

# A Propagation Model for Integrating Web of Things and Social Networks

Lina Yao

Supervised by: Michael Sheng

School of Computer Science

The University of Adelaide, Adelaide SA 5005, Australia

lina.yao@adelaide.edu.au

**Abstract.** Modeling interest of a user for services recommendation and friendship between users is the major activity of social networks. The information used by social networks such as user profiles is unfortunately easy to be faked and misled by the users, which often results in poor service recommendation and friendship prediction. In this paper, we propose a propagation model that integrates the emerging Web of Things (resource/services networks) and social networks together so that better service recommendation and friendship prediction can be achieved by considering interactions between people and things.

## 1 Introduction

Over the years, the Web has gone through many transformations, from traditional linking and sharing of computers and documents, to current connecting of people via numerous social networks such as Facebook and LinkedIn. The main goals of social networks are to effectively model the interest of a user and the friendship between users, which support a number of useful applications such as service recommendation based on user's interest and friendship prediction [14]. With the recent advances in radio-frequency identification (RFID) technology and Web services [11], the Web will continue the transformation and will be slowly evolving into the so-called "Web of Things and Services". This future Web will provide an environment where everyday physical objects such as buildings, sidewalks, and commodities are readable, recognizable, addressable, and even controllable using services via the Web.

Existing research on social networks are mainly based on mining i) the users' profiles, which are the meta-data from users when they register on the social networks, ii) comments and topics posted by the users, and iii) the interactions between users (e.g., news the users share and games the users play). All these sources are unfortunately easy to be faked and misled by the users. This often results in poor service recommendation and friendship prediction. With the emerging Web of Things, a new additional dimension becomes possible to improve service recommendation and friendship prediction by studying the relationship between things and people. Based on the social phenomenon of homophily [5], people who access similar resources tend to have similar interest and are more likely to be friends.

In this paper, we propose to connect Web of things and services with Web of people (i.e., social networks) together. This connection enables better service recommendation

and friendship and human behavior prediction based on analyzing interactions between people and things. Two major techniques are widely used in social networks analysis. Content-based recommender systems make recommendations by analyzing the content of textual information and finding regularities in the content, while collaborative filtering (CF) techniques use user-item ratings data to make predictions and recommendations [12]. Both approaches have limitations. CF systems do not explicitly incorporate feature information and content-based systems do not necessarily incorporate the information in preference similarity across individuals. Hybrid CF techniques (e.g., the content-boosted CF algorithm [6] and Personality Diagnosis (PD) [9]), which combine CF and content-based techniques, hold the hope to avoid the limitations and thereby improve recommendation performance. In this paper, we propose to use a hybrid methodology that combines collaborative filtering techniques with knowledge-based recommending techniques, to model the interactions between people and things.

This paper discusses the research problem and significance (Section 2), our proposed model to study three relationships among people and things, as well as some future work ahead (Section 3).

## 2 Problem Description and Research Significance

In a Web of Things environment, physical things connect to the Internet offering resources and services. Social networks are designed to connect people over the Internet where people can interact with each other, share resources and create their own community. We propose to establish an integrated network that links social networks and Web of things together. In such a network, users not only find their interested people, but also their interested services and resources provided by the things. Users' interest and friendship based on social networks can propagate to the Web of Things.

There are two kinds of entities considered in this work: *people* and *things*, which generate three kinds of interactions, namely *people to people*, *people to things*, and *things to things*. Our goal is to propose an integrated framework that can better predict and recommend the most appropriate services and things to users. In particular, we will analyze two behaviours: *user behaviour* and *thing behaviour*. User behaviour can be learned from the people to people network (e.g., friendships in social networks) and people to things network (e.g., interactions between users and things), while thing behaviour can be extracted from the people to things network and things to things network (i.e., linkage among things).

Unfortunately, the information needed to deduce user or thing behaviors are usually noisy, in large volume, and even missing. Inferring missing elements (also called latent variables) from observable elements is a challenging task. The key task in our work is to associate the latent variables with both users and things, to define coupled models that encode relationships between people and things. In particular, we will define a shared latent variable to assure dynamical and evolutionary interaction between Web of things and social networks during the learning process.

## 2.1 Research Significance

Existing researches in social networks focus on analyzing users' interests by profiling their information from the users' profile meta-data or inducing users' interests by digging their behaviors among the social networks. Recently, researchers have begun to consider the effects from user-user friendships, which have the impact on services recommendation. For example, user-user friendships are taken into account to recommend the most appropriate services, e.g., advertisements and news items [14]. However, very few work involve things in the services recommendation applications.

Existing recommendation systems work well in the traditional services-user environment, like collaborative filtering (CF) techniques. However, the precondition of such systems is that users are independent. In our model, users' interactions are an important factor and they are correlated other than propositional entities, i.e., the users' interactions can affect the decision making process of the recommendation results. In addition, things can be composed to complete a certain service. Things can come and go, and are very uncertain in terms of service offering. This will result in well known *cold-start* problem in recommendation systems. Cold-start problems occur when a new person or new thing comes into the network, there is no or little initial information available for them [13].

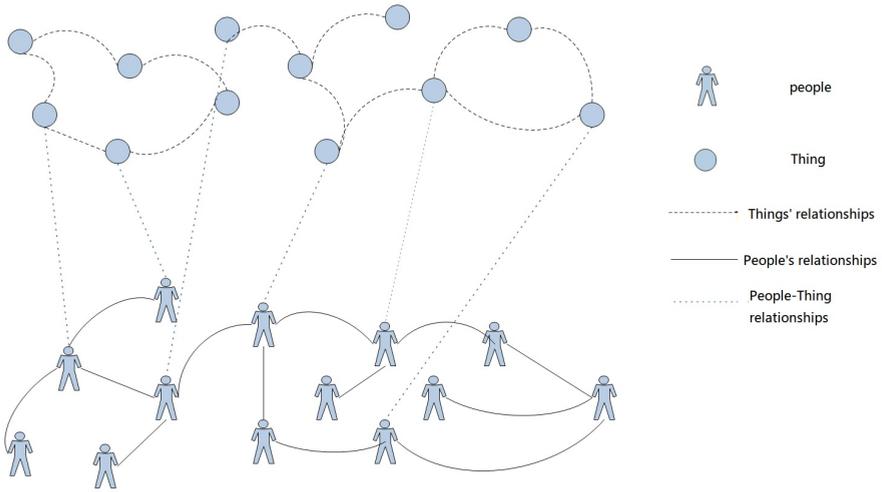
The key idea of our work is to associate the latent variables with both people and things and encode information about the people to people friendship, people to thing interactions and thing to thing linkage. With our result, it is possible to further perform mining—based on the interactions and relationship between people and things—to predict human activities and behaviors. The relationships can be understood as (i) behavior of user by using things, (ii) interactions between users and things (e.g., where the user shows up and her location information), and (iii) correlations among things.

## 3 Proposed Solution

The nontrivial correlation between Web of things and social networks motivates joint modeling of both sources of evidence. The thing to people propagation model simultaneously encodes the three heterogeneous types of dyadic relationships, namely *people to things* ( $\{\exists y_{i,j}, \forall i \in \mathcal{P}, j \in \mathcal{S}\}$ ), *people to people* ( $\{\exists r_{i,i'}, \forall i, i' \in \mathcal{P}\}$ ), and *things to things* ( $\{\exists t_{i,i'}, \forall i, i' \in \mathcal{S}\}$ ), where  $\mathcal{P}$  and  $\mathcal{S}$  denote people and things respectively. We consider all three relationships as directed graph, and model them based on the latent variable graphical model.

### 3.1 Modeling People to Things

This can be considered as a recommendation system and its aim is to offer a user  $p_i$  the most appropriate things/services  $s_j$ . We propose to use the collaborative filtering (CF) techniques in the modeling, which is based on learning the past interactions between people and things. In particular, neighborhood-based collaborative filtering is one of prevalent memory-based CF that assumes the interactions between user  $p_i$  and a thing  $s_j$  depends on the observations of neighbouring users or things [10]. Similar things or services can be propagated to a particular user [14].



**Fig. 1.** Integrating Web of things and social networks: all the connections include unipartite edges within the people to people network, bipartite user to things and unipartite things to things

We assume that for each person  $p_i$  and each thing or service  $s_j$ , their observable variables are  $x_i$ , which includes the person’s registration information and profile etc, and  $x_j$ , which is the thing’s description respectively. Moreover, the hidden variable  $z_i$  refers to the person’s interests that are not observed directly and  $z_j$  refers to the thing’s semantic information or topics. So  $p(y_{i,j})$  depends on the latent and observable variables which is denoted by  $p(y_{i,j}|i, j) \sim p(y_{i,j}|x_i, x_j, z_i, z_j; \Theta)$ , where  $\Theta$  is a collection of hyper parameters. It should be noted that we must consider the cold-start problem since new things keep adding easily to the Internet while old things may disappear. We will refer to [1] to deal with the cold-start problem.

We define  $x_i$  and  $x_j$  are the observable variables for the people and things based on its defined dependency, their relationship will be:  $z_i \sim p(z_i|x_i)$ ,  $z_j \sim p(z_j|x_j)$ , and  $z_i$  and  $z_j$  use their inner product  $z_i^T z_j$ . By coupling the neighbourhood dependency with the latent variable model together based on [2,3], we have:

$$\hat{z}_i = \frac{\sum_{i' \in \Gamma_i} w_{i,i'} z_{i'}}{\sum_{i' \in \Gamma_i} w_{i,i'}}$$

and  $p(y_{i,j}|i, j) \sim p(y_{i,j}|x_i, x_j, \hat{z}_i z_j; \Theta)$ , where  $w_{i,i'}$  is measure the similarity between people  $i$  and its neighbour  $i' \in \Gamma_i$ .

### 3.2 Modeling People to People

The people to people relationship is modeled using the Random Walk on graph. A random walk is a finite Markov chain that is time reversible [4], which starts at a node (each person can be considered as a node)  $i$  and iteratively moves to a neighbor of  $i$  chosen uniformly at random. The hitting time  $h_{i,i'}$  from  $i$  to  $i'$  is the expected number

of steps required for a random walk starting at  $i$  to reach  $i'$ . In our model, we define the graph to be a undirected graph, so that the commute time  $c_{i,i'} := h_{i,i'} + h_{i',i}$ . We normalize the hitting time, so these measures can be considered as proximity (similarity) measure  $r_{i,i'} := h_{i,i'} \cdot \pi_{i'}$  for each pair of people  $(i, i')$ . Moreover, in our model, the transition probability  $p(i, i') = r_{i,i'}/d_i$  and  $p(j, j') = t_{j,j'}/d_j$ , where  $d_i$  and  $d_j$  are the degree of  $i$  and  $j$  respectively. Based on the same assumption [14], assuming user  $i$  is fully characterized by observable features  $x_i$  and latent variable  $z_i$ , we have  $r_{i,i'} \sim p(r_{i,i'} | x_i, x_{i'}, z_i, z_{i'}; \Theta)$ .

### 3.3 Modeling Things to Things

We consider the things to things relationship as a graph. The similarity between things is measured based on the non-functionalities of things. We construct the similarity graphs using Preferential attachment to model thing to thing relationship proximity. Its basic idea is the probability that a new connection involves node (thing)  $j$  is proportional to  $|\Gamma(j)|$ , the current number of neighbors of  $j$  [7], so the  $s_{j,j'} := |\Gamma(j)| \cdot |\Gamma(j')|$  [8].

For the things  $s_j$  and  $s'_{j'}$ , we connect  $s_j$  and  $s'_{j'}$  if  $s'_{j'}$  is among the  $k$ -nearest neighbors of  $s_j$ . We further measure their pairwise similarities  $t_{j,j'} = t\{s_j, s'_{j'}\}$  by the similarity function we defined which is symmetric and non-negative, and so that the corresponding similarity matrix  $\mathcal{Q} = t_{j,j'}$ , where  $j, j' \in \{1, n\}$ . We also consider each thing depends on its observable factors  $x_j$  and latent factors  $z_j$ . Finally, we have  $t_{j,j'} \sim p(t_{j,j'} | x_j, x_{j'}, z_j, z_{j'}; \Theta)$ .

Our model integrates the people-people relationship model  $(i, i', r_{i,i'})$ , people-thing relationship model  $(i, j, y_{i,j})$  and thing-thing relationship model  $(j, j', t_{j,j'})$  together and there are two tuples of sharing factors, namely  $(i, x_i, z_i)$  and  $(j, x_j, z_j)$ . Our model should learn the shared parameters and latent variables from integrated relationships. In general, latent variable graphical models have a significant difficulty for model selection because one may not know the number of relevant latent variables, nor the relationship between these variables and the observed variables. Convex formulations require the manipulation of a full matrix which is impractical for anything beyond thousands of participants in our model. In addition, relationships between users and things change over time, which calls for a learning algorithm that are efficient and amendable to dynamic data streams.

We will build up a testbed to verify and test our latent graphical model, based on popular social networks such as Facebook and LinkedIn, as well as things network connected by RFID and sensors. We will also conduct some comparison experiments to demonstrate that our model can produce more accurate and efficient recommendations.

## 4 Conclusion

The emerging Web of things and services provides an exciting opportunity to improve the poor performance of services recommendation and friendship prediction in today's social networks. This PhD work focuses on investigating ways to efficiently integrate social networks and Web of things together by considering people to people relationships, things to things relationships, as well as people to things relationships. In this

paper, we have presented some initial ideas of this work, as well as some preliminary design. Currently, we are constructing the integrated model. Our future work also includes verifying the proposed model using real social networks.

## References

1. Agarwal, D., Chen, B.: Regression-based latent factor models. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 19–28. ACM (2009)
2. Koren, Y.: Factorization meets the neighborhood: A multifaceted collaborative filtering model. In: Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 426–434. ACM (2008)
3. Koren, Y.: Factor in the neighbors: Scalable and accurate collaborative filtering. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 4(1), 1 (2010)
4. Lovász, L.: Random walks on graphs: A survey. *Combinatorics, Paul Erdos is Eighty* 2(1), 1–46 (1993)
5. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27(1), 415–444 (2001)
6. Melville, P., Mooney, R., Nagarajan, R.: Content-boosted collaborative filtering for improved recommendations. In: Proceedings of the National Conference on Artificial Intelligence, pp. 187–192. MIT Press (2002)
7. Mitzenmacher, M.: A brief history of generative models for power law and lognormal distributions. *Internet Mathematics* 1(2), 226–251 (2004)
8. Newman, M.: Clustering and preferential attachment in growing networks. *Physical Review E* 64(2), 025102 (2001)
9. Pavlov, D., Pennock, D.: A maximum entropy approach to collaborative filtering in dynamic, sparse, high-dimensional domains. In: *Advances in Neural Information Processing Systems*, pp. 1465–1472. MIT Press (2002)
10. Sarwar, B., Karypis, G., Konstan, J., Reidl, J.: Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th International Conference on World Wide Web (WWW 2010), pp. 285–295. ACM (2001)
11. Sheng, Q.Z., Zeadally, S., Luo, Z., Jung, J.-Y., Maamar, Z.: Ubiquitous rfid: Where are we? *Information Systems Frontiers* 12(5), 485–490 (2010)
12. Si, L., Jin, R.: Flexible mixture model for collaborative filtering. In: Proceedings of the 20th International Conference on Machine Learning (ICML 2003), vol. 20, p. 704 (2003)
13. Su, X., Khoshgoftaar, T.M.: A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, 4:2 (January 2009)
14. Yang, S., Long, B., Smola, A., Sadagopan, N., Zheng, Z., Zha, H.: Like like alike: Joint friendship and interest propagation in social networks. In: Proceedings of the 20th International Conference on World Wide Web, pp. 537–546. ACM (2011)