# Interaction Network Representations for Human Behavior Prediction

Amnay Amimeur[1], NhatHai Phan[1], Dejing Dou[1], David Kil[2], Brigitte Piniewski[3]
{aamimeur,haiphan,dou}@cs.uoregon.edu, david.kil@healthmantic.com, BPiniewski@peacehealthlabs.org
[1] University of Oregon, USA; [2] Health Mantic Inc., USA; [3] PeaceHealth Laboratories, USA

***Abstract*— Human behavior prediction is critical to studying how healthy behavior can spread through a social network. In this work we present a novel user representation based human behavior prediction model, the *User Representation-based Socialized Gaussian Process* model (UrSGP). First, we present the *Deep Interaction Representation Learning* (Deep Interaction) model for learning latent representations of interaction social networks in which each user is characterized by a set of attributes. In particular, we consider social interaction factors and user attribute factors to build a bimodal, fixed representation of each user in the network. Our model aims to capture the evolution of social interactions and user attributes and learn the hidden correlations between them. We then use our latent features for human behavior prediction via the UrSGP model. An empirical experiment conducted on a real health social network demonstrates that our model outperforms baseline approaches for human behavior prediction.**

## I. Introduction

Research has identified physical inactivity as a leading risk factor associated with a variety of cardiovascular diseases and other chronic conditions such as obesity, diabetes, and joint problems [1]. Online social networks have garnered interest from the health informatics community because network users can influence each other to promote healthy behavior.

Several studies on Internet-delivered interventions have reported positive behavioral outcomes [2], [3]. Online social networks can help people interact and participate in various physical activities to promote physical activities with affordable cost. However, there has been a lack of quantitative study on how social networks may encourage users to engage in physical activity.

Human behavior prediction in online health social networks aims to predict an individual's behavior given their past behavior where behavior, generally, includes personal behavior history, social correlation, and social influence. Current state-of-the-art human behavior prediction models do not directly take into account information regarding social interaction between users and how this social interaction evolves over time. The homophily principle states that people tend to interact with those who share similar characteristics or attributes. Thus, we expect a strong correlation between the evolution of user attributes and the evolution of social interaction over time in health social networks. In addition, the benefit of declaring friend connections is marginal compared with actual interaction [4]. Therefore, we should exploit interaction in order to build accurate human behavior prediction models. To this end, we incorporate into state-of-the-art human behavior prediction models latent representations of network users that capture both the evolution of social interaction and user attributes over time.

Network representation learning has emerged as a popular topic within the social and information network analysis. It aims to learn low-dimensional latent representations of users that implicitly capture the social semantics of the network. Recent work in network representation learning includes the DeepWalk algorithm [5] which learns social representations of vertices of a graph by modeling short random-walks along the graph topology. DeepWalk is the first application of *deep feature learning* to network representation learning and inspires our own representation learning method. However, much of the work to date [5]–[9] on network representation learning attempts to gain insight into social dynamics by only considering the static friend connections within the network and not the dynamic social interaction layer which explicitly documents the transfer of information between users over time. Ignoring the social interaction layer can result in a substantial loss of information.

This work improves the current state-of-the-art in human behavior prediction by incorporating latent social representations into the Socialized Gaussian Process model (SGP) [10]. We call the new model the User Representation-based Socialized Gaussian Process Model (UrSGP). We also improve upon existing network representation learning methods by learning directly from the dynamic interaction layer of social networks instead of the static friend topology. Furthermore, we apply a multimodal deep learning approach to learn features that capture the correlation between user interaction and user attributes. Empirical testing of our new human behavior prediction model, conducted on a real health social network, demonstrates the effectiveness of the learned features, showing we are able to significantly improve prediction accuracy for human behavior prediction.

Our contributions are as follows:

- We capture the evolving structure of social interaction in networks over time by employing an interaction random walk through the interaction layer of social networks.
- We construct our features by learning a fixed, bimodal representation that relates the evolution of a user's social interaction to their attributes.
- We improve upon existing baseline approaches for human behavior prediction by incorporating our user representations into the Socialized Gaussian Process Model.

The rest of the paper is organized as follows. Section 2 discusses related work. Section 3 provides important background information needed for clarifying our problem definition and contributions. In section 4, we present our Deep Interaction model and our User Representation-based Socialized Gaussian Process model. Section 5 presents our results for human behavior prediction via interaction network representations.

## II. Related Work

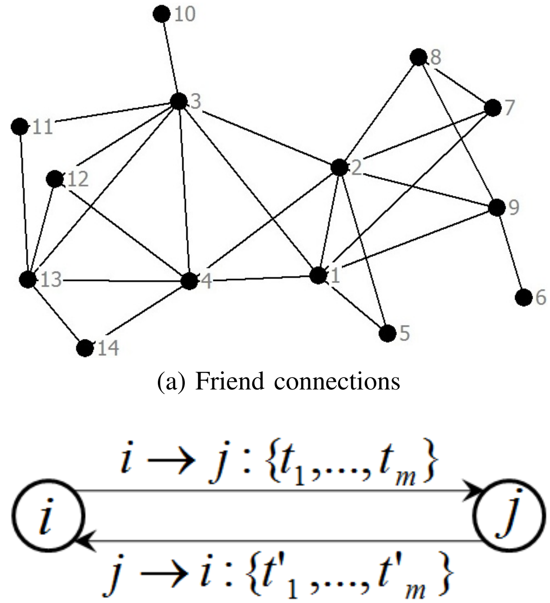
Prediction of human social behavior has been studied recently. In this work, we improve upon the social correlation factor of the Socialized Gaussian Process model [10], described in detail in Section 4.3, which is not designed to work with user attributes. Phan et al. [11], [12] propose leveraging Restricted Boltzmann Machines for human behavior prediction by structuring the visible layer to incorporate users' features both historical and current at prediction time. However, they are not able to include information regarding social interaction.

Representation learning is widely used in computer vision, natural language processing, and knowledge representation learning [13]. Some researchers have focused on network representation learning (NRL) [5], [9], [14]–[16] but none consider the most fundamental phenomenon of social networks, the evolution of social interaction and user attributes over time. The closest work to our paper is the recently proposed DeepWalk model [5]. DeepWalk learns social representations of vertices of graphs by modeling short random-walks. In [16], the authors propose a Deep Graph Kernel (Deep GK) method to learn similarities between structured objects, such as graphs and strings. In [14], Cao et al. propose a model, named GraRep, for learning graph representations for knowledge management. The model captures $k$-step relational information with different values of $k$ amongst vertices from the graph by manipulating different global transition matrices. Tang et al. [15] present an efficient edge-sampling algorithm for preserving first order and second order proximity in graph embeddings. Ahmed et al. [17] introduce graph embedding via matrix factorization by representing a graph as an affinity matrix. This technique can only be applied to undirected graphs. Another improvement of the DeepWalk model is the text-associated DeepWalk (TADW) [9] model. The TADW model incorporates text features of vertices into network representations under the framework of matrix factorization.

The related works discussed thus far use only shallow models. Wang et al. [18] proposed a deep framework for representation learning called Structural Deep Network Embedding in order to capture the very highly nonlinear, complex, and sparse nature of the underlying structure of social networks.

Some work focuses on using latent representations or social dimensions for collective behavior prediction. Tang et al. [19] apply various edge clustering approaches for community detection and convert the resulting edge partitions into a social dimension representation. They consider behavior prediction as a multi-label classification task but do not account for information outside of the network topology, ignoring historical behavior, user attributes, and user interaction.

Our research differs from state-of-the-art models in several aspects. We learn the similarities between users by: 1) Considering the evolution of the social network over time; 2) Capturing the evolution of social interactions; 3) Capturing the change in user attributes; and 4) Learning the hidden correlations between social interactions and user attributes. We expect latent representations that capture both a user's interaction with the network and the correlations between social interaction and attributes to provide a strong indication of social correlation among two users which we leverage to create our UrSGP model.




(a) Friend connections

(b) Received messages

Fig. 1. Static topology versus interaction layer

## III. Background and Problem Definitions

### A. Social Interaction Networks

*Social network* is an umbrella term used to describe a set of networks layered together to capture and facilitate social interaction between users. Among these layers, the *friend topology* or *friend network* can be defined by a set of edges denoting friendship between users. Accompanying the friend topology is the underlying *interaction network* which differs from the friend topology by having edges defined by a set of timestamps of interaction. In other words, given a graph $\mathcal{G} = (V, E, \mathcal{A})$, where $V$ is the set of users, $E$ is the set of edges, and $\mathcal{A}$ is the set of attributes of the network, the interaction network can be defined as a set of edges written as [20]

$$E = \{(u_i, v_i, t_i) | (u_i, v_i) \in V \ and \ t_i \in \mathbb{R}\} \tag{1}$$

This would suggest that the interaction network may change over time whereas the friend topology is static as shown in Figure 1. Interaction networks can be considered an implicit encoding of the dynamics of social interaction between users or groups of users in the network, and as a result, can provide more insight into these dynamics than the friend topology.

### B. Attribute Vectors

We aim to take advantage of the growing richness of social network data by incorporating a set of characteristic user attributes into our model, denoted as an *attribute vector*. Our goal is to generate a set of features that capture both the dynamic of user interaction in the network and the correlation between interaction and user attributes. We use the term attribute vector to generally denote any set of characteristic features, both static and dynamic, that could describe a user at a particular time. In our healthcare application domain, an attribute vector may include biomarker measurements like HDL and LDL cholesterol levels or biometric measures such as height, weight, and body mass index (BMI) where all measurements are taken at what can be considered the same time. Given a network $\mathcal{G} = (V, E, \mathcal{A})$, the attributes of the network are defined as $\mathcal{A} = \{(v_i, A_t^i) | v_i \in V\}$ where the attribute vector $A_t^i = (a_1^i, a_2^i, ..., a_M^i)$ and $M$ is the dimension of the attribute vector and $a_m^i$ is an attribute measured at time $t$ corresponding to user $i$.

### C. Learning Representations of Words and Paragraphs

Representation learning and hierarchical feature extraction have developed into a subfield of machine learning and data mining due in part to the emergence and popularity of deep learning over the past decade.

Recently, the language modeling community has taken advantage of unsupervised feature learning techniques to generate semantic representations of words in a context. The popular SkipGram and Continuous Bag-of-Words models [21] accomplish this by training neural networks to maximize the probability of words in a context given a word and the probability of a word given its context, respectively. The key to these models is the softmax layer on top of the final hidden layer of the network which forces the network output to sum to one, effectively creating a probability distribution over network weights. Given and input word $w_I$, we define the probability of a word in the vocabulary $w_o$ given $w_I$ as:

$$P(w_o|w_I) = \frac{exp(v'^T_{w_o} v_{w_I})}{\sum_{w \in W} exp(v'^T_{w} v_{w_I})} \tag{2}$$

where $v'^T_{w_o} v_{w_I}$ is output of the node in the output layer of the SkipGram network corresponding to word $w_o$ in the vocabulary. The energy function of the SkipGram network accordingly becomes:

$$E = -\log P(w_{o,1}, ..., w_{o,C} | w_I) \tag{3}$$

where $w_{o,C}$ is the $C$-th word of the output context.

By maximizing the probability of output weights given hidden weights, the network can learn a low dimensional vector representation of each word in the vocabulary.

The Distributed Memory Model of Paragraph Vectors (PV-DM) [22] is an extension of the SkipGram model. The PV-DM is quite similar to the SkipGram model with one exception. The model learns representations for paragraphs in documents by concatenating word vectors with paragraph vectors to build the hidden layer of the neural network.

The PV-DM model inspires the inclusion of user attribute vectors with social interaction representations to build a unified representation that captures the evolution of user attributes with social interaction, as detailed in Section 4.2.

### D. Problem Definitions

**Learning Latent Social Interaction Representations.** Let $\mathcal{G} = (V, E, \mathcal{A})$ be an interaction network where $V$ is the set of users in the network, $E$ is the set of all interactions between users over all time such that $\forall e_{ij} \in E, e_{ij} = \{t_1, ..., t_n\}$ where $t_m$ is a time of interaction between users $i$ and $j$, and $A$ is the set of attribute vectors of all users over all time. We aim to learn a fixed representation, $H$, of the evolution of social interaction and attribute vectors over time, where $H \in \mathbb{R}^{|V| \times d}$ and $d$ is the dimension of the representation.

**Human Behavior Prediction.** Given the social network $\mathcal{G}$ and individuals' past behaviors until day $t$, $\mathcal{X}_{1..t} = (X^1_{1..t}, X^2_{1..t}, ..., X^N_{1..t})$, where $X^i_{1..t} = (x^i_1, x^i_2, ...x^i_t)$ with $x^i_t \in (-1, 0, 1)$. Note, $x^i_t = 1$ indicates user $u_i$ plays sports at day $t$, $x^i_t = -1$ indicates whether a user $u_i$ does not play sports at day $t$, while $x^i_t = 0$ indicates user $u_i$'s record is missing at day $t$. $N$ is the number of users in the social network. The socialized human behavior prediction problem is to predict the individual's behaviors at day $t+1$, i.e., $\mathcal{X}_{t+1}$.

## IV. Feature Learning and Behavior Prediction

In this section, we present our Deep Interaction Representation Learning model to learn joint representations between social interactions and user attributes. We then show how we improve the Socialized Gaussian Process model by using our latent representations to create the User Representation-based Socialized Gaussian Process model.

### A. Interaction Random Walk

Perozzi et al. [5] proposed choosing streams of short random walks as a primitive for capturing graph topology structure. By capturing graph structure as a stream of random walks, they are able to apply new natural language modeling techniques to learn fixed low dimensional representations of nodes in a graph. We adopt this primitive but make key changes to accommodate walking through a dynamic network. We define the probability of transitioning from the current vertex $u$ to the next vertex $v$ as a distribution over the difference between the current time and the next timestamp along the edge connecting the two vertices. More formally we define a transition probability from vertex $u$ to vertex $v$ along edge $k$ at time $t$ as:

$$P_{uv,t} = 1 - \frac{t_{kv} - t}{\sum_{i \in T} t_i - t} \tag{4}$$

where $t_{kv}$ is the next timestamp along edge $k$ and $T$ is the set of the next timestamps of interaction along all edges connected to $u$.

In order to choose the next step in the walk, we repeat Bernoulli trials independently for each edge in the set defined above and choose to travel along the edge with the highest number of successful outcomes.

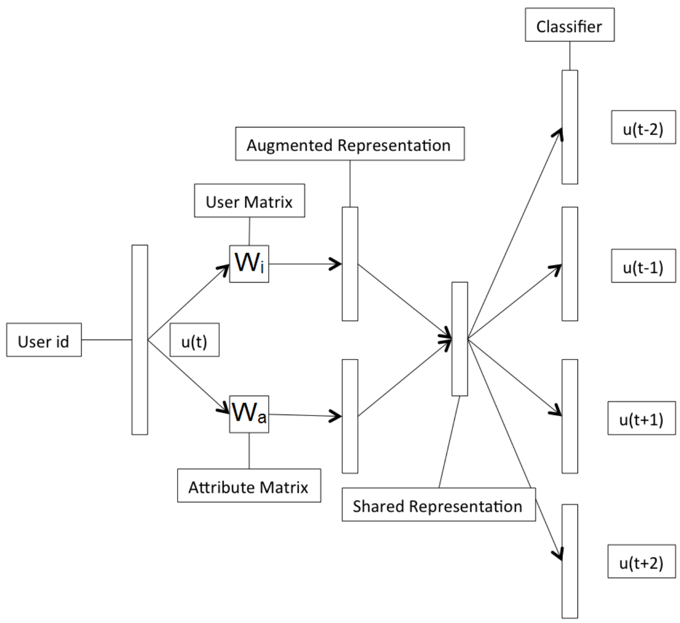



Fig. 2. Deep Interaction Representation Learning Model

*B. Deep Interaction Representation Learning*

In this section, we present our Deep Interaction Representation Learning (Deep Interaction) model to learn joint representations between social interactions and user attributes. First we present the model then detail how we capture social structure through the dynamic interaction layer of a social network. The objective of our model is to learn low dimensional fixed representations of users from the dynamic of social interaction within a network as well as users' attributes. We capture the social structure of a network by employing streams of random walks along the interaction layer of the social network. Given a random walk $\mathcal{R}_{u_i} = (u_{i-r}, ..., u_{i+r})$ where $r$ is the length of the walk we define the context of $\mathcal{R}_{u_i}$ as $\{u_{O,1}, ..., u_{O,C}\}$ where $C$ is $w-1$. Our aim is to apply language modeling techniques to learn a representation of $u_i$ given its context, however we make a few changes to the SkipGram model used by [5] to incorporate user attributes. We follow work on multimodal deep learning [23] in order to learn joint representations of user interaction and user attributes which are used for prediction.

The input to our model (Figure 2) is a one-hot encoded vector for selecting the appropriate user representation and attribute vector from the hidden weights. The two representations are averaged/concatenated and joined once more via an additional hidden layer to learn the joint representation of interest. The final hidden layer is a softmax layer, and in practice we use a hierarchical softmax layer to speed up training time.

More formally, given matrices $\mathbf{W_i}$ and $\mathbf{W_a}$ as shown below, where matrix $\mathbf{W_a}$ is the set of attributes for the network and matrix $\mathbf{W_i}$ holds the user interaction representations. The components of the augmented hidden layer are

$$\mathbf{h_i} = \mathbf{x}^{\mathbf{T}}\mathbf{W_i} \tag{5}$$

$$\mathbf{h_a} = \mathbf{x}^{\mathbf{T}}\mathbf{W_a} \tag{6}$$

Note that the attribute matrix $\mathbf{W_a}$ is $|V| \times k$ where $|V|$ is the number of users in the network and $k$ is the dimension of the historical attribute vector of each in the network. Historical attribute vectors are formed by concatenating all attribute vectors over all time for each user in $\mathcal{G}$. Therefore, $\mathbf{W_a}$ can be considered a historical record of all attributes of all users over all time in the network. $\mathbf{h_i}$ denotes the user interaction component of the augmented hidden layer selected from the $|V| \times b$ weight matrix $\mathbf{W_i}$ where $b$ is the dimension of $\mathbf{h_i}$. $\mathbf{h_a}$, of dimension $k$, denotes the user attribute component of the augmented hidden layer. The shared hidden layer is constructed as

$$\mathbf{h_s} = [\mathbf{h_i}|\mathbf{h_a}]^{\mathbf{T}}\mathbf{W_s} \tag{7}$$

where $\mathbf{W_s}$ is the $|V| \times d$ joint representation matrix which holds our final user representations. The objective of the Deep Interaction model is to minimize the following energy function:

$$E = -\log p(u_{O,1}, ..., u_{O,C}|u_i) \tag{8}$$

The output layer of our model is a mixture of $C$ multinomial distributions, one for each user in the context. We can consider the output layer to be composed of C concatenated panels, each a multinomial distribution over all $N$ users in the network as shown in Figure 2. We can then define the output of the $j$-th node of the $c$-th panel as the following.

$$y_{c,j} = p(u_{c,j} = u_{O,C}|u_i) = \frac{exp(g_{c,j})}{\sum_{j'=1}^{|V|} exp(g_{c,j'})} \tag{9}$$

where

$$g_{c,j} = \mathbf{h_s}^{\mathbf{T}}\mathbf{W_{s'}} \tag{10}$$

and $W'_s$ is the weight matrix from the shared hidden layer to the output layer. $g_{c,j'}$ corresponds to the pre-normalized output of the $j'$-th node on the $c$-th panel of the output layer calculated in a similar fashion as $g_{c,j}$. Note that Equation 6 is exactly the softmax function where the denominator serves to normalize the output of the $j$-th node of the $c$-th panel in order to create a multinomial probability distribution.

The energy function accordingly becomes:

$$E = -\log \prod \frac{exp(g_{c,j})}{\sum_{j'=1}^{|V|} exp(g_{c,j'})} \tag{11}$$

To train our model, stochastic gradient descent and back-propagation are used minimize the energy function given in Eq. 11. The joint representations are learned in two steps. First the interaction representations are pre-trained independently via the SkipGram model. Once the interaction representations are learned, they are used as input to the next layer of the model which learns the joint representations by stochastic gradient descent where the gradient is obtained via backpropagation. As in Le et al. [22], all other parameters, i.e. the interaction representations and softmax parameters, are fixed. In other words, user attributes and interaction representations are the input to joint layer where the inference objective is as defined in Eq. 8.

As shown in Figure 2, the model has two hidden layers and a softmax output layer. In practice we use a hierarchical softmax layer which determines overall depth of the neural network since it is implemented as a binary tree with $|V|$ leaf nodes.

### C. Latent Features for Behavior Prediction

To improve the Social Gaussian Process (SGP) model [10], which is known as an efficient human behavior prediction model, we focus on improving the social correlation and implicit social influence parameters. Shen et al. [10] propose defining social correlation as inversely proportional to the square distance between the historical behaviors of two users. Instead of only using statuses of users in estimating the social correlations among users, we incorporate user representations into the social correlation function in the SGP model as follows:

$$\Omega_{ij,t} = \frac{1}{T} exp\Big[ - \sum_{m=t-T\ldots t} (X_m^i - X_m^j)^2 \Big] + \Big( 1 - cosine(\mathbf{h_{s,i}}, \mathbf{h_{s,j}}) \Big) \quad (12)$$

where $\mathbf{h_{s,i}}$ and $\mathbf{h_{s,j}}$ are the shared hidden features $\mathbf{h_s}$, also called user representations of user $i$ and user $j$, constructed from the shared hidden layer in Eq. 7. $T$ is the number of timestamps in the training set. $\Omega_{ij,t}$ is the social correlation between user $i$ and user $j$ at timestamp $t$. We follow [21] in choosing cosine similarity as a distance metric for measuring similarity between two user representations. By using latent social representations in the social correlation parameter, we are able to leverage the social interaction and user information encoded in the social network to strengthen our notion of correlation between two users and their behaviors, which we show by comparing against relevant baselines. The model is called the *User Representation-based Socialized Gaussian Process* (UrSGP) model.

## V. Experiments

**Human Physical Activity Dataset.** The dataset was collected via a collaboration among several health laboratories and universities to help people maintain active lifestyles and lose weight. The dataset is collected from 254 users, including personal information, a social network, and their daily physical activities over ten months.

The initial physical activity data, collected from each user via special electronic equipment for each user, records information such as the number of walking and running steps. Since some users' daily records are missing, we show the basic analysis on the distribution of physical activity record numbers in Figure 3. In the Figure 3, there are 14 users with their daily physical activity record number smaller than 10, and 8 users with their record number larger than 10 but smaller than 20. Thus, to clean the data, we filtered the users whose daily physical activity record number is smaller than 80. In addition, we only consider users who contribute to the social

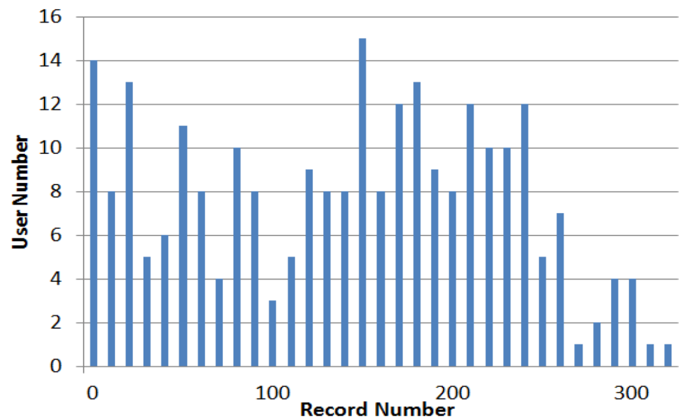



Fig. 3. Distribution of the record number and user number in the health social network.

communication (i.e., users must send (resp., receive) messages to (resp., from) other users). We only consider users who interact with others in the network (i.e., users must send (resp., receive) messages to (resp., from) other users). We have 123 users with 2,766 inbox messages for our experiments. Figure 4 illustrates the distributions of friend connections and the number of received messages in the health social network, showing they clearly follow the Power law distribution.

To validate our Deep Interaction model, we leverage the Socialized Gaussian Process model [10] by using the user representations learned to improve the social correlation modeling. Our UrSGP model is used for human behavior prediction on a real health social network. We first elaborate on the experiment configurations, evaluation metrics, and baseline approaches. Then, we introduce the experimental results.

**Dataset and Experiment Configurations.** In total, we have 30 features taken into account (Table III). All the features are weekly summarized. We use random walks of length five with ten walks per node to generate the ”corpus” from which we train the model in Figure 2. The dimension of the joint representation learned is 100. The weights are randomly initialized from a zero-mean Gaussian with a standard deviation of 0.01 with learning rates set to $10^{-3}$.

**Evaluation metric.** To verify the effectiveness of our novel, state-of-the-art human behavior prediction model, we predict the individual's future activities according to their past behaviors and social network information. In the experiment, we select two weeks as the unit for prediction, i.e., leveraging the previous 10 weeks' daily records to predict the $11^{th}$ and $12^{th}$ weeks' behaviors of users. We use the metric **accuracy** to measure the prediction quality between week $t$ and $t+1$.

$$accuracy = \frac{\sum_{i=1..N} \sum_{d \in \{t,t+1\}} I((X_d^i \neq 0) = \widetilde{X}_d^i)}{\sum_{i=1..N} \sum_{d \in \{t,t+1\}} I(X_d^i \neq 0)} \quad (13)$$

where $X_d^i$ is the true user activity at day $d$ for $u_i$, and $\widetilde{X}_d^i$ denotes the predicted value. $(X_d^i \neq 0)$ indicates the physical activity record is not missing. $I$ is the indication function, where $I(X) = 1$ when $X$ is true, otherwise $I(X) = 0$. $N$ is the number of users.

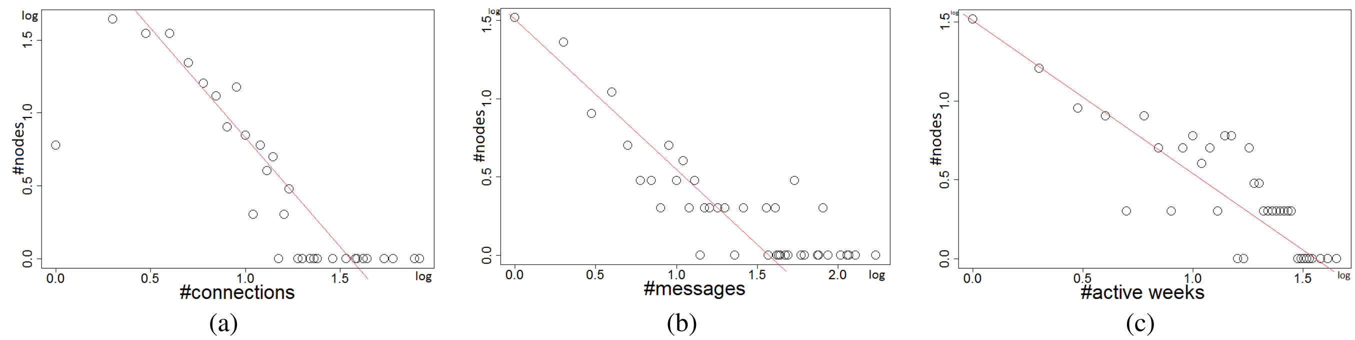



Fig. 4. The distributions of friend connections (a), inbox messages (b), and active users (c) in the health social network dataset.

TABLE I

PREDICTION ACCURACY COMPARISON WITH DIFFERENT MODELS(T7-T38)

| Weeks | LAR | PGP | SLAR | BPS | RF | SGP | SRF | UrSGP |
|---|---|---|---|---|---|---|---|---|
| T7-T8 | 0.56122 | 0.62448 | 0.55918 | 0.53877 | 0.68571 | 0.71836 | 0.73244 | **0.75306** |
| T9-T10 | 0.60404 | 0.63872 | 0.62716 | 0.5881 | 0.63873 | 0.64595 | **0.69717** | 0.69075 |
| T11-T12 | 0.64591 | 0.64177 | 0.62102 | 0.5988 | 0.70393 | 0.67731 | **0.70731** | 0.70539 |
| T13-T14 | 0.67729 | 0.71812 | 0.69023 | 0.6344 | 0.71833 | 0.74402 | 0.74808 | **0.76494** |
| T15-T16 | 0.69497 | 0.71923 | 0.71490 | 0.6603 | 0.71833 | 0.74870 | 0.74982 | **0.76603** |
| T17-T18 | 0.72879 | 0.74734 | 0.7526 | 0.70053 | 0.74812 | 0.75088 | 0.74988 | **0.76855** |
| T19-T20 | 0.73877 | 0.73518 | 0.7441 | 0.6921 | 0.73454 | 0.76211 | 0.75896 | **0.78007** |
| T21-T22 | 0.72760 | 0.74680 | 0.7239 | 0.7029 | 0.74755 | 0.75411 | 0.75666 | **0.77240** |
| T23-T24 | 0.72310 | 0.73526 | 0.71562 | 0.6604 | 0.73727 | 0.74555 | 0.75535 | **0.76427** |
| T25-T26 | 0.71372 | 0.71567 | 0.70001 | 0.6679 | 0.71691 | 0.73417 | 0.72741 | **0.75365** |
| T27-T28 | 0.72745 | 0.76089 | 0.7173 | 0.693 | 0.74983 | 0.77001 | 0.75322 | **0.79027** |
| T29-T30 | 0.72866 | 0.71553 | 0.73304 | 0.6969 | **0.76973** | 0.74288 | 0.74002 | 0.76477 |
| T31-T32 | 0.78965 | 0.79083 | 0.7755 | 0.7403 | 0.72259 | 0.80846 | 0.74575 | **0.83196** |
| T33-T34 | 0.74007 | 0.75416 | 0.7387 | 0.7055 | 0.76347 | 0.76056 | 0.76382 | **0.78617** |
| T35-T36 | 0.71059 | 0.73913 | 0.7105 | 0.6752 | 0.73051 | 0.74048 | 0.73894 | **0.76766** |
| T37-T38 | 0.74616 | 0.76490 | 0.7427 | 0.712 | 0.74383 | 0.76320 | 0.75723 | **0.79727** |

TABLE II

PAIRED T-TEST(2-TAIL) RESULTS

| | SGP | SRF | RF | PGP | SLAR | LAR | BPS |
|---|---|---|---|---|---|---|---|
| UrSGP | 0.009 | 0.009 | 0.006 | $1.36e-3$ | $2.77e-4$ | $2.02e-4$ | $5.89e-8$ |

**Competitive Prediction Models.** We compare the UrSGP model with the conventional methods reported in [10]. The competitive methods are divided into two categories: personalized behavior prediction methods and socialized behavior prediction methods. Personalized methods only leverage an individual's past behavior records for future behavior predictions. Socialized methods use both an individual's past behavior records and his or her friends' past behaviors for predictions. Specifically, the five models reported in [10] are Socialized Gaussian Process (**SGP**) model, Socialized Logistical Autoregression (**SLAR**) model, Socialized Random Forest (**SRF**) model, Random Forest (**RF**) model [24], Personalized Gaussian Process (**PGP**) model, Logistical Autoregression (**LAR**) model, and Behavior Pattern Search (**BPS**) model. Note that the Socialized Random Forest (**SRF**) model incorporates not only the personalized historical behaviors but also the user's social network information. When creating random forest trees, it combines the user's physical behavior and his or her friends' physical behaviors at each time period as a set of different features. Then the model picks whichever one has the best information gain, and splits the tree.

TABLE III

PERSONAL CHARACTERISTICS.

| | | |
|---|---|---|
| **Behaviors** | #joining competitions | #exercising days |
| | #goals set | #goals achieved |
| | $\sum$(distances) | avg(speeds) |
| **Social Communications (the number of inbox messages)** | Encouragement | Fitness |
| | Followup | Games |
| | Competition | Personal |
| | Study protocol | Technique |
| | Progress report | Meetups |
| | Social network | Goal |
| | Wellness meter | Feedback |
| | Heckling | Explanation |
| | Invitation | Notice |
| | Technical fitness | Physical |
| **Biomarkers** | Wellness Score | BMI |
| | Wellness Score slope | BMI slope |

**Experimental Results.** We report the performance of different human behavior models for predicting an individual's future behaviors. The individual's behavior records are divided according to the time series, e.g., $T1-T8$ indicates the records from the first week to the eighth week. Therefore, we can evaluate the models at different time periods. As shown in the

Table I, we compare accuracy across the eight human behavior prediction models.

Table I shows the User Representation-based Socialized Gaussian Process model (UrSGP) outperforms the other baseline methods and the SGP model. The proposed UrSGP model achieves further improvement based on SGP by incorporating the user presentations learned from our social interaction network into the dynamic social correlation information. In terms of accuracy, the UrSGP method improves the performance in average as high as 2.34%, 2.43%, 3.92%, 4.55%, 7.38%, 7.76% and 16.09% in contrast to SRF, SGP, RF, PGP, SLAR, LAR and BPS respectively.

Finally, to validate the statistical significance of our experiments, we perform the paired t-test (2-tail) over the accuracy of the experimental results. As shown in Table II, all the t-test results are less than 0.01, which means the improvements of UrSGP over other methods are statistically significant.

## VI. Conclusion

In this paper, we proposed a new human behavior prediction model, the User Representation-based Socialized Gaussian Process model, which utilizes latent social interaction representations to capture social correlation among users in a social network. We also present a novel Deep Interaction Representation Learning (Deep Interaction) model to learn the social network representation given the evolution of social interactions and user attributes. An empirical experiment conducted on a real health social network demonstrates that our user representations can be used to significantly improve the accuracy of human behavior prediction. However, as the size of a social network grows and the sparsity of interaction increases, our method cannot generate enough walks to cover the entire network. In future works, we explore ways to solve this problem in order to scale our method to large networks. We will conduct more experiments using our feature learning method on other social network analysis tasks, such as community detection, link recommendation, and graph kernel learning in order to improve our understanding of health and social behaviors and their propagation in online social networks.

## Acknowledgment

This work is supported by the NIH grant R01GM103309 to the SMASH project. We thank Xiao Xiao and Rebeca Sacks for their contributions. Dr. Nhathai Phan has now joined the New Jersey Institute of Technology faculty as an Assistant Professor.

## References

[1] D. E. Warburton, C. W. Nicol, and S. S. Bredin, "Health benefits of physical activity: the evidence," *Canadian medical association journal*, vol. 174, no. 6, pp. 801–809, 2006.

[2] B. Marcus, C. Nigg, D. Riebe, and L. Forsyth, "Interactive communication strategies: implications for population-based physical activity promotion." *American Journal of Preventive Medicine*, vol. 19, no. 2, pp. 121–6, 2000.

[3] C. Vandelanotte, K. Spathonis, E. Eakin, and N. Owen, "Website-delivered physical activity interventions: A review of the literature." *American Journal of Preventive Medicine*, vol. 33, no. 1, pp. 54–64, 2007.

[4] N. Phan, J. Ebrahimi, D. Kil, B. Piniewski, and D. Dou, "Topic-aware physical activity propagation with temporal dynamics in a health social network," *(Accepted by) ACM TIST*, 2016.

[5] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014, pp. 701–710.

[6] S. Cao, W. Lu, and Q. Xu, "Grarep: Learning graph representations with global structural information," in *CIKM '15*. ACM, 2015, pp. 891–900.

[7] L. Deng and D. Yu, "Deep learning: Methods and applications," *Foundations and Trends in Signal Processing*, vol. 7, no. 3–4, pp. 197–387, 2014.

[8] P. Yanardag and S. Vishwanathan, "Deep graph kernels," in *KDD '15*. ACM, 2015, pp. 1365–1374.

[9] C. Yang, Z. Liu, D. Zhao, M. Sun, and E. Y. Chang, "Network representation learning with rich text information," in *Proceedings of the 24th International Conference on Artificial Intelligence*, ser. IJCAI'15, 2015, pp. 2111–2117.

[10] Y. Shen, R. Jin, D. Dou, N. Chowdhury, J. Sun, B. Piniewski, and D. Kil, "Socialized gaussian process model for human behavior prediction in a health social network," in *Data Mining (ICDM), 2012 IEEE 12th International Conference on*. IEEE, 2012, pp. 1110–1115.

[11] N. Phan, D. Dou, B. Piniewski, and D. Kil, "Social restricted boltzmann machine: Human behavior prediction in health social networks," in *ASONAM '15*. ACM, 2015, pp. 424–431.

[12] N. Phan, D. Dou, H. Wang, D. Kil, and B. Piniewski, "Ontology-based deep learning for human behavior prediction in health social networks," in *BCB*. ACM, 2015, pp. 433–442.

[13] L. Deng and D. Yu, "Deep learning: Methods and applications, Tech. Rep. MSR-TR-2014-21, May 2014.

[14] S. Cao, W. Lu, and Q. Xu, "Grarep: Learning graph representations with global structural information," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, ser. CIKM '15, 2015, pp. 891–900.

[15] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei, "Line: Large-scale information network embedding," in *WWW '15*. International World Wide Web Conferences Steering Committee, 2015, pp. 1067–1077.

[16] P. Yanardag and S. Vishwanathan, "Deep graph kernels," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '15, 2015, pp. 1365–1374.

[17] A. Ahmed, N. Shervashidze, S. Narayanamurthy, V. Josifovski, and A. J. Smola, "Distributed large-scale natural graph factorization," in *WWW '13*. International World Wide Web Conferences Steering Committee, 2013, pp. 37–48.

[18] D. Wang, P. Cui, and W. Zhu, "Structural deep network embedding," in *KDD '16*. ACM, 2016.

[19] L. Tang and H. Liu, "Toward predicting collective behavior via social dimension extraction," *Intelligent Systems, IEEE*, vol. 25, no. 4, pp. 19–25, 2010.

[20] P. Rozenshtein, N. Tatti, and A. Gionis, "Discovering dynamic communities in interaction networks," in *Machine Learning and Knowledge Discovery in Databases*. Springer, 2014, pp. 678–693.

[21] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111–3119.

[22] Q. V. Le and T. Mikolov, "Distributed representations of sentences and documents," *arXiv preprint arXiv:1405.4053*, 2014.

[23] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, "Multimodal deep learning," in *Proceedings of the 28th international conference on machine learning (ICML-11)*, 2011, pp. 689–696.

[24] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.