Robustness Analysis of Mobile Ad Hoc Networks Using Human Mobility Traces

Dongsheng Zhang and James P.G. Sterbenz*

*Information and Telecommunication Technology Center
Department of Electrical Engineering and Computer Science
The University of Kansas, Lawrence, KS, 66045, USA
{dzhang, jpgs}@ittc.ku.edu
†School of Computing and Communications (SCC) and InfoLab21
Lancaster LA1 4WA, UK
jpgs@comp.lancs.ac.uk
www.ittc.ku.edu/resilinets

Abstract—With the rapid advancement of wireless technology and the exponential increase of wireless devices in the past decades, there are more consumer applications for MANETs (mobile ad hoc networks) in addition to the traditional military uses. A resilient and robust MANET is essential to high service quality for applications. The dynamically changing topologies of MANETs pose a huge challenge to normal network operations. Furthermore, malicious attacks against critical nodes in the network could result in the deterioration of the network. In this paper, we employ several real-world human mobility traces to analyze network robustness in the time domain. We apply attacks against important nodes of the human topology and compare the impact of attacks based on different centrality measures. Our results confirm that nodes with high betweenness in a well-connected large dynamic network play the most pivotal roles in the communication between all node pairs.

Keywords—Robustness; Dynamic networks; Human mobility traces; Resilience and survivability; MANETs; Graph centrality

I. INTRODUCTION AND MOTIVATION

MANETs (mobile ad hoc networks) have existed for more than two decades and were originally used in military tactical networks. Mobile devices including cellphones, tablets, and wearable devices have proliferated in an unprecedented way in the past several years. Smart phones have become indispensable in peoples’ daily lives for the purposes such as shopping, personal banking, and communicating. MANETs now have applications for a variety of situations including sensor networks, emergency services, education, and entertainment. MANETs now have applications for a variety of situations including sensor networks, emergency services, education, and entertainment. MANETs are established in such a way that mobile nodes can dynamically self-organize into an ad hoc network without infrastructure. Communication between node pairs relies on multihop traffic forwarding from other nodes. The robustness of the underlying wireless dynamic topology is the key to the normal functioning of the network applications in mobile devices.

MANETs inherit some problems from traditional wireless networks due to the open channel of wireless medium. Information about device positions can be obtained using passive eavesdropping [1]. This is difficult to detect as it does not produce traffic within the network. Malicious attackers could determine the most significant nodes in a network in terms of their roles of providing connectivity to other nodes. We previously proposed an approach to model malicious attacks in MANETs [2], [3]. Dynamic networks within a certain window size are aggregated and represented by a weighted static adjacency matrix. Weighted centrality metrics are used as node significance indicators. Based on calculated weighted centrality metrics of each node, we apply attacks against nodes with high centrality values prior to low centrality nodes. A major constraint of this approach is that node positions are assumed to be known in advance so that weighted centrality values can be calculated. This makes the modeling approach only applicable in very limited scenarios such as pre-programmed networks.

There are many scenarios in which mobile devices are attached to or operated by human beings. In this paper, we study and analyze an important subset of MANET applications of ad hoc communications in conferences, campuses, local fairs, and theme parks. The publicly available GPS traces representing humans’ walking provide a realistic moving pattern for our analysis [4]. As there is a certain level of correlation between consecutive steps, the high centrality nodes of the current time instance can be used to predict the significant nodes in the near future. Based on real-world human mobility traces, we further examine the correlation between dynamic topologies within a certain window size. We analyze network robustness from the perspectives of structural change of dynamic topologies and network flow robustness under malicious attacks. We assume the attackers are able to obtain the global information and attack nodes with more essential roles. After identifying the significant nodes, topology control techniques used in wireless networks such as adapting transmission power [5] can be exploited to minimize the variances of node roles. The rest of the paper is arranged as follows. In Section II, we introduce the background and related work for this paper, including MANET security, robustness measurements, and mobility models. Section III provides details of the data sets used for this paper. We provide a comprehensive robustness analysis for several selected scenarios in Section IV. Finally,
we summarize this paper and discuss the steps for future work in Section V.

II. BACKGROUND AND RELATED WORK

We have used synthetic mobility models such as Random Waypoint [6] and Gauss-Markov [7] to evaluate network robustness of MANETs [2], [3]. However, the way a person moves is not random, and synthetic mobility models cannot accurately simulate the real-world moving patterns of human beings. The previous work is based on the knowledge of complete historical topology information, which makes it very limited to real-world applications. In this section, we first introduce potential security threats in MANETs, which could be exploited by malicious attackers to obtain global network topology. Next, we present and compare existing robustness metrics and select flow robustness as the metric for the analysis of dynamic mobile networks [8]. Finally, we present different mobility models used in MANET research and introduce the real-world mobility of human movement in this work.

A. Threat to MANETs

Compared to fixed wireless networks, MANETs are more vulnerable to information and physical attacks [1]. MANETs inherit the traditional problems in wireless networks, including the open wireless medium, unprotected channel, and hidden-terminal problems. Due to the mobility of each node, dynamic and intermittent connectivity poses even greater challenges for the normal operation of MANETs [1]. Malicious attackers can bypass intrusion detection systems and impersonate part of the network undetected. An approach has been proposed to infer routing topology for wireless sensor networks based on the measurement received in the sink [9]. Empirical human contact networks were shown to be predictable [10], which could be utilized by malicious attackers to disrupt the normal operation of MANETs. A mobile multi-agent-based framework has been proposed to discover topology in the ad hoc wireless environment [11]. Topology prediction in mobile wireless networks has been used to improve network throughput and delay [12]. Malicious attackers that might camouflage themselves as a member of the network or hijack the existing node in the network can use a similar approach to obtain and predict the network topology [13]–[15].

B. Robustness metrics

$k$-connectedness describes how well networks survive node or link failures without being partitioned [16]. However, the algorithmic complexity for calculating the $k$ is NP-complete. Algebraic connectivity and other graph spectrum metrics also provide informative characterization of network robustness [17]–[20]. One of the main disadvantages of using these metrics to measure robustness is that they provide the same measurement values if the network is disconnected. Particularly in a MANET environment, the network might be partitioned into several small components from time to time. Giant component size captures the size of the largest connected component [21]; however, when the network consists of several node clusters of similar size, giant component size cannot provide an accurate description of network connectivity. The flow robustness metric can capture both the number of components and the size of each component [8]. It is computed as the ratio of the number of reliable flows to the number of total flows in the network [8]. A flow is considered reliable if there exist at least one path between source and destination. The total number of traffic flows is $n(n-1)$ for a connected $n$-node network. This metric captures the ability of network nodes to communicate with each other. The value range for flow robustness is $[0, 1]$, where 1 indicates that all the nodes can communicate and 0 means there is no node-pair communication within the entire network. In this paper, we focus on the analysis of MANETs, and flow robustness is the best metric fit for our scenario since it can accurately describe communication ability between node pairs in a disconnected network.

C. Mobility models

A survey of synthetic mobility models for MANETs research and simulation has showed that the selection of mobility model has great impact on the performance of MANET protocols [22]. Mobility models including Random Walk [23], Random Waypoint [6], Gauss-Markov [7] might provide movement patterns that match the real-world scenarios if appropriate parameters are chosen [22]. However, the selection of parameters could be very difficult as there is no guidance indicating what parameters should be used to represent a certain real-world scenario. Furthermore, most synthetic traces have not been validated against real-world movements. The publicly available real-world mobility traces collected by different organizations and institutes have been summarized [24]. A trace-validated mobility model called SLAW (Self-similar Least Action) that simulates human walks has been proposed [25]. Lévy-walk nature of human mobility has been validated with heavy-tail distribution of flight length and pause-time [4], which cannot be captured by the existing synthetic mobility models such as Random Waypoint and Gauss-Markov. In this paper, instead of using synthetic mobility traces, we analyze the robustness of several real-world traces publicly available from CRAWDAD [26]. In the next section, we will provide more details regarding these traces.

III. HUMAN MOBILITY TRACES

We choose human mobility traces collected from five different sites [26]. These traces were originally used to study the statistics of human mobility pattern and similarity between human walks and Lévy Walks [4]. The data sets provided are 30 seconds average of GPS coordinates recorded using Garmin 60CSx handhold devices. Table I presents basic information regarding the five traces.

A. Data sets

There are occasional cases that GPS signals cannot be received when GPS holders move indoors in the original data
sets. In addition, the average speed during a 30-second window for some nodes is calculated as high as 200 m/s based on the original trace. We remove the traces from the data sets if there is any occurrence of velocity higher than 50 m/s during a 30-second time window. NewYork traces were collected from volunteers living in the Manhattan NY area, and they traveled by cars and buses. The KAIST and NCSU traces were collected in two campuses, one in Korea Advanced Institute of Science and Technology and the other in North Carolina State University. The Orlando traces were obtained from volunteers who spent their holidays in the Disney World. The StateFair traces were obtained from participants who went to the North Carolina State Fair. This set of traces were collected outdoors, and the size of the StateFair site is the smallest of all. As each trace in the same site lasts for different durations of time, we truncate all traces based on the minimum trace time for each site. The trace duration used for each site is listed in second column of Table I. The first column provides the number of traces collected for each site.

### B. Trace statistics

In real-world MANET scenarios including the five sites used here, the actual network density depends on the number of nodes and transmission range of each node. The number of nodes in MANETs is based on the number of traces provided in the data sets even though the node number in a real case could be significantly higher since only a small portion of people are selected as candidates for trace collection. Theoretically, transmission range can be adjusted by increasing the radio power of the handheld devices. Noting that handheld devices are mostly battery-powered, the transmission range cannot be increased infinitely. We choose 250 meters to 1000 meters as an acceptable range.

In Table I, we present the average pause-time of all nodes for each site. Pause-time is the duration a person halts before moving to another location. The StateFair site has the smallest average pause-time, which accounts for about 7% (380/5340) of the entire trace time. In contrast, the pause-time of KAIST and NCSU account for a very high percentage of entire trace duration, which indicates the dynamic topologies remain stable most of the time. We analyze the distribution of node velocities for all 30-second windows. The CCDF (Complementary Cumulative distribution function) of node velocities are presented in Figure 1. Both x and y axes use log-scale. It is apparent that node velocities follow a non-linear distribution. For NCSU and StateFair sites, almost 90 percent of velocities are distributed below 1 m/s while the maximum velocity for NCSU is around 20 m/s. The velocity of the New York site is generally higher than the other 4 sites, and probability of velocity higher than 5 m/s is approximately 0.1. About 80% of velocities in the KAIST site is less than 1 m/s while the maximum velocity could be as high as 50 m/s. The non-uniform distribution of node velocities cannot be captured by synthetic mobility models such as random waypoint and Gauss-Markov. This also poses a big challenge for MANET robustness as whenever nodes starts moving with high speed, the network performance cannot be guaranteed.

![CCDF of average node velocities](image)

We also present average flow robustness, node degree, average giant component size of each site using 3 different transmission ranges in Table I. With the same transmission range, the StateFair site has the highest average flow robustness of all. When the transmission range is 1000 m, the network becomes a full-mesh as the average giant component size is equal to the total number of nodes. The NewYork site has the lowest average flow robustness. Even when the transmission range is set to 1000 m, the average node degree is 2.8 noting that there are 39 nodes in total. The average flow robustness of the Orlando site is slightly higher than NewYork but still presents a very low network connectivity. The average giant component size of the KAIST site with 500 and 1000 m transmission range are 73.1 and 75.2, while the average degree almost doubles from 27.8 to 53.9. This means that there are several nodes far apart from the largest node clusters, and the increase of transmission range only leads to a higher connectivity within a local cluster with other nodes.

### Table I

<table>
<thead>
<tr>
<th>Site</th>
<th>Min. trace duration [s]</th>
<th>Pause-time [s]</th>
<th>Avg. flow robustness</th>
<th>Avg. node degree</th>
<th>Avg. giant component size</th>
</tr>
</thead>
<tbody>
<tr>
<td>KAIST (83)</td>
<td>15180</td>
<td>5440</td>
<td>0.509</td>
<td>0.777</td>
<td>0.822</td>
</tr>
<tr>
<td>Orlando (41)</td>
<td>7860</td>
<td>1546</td>
<td>0.167</td>
<td>0.208</td>
<td>0.233</td>
</tr>
<tr>
<td>NewYork (39)</td>
<td>4440</td>
<td>1382</td>
<td>0.022</td>
<td>0.041</td>
<td>0.190</td>
</tr>
<tr>
<td>NCSU (35)</td>
<td>6180</td>
<td>2490</td>
<td>0.112</td>
<td>0.238</td>
<td>0.488</td>
</tr>
<tr>
<td>StateFair (19)</td>
<td>5340</td>
<td>380</td>
<td>0.825</td>
<td>0.993</td>
<td>1.000</td>
</tr>
</tbody>
</table>

The average flow robustness of all sites

The average giant component size of each site using 3 different transmission ranges in Table I.
still isolated from the giant component.

The normal functioning of MANETs requires network connectivity to remain above certain level since all routes are established in real time. For the flow robustness analysis, we will focus on StateFair, NCSU, and KAIST due to space limit. Snapshots of the three sites are provided in Figure 2. Both x and y coordinates are the distance from a reference in meters. Each color represents the snapshot of start (blue), middle (red), and end point (yellow) of the traces respectively. StateFair traces are confined within a relatively small area (1200 × 1000 m²). Both NCSU and KAIST site span a large area with most nodes clustering around a particular part of the map, as shown in Figures 2b and 2c.

IV. DYNAMIC TOPOLOGY ANALYSIS

In this section, we first examine autocorrelation of time-varying flow robustness. Next, we evaluate the change of high centrality nodes over time. Finally, we apply attacks against high centrality nodes using different window sizes. All the analyses performed in this paper are from the topological perspective, which does not include network simulation involving protocols of multiple network stacks. We want to understand how network robustness is affected by the dynamic topologies.

Note that the nodes are mobile most of the time, but topology does not necessarily change as long as nodes stay within transmission range of each other. Furthermore, in terms of identifying the significant nodes in the network, high centrality nodes might remain the same if the topology changes only slightly.

A. Time-varying flow robustness

Network robustness changes over time as nodes disconnect and reconnect to others constantly. We compute the flow robustness for each 30 s topology snapshot, and then calculate the autocorrelation between time-varying flow robustness. In order to compare fairly among three different sites, we only analyze first 4500 s trace data. Figure 4 presents time-varying flow robustness for StateFair, NCSU, and KAIST sites. Figure 5 shows the corresponding autocorrelation coefficient for 3 sites ($R_{StateFair}$, $R_{NCSU}$, $R_{KAIST}$) using different transmission ranges. For the StateFair site, flow robustness is always 1 for 1000 m transmission range; hence we do not provide the autocorrelation as the variance is 0. For 500 m transmission range, StateFair flow robustness falls below 100% for a short period of time after 3600 s and remains at 100% for the rest of time. This indicates a very high network connectivity as we
can also see that average giant component size is 18.9 from Table I. For StateFair traces with 250 m transmission range, flow robustness fluctuates between 50% and 100%. $R_{\text{StateFair}}$ displays a linear decrease to 0 within 300 s, and certain levels of periodicity are observed since participants of North Carolina State Fair are moving in a relatively confined area. Flow robustness of the NCSU site increases during specific short time windows as shown in Figure 4b. $R_{\text{NCSU}}$ with a 250 m transmission range is weaker than 500 and 1000 m transmission ranges. As shown in Figure 2b, the majority of nodes move slowly within a small area in the map, and a smaller transmission range causes the nodes to disconnect from others more frequently.

Flow robustness of the KAIST site shows the strongest autocorrelation of all three sites. Similar to $R_{\text{NCSU}}$, $R_{\text{KAIST}}$ is higher with longer transmission range. $R_{\text{KAIST}}$ with 500 and 1000 m transmission ranges presents a more linear decrease of autocorrelation, and time-varying flow robustness is strongly correlated within 1800 s. $R_{\text{KAIST}}$ with 250 m transmission range decreases to 0 significantly faster than with 500 and 1000 m transmission ranges. As shown in Figure 4c, flow robustness oscillates far more frequently between 0.35 and 0.7 with 250 m transmission range. The giant component gets partitioned into half of the original size. Whenever flow robustness goes up or down to a new state, the network connectivity remains for a certain period of time. Remediation measures can be taken to improve network connectivity, such as adjusting transmission power of certain nodes or adding extra static nodes to bridge the network if there is a repeated pattern of network disconnection. For the rest of the section, we narrow down our scenarios to 250 m transmission range for StateFair, 500 m for KAIST, and 1000 m for NCSU.

B. High centrality nodes change over time

Centrality metrics (degree, closeness, and betweenness) have been used to measure relative node significance within a network in various areas [27]–[29]. Node degree centrality is a measure of local node communication ability. Both betweenness and closeness centrality are related to the shortest paths between node pairs. Node betweenness is a measure of the degree to which it enables communication between other node pairs. A disadvantage of using betweenness is that if none of the nodes falls on the shortest paths of other node pairs, they receive the same 0 betweenness value. A node’s closeness is a measure of the extent to which its communication capabilities are independent of the functioning (or malfunctioning) of others. We use these three centrality metrics to measure relative node significance for each snapshot of topology. We compute the top 20% highest centrality nodes for each 30 s snapshot and round it up to an integer value. Node centrality metrics are calculated adaptively after the removal of each node [30]. Then we compare how many common high-centrality nodes between two different snapshots. For example, in the StateFair trace, we calculate a set of 4 nodes with highest degree, closeness,
and betweenness respectively. We compare the node sets with a range of time differences and then calculate the average number of common nodes for each range. In Figure 6a, for two snapshots of 30 s time difference, the average number of common nodes between them is approximately 3 for all three centrality metrics, which indicates a relatively high similarity. It is apparent that with the increase of time difference, each node has a larger deviation from the original position. When window size is increased to 300 s, the average number of common nodes decreases to less than 2. For window sizes that are larger than 600 s, the average common nodes is approximately 1 with no more decrease. This can be explained as nodes in StateFair site move within a confined area and have a high chance to meet each other repeatedly after a certain period of time.

For the NCSU site, the top 7 nodes with the highest centrality are compared across different time window sizes. There are more than 6 nodes shared between the top centrality nodes with 30 s window size, which indicates extremely slight change of topology structure. Even with a window size of 900 s, an average 5 out of 7 nodes are the same. This would explain the difference of centrality-based attacks using different window sizes in the next subsection.

The average number of common high centrality nodes within a 30 s window size for the KAIST site is about 15 out of 17 for all centrality metrics, which also shows a high correlation within a 30 s window. When the window size increases, the common nodes with high betweenness decreases faster. This is because shortest paths between all nodes pairs are more sensitive to the change of node positions in a comparatively larger network and node betweenness highly relies on the count of shortest paths. When the window size is increased to 900 s, an average of approximately 10 of 17 nodes are the same for two snapshots of topologies.

C. Flow robustness under centrality-based attacks

We apply centrality-based attacks against the above three scenarios using window sizes of 30, 300, and 900 s. We compare the average network flow robustness of centrality-based attacks with random failures. Five different levels of damage are applied adaptively with up to 50% of the total number of nodes being removed in each scenario.

For StateFair scenarios, betweenness-based attacks have the heaviest impact on network flow robustness. With 300 s window size, the gap between random failures and centrality-based attacks decreases compared to attacks using 30 s window size. With 900 s window size, the difference between random failures and centrality-based attacks becomes smaller. In Figure 8, the baseline flow robustness is only 50% for the NCSU site with 1000 m transmission range. As shown in Figures 2b and 2c, nodes in this site span a large campus area, while the majority of them construct a connected component with the rest being isolated most of the time. The difference between the impact of each centrality-based attack on the network is modest for 30, 300, and 900 window sizes. This is because network structure remains relatively stable within the giant components of the NCSU site shown in Figure 6b. In addition, with failure rate higher than 0.3, betweenness-based attacks have less impact on the network than attacks based on degree and closeness. When the network is partitioned into small fully-connected components, all the remaining nodes have the same betweenness of 0, which makes them indistinguishable from each other. With a total of 83 nodes in the KAIST site, a significant difference among betweenness-based attacks and other metrics for 10% and 20% node removal is observed in Figure 9a. Degree and closeness become better node significance indicators when there are more than 30% of nodes removed from the network. As the calculation of node degree is based on its neighborhood, it makes sense that local network structural change becomes dominant in overall connectivity of the entire network. Both betweenness and closeness metrics are calculated based on the global shortest paths of the entire network. However, the betweenness metric provides more accurate indication of node significance in terms of providing connectivity of the entire network. The global influence of node closeness is less than node betweenness because the sum of the inverse ofarness to every other node also takes into account all local nodes within the neighbors.

Figure 10 presents the giant component component sizes under centrality-based attacks with a 30 s window size. The relationship among giant component sizes under different cen-
Fig. 7. Centrality-based attacks using different window sizes for StateFair trace

Fig. 8. Centrality-based attacks using different window sizes for NCSU trace

Fig. 9. Centrality-based attacks using different window sizes for KAIST trace

Fig. 10. Giant component size under centrality-based attacks with 30 s window size
tality attacks are similar to the corresponding flow robustness under attacks. When the percentage of nodes being attacks reaches 0.3, the slope of the curve for betweenness-based attacks becomes less sharp than the degree- and closeness-based attacks. Even though average giant component size is more than 20 as shown in Figure 10c, all nodes in each fully-connected component have the same betweenness value of 0.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we evaluated the network robustness of real-world humans’ movement traces. We present the time-varying flow robustness and its autocorrelation. Periodical patterns are observed within confined areas. We study how high centrality values of nodes change over different time window sizes. The highest betweenness centrality change more frequently in large networks. High betweenness nodes play a significant role in providing connectivity to the communications of other nodes in the network, particularly with relatively large network size, whereas node betweenness fails to differentiate node significance in networks consisting of isolated fully-connected components. Future work includes the analysis of network performance using ns-3 by running different MANET routing protocols [31]. In addition, robustness analysis of humans walking using DTN (delay-tolerant network) routing protocols will be conducted.

ACKNOWLEDGMENTS

We would like to thank Mohammed Alenazi for the discussion of this work. This research was supported in part by NSF grant CNS-1050226 (Multiayer Network Resilience Analysis and Experimentation on GENI) and by NSF grant CNS-1219028 (Resilient Network Design for Massive Failures and Attacks).

REFERENCES