A Social-Relation-Based Game Model for Distributed Clustering in Cooperative Wireless Networks

Li Bowen
Institute of Information Engineering.CAS
School of Cyber Security.UCAS
Beijing,China
libowen@iie.ac.cn

Zhang Shunliang*
Institute of Information Engineering.CAS
Beijing,China
zhangshunliang@iie.ac.cn

Shan Xu
Institute of Information Engineering.CAS
Beijing,China
shanxu@iie.ac.cn

Wang Yongming
Institute of Information Engineering.CAS
Beijing,China
wangyongming@iie.ac.cn

Gao Zhenxiang
Department of Computer Science and Engineering, UConn
Storrs, CT, USA
zhenxiang.gao.mail@gmail.com

ABSTRACT

In this paper, a novel framework for cluster detection in cooperative wireless networks is proposed. This framework is modeled by a dynamic game with incomplete information, in which each player in the game aspires to improve its position in the network by forming cooperative groups. Instead of static systems, the attention we paid in this paper is highly dynamic networks, where the users' high mobility brings a huge challenge in clustering. In order to mitigate that impact, this paper is from the perspective of social relations to clustering, extracting users' social nature from their mobile patterns and designing distributed cluster strategy based on game model. The introduction of social nature with generally long-term characteristics makes clustering framework more predictive and stable. Simulations on real-world networks show that the proposed approach performs well in clustering in cooperative wireless networks.

*Corresponding author

This work was supported by the Integrated System Architecture and Network Security Technology under Grant No. Y9X0021104.

MobiQuitous, November 12–14, 2019, Houston, TX, USA

ACM ISBN 978-1-4503-7283-1/19/11.
https://doi.org/10.1145/3360774.3360820

CCS CONCEPTS

• **Computing methodologies** → **Self-organization**; • **Networks** → *Network dynamics*; *Mobile ad hoc networks*; • **Human-centered computing** → *Mobile devices.*

KEYWORDS

5G, Cooperative wireless networks, Distributed cluster, Social relations, Game theory.

ACM Reference Format:
Li Bowen, Zhang Shunliang, Shan Xu, Wang Yongming, and Gao Zhenxiang. 2019. A Social-Relation-Based Game Model for Distributed Clustering in Cooperative Wireless Networks. In *16th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous), November 12–14, 2019, Houston, TX, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3360774.3360820

1 INTRODUCTION

Ever rapid proliferation of wireless devices and services has been bringing exponential growth in mobile Internet, which poses potential challenges for mobile operators, especially in terms of quality of service provisioning [16]. As a promising paradigm in 5G [15], cooperation in the wireless network can efficiently improve overall performance by resource sharing, including decreasing device energy consumption, balancing network load, increasing network scalability and reliability, and so forth. However, such a paradigm needs one primary requirement that an efficient technique to form cooperative groups, i.e. clusters.

A clustering algorithm is defined by a set of rules that form a heuristic to solve the problem of splitting a network into different partitions (called clusters, groups or coalitions). One of the most widely-used methods in the state of the art

is Voronoi-based clustering. Those heuristics partition the clusters by the influence spheres of cluster heads (CHs), i.e, detecting CHs first and forming clusters latter. For example, Lowest-ID algorithm provided in [1] allows nodes with the lowest identifier to be selected as CHs and the rest of nodes choose to affiliate to their nearest neighbor CH. On the basis of this work, there are varieties of improved algorithms for improving cluster performance. For instance, Mario et al. [4] introduce CHs detecting rules for improving the cluster reliability, Chiang et al. [2] introduce re-cluster control rules for decreasing cluster complexity, and Li et al. [11] introduce cluster size rules for improving cluster effectiveness. However, the principle of Voronoi-based methods, i.e., CH-selecting first, lacks controlling on the cluster itself, which is no easy to complete by patching rules.

Many researchers study the clustering problem from the cluster itself rather than based on CH selection. S-WEB is a cluster-first attempt, which groups nodes based on their geographic location relative to the base station (BS). The flexibility of S-WEB [9], the cluster size can be controlled by distance and angle parameters, makes it useful in static networks. However, S-WEB is not ideal in dynamic systems because of the instability of cluster members caused by mobilities. In this context, some researchers propose mobility-aware clustering techniques to cluster nodes with similar mobility patterns [14, 17]. Those methods utilize mobility information, collected from GPS or environmental information like signal strength (RSS), signal frequency (Doppler effect), as input to predict nodes movement and form stable clusters. However, the performance of those mobility-aware clustering is limited because *(i)* not all mobile devices are equipped with a GPS; *(ii)* GPS consumes much energy and presents inaccuracy in indoor; *(iii)* unpredictable channel conditions produce inaccurate signal characteristics [15].

In this paper, we investigate the clustering problem in cooperative wireless networks from the perspective of social nature. It's immediate that nowadays most mobile devices are carried by humans, such as smartphones, wearables, and vehicles, that is, the mobilities of those devices are indeed controlled by humans. In addition, the mobilities of humans are not chaotic but are strongly impacted by their social relations and these social relations have stable long-term characteristics [3, 5]. Thus, stable clusters can be formed by grouping nodes with similar social relations. To achieve that, this paper first explores the social nature of nodes and proposes a metric to extract social relationship among users from their contact patterns, as well as a metric to detect social community of each node. The validity of metrics are proved by theoretical derivations. Note that the analysis of social nature is no need for accurate characteristics of the signal, but only binary information (contact or not) to obtain contact patterns. Inspiring by game theory, we formulate the clustering problem as a dynamic game with incomplete information and then design a distributed clustering algorithm on the basis of the social natures. Instead of most existing methods, this algorithm is no need of CH detection. Extensive simulations on real-world mobility networks show that the proposed clustering algorithm can well form cooperative groups.

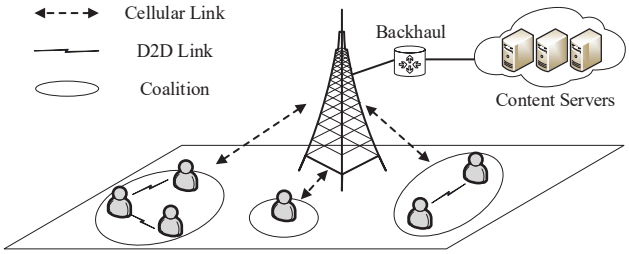


Figure 1: A paradigm of cooperative wireless networks.

The rest of the paper is organized as follows: In Section 2, we briefly depict the network model. In Section 3, we detail our framework for distributed cluster detection. The detailed social parameters analysis and design are described in Section 4, and the experiment results are illustrated in Section 5. Finally, we conclude the paper along with insights into future directions in Section 6.

2 NETWORK MODEL

Consider a cooperative wireless network (as shown in Fig. 1), which consists of N mobile devices (MDs). Each MD can directly communicate with others over short-range radio links when they are within the direct transmission range, and can communicate with base stations (BSs) via the cellular links. In this paper, we assume the cooperation is for data acquisition and all MDs in the network are interested in these data. For the sake of clarity, the dynamic cooperative wireless network in this paper is modeled as a temporal graph set.

Definition 1. *A global graph $\mathcal{G}_g := \{G_0, \ldots, G_T\}$ is composed by series of transient graph $G_t := (\mathcal{V}_t, \mathcal{E}_t)$, where $\mathcal{V}_t := \{v_i | \, 1 \le i \le N\}$ is a set of nodes representing all MDs and $\mathcal{E}_t := \{(v_i, v_j)_t \, | \, d(v_i, v_j)_t \le D, \, v_i \in \mathcal{V}_t, \, v_j \in \mathcal{V}_t\}$ is a set of edges representing all interaction states among the MDs at the time slot t. If and only if two MDs are in physical proximity, i.e., their distance $d(v_i, v_j)_t$ is less than the maximum wireless transmission distance D, they can communicate with each other via D2D link. Note that the symbol $:=$, in this paper, stands for by definition.*

However, global network information is fuzzy for nodes in the distributed system — local nodes are unable to grasp the information beyond their surroundings. Thus, in terms of the information local users have, we model it as the local network graphs.

Definition 2. *A local graph is defined as* $\{\mathcal{G}_l(v_i)\}_t := [\mathcal{V}_t\{v_i\}, \mathcal{W}_t\{v_i\}_{2\times M}, \mathcal{H}_t\{v_i\}]$. $\mathcal{V}_t\{v_i\} := \{v_j | \left(v_i, v_j\right)_t \in \mathcal{E}_t\}$ *is a contacting-nodes set, recording the contact feature of node* v_i *at each time.* $\mathcal{W}_t\{v_i\}_{2\times M}$ *is a two-dimensional vector, recording the relationships with contacted nodes, where* $\mathcal{W}_t\{v_i\}_{1\times m} := \bigcup_{k=1}^{t} \mathcal{V}_t\{v_i\}$ *records all nodes ID* v_i *has contacted and* $\mathcal{W}_t\{v_i\}_{2\times m} := \{w_t(v_i, v_j) | v_j \in \mathcal{W}_t\{v_i\}_{1\times M}\}$ *records the relative relationships calculated by social-relation-extraction algorithm* $w_t(v_i, v_j) := f(w_{t-1}(v_i, v_j), \mathcal{V}_t\{v_i\})$. $\mathcal{H}_t\{v_i\} := \{v_j | \mathfrak{C}_{v_j}(w_t(v_i, v_j)), v_j \in \mathcal{W}_t\{v_i\}_{1\times M}\}\}$ *is the community structure, recording the* v_i*'s close nodes, where* $\mathfrak{C}_{v_j}(w_t(v_i, v_j))$ *is close condition related to current relationships. The social-relation-extraction algorithm* $f(\cdot)$ *and the close condition* $\mathfrak{C}$ *will be introduced in section 4.*

3 CLUSTERING FRAMEWORK

This section details our framework of distributed clustering technique, comprising of two parts: dynamic game model and distributed clustering algorithm design.

3.1 Game Model

In consideration of the self-organizing nature of nodes (i.e. humans), we model each node in the network as a game player. Specifically, every node wants to cooperate by forming coalitions in order to improve their positions, i.e, fast data acquisition, in the game. Notice that the game in our model is played over a number of periods due to the dynamic nature of the network. Each player is available to join any coalition as they want at any time. Formally, the dynamic game is denoted as follow:

Definition 3. *The dynamic game is denoted as* $\mathcal{Q} = [\mathcal{N}_t, \{\mathcal{P}_t^n\}_{n\in\mathcal{N}_t}, \{\mathcal{U}_t^n\}_{n\in\mathcal{N}_t}]$, *and the formulation of this game is as follow:*

- *Players:* $\mathcal{N}_t$ *is the set of game players, that is, all MDs at time* t.
- *Strategy:* $\mathcal{P}_t^n = \{p_1, p_2 \cdots, p_{\max}\}$ *is the set of available coalition selection strategies for player* n *at time* t. *A strategy profile of all players is a vector* $\mathbf{P}_t = (p_t^1, p_t^2, \cdots, p_t^N) \in \mathcal{P}_t$, *where* p_t^n *stands for player* n*'s decision at time* t *and* $\mathcal{P}_t = \mathcal{P}_t^1 \otimes \mathcal{P}_t^2 \otimes \cdots \otimes \mathcal{P}_t^N$ *stands for the joint strategy space for all players at time* t.
- *Utility:* $\mathcal{U}_t^n$ *is the utility function for measuring the avenue of each decision according player* n*'s knowledge about coalitions at time* t. *Notice that each player's knowledge is not complete, even not right sometimes.*
- *Coalition structure:* $\mathcal{O}_t = \{\mathcal{S}_1, \ldots, \mathcal{S}_l, \ldots, \mathcal{S}_L\}$ *is the transient coalition structure at time* t, *where* $\mathcal{S}_l$ *is a coalition verifying* $\bigcup_1^L \mathcal{S}_l = \mathcal{V}_t$. *The coalition structure is initialized with each player as a single coalition, denoted as* $\mathcal{O}_0 = \{\mathcal{S}_l | \mathcal{S}_l = v_l, 1 \le l \le N\}$. *Noted that the game processes are asynchronous and incomplete because of distributed system. Thus there exists divergence for different players about one coalition, formally,* $\exists \mathcal{S}_l \in \mathcal{O}_t, t \ge 0, \mathcal{S}_l(\mathfrak{P}) \ne \mathcal{S}_l(\mathfrak{O})$, *where* $\mathcal{S}_l(\mathfrak{P})$ *represents the coalition structure player* $\mathfrak{P}$ *grasps, and there exist two coalitions including the same player at a time* t, *formally,* $\exists \mathcal{S}_l, \mathcal{S}_m \in \mathcal{O}_t, t \ge 0, \mathcal{S}_l \cap \mathcal{S}_m \ne \emptyset$.

In our model, the decisions what each time the players need to make are to choose whether participates in a new coalition or keeps constant. We call these processes as switch operations.

Definition 4. *A switch operation* $\sigma_{l\mapsto k}(\mathfrak{P})$ *is defined as the transfer of player* $\mathfrak{P}$ *from* $\mathcal{S}_l \in \mathcal{O}_t$ *to* $\mathcal{S}_k \in \mathcal{O}_t \cup \{\emptyset\}$, $\sigma_{l\mapsto k}(\mathfrak{P}) : \mathcal{S}_l \mapsto \mathcal{S}_l/\mathfrak{P}$ *and* $\mathcal{S}_k \mapsto \mathcal{S}_k \cup \mathfrak{P}$.

Note 1: when $\mathcal{S}_k = \emptyset$, $\sigma_{l\mapsto k}(\mathfrak{P})$ leads to a new coalition forming. That special process is also noted $\sigma_{l\mapsto\emptyset}(\mathfrak{P})$.

Note 2: when $\mathcal{S}_k = \mathcal{S}_l$, $\sigma_{l\mapsto k}(\mathfrak{P})$ leads to coalition holding, i.e., no change compared to previous decisions. That special process is also noted $\sigma_{l\mapsto l}(\mathfrak{P})$.

In order to determine which process should be executed, participating in aother coalition, establishing a new coalition, or keeping constant, we define a reference indicator — switch operation gain — as follow:

Definition 5. *The switch operation gain* $g(\sigma_{l\mapsto k}(\mathfrak{P}))$ *associated with* $\sigma_{l\mapsto k}(\mathfrak{P})$ *is defined as:*

$$g\left(\sigma_{l\mapsto k}\left(\mathfrak{P}\right)\right) = \mathcal{U}_t^{\mathfrak{P}}\left(\mathcal{S}_k^{(t-1)} \cup \mathfrak{P}\right) - \mathcal{U}_t^{\mathfrak{P}}\left(\mathcal{S}_l^{(t-1)}\right).$$

In this paper, the utility function $\mathcal{U}$ is designed according to nodes' social features. Each node prefers forming in a coalition with nodes who have similar social behaviors, formally, the $\mathcal{U}^{\mathfrak{P}}$ is defined as:

$$\mathcal{U}^{\mathfrak{P}}\left(\mathcal{S}_l\right) = \frac{|\mathcal{S}_l \cap \mathcal{H}\{\mathfrak{P}\}|}{|\mathcal{S}_l|},$$

where $|\cdot|$ denotes the number of elements in a set. Noted that in our model, the cooperation is preferable for each player and no one wants to stay in alone. But if someone is unsatisfied with its current coalition and finds some peers who also unsatisfied with their own coalitions, it can establish a coalition and wait for peers to join in. The satisfied condition C is defined as the utility of minimum coalition unit, formally, $C = \frac{|\{\mathfrak{P},\mathfrak{Q}\}\cap\mathcal{H}\{\mathfrak{P}\}|}{|\{\mathfrak{P},\mathfrak{Q}\}|} = 1/2$, $\mathfrak{Q} \in \mathcal{H}\{\mathfrak{P}\}$. On the basis of this condition, we can derive the convergence of coalitions.

Proposition 1. *The coalitions O_t is convergence: the maximum coalition $\mathcal{S}_m$ is no more than* $2\times\max\{|\mathcal{H}\{\mathfrak{P}\}|\}$, $\mathfrak{P}\in\mathcal{S}_m$ *elements and the stable coalition $\mathcal{S}_s$ is no more than* $2\times\min\{|\mathcal{H}\{\mathfrak{P}\}|\}$, $\mathfrak{P}\in\mathcal{S}_s$ *elements.*

Proof. See Appendix A. □

3.2 Algorithm Design

In this section, a distributed coalition evolution algorithm is proposed to cluster the network. The detail parameters design is displayed in section 4, and the experienced results of this algorithm are displayed in section 5.

The proposed algorithm is based on switch operations. Specifically, every time when any player $\mathfrak{P}$ meets another player $\mathfrak{Q}$, they invite each other to join in their own coalition. Then each node needs to make a decision according to the receiving invitations (i.e. strategies). Notice that even some nodes do not receive an invitation, they still have one strategy — leaving the current coalition and establishing a new. The detailed algorithm is shown in Algorithm 1 and mainly consists of four parts:

- *Collection of candidate strategies* (line 1-3): when any player meets the player $\mathfrak{P}$, they will send out invitations on behalf of their coalitions. All invitations, as well as its current coalition, constitute the strategy profile $\mathcal{P}_t^{\mathfrak{P}}$ of player $\mathfrak{P}$ at this time.
- *Selection of best strategy* (line 4): The best switch operation strategy $\sigma_{l\mapsto k^*}(\mathfrak{P})$ is determined according to the switch operation gain of each strategy.
- *Implementation of switch operation* (line 5-10): According to the determined strategy, each node implements a relative switch operation, i.e., transferring to a new coalition, establishing a new coalition or staying constant. Note that even there is no need to leave the current coalition, each node will update some information associated with the existing coalition in order to future games.
- *Intention of leaving* (line 11-12): if there exist some players participating in one coalition but dissatisfying about it, then they will send out a signal $\{\emptyset\}$ — intention to form a new coalition — to seek peers.

4 THEORY ANALYSIS AND DESIGN

The social feature is the basis of algorithm 1 design but is not obvious for each node in the distributed system. In this section, we analyze the social nature of networks from the decentralized perspective, and propose a distributed metric for social relation extracting and a distributed measure for community detecting.

Algorithm 1: Clustering algorithm

Data: Current coalition $\mathcal{S}_l$ of player $\mathfrak{P}$
Result: Imminent coalition $\mathcal{S}_l$ of player $\mathfrak{P}$
// Collecting candidate strategies
1 Set $\mathcal{P}_t^{\mathfrak{P}}=\mathcal{S}_l$
2 **foreach** $\mathfrak{Q}\in\{v_i|\ (v_i,\mathfrak{P})_t\in\mathcal{E}_t\}$ **do**
3 $\quad\mathcal{P}_t^{\mathfrak{P}}=\mathcal{P}_t^{\mathfrak{P}}\cup\mathcal{S}_k$, $\mathcal{S}_k$ is the coalition invited by player $\mathfrak{Q}$.
// Selecting one best strategy
4 Find $(k^*,\mathcal{S}_{k^*})$ such that $g(\sigma_{l\mapsto k^*}(\mathfrak{P}))\geq g(\sigma_{l\mapsto k}(\mathfrak{P}))$, $\forall\mathcal{S}_k\in\mathcal{M}$
// Implementing switch operation
5 **if** $k^*\mapsto l$ **then**
6 $\quad$remain current coalition and update coalition information, $\mathcal{S}_l=\mathcal{S}_l^*$
7 **else if** $k^*\mapsto\emptyset$ **then**
8 $\quad$establish a new coalition, $\mathcal{S}_l=\{\mathfrak{P}\}$
9 **else**
10 $\quad$participate in a new coalition, $\mathcal{S}_l=\mathcal{S}_{k^*}$
// Forming a leaving signal
11 **if** *leaving-condition C is satisfied* **then**
12 $\quad\mathcal{S}_l=\mathcal{S}_l\cup\{\emptyset\}$

4.1 Relation Metric

To extract the social relations among nodes, the straight idea is to find some features to represent them. Original methods compare total contact duration (TCD), contact frequency (CF), maximum contact time (Max-CT) or minimum separation time (Min-ST), etc, to differentiate the relations arising from different encounter patterns. However, the straightforward usages of those metrics are unable to yield accurate results [10]. Consequently, the methods of mixing several features are prevalent, e.g., wide-used ASP (average separation period)-based methods [10, 18]. Those methods evaluate the effect of TCD and CF by the aid of the ASP, and contact irregularities by the aid of a penalty factor. However, the ASP is unable to accurately reflect the contact information. For example, consider the three cases in fig. 1, the ASP-based methods are unable to differentiate them due to the same ASP, while case (a) is preferable to case (b) due to the stable and long-last contacts, and case (b) is preferable to case (c) due to the regular contact duration.

In order to quantify the relations more accurately, we take the aforementioned three factors: frequency, duration, and regularity, into consideration from the contact perspective and propose a new metric based on power function. Formally,

$$w_{i,j}=f\left(\Psi\left(v_i,v_j\right)\right)=\frac{\sum_{\psi_k\in\Psi(v_i,v_j)}s(\psi_k)}{s(T)} \tag{1}$$

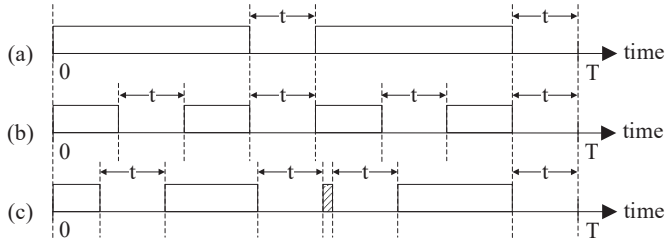


Figure 2: Three contact cases between nodes during $[0, T]$, where shaded boxes represent the contact durations.

where $s(\cdot)$ is a power function, $s(\psi_k) = {\psi_k}^{\alpha}$, $\alpha \in [0, 1]$, and $\Psi(v_i, v_j) = \{\psi_k\}$ is the contact records of nodes v_i, v_j during recording time $[0, T]$, $0 \le \psi_k \le T$, $0 \le \sum_{\psi_k \in \Psi(v_i, v_j)} \psi_k \le T$. Since power function is monotonically increasing concave function over the coordinate origin, we attain Theorem 1, Theorem 2 and Theorem 3.

Theorem 1. *$w_{i,j}$ is positively correlated with the frequency, the longevity, and the regularity.*

Proof. See Appendix B. □

Theorem 2. *Parameter α can adjust the preference between frequency and longevity. When α is close to 0, $w_{i,j}$ is more affected by frequency, while when α is close to 1, $w_{i,j}$ is more affected by longevity.*

Proof. See Appendix C. □

Theorem 3. *$w_{i,j}$ converges to [0,1] and the closer relationship yield out bigger $w_{i,j}$.*

Proof. See Appendix D. □

Metric 1 can be easy to extend to distributed system (proved in Appendix E), expressed as follow:

$$w_{i,j}^{(n)} = \frac{s\left(T^{(n-1)}\right) \cdot w_{i,j}^{(n-1)} + s(\psi_k)}{s\left(T^{(n)}\right)}, \tag{2}$$

where $T^{(n-1)}$ and $T^{(n)}$ are the last and current encounter-end time, $w_{i,j}^{(n-1)}$ and $w_{i,j}^{(n)}$ are the the last and current social relation.

4.2 Community Detection

According to the proposed metric, each node in the network can acquire social relations with its adjacent nodes. Since the value of the metric is proportional to the social relations, we say the nodes with high social relations to a node are its intimate neighbors (or apparent influence sphere), which is denoted to

$$\mathcal{M}_A(v_i) = \{v_j | w_{i,j} > \tau\}, \tag{3}$$

where τ is the threshold of intimate relations. However, except the apparent members, there are many potential nodes that can indirectly influence each other through the mutual intimate nodes as long as mutual intimate nodes are enough more. Consequently, we defined the potential members as

$$\mathcal{M}_B(v_i) = \{v_j | \, |\mathcal{M}_A(v_i) \cap \mathcal{M}_A(v_j)|,\ v_j \notin \mathcal{M}_A(v_i)| > \delta\}, \tag{4}$$

where δ is the threshold of mutual intimate nodes number. Since the relation distribution is approximated by a power law, we can attain Proposition 2.

Proposition 2. *The threshold of number of mutual intimate nodes δ is equal to $\lceil \log_{1-\beta^2}(1-\tau) \rceil$, $\beta = \frac{\int_C w \cdot g(w)\, dw}{\int_C g(w)\, dw}$, where $\lceil \cdot \rceil$ represents rounding element to the nearest integer greater than or equal to that element, $g(w) = \frac{b}{a}\left(\frac{w}{a}\right)^{b-1} e^{(-w/a)^b}$ is the distribution of the relation and $C = \{w | \, w > \tau\}$ is the integral interval.*

Proof. Given two nodes v_i and v_j in the network, they are not imitate but have $|\mathcal{M}_A(v_i) \cap \mathcal{M}_A(v_j)|$ mutual intimate adjacencies. Therefore, the two can influence each other through mutual intimate adjacencies. The weight of the potential relations v_i and v_j can be derived as

$$w_{i,j}^{p} = 1 - \prod_{v_k \in \mathcal{M}_A(v_i) \cap \mathcal{M}_A(v_j)} \left(1 - w_{i,k} \cdot w_{k,j}\right). \tag{5}$$

Here, the $w_{i,k}$ and $w_{k,j}$ are the social relations between v_i, v_j with the intermediate nodes, actually, which are two mutually independent random variables subject to a power law distribution restricted to $[\tau, 1]$. Most existing researches utilize Exponential distribution to fit user's social relations [7, 8], but this paper utilizes Weibull distribution, which is proved more accurate in section 5. Weibull distribution is expressed as

$$g(w) = \frac{b}{a}\left(\frac{w}{a}\right)^{b-1} e^{(-w/a)^b}, \tag{6}$$

where $a > 0$ is shape parameter and $b > 0$ is the scale parameter. Consequently, the expected value of w_{v_i, v_j}^{p} can express as

$$\begin{aligned} \mathbb{E}\left(w_{i,j}^{p}\right) &= \mathbb{E}\left(1 - \prod_{v_k \in \mathcal{M}_A(v_i) \cap \mathcal{M}_A(v_j)} \left(1 - w_{i,k} \cdot w_{k,j}\right)\right) \\ &= 1 - \prod_{v_k \in \mathcal{M}_A(v_i) \cap \mathcal{M}_A(v_j)} \left(1 - \mathbb{E}\left(w_{i,k}\right) \cdot \mathbb{E}\left(w_{k,j}\right)\right) \\ &= 1 - (1 - \beta^2)^{\eta}, \end{aligned} \tag{7}$$

where $\eta = |\mathcal{M}_A(v_i) \cap \mathcal{M}_A(v_j)|$ and β is the expected value of $w_{i,j}$,

$$\beta = \frac{\int_C w g(w)\, dw}{\int_C g(w)\ dw},\ b > 0,\ C = \{w | \, w > \tau\}. \tag{8}$$

We say node v_i and node v_j are potential intimate when $\mathbb{E}\left(w_{i,j}^p\right)$ is larger than intimate threshold τ, that is,

$$\mathbb{E}\left(w_{i,j}^p\right) = 1 - (1-\beta^2)^\eta > \tau, \tag{9}$$

$$\eta > \log_{1-\beta^2}(1-\tau). \tag{10}$$

Thus, the threshold of the number of mutual intimate nodes is set as the rounded-up number of η, i.e., $\delta = \lceil \log_{1-\beta^2}(1-\tau) \rceil$. □

Consequently, the community of each node is structured as its apparent and potential influence sphere and denoted as

$$\mathcal{H}_t\{v_i\} = \{v_j | \mathfrak{C}_{v_j}(w_t(v_i, v_j)), v_j \in \mathcal{W}_t\{v_i\}_{1\times M}\}\},$$
$$\mathfrak{C}_{v_j}(w_t(v_i, v_j)) := v_j \in \mathcal{M}_B(v_i) \cup \mathcal{M}_A(v_i). \tag{11}$$

5 EXPERIMENT RESULTS

In order to verify the performance of the proposed algorithm, in this paper, we do the simulations on real-world mobility networks and make comparisons with two existing algorithms: CLQ [12], VOTE [11].

5.1 Dataset

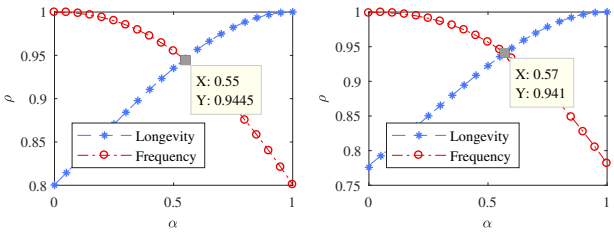
The simulations in this work are based on two widely used real-world datasets: *(i)* MIT Reality Mining Data (Reality) [3], and *(ii)* UCSD Wireless Topology Discovery Trace (WTD) [13]. In Reality, Bluetooth data are recorded by 97 smartphones deployed on students and staff at MIT over 246 days. In WTD, WiFi data are recorded by 275 PDAs carried by freshmen students at UCSD over 11 weeks. The details about the datasets are illustrated in Table. We extract the contact records from the partial data of the two datasets for our simulations. Each processed records includes the start and end time of each encounter and the IDs of the nodes in contact.

Table 1: Characteristics of Two Datasets

Dataset	Reality	WTD
Device	Phone	PAD
Network type	Bluetooth	WiFi
Contact type	direct	Ap-based
Duration (days)	246	77
Number of nodes	97	275
Number of contacts	54,667	135,364

5.2 Social relation Analysis

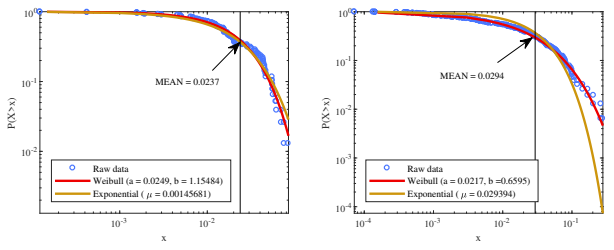
First, we verify the flexibility of social-relation-extraction algorithm, that is, the weight of social relationship can be adjusted by parameter α. The results are shown in Fig. 3, where the X-axis is the parameter α while the Y-axis is the Pearson correlation coefficient between social relations and contact longevity as well as frequency. Both experiments on Reality and WTD, it is obvious that with parameter α increasing, the bias is shifted from frequency to longevity. In subsequent experiments, WLOG, we regard the factor of frequency and longevity as the same, i.e., setting $\alpha = 0.55$ in Reality experiments and setting $\alpha = 0.57$ in WTD experiments.



(a) MIT (b) WTD

Figure 3: The flexibility analysis of social-relation-extraction metric in both datasets.

Second, the distributions of social relations are analyzed and shown in Fig. 4, where Fig. 4a is for Reality while Fig. 4b is for WTD. The distributions of both relations are fitted by Experimental distribution (most researchers used) and Weibull distribution. No matter which dataset shows that Weibull distribution has better fitting accuracy. Noted that even though some data subject to Exponential distribution, the Weibull distribution can also work well, just setting scale parameter $a = 1$. In subsequent experiments, WLOG, we set the intimate threshold τ is equal to the mean of social relations. Thus the threshold of number of mutual intimate nodes δ can be calculated according to Proposition 2, tuning out $\delta = 13$ in Reality, $\delta = 6$ in WTD.

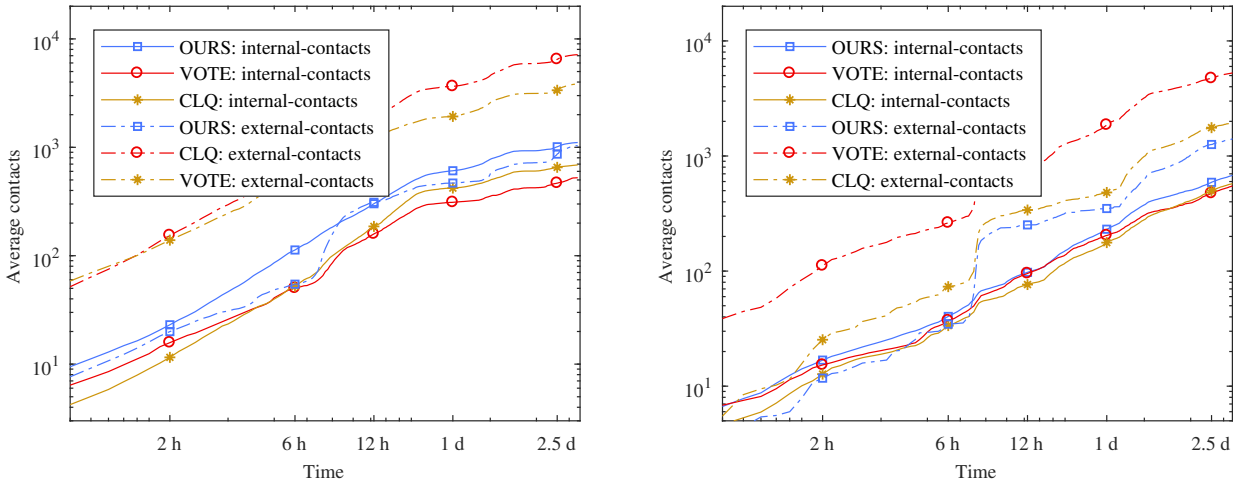


(a) MIT (b) WTD

Figure 4: Distribution analysis of social relations in both datasets.

5.3 Cluster verification

In order to verify the performance of proposed methods in clustering, we do the comparisons with two existing methods, CLQ and VOTE. It is immediate that for an ideal cluster

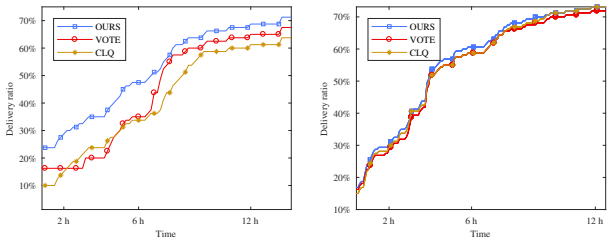


(a) MIT (b) WTD

Figure 5: Community Properties.

structure, there is high internal-contact densities and low external-contact densities. Thus we record the average internal and external contacts in the clusters detected by three methods respectively, as shown in Fig. 5. We can see the cluster structure detected by our method has the maximum internal contacts while minimal external contacts, especially in Reality dataset, most of the time there are more internal contacts than external contacts as shown in Fig. 5a. It well illustrates our good performance in clustering.

In addition, we also test our model in data dissemination occasion, where there are few nodes (called seeds) retrieving data from BS through cellular links and others retrieved data from neighbors through D2D links. Noted that seeds are selected based on cluster structure — the most influenced node in a community is selected as seed — and the influence is measured by degree centrality. The results are shown in Fig. 6. Obviously, both in Reality and WTD, the usage of our method can spread the content faster and more widely.



(a) MIT (b) WTD

Figure 6: Dissemination simulations base on Reality and WTD.

6 CONCLUSION

In this paper, we formulate the clustering process as a dynamic game with incomplete information and propose a distributed clustering algorithm. In order to partition stable clusters, social nature is utilized. Specifically, a metric to extract social relations is proposed to represent the similarities of mobility among nodes and a measure to detect social community is proposed to represent the influence sphere of each node. Through theoretical derivations and experimental verification, the proposed metric can well extract the social natures and have good flexibility, and the proposed clustering algorithm shows superior performance in distributed clustering. Despite the promising results, our model is more suitable for delay-insensitive services. In the future, we will attempt to leverage the results of our framework — social clustering — and combine real-time information to redesign the low-latency systems.

APPENDIX

Appendix A: Proof of the Proposition 1

For a stable coalition, no one internal is unsatisfied with this coalition, that is, $\forall \mathfrak{P} \in \mathcal{S}_l$, $\frac{|\mathcal{S}_l \cap \mathcal{H}\{\mathfrak{P}\}|}{|\mathcal{S}_l|} \geq 1/2$ so that

$$|\mathcal{S}_l| \leq |\mathcal{S}_l \cap \mathcal{H}\{\mathfrak{P}\}|, \forall \mathfrak{P} \in \mathcal{S}_l. \tag{A.1}$$

Thus maximum quantity of stable coalition $|\mathcal{S}_s| \leq 2 \times \min\{|\mathcal{H}\{\mathfrak{P}\}|\}$. Actually, not all players met the satisfied condition in a coalition. There exist some players, who are unsatisfied with their existing coalition but have not found a preferable one, having to stay in the current coalition, thus extending coalition quantity. However, the maximum element quantify is no more than $2\times\max\{|\mathcal{H}\{\mathfrak{P}\}|\}$. Because if $\forall\mathfrak{P} \in \mathcal{S}_l$, $\frac{|\mathcal{S}_l \cap \mathcal{H}\{\mathfrak{P}\}|}{|\mathcal{S}_l|} < 1/2$, the close partners will establish a new when they contact next time.

Appendix B: Proof of the Theorem 1

Since $s(\cdot)$ is a monotonically increasing concave function over the coordinate origin, social-relation-extraction metric w is positively correlated with the frequency, the longevity, and the regularity of contacts,

$$w_{i,j} = \frac{\sum_{\psi_k \in \Psi(v_i, v_j)} s(\psi_k)}{s(T)}. \tag{B.1}$$

Here, the frequency $\mathcal{F}$, the longevity $\mathcal{L}$, and the regularity $\mathcal{R}$ are respectively denoted as

$$\mathcal{F}_{i,j} = |\Psi(v_i, v_j)|, \tag{B.2}$$

$$\mathcal{L}_{i,j} = \sum_{k=1}^{\mathcal{F}_{i,j}} \psi_k, \tag{B.3}$$

$$\mathcal{R}_{i,j} = \frac{\sum_{k=1}^{\mathcal{F}_{i,j}} \|\psi_k - \zeta\|}{\mathcal{L}_{i,j}}, \tag{B.4}$$

where ζ is an arbitrary number in $\Psi(v_i, v_j)$. Proceeding stepwise, we shall prove:

1) Let $\mathcal{X} = \{x_1, x_2, \cdots, x_N\}$ be finitely non-repeating points from the domain of $s(\cdot)$, $I \subseteq \mathbb{R}$. $\forall\zeta \in \mathcal{X}$, the concave function $s(x_n)$ can be expressed as $s(\zeta) = v\zeta + \mu + \sum_{n=1}^{N} \kappa_n \|x_n - \zeta\|$, where v, μ are two real constants and κ_n are series of non-positive constants.
2) w is positively correlated with $\mathcal{R}$.
3) w is positively correlated with $\mathcal{L}$.
4) w is positively correlated with $\mathcal{F}$.

1) we WLOG assume that $x_1 < x_2 < \cdots < x_N$ (if not, sort $\mathcal{X}$ first). Then, for every $n \in [1, N]$, according to the proof, "Absolute values interpolate convex functions" in [6], the function $s(\cdot)$ can be expressed as

$$s(\zeta) = v\zeta + \mu + \sum_{n=1}^{N} \kappa_n \|x_n - \zeta\|, \ \forall\zeta \in \mathcal{X}, \tag{B.5}$$

where

$$v = e[x_2, x_1] + \sum_{n=1}^{N} \kappa_n, \tag{B.6}$$

$$\mu = s(x_1) - x_1 \cdot e[x_2, x_1] - \sum_{n=1}^{N} \kappa_n x_n, \tag{B.7}$$

$$\kappa_n = \begin{cases} \frac{1}{2}e[x_{n+1}, x_n] - \frac{1}{2}e[x_n, x_{n-1}] & 0 < n < N, \\ 0 & \text{otherwise.} \end{cases} \tag{B.8}$$

$$e[x_m, x_n] = \frac{s(x_m) - s(x_n)}{x_m - x_n}. \tag{B.9}$$

Since the function $s(\cdot)$ is concave on I, we have

$$\frac{\frac{1}{x_{n+1}-x_n}s(x_{n+1}) + \frac{1}{x_n - x_{n-1}}s(x_{n-1})}{\frac{1}{x_{n+1}-x_n} + \frac{1}{x_n - x_{n-1}}} \leq s\left(\frac{\frac{1}{x_{n+1}-x_n}x_{n+1} + \frac{1}{x_n - x_{n-1}}x_{n-1}}{\frac{1}{x_{n+1}-x_n} + \frac{1}{x_n - x_{n-1}}}\right). \tag{B.10}$$

Since $\frac{\frac{1}{x_{n+1}-x_n}x_{n+1} + \frac{1}{x_n - x_{n-1}}x_{n-1}}{\frac{1}{x_{n+1}-x_n} + \frac{1}{x_n - x_{n-1}}} = x_n$, (B.10) simplifies to

$$\frac{\frac{1}{x_{n+1}-x_n}s(x_{n+1}) + \frac{1}{x_n - x_{n-1}}s(x_{n-1})}{\frac{1}{x_{n+1}-x_n} + \frac{1}{x_n - x_{n-1}}} \leq s(x_n), \tag{B.11}$$

so that

$$\frac{s(x_{n+1}) - s(x_n)}{x_{n+1} - x_n} \leq \frac{s(x_n) - s(x_{n-1})}{x_n - x_{n-1}}, \tag{B.12}$$

i.e., $e[x_{n+1} - x_n] - e[x_n - x_{n-1}] \leq 0$, $\kappa_n \leq 0$. It is immediate to prove v and μ are real constants since $\mathcal{X}$ is a finite sequence. Thus, the proof of **1)** is completed.

2) Let $\mathcal{X} = \{x_1, x_2, \cdots, x_N\}$, $\mathcal{Y} = \{y_1, y_2, \cdots, y_M\}$ be $M + N$ finite points on I. In order to confirm the correlation between the metric w with the regularity $\mathcal{R}$, we WLOG assume $\mathcal{F}$ and $\mathcal{L}$ are the same for the two sequences, i.e., $\mathcal{F}^{\mathcal{X}} = \mathcal{F}^{\mathcal{Y}}$ and $\mathcal{L}^{\mathcal{X}} = \mathcal{L}^{\mathcal{Y}}$. Assuming that $\mathcal{R}^{\mathcal{X}} \leq \mathcal{R}^{\mathcal{Y}}$, we have

$$\frac{\sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} \|x_n - \zeta\|}{\mathcal{L}^{\mathcal{X}}} \leq \frac{\sum_{m=1}^{\mathcal{F}^{\mathcal{Y}}} \|y_m - \zeta\|}{\mathcal{L}^{\mathcal{Y}}}, \ \forall\zeta \in \{\mathcal{X}, \mathcal{Y}\}. \tag{B.13}$$

According to **1)**, there exist two real constants v and μ and $\mathcal{F}^{\mathcal{Y}} + \mathcal{F}^{\mathcal{Y}}$ non-positive constants κ_ι such that

$$s(\zeta) = v\zeta + \mu + \sum_{\iota=1}^{\mathcal{F}^{\mathcal{Y}} + \mathcal{F}^{\mathcal{Y}}} \kappa_\iota \|z_\iota - \zeta\|, \ \forall\zeta \in \mathcal{Z} = \{\mathcal{X}, \mathcal{Y}\} \tag{B.14}$$

Consequently, we have (B.15) to (B.18), which means the metric value increases with the irregularity of interactions decreases, i.e., w is positively correlated with $\mathcal{R}$.

3) Let $\mathcal{X} = \{x_1, x_2, \cdots, x_N\}$, $\mathcal{Y} = \{y_1, y_2, \cdots, y_M\}$ be $M + N$ finite points on I. In order to confirm the correlation between the metric w with the regularity $\mathcal{L}$, we WLOG assume $\mathcal{F}$ and $\mathcal{R}$ are the same for the two sequences, i.e., $\mathcal{F}^{\mathcal{X}} = \mathcal{F}^{\mathcal{Y}}$ and

$$\frac{\sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} \|x_n - \zeta\|}{\mathcal{L}^{\mathcal{X}}} = \frac{\sum_{m=1}^{\mathcal{F}^{\mathcal{Y}}} \|y_m - \zeta\|}{\mathcal{L}^{\mathcal{Y}}}, \ \forall\zeta \in \{\mathcal{X}, \mathcal{Y}\}. \tag{B.19}$$

When $\mathcal{L}^{\mathcal{X}} \leq \mathcal{L}^{\mathcal{Y}}$, it is immediate that $w^{\mathcal{X}} \leq w^{\mathcal{Y}}$ according to (B.17) and (B.18), otherwise $w^{\mathcal{X}} > w^{\mathcal{Y}}$. Thus, the proof of **3)** is completed.

$$w^{\mathcal{X}} = \frac{\sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} s\left(x_n\right)}{s\left(T\right)} = \frac{\sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} \left(\nu x_n + \mu + \sum_{\iota=1}^{\mathcal{F}^{\mathcal{X}}+\mathcal{F}^{\mathcal{Y}}} \kappa_\iota \|z_\iota - x_n\|\right)}{s\left(T\right)} \tag{B.15}$$

$$= \frac{\nu \cdot \sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} x_n + \sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} \mu + \sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} \sum_{\iota=1}^{\mathcal{F}^{\mathcal{X}}+\mathcal{F}^{\mathcal{Y}}} \kappa_\iota \|z_\iota - x_n\|}{s\left(T\right)} \tag{B.16}$$

$$= \frac{\nu \cdot \mathcal{L}^{\mathcal{X}}}{s\left(T\right)} + \frac{\mu \cdot \mathcal{F}^{\mathcal{X}}}{s\left(T\right)} + \sum_{\iota=1}^{\mathcal{F}^{\mathcal{X}}+\mathcal{F}^{\mathcal{Y}}} \left(\kappa_\iota \cdot \mathcal{L}^{\mathcal{X}} \cdot \sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} \frac{\|x_n - z_\iota\|}{s\left(T\right) \cdot \mathcal{L}^{\mathcal{X}}}\right) \tag{B.17}$$

$$\geq \frac{\nu \cdot \mathcal{L}^{\mathcal{Y}}}{s\left(T\right)} + \frac{\mu \cdot \mathcal{F}^{\mathcal{Y}}}{s\left(T\right)} + \sum_{\iota=1}^{\mathcal{F}^{\mathcal{X}}+\mathcal{F}^{\mathcal{Y}}} \left(\kappa_\iota \cdot \mathcal{L}^{\mathcal{Y}} \cdot \sum_{m=1}^{\mathcal{F}^{\mathcal{Y}}} \frac{\|y_m - z_\iota\|}{s\left(T\right) \cdot \mathcal{L}^{\mathcal{Y}}}\right) = \frac{\sum_{m=1}^{\mathcal{F}^{\mathcal{Y}}} s\left(y_m\right)}{s\left(T\right)} = w^{\mathcal{Y}} \tag{B.18}$$

$$\left(\text{since } \mathcal{L}^{\mathcal{X}} = \mathcal{L}^{\mathcal{Y}},\ \mathcal{F}^{\mathcal{X}} = \mathcal{F}^{\mathcal{Y}},\ \frac{\sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} \|x_n - \zeta\|}{\mathcal{L}^{\mathcal{X}}} \leq \frac{\sum_{m=1}^{\mathcal{F}^{\mathcal{Y}}} \|y_m - \zeta\|}{\mathcal{L}^{\mathcal{Y}}},\ \forall\, \zeta \in \{\mathcal{X},\ \mathcal{Y}\} \text{ and } \kappa_\iota \leq 0\right)$$

4) Let $\mathcal{X} = \{x_1, x_2, \cdots, x_N\}$ be N finite points on I. Given a new sequence $\mathcal{X}'$, which is extending of $\mathcal{X}$,

$$\mathcal{X}' = \{\underbrace{0, \cdots, 0}_{h}, x_1, x_2, \cdots, x_N\}. \tag{B.20}$$

It is easy to prove that

$$\frac{\sum_{n=1}^{\mathcal{F}^{\mathcal{X}}} \|x_n - \zeta\|}{\mathcal{L}^{\mathcal{X}}} \leq \frac{\sum_{m=1}^{\mathcal{F}^{\mathcal{X}'}} \|x'_m - \zeta\|}{\mathcal{L}^{\mathcal{X}'}},\ \forall\, \zeta \in \{0,\ \mathcal{X}\}. \tag{B.21}$$

Given another sequence $\mathcal{Y} = \{y_1, y_2, \cdots, y_M\}$ on I satisfying that $\mathcal{L}^{\mathcal{X}} = \mathcal{L}^{\mathcal{Y}}$, $\mathcal{R}^{\mathcal{X}} = \mathcal{R}^{\mathcal{Y}}$ and $M = h + N$. Consequently,

$$\frac{\sum_{m=1}^{\mathcal{F}^{\mathcal{Y}}} \|y_n - \zeta\|}{\mathcal{L}^{\mathcal{Y}}} \leq \frac{\sum_{m=1}^{\mathcal{F}^{\mathcal{X}'}} \|x'_m - \zeta\|}{\mathcal{L}^{\mathcal{X}'}},\ \forall\, \zeta \in \{0,\ \mathcal{X}, \mathcal{Y}\}. \tag{B.22}$$

Since $\mathcal{L}^{\mathcal{Y}} = \mathcal{L}^{\mathcal{X}'}$ and $\mathcal{F}^{\mathcal{Y}} = \mathcal{F}^{\mathcal{X}'}$, according to **2)**, we have $w^{\mathcal{Y}} \geq w^{\mathcal{X}'} = w^{\mathcal{X}'}$ (due to $s(0) = 0$). Thus, the proof of **4)** is completed.

Appendix C: Proof of the Theorem 2

When $\alpha = 0$, $w_{i,j} = \sum_{\psi_k \in \Psi(v_i, v_j)} (\psi_k)^\alpha / s\left(T\right) = \mathcal{F}_{i,j}$, while when $\alpha = 1$, $w_{i,j} = \sum_{\psi_k \in \Psi(v_i, v_j)} (\psi_k)^\alpha / s\left(T\right) = \mathcal{L}_{i,j}$. In addition, $\forall \alpha \in [0, 1]$, $\psi \in [1, \infty]$, $\frac{\partial \psi^\alpha}{\partial \alpha} > 0$, thus the bias will shift from frequency to longevity with α increasing.

Appendix D: Proof of the Theorem 3

According to Appendix B, it is immediate that value $w_{i,j}$ is positively correlated with the relationship nodes i and j. Since for any set of contact periods, $\sum(\psi_k)^\alpha \leq (\sum \psi_k)^\alpha$, $\alpha \in [0, 1]$ according to concavity, thus $w_{i,j}$ is no more than 1. While when there is not any contact between nodes, it is immediate that $w_{i,j}$ reaches to the minimum 0.

Appendix E: Proof of the Distributed Social-Relation-Extracting Metric

According to (2), we have

$$w_{i,j}^{(n)} = \frac{s\left(T^{(n-1)}\right) \cdot w_{i,j}^{(n-1)} + s(\psi_k)}{s\left(T^{(n)}\right)} \tag{E.1}$$

$$= \frac{s\left(T^{(n-1)}\right) \cdot \frac{s\left(T^{(n-2)}\right) \cdot w_{i,j}^{(n-2)} + s(\psi_{k-1})}{s\left(T^{(n-1)}\right)} + s(\psi_k)}{s\left(T^{(n)}\right)} \tag{E.2}$$

$$= \frac{s\left(T^{(n-2)}\right) \cdot w_{i,j}^{(n-2)} + s(\psi_{k-1}) + s(\psi_k)}{s\left(T^{(n)}\right)} \tag{E.3}$$

$$= \cdots \tag{E.4}$$

$$= \frac{s(\psi_1) + \cdots + s(\psi_{k-1}) + s(\psi_k)}{s\left(T^{(n)}\right)}, \tag{E.5}$$

which is equal to (1). Thus (2) is the distributed form of (1).

REFERENCES

[1] Dennis Baker and Anthony Ephremides. 1981. The architectural organization of a mobile radio network via a distributed algorithm. *IEEE Transactions on communications* 29, 11 (1981), 1694–1701.

[2] Ching-Chuan Chiang, Mario Gerla, and Lixia Zhang. 1998. Forwarding group multicast protocol (FGMP) for multihop, mobile wireless networks. *Cluster Computing* 1, 2 (1998), 187–196.

[3] Nathan Eagle, Alex Sandy Pentland, and David Lazer. 2009. Inferring friendship network structure by using mobile phone data. *Proceedings of the national academy of sciences* 106, 36 (2009), 15274–15278.

[4] Mario Gerla and Jack Tzu-Chieh Tsai. 1995. Multicluster, mobile, multimedia radio network. *Wireless networks* 1, 3 (1995), 255–265.

[5] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. 2008. Understanding individual human mobility patterns. *nature* 453, 7196 (2008), 779.

[6] Darij Grinberg. 2008. Generalizations of Popoviciu's inequality. *arXiv e-prints*, Article arXiv:0803.2958 (Mar 2008), arXiv:0803.2958 pages. arXiv:math.FA/0803.2958

[7] Pan Hui, Augustin Chaintreau, James Scott, Richard Gass, Jon Crowcroft, and Christophe Diot. 2005. Pocket switched networks

and human mobility in conference environments. In *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*. ACM, 244–251.

[8] Thomas Karagiannis, Jean-Yves Le Boudec, and Milan Vojnovic. 2010. Power law and exponential decay of intercontact times between mobile devices. *IEEE Transactions on Mobile Computing* 9, 10 (2010), 1377–1390.

[9] Hanh Le, Doan Hoang, and Ravi Poliah. 2008. S-Web: An efficient and self-organizing wireless sensor network model. In *International Conference on Network-Based Information Systems*. Springer, 179–188.

[10] Feng Li and Jie Wu. 2009. LocalCom: a community-based epidemic forwarding scheme in disruption-tolerant networks. In *2009 6th annual IEEE communications society conference on sensor, mesh and ad hoc communications and networks*. IEEE, 1–9.

[11] Fei Li, Shile Zhang, Xin Wang, Xiangyang Xue, and Hong Shen. 2004. Vote-based clustering algorithm in mobile ad hoc networks. In *International Conference on Information Networking*. Springer, 13–23.

[12] Raphaël Massin, Christophe J Le Martret, and Philippe Ciblat. 2017. A coalition formation game for distributed node clustering in mobile ad hoc networks. *IEEE Transactions on Wireless Communications* 16, 6 (2017), 3940–3952.

[13] Marvin McNett and Geoffrey M Voelker. 2005. Access and mobility of wireless PDA users. *ACM SIGMOBILE Mobile Computing and Communications Review* 9, 2 (2005), 40–55.

[14] Omer Narmanlioglu and Engin Zeydan. 2018. Mobility-aware cell clustering mechanism for self-organizing networks. *IEEE Access* 6 (2018), 65405–65417.

[15] Victor Sucasas, Ayman Radwan, Hugo Marques, Jonathan Rodriguez, Seiamak Vahid, and Rahim Tafazolli. 2016. A survey on clustering techniques for cooperative wireless networks. *Ad Hoc Networks* 47 (2016), 53–81.

[16] Xiaofei Wang, Yuhua Zhang, Victor CM Leung, Nadra Guizani, and Tianpeng Jiang. 2018. D2D big data: Content deliveries over wireless device-to-device sharing in large-scale mobile networks. *IEEE Wireless Communications* 25, 1 (2018), 32–38.

[17] Zhigang Wang, Lichuan Liu, MengChu Zhou, and Nirwan Ansari. 2008. A position-based clustering technique for ad hoc intervehicle communication. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38, 2 (2008), 201–208.

[18] Huan Zhou, Victor CM Leung, Chunsheng Zhu, Shouzhi Xu, and Jialu Fan. 2017. Predicting temporal social contact patterns for data forwarding in opportunistic mobile networks. *IEEE Transactions on Vehicular Technology* 66, 11 (2017), 10372–10383.