

Fine-grained Social Relationship Extraction from Real Activity Data under Coarse Supervision

Kota Tsubouchi
Yahoo! JAPAN Research
Tokyo, Japan
ktsubouc@yahoo-corp.jp

Osamu Saisho
The University of Tokyo
Tokyo, Japan
saisho@ics.t.u-tokyo.ac.jp

Junichi Sato
The University of Tokyo
Tokyo, Japan
jsato@ics.t.u-tokyo.ac.jp

Seira Araki
The University of Tokyo
Tokyo, Japan
araki@ics.t.u-tokyo.ac.jp

Masamichi Shimosaka
Tokyo Institute of Technology
Tokyo, Japan
simosaka@miubiq.cs.titech.ac.jp

ABSTRACT

Understanding social relationships plays an important role in smooth information sharing and project management. Recently, extracting social relationships from activity sensor data has gained popularity, and many researchers have tried to detect close relationship pairs based on the similarities between activity sensor data, namely, unsupervised approaches. However, there is room for further research into social relationship analysis of sensor data in terms of extraction performance. We therefore focus on improving the accuracy of detection and propose a novel fine-grained social relationship extraction from coarse supervision data by supervised approach based on multiple instance learning. In this paper, *fine-grained* relationship means the relationship including information about the time and duration they are together, and *coarse supervision data* is the data containing only information about whether they are together in a day. In this research, we evaluate the feasibility of our extraction method and analyze the extracted fine-grained social relationships. Our approach improve detection accuracy and achieve extraction of fine-grained relationships from coarse supervision data.

Author Keywords

Working Relationship, Detection of People Close to Each Other, Coarse Supervision Data, Activity Sensor

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Extracting social relationships supplies subtle and useful information to organizations, so understanding social relationships in companies could improve companies' performance.

For the last decades, many researchers have investigated how to extract social networks from people's connections. Recent improvement in sensing and wearable computing technologies enables a new level of social relationship extraction researches for real world or offline social relationship analysis [6]. In this research, we target only "offline social relationships". These relationships can be extracted by various approaches. Approaches with GPS and Wi-Fi location data [4] and proximity sensors such as RFID [8], infrared [3] and Bluetooth [9] assume that closer proximity of two users shows a stronger relationship between them. Approaches with activity sensors assume that stronger similarity of the activities of two people shows a stronger relationship between them. Tsubouchi et al. [10] proposed a method to estimate social relationships from the correlation of the number of steps.

The existing unsupervised approaches with sensor data have been more practical these days because the activity sensor data logs of smartphones or wearable sensors are becoming more accessible. However, these unsupervised approaches cannot estimate precisely in ordinary situations. For example, two people belonging to the same organization that are often near each other are estimated to get along well because their proximity is strong even though their relationship is actually weak. For more precise estimation, we should divide the data in the situation and calculate the correlation with different weights. For example, lunches in the daytime and drinking parties after work are important situations to estimate relationships because interaction outside offices reflects closer relationships than interaction in offices.

In this research, we extract precise social relationships by a supervised approach. Our supervised approach does not require complicated and troublesome questionnaires with questions such as "Did you eat lunch with Mr. X yesterday? If so, state the time and duration." Such troublesome questionnaires are unreliable and unfeasible. By considering the feasibility of questions, we propose a novel approach to detect specific social relationships from much simpler questions such as "Were you together with Mr. X yesterday?"

Our contributions are summarized as follows. First, we propose a novel supervised approach to extract a working rela-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
ISWC '15, September 7–11 2015, Osaka, Japan.

Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3578-2/15/09...\$15.00.
<http://dx.doi.org/10.1145/2802083.2808402>

tionship and show that the result is more precise than the result of the existing unsupervised approach. Second, with our proposed approach, the time and duration they are together are also estimated from coarse supervised data that do not include the time and duration. The proposed method can also be applied to other sensor data like data collected with Bluetooth in existing researches though we focus on the number of steps in this research.

RELATED WORK

Many researchers analyze social relationships by unsupervised approaches from easily accessible sensor data. Cranshaw et al. clarify social relationships from GPS and Wi-Fi location data of the users' smartphones and personal computers [4]. To focus on the distance between users, Bluetooth [9] and radio frequency identifier (RFID) [8] are often used. Tsubouchi et al. [10] have proposed drawing social relationships from the similarity of the number of steps of users. In these unsupervised approaches, we can obtain working relationships at any resolution in accordance with our objectives. For example, by extracting the morning data from all data, researchers can acquire the working relationship in the morning easily. On the other hand, unsupervised approaches have drawbacks in extraction performance. For example, two people of the same organization are often near each other, and their GPS and Wi-Fi data have lots of similarities even if their working relationship is weak. It is impossible to eliminate such cases by unsupervised approach. The result of working relationship analysis by the number of steps is totally different due to the setting of the calculation parameters.

For the totally supervised approach, concrete questions such as "when, with whom, and for what activity were you together?" are required. This is because the precise relationship in the supervised model can only estimate relationships that are asked for. If we want to know the relationship in the morning, we have to ask them about their morning activity. This complicates the questionnaire and is not feasible because of two problems: answer reliability and troublesomeness. With regard to answer reliability, most people cannot remember the accurate time and duration of activity such as "I ate lunch with Mr. X from 12:30 to 13:12", so the answer is not reliable. With regard to troublesomeness, it is really troublesome to collect and analyze the answer data from the questionnaires. The feasible supervised methods use reliable questionnaire data and questions that are easy to ask such as "were you together with Mr. A yesterday?". Such simple questions are reliable and easy to analyze, but there is no detailed information such as the time and duration they were together.

In this work, we use a multiple instance learning (MIL) [5] framework to estimate the time and duration of being together with the coarse questionnaire result. MIL is a well known supervised learning algorithm in vision research. MIL can produce a cognitive model from the coarse training data by labeled bags and instances. For example, a model that judge "whether the horse is in the target frame" is learned by coarse training data such as "whether this video has the horse in any frames." Each frame is an instance, and the video is a bag.

OUR FRAMEWORK

Fine-grained social relationship extraction

In this research, we extract fine-grained social relationships from the similarity between the activity data of two people. All social relationships are not equal, for instance, the relationship among co-workers on the job and that among co-workers drinking after work will be different. When considering the quality of such relationships, the time and duration they are together is an important indicator. In this paper, we focus on *fine-grained* social relationships, that is, social relationships including detailed information about the time and duration they are together. To extract the time and duration two people are together, we applied the scan window method to time-series activity data of two people. Figure 1 shows the overview of extraction. x_i and x_j in Figure 1 are the

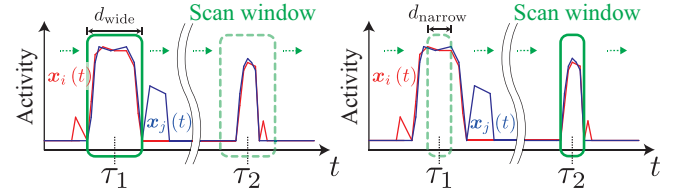


Figure 1: Overview of fine-grained relationship extraction from activity data. Proposed method enables detection of being together whenever and for however long.

time-series activity data of one user and the other user, respectively. We prepare time windows with various window widths and scan the data with the windows. Then, we extract features from the data for each time τ with each window width d and estimate their being together for each time. Figure 1 shows an example of detection with two time windows whose widths are d_{wide} and d_{narrow} respectively. In the left graph in Figure 1, being together for a long time is detected with the wide time window at τ_1 ; on the other hand, the narrow window detects being together for a brief time at τ_2 in the right graph. Thus, whenever and for however long they are together, we can detect this with this method and extract fine-grained social relationships.

Learning model under coarse supervision

Supervised approaches can be divided into two types: learning the model under dense supervision or coarse supervision. In this paper, we define the difference between the two types as presence or absence of the information about the time they are together. Considering the reliability of the answers and counting them, we ask each user "were you together with Mr. X yesterday?" in the coarse questionnaire. The training data acquired from the questionnaire, therefore, is coarse supervision, that is, whether Mr. X and Mr. Y were together for even a moment in a day. Such coarse data shows that they were together at *certain times* of the day; however, we cannot know the time and duration they were together from such coarse data directly.

Considering such characteristics of the data, this research utilizes the MIL [5] framework. Each instance is a feature vector extracted from steps-per-minute data of two people at a certain time of a day in a certain time window. That is, the feature vector ϕ is represented as $\phi(x_i, x_j, d, \tau)$ in Figure 1.

Multiple factors can be included in the feature, such as beginning or end of activity and similarity of activity. A bag, which is a set of instances, is a set of all feature vectors extracted from step data of two people in a certain day. With coarse training data obtained from questionnaires, we cannot know the label of each instance; however, we can get the correct label of each bag. MIL is suited for such coarse supervision because the model can be learned only with the bag label in the MIL framework.

In this paper, we leverage mi-SVM [2] for learning. mi-SVM is one of the extensions of support vector machine (SVM) learning algorithm for MIL proposed by Andrews et al. The mi-SVM algorithm can estimate not only the label of each bag but also that of each instance, so we can detect when and for how long two people were together from the coarse training data. Thus, we can extract fine-grained social relationships. In the prediction step, we detect being together by thresholding of the inner product between a feature vector $\phi(\mathbf{x})$ and the weight vector \mathbf{w} which is learnt with mi-SVM:

$$\begin{cases} \text{Being together} & \text{if } \mathbf{w}^T \phi(\mathbf{x}) > \theta \\ \text{Not being together} & \text{otherwise} \end{cases}$$

where θ denotes a threshold value.

EXPERIMENT

In this section, we evaluate the detection accuracy of the proposed coarse supervised method. Then, we compare the detection accuracy of our method with that of the unsupervised method. In addition, this section shows that the proposed method can estimate the time and duration the subjects were together, which is too detailed a question to ask in everyday questionnaires.

Experimental method

Table 1 shows details of participants, devices for obtaining data, and experimental period. All participants wore the Fit-

Table 1: Overview of experiment.

Device	Fitbit One[1]
Period	Weekdays from 2013/6/26 to 2013/7/10
# of participants	31
# of organizations	5
Age range	21-46 yrs.

bit One, which is a well known commercial wearable activity monitor, for about two weeks and answered the question “were you together with Mr. X yesterday?” every day. In addition, we asked participants “did you eat lunch / dinner with Mr. X yesterday?” as reference data for the evaluation of detection accuracy, that is, we didn’t use these data for learning and prediction. This question considers the fact that most people cannot remember the accurate time of an activity. In fact, although the question is easy to answer, the answers conflicted with each other in some pairs. These cases show the difficulty in acquiring correct training data. Moreover, this fact indicates that it is totally impractical to obtain more detailed (e.g., hourly) reference data even in a short-term experiment. In this paper, “a pair of participants who were together”

is defined as a pair of participants both of whom answered they were together with each other in the questionnaire.

Parameter settings

We set up seven types of scan windows, namely, a window width d_k of $\{5, 10, 30, 60, 90, 120, 180[\text{mins.}]\}$. The dimension of steps-per-minute vector \mathbf{x}_i and \mathbf{x}_j in a day shown in Figure 1 is $60 [\text{mins.}] \times 24 [\text{hrs.}] = 1440$. Thus, variable time τ ranges from $[d_k/2, 1439 - (d_k/2)]$ when the width of a scan window is d_k . Therefore, the number of feature vectors generated from steps-per-minute data of two people in a day, which means the number of instances in a bag, is $\sum_{k=1}^7 (1440 - d_k)$.

Feature

We generated five types of features from steps-per-minute data of two users. One type of features is the same utilized by Tsubouchi et al. [10]. The other four types are similarity of timing of beginning / end of walking, similarity of change rate of the number of steps, constancy of difference between the number of steps of users, and activity level. To demonstrate the effectiveness of the proposed method clearly, we calculate the performance of our supervised method using only the first of these features in comparison with the unsupervised method of Tsubouchi et al. [10]. Since our method automatically optimizes parameters, this promotes us to add various features such as the ones above to the model.

Result

A total of 283 pairs of participants answered they were together, and 4481 pairs answered they were not. It should be noted that, if a pair answered they were together more than once in a day, we treated them as more than one pair. First, we evaluated the detection accuracy of proposed approach. Figure 2 shows the precision-recall curves of the proposed supervised method and the conventional unsupervised method [10]. 200 points are shown in each curve in accordance with the linearly increasing threshold value, which determines whether each of all pairs was together. For example, if the threshold value is negative infinity, all pairs will be judged being together. For a fair comparison, we also plot

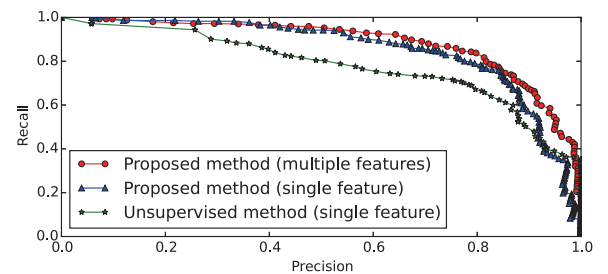


Figure 2: Precision-recall curves of proposed supervised method and conventional unsupervised method.

the precision-recall curve of the proposed supervised method using only the same feature as Tsubouchi et al. [10] in Figure 2. The area under the precision-recall curve (AUC, which is commonly used to evaluate the performance of binary classifier) of the proposed supervised method with multiple features, that with a single feature, and that of the conventional unsupervised method with a single feature were 0.88, 0.86,

and 0.78, respectively. The result that the AUC of the blue curve is larger than that of the green curve shows that our coarse supervised method outperforms the conventional unsupervised method. Furthermore, the high performance of the red curve indicates that we can detect being together more precisely with various additional features, which can be easily added only in the proposed method.

Additionally, we extracted social relationships in the daytime (11:00–14:00) and in the evening (17:00–24:00) for understanding fine-grained social relationships. We can extract social relationships at *any* time interval with the proposed method; however, we selected the time described above due to the difficulty of obtaining reference data other than that of lunch or dinner with a questionnaire. The beginning and end time of the daytime and the evening are designed to include all timings of lunch and dinner of the reference data. Figure 3 shows graphs of the social relationships generated from reference questionnaire or extracted from steps-per-minute data. Each node shows a participant, and two nodes linked with

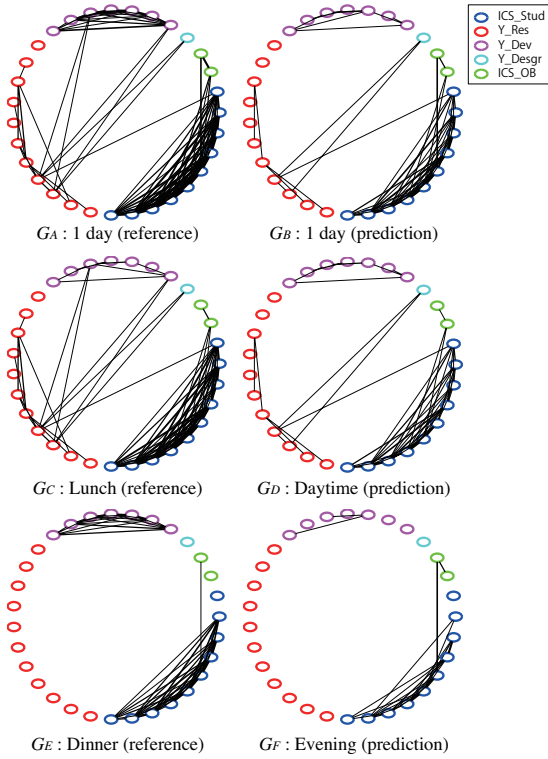


Figure 3: Graphs of social relationships generated from reference questionnaire (the three graphs in left side) or extracted from steps-per-minute data (the others in right side).

each other show that the two participants answered they were together. Also, the organizations that they belonged to were classified by different colors. G_A in Figure 3 shows the social relationships generated from coarse training data (i.e., answers to “were you together with Mr. X yesterday?”) and G_C and G_E show those generated from the fine-grained reference data (i.e., answers to “did you eat lunch / dinner with Mr. X yesterday?”). G_B , G_D and G_F show the relationships extracted from steps-per-minute data with the proposed method

in 1 day (00:00–24:00), in the daytime (11:00–14:00) and in the evening (17:00–24:00), respectively.

For quantitative evaluation, the similarity score of G_X and G_Y , which are the graphs of social relationships whose representation matrices are X and Y respectively, is formulated as $\frac{\text{tr}(XY)}{\|X\|_{\text{fro}}\|Y\|_{\text{fro}}}$ where $\text{tr}(\cdot)$ denotes the trace of the matrix and $\|\cdot\|_{\text{fro}}$ denotes the Frobenius norm. This score is generally used as a matrices’ correlation score [7]. X and Y are symmetric because the graph of social relationships is undirected graph. G_C and G_E in Figure 3 indicate that social relationship changes drastically depending on the time of day. From a quantitative perspective, these two graphs are far from analogous because the similarity score of G_C and G_E is 0.65. Then, we demonstrate that the proposed method is able to provide such social relationships which vary with the time of day. G_D and G_F in Figure 3 show the fine-grained social relationships extracted from steps-per-minute data with the proposed coarse supervised method. The organizations labeled “ICS_***” are friendly laboratories of a university. The organizations labeled “Y_***” are groups of a company.

Table 2 shows the calculated similarity scores. Table 2 reveals that G_C and G_E is analogous with G_D and G_F respectively. Therefore, it is clear that our method can extract fine-grained social relationships precisely. That is, the social relation-

Table 2: Similarity scores.

		prediction	
		G_D : Daytime	G_F : Evening
reference	G_C : Lunch	0.82	0.55
	G_E : Dinner	0.60	0.75

ships extracted from step data and the coarse questionnaire data are qualitatively and quantitatively similar to the relationships generated from the detailed reference questionnaire data. Thus, we can extract fine-grained social relationships only from activity data and coarse questionnaire data.

CONCLUSION

In this research, we presented a novel approach to extract fine-grained working relationships from coarse supervised data utilizing MIL. We compared the accuracy of our supervised method with the conventional unsupervised method. Moreover, we also compared the human relationships extracted from the activity sensor data with those from the questionnaire. The experimental results show that our approach with coarse supervision data achieves more precise detection than the existing unsupervised approach and extracts working relationships with detailed information such as time when they are together even if such detailed information is not available from questionnaires.

These contributions make it more easy and precise to extract a working relationship from activity sensor data. Our framework improves accuracy and reduces the load of questionnaires by asking a simple question once a day. However, the framework should need only supervision data that are coarser i.e. asking less than once a day and asking only a part of the members. Our future work is improving the trade-off between the coarseness of data and the accuracy of the detection.

REFERENCES

1. Fitbit one. <http://www.fitbit.com/one>.
2. Andrews, S., Tsochantaridis, I., and Hofmann, T. Support vector machines for multiple-instance learning. In *Adv. NIPS 2002* (2002), 561–568.
3. Choudhury, T., and Pentland, A. The sociometer: A wearable device for understanding human networks. In *CSCW 2002* (2002).
4. Cranshaw, J., Toch, E., Hong, J., Kittur, A., and Sadeh, N. Bridging the gap between physical location and online social networks. In *Proc. of UbiComp 2010* (2010), 119–128.
5. Dietterich, T., Lathrop, R., and Lozano-Pérez, T. Solving the multiple instance problem with axis-parallel rectangles. *Artificial Intelligence* 89, 1 (1997), 31–71.
6. Eagle, N., Pentland, A. S., and Lazer, D. Inferring friendship network structure by using mobile phone data. In *Proc. of NAS 2009* (2009), 15274–15278.
7. Herdin, M., Czink, N., Ozelik, H., and Bonek, E. Correlation matrix distance, a meaningful measure for evaluation of non-stationary mimo channels. In *VTC 2005-Spring*, vol. 1 (2005), 136–140.
8. Konomi, S., Inoue, S., Kobayashi, T., Tsuchida, M., and Kitsuregawa, M. Supporting colocated interactions using rfid and social network displays. *Pervasive Computing, IEEE* 5, 3 (2006), 48–56.
9. Staiano, J., Lepri, B., Aharony, N., Pianesi, F., Sebe, N., and Pentland, A. Friends don't lie: Inferring personality traits from social network structure. In *Proc. of UbiComp 2012* (2012), 321–330.
10. Tsubouchi, K., Kawajiri, R., and Shimosaka, M. Working-relationship detection from fitbit sensor data. In *Proc. of UbiComp 2013* (2013), 115–118.