2007), "Trust based recommender system for semantic web", *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007)*, pp. 2677-2682.

Boriah, S., Chandola, V. and Kumar, V. (2008), "Similarity measures for categorical data: a comparative evaluation", *Proceedings of the 8th SIAM International Conference on Data Mining*, pp. 243-254.

Centola, D. (2010), "The spread of behavior in an online social network experiment", *Science*, Vol. 329 No. 5996, pp. 1194-1197.

Chai, J.Y., Ngai, E.W.T. and Liu, J.N.K. (2014), "Dynamic tolerant skyline operation for decision making", *Expert Systems with Applications*, Vol. 41 No. 15, pp. 6890-6903.

Chang, T., Chopra, V., Zhang, C. and Woolford, S.J. (2013), "The role of social media in online weight management: systematic review", *Journal of Medical Internet Research*, Vol. 15 No. 11, p. e262, doi:10.2196/jmir.2852.

Chen, J., Geyer, W., Dugan, C., Muller, M. and Guy, I. (2009), "'Make new friends, but keep the old' – recommending people on social networking sites", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 09)*, pp. 201-210.

Christakis, N.A. and Fowler, J.H. (2007), "The spread of obesity in a large social network over 32 years", *The New England Journal of Medicine*, Vol. 357 No. 4, pp. 370-379.

Curras-Perez, R., Ruiz-Mafe, C. and Sanz-Blas, S. (2014), "Determinants of user behavior and recommendation in social networks: an integrative approach from the uses and gratifications perspective", *Industrial Management & Data Systems*, Vol. 114 No. 9, pp. 1477-1498.

Forslund, H.B., Torgerson, J.S., Sjöström, L. and Lindroos, A.K. (2005), "Snacking frequency in relation to energy intake and food choices in obese men and women compared to a reference population", *International Journal of Obesity*, Vol. 29 No. 6, pp. 711-719.

Goel, S. and Goldstein, D.G. (2014), "Predicting individual behavior with social networks", *Marketing Science*, Vol. 33 No. 1, pp. 82-93.

Gorin, A., Phelan, S., Tate, D., Sherwood, N., Jeffery, R. and Wing, R. (2005), "Involving support partners in obesity treatment", *Journal of Consulting and Clinical Psychology*, Vol. 73 No. 2, pp. 341-343.

Hu, F.B., Li, T.Y., Colditz, G.A., Willett, W.C. and Manson, J.E. (2003), "Television watching and other sedentary behaviors in relation to risk of obesity and type-2 diabetes mellitus in Women", *The Journal of the American Medical Association*, Vol. 289 No. 14, pp. 1785-1791.

Hwang, K.O., Ottenbacher, A.J., Green, A.P., Cannon-Diehl, M.R., Richardson, O., Bernstam, E.V. and Thomas, E.J. (2010), "Social support in an internet weight loss community", *International Journal of Medical Informatics*, Vol. 79 No. 1, pp. 5-13.

Iannotti, R.J. (2013), *Health Behavior in School-Aged Children (HBSC), 2009-2010*, Inter-University Consortium for Political and Social Research, Ann Arbor, MI.

Jones, K. (1993), "A statistical interpretation of term specificity and its application in retrieval", *Journal of Documentation*, Vol. 28 No. 1, pp. 11-21.

Kiernan, M., Moore, S.D., Schoffman, D.E., Lee, K., King, A.C., Taylor, C.B., Kiernan, N.E. and Perri, M.G. (2012), "Social support for healthy behaviors: scale psychometrics and prediction of weight loss among women in a behavioral program", *Obesity*, Vol. 20 No. 4, pp. 756-764.

Koutrika, G. and Ioannidis, Y. (2010), "Personalizing queries based on networks of composite preferences", *ACM Transactions on Database Systems*, Vol. 35 No. 2, pp. 1-50, available at: http://doi.acm.org/10.1145/1735886.1735892 (accessed January 1, 2015).

Kim, Y.S. (2011), "Text recommender system using user's usage patterns", *Industrial Management & Data Systems*, Vol. 111 No. 2, pp. 282-297.

Li, J.S., Barnett, T.A., Goodman, E., Wasserman, R.C. and Kemper, A.R. (2013), "Approaches to the prevention and management of childhood obesity: the role of social networks and the use of social media and related electronic technologies: a scientific statement from the American heart association", *Circulation*, Vol. 127, pp. 260-267.

Ma, Y., Bertone, E.R., Stanek, E.J., Reed, G.W., Hebert, J.R., Cohen, N.L., Merriam, P.A. and Ockene, I.S. (2003), "Association between eating patterns and obesity in a free-living US adult population", *American Journal of Epidemiology*, Vol. 158 No. 1, pp. 85-92.

Maher, C.A., Lewis, L.K., Ferrar, K., Marshall, S., Bourdeaudhuij, I.D. and Vandelanotte, C. (2014), "Are health behavior change interventions that use online social networks effective? A systematic review", *Journal of Medical Internet Research*, Vol. 16 No. 2, p. e40, doi:10.2196/jmir.2952.

Ngai, E.W.T., Moon, K.K., Lam, S.S., Chin, E.S.K. and Tao, S.S.C. (2015), "Social media models, technologies, and applications: an academic review and case study", *Industrial Management & Data Systems*, Vol. 115 No. 5, pp. 769-802.

Pagoto, S. (2012), "How social media can help you lose weight", *Psychology Today*, October 10.

Shoham, D.A., Tong, L.P., Lamberson, P.J., Auchincloss, A.H., Zhang, J., Dugas, L., Kaufman, J.S., Cooper, R.S. and Luck, A. (2012), "An actor-based model of social network influence on adolescent body size, screen time, and playing sports", *PLoS One*, Vol. 7 No. 6, pp. 1-11.

Spertus, E. and Sahami, M. (2005), "Evaluating similarity measures: a large scale study in the Orkut social network", *Proceeding of the 11the ACM International Conference on Knowledge Discovery in Data Mining (SIGKDD)*, pp. 678-684.

Trogdon, J.G., Nonnemaker, J. and Pais, J. (2008), "Peer effects in adolescent overweight", *Journal of Health Economics*, Vol. 27 No. 5, pp. 1388-1399.

Vosecky, J., Hong, D. and Shen, V.Y. (2009), "User identification across social networks using the web profile and friend network", *International Journal of Web Applications*, Vol. 2 No. 1, pp. 23-34.

World Health Organization (WHO) (2003), "Diet, nutrition and the prevention of chronic diseases", report of a joint FAO/WHO Expert Consulation, World Health Organization, Geneva.

World Health Organization (WHO) (2010), "Global recommendations on physical activity for health", Report of World Health Organization, World Health Organization, Geneva.

Yang, X.W., Steck, H. and Liu, Y. (2012), "Circle-based recommendation in online social networks", *Proceedings of the 18th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD)*, pp. 1267-1275.

Corresponding author
Dr Junyi Chai can be contacted at: donchaiam@gmail.com

This article has been cited by:

1. Yafei Wang, Ingmar Weber, Prasenjit MitraQuantified Self Meets Social Media 93-97. [CrossRef]