On the Impact of Graph Structure on Mobility in Opportunistic Mobile Networks

Christoph P. Mayer and Oliver P. Waldhorst
Institute of Telematics, Karlsruhe Institute of Technology (KIT), Germany
{mayer,waldhorst}@kit.edu

Abstract—Opportunistic mobile networks are a promising way to offload infrastructure networks, or provide communication in case of insufficient or non-existent infrastructure coverage. Understanding of the mobility process that drives such networks is crucial for design, analysis, and configuration. Generally, this mobility process is modeled on a plain playground where devices can move freely; both in case of simulation, and analysis of real-world traces. Graph-based playgrounds provide more realistic models but their impact on mobility is insufficiently understood. We provide a methodology to analyze the impact of the underlying graph on inter-contact time using methods from spectral graph theory. We gather the inter-contact times for a random mobility process. Inter-contact times provide the most important mobility characteristic for mobile opportunistic networks; defined in [6] as “the time elapsed between two successive contact periods for a given pair of devices”. Chaintreau et al. [6] have confirmed through simulations that the inter-contact time “directly impacts routing delays in DTN” [11]. While some conjectures on the relationship between inter-contact times and the underlying graph where mobile devices move on exist, e.g. [11], to the best of our knowledge there is no sound theoretical framework to provide evidence for or against the existence of such relationship.

In this paper, we use results from spectral graph theory to derive a correlation between graph characteristics and inter-contact times. Through simulation of a social mobility model and a random mobility model on synthetic grid-based graphs and real-world city graphs we gather data on inter-contact times. Parameters for models of inter-contact time distribution are acquired by model fitting. We derive that the spectral gap—determined by the second largest eigenvector of a graph’s random walk—directly determines inter-contact times for certain types of graphs and mobility models and analyze the correlation between the spectral gap and the parameters of the inter-contact time distributions. Our main findings are:

- Inter-contact times are strongly influenced by grid-based graph structure in random and social mobility models.
- Real-world city maps do not exhibit sufficient difference in structure to effectively influence inter-contact time of a social mobility model. Difference in structure is strong enough to influence a random mobility model.

In Section II we present city- and grid-graphs with the graph model we use. Simulations of mobility models on the graphs are described in Section III. In Section IV we perform model fitting to determine the parameters for the mobility model based on simulative data. Analysis through random walks and spectral graph theory is provided in Section V to derive a simple graph metric that we use in Section VI for correlation with parameters from model fitting. Related work is given in Section VII, summary and concluding remarks in Section VIII.

I. INTRODUCTION

The rapid growth in number of mobile devices like smartphones enables new forms of infrastructure-less communication through opportunistic mobile networks such as Delay Tolerant Networks (DTN). Such networks can offload congested infrastructure networks [1], [2], provide communication when infrastructure coverage is unavailable [3], [4], and provide infrastructure-less content dissemination [5].

Mobile opportunistic networks exploit characteristics of human mobility for routing, which have shown to follow a power-law distribution in inter-contact time [6], [7], [8]. Based on these observations—which have been analyzed from mobility traces obtained through real-world experiments—mobility models have been proposed that can be used for simulative studies [9], [10]. Typically such mobility models are applied on a plain playground, allowing mobile devices to move to an arbitrary destination on a straight line. In reality, however, movement of mobile devices is restricted by the street structure of a city.

Inter-contact times provide the most important mobility characteristic for mobile opportunistic networks; defined in [6] as “the time elapsed between two successive contact periods for a given pair of devices”. Chaintreau et al. [6] have shown analytically that the slope of the inter-contact time distribution follows a power-law distribution which directly affects the expected delay in DTNs—and therewith delivery probability if messages are time-to-live restricted. Rhee et al. have confirmed through simulations that the inter-contact time “directly impacts routing delays in DTN” [11].

While some conjectures on the relationship between inter-contact times and the underlying graph where mobile devices move on exist, e.g. [11], to the best of our knowledge there is no sound theoretical framework to provide evidence for or against the existence of such relationship.

In this paper, we use results from spectral graph theory to derive a correlation between graph characteristics and inter-contact times. Through simulation of a social mobility model and a random mobility model on synthetic grid-based graphs and real-world city graphs we gather data on inter-contact times. Parameters for models of inter-contact time distribution are acquired by model fitting. We derive that the spectral gap—determined by the second largest eigenvector of a graph’s random walk—directly determines inter-contact times for certain types of graphs and mobility models and analyze the correlation between the spectral gap and the parameters of the inter-contact time distributions. Our main findings are:

- Inter-contact times are strongly influenced by grid-based graph structure in random and social mobility models.
- Real-world city maps do not exhibit sufficient difference in structure to effectively influence inter-contact time of a social mobility model. Difference in structure is strong enough to influence a random mobility model.

In Section II we present city- and grid-graphs with the graph model we use. Simulations of mobility models on the graphs are described in Section III. In Section IV we perform model fitting to determine the parameters for the mobility model based on simulative data. Analysis through random walks and spectral graph theory is provided in Section V to derive a simple graph metric that we use in Section VI for correlation with parameters from model fitting. Related work is given in Section VII, summary and concluding remarks in Section VIII.

II. CITY AND GRID GRAPHS

We have selected ten metropolitan cities—shown in Figure 1 and listed in Table I—for our analysis that cover a large set of different street graph structures. Our data is based on OpenStreetMap1 and has undergone conversion by custom tools2 for analysis and simulation. For geographic extraction

1http://www.openstreetmap.org
2http://www.tm.kit.edu/~mayer/osm2wkt
we selected the city center and cut a $2 \times 2$ km square area of the city. Post-processing steps have been performed to remove small unconnected partitions and map-data errors, and conversion for simulation and analysis.

We model the city street graph $G = (V, E)$ as follows: street crossings are vertices $V$ and street parts between crossings are edges $E$. A complete street $s_t \in S$ is made up of multiple street parts $s_t = \{e_q, \ldots, e_p\} = \{(v_i, v_j), \ldots, (v_k, v_l)\}$. The length of a street part $e_i$ is denoted $L(e_i)$, and the length of the complete street $s_t$ as $L(s_t) = \sum_{e_i \in s_t} L(e_i)$. We focus on street parts rather than complete streets as they provide a clearer definition through the graph’s edges. Table I shows different graph-related properties of the city graphs, we will come back to these later in Section VI.

Besides city street graphs we use synthetic grid-based maps. A $k \times k$ grid is a graph $G = (V, E)$ with $V = \{1, \ldots, k\}^2$ and $E = \{(i, j), (i', j') : |i - i'| + |j - j'| = 1\}$, i.e. a vertex has edges to its one-hop neighboring vertices. We name a $k \times k$ grid simple $k$-grid in the following, and $k$ the grid-id. Note, that equal to the city graphs we always use a $2 \times 2$ km area for the grids, i.e. the grids are always laid out in the same spatial relation with respect to the outer vertices, but with different density of the grid. Therefore, the distance between two adjacent vertices in a $k$-grid is always $2$ km/$(k - 1)$.

### III. SIMULATION

Our reason to perform simulation over pure modeling for analysis of mobility on graphs is twofold: First, movement speed of devices and spatial layout of the graph is inherently considered in simulations. Second, communication range of devices and resulting effects become naturally apparent. While such effects can be analytically captured, the methodology quickly becomes unmanageable. We perform simulation of mobile devices that can only move on a given graph to gather the mobility behavior on different city- and grid-graphs, in combination with two mobility models based on social behavior and random movement. As simulator we use ONE [12], as social mobility model Small World in Motion (SWIM) [10], and as random mobility model random walk as described in Section V. ONE can simulate a large number of mobile devices through mobility models on graphs, and is normally used for DTN evaluation. SWIM is a social mobility model where devices select their next destination based on distance to their home location, and the number of other devices they have met in specific locations. It has been shown in [10] that SWIM can generate realistic inter-contact time distributions of power-law and exponential decay that have been observed in real-world traces [6], [7]. Our main interest is in the pair-wise inter-contact time behavior of the mobile devices, i.e. the time it takes until the exact pair of devices comes into contact again. Therefore, the number of devices in the simulation is irrelevant, but a minimum of two devices required. However, to gather more inter-contact time samples we use 50 devices in the simulation. Furthermore, SWIM requires a certain number of devices to build up its social context. In both mobility models devices randomly select their movement speed uniformly from $[1, 3]$ m/s. SWIM is configured with $\alpha = 0.1$ to get a strong power-law slope (cf. Figure 6c), waiting-time slope 1.45, and waiting-time cut-off $12$ h (see [10]). Two devices are in contact when the distance between them is equal or less than the communication range of $50$ m. The simulation performs a warmup phase of $7$ days and actual measurement of inter-contact behavior in the $7$ days following the warmup. During warmup the device distribution in random walk becomes stationary, and SWIM can build up its social context.

We perform simulations for each city and grid listed in Table II. Each graph is simulated with $10$ independent seeds with both SWIM and random walk. The complimentary cumulative distribution function (CCDF) of inter-contact times is calculated for each seed and mean values are obtained. In case a seed’s CCDF does not contain a sample of a required inter-contact time value we perform linear interpolation to allow for calculation of mean values over the $10$ seeds$^3$. Therefore, for each combination of graph (cities $\cup$ grids) and mobility model (SWIM $\cup$ random walk) we get a CCDF inter-contact time distribution based on mean over $10$ seeds.

### IV. MODEL FITTING

To compare the inter-contact time samples obtained through simulation—and especially the resulting CCDF—we fit the CCDF to a model that describes the distribution with few parameters. In case of the random walk the CCDF exhibits a power-law with light slope and a smooth transition to exponential decay [7]. We can therefore model the inter-contact time distribution of random walk through:

$$f_{\text{random}}(x; a, b, c) \propto (a \cdot x)^{-b} \cdot e^{-c \cdot x} \quad (1)$$

In case of SWIM we observe a much sharper transition from a strong power-law to fast exponential drop. This point of transition equals the waiting-time cut-off configured in SWIM for $12$ h (cf. Section III). Karagiannis et al. name this point the characteristic time and observe it to be roughly at around half a day in real-world traces [7]; which is in conformance with our simulative results. As our domain size is always of $2 \times 2$ km, this transition point is fixed, resulting from the results

$^3$CCDF values obtained through simulation are very dense on the x-axis and interpolation only necessary on very small intervals between two consecutive inter-contact CCDF samples

| City   | $|V|$  | $|E|$  | $L(\epsilon)$  | $d(\epsilon)$ | $1 - \lambda_2$ |
|--------|-------|-------|----------------|---------------|-----------------|
| Cairo  | 1024  | 1580  | 63 m           | 3.08          | 0.000648        |
| Chennai| 822   | 1255  | 70 m           | 3.05          | 0.001153        |
| Karlsruhe| 2902 | 4853  | 26 m           | 3.34          | 0.000569        |
| Los Angeles | 480 | 775   | 119 m          | 3.22          | 0.002909        |
| Manhattan | 1032 | 1852  | 64 m           | 3.58          | 0.001651        |
| New Delhi | 387  | 590   | 101 m          | 3.04          | 0.004383        |
| Richmond | 569   | 927   | 87 m           | 3.25          | 0.001998        |
| San Francisco | 792 | 1287  | 73 m           | 3.25          | 0.002200        |
| Tokyo   | 1888  | 2577  | 31 m           | 2.72          | 0.000853        |
| Venice  | 4750  | 6459  | 21 m           | 2.71          | 0.000324
of Cai et al. [13]. Therefore, we model the inter-contact times distribution of SWIM through a piecewise function:

\[ f_{social}(x; a, b, c) \propto \begin{cases} 
  (a \cdot x)^{-b} & \text{if } x < 12 \text{ hours} \\
  e^{-c \cdot x} & \text{if } x \geq 12 \text{ hours} 
\end{cases} \] (2)

In the following we name the parameters that require fitting as “shift” \( a \), “power-law slope” \( b \), and “exponential” \( c \).

For fitting \( f_{random} \) and \( f_{social} \) to the inter-contact times CCDF obtained through simulations we perform least-squares fitting. Parameters obtained for \( f_{random} \) and \( f_{social} \) are given in Table II. In Section VI we will correlate these parameters to properties of the graph that can be obtained easily and therewith analyze whether the underlying graph or the mobility model is the dominating force for the resulting inter-contact distribution.

V. RANDOM WALKS AND SPECTRAL GRAPH ANALYSIS

In the following we use random walks on graphs modeled by Markov chains to survey the properties of the underlying graph to the mobility behavior (see [14] for an excellent introduction and survey on random walks). Note that the random walk analysis does not describe the spatial layout of the graph and therewith does not account for movement speed of devices and communication range, for this we performed simulations in Section III. We can, however, analyze graph properties and their relation with mobility much better and will correlate the analytical results with simulation in Section VI.

Recall that we model the graph where mobile devices move as \( G = (V, E) \). A random walk on \( G \) starts at vertex \( v_0 \in V \) and is at \( v_t \) in the \( t \)-th step. In each step the random walk selects a neighboring vertex with probability \( 1/d(v_t); d(v_t) \) being the degree of vertex \( v_t \). \( P_t(i) \) is the probability of the random walk of residing on \( v_t \) at step \( t \). The sequence of random locations \( (v_t : t = 0, 1, \ldots) \) is a Markov chain. \( \mathcal{A}_G \) is the adjacency matrix of the street graph \( G \), and \( D \) a diagonal matrix with the \( i \)-th diagonal element set to \( 1/d(i) \). \( M = D\mathcal{A}_G = (p_{ij})_{i,j \in V} \) is the transition probability matrix of the Markov chain. The \( t+1 \)-th step of the random walk is defined through \( P_{t+1} = M^TP_t \), and \( P_t = (M^T)^TP_0 \). I.e. the probability distribution \( P_t \) of the random walk can be calculated through the initial distribution and the transition probabilities obtained from \( G \). If \( P_t = M^TP_t \), the probability distribution is stationary and \( \pi = P_t \) is called steady-state. The mixing time is the number of steps \( t \) until \( P_t \) reaches \( \pi \), therefore it measures how fast the random walk converges to a limiting distribution.

The hitting time \( H(i, j) \) is the expected number of steps a random walk takes starting from \( i \) requires until \( j \) is reached, and \( \kappa(i, j) = H(i, j) + H(j, i) \) is called the commute time. Based on this we can introduce the return time: the expected number of steps a random walk requires until it returns to its origin (sometimes called recurrence time in literature). Polya showed 1921 in [15] that the probability of returning to the starting point converges to 1 as \( t \to \infty \) for a random walk on a \( d \)-dimensional lattice with \( d = 1 \) and \( d = 2 \), but converges to a value \( < 1 \) for \( d \geq 3 \) (e.g. \( p(3) = 0.340537 \ldots \) as shown later by Watson, 1939). Note, that the graph structures we analyze are 2-dimensional, i.e. \( d = 2 \), and our graphs are finite, therefore a return time is guaranteed to exist. Since our goal is, however, to draw conclusions on the opportunistic networks, we are not interested in the return time of a single random walk, but rather in the inter-contact time between two random walks, as described above.
Therefore, for a two walkers in two dimensions who meet after \( n \) steps can be considered as being a single path that starts at \((0,0)\) and returns to \((0,0)\) after \( 2n \) steps. This means that the probability that two random walkers in two dimensions meet is the same as the probability that a single walker in two dimensions ever returns to the starting point". We can therefore transform our analysis of inter-contact time to the analysis of return time. This is in line with the derivation of return time and inter-contact time for random walk on a circuit in [7], where return time and inter-contact time only differ by a constant factor. We conclude from [16, p. 17] and [15] that on a 2-dimensional graph two random walkers are sure to meet again after a contact, i.e. inter-contact time is sure to exist and can be modeled through return time, based on hitting time.

For a regular graph the return time equals \(|V|\) (see [14]). Therefore, for a \( k \)-grid (with wrapped borders to keep regularity) where \(|V| = k^2\) the return time is \(1/\sqrt{k} = k^{-0.5}\); a power-law correlation between grid-id and return time; we come back to this in Section VI. In face of finite graphs, special structures expose a strong heavy tail in inter-contact times from random walks, e.g. circles or lines [7]. Cai and Eun showed that the exponential drop of the inter-contact time is due to the bounded and therewith finite domain [13], i.e. the finite number of vertices for a random walk. Karagiannis et al. first reported on this dichotomy of power-law and exponential decay in [7], confirming and refining the findings of Chaintreau et al. [6] who were the first to report on the power-law nature of inter-contact times.

We now analyze the properties of the graph that affect the hitting time, and therewith the return time which we use to model inter-contact time. Analyzing the spectral properties of a graph in terms of eigenvalues and eigenvectors, the hitting and commute times can be calculated. Notably, hitting time \( H(s,t) \), steady-state distribution \( \pi \), and eigenvalues \( \lambda_k \) can be put into relation as follows [14]:

\[
\sum_i \pi(t)H(s,t) = \sum_{k=2}^{n} \frac{1}{1 - \lambda_k}
\]

(3)

The second largest eigenvalue \( \lambda_2 \) is of high importance due to the fact that (1) it is closest to 1 and therewith the denominator smallest and the fraction largest over all \( k \) in Eq. 3, (2) it describes nicely the distribution of the other eigenvalues \( \lambda_k \) through an upper bound, as shown in Figure 2. In the following we exploit the line of arguments that

1) we can reside to return time when modeling inter-contact time [16, p. 17],
2) the return time is based on hitting time [14],
3) \( \lambda_2 \) has strongest impact on hitting time and bounds the other eigenvalues \( \lambda_k \) for \( k > 2 \) (cf. Eq. 3, [14]).

In the following we use the spectral gap as a correlation for city graphs and grid-based graphs with the inter-contact times resulting the random and social mobility models. The spectral gap is defined as \( 1 - \lambda_2 \) (cf. denominator in Eq. 3). Calculation of the spectral gap is directly coupled with a random walk: Let \( A_G \) be the adjacency matrix of \( G \), let \( D \) be a diagonal matrix with \( (D)_{ii} = 1/d(i) \), and \( N = D^{1/2}A_G D^{1/2} = D^{-1/2}M D^{1/2} \). Matrix \( N \) has the same eigenvalues as \( M = DA_G \), but is symmetric and allows for easier calculation. Remember, that \( M \) is the transition probability matrix of the Markov chain. Note, that as \( N \) is symmetric its eigenvalues \( \lambda_k \) are real and it follows from the Perron-Frobenius theorem: \( 1 = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq -1 \). The spectral gap is defined through the second largest eigenvalue as \( 1 - \lambda_2 \). Note that the spectral gap results from the random walk on \( G \). Figure 4 shows the spectral gap of the analyzed cities together with the spectral gap of grid graphs. Here, we map a city graph to a grid graph based on the spectral gap. In the following section we use the relation between spectral gap and inter-contact time—described above and in Eq. 3—to correlate the spectral gap with parameters obtained for the mobility model through fitting of simulation data.

VI. CORRELATION BETWEEN GRAPH AND MOBILITY

Based on the simulations performed in Section III we extracted parameters through model fitting in Section IV that capture the important metric of inter-contact time with few comparable variables shape \( a \), power-law slope \( b \), and exponential \( c \). In the previous Section V an analytical analysis of random walks on graphs has been performed and shown that the spectral gap has strong impact on the inter-contact time and is easy to calculate based on eigenvalue analysis. In this section we combine these results and show the correlation between the fitted parameters and the spectral gap.

The most important parameter of inter-contact time is the power-law slope (parameter \( b \) in Eq. 1 and Eq. 2). Chaintreau et al. showed in [6] how this value affects the forwarding performance of opportunistic networks and underlined its importance. Figure 5 shows the correlation between spectral gap
and power-law slope for all combinations of grid graph and city graph with social mobility and random mobility. For grid-based graphs the spectral gap provides a smooth correlation for the power-law slope, for both the social and random mobility model. Grid-id on lower x-axis and spectral gap on upper x-axis have been aligned, and city graphs aligned to the upper x-axis based on their spectral gap (see Table I). The correlation for the random mobility model is based on the argumentation in Section V, the correlation with the social mobility model, however, is surprising but makes our approach more powerful and generally applicable. The minimum at grid-id 40 in the social mobility model results from the transmission range of devices which is configured to be 50 m (cf. Section III), i.e. the power-law slope decreases when the grid becomes more dense, up until the point where distance between two parallel edges equals the communication range and two devices on parallel edges can be in contact. Figure 5 further shows the city graphs, in the order of Figure 4 and aligned on the spectral gap (upper x-axis). We see that, again, the spectral gap is correlated to the slope for random mobility on the city graphs (lower short line), but provides no correlation for the social mobility model on city graphs (upper short line). We can conclude that the graph structures of cities are not sufficiently diverse to impact the mobility model as in the other cases where we see a clear correlation between spectral gap (based on the graph) and power-law slope (resulting mobility).

In Figure 6 we show further correlations: for shift a we have a correlation for grid-based graphs shown in Figure 6a, but not for city graphs, there the shift is very stable on a single value. Exponential c can be correlated nicely with spectral gap in both grid graphs shown in Figure 6b and to some degree with city graphs. Recall that we easily deduced a power-law relation between return time and grid-id for random walk in Section V for regular grid-based graphs with wrapped borders. Figure 3 shows the power-law relation for grid-based graphs between grid-id and spectral gap. Although our grids are not completely regular (at the borders) the relation sufficiently holds for our purpose. We can conclude that the correlation of return time with spectral gap holds sufficiently for our graphs, and inter-contact time is correlated with the spectral gap—which supports our line of arguments proposed earlier. For completeness, we show the power-law slope $b$ with variation of the $\alpha$ parameter for the SWIM model in Figure 6c.

We conclude that the underlying graph has strong impact on resulting inter-contact time for random walk mobility on grid-based graphs and on real-world city graphs. Therefore, the choice of the graph for simulations is of high importance. While the social mobility is clearly influenced over different grid-based graphs, real-world city graphs do not differ sufficiently in structure to influence the inter-contact behavior of a social mobility model. Specifically, while the city graph structure does impact the mobility, real-world cities are not sufficiently diverse to further impact the mobility. Therefore we can give a positive answer on the assumption of Rhee et al. who stated in [11] that “[…] these tendencies are likely caused by human intentions in deciding travel directions […] but not by geographical constraints such as roads […]”.

VII. RELATED WORK

The first work to identify the power-law distribution of inter-contact times has been performed by Chaintreau et al. based on analysis of real-world traces [6]. Karagiannis et al. refined these results in that the distribution of inter-contact times does not only follow a power-law, but exhibits an exponential decay, starting from roughly half a day [7]. Furthermore, they showed that already simple mobility models can exhibit power-law and exponential decay in inter-contact times, e.g. random walk on a circle. Cai and Eun found that the time where the distribution turns exponential is due to the finite domain where mobile devices move [13]. Several mobility models have been proposed that can generate the behavior of power-law and exponential inter-contact times [9], [10]. While work has found that the impact of the underlying graph where mobile devices move is of high importance for routing protocols, e.g. [17], we are to the best of our knowledge the first to provide insight into the correlation between graph and inter-contact behavior.

Large number of work has been done in the analysis of cities for understanding how the structure of city impacts social
development (e.g., slums or crime) in different cities. E.g., Volchenkov and Blanchard analyze city structures through random walks to find accessible parts of the city and compare flow-based properties of the graphs to get an understanding on social behavior [18]. An excellent overview on city graphs gives the book by the same authors [19]. Random walks have been studied in many context, a great survey on random walks on graphs is given by Lovász in [14].

VIII. CONCLUSION AND OUTLOOK

Understanding mobility is crucial for design, analysis, and deployment of mobile opportunistic networks. Whereas most mobility models are defined on a plain playground, a graph-based playground that restricts movement of devices is more realistic. We used spectral graph analysis to study the impact of the underlying graph on the mobility behavior using synthetic grid-based graphs and real-world city street graphs. Through simulative evaluation of random and social mobility we gathered data and performed fitting to extract parameters of inter-contact time distribution. We analyzed the impact of the graph structure from an analytical point of view and found that the spectral gap has a strong correlation to the parameters of the inter-contact time distribution.

While we studied impact on inter-contact time, other mobility measures are of importance, e.g. the graph structure can influence the frequency of contacts, which impacts the probability of multi-hop paths.

ACKNOWLEDGMENTS

Christoph P. Mayer is supported by the Landesstiftung Baden Württemberg gGmbH within the SpVoNet project. Oliver P. Waldhorst is supported by the “Concept of the Future” of KIT within the framework of the German Excellence Initiative. The authors thank Prateek Gaur who worked on cleaning up the city maps during his DAAD-funded internship at KIT. Thanks to Robert Görke for sharing his insightful ideas on graph metrics.

REFERENCES


