Understanding mobility in networks: A node embedding approach

Matheus F. C. Barros
Universidade Federal de Ouro Preto

Carlos H. G. Ferreira
Universidade Federal de Ouro Preto
Universidade Federal de Minas Gerais
Politecnico di Torino

Bruno Pereira dos Santos
Universidade Federal de Ouro Preto

Lourenço A. P. Júnior
Instituto Tecnológico de Aeronáutica

Marco Mellia
Politecnico di Torino

Jussara M. Almeida
Universidade Federal de Minas Gerais

ABSTRACT

Motivated by the growing number of mobile devices capable of connecting and exchanging messages, we propose a methodology aiming to model and analyze node mobility in networks. We note that many existing solutions in the literature rely on topological measurements calculated directly on the graph of node contacts, aiming to capture the notion of the node's importance in terms of connectivity and mobility patterns beneficial for prototyping, design, and deployment of mobile networks. However, each measure has its specificity and fails to generalize the node importance notions that ultimately change over time. Unlike previous approaches, our methodology is based on a node embedding method that models and unveils the nodes' importance in mobility and connectivity patterns while preserving their spatial and temporal characteristics. We focus on a case study based on a trace of group meetings. The results show that our methodology provides a rich representation for extracting different mobility and connectivity patterns, which can be helpful for various applications and services in mobile networks.

Keywords

Network Mobility, Node Embedding, Network Management

1. INTRODUCTION

Analyzing and modeling mobility patterns have fundamental roles in mobile network prototyping, design, and deployment. To study such aspect, we can use mobility data gathered directly from the devices or the infrastructure that they are attached to. Examples of such data are GPS trace, call detail records (from phones), location-based social media, etc [15]. With these data, one is able to study the behavior of mobile entities and model their mobility patterns, which can be used to evaluate the performance of newly designed networking protocols and solutions [10, 19, 22].

One way to model mobility is through traditional temporal graph theory. In this approach, nodes are the network entities, and the edges are interactions (i.e., calls, message or data exchanges) over time. Conveniently, well-known graph theory models [3, 28] are ready to use. Armed with this, several proposals design solutions for, e.g., networking routing by exploring node importance criteria such as centrality metrics, community structure, and social metrics [3, 6, 13, 23, 24, 29, 44]. However, each metric captures a specific notion of node importance, indicating which nodes are central elements in the network structure. Although some earlier studies and solutions on network mobility have explored

mobility modeling based on such metrics, they cannot generalize a network nodes' importance criteria and, most important, do not consider the temporal dimension [11, 26]. These limitations are especially harmful, for example, to the validation of opportunistic routing protocols since none of the existing modelings approaches fully cover the nodes' importance criteria on the network connectivity [18, 22].

Recently, *node embedding* techniques have been proposed as an alternative to studying mobility rather than traditional graph theory. There exists a range of mobility-related applications of embeddings such as Point of Interest (PoI) prediction and recommendation [1, 20, 35, 36], and urban mobility modeling and planning [31, 33, 42]. In a nutshell, *node embedding* encodes a graph structure to low dimensional vector space (a.k.a. latent space), unveiling hidden patterns that are hard to catch using only the graph topology [16]. Although such features are desirable for exploring mobility patterns to build solutions in a network context, few efforts use *embeddings* to understand mobility patterns and capture the importance of connections established by mobile nodes, specifically valuable to design routing and information dissemination solutions in mobile networks [11].

Aiming to address the aforementioned issues, we propose using *node embedding* to model and analyze the mobility pattern over spatial and temporal dimensions. Our *embedding*-based methodology goes beyond capturing the topological aspects of the network, the mobility, and it factors the evolution of such patterns through a sequence of networks over time too. We present a case study related to the regular meeting of groups of people that applies our approach. The results show that we can extract and analyze mobility patterns, identifying strategical nodes for network connectivity and tracking the evolution of such properties over time. Our approach, albeit preliminary, can be seen as a first alternative to support the development of communication protocols, framework design, share and allocation of network resources, or applications [28, 32] which considers node stability over time.

This work is organized as follows: Section 2 summarizes the main efforts in the literature that employ *embeddings* approaches for studying network mobility, Section 3 presents our case study, while Section 4 describe the methodology. Next, the results are presented in Section 5. Finally, Section 6 discusses the main findings, offers conclusions, and future directions.

2. RELATED WORK

Modeling mobility and interactions between entities has rich applicability. For example, mobility-aware solutions were proposed focusing on routing metrics [4, 22], protocols and actions for coordination [2], and analysis of individu-

als' behavior on a mobile network [14]. However, solutions based on *node embedding* techniques have been little explored. Following, we review some prior work that model mobility pattern using embedding methods.

Yu et al. [40] proposed inferring the relationship strength between users as well as understanding the motivation behind them. To do this, they model a contacts frequency network between two users whose edge weight is adjusted based on contextual information from the contact's location. Yang et al. [35] focused on the tasks of predicting new friendships and predicting the location of users. First, the authors proposed LBSN2Vec, an embedding approach that receives as input a network of contacts modeled as a hypergraph from Location-Based Social Networking (Location-Based Social Networking (LBSBN)) data. In another work [36], the authors introduced contextual information made possible by the creation of hyperedges considering not only users' contacts but also points of interest (PoI) and time.

Embeddings techniques in the context of mobility have been applied in information retrieval tasks. Specifically, in the context of PoI recommendation, Rahmani *et al.* [20] proposed CATAPE, a method that simultaneously incorporates sequences of user locations and categories of PoI's visited by them to recommend potential new points of interest. Likewise, Yuan *et al.* combined information from historical and contrasting users' queries to find mobility similarity [41]. Such approach focuses on recommending a PoI to a user through keywords in an incomplete query. Wang *et al.* focused simultaneously on recommending and predicting users' PoIs [31].

Motivated by urban mobility planning, Zhang *et al.* [43] used an embedding strategy to extract mobility patterns on transport lines using user mobility data in the transport system. Later, the authors considered data from PoI's for the construction of urban mobility networks for users in order to find mobility patterns among users considering both data sources [42]. Similarly, Wang *et al.* proposed a framework based on embedding whose focus is to analyze the dynamics of communities formed by mobility between PoIs in different periods (i.e., weekdays and weekends) [30].

Unlike some previous work [35, 36, 40], our approach does not consider contextual information. We notice that such information offers greater possibilities for pattern analysis and extraction. On the other hand, it poses extra challenges such as data collection, storage, and processing. In addition, collect such data demands financial, operational, privacy, and security issues. Therefore, unlike these works, our approach is independent of contextual information and is based exclusively on a sequence of contact graphs that represent the topological organization of the network. This is motivated by the problem addressed here, which consists of modeling the importance of users for network connectivity considering their mobility. Therefore, we only use network' topological information in this initial effort, although we can extend it to consider contextual information.

Finally, we discuss works related to improving urban mobility and networking design [31, 33, 42]. Concerning these efforts, our proposal can be seen as an initial attempt towards planning and designing applications inherent to effective routing in opportunistic networks, which offers a range of applications in smart cities such as dissemination/collection of information, location, vehicular communication, among others [8, 25].

3. DATASET

One of the greatest challenges of performing mobility analysis in networks is the lack of reliable and open data in such field. This is due to privacy issues, scale (space, time, # entities) or costs needed to collect and store such data (typically huge amounts). Fortunately, efforts have

Table 1: GRM input parameters.

Parameters	Values
# of Nodes	100
# of Groups	500
Sim. Duration	87 days
K	70% -24h; 15% -7 days; 15% - 6h
Grid	30 x 30
Cell Size	50m^2
α_{gmt}	3
β_{gmt}	30 days
α_{dur}	3
β_{dur}	30 days
α_{size}	2.24
β_{size}	30
Social Network	Gaussian Random Partition [7]

been done to build mobility models able to generate synthetic trajectories for mobile entities by using real data. Synthetic approaches offer a cheap and enabler alternative for quick and scalable performance evaluation of networking solutions in a near real environment with different scales of space, time, and number of entities [38]. Also, synthetic models enable reproducibility and fair comparison of different methods. Given this, in this work we check the potential of node embeddings using synthetic mobility traces. In detail, we use the Group Regularity Mobility Model (GRM) [19]. GRM models group meeting dynamics and regularity in human mobility. It simulates the human mobility pattern by considering the cyclical and regular behavior commonly exhibited by humans, superposing the effect of sporadic groups meetings such as sport events, presence of places for leisure, parties, etc. The results is a trace where nodes moves, replicating the periodic behavior with the superposition of those occasional meetings. These features differentiate GRM from other models in the literature.

To produce a trace (nodes trajectories), GRM receives as input the area, duration, number of nodes, and groups. Some statistical parameters offer the ability to generate different scenarios: (i) *Group Meeting Time* that defines the time interval between meetings; such interval is governed by a power-law with exponential cut-off with parameters α_{gmt} and β_{gmt}; (ii) *Group Meetings Duration* that defines the time for which nodes will spend together; this duration follow a truncated power law distribution of parameters α_{dur} and β_{dur}; (iii) *Group Structure* that decides which nodes will be at each meeting; the grouping approach follows also a power law with exponential cuts tuned by α_{size} and β_{size}; (iv) *Social Context* is a network of social edges ("friends") between nodes; GRM consider social relationship into account because people have higher probability to attend meetings in which friends' nodes participate into. For this, GRM assumes that each group has a regularity factor defined by K. For example, groups with a factor $k = 24h$ meet every 24, 48, or 72h. This parameter works as a multiplier that generates the periodic behavior of the traits. Here, we use the same parameters as the GRM's authors to generate our dataset trace (see Table 1). It is worthy to highlight that the choice of this trace, in particular, does not prevent our approach from being used with other traces.

4. METHODOLOGY

This section describes the methodology we propose in our study, starting with our network model (Section 4.1) followed by the spatial-time aware node embedding approach applied to model a latent space (Section 4.2).

4.1 Network Model

To model the nodes' contact network observed in graphs, we first discretize the period analyzed in non-overlapping

time windows $\{\Delta_t\}$ of duration $24\,\mathrm{h}$[1] represented by the set $T = \{1, 2, ..., 87\}$. For each time window Δ_t we build an undirected graph G_{Δ_t}, with $G_{\Delta_t} = (V_{\Delta_t}, A_{\Delta_t})$, such that $V_{\Delta_t} = \{v_1, v_2, v_3, ..., v_i\}$ is the set of nodes representing individuals at the instant of time Δ_t; and $A_{\Delta_t} = \{a_1, a_2, a_3, ..., a_n\}$ a set of edges connecting any two nodes v_i and v_j ($v_i \neq v_j$). We assume two nodes are connected if they are within a radius of 100 meters during the considered Δ_t. In practice, this could represent an exchange or a potential exchange of messages between any two nodes for specific scenarios and applications such as those provided by the Wi-Fi 6 [34,39] technology. It is important to highlight that such parameters can be configured according to the application.

With the definition of the graph, we consider some purely structural measures typically in the literature in mobile network applications that provide knowledge of node importance connectivity. Here, we can compare them to the patterns captured from our methodology. Among the various possibilities, we selected some of the primary measures of centrality (Degree, Betweenness, Closeness, Eigenvector), and *clustering coefficient*, previously adopted in various solutions in network management [13, 18, 23, 24, 29, 44]. For brevity, we recommend reading [3,17] for more details about such metrics.

4.2 DynamicNode2Vec

We use our prior DynamicNode2Vec technique to extract mobility patterns from mobile nodes [11]. Such technique is able to temporally represent nodes from a sequence of graph networks (as modeled in Section 4.1) in a low-dimensional vector space (embeddings) while preserving the existing network properties (e.g. neighborhood). *DynamicNode2vec* uses a *biased random walk* algorithm based on a state-of-art method for static *node embedding* known as *Node2Vec* [12]. The intuition behind those walks is to traverse the graph randomly and produce a sequence (i.e., vectors) of visited nodes. The extraction of several random sequences generates a sampling process, where the appearance of nodes depends on the original graph. The algorithm defines two parameters: (i) nw defined as the number of walks (or extracted sequences) per node; (ii) wl the number of steps of each walk (or the length of each sequence).

To configure how the sampling will be performed, the algorithm can use Depth-First Sampling (DFS) or Breadth-First Sampling (BFS) through the p and q parameters. In details, p determines the probability of immediately returning to an already visited node; q controls whether the tour is close to the origin node (for BFS) or it goes further away (for DFS). Here, we are interested in BFS technique, which captures the proximity of nodes using a neighborhood-first strategy. In a nutshell, by sampling preferably neighbor nodes, the generated walks tend to include all nodes that belong to the same portion of the graph. Therefore, we tend to encode the importance of connections between nodes once mapped in the low-dimensional vector space. For this, we set $p = 1$ and $q = 0.5$ to get this behavior (see authors recommendations in [12]).

Given we have a different graph at each time window Δ_t, we generate the paths (walks) for each of these graphs. As result, nodes that are closer to each other in a given time window have a high number of shared neighbors, and important edges for the network connectivity are more likely to appear close together in the multiple paths sampled in the respective graph G_{Δ_t}. After, *DynamicNode2Vec* incorporates an optimization model that efficiently observes the node association and their changes over time as they appear on the sampled paths, mapping them into a latent space

temporally equivalent [37]. In summary, this is achieved as follows.

First, for each observed time window we compute a node co-occurrence matrix through a sliding window (of size five as suggested in [37]) that runs over the sampled paths. That way, it counts how many times two nodes are co-visited within that sliding window. The greater the co-occurrence between these two nodes, the closer they are in the network, and the more important is the connection between them. This matrix is known as Positive Pointwise Mutual Information (PPMI). At the next step, we create a compact representation in d dimensions from each PPMI matrix derived from the sequences of observed networks. This step is done by a low-rank matrix factorization obtained by solving an optimization problem. We consider two factors in the objective function design to handle, respectively, *overfitting* (λ) and *alignment* (τ). Their definition is based on a training process. For the sake of brevity, we recommend reading the original reference for more detail [11]. For our study, we set the parameters as follows: $nw = 4$, $nw = 8$, $\lambda = 50$, $tau = 15$ and $d = 50$.

With the obtained embeddings, we can use two metrics to quantify the degree of mobility of nodes on the network, and the importance of the connections established by them: [11, 37]:

Cosine Distance: Quantifies the similarity between two vectors in the same vector space. Given two vectors generated by a node (user) v_i in any two time windows Δt_1 and Δt_2, the cosine distance is defined as $cos(x, y) = 1 - x \cdot y/(\|x\|\|y\|)$. The cosine distance ranges from 0 to 1. Values close to 0 indicate that the node is always close to the same nodes in the two compared time windows; values close to 1 indicate that the node has drastically changed its network of contacts in the period, with total different neighbors.

Vector norm: By definition, the more a node appears in the sampled paths in the graph, the greater is the vector norm of its vector, taking into account all vector space of observed time. This way, it is possible to identify those nodes that exhibit an important connectivity pattern since they are frequently inserted in the sampled paths. The main advantage of using the vector norm regarding the frequency with which nodes appear on the sampled paths is that it is more robust to sporadic disturbances [37].

5. RESULTS

In this section, we present the main results of our work. First, we illustrates our embedding representation of the mobile network considering space and time characteristics. Next, the node mobility and importance is analyzed through the case study are presented.

5.1 Exemplifying the spatio-temporal model

To exemplify how our proposal works, we pick just the first six time windows. Recall the *DynamicNodeVec* produce a node embedding representation of 50 dimensions. To be human-visible, we reduce from 50 to 2 dimensional representation by using t-SNE [27], a dimensionality reduction technique for data visualization. Furthermore, to illustrate how the node dynamics can be captured in the latent space modeled, we select only two nodes that presented the highest and lowest average day-to-day mobility (review Section 4.2) according to the first six time windows. We compute the cosine distance day after day for all nodes. Then, we compute the average of these distances for each node.

Figure 1 displays the nodes that have the highest and lowest average mobility in the period. Their average cosine distance were 0.10 (min) and 0.43 (max), respectively. The

---

[1]The choice of daily time windows is based on the meetings periodicity presented in Table 1, but can be configured according to the scenario. It should be noticed that 85% of meeting occurs on a daily scale.
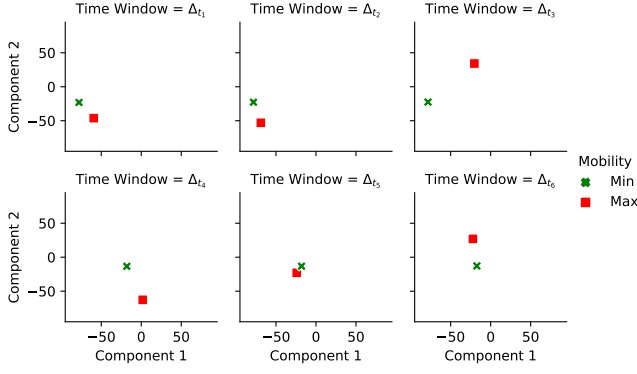
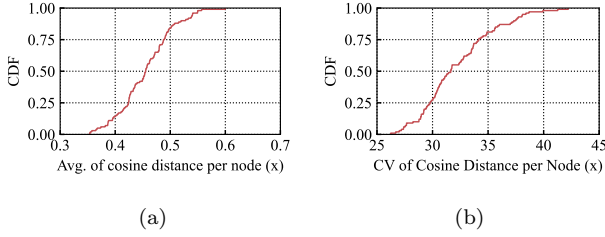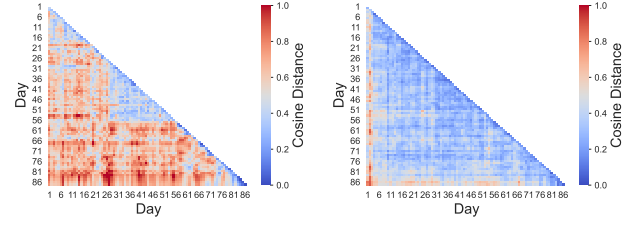Figure 1: t-SNE Two-dimensional representation of node embeddings for the first six-time windows.

Figure 2: Cumulative Distribution Function (CDF) of average and CV of mobility shifting of nodes (measured w.r.t. cosine distance).

(a) Highest avg. cosine distance (b) Lowest avg. cosine distance

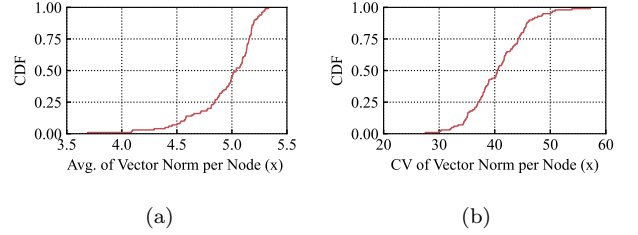Figure 3: Mobility shifting of the nodes with lowest and highest cosine distance over consecutive days.

Figure 4: CDF of average and CDF of the nodes' connection importance (measured w.r.t. nodes' vector norm).

node with the lowest mobility presented stable (few topological changes) behaviors in the first three time windows (Figure 1 top plots). However, between the time windows Δt_3 and Δt_4, a shifting is observed, indicating a change in connectivity of this node. Afterward, the node remains stable until the window Δt_6. On the other hand, it is possible to notice a high day-to-day variability for the node with the highest average mobility in the period analyzed. In practice, this indicates that this node has significant neighborhood changes in the network. In this way, we can distinguish nodes with different levels of mobility and, if necessary, separate them into specific periods as needed by several network applications.

5.2 Unveiling the mobility and connectivity importance

In order to understand the mobility patterns, we start by investigating the mobility level of all nodes throughout the period observed on the latent space modeled. Some of the possible applications of our approach include identifying nodes with high mobility degrees for routing in opportunistic or delay-tolerant networks [13, 18, 24]. Therefore, we analyze the nodes' mobility levels assuming that new connections - represented by topological changes in the network and captured in the latent space - can occur with an acceptable delay. Thus, we calculate the cosine distance of a time window Δt_i for all subsequent time windows Δt_j, such that, $i < j$. In other words, we evaluate the degree of mobility and the potential that a node has to establish new connections throughout the analyzed period.

Figure 2a shows the nodes cosine distance average in a CDF. Looking at the 10% of the most stable nodes, in terms of connectivity, the average cosine distance is at most 0.38. On the other hand, the 10% of the nodes that change the most shows a mobility average greater than 0.54, exhibiting a 42% higher mobility level. In order to check to what extent the mobility changes of the nodes are stable or characterized by bursts in specific periods, we compute the Coef-

ficient of Variation (CV) as a measure of dispersion[2]. The CV distribution is shown in Figure 2b. In practice, values greater than 30 are considered high coefficients of variation [21]. Therefore, it is possible to observe that more than 75% of the nodes in the network show a greater dynamic in specific periods.

We then select nodes with the highest and lowest average cosine distance to investigate mobility changes over consecutive days. Figures 3a and 3b show heatmaps where one is able to compare day to day changes over the observed period. The axes x and y indicate the cosine distance to two given time windows. The color indicates the topological change level encoded in the latent space, revealing their degree of mobility according to the cosine distance. The more red and intense the cell, the greater the difference in node connections over the two days observed. Conversely, the more blue and intense, the smaller this difference. It is possible to notice that the node with the high average cosine distance (Figure 3a) tends to present high mobility levels for most pairs of days observed, which indicates constant topological changes. We also investigated the node's behavior according to the highest and lowest average CV and found that they exhibit similar behavior at mobility levels to the cosine distance average but with more dispersed changes in even more specific periods.

Recall that the cosine distance indicates node mobility in terms of connectivity change. However, we can not claim that a node with a higher mobility level provides the most important connections, for example, to allow a message to reach the entire network. For this reason, we evaluate the vector norm (see details in Section 4.2), which captures the importance of a node in terms of connectivity. Figure 4a shows the CDF of the average vector norm for all nodes. The 10% with the lowest value has an average norm smaller than or equal to 4.45. Conversely, the 10% with the highest average value present an average norm greater than or equal to 5.3, resulting in a difference of almost 20%. Note that nodes with higher vector norm are more important and central in terms of connectivity. We also evaluate the CV of the norms obtained per node and present the distribution in Figure 4b. Unlike cosine distance, here, 98% of nodes have a high CV (greater than 30). This great vari-

---

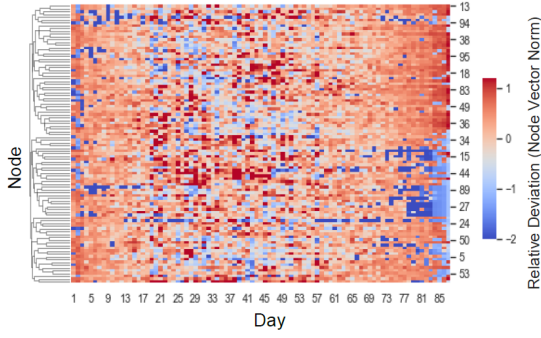[2]CV is computed by the ratio of the standard deviation to the mean and expressed in percentage.

Figure 5: Heatmap of the node vector norm over the observed period.

Table 2: Correlations between cosine distance and the vector norm.

Metrics		Correlation
Avg. of Cosine Distance	CV of Cosine Distance	-0.04
Avg. of Cosine Distance	CV of Vector Norm	0.70
Avg. of Vector Norm	CV of Cosine Distance	-0.95
Avg. of Vector Norm	CV of Vector Norm	-0.10

ability suggests that some nodes are more important for exchanging information in the network at specific times as network connections evolve.

The vector norm of each node in each day can be used to understand nodes' connectivity importance in the temporal aspect. Figure 5 presents a day-to-day heatmap for all nodes' norm vectors. Rows represent the nodes and columns represent the days. Each day was normalized by using the *z-score*[3]. Thus, one can see how much node's importance deviates relatively from the average importance over days. The color of each cell reflects the normalized norm value. Also, we compute a dendrogram to group nodes that present similar importance on specific days. The Pearson correlation between the node norms was employed as dendrogram grouping metric [5]. We can observe that some nodes present greater importance than the others (indicated by red cells) in specific periods, confirming our finding about the high CV of the norm.

As the cosine distance and the vector norm provide complementary analyses, we investigate how these measures are related. Intuitively, we aim to understand how the mobility of nodes is related to the importance of their connections. To do so, we apply the Pearson's correlation between the average and CV of both measures, the result is summarized in Table 2. The strongest correlation is negative and occurs between the average of the norm and the CV of the cosine distance. This suggests that nodes with more important connections have a stable mobility pattern. Then, we observe a moderate correlation between the average cosine distance and the CV of the norm, which means that nodes who consistently make abrupt connectivity changes tend to maintain connections that are sometimes important and sometimes not. The other correlations are not significant for $p = 0.01$.

We also investigate the Person's correlation between our approach and traditional graph topological importance measures [9, 13, 18, 29, 44]. To do it, we consider the main well-known topological measures used in these works. Remember, they capture the most varied notions of nodes' centrality and connectivity in the network, and, therefore, different notions of importance. Table 3[4] depicts the correlations between these measures.

Turning the focus on the centrality measures (Degree, Betweenness, Closeness, and Eigenvector). It is possible to observe a moderate negative correlation between traditional

Table 3: Correlations between topological and embedding measurements.

Embedding	Topological	Correlation
Avg. of Cosine Distance	Avg. of Degree	-0.64
Avg. of Cosine Distance	Avg. of Betweenness	-0.69
Avg. of Cosine Distance	Avg. of Closeness	-0.64
Avg. of Cosine Distance	Avg. of Eigenvector	-0.65
Avg. of Cosine Distance	Avg. of Clustering Coefficient	0.51
Avg. of Vector Norm	Avg. of Degree	0.33
Avg. of Vector Norm	Avg. of Betweenness	0.50
Avg. of Vector Norm	Avg. of Closeness	0.28
Avg. of Vector Norm	Avg. of Eigenvector	0.32
Avg. of Vector Norm	Avg. of Clustering Coefficient	-0.55

graph measures and the average of the cosine distance. It suggests that the more dynamic a node is, the lower its centrality according to different views. The average vector norm presents slightly positive correlation with respect to traditional measures, a possible explanation is that centrality measures typically capture unique notions of centrality and importance, and do not consider the temporal factor [26]. Hence, they cannot generalize a concept of centrality that is universally important in the graph. Finally, we consider the clustering coefficient, which shows a low-to-moderate positive and negative correlation, respectively, with both node metrics in latent space. The positive case suggests that nodes with neighbors more connected (high clustering coefficient) tend to change their connections collectively (high avg. of cosine distance), which may reflect our case study's regular group meetings property. On the other hand, nodes with less-connected neighborhoods tend to have essential links for network connectivity (high avg. of vector norm), explained by bridges and peripheral connections in the networks.

6. CONCLUSION

Understanding mobility in mobile networks includes examining how entities move along with their role for network connectivity. Those are patterns notably relevant to design networking protocols and solutions. This work proposed a methodology for modeling and analyzing mobile networks by employing a node embedding approach capable of unveiling patterns related to network connectivity, mobility, and their natural evolution. We noticed that most solutions designs in the literature rely on graph theory metrics computer on the network. Consequently, they tend to privilege particular characteristics indicating which nodes are central points in the structure. Alternatively, our approach offers a generalized notion of connectivity importance based on the existing links in each node's network tracking it while the connections in the network change.

By applying our methodology in a case study as a proof of concept, we showed that our approach could reveal nodes with different mobility and connectivity importance levels, whereas the network topology evolves. Finally, we contrasted the notion of connectivity importance provided by our methodology with some topological network measurements used network solutions present in the literature. We then observed that these measures could not generalize the patterns of connectivity captured by our methodology.

As future works, our methodology offers a range of possibilities. An immediate opportunity is to apply our methodology to design new solutions that require an efficient node selection strategy for dissemination/collection of information, location, vehicular communication, among other applications in the context of smart cities. It also includes evaluating the performance of such protocols, comparing purely topological measures to those proposed here. Finally, another possibility is to extend the network modeling to include contextual information about the nodes' visitation locations. It would enable new tasks to be performed, including prediction and recommendation of points of interest to users.

---

[3]Each value is subtracted by the column average and divided by the column standard deviation.

[4]All observations are significant for $p = 0.01$.

7. REFERENCES

[1] S. Agrawal, D. Roy, and M. Mitra. Tag embedding based personalized point of interest recommendation system. *Information Processing & Management*, 58(6), 2021.

[2] M. Alajeely, R. Doss, and A. Ahmad. Routing protocols in opportunistic networks–a survey. *IETE Technical Review*, 35(4):369–387, 2018.

[3] A.-L. Barabási et al. *Network Science*. Cambridge University Press, 2016.

[4] S. Batabyal and P. Bhaumik. Mobility models, traces and impact of mobility on opportunistic routing algorithms: A survey. *IEEE Communications Surveys & Tutorials*, 17(3):1679–1707, 2015.

[5] J. Benesty, J. Chen, Y. Huang, and I. Cohen. Pearson correlation coefficient. In *Noise reduction in speech processing*, pages 1–4. Springer, 2009.

[6] S. P. Borgatti and M. G. Everett. A graph-theoretic perspective on centrality. *Social networks*, 28(4):466–484, 2006.

[7] U. Brandes, M. Gaertler, and D. Wagner. Experiments on graph clustering algorithms. In *European Symposium on Algorithms*, pages 568–579. Springer, 2003.

[8] Y. Cai, H. Zhang, Y. Fan, and H. Xia. A survey on routing algorithms for opportunistic mobile social networks. *China Communications*, 18(2), 2021.

[9] E. M. Daly and M. Haahr. Social network analysis for routing in disconnected delay-tolerant manets. In *Proceedings of ACM international symposium on Mobile ad hoc networking and computing*, 2007.

[10] C. H. Ferreira, L. H. Nunes, L. A. Pereira, L. H. Nakamura, J. C. Estrella, and S. Reiff-Marganiec. Peesos-cloud: a workload-aware architecture for performance evaluation in service-oriented systems. In *2016 IEEE World Congress on Services (SERVICES)*, pages 118–125. IEEE, 2016.

[11] C. H. G. Ferreira, F. M. Ferreira, B. de Sousa Matos, and J. M. de Almeida. Modeling dynamic ideological behavior in political networks. *The Journal of Web Science*, 7, 2019.

[12] A. Grover and J. Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the International Conference on Knowledge discovery and data mining*, 2016.

[13] X. Guo, C. Dong, and W. Wen. Dynamic computation offloading strategy with dnn partitioning in d2d multi-hop networks. In *2021 9th International Conference on Communications and Broadband Networking*, pages 172–178, 2021.

[14] A. Hess, K. A. Hummel, W. N. Gansterer, and G. Haring. Data-driven human mobility modeling: a survey and engineering guidance for mobile networking. *ACM Computing Surveys (CSUR)*, 48(3):1–39, 2015.

[15] A. Hess, K. A. Hummel, W. N. Gansterer, and G. Haring. Data-driven human mobility modeling: a survey and engineering guidance for mobile networking. *ACM Computing Surveys*, 48(3), 2016.

[16] Y. Lu, X. Wang, C. Shi, P. S. Yu, and Y. Ye. Temporal network embedding with micro-and macro-dynamics. In *International Conference on Information and Knowledge Management*, 2019.

[17] M. E. Newman. A measure of betweenness centrality based on random walks. *Social networks*, 27(1), 2005.

[18] I. O. Nunes, C. Celes, I. Nunes, P. O. V. de Melo, and A. A. Loureiro. Combining spatial and social awareness in d2d opportunistic routing. *IEEE Communications Magazine*, 56(1):128–135, 2018.

[19] I. O. Nunes, C. Celes, M. D. Silva, P. O. Vaz de Melo, and A. A. Loureiro. Grm: Group regularity mobility model. In *Proceedings of the 20th ACM International Conference on Modelling, Analysis and Simulation of Wireless and Mobile Systems*, pages 85–89. ACM, 2017.

[20] H. A. Rahmani, M. Aliannejadi, R. Mirzaei Zadeh, M. Baratchi, M. Afsharchi, and F. Crestani. Category-aware location embedding for point-of-interest recommendation. In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval*, pages 173–176, 2019.

[21] G. F. Reed, F. Lynn, and B. D. Meade. Use of coefficient of variation in assessing variability of quantitative assays. *Clinical and Vaccine Immunology*, 9(6):1235–1239, 2002.

[22] B. P. Santos, O. Goussevskaia, L. F. Vieira, M. A. Vieira, and A. A. Loureiro. Mobile matrix: routing under mobility in iot, iomt, and social iot. *Ad Hoc Networks*, 78:84–98, 2018.

[23] B. P. Santos, L. F. M. Vieira, and A. A. F. Loureiro. Routing and Mobility Management in the Internet of Things. In *Brazilian Symposium on Computer Networks and Distributed Systems*, 2020.

[24] C. Sobin, V. Raychoudhury, G. Marfia, and A. Singla. A survey of routing and data dissemination in delay tolerant networks. *Journal of Network and Computer Applications*, 67:128–146, 2016.

[25] A. S. Syed, D. Sierra-Sosa, A. Kumar, and A. Elmaghraby. Iot in smart cities: A survey of technologies, practices and challenges. *Smart Cities*, 4(2):429–475, 2021.

[26] T. W. Valente, K. Coronges, C. Lakon, and E. Costenbader. How correlated are network centrality measures? *Connections (Toronto, Ont.)*, 28(1):16, 2008.

[27] L. Van der Maaten and G. Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[28] N. Vastardis and K. Yang. Mobile social networks: Architectures, social properties, and key research challenges. *IEEE Communications Surveys & Tutorials*, 15(3), 2013.

[29] L. Wan, G. Han, J. Jiang, C. Zhu, and L. Shu. A doa estimation approach for transmission performance guarantee in d2d communication. *Mobile Networks and Applications*, 22(6):998–1009, 2017.

[30] P. Wang, Y. Fu, J. Zhang, X. Li, and D. Lin. Learning urban community structures: A collective embedding perspective with periodic spatial-temporal mobility graphs. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 9(6):1–28, 2018.

[31] Y. Wang, H. Sun, Y. Zhao, W. Zhou, and S. Zhu. A heterogeneous graph embedding framework for location-based social network analysis in smart cities. *IEEE Transactions on Industrial Informatics*, 16(4):2747–2755, 2019.

[32] F. Xia, L. Liu, B. Jedari, and S. K. Das. Pis: A multi-dimensional routing protocol for socially-aware networking. *IEEE Transactions on Mobile Computing*, 15(11):2825–2836, 2016.

[33] F. Xu, Z. Lin, T. Xia, D. Guo, and Y. Li. Sume: Semantic-enhanced urban mobility network embedding for user demographic inference. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(3), 2020.

[34] M. B. Yaakop, I. A. Abd Malik, Z. bin Suboh, A. F. Ramli, and M. A. Abu. Bluetooth 5.0 throughput comparison for internet of thing usability a survey. In

2017 International Conference on Engineering Technology and Technopreneurship (ICE2T), pages 1–6. IEEE, 2017.

[35] D. Yang, B. Qu, J. Yang, and P. Cudre-Mauroux. Revisiting user mobility and social relationships in lbsns: a hypergraph embedding approach. In *The world wide web conference*, pages 2147–2157, 2019.

[36] D. Yang, B. Qu, J. Yang, and P. Cudré-Mauroux. Lbsn2vec++: Heterogeneous hypergraph embedding for location-based social networks. *IEEE Transactions on Knowledge and Data Engineering*, 2020.

[37] Z. Yao, Y. Sun, W. Ding, N. Rao, and H. Xiong. Dynamic word embeddings for evolving semantic discovery. In *Proceedings of the eleventh acm international conference on web search and data mining*, pages 673–681, 2018.

[38] S. Yi, Z. Qin, and Q. Li. Security and privacy issues of fog computing: A survey. In *International conference on wireless algorithms, systems, and applications*, pages 685–695. Springer, 2015.

[39] J. Yin, Z. Yang, H. Cao, T. Liu, Z. Zhou, and C. Wu. A survey on bluetooth 5.0 and mesh: New milestones of iot. *ACM Transactions on Sensor Networks (TOSN)*, 15(3):1–29, 2019.

[40] Y. Yu, H. Wang, and Z. Li. Inferring mobility relationship via graph embedding. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3):1–21, 2018.

[41] Z. Yuan, H. Liu, Y. Liu, D. Zhang, F. Yi, N. Zhu, and H. Xiong. Spatio-temporal dual graph attention network for query-poi matching. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 629–638, 2020.

[42] T. Zhang, X. Duan, and Y. Li. Unveiling transit mobility structure towards sustainable cities: An integrated graph embedding approach. *Sustainable Cities and Society*, 72:103027, 2021.

[43] T. Zhang, J. Wang, C. Cui, Y. Li, W. He, Y. Lu, and Q. Qiao. Integrating geovisual analytics with machine learning for human mobility pattern discovery. *ISPRS International Journal of Geo-Information*, 8(10):434, 2019.

[44] Z. Zhang, P. Zhang, D. Liu, and S. Sun. Srsm-based adaptive relay selection for d2d communications. *IEEE Internet of Things Journal*, 5(4), 2017.