Intercommunity Detection Scheme for Social Internet of Things: Compressive Sensing Over Graphs Approach

Tengjiao Wang, Jingbo Tan, Wenbo Ding, *Member, IEEE*, Yanru Zhang, *Member, IEEE*, Fang Yang, *Senior Member, IEEE*, Jian Song, *Fellow, IEEE*, and Zhu Han, *Fellow, IEEE*

***Abstract*—As a cluster between Internet of Things (IoT) and mobile social networks (MSNs), social IoT allows close interaction between humans and things. Many applications of IoT and MSN are based on the community structure to realize efficient services, and thus the detection of the community structure has become a key problem and attracted much academic attention. In this paper, a compressive sensing (CS) over graphs-based intercommunity detection scheme is proposed for the social IoT. By exploiting the probability of two nodes encountering each other, the encounter probability and the edge-clustering coefficient are utilized to define a novel metric for each connection and construct the measurement matrix. Then a CS-based detection algorithm is proposed to detect the intercommunity links. Moreover, the edge-clustering coefficient is exploited as the prior information to further improve accuracy and reduce complexity. Simulations show that the proposed detection scheme outperforms the conventional random walk-based scheme in the social IoT.**

***Index Terms*—Compressive sensing (CS) over graphs, edge clustering coefficient, Internet of Things (IoT), mobile social networks (MSNs), prior information, smart cities.**

I. Introduction

INTERNET of Things (IoT) is a recent communication paradigm [1], which connects a large variety of physical devices and provide data services among them [2], [3]. The mobile social network (MSN) is one emerging form of social network, which utilizes the social relationship information to improve efficiency and effectiveness of network services. MSN provides a variety of data delivery services involving the social relationship among users [4], [5]. MSN facilitates many social network-aware applications, such as social network-aware routing [6], [7], context-aware information diffusion [8], [9], and role determination [10], [11].

Recently, with more and more humans utilizing smart devices and services every day to improve their life quality, the concept of the social IoT has been introduced to combine IoT with MSNs to meet the increasing demands [12]. As the cluster of IoT and MSN, the social IoT considers not only thing-to-thing connection and human-to-human connection, but also the close interaction between things and humans. Therefore, the social IoT has been considered as one of the trends for IoT and building blocks for smart cities [13].

In the social IoT, many applications and services are based on communities. The community-based structure helps the data-centric services [3], the community-level services [14], [15], and the trust management services [16], [17] in IoT. Furthermore, many data forwarding algorithms are also based on communities, including Simbet [18], Bubble Rap [19] and so on. Therefore, the detection of intercommunity links in the social IoT deserves further research beyond its first investigation in the field of conventional social networks. Girvan and Newman (GN) introduced a divisive algorithm in which the edges are selected and removed one-by-one based on the value of their "edge betweenness," calculated by the shortest path tree [20]. The GN algorithm represents a significant step forward on the community detection in networks. However, the order of the removal process reveals the community structure of a network. It therefore suffers from high complexity for the large-scale social IoT [21]. Then, Radicchi *et al.* [22] proposed a method based on the edge-clustering coefficient. In this method, they defined the edge-clustering coefficient based on the triangle or quadrilateral links, and used it to achieve higher detection speed and lower complexity [22]. From then on, the edge-clustering coefficient has been used in many other community detection schemes to enhance the performance, such as the community detection in complex networks [23] and in protein-protein interaction networks [24].

As a branch of compressive sensing (CS) [25]–[27], CS over graphs, which takes advantage of the sparsity over the

Manuscript received January 15, 2018; revised April 2, 2018; accepted May 10, 2018. Date of publication May 15, 2018; date of current version January 16, 2019. This work was supported by the National Key Research and Development Program of China under Grant 2017YFB0403402. *(Corresponding author: Fang Yang.)*

T. Wang, J. Tan, F. Yang, and J. Song are with the Department of Electronic Engineering, Research Institute of Information Technology, Tsinghua University, Beijing 100084, China, and also with the Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China (e-mail: wangtj16@mails.tsinghua.edu.cn; tanjb17@mails.tsinghua.edu.cn; fangyang@tsinghua.edu.cn; jsong@tsinghua.edu.cn).

W. Ding is with the School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA 30318 USA (e-mail: vincentdwb@gmail.com).

Y. Zhang is with the Department of Information Engineering, Chinese University of Hong Kong, Hong Kong (e-mail: yrzhang@ie.cuhk.edu.hk).

Z. Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 17104, South Korea (e-mail: zhan2@uh.edu).

Digital Object Identifier 10.1109/JIOT.2018.2837048

2327-4662 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

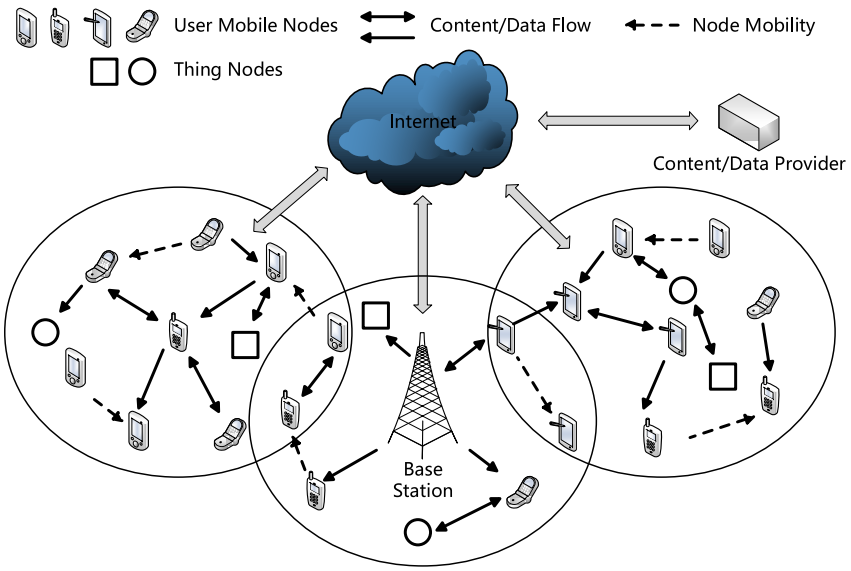


Fig. 1. Basic architecture of the social IoT.

network, has provided another approach in community detection [28], [29]. In the standard CS over graphs model, the sparse value should be additive, e.g., the network latency or the logarithm of the packet loss probability. However, under many conditions in online social networks or MSN, it is difficult to find such additive values. The CS approach for intercommunity detection (CS-ComDet) [30] constructs an additive value via the edge-clustering coefficient to resolve this problem in the traditional social networks and achieves good performance. However, to our best knowledge, there is no relevant work in the social IoT and the direct application of CS over graphs-based methods cannot achieve ideal performance. Some inherent characteristics in social IoT motivate us to find more efficient approaches.

In this paper, the intercommunity detection issue is investigated in the social IoT scenarios. First, the edge-clustering coefficient and the encounter probability in the social IoT are investigated to provide information of the whole network from different aspects. Second, a novel metric is proposed based on the edge-clustering coefficient and the encounter probability to better construct the measurement matrix. Third, an efficient three-step detection method is proposed based on the CS over graphs theory to detect the community structure. Last but not least, the edge-clustering coefficient is exploited as the prior information to reduce complexity of the conventional CS algorithms and improve the recovery performance as well. Simulations are carried out on real MSN traces and the results demonstrate that the proposed method could achieve higher accuracy with fewer measurements compared to the random walk-based method. After the detection, the community structure can be derived to benefit various applications and services in IoT [14]–[17].

The rest of this paper is organized as follows. In Section II, the system model, the traditional random walk method based on CS over graphs and the edge-clustering coefficient are explained, respectively. In Section III, the proposed intercommunity detection method and the necessary analysis is addressed. The numerical simulation results are provided in Section IV. Finally, the conclusions are drawn in Section V.

Notation: Column vectors and matrices are denoted by lowercase and uppercase boldface letters; $(\cdot)^{\mathrm{H}}$ and $(\cdot)^{\dagger}$ denote the conjugate transposition and pseudo-inversion, respectively; $|\Phi|$ denotes the cardinality of set Φ; $\mathbf{v}_i$ represents the ith element of vector $\mathbf{v}$; and $\mathbf{M}_{(i,j)}$ represents the element in the ith row, and jth column of matrix $\mathbf{M}$.

II. System Model

In this section, the traditional random walk-based CS over graphs, the encounter probability, and the edge-clustering coefficient will be investigated.

A. Compressive Sensing Over Graphs Detection

In the social IoT, the social network of all the entities including humans and things are taken into consideration for the evolution beyond IoT [31], where the basic architecture of the social IoT is shown in Fig. 1. The human mobile nodes and the thing nodes converse and connect with each other, while they are connected to the Internet and the content provider is used to provide extra data and content. In order to meet a variety of communication requirements, the network architectures can be centralized networks or opportunistic networks.

Generally, the densely connected groups of nodes and sparse connections between groups define a common comprehension of the community structure in a social IoT network. These groups of nodes and the links between them are called "Community," and the connections between groups are called "Intercommunity links" [32]. As illustrated in Fig. 2, the dashed lines represent the intercommunity links.

We consider a network which can be described by an undirected connected graph $G = (V, E)$, where V represents the set of nodes with cardinality $|V| = n$ and E represents the set

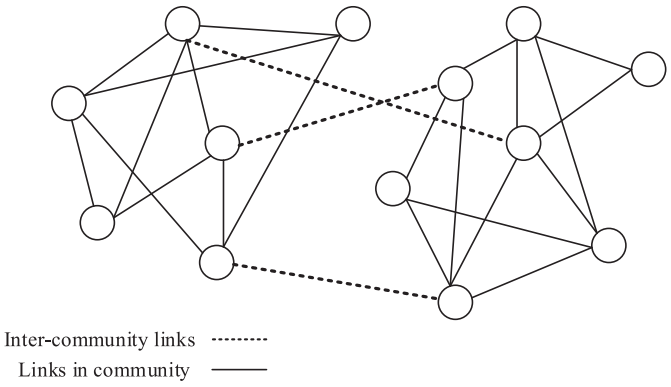


Fig. 2. Community and the intercommunity links in MSN.

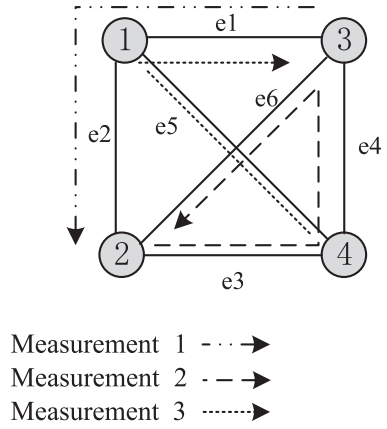


Fig. 3. Example of measurement in graph.

of edges with cardinality $|E| = k$. The weight of each link can be 1 or any other meaningful value in a certain network and the degree of every node is denoted as $\deg(N_i)$. In the graph, we consider a connected path with length t and the measurement of it is the sum of every link's weight on the path. Now, suppose we have m measurements in total over the whole graph. $\mathbf{x}$ represents a vector with length k whose jth element is the weight of the jth edge and $\mathbf{y}$ represents a vector with length m whose jth element is the result of the jth measurement. Matrix $\mathbf{A}$ with size $m \times k$ is called measurement matrix, which reflects the composition of the m measurements whose element in the ith row and jth column is 1 if the jth edge is on the ith measurement path, and 0 otherwise. Thus, we have

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{1}$$

Fig. 3 gives an example of the construction of the measurement matrix. We consider a network with $|V| = 4$ nodes, $|E| = 6$ edges, and 3 measurements. In measurement 2, the path goes through the e3, e4, and e6 in order. Correspondingly, the third, fourth, and sixth elements in the second row of the matrix $\mathbf{A}$ are 1's. The measurement matrix $\mathbf{A}$ is

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{2}$$

The typical model in the CS theory has the same expression as (2) [33]. It has been proved that if m and t satisfy $m = \mathcal{O}(s\log(n))$, $t = \mathcal{O}(|E|/s)$ and the matrix $\mathbf{A}$ satisfies the restricted isometry property condition, we can recover the *s-sparse* $\mathbf{x}$ through CS recovery algorithms with an expected accuracy, in which every measurement of matrix $\mathbf{A}$ or every connected link is generated by a "good start" regularized random walk [28], [29]. Furthermore, some prior information determined in advance can reduce the complexity and improve the accuracy of CS [34]. It is intuitive that if the weight of intercommunity links are sparse, the CS over graphs can be used on detection.

B. Encounter Probability

Learning from [9], we are going to use the parametric method here to generate the network structure. We see that the encounter durations of users follow a gamma distribution as:

$$X \sim \Gamma(k, \theta) = \Gamma\left(\frac{M_{i,j}^2}{I_{i,j}}, \quad \frac{I_{i,j}}{M_{i,j}}\right) \tag{3}$$

where $k = M_{i,j}^2/I_{i,j}$, $\theta = I_{i,j}/M_{i,j}$, $M_{i,j}$, and $I_{i,j}$ denote the mean and the variance of the encounter duration between users i and j, respectively.

Given the encounter histories of a group of users with the contact duration X_n and the number of encounters $N_{i,j}$, we can calculate the mean of the encounter duration $M_{i,j}$ as

$$M_{i,j} = \frac{\sum_n X_n}{N_{i,j}}. \tag{4}$$

Similarly, variance $I_{i,j}$ of the encounter duration can be obtained to represent the fluctuation of the encounter period

$$I_{i,j} = \frac{\sum_n \left(X_n - M_{i,j}\right)^2}{N_{i,j}}. \tag{5}$$

Thus, the probability density function of the encounter duration is given by

$$f(x;\ k,\ \theta) = \frac{1}{\theta^k}\frac{1}{\Gamma(k)}x^{k-1}e^{-\frac{x}{\theta}} \tag{6}$$

where $\Gamma(k) = \int_0^\infty t^{k-1}e^{-t}dt$.

In order to guarantee a successful data transmission, we assume that two users can be regarded as encountering with each other if and only if they have been detected by each other for at least $X_{\min}$. Thus, the probability of any two users encountering with each other is $p_{i,j}$, which can be represented as

$$p_{i,j} = 1 - \int_0^{X_{\min}} f(u; k, \theta)du = 1 - \frac{\gamma\left(k,\ \frac{X_{\min}}{\theta}\right)}{\Gamma(k)} \tag{7}$$

where $\gamma(k, [X_{\min}/\theta])$ is the lower incomplete gamma function, written as

$$\gamma\left(k, \frac{X_{\min}}{\theta}\right) = \int_0^{\frac{X_{\min}}{\theta}} t^{k-1}e^{-t}dt. \tag{8}$$

C. Edge-Clustering Coefficient

The edge-clustering coefficient is defined as the number of triangles including a given edge, divided by the possible maximum number of these triangles. Every edge has an edge-clustering coefficient and the possible maximum number of triangles can be calculated by the degree of two nodes linked by the edge [22]. For the edge between the ith node and jth node, its edge-clustering coefficient is

$$\mathbf{C}_{(i,j)} = \frac{z_{i,j}+1}{\min\left[\deg(\mathbf{N}_i)-1,\ \deg\left(\mathbf{N}_j\right)-1\right]} \tag{9}$$

where $z_{i,j}$ represents the number of triangles including that edge and $\min[\deg(\boldsymbol{N}_i)-1, \deg(\boldsymbol{N}_j)-1]$ is the maximum possible number of the triangles. The idea behind the coefficient is that the sparse intercommunity links have a low possibility to be included in a triangle and the links in the community are on the contrary. Specifically, the intercommunity links tend to have a smaller edge-clustering coefficient than the links in the community.

Thus, the edge-clustering coefficient indicates the intrinsic sparsity of the intercommunity links to a certain degree which can be used as the weight in the method of graph detection toward CS. The edge with the smallest edge-clustering coefficient has a large possibility to be an intercommunity link which can be served as the prior information in CS to reduce complexity.

Thus, combining edge-clustering coefficient with the CS-based intercommunity detection method leads to lower complexity and higher accuracy.

III. Proposed Intercommunity Detection Scheme

In this section, we propose a novel intercommunity detection scheme for the social IoT under the framework of CS over graphs [28], [29]. We define a novel metric by exploiting the encounter probability and the edge-clustering coefficient to construct the measurement matrix better. Furthermore, the edge-clustering coefficient is utilized as the prior information to reduce the complexity of the standard CS recovery algorithms and enhance the recovery performance as well.

The whole algorithm consists of three steps.

A. Step 1: Edge-Clustering Coefficient Calculation

The edge-clustering coefficient can be simply calculated based on the adjacent matrix of the graph. First, we measure the number of triangles that each link is included in the matrix $\mathbf{Z}$ whose element in the ith row and the jth column represents the number of triangles of the edge between $\boldsymbol{N}_i$ and $\boldsymbol{N}_j$. Then, we use (9) to derive the edge-clustering coefficient in matrix $\mathbf{C}$. Finally, we choose the two edges with the smallest edge-clustering coefficient as the prior information which is used in the following steps. Also, the labels of the two chosen edges are restored in vector $\boldsymbol{\eta}$.

B. Step 2: Measurement Matrix Construction

As explained in Section II-B, the encounter probability between two users is one of the important parameters to describe a social network. Different from the random walk method, we define a novel metric to describe the weight on a node, i.e., the degree of node $\boldsymbol{N}_i$ multiplies the encounter probability of all the neighbors of $\boldsymbol{N}_i$, which is written as

$$\mathbf{S}_i = \deg(\mathbf{N}_i) \times \left(\sum_{j \in \text{neighbor}_i} \mathbf{U}_{(i,j)} \right) \tag{10}$$

where $\mathbf{U}_{i,j}$ denotes the encounter probability between the nodes i and j, considering that intercommunity links have higher probabilities to connect nodes with large degrees and frequent encounters.

Algorithm 1 Enhanced CS Over Graphs-Based Intercommunity Detection

Input:
1) graph of the network $G(V, E)$;
2) measurement number m;
3) number of measurement steps t;
4) the intercommunity links' sparsity s;
5) iteration times T;
6) encounter possibility matrix U;

Output: The s-sparse estimate vector $\mathbf{x}$;

Step 1: Edge-Clustering Coefficient Calculation:

$\mathbf{Z}_{(i,j)}$ = numbers of triangles between i and j;

$\mathbf{C}_{(i,j)} = \dfrac{z_{i,j}+1}{\min[\deg(\boldsymbol{N}_i)-1, \deg(\boldsymbol{N}_j)-1]}$;

$\boldsymbol{\eta}$ = indexes of the minimum elements in $\mathbf{C}$;

Step 2: Measurement Matrix Construction:

$\mathbf{S}_i = \deg(\boldsymbol{N}_i) \times \left(\sum_{j \in \text{neighbor}_i} \mathbf{U}_{(i,j)}\right)$;

$\mathbf{P}_{first,i} = \dfrac{\mathbf{S}_i}{\sum_i \mathbf{S}_i}$

for $i = 1 : m$ **do**
  $\boldsymbol{V}_t$ = select related from $\mathbf{P}_{first}$;
  add $\boldsymbol{V}_t$ as the first node of $\boldsymbol{\Omega}_{temp}$;
  for $j = 1 : t$ **do**
    if $\text{neighbor}_{\boldsymbol{V}_t} \neq \boldsymbol{\phi}$ **then**
      $\mathbf{P}_{trans,i,j} = \dfrac{\mathbf{S}_j}{\sum_{k \in \text{neighbor}_i} \mathbf{S}_k}$;
      $\boldsymbol{V}_{next}$ = select in $\text{neighbor}_{\boldsymbol{V}_t}$ related from $\mathbf{P}_{trans}$;
    else
      $\boldsymbol{V}_{next}$ = the previous node;
    end if
    delete the link between $\boldsymbol{V}_t$ and $\boldsymbol{V}_t$, update $\mathbf{S}$;
    add $\boldsymbol{V}_{next}$ into $\boldsymbol{\Omega}_{temp}$;
    $\boldsymbol{V}_t = \boldsymbol{V}_{next}$;
  end for
  add $\boldsymbol{\Omega}_{temp}$ into the ith row of Matrix $\mathbf{A}$;
end for

Step 3: Compressive Sensing:

$$\mathbf{W}_{(i,j)} = \begin{cases} \dfrac{1}{\mathbf{C}_{(i,j)}}, & \mathbf{C}_{(i,j)} < \boldsymbol{\varepsilon}(\mathbf{C}); \\ 0, & \mathbf{C}_{(i,j)} > \boldsymbol{\varepsilon}(\mathbf{C}); \end{cases}$$

transmit $\mathbf{W}$ into $\mathbf{w}$;

$\mathbf{y} = \mathbf{A}\mathbf{w}$;

$\mathbf{x} = CS(\mathbf{A}, \mathbf{y}, T, \boldsymbol{\eta})$.

Then we choose the nodes one-by-one with the possibility according to $\mathbf{S}$. After a node is chosen, we delete the link between V_{next} and V_t, and update $\mathbf{S}$. In this step, we operate m times walk with length t and add the ith measurement to the ith row of matrix $\mathbf{A}$.

C. Step 3: Compressive Sensing

In the model of CS over graphs, shown in (1), we assume that *s-sparse* vector $\mathbf{x}$ represents the weight of every link in the graph. However, in the intercommunity detection, we assume that the nonzero elements of $\mathbf{x}$ represent the intercommunity links. As a result, there is no harm to suppose $\mathbf{x}$ in the easiest form whose ith element is positive when the ith link is an intercommunity link, and 0 otherwise. Consequently, to approach the measurement result $\mathbf{y}$, we transform the matrix $\mathbf{W}$ to the vector $\mathbf{w}$ whose ith element is 0 if the ith link's edge-clustering coefficient is greater than average, and $1/\mathbf{C}_{(i,j)}$ otherwise, which means links with a small edge-clustering coefficient have a higher possibility to be the intercommunity links. Recognizing the measurement matrix $\mathbf{A}$ is generated with a high possibility to pass the intercommunity links, we have

$$\tilde{\mathbf{y}} = \mathbf{A}\mathbf{w} \approx \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{11}$$

Thus, we use $\mathbf{A}$ and $\mathbf{w}$ to reconstruct $\mathbf{y}$. Finally, the CS algorithm is applied to recover the vector $\mathbf{x}$, with the observation vector $\mathbf{y}$, measurement matrix $\mathbf{A}$, the iteration time T, and the prior information $\boldsymbol{\eta}$. Various conventional CS recovery algorithms can be used to solve the problem, such as orthogonal matching pursuit [35], compressive sampling matching pursuit (CoSaMP) [34], and subspace pursuit (SP) [36]. The comparisons of the complexity and the number of minimum measurements of these algorithms are summarized in Table I.

In our proposed scheme, since the edges with smaller edge-clustering coefficient have a larger possibility to be intercommunity links, indexes $\boldsymbol{\eta}$ of the minimum elements in the edge-clustering coefficient matrix are utilized as prior information to reduce the complexity of the conventional CS algorithms. The initial support set is derived according to $\boldsymbol{\eta}$, thus there are only $(s - |\boldsymbol{\eta}|)$ remaining nonzero entries need to be detected. The cardinality of $\boldsymbol{\eta}$ can be chosen according to the scale of the network. If the SP recovery algorithm is utilized to detect the intercommunity links, the complexity is reduced to $\mathcal{O}((s - |\boldsymbol{\eta}|)mn)$. Since the CS recovery algorithm contributes the major complexity, the proposed algorithm has much lower complexity than conventional CS over graphs-based algorithms.

The pseudo-code of our proposed detection scheme is shown in Algorithm 1.

IV. Simulation Results

This section investigates the performance of the proposed scheme. We verify our proposed methods on one traditional social network trace and two MSN traces to demonstrate the performance of the proposed scheme. The traditional social network is simulated as a special example in which the encounter probability is 1 between every pair of users. Among the two MSNs, one is a raw data set collected by CRAWDAD team [37], the other is Infocom data collected during Infocom 2005 [38]. These sets of network data are used to illustrate the outstanding performance of the proposed method on intercommunity detection compared with the random walk-based method (without the prior information). Besides, we choose the CoSaMP algorithm to perform the sparse signal recovery in the proposed method [34].

TABLE I
Complexity and Minimum Measurement Requirement of Some CS Recovery Algorithms

Algorithms	Complexity	Minimum Measurement
OMP	$\mathcal{O}(smn)$	$\mathcal{O}(s\log n)$
CoSaMP	$\mathcal{O}(mn)$	$\mathcal{O}(s\log n)$
SP	$\mathcal{O}(smn)$	$\mathcal{O}(s\log(n/s))$

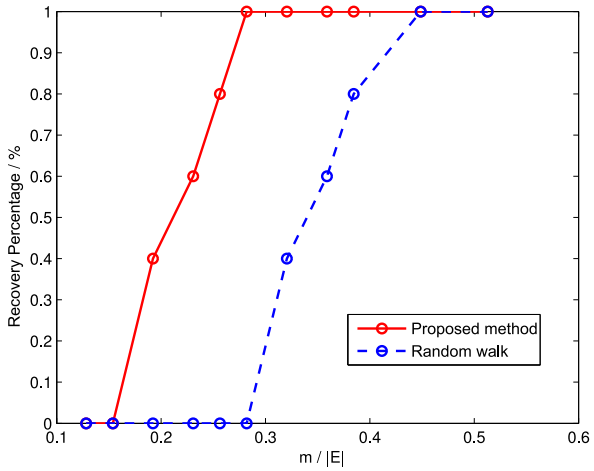


Fig. 4. Karate Club network intercommunity detection recovery percentage with $s = 5$.

A. Traditional Social Network

To evaluate the performance gain of the prior information, we choose a traditional social network to simulate our algorithm on, the Zachary's Karate Club network [39]. The Karate social network has 34 nodes and 78 edges. The random walk-based method is also provided here for comparison.

As is shown in Fig. 4, the proposed method achieves better performance than that of the random walk-based method, i.e., has the higher recovery rate with the same number of measurements. This improvement may benefit from the prior information based on the edge-clustering coefficient.

B. CRAWDAD Data

The CRAWDAD data is collected by T-mote invent devices which can detect invent-to-invent encounters anywhere throughout the town of St. Andrews and beyond. The raw data is expressed in a 25×25 matrix $\mathbf{R}$ whose entry in the ith row and the jth column represents the probability of the ith and jth users encountering with each other for at least 20 s [9]. To process the data and analyze the community structure, we set up a threshold at 0.8 to distinguish the relationship between

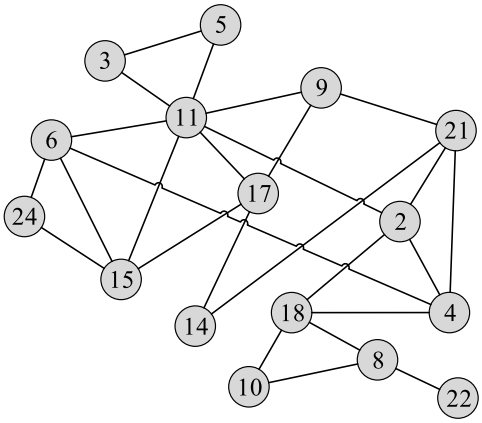


Fig. 5. Trunk of the CRAWDAD MSN.

TABLE II
INDEXES OF THE LINKS IN THE CRAWDAD MSN

Link index	Nodes connection	Link index	Nodes connection
1	2, 4	14	8, 10
2	2, 11	15	8, 18
3	2, 18	16	8, 22
4	2, 21	17	10, 18
5	3, 5	18	9, 11
6	3, 11	19	9, 17
7	4, 6	20	9, 21
8	4, 18	21	11, 15
9	4, 21	22	11, 17
10	5, 11	23	14, 17
11	6, 11	24	14, 21
12	6, 15	25	15, 17
13	6, 24	26	15, 24

users meaning that if the probability is larger than 0.8, we account the two people have a tight social connection. So, we get new adjacent matrix $\mathbf{G}$ and encounter matrix $\mathbf{U}$ from the following formula:

$$\mathbf{G}_{(i,j)} = \begin{cases} 1, & \mathbf{R}_{(i,j)} > 0.8 \\ 0, & \mathbf{R}_{(i,j)} < 0.8 \end{cases} \tag{12}$$

$$\mathbf{U}_{(i,j)} = \begin{cases} \mathbf{R}_{(i,j)}, & \mathbf{R}_{(i,j)} > 0.8 \\ 0, & \mathbf{R}_{(i,j)} < 0.8. \end{cases} \tag{13}$$

Actually, this operation with the threshold may wipe off the people who have few number of connections with others and push the intercommunity detection to concentrate on the trunk of the MSN. After the operation, 26 links remain in the trunk of the MSN, which is shown in Fig. 5. The indexes of the corresponding links are also illustrated in Table II.

Then, the proposed method is applied to the matrix G with different sparsity from $s = 2$ to $s = 4$. The detected indexes of the intercommunity links are shown in Table III. An illustration is also provided to explain the simulation results more intuitively. In Fig. 6, we can see the 2nd, 7th, 20th, and 24th links, which are represented by the dotted lines, dashed lines, and two types of dotted and dashed lines, respectively, connecting the two communities in this network, which is a nice result proving that our proposed method works efficiently on the community detection. The proposed method is compared with the random walk method with different m. The results are

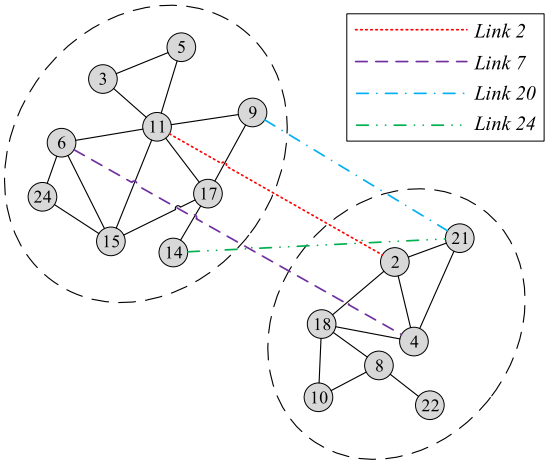


Fig. 6. CRAWDAD intercommunity detection results.

TABLE III
CRAWDAD MSN INTERCOMMUNITY DETECTION RESULTS WITH DIFFERENT s

Sparsity	Indexes of the inter-community links
2	2, 7
3	2, 7, 20
4	2, 7, 20, 24

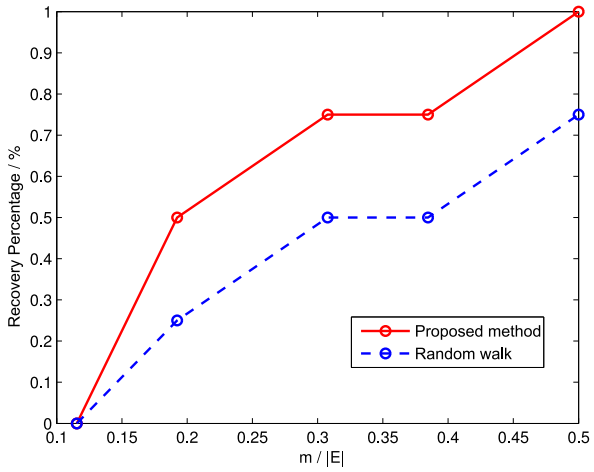


Fig. 7. CRAWDAD intercommunity detection recovery percentage with $s = 4$.

shown in Fig. 7 at the condition of $s = 4$ and $t = 15$. It can be seen from the results that the proposed method achieves higher accuracy even with fewer measurements.

C. Infocom Data

Another data set is the Infocom data collected by [38]. The experiments were conducted during Infocom 2005 for around three days. In the year of 2005, the traces includes the encounter history of 41 mobile iMotes with a wireless range around 30 m, which were worn by students attending the conference. We simulated the proposed method on the Infocom 2005 data after the same preprocessing as the CRAWDAD data, shown in (12) and (13). The results are illustrated in Fig. 8. It can be seen from the figure that the

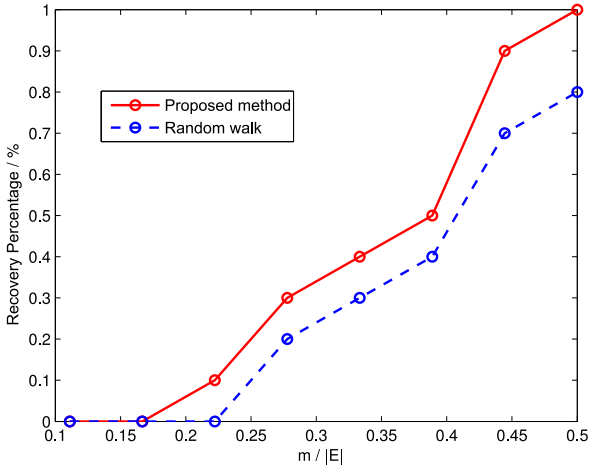


Fig. 8. Infocom 2005 intercommunity detection recovery percentage with $s = 10$.

recovery percentage increase with the increase of $m/|E|$, and the proposed method enjoys better recovery percentage in all range of $m/|E|$. The proposed enhanced CS over graphs-based intercommunity detection method still demonstrates better precision even with fewer measurements.

V. Conclusion

Facing the new trend of combining IoT and MSN to provide close interaction between humans and things for smart cities, we considered the intercommunity detection issue in this paper. A CS over graphs-based intercommunity detection scheme is first proposed for the social IoT scenarios. The proposed scheme utilizes the encounter probability and the edge-clustering coefficient to define a novel additive metric to make the CS algorithms applicable. Moreover, the edge-clustering coefficient is exploited as the prior information to reduce the complexity and enhance the recovery performance of the conventional CS algorithms. Simulation results demonstrate that the proposed scheme enjoys better performance than the traditional random walk-based one.

References

[1] J. A. Stankovic, "Research directions for the Internet of Things," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 3–9, Feb. 2014.

[2] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

[3] H. Yue, L. Guo, R. Li, H. Asaeda, and Y. Fang, "Dataclouds: Enabling community-based data-centric services over the Internet of Things," *IEEE Internet Things J.*, vol. 1, no. 5, pp. 472–482, Oct. 2014.

[4] Z. Pei-Kun, Z. Juan-Juan, and W. Wu, "Division of mobile social network based on user behavior," in *Proc. Int. Conf. Wavelet Analy. Pattern Recognit. (ICWAPR)*, Tianjin, China, Jul. 2013, pp. 148–152.

[5] N. Kayastha, D. Niyato, P. Wang, and E. Hossain, "Applications, architectures, and protocol design issues for mobile social networks: A survey," *Proc. IEEE*, vol. 99, no. 12, pp. 2130–2158, Dec. 2011.

[6] Y. E. Sagduyu and Y. Shi, "Navigating a mobile social network," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 122–128, Oct. 2015.

[7] A. Sapuppo, "Spiderweb: A social mobile network," in *Proc. Wireless Conf. (EW)*, Lucca, Italy, Apr. 2011, pp. 475–481.

[8] G. Araniti, A. Orsino, L. Militano, L. Wang, and A. Iera, "Context-aware information diffusion for alerting messages in 5G mobile social networks," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 427–436, Apr. 2017.

[9] Y. Zhang, E. Pan, L. Song, W. Saad, Z. Dawy, and Z. Han, "Social network aware device-to-device communication in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 177–190, Jan. 2015.

[10] M. I. Akbas, M. R. Brust, and D. Turgut, "Social network generation and role determination based on smartphone data," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Orlando, FL, USA, Mar. 2012.

[11] M. I. Akbas, R. N. Avula, M. A. Bassiouni, and D. Turgut, "Social network generation and friend ranking based on mobile phone data," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Budapest, Hungary, Jun. 2013, pp. 1444–1448.

[12] A. M. Ortiz, D. Hussein, S. Park, S. N. Han, and N. Crespi, "The cluster between Internet of Things and social networks: Review and research challenges," *IEEE Internet Things J.*, vol. 1, no. 3, pp. 206–215, Jun. 2014.

[13] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[14] R. Jalali, K. El-Khatib, and C. McGregor, "Smart city architecture for community level services through the Internet of Things," in *Proc. Int. Conf. Intell. Next Gener. Netw. (ICIN)*, Paris, France, Feb. 2015, pp. 108–113.

[15] Y. L. Lin, W. C. Chung, C. H. Hsu, and Y. C. Chung, "Community-based M2M framework using smart/HetNet gateways for Internet of Things," in *Proc. IEEE Int. Conf. Data Sci. Data Intensive Syst. (DSAA)*, Sydney, NSW, Australia, Dec. 2015, pp. 688–695.

[16] D. Hussein, E. Bertin, and V. Frey, "A community-driven access control approach in distributed IoT environments," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 146–153, Mar. 2017.

[17] O. B. Abderrahim, M. H. Elhdhili, and L. Saidane, "TMCoI-SIOT: A trust management system based on communities of interest for the social Internet of Things," in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Valencia, Spain, Jun. 2017, pp. 747–752.

[18] E. M. Daly and M. Haahr, "Social network analysis for information flow in disconnected delay-tolerant MANETs," *IEEE Trans. Mobile Comput.*, vol. 8, no. 5, pp. 606–621, May 2009.

[19] P. Hui, J. Crowcroft, and E. Yoneki, "BUBBLE rap: Social-based forwarding in delay-tolerant networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 11, pp. 1576–1589, Nov. 2011.

[20] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 12, pp. 7821–7826, Jun. 2002.

[21] A. Didwania and Z. Narmawala, "A comparative study of various community detection algorithms in the mobile social network," in *Proc. 5th Nirma Univ. Int. Conf. Eng. (NUiCONE)*, Ahmedabad, India, Nov. 2015, pp. 1–6.

[22] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and identifying communities in networks," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 9, pp. 2658–2663, Mar. 2004.

[23] R. Zhang, L. Li, C. Bao, L. Zhou, and B. Kong, "The community detection algorithm based on the node clustering coefficient and the edge clustering coefficient," in *Proc. 11th World Congr. Intell. Control. Autom. (WCICA)*, Shenyang, China, Jun. 2014, pp. 3240–3245.

[24] J. Wang, M. Li, H. Wang, and Y. Pan, "Identification of essential proteins based on edge clustering coefficient," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 9, no. 4, pp. 1070–1080, Jul./Aug. 2012.

[25] J. Haupt, R. Castro, R. Nowak, G. Fudge, and A. Yeh, "Compressive sampling for signal classification," in *Proc. 6th Fortieth Asilomar Conf. Signals Syst. Comput. (ACSSC)*, Pacific Grove, CA, USA, Oct. 2006, pp. 1430–1434.

[26] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.

[27] Z. Han, H. Li, and W. Yin, *Compressive Sensing for Wireless Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[28] W. Xu and A. Tang, "Compressive sensing over graphs: How many measurements are needed?" in *Proc. Int. Conf. Wavelet Anal. Pattern Recognit. (ICWAPR)*, Allerton, IL, USA, Sep. 2010, pp. 615–619.

[29] W. Xu, E. Mallada, and A. Tang, "Compressive sensing over graphs," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Shanghai, China, Apr. 2011, pp. 2087–2095.

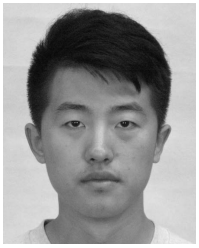
[30] H. Mahyar, H. R. Rabiee, A. Movaghar, E. Ghalebi, and A. Nazemian, "CS-ComDet: A compressive sensing approach for inter-community detection in social networks," in *Proc. IEEE/ACM Int. Conf. Adv. Soc. Netw. Anal. Min. (ASONAM)*, Paris, France, Aug. 2015, pp. 89–96.

[31] L. Atzori, A. Iera, and G. Morabito, "SIoT: Giving a social structure to the Internet of Things," *IEEE Commun. Lett.*, vol. 15, no. 11, pp. 1193–1195, Nov. 2011.
[32] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, nos. 3–5, pp. 75–174, Feb. 2010.
[33] E. Candes, "Compressive sampling," in *Proc. Int. Congr. Math.*, Madrid, Spain, Sep. 2006, pp. 1433–1452.
[34] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmonic Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.
[35] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[36] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
[37] D. Kotz and T. Henderson, "CRAWDAD: A community resource for archiving wireless data at dartmouth," *IEEE Pervasive Comput.*, vol. 4, no. 4, pp. 12–14, Oct./Dec. 2005.
[38] A. Chaintreau *et al.*, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Trans. Mobile Comput.*, vol. 6, no. 6, pp. 606–620, Jun. 2007.
[39] W. W. Zachary, "An information flow model for conflict and fission in small groups," *J. Anthropol. Res.*, vol. 33, no. 4, pp. 452–473, 1997.



Tengjiao Wang received the B.S. degree from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2016, where he is currently pursuing the Ph.D. degree at the DTV Technology Research and Development Center.

His current research interests include 5G wireless technologies, multiple-input multiple-output systems, nonorthogonal multiple access, and sparse signal processing theory in wireless communications.



Jingbo Tan received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2017, where he is currently pursuing the Ph.D. degree at the Department of Electronic Engineering.

His current research interests include wireless communications, with the emphasis on massive multiple input multiple output and millimeter-wave communications.



Wenbo Ding (S'15–M'16) received the B.S.E. degree (with Distinction) from the Department of Electronic Engineering, Tsinghua University, Beijing China, in 2011, where he is currently pursuing the Ph.D. degree at the DTV Technology Research and Development Center.

His current research interests include sparse signal processing and its applications in the field of the power line communication, visible light communication, smart grid and future 5G wireless communications. He has authored or co-authored over 30 journal and conference papers.

Dr. Ding was a recipient of the IEEE Scott Helt Memorial Award for the best paper published in the IEEE TRANSACTIONS ON BROADCASTING in 2015, the Best Paper Award of the National Doctoral Student Academic Conference, the Tsinghua Top Grade Scholarship, and the Academic Star of Electronic Engineering Department in Tsinghua University in 2015.



Yanru Zhang (S'13–M'16) received the B.S. degree in electronic engineering from the University of Electronic Science and Technology of China, in 2012, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Houston, Houston, TX, USA, in 2016.

She is currently a Post-Doctoral Fellow with the Department of Information Engineering, Chinese University of Hong Kong, Hong Kong. Her current research interests include the contract theory and matching theory in network economics, Internet and applications, wireless communications, and networking.

Dr. Zhang was a recipient of the Best Paper Award of the IEEE ICCS 2016 and ICCC 2017.



Fang Yang (M'11–SM'13) received the B.S.E. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2005 and 2009, respectively.

He is currently an Associate Professor with the Research Institute of Information Technology, Tsinghua University. He has authored or co-authored over 120 peer-reviewed journal and conference papers. He holds over 40 Chinese patents and 2 PCT patents. His current research interests include channel coding, channel estimation, interference cancellation, and signal processing techniques for communication system, especially in power line communication, visible light communication, and digital television terrestrial broadcasting.

Dr. Yang was a recipient of the IEEE Scott Helt Memorial Award (Best Paper Award of the IEEE TRANSACTIONS ON BROADCASTING) in 2015. He is the Secretary General of Subcommittee 25 of the China National Information Technology Standardization (SAC/TC28/SC25).



Jian Song (M'06–SM'10–F'16) received the B.Eng. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1990 and 1995, respectively.

He was with Tsinghua University. He was with the Chinese University of Hong Kong, Hong Kong, and the University of Waterloo, Waterloo, ON, Canada, in 1996 and 1997, respectively. He has been with Hughes Network Systems, Germantown, MD, USA for seven years. In 2005, he joined the faculty team with Tsinghua University, as a Professor, where he is currently the Director of Tsinghua University's DTV Technology Research and Development Center. He has performed research in many different areas of fiber-optic, satellite, and wireless communications, as well as the power line communications. His current research interests include digital TV broadcasting. He has authored or co-authored over 200 peer-reviewed journal and conference papers. He holds 2 U.S. and over 40 Chinese patents.

Dr. Song is a Fellow of the IET.



Zhu Han (S'01–M'04–SM'09–F'14) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland at College Park, College Park, MD, USA, in 1999 and 2003, respectively.

From 2000 to 2002, he was a Research and Development Engineer with JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland at College Park. From 2006 to 2008, he was an Assistant Professor with Boise State University, Boise, ID, USA. He is currently a Professor with the Electrical and Computer Engineering Department and the Computer Science Department, University of Houston, Houston, TX, USA. His current research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grids.

Dr. Han was a recipient of the NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award of the *Journal on Advances in Signal Processing* in 2015, the IEEE Leonard G. Abraham Prize in the field of communications systems (Best Paper Award of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS) in 2016, and several Best Paper Awards of IEEE conferences. He is currently an IEEE Communications Society Distinguished Lecturer.