Abstract—In this paper, we examine how the paths between any two nodes in a mobile opportunistic network change with increasing number of hops a message can follow. At the core of our analysis is the all hops optimal path (AHOP) problem that has been explored in the context of network capacity analysis of static networks. We represent the total delay of a route from a source node to a destination node as additive weight and use the number of encounters as a representation of bottleneck weight. First, we construct a static (contact) graph from the meetings recorded in a human contact trace and then analyze the change in these two metrics with increasing number of hops. Next, we aggregate all the contact events in a time interval and construct several time-aggregated graphs over which we calculate the capacity metrics. Although, we observe differences in the properties of the static and the time-aggregated networks (e.g., higher connectivity and average degree in static graph), our analysis shows that second hop brings most of the benefits of multi-hop routing. However, the optimal paths are achieved at further hops, e.g., hop count \( h \approx 4 \). Our finding, which is also verified by simulations, is paramount as it puts an upper bound on the hop count for the hop-restricted routing schemes by discovering the optimal hop count for both additive and bottleneck weights.

I. INTRODUCTION

Mobile opportunistic networks rely on short-range radios (e.g., Bluetooth, WiFi Direct) to transmit data between two nodes and exploit the mobility of nodes to physically carry messages from one location to another. This kind of operation is favourable for several reasons: (i) interconnection without dependency on the infrastructure, (ii) hop gain due to direct link between the transmitter and the receiver [1], (iii) spectrum reuse gain, and (iv) mobile data offloading. On the other hand, it is more challenging to provide guaranteed performance in such a dynamic network. Mobility of nodes results in intermittent connections and raises uncertainty in the network topology. The major challenge is hence the lack of global knowledge at the nodes to decide on the optimal forwarding paths or other networking tasks.

Epidemic protocol [2] is the simplest protocol that does not rely on any information about the network (e.g., about nodes and connections) but greedily replicates a message to every node that does not have it. Albeit being desirable due to its simplicity, epidemic protocol over-consumes network resources (e.g., waste of bandwidth due to too many replications) and may not perform well under resource-restricted networks (e.g., short contact duration or small buffer capacity).

A less greedy solution is hop-limited protocol [3], [4] which limits the journey of a message in the network to maximum \( h \) hops. More sophisticated protocols aim to balance the tradeoff between delivery ratio and resource consumption by tuning the protocol parameters, e.g., the maximum number of replications [4], lifetime of a message in the network [5], replication/forwarding logic [6], and so on. However, no matter how optimized the protocol is, the performance of a mobile opportunistic network also strongly depends on the node mobility [7]. More specifically, two properties related to node mobility are paramount: contact duration and inter-contact time duration. Contact duration is the time two nodes stay connected while inter-contact time is the time elapsed between two consequent contacts among two nodes. Both determine the transmission capacity (i.e., how much data can be transmitted) as well as the speed of change in the network topology.

Although opportunistic routing protocols have been studied extensively from many perspectives, the effect of hop limitation is not yet fully understood. In this work, we focus on a general hop-limited protocol and analyze how hop count limitation affects the network performance. We do not assume any sophisticated protocol as we attempt to understand the effect of node mobility rather than other parameters (e.g., number of replicas, time-to-live duration, effect of social-structure). Instead of modelling the node mobility analytically, we focus on the change of the network graph as a consequence of node mobility using the real human traces. First, given a mobility trace we construct a connectivity graph and explore the change in the network performance indicators, e.g., fraction of reached nodes, using the solution for all hops optimal path problem (Section III). This static graph aggregates all the contacts into a contact graph.

Second, instead of aggregating all the contact history into a single connectivity graph, we sample the network on regular time intervals and analyze each snapshot separately (Section IV). Social-aware opportunistic routing schemes largely tend to aggregate contacts either using a sliding window approach (e.g., BubbleRap [8]) or accounting for the whole past events (e.g., SimBet [9]) to derive metrics about the nodes, e.g., centrality, average degree. But, [10] shows that both short and long time windows fail to differentiate the nodes according to their social metrics. Similarly, we compare these two approaches for hop analysis to see how local observations
agree/disagree with the analysis of the static graph.

Finally, we simulate a hop-limited routing scheme to see the change in performance during operation (Section V). Obviously, the third analysis explains the real effect of the hop count. However, the formers are also useful for estimating the network performance if they can provide similar conclusions to the third analysis.

The main contributions of our paper are to answer the following questions using AHOps algorithm on static graph and time-aggregated graphs, and compare them with simulation results:

- Q1: How is the average time to send a packet from one arbitrary node to another arbitrary node affected by hop restriction \( h \)?
- Q2: How is the fraction of nodes reachable from one arbitrary node to another arbitrary node affected by \( h \)?
- Q3: How is the delivery ratio from one arbitrary node to another arbitrary node affected by \( h \)?

II. SYSTEM MODEL

We consider a network of \( n \) mobile nodes. A contact starts when two nodes come in transmission range of each other, and ends when they can not maintain a connection. Nodes can exchange messages only during contact time.

In a hop-limited routing scheme, each message has a hop count field that shows the number of routers this message has followed. When a message is forwarded from one node to another, the hop count of this message is incremented by 1 at the receiving node. If the hop limitation is \( h \), the message can only be forwarded \( h \) hops.

We use the following human contact traces in our analysis Infocon01 [11], Cambridge and Infocon06 (collected by Haggie project [12] and downloaded from [13]), whose basic properties are listed in Table I.

<table>
<thead>
<tr>
<th>Trace</th>
<th>Duration</th>
<th># of devices(n)</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infocon01</td>
<td>( \approx ) 3 days</td>
<td>41</td>
<td>Conference participants</td>
</tr>
<tr>
<td>Cambridge</td>
<td>( \approx ) 11 days</td>
<td>36</td>
<td>1st and 2nd year undergraduates</td>
</tr>
<tr>
<td>Infocon06</td>
<td>( \approx ) 4 days</td>
<td>78 (and 20 stationary devices)</td>
<td>Conference participants</td>
</tr>
</tbody>
</table>

III. ALL HOPS OPTIMAL PATHS (AHOP)

One common practise in exploring the network capacity is to analyze the network topology. In opportunistic networks, the topology changes frequently. However, human contact traces that are collected for a long duration can reflect the stationary contact probabilities which then can be used for constructing the expected network topology. Given the network’s stationary meeting characteristics, we can model the encounter events as a graph \( G = (V, E) \) where vertex set \( V \) is the set of nodes and edge set \( E \) is the set of all possible pairwise contacts among nodes. In this graph, there are \( n = |V| \) vertices and maximum \( |E| = O(n^2) \) edges for which we assign weights \( (w) \) based on the contact characteristics of the two connecting vertices.

Given a source node \( s \), our aim is to find the shortest (e.g., fastest) path to a destination \( d \) that has at most \( k \) hops for each \( k, 1 \leq k < h \). In the most general form, this problem is known as all hops optimal path problem (AHOps) [14]. Fig. 1(a) shows an example connectivity graph. An edge in this graph indicates that the two nodes has met at least once. Fig. 1(b) depicts some of the paths from \( A \) to \( L \) (the leaf nodes). Number near each leaf shows the hop count of the path. The optimal path among all these paths depends on the edge weights.

Let \( p \) be a path from \( s \) to \( d \) consisting of vertices \( \langle s = n_0, n_1, \ldots, d = n_k \rangle \) in the given order, and let \( w(n_i, n_j) \) be the weight of the edge between \( n_i \) and \( n_j \). Denote by \( w(p) \) the weight of a path \( p \) and define it as a function of weights of edges along this path:

\[
w(p) = f(w(n_0, n_1), w(n_1, n_2), \ldots, w(n_{k-1}, n_k))
\]

where \( k \leq h \) and \( n_i \in V \) for all \( i \). If \( f \) is additive, the weights are said to be additive weights; if \( f \) is maximum/minimum, the weights are bottleneck weights [14]–[16]. Both metrics are of our interest as they reflect different performance requirements in networks; bottleneck weight matches the minimum bandwidth requirements needed for a service whereas additive weight is more appropriate to measure the total delay of a path. As for opportunistic forwarding, we consider both weights:

(i) additive weight: expected time to reach a target node is the total expected inter-contact time of the selected path. Hence, we compute the total delay between \( s \) and \( d \) as the additive weight of the path \( p \):

\[
w(p) = \sum_{(n_i, n_j) \in p} w(n_i, n_j).
\]

(ii) bottleneck weight: as encounters are probabilistic, we aim to select the edges that will appear with high probability. In a sense, given the number of encounters among nodes, a routing scheme tries to decrease the risk of relying on less probable encounters by selecting the most probable paths. An additive weight may not properly choose the highest probable paths if only total delay is considered. Hence, we define the weight of an edge as the inverse of the number of encounters between the corresponding nodes. Then, the weight of \( p \) is
the maximum weight of the edges along this path:

$$w(p) = \max_{(n_i, n_j) \in P} w(n_i, n_j).$$

Let $l(p)$ denote the length of the path, i.e., number of edges. Then, we define an $h$-hop constrained optimal shortest path as the path with at most $h$ hops and yields the minimum weight among all paths between $s$ and $d$. More formally, $p_h^*$ is defined as follows:

$$p_h^* = \arg \min_p w(p) \text{ and } l(p) \leq h.$$

In [14], it was shown that Bellman-Ford algorithm provides the best solution with complexity $O(h|E|)$ for additive weights whereas for the simpler case of bottleneck weights authors propose an improvement upon Bellman-Ford with lower complexity. We should note that for human contact networks $h \ll n$ due to the small world structure [17]. Hence, although the worst case complexity is cubic for additive weights, average complexity would be much lower, e.g., $O(n^2)$. We compute the solution for AHOP by modifying the Bellman-Ford algorithm for hop restricted shortest paths [18]. We do not focus on efficiency of this calculation, but we should note that the running time of the algorithm can be decreased by pruning some of the edges via thresholding.

IV. ANALYSIS ON THE NETWORK SNAPSHOTS

Solving the AHOP problem as discussed above lets us discover the shortest paths and their capacities given hop restrictions. With this knowledge, we can grasp the network characteristics and the capacity limitations. One of the shortcomings of such an analysis is the loss of temporal network dynamics [19], [20]. In other words, we aggregate the snapshots of the network as if an occurring edge always exists there. On the other hand, temporal reachability graph [21] is not identical to the aggregated connectivity graph. In the above AHOP analysis, the time of appearing of the edges is lost. Take the graph in Fig.1(a) as example. Node $B$ can be a good relay for $A$ in reaching $E$, only if $B$ meets $E$ after meeting $A$. Fig. 2 illustrates two snapshots of the same graph. Neither of the networks is connected contrary to the network that is formed by aggregating these two snapshots. The encounter between $A$ and $B$ occurs after $B$-$E$ meeting which makes the path $A \rightarrow B \rightarrow E$ impossible. The static network derived from the encounters does not express this property of the network. Hence, some of the paths discovered by our AHOP solution may be causally impossible due to the time ordering of the edges in the temporal network. As a result, the connectivity of the network is overestimated and thereby also the resulting network capacity [10].

Another approach is to decrease the time window of the analysis. Instead of aggregating the whole contact events, we analyze the network several times and average the statistics over these observations. In this approach, events occurring inside a time window are aggregated, and at the end of the time window we have a snapshot of the network. The proper choice of the time window size is paramount and it depends on the network dynamics. If the network is changing abruptly, then the time window should be shorter compared to that of a more stable network. Understanding the effect of time window length is of our interest as many opportunistic routing protocols collect encounter information to predict future meetings or make informed forwarding decisions. However, the collected information becomes stale due to the change in the network dynamics. Therefore, the metrics that are derived from the aggregated network may be misleading for these protocols and result in performance degradation. Hence, time-varying graph (TVG) and related metrics (e.g., temporal shortest path) are argued to be more powerful [20], [21].

Time-varying graphs (TVG) have been studied under different contexts and with different terminology (see [20]). For example, ad hoc networks [22], animal networks such as ant colonies [23], and social networks [24] entail connections among the network entities that are changing (appearing, disappearing) over time and are complicated compared to the static graphs. TVGs extend the static graphs on time dimension by a presence function that shows if a particular edge exists at a specific time. An opportunistic network is obviously a TVG and previous works [17], [21] highlighted the benefits of TVG modelling for opportunistic networks. In this work, we take a midway approach and analyze the traces by breaking down the whole history of contacts into regular time intervals. This approach is obviously not as accurate as TVGs, however it alleviates the deficiency of the information loss by static graphs to some extent.

Suppose that the aggregation time is $T_{agg}$ for a trace of
time units. Then, we have $|T/T_{agg}|$ intervals. A contact event occurring in a specific interval is reflected as a link in the corresponding contact graph $G_i$, where $i$ is the index of the graph. For each link in $G_i$, we calculate average inter-contact time and number of meetings. Finally, we apply the previous analysis in Sec. III to each $G_i$. Fig. 3 summarizes the time granularity of static, TVGs, and time-aggregated graphs.

Lastly, we observe the change in network capacity while the network is under operation. We simulate a flooding-based forwarding protocol that is subject to hop limitations.

V. NUMERICAL EVALUATION

We analyze the traces using R and use ONE [25] for simulations. We generate network snapshots using Timeordered package [23] from a trace and a list of time intervals.

Fig. 4 shows the effect of hop count for all traces. We normalize the bottleneck and additive weights for each trace as we aim to show the change in the network dynamics rather than the actual values of the related metrics. As Figs.4(a) and 4(b) show, relaxing the hop limitation improves the network capacity: higher bottleneck capacity indicating the higher probability of path’s existence and lower additive weight indicating lower delays among the nodes. The most significant gain is achieved by letting two hop routing rather than a direct delivery (i.e., $h = 1$). Increasing to $h = 4$ still brings benefits especially for bottleneck capacity, however after $h \approx 4$ change in the considered metrics is negligible. Please note the consistency of the behaviour for all traces. Infocom05 has the highest delays followed by Cambridge and Infocom06. Regarding connectivity, we observe that only two hops are sufficient to reach all the nodes in the network from any node (not depicted). For $h = 1$, the reached fraction of nodes is (0.96, 0.82, 0.88) for Infocom05, Cambridge, and Infocom06, respectively. However, the existing path is not optimal as shown in Figs. 4(c) and 4(d).

Recall that in AHOP an $h$-hop optimal path is the path with the minimum path weight and is shorter than or equal to $h$ hops. We refer to the hop count that achieves the optimal path as the optimal hop count. From Fig. 4(c) and Fig. 4(d) we have two observations: all networks represented by the traces achieve optimal operation at very few hops, three to four. Second, generally speaking, the optimal bottleneck capacities are achieved at higher hops compared to additive capacity. This is due to the strength of weak ties [26] in case of additive weights; in bottleneck capacity calculation our algorithm avoids the weak links (links with low number of meetings) whereas in additive weights these weak links may be a fast move towards the target node.

In summary, for the research questions we listed in Section II, we have the following conclusions from the AHOP analysis. As for (Q1), nodes can be reached faster by relaxing hop count, however the improvement vanishes after several hops. More specifically, optimal hop counts considering the total path delay are $h \approx 3$ for Infocom05, $h \approx 2$ for Cambridge, and $h \approx 2.6$ for Infocom06. As for (Q2), two hops are sufficient to reach every node from every other node.

As for (Q3), indirectly from our bottleneck weight analysis, we can conjecture that delivery ratio increases significantly if at least two hops are allowed. However, for all traces the performance increase tends to stabilise after $h \approx 4$.

Next, we set $T_{agg} = \{1, 6, 24\}$ hours to analyze how time-aggregation affects the network dynamics. We refer to these settings as short, medium, and long aggregation windows. Before presenting the average capacity and optimal hop count, let us see how optimal hop count changes over time for each trace. In Fig. 5(a), we plot the results for short $T_{agg}$. As Infocom05 trace is approximately 3 days long, we have 70 snapshots while Cambridge has 274, and Infocom06 has 93 snapshots. The figures show the hop counts providing the best bottleneck weights for $h = 3$ and $h = 8$ for both the time-aggregated graph and static graph. The reachable fraction of nodes is not presented here due to space restrictions, but our results show that static graph overestimates the connectivity of the network. All nodes can be reached in two hops if waited sufficiently long. However, the actual reachable fraction is much lower according to the results of our snapshot analysis in Fig. 6(a). This is not primarily because of the hop restriction but the inherent network dynamics. Albeit setting $h = 8$ leads to higher connectivity than that of $h = 3$, the connectivity is still drastically lower than the static graph. Figures also capture the difference in time: during daytime connectivity is higher compared to the night times. This change in the connectivity manifests itself as the change in optimal hop count over time. During some periods, connectivity is so low that although the
In reality, but increasing the bottleneck capacity is due to the added edges that may violate the time ordering.

Optimal hops are on the average higher for higher benefits diminish after. Similar to our previous results, the increasing hop count (nodes in the given time period). Regarding bottleneck capacity, dynamics of the network put a limit on the ratio of reachable nodes does not hold. Increasing hop count does not help as the two hops are sufficient to provide full connectivity among all nodes in the given time period. Regarding bottleneck capacity, our previous conclusion that of multi-hop routing. However, our previous AHOP analysis shows, reachable fraction of nodes in static graph is higher than what actual graph $G_t$ could provide. The overestimation is clear in Fig. 6(a): static and long $T_{agg}$ have the highest reachable fraction ($\approx 0.9 - 1$) whereas short $T_{agg}$ is significantly lower ($\approx 0.3$). Regarding the improvement with increasing $h$, all exhibit the same behaviour. In line with the results of Sec.III, we see that $h = 2$ provides most of the benefits of multi-hop routing. However, our previous conclusion that two hops are sufficient to provide full connectivity among all nodes does not hold. Increasing hop count does not help as the dynamics of the network put a limit on the ratio of reachable nodes in the given time period.

Fig. 5. Change in optimal hops over time.

Fig. 6. AHOP analysis for $T_{agg} = \{1, 6, 24\}$ hours. Infocom05 trace.

protocol lets 8 hops, single hop routing achieves the optimal performance (which is very low). This change in hop count may call for protocols that adapt the parameters according to the time of the operation. Regarding the amount of change, the standard deviation is 0.54 hops for $h = 2$ and 0.77 hops for $h = 8$ in Infocom05, while they are (0.45, 0.62) hops in Cambridge and (0.42, 0.83) hops in Infocom06.

In Fig. 7, we plot the simulation results: the delivery ratio, delivery delay, and average hop count of the delivered messages for Infocom05 and Infocom06 traces. In these scenarios, we set time-to-live (TTL) of a message to $ttl = \{1, 6, 24\}$ hours. A message exceeding its ttl is dropped. For the sake of comparison, we set buffer capacity and contact capacity large so that the performance is not restricted by these factors. From Fig. 7(a), we can see the improvement facilitated by increasing $h$. In all scenarios, we see two regions; in the first region the delivery ratio increases with increasing $h$ which later changes marginally in the second region. The turning point is $3 - 4$ hops. This behaviour agrees our previous AHOP analysis.

Unsurprisingly, delivery delay in Fig. 7(b) shows a similar trend to that of additive weights in Fig. 4(b). In this figure, the results are not normalized to give an idea about the delay of communication in such opportunistic networks. For $ttl = 1$ h, the delay slightly changes across different $h$, which indicates that the performance is primarily determined by the time restriction (and network mobility) rather than the hop count restriction. Lower average hop count in Fig. 7(c) corroborates this claim. For other settings, we observe a significant drop in delivery delay from $h = 1$ to $h = 2$. Another thing to note is that messages are routed in $h \approx 2.3$ hops in Infocom05 trace independent of ttl, whereas $h \approx 3.2$ for Infocom06.

Fig. 6 demonstrates the average results collected from multiple snapshots of the network. We present only results for Infocom05 as the rest behave similarly. As our simple example showed, reachable fraction of nodes in static graph is higher than what actual graph $G_t$ could provide. The overestimation is clear in Fig. 6(a): static and long $T_{agg}$ have the highest reachable fraction ($\approx 0.9 - 1$) whereas short $T_{agg}$ is significantly lower ($\approx 0.3$). Regarding the improvement with increasing $h$, all exhibit the same behaviour. In line with the results of Sec.III, we see that $h = 2$ provides most of the benefits of multi-hop routing. However, our previous conclusion that two hops are sufficient to provide full connectivity among all nodes does not hold. Increasing hop count does not help as the dynamics of the network put a limit on the ratio of reachable nodes in the given time period. Regarding bottleneck capacity, increasing hop count ($h > 2$) leads to higher bottleneck capacity for all settings. Similar to our previous results, the benefits diminish after $h \approx 4$. As Fig. 6(c) illustrates, the optimal hops are on the average higher for higher $T_{agg}$. This is due to the added edges that may violate the time ordering in reality but increasing the bottleneck capacity.

VI. CONCLUSIONS

In this paper, we have explored the effect of hop count restriction on opportunistic routing schemes. We have used
the modified Bellman-Ford’s algorithm which is proposed for all hops optimal path problem (AHOP) and is a good fit for our research question. We analyzed human contact traces using three approaches: first we derived the contact graph from the traces and computed the AHOP on the static graph. This approach is optimistic as it disregards the time ordering of the links. To decrease the deviation from the actual graph, we observed the network on multiple time points and aggregated all events occurring in a time interval into a network snapshot. Next, we derived the performance by averaging related metrics over the snapshots. Finally, we simulated a hop-limited routing protocol. Although characteristics (e.g., connectivity) of the network snapshots may be different from the static graph, they are similarly affected by the hop count limitation. Connectivity of the network as well as the performance (e.g., lower delay) improves with increasing hop count. Generally speaking, while two hops are sufficient to achieve the highest connectivity the network can provide given the time restrictions, routing protocols can perform better by increasing hop count to $h \approx 4$. After this point, the performance tends to stabilize.

In this work, we studied small-world networks that exhibit some community structure which reflects this property as low number of hops achieving good connectivity for the static graph. As future work, we plan to provide a formal framework which can guide us better in understanding the effect of hop count for general opportunistic networks.

REFERENCES


