Received March 26, 2018, accepted April 17, 2018, date of publication April 30, 2018, date of current version May 24, 2018. *Digital Object Identifier* 10.1109/ACCESS.2018.2831247

A Time-Ordered Aggregation Model-Based Centrality Metric for Mobile Social Networks

HUAN ZHOU^{®1}, (Member, IEEE), MENGNI RUAN¹, CHUNSHENG ZHU^{®2}, (Member, IEEE), VICTOR C. M. LEUNG^{®2}, (Fellow, IEEE), SHOUZHI XU^{®1}, AND CHUNG-MING HUANG^{®3}, (Senior Member, IEEE)

¹College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China

²Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC V6T 1Z4, Canada

³Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 701, Taiwan

Corresponding author: Chunsheng Zhu (chunsheng.tom.zhu@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61602272 and in part by the Natural Science Foundation of Hubei Province of China under Grant 2017CFB594.

ABSTRACT How to measure the centrality of nodes is a significant problem in mobile social networks (MSNs). Current studies in MSNs mainly focus on measuring the centrality of nodes in a certain time interval based on the static graph that do not change over time. However, the network topology of MSNs is changing very rapidly, which is the main characteristic of MSNs. Therefore, it will not be accurate to measure the centrality of nodes in a certain time interval by using the static graph. To solve this problem, this paper first introduces a new centrality metric named cumulative neighboring relationship (CNR) for MSNs. Then, a time-ordered aggregation model is proposed to reduce a dynamic network to a series of time-ordered aggregation model, this paper proposes three particular time-ordered aggregation methods and combines with the proposed centrality metric CNR to measure the importance of nodes in a certain time interval. Finally, extensive trace-driven simulations are conducted to evaluate the performance of our proposed time-ordered aggregation model-based centrality metric time-ordered aggregation method can measure *TCNR* centrality in a certain time interval more accurately than other aggregation methods, and our proposed time-ordered aggregation model-based centrality metric *TCNR* outperforms other existing temporal centrality metrics.

INDEX TERMS Centrality, mobile social networks, dynamic network, time-ordered aggregation model, trace-driven simulation.

I. INTRODUCTION

FEE Access

Recently, with the rapid popularity of mobile devices (such as i-pad, PDAs, and smart-phones), Mobile Social Networks (MSNs) have began to emerge, which combine opportunistic mobile networks and social network analysis technologies together [1]–[3]. Specifically, social network analysis technologies (e.g., node centrality, community, similarity and so on) are utilized in MSNs to exploit mobile nodes' social relationships [4], [5]. The motivation is that the social features of mobile nodes stand for stable and long relationships among them, which can be used for designing efficient data forwarding strategies for MSNs. MSNs are promising to deal with data sharing and dissemination among mobile nodes in delay-tolerant scenarios by using mobile devices equipped with wireless interfaces (e.g., Bluetooth, Wi-Fi) when they are within the communication range of each other. Under this situation, an end-to-end transmission path between mobile nodes might not exist due to the lack of infrastructure and the time-varying network topology. Therefore, mobile nodes in MSNs employ a store-carry-and-forward scheme to forward messages [6]–[8].

To improve the performance of data forwarding in MSNs, a significant problem in the research of MSNs is to measure the centrality (importance) of nodes in the network [9]–[11]. For example, if we want to disseminate messages to more nodes in the network in a certain time interval, an intuitive way is to choose nodes which have higher centrality values as the source nodes. Considering the application, previous studies have proposed diverse centrality metrics to measure the relative importance of nodes in the network. However, the current studies in MSNs focused on measuring these centrality metrics in a certain time interval based on the static

graph that do not change over a long period of time. Actually, network topology in MSNs is changing very rapidly, which is driven by natural social behavior of people [12]. Therefore, it will not be accurate to measure the centrality of nodes in a certain time interval by using the static graph.

The initial attempts had tried to address this problem by introducing the time-ordered graph, but did not provide effective methods for MSNs [13]. In this paper, we first propose a new centrality metric named Cumulative Neighboring Relationship (CNR) based on the neighboring relationship of pair-wise nodes for MSNs. Then, we introduce a time-ordered aggregation model, which reduces a dynamic network to a series of time-ordered networks [14]. Specifically, there can be many different methods to aggregate time-ordered network graph and each has its advantages and disadvantages. Without loss of generation, we consider three particular time-ordered aggregation methods: the Average Time-ordered Aggregation Method, the Linear Time-ordered Aggregation Method, and the Exponential Time-ordered Aggregation Method, and combine with our proposed centrality metric CNR to measure the importance of nodes in a certain time interval. The Average Timeordered Aggregation Method aggregates temporal network graph by assigning equal weights to each time window. The Linear Time-ordered Aggregation Method assigns linearly decreased (or increased) weights to each time window, and the Exponential Time-ordered Aggregation Method assigns exponentially decreased (or increased) weights to each time window.

It is worth noticing that compared with nodes which are active at the end of a time interval, nodes active at the start time of the time interval can disseminate messages to other nodes in the network more quickly. Therefore, the centrality values at time windows which are closer to the start time of the time interval should be assigned with larger weights. Finally, the contributions of this paper are summarized as follows:

- To measure the importance of nodes in the network more accurately, we propose a new centrality metric named Cumulative Neighboring Relationship (CNR) based on the neighboring relationship of pair-wise nodes for MSNs.
- 2) We introduce a time-ordered aggregation model to reduce a dynamic network to a series of time-ordered networks. Based on the time-ordered aggregation model, we propose three particular time-ordered aggregation methods: the Average Time-ordered Aggregation Method, the Linear Time-ordered Aggregation Method, the Exponential Time-ordered Aggregation Method, and combine with our proposed centrality metric CNR to measure the importance of nodes in a certain time interval.
- 3) Extensive real trace-driven simulations are conducted to evaluate the performance of our proposed time-ordered aggregation model-based centrality metric *TCNR*.

The remainder of this paper is organized as follows. Section II introduces the related work. Section III introduces the network model related to this paper and the motivation of this paper. Section V introduces the time-ordered aggregation model, and three particular time-ordered aggregation methods. Through extensive real trace-driven simulations, Section VI evaluates the performance of our proposed timeordered aggregation methods under our proposed centrality metric CNR, and then compares our proposed time-ordered aggregation model-based centrality metric with other existing temporal centrality metric. Section VII concludes the paper.

II. RELATED WORK

Previous studies have proposed diverse centrality metrics to measure the relative importance of nodes in the network, such as Betweeness Centrality, Degree Centrality, Closeness Centrality, and so on [15]–[18]. These centrality metrics have been widely used to design efficient routing protocols for MSNs. For example, SimBet proposed in [19] combines the Betweenness centrality in the Ego network and similarity to select relays nodes, so as to increase the data forwarding performance in MSNs. Similarly, BUBBLE Rap proposed in [20] combines the Betweenness centrality in the Ego network with social communities to increase the data forwarding performance in MSNs. Gao et al. [22] proposed the Cumulative Contact Probability (CCP) as the centrality metric to select relay nodes for multicasting in MSNs. Fan et al. [23] proposed a centrality metric geocentrality to measure the user density of each geocommunity in MSNs. Based on the Degree centrality, Socievole and Rango [24] proposed a novel an Energy-aware Centrality- based Forwarding strategy (ECF) for MSNs. The proposed ECF protocol modulates Degree centrality with nodes' energy level so as to prolong the network lifetime. To improve the opportunistic forwarding efficiency, Yuan et al. [25] exploit the relative importance (called partial centrality) of a node with respect to a group of nodes, and design a new opportunistic forwarding strategy, Opportunistic Forwarding with Partial Centrality (OFPC). Furthermore, they also theoretically quantify the impact of the partial centrality on the data forwarding performance using graph spectrum. However, the current studies in MSNs focused on measuring these centrality metrics in a certain time interval based on the static graph that do not change over a long period of time, they do not consider the dynamic change of network topology. Therefore, it is not accurate to measure the centrality of nodes in a certain time interval by using the static network graph.

The initial attempts had tried to address this problem by introducing the time-ordered graph. Kim and Anderson [13] proposed temporal node centrality metrics (degree, closeness, betweenness) that captured the temporal characteristics of dynamic networks based on the time-ordered graph. However, social natures, movements and activity patterns which are strongly impacted by their social relationships are not considered in this paper. Similarly, Gao *et al.* [26] proposed a temporal evolution graph model to more accurately capture



FIGURE 1. Illustration of time windows in time interval $[t_s, t_e]$.

the topology dynamics of the mobile social network over time. Based on the proposed model, they explore human social relations and mobility patterns to redefine three common centrality metrics: degree centrality, closeness centrality and betweenness centrality. Zhou *et al.* [27] proposed a data forwarding strategy called TCCB based on the predicted temporal social contact patterns, e.g., temporal closeness and temporal centrality. Different from previous studies, in this paper we propose a new centrality metric by exploiting the social relationship between node pairs, and propose a more general time-ordered aggregation model based on the timeordered graph for MSNs.

III. PRELIMINARIES

In this section, we will first introduce the network model related to this paper, and then introduce the motivation of this paper.

A. NETWORK MODEL

In this part, we introduce the network model related to this paper. We assume that the time during which a network is observed is finite, from t_s to t_e . The dynamic network graph $G(t_s, t_e) = (V, E_{t_s, t_e})$ in the time interval $[t_s, t_e]$ consists of a set of vertices V and a set of temporal edges E_{t_s, t_e} , where stochastic contact process between a node pair $i, j \in V$ on a time interval $[t_a, t_b]$ ($t_s \leq t_a \leq t_b \leq t_e$) is modeled as an temporal edge $e_{t_a, t_b}^{ij} \in E_{t_s, t_e}$ [27], [28].

In this paper, we also focus on measuring centrality metrics based on the time-ordered graph. As shown in Fig. 1, the time interval $T = t_e - t_s$ is divided into fixed discrete time windows $\{1, 2, ..., n\}$. Here, $w = \frac{T}{n} = \frac{t_e - t_s}{n}$ is used to denote the size of each time window, expressed in some time unites (e.g., minutes or hours). In other words, a series of time-ordered graph, $G_1, G_2, ..., G_n$ can be used to represent the dynamic network graph $G(t_s, t_e)$, and G_k $(1 \le k \le n)$ represents the *k*-th temporal snapshot of the dynamic network graph $G(t_s, t_e)$ during the *k*-th time window.

B. MOTIVATION

In this part, we introduce the motivation of this paper. To make the motivation of this paper clearly, we use a simple example to show the difference between calculating centrality metric based on the static graph and the time-ordered graph. Fig. 2 shows the static graph and the time-ordered graph in G(1, 3), and w = 1, respectively. The left side in Fig. 2 shows the static graph in G(1, 3), and the right side in Fig. 2 shows the time-ordered graph in G(1, 3). Unlike the static graph, G(1, 3).



FIGURE 2. Illustration of the static graph and the time-ordered graph.

Time Interval	Time-ordered	Static
G_1	1	
G_2	0	0
G ₃	0	
G(1,3)	1	0

FIGURE 3. Comparison between calculating Betweeness centrality based on the static graph and the time-ordered graph.

a series of time-ordered graphs, G_1 , G_2 and G_3 represent temporal edge relationships among nodes A, B, C, and D.

Taking node D in Fig. 2 and Betweeness centrality as an example, as shown in Fig. 3, it can be found that it is totally different to calculate the Betweeness centrality based on the static graph and the time-ordered graph. As if we calculate using the time-ordered graph, the Betweeness centrality value of node D in G_1 is 1, then the aggregate Betweeness centrality value of node D in G(1, 3) is 1, but if we calculate using the static graph, the Betweeness centrality value of node D in G(1, 3) is 0. Therefore, how to accurately measure the importance of of nodes in a certain time interval is a significant problem for MSNs. In the next sections, we will first propose a new centrality metric, and then combine it with the proposed time-ordered aggregation model to design a time-ordered aggregation model-based centrality metric for MSNs.

IV. A NEIGHBORING RELATIONSHIP-BASED CENTRALITY METRIC

In this section, to measure the importance of nodes in the network more accurately, we propose a new centrality metric named Cumulative Neighboring Relationship (CNR) based on the neighboring relationship of pair-wise nodes for MSNs. Centrality refers to a group of metrics that aim to quantify the "importance" or "influence" of a particular node (or group) within a network.

In MSNs, nodes in the network usually have the knowledge of their past contact information with other nodes, also called contact history, e.g., inter-contact time, contact duration, separating time, and so on [29]–[31]. Since the separating time contains both the frequency and duration of the pair-wise contact, it will be more accurate to depict the neighboring relationship by using the separating time, than only by using the inter-contact time or contact duration. It is easy to find that a shorter average separating time reflects a closer relationship. At the same time, the variance of the separating time



FIGURE 4. Illustration of the separating time.

is also recorded to reflect the irregularity of the relationship. Therefore, we deduce a single metric called Neighboring Relationship (NR) based on the average separating time and the variance of the separating time, which depicts the social relationship between each pair of nodes in MSNs.

We use α_{ij} and β_{ij} to represent the average separating time and the variance of the separating time between two nodes *i* and *j* in a certain time window, respectively. A simple example is given in Fig. 4 to show the separating time S(ij)between two nodes *i* and *j*. Then, it is easy to obtain that $\beta_{ij} = [S_1(ij) + S_2(ij) + S_3(ij)]/3$, and $\beta_{ij} = [|\alpha_{ij} - S_1(ij)| + |\alpha_{ij} - S_2(ij)| + |\alpha_{ij} - S_3(ij)|]/3$, where $S_1(ij), S_2(ij)$, and $S_3(ij)$ are separating times between nodes *i* and *j* in the time window.

An exponential function is used to normalize the average separating time in the time window, and the resulting metric is denoted as *ASep*, which is expressed as:

$$ASep_{ij} = e^{1 - \frac{\alpha_{ij}}{w}},\tag{1}$$

An exponential function is also used to normalize the variance of the separating time in the time window, and the resulting metric is denoted as *VSep*, which is expressed as follows:

$$VSep_{ij} = e^{\frac{\beta_{ij}}{w}},\tag{2}$$

Taking both the average separating time $ASep_{ij}$ and the variance of the separating time $VSep_{ij}$ into consideration, the Neighboring Relationship *NR* between two nodes *i* and *j* is expressed as follows:

$$NR(ij) = ASep_{ij} - gVSep_{ij},$$
(3)

where g is a penalty parameter in the range of [0, 1], which decides the penalty of the irregularity metric and should be sufficiently small.

The above definition only considers the direct neighboring relationship. Actually, some nodes may never encounter each other before, so they do not have a direct neighboring relationship. However, there may exist a multi-hop path between them, via which a data can be delivered easily. Therefore, the definition of the neighboring relationship should not exclude such indirect neighboring relationships. Considering the k-hop neighbors, the neighboring relationship NR between two nodes i and j can be expressed as follows:

$$NR(ij) = max_{p \in P} \{ \prod_{(u,v) \in p} NR(uv) \},$$
(4)

where *P* is the set of *k*-hop paths from node *i* to node *j*, and *p* is one of the paths from *i* to *j*.

VOLUME 6, 2018

The centrality value of a node measures the structural importance of the node in the network. Nodes with higher centrality values have stronger capability of connecting other nodes in the network. Taking all nodes in the network into account, a new centrality metric named Cumulative Neighboring Relationship (CNR) based on the neighboring relationship of pair-wise nodes is proposed for MSNs, shown as follows:

$$CNR(i) = \frac{1}{N-1} \sum_{j \in V, j \neq i} NR(ij)$$
(5)

V. TIME-ORDERED AGGREGATION MODEL-BASED CENTRALITY METRIC

In this section, we introduce the proposed time-ordered aggregation model-based centrality metric *TCNR*, which combines the time-ordered aggregation model with the proposed centrality metric CNR. The time-ordered aggregation model takes a variable vector CNR(i) as input and returns a single numeric evaluation of the temporal information contained in CNR(i). The centrality vector CNR(i) of a node *i* in $G(t_s, t_e)$ is denoted as $CNR(i) = (CNR_1(i), CNR_2(i), \ldots, CNR_n(i))$. Then, for an input vector CNR(i) with weight vector θ , we use $TCNR_{G(t_s, t_e)}(i)$ to represent the time-ordered aggregation model-based CNR in $G(t_s, t_e)$, which is expressed as:

$$TCNR_{G(t_s,t_e)}(i) = \theta_1 CNR_1(i) + \theta_2 CNR_2(i) + \ldots + \theta_n CNR_n(i)$$

= $\theta \cdot CN\mathcal{R}(i)$ (6)

with regularization condition $\sum_{k=1}^{n} \theta_k = n$.

Based on the above equation, it can be found that the weight vector θ is used to tune the centrality vector $\mathcal{CNR}(i)$ so that centrality values in the time interval are counted toward the value of the time-ordered aggregation method. The sum of the weight vector over the time interval must be equal to n. This is referred to as the regularization condition of the weight vector. As introduced in Section I, to further accurately evaluate the importance of nodes in a certain time interval, a good time-ordered aggregation method will assign larger θ_k to centrality values in time windows which are closer to the start time of the time interval, and assign smaller θ_k to centrality values in time windows which are distant to the start time of the time interval. Therefore, according to the time-ordered aggregation model, three particular timeordered aggregation methods are proposed below to generate weight vector θ , so as to tune the centrality vector $\mathcal{CNR}(i)$ in a certain time interval.

A. AVERAGE TIME-ORDERED AGGREGATION METHOD

The main idea of the Average Time-ordered Aggregation Method is to aggregate centrality values in each time window by simply averaging over it. As there are *n* time windows in $G(t_s, t_e)$, and each time window is assigned with the equal weight, according to Eq. (6), then the expression of the Average Time-ordered Aggregation Method-based CNR is shown

as:

$$TCNR_{G(t_s,t_e)}(i)$$

$$= CNR_1(i) + CNR_2(i) + CNR_3(i) + \ldots + CNR_n(i)$$

$$= \sum_{k=1}^n CNR_k(i)$$
(7)

B. LINEAR TIME-ORDERED AGGREGATION METHOD

According to the definition above, the Average Time-ordered Aggregation Method lacks the relative differences of centrality values in different time windows. Therefore, in this part, the Linear Time-ordered Aggregation Method assigns different weights to centrality values in different time windows based on how close it is to the start time of the time interval. It is worth noticing that centrality values in more recent time windows should be assigned with larger weights, compared with centrality values in more distant time windows. As a result, the weights decrease linearly from the recent time windows to the distant time windows in the Linear Timeordered Aggregation Method. For the Linear Time-ordered Aggregation Method with *n* time windows, the weight vector can be obtained as $\theta = (nu, (n-1)u, \dots, 2u, u)$. Using the regularization method, we can solve $u = \frac{2}{1+n}$. Substituting it back to the weight vector, we can obtain that $\theta = (n, n-1, ..., 2, 1) \frac{2}{1+n}$. Based on the regularized weight vector and Eq. (6), the expression of the Linear Time-ordered Aggregation Method-based CNR is shown as:

$$TCNR_{G(t_s,t_e)}(i) = \frac{2}{1+n} \sum_{k=1}^{n} (n-i+1)CNR_k(i)$$
 (8)

C. EXPONENTIAL TIME-ORDERED AGGREGATION METHOD

The Exponential Time-ordered Aggregation Method assigns exponentially decreased weights to centrality values in different time windows. Similar to the Linear Time-ordered Aggregation Method, centrality values in time windows which are closer to the start time of the time interval will be assigned with larger weights in the weight vector. Different from the Linear Time-ordered Aggregation Method, the Exponential Time-ordered Aggregation Method decreases the weight vector exponentially. As a result, the importance of temporal centrality values in the Exponential Time-ordered Aggregation Method decreased more rapidly overtime than that in the Linear Time-ordered Aggregation Method. Then, according to Eq. (6), the definition of the Exponential Time-ordered Aggregation Method-based CNR is shown as:

$$TCNR_{G(t_s,t_e)}(i) = n\gamma^{n-1}CNR_n(i) + \sum_{k=1}^{n-1} n(1-\gamma)\gamma^{k-1}CNR_k(i)$$
(9)

It is worth noticing that the parameter γ is in the range of (0, 1), which controls the main exponential component of the weight vector. It is easy to find that the sum of the weight vector is also equal to *n*, which meets the

TABLE 1. Trace characteristics.

MIT Reality	Infocom 06
Nokia 6600	iMote
Bluetooth	Bluetooth
246	3
300	120
114,046	182,951
97	78
0.024	6.7
	MIT Reality Nokia 6600 Bluetooth 246 300 114,046 97 0.024

regularization condition. The weight of the first time window is slightly different from the consequent ones. However, all the values in the weight vector are dominated by γ^{k-1} , which is very large at recent time windows but decays exponentially when γ ranges from 0 to 1.

VI. PERFORMANCE EVALUATION

In this section, we focus on evaluating the performance of the proposed time-ordered aggregation model-based centrality metric using different real mobility traces.

A. SIMULATION SETUP

Two real mobility traces, *Infocom 06* [32] and *MIT Reality* [33] collected from realistic environments are used to evaluate the performance of the proposed time-ordered aggregation model-based centrality metric. The Infocom 06 trace includes 78 participants that all carry iMote nodes with Bluetooth interface to attend the IEEE INFOCOM 2006 conference, while the MIT Reality trace include 97 participants that all carry Nokia 6600 in the MIT university. The traces cover various types of corporate environments and have various experiment periods. Some characteristics of the traces are summarized in Table 1.

In the simulation, flooding is chosen as the routing protocol to disseminate messages [34]. For the Exponential Time-ordered Aggregation Method, we set γ as 0.8. To evaluate the performance of the proposed centrality, two performance metrics are defined in this paper: the average propagation delay D(i) and the propagation ratio P(i). D(i) is used to quantify how quickly node *i* can disseminate messages to all other nodes on the time interval, and P(i) is used to quantify how many nodes node *i* can disseminate to in the time interval, which are introduced as follows.

1) The average propagation delay D(i): The average propagation delay from node *i* to all the other nodes in the choosing mobility traces, which can be computed as follows:

$$D(i) = \frac{1}{|V| - 1} \sum_{j \in V, \neq i} \frac{1}{D(i, j)}$$
(10)

2) The propagation ratio P(i): the ratio of nodes in the network being successfully propagated by node *i*.

Furthermore, we compare our proposed centrality metric with the following two temporal centrality metrics:



FIGURE 5. The average Pearson correlation coefficients between P(i) (D(i)) and *i*'s TCNR values under different aggregation methods when the window size is different in the *MIT Reality* trace. (T = 120 hours). (a) P(i)-*i*'s TCNR Values. (b) D(i)-*i*'s TCNR Values.

1) The temporal degree $TDeg_{G(t_s,t_e)}(i)$ for a node $i \in V$ on a time interval $[t_s, t_e]$ is expressed as:

$$TDeg_{G(t_s,t_e)}(i) = \frac{\sum_{k=1}^{n} 2D_k(i)}{2(|V| - 1)m}$$
(11)

where $D_k(i)$ is the degree of node *i* in G_k .

2) The temporal betweeness $TBet_{G(t_s,t_e)}(i)$ for a node $i \in V$ on a time interval $[t_s, t_e]$ is expressed as:

$$TBet_{G(t_s,t_e)}(i) = \sum_{1 \le k < n} \sum \frac{\sigma_{k,n}(s,d,i)}{\sigma_{k,n}(s,d)}$$
(12)

where $\sigma_{k,n}(s, d, i) = |\mathcal{P}_{k,n}(s, d, i)|$, $\sigma_{k,n}(s, d) = |\mathcal{P}_{k,n}(s, d)|$. Here, $\mathcal{P}_{k,n}(s, d)$ denotes the set of temporal shortest paths from source *s* to destination *d* on the time interval $[G_1, G_n]$ and $\mathcal{P}_{1,n}(s, d, i)$ denotes the subset of $\mathcal{P}_{k,n}(s, d)$ consisting of paths that have *i* in their interior.

Pearson correlation coefficient is used to test whether D(i) (or P(i)) increases with node *i*'s centrality value under



FIGURE 6. The average Pearson correlation coefficients between P(i) (D(i)) and *i*'s TCNR values under different aggregation methods when the window size is different in the *Infocom 06* trace. (T = 15 hours). (a) P(i)-*i*'s TCNR Values. (b) D(i)-*i*'s TCNR Values.

different situations [35], which can be expressed as:

$$\rho_{X,Y} = Cor(X,Y) = \frac{Cov(X,Y)}{\Delta_X \Delta_Y}$$
(13)

where Cov is the covariance, Δ_X is the standard deviation of X, and Δ_Y is the standard deviation of Y. It is worth noticing that larger correlation coefficient value means the corresponding method can better measure node's centrality metrics in the time interval, and vice-versa.

B. PERFORMANCE COMPARISON

In this part, we will first evaluate the performance of different time-ordered aggregation methods-based centrality metrics, and then compare our proposed time-ordered aggregation model-based centrality metric *TCNR* with other existing temporal centrality metrics in the MIT Reality and Infocom 06 traces, respectively.



FIGURE 7. The average Pearson correlation coefficients between P(i) (D(i)) and *i*'s TCNR values under different aggregation methods when the simulation time is different in the *MIT Reality* trace.(w = 1 hour). (a) P(i)-*i*'s TCNR Values. (b) D(i)-*i*'s TCNR Values.

1) DIFFERENT TIME-ORDERED AGGREGATION METHODS

In this part, we evaluate the performance of different timeordered aggregation methods in the MIT Reality and Infocom 06 traces, respectively.

Figs. 5 and 6 represent the average Pearson correlation coefficients between D(i) (P(i)) and the corresponding centrality values under different aggregation methods when the window size is different in the MIT Reality and Infocom 06 traces, respectively. Here, we use *Sta.* to denote the Static Aggregation Method, *Ave.* to denote the Average Time-ordered Aggregation Method, *and Exp.* to denote the Exponential Time-ordered Aggregation Method, and *Exp.* to denote the Exponential Time-ordered Aggregation Method, and *Exp.* to denote the Exponential Time-ordered Aggregation Method. It is easy to find that with the increase of the window size, correlation coefficients of the Exponential Time-ordered Aggregation methods, not only in the MIT Reality trace, but also in the Infocom 06 trace, which means that the Exponential Time-ordered Aggregation Method can measure *TCNR* centrality of nodes



(b)

FIGURE 8. The average Pearson correlation coefficients between P(i)(D(i)) and *i*'s TCNR values under different aggregation methods when the simulation time is different in the *Infocom 06* trace.(w = 0.5 hour). (a) P(i)-*i*'s TCNR Values. (b) D(i)-*i*'s TCNR Values.

in the time interval more accurately than other aggregation methods. Furthermore, *Sta.* performs worst not only in the MIT Reality trace, but also in the Infocom 06 trace, which demonstrates that the time-ordered aggregation methods can measure *TCNR* centrality of nodes in the time interval more accurately than the Static Aggregation Method. It is worth noticing that correlation coefficients of the time-ordered aggregation methods decrease with the increase of the window size, especially *Exp.*. This is reasonable because if the window size is large, the number of time windows will be small, then the calculation of *TCNR* centrality in the time interval will not be accurate. Under the extreme situation, if the the window size is equal to the measured time interval, then the performance of time-ordered aggregation methods is the same as the Static Aggregation Method.

Figs. 7 and 8 represent the average Pearson correlation coefficients between D(i) (P(i)) and the corresponding *TCNR* centrality values under different aggregation methods when the simulation time is different in the MIT Reality and



FIGURE 9. The average Pearson correlation coefficients between P(i) (D(i)) and *i*'s different centrality values when the window size is different in the *MIT Reality* trace.(T = 120 hours). (a) P(i)-Different Centrality Metrics. (b) D(i)-Different Centrality Metrics.

Infocom 06 traces, respectively. Similar to the results in Figs. 5 and 6, it can be found that with the increase of the simulation time, the Exponential Time-ordered Aggregation Method also performs best not only in MIT Reality trace, but also in the Infocom 06 trace, which means the Exponential Time-ordered Aggregation Method can measure TCNR centrality of nodes in the time interval more accurately than other aggregation methods. Moreover, correlation coefficients of the time-ordered aggregation methods are much larger than that of the Static Aggregation Method not only in the MIT Reality trace, but also in the Infocom 06 trace, which demonstrates that the time-ordered aggregation methods can measure TCNR centrality of nodes in the time interval more accurately than the Static Aggregation Method. It is worth noticing that with the increase of the simulation time, correlation coefficients of the time-ordered aggregation methods are more stable than that of Sta., especially Exp.. The main reason is that the Time-ordered Aggregation Model assigns different weights to CNR values in different time windows,



FIGURE 10. The average Pearson correlation coefficients between P(i) (D(i)) and *i*'s different centrality values when the window size is different in the *Infocom* trace. (T = 15 hours). (a) P(i)-*i*'s different Centrality Values. (b) D(i)-*i*'s different Centrality Values.

and CNR values in time windows which are closer to the start time of the time interval are assigned with larger weights. Therefore, our proposed time-ordered aggregation methods still performs well when the simulation time increases.

To summarize, the Exponential Time-ordered Aggregation Method can measure *TCNR* centrality in the time interval more accurately than other aggregation methods, not only in the MIT Reality trace, but also in the Infocom 06 trace. Therefore, we recommend to use the Exponential Timeordered Aggregation Method to measure *TCNR* centrality in the time interval. Furthermore, the time-ordered aggregation methods can measure *TCNR* centrality in the time interval more accurately than the Static Aggregation Method, which demonstrates the effectiveness of our proposed time-ordered aggregation model.

2) DIFFERENT CENTRALITY METRICS

After evaluating the performance of different time-ordered aggregation methods, in this part, we compare our proposed



FIGURE 11. The average Pearson correlation coefficients between P(i) (D(i)) and *i*'s different centrality values when the simulation time is different in the *MIT Reality* trace. (w = 1 hour). (a) P(i)-*i*'s different Centrality Values. (b) D(i)-*i*'s different Centrality Values.

time-ordered aggregation model-based centrality metric *TCNR* with other existing temporal centrality metrics in the MIT Reality and Infocom 06 traces, respectively. Since the Exponential Time-ordered Aggregation Method measures TCNR centrality in the time interval more accurately than other time-ordered aggregation methods, we choose to use the Exponential Time-ordered Aggregation Method-based CNR to represent our proposed time-ordered aggregation model-based centrality metric *TCNR*.

Figs. 9 and 10 show the average Pearson correlation coefficients between D(i) (P(i)) and the corresponding values of different centrality metrics when the the window size is different in the MIT Reality and Infocom 06 traces, respectively. It can be found that with the increase of the window size, *TCNR* performs much better than the temporal degree centrality and the temporal betweeness centrality, and the temporal degree centrality performs worst in the MIT Reality trace. The main reason is that our proposed *TCNR* centrality takes both the neighbouring relationship of pair-wise nodes



FIGURE 12. The average Pearson correlation coefficients between P(i) (D(i)) and *i*'s different centrality values when the simulation time is different in the *Infocom 06* trace. (w = 0.5 hour). (a) P(i)-*i*'s different Centrality Values. (b) D(i)-*i*'s different Centrality Values.

and the time-ordered aggregation model into consideration. It is worth noticing that TCNR in the Infocom 06 trace does not perform as well as that in the MIT Reality trace. Correlation coefficients of TCNR is larger than that of TBet and TDeg only when the window size is small (w = 0.5 hour), and the performance of TCNR decreases rapidly when the window size increases in the Infocom 06 trace. The main reason is that TCNR is based on the neighbouring relationship of pair-wise nodes, but neighbouring relationship of pairwise nodes is not obvious in Infocom 06 trace. As introduced above, the Infocom 06 trace is collected by volunteers who attending the IEEE INFOCOM 2006 conference, and the MIT Reality trace is collected by teachers and students in the MIT campus. Volunteers who attending the conference are more likely to seek out new colleagues to talk to at the breaks between sessions, rather than socializing with the same people. However, teachers and students in the MIT campus are more likely to meet the same people when they are taking classes or walking in the campus. Therefore, it is harder to

depict the neighbouring relationship of pair-wise nodes in the Infocom 06 trace than in the MIT Reality trace.

Figs. 11 and 12 show the average Pearson correlation coefficients between D(i) (P(i)) and the corresponding centrality values under different centrality metrics when the simulation time is different in the MIT Reality and Infocom 06 traces, respectively. Similar to the results in Figs. 9 and 10, with the increase of the simulation time, TCNR also performs much better than the temporal degree centrality and the temporal betweeness centrality in the MIT Reality trace, and correlation coefficients of the TCNR are more stable than that of TBet and TDeg. Furthermore, TCNR and the temporal betweeness centrality both perform well in the Infocom 06 trace. Although the temporal betweeness centrality outperforms TCNR under some situations in the Infocom 06 trace, correlation coefficients of TCNR is very close to that of TBet in the Infocom 06 trace. As shown in Fig. 12, with the increase of the simulation time, the average Pearson correlation coefficients of TCNR are larger than 0.9.

To summarize, although under some situations, the temporal betweeness centrality outperforms our proposed centrality metric *TCNR* in the Infocom 06 trace, but correlation coefficients of *TCNR* is very close to that of TBet in the Infocom 06 trace, and *TCNR* performs much better than the temporal degree centrality and the temporal betweeness centrality in the MIT Reality trace. Therefore, our proposed timeordered aggregation model-based centrality metric *TCNR* outperforms the temporal degree centrality and the temporal betweeness centrality, which demonstrates the effectiveness of our proposed centrality metric.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we first introduced a new centrality metric named Cumulative Neighboring Relationship (CNR) based on the neighboring relationship of pair-wise nodes for MSNs. Then, we introduced a time-ordered aggregation model, which reduces a dynamic network to a series of time-ordered networks. Based on the time-ordered aggregation model, we proposed three particular time-ordered aggregation methods: the Average Time-ordered Aggregation Method, the Linear Time-ordered Aggregation Method, the Exponential Time-ordered Aggregation Method, and combined with our proposed centrality metric CNR to measure importance of nodes in a certain time interval. Extensive trace-driven simulations are conducted to evaluate the performance of our proposed time-ordered aggregation model-based centrality metric TCNR. The results show that the Exponential Timeordered Aggregation Method can measure TCNR centrality in the time interval more accurately than other aggregation methods, not only in the MIT Reality trace, but also in the Infocom 06 trace. Therefore, we recommend to use the Exponential Time-ordered Aggregation Method to measure TCNR centrality metrics in the time interval. Furthermore, the proposed time-ordered aggregation model-based centrality metric TCNR outperforms other existing temporal centrality metrics, which demonstrates the effectiveness of our

REFERENCES

- H. Zhou, J. Chen, J. Fan, Y. Du, and S. K. Das, "ConSub: Incentive-based content subscribing in selfish opportunistic mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 669–679, Sep. 2013.
- [2] A. Rahim et al., "Vehicular social networks: A survey," Pervasive Mobile Comput., vol. 43, pp. 96–113, Jan. 2018.
- [3] H. Zhou, H. Zheng, J. Wu, and J. Chen, "Energy efficiency and contact opportunities tradeoff in opportunistic mobile networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3723–3734, May 2016.
- [4] Q. Liu, G. Wang, F. Li, S. Yang, and J. Wu, "Preserving privacy with probabilistic indistinguishability in weighted social networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 28, no. 5, pp. 1417–1429, May 2017.
- [5] G. Yang, S. He, Z. Shi, and J. Chen, "Promoting cooperation by social incentive mechanism in mobile crowdsensing," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 86–92, Mar. 2017.
- [6] X. Kong *et al.*, "Mobility dataset generation for vehicular social networks based on floating car data," *IEEE Trans. Veh. Technol.*, to be published.
- [7] F. Xia, A. Rahim, X. Kong, M. Wang, Y. Cai, and J. Wang, "Modeling and analysis of large-scale urban mobility for green transportation," *IEEE Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1469–1481, Apr. 2018.
- [8] H. Zhou, S. Xu, D. Ren, C. Huang, and H. Zhang, "Analysis of event-driven warning message propagation in vehicular ad hoc networks," *Ad Hoc Netw.*, vol. 55, pp. 87–96, Feb. 2017.
- [9] L. C. Freeman, "Centrality in social networks conceptual clarification," *Soc. Netw.*, vol. 1, no. 3, pp. 215–239, 1979.
- [10] F. Li and J. Wu, "LocalCom: A community-based epidemic forwarding scheme in disruption-tolerant networks," in *Proc. IEEE SECON*, Jun. 2009, pp. 1–9.
- [11] H. Zhou, H. Wang, X. Li, and V. Leung, "A survey on mobile data offloading technologies," *IEEE Access*, vol. 6, no. 1, pp. 5101–5111, Jan. 2018.
- [12] H. Kim, J. Tang, R. Anderson, and C. Mascolo, "Centrality prediction in dynamic human contact networks," *Comput. Netw.*, vol. 56, no. 3, pp. 983–996, 2012.
- [13] H. Kim and R. Anderson, "Temporal node centrality in complex networks," *Phys. Rev. E, Stat. Phys.*, vol. 85, no. 2, p. 026107, 2012.
- [14] H. Zhou, L. Tong, S. Xu, C. Huang, and J. Fan, "Predicting temporal centrality in opportunistic mobile social networks based on social behavior of people," *Pers. Ubiquitous Comput.*, vol. 20, no. 6, pp. 885–897, 2016.
- [15] P. V. Marsden, "Egocentric and sociocentric measures of network centrality," Soc. Netw., vol. 24, no. 4, pp. 407–422, 2002.
- [16] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.
- [17] N. Magaia, A. P. Francisco, P. Pereira, and M. Correia, "Betweenness centrality in delay tolerant networks: A survey," *Ad Hoc Netw.*, vol. 33, pp. 284–305, Oct. 2015.
- [18] H. Zhou, L. Tong, T. Jiang, S. Xu, J. Fan, and K. Lv, "Maximum data delivery probability-oriented routing protocol in opportunistic mobile networks," *Peer-Peer Netw. Appl.*, vol. 10, no. 3, pp. 500–509, 2017.
- [19] E. M. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant MANETs," in *Proc. 8th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2007, pp. 32–40.
- [20] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: Social-based forwarding in delay tolerant networks," in *Proc. 9th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2008, pp. 241–250.
- [21] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: A social network perspective," in *Proc. 10th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2009, pp. 299–308.
- [22] J. Fan, J. Chen, Y. Du, W. Gao, J. Wu, and Y. Sun, "Geocommunitybased broadcasting for data dissemination in mobile social networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 4, pp. 734–743, Apr. 2013.
- [23] A. Socievole and F. De Rango, "Energy-aware centrality for information forwarding in mobile social opportunistic networks," in *Proc. IEEE IWCMC*, Aug. 2015, pp. 622–627.

- [24] P. Yuan, P. Liu, and S. Tang, "Exploiting partial centrality of nodes for data forwarding in mobile opportunistic networks," in *Proc. 17th Int. Conf. Comput. Sci. Eng. (CSE)*, Dec. 2014, pp. 1435–1442.
- [25] Z. Gao, Y. Shi, and S. Chen, "Measures of node centrality in mobile social networks," *Int. J. Modern Phys. C*, vol. 26, no. 9, p. 1550107, 2015.
- [26] H. Zhou, C. M. V. Leung, C. Zhu, S. Xu, and J. Fan, "Predicting temporal social contact patterns for data forwarding in opportunistic mobile networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10372–10383, Nov. 2017.
- [27] H. Chen and W. Lou, "Contact expectation based routing for delay tolerant networks," *Ad Hoc Netw.*, vol. 36, pp. 244–257, Jan. 2016.
- [28] W. Jiang, G. Wang, M. Z. A. Bhuiyan, and J. Wu, "Understanding graphbased trust evaluation in online social networks: Methodologies and challenges," ACM Comput. Surv., vol. 49, no. 1, 2016, Art. no. 10.
- [29] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for big data," *Inf. Fusion*, vol. 42, pp. 146–157, Jul. 2018.
- [30] C. Zhu, H. Zhou, V. C. M. Leung, K. Wang, Y. Zhang, and L. T. Yang, "Toward big data in green city," *IEEE Commun. Mag.*, vol. 55, no. 11, pp. 14–18, Nov. 2017.
- [31] H. Zhang, Y. Qi, H. Zhou, J. Zhang, and J. Sun, "Testing and defending methods against dos attack in state estimation," *Asian J. Control*, vol. 19, no. 3, pp. 1295–1305, 2017.
- [32] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "Crawdad data set Cambridge/haggle (v. 2009-05-29)," Dept. Phys., Cambridge Univ., Cambridge, U.K., Tech. Rep., 2009.
- [33] N. Eagle, A. Pentland, and D. Lazer, "Inferring social network structure using mobile phone data," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 36, pp. 15274–15278, 2009.
- [34] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Dept. Comput. Sci., Duke Univ., Durham, NC, USA, Tech. Rep. CS-200006, 2000.
- [35] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise Reduction in Speech Processing*. Heidelberg, Germany: Springer, 2009, pp. 1–4.



HUAN ZHOU (M'14) received the Ph.D. degree from the Department of Control Science and Engineering from Zhejiang University. He was a Visiting Scholar with Temple University from 2012 to 2013 and a CSC supported Post-Doctoral Fellow with The University of British Columbia from 2016 to 2017. He is currently a school-hired Professor with the College of Computer and Information Technology, China Three Gorges University. His research interests include mobile social

networks, VANETs, opportunistic mobile networks, and wireless sensor networks. He was a TPC Member of the IEEE WCSP 2013 and 2014, CCNC 2014 and 2015, ICNC 2014 and 2015, ANT 2015 and 2016, the IEEE Globecom 2017, and ICC 2018. He was a recipient of the Best Paper Award at I-SPAN 2014. He was the Special Session Chair of the 3rd International Conference on Internet of Vehicles in2016. He was a Lead Guest Editor of *Pervasive and Mobile Computing*. He is serving as an Associate Editor for the IEEE Access and the Local Arrangement Co-Chairs of I-SPAN 2018.



MENGNI RUAN received the bachelor's degree from the Hubei University of Arts and Science. She is currently a Graduate Student with the College of Computer Information and Technology, China Three Gorges University. Her main research interests are mobile data offloading and mobile social networks.



CHUNSHENG ZHU (M'16) received the Ph.D. degree in electrical and computer engineering from The University of British Columbia, Canada. He is currently a Post-Doctoral Research Fellow with the Department of Electrical and Computer Engineering, The University of British Columbia. He has authored over 100 publications published by refereed international journals, including the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, the IEEE TRANSACTIONS ON

COMPUTERS, the IEEE TRANSACTIONS ON INFORMATION FORENSICS and SECURITY, the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING, the IEEE TRANSACTIONS ON CLOUD COMPUTING, ACM *Transactions on Embedded Computing Systems*, and ACM *Transactions on Cyber-Physical Systems*, magazines including *IEEE Communications Magazine*, *IEEE Wireless Communications Magazine*, and *IEEE Network Magazine*), and conferences, including the IEEE INFOCOM, the IEEE IECON, the IEEE SECON, the IEEE DCOSS, the IEEE ICC, and the IEEE GLOBECOM. His research interests mainly include Internet of Things, wireless sensor networks, cloud computing, big data, social networks, and security.



VICTOR C. M. LEUNG (S'75–M'89–SM'97– F'03) received the B.A.Sc. degree (Hons.) in electrical engineering from The University of British Columbia (UBC) in 1977 and the Ph.D. degree in electrical engineering from the University of British Columbia Graduate School in 1982. He received the APEBC Gold Medal as the Head of the graduating class in the Faculty of Applied Science. He attended the Graduate School at UBC on a Canadian Natural Sciences and Engineering

Research Council Postgraduate Scholarship.

From 1981 to 1987, he was a Senior Member of Technical Staff and a Satellite System Specialist with MPR Teltech Ltd., Canada. In 1988, he was a Lecturer with the Department of Electronics, The Chinese University of Hong Kong. He returned to UBC as a Faculty Member in 1989, where he is currently a Professor. He has co-authored over 950 technical papers in international journals and conference proceedings, 37 book chapters, and co-edited 12 book titles. Several of his papers had been selected for best paper awards. His research interests are in the areas wireless networks and mobile systems. He is the TELUS Mobility Research Chair in advanced telecommunications engineering with the Department of Electrical and Computer Engineering, UBC.

Dr. Leung is a fellow of the Royal Society of Canada, the Engineering Institute of Canada, and the Canadian Academy of Engineering. He was a recipient of the IEEE Vancouver Section Centennial Award and the 2012 UBC Killam Research Prize, and the 2017 Canadian Award for telecommunications research. He has co-authored papers, that received the 2017 IEEE ComSoc Fred W. Ellersick Prize and the 2017 IEEE SYSTEMS JOURNAL Best Paper Award. He was a Distinguished Lecturer of the IEEE Communications Society. He served on the editorial boards of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS-Wireless Communications Series and Series on Green Communications and Networking, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE TRANSACTIONS ON COMPUTERS, and the Journal of Communications and Networks. He has guest-edited many journal special issues and provided leadership to the organizing committees and technical program committees of numerous conferences and workshops. He is serving on the Editorial Board of the IEEE WIRELESS COMMUNICATIONS LETTERS, the IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING, the IEEE Access, Computer Communications, and several other journals. He is a Registered Professional Engineer in the Province of British Columbia, Canada.

IEEE Access



SHOUZHI XU received the Ph.D. degree in information and communication engineering from the Huazhong University of Science and Technology in China. He is currently a Full Professor with the College of Computer and Information Technology, China Three Gorges University, Yichang, China. His current research interests include wireless sensor network, vehicle ad hoc network, and network security. His research has been funded by Natural Science Foundation of China and National Tech-

nology Research and Development Program of China. He is a fellow of the China Computer Federation.



CHUNG-MING HUANG received the B.S. degree in electrical engineering from National Taiwan University in 1984 and the M.S. and Ph.D. degrees in computer and information science from The Ohio State University in 1988 and 1991, respectively. He was the Director of the Institute of Medical Informatics, National Cheng Kung University, Taiwan. He is currently a Distinguished Professor with the Department of Computer Science and Information Engineering, National Cheng Kung

University, Tainan, Taiwan. He has authored over 300 referred journal and conference papers in wireless and mobile communication protocols, interactive multimedia systems, audio and video streaming, and formal modeling of communication protocols. His research interests include wireless and mobile network protocol design and analysis, green computing and communication, media processing and streaming, and innovative network applications and services. He was the Chair of the Department of Computer Science and Information Engineering. He was the Chair of the IEEE Vehicular Technology Society Tainan Chapter. He has served as the steering chair, the general chair, the program chair or the vice program chair of some international conferences.