

## Research Article

# Evaluation of Congestion Aware Social Metrics for Centrality-Based Routing

Muhammad Arshad Islam <sup>1</sup>, Muhammad Azhar Iqbal,<sup>2</sup> Muhammad Aleem <sup>1</sup>,  
Zahid Halim,<sup>3</sup> Gautam Srivastava,<sup>4,5</sup> and Jerry Chun-Wei Lin <sup>6</sup>

<sup>1</sup>National University of Computer and Emerging Sciences-FAST, Islamabad, Pakistan

<sup>2</sup>School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu, China

<sup>3</sup>Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Swabi, Pakistan

<sup>4</sup>Department of Math and Computer Science, Brandon University, Canada

<sup>5</sup>Research Centre for Interneural Computing, China Medical University, Taiwan

<sup>6</sup>Western Norway University of Applied Sciences, Norway

Correspondence should be addressed to Jerry Chun-Wei Lin; [jerrylin@ieee.org](mailto:jerrylin@ieee.org)

Received 13 February 2021; Accepted 4 June 2021; Published 21 June 2021

Academic Editor: Nawab Muhammad Faseeh Qureshi

Copyright © 2021 Muhammad Arshad Islam et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Opportunistic networks utilize pocket switching for routing where each node forwards its messages to a suitable next node. The selection of the forwarder node is crucial for the efficient performance of a routing protocol. In any opportunistic network, some nodes have a paramount role in the routing process and these nodes could be identified with the assistance of the existing centrality measures available in network theory. However, the central nodes tend to suffer from congestion because a large number of nodes that are relatively less central attempt to forward their payload to the central nodes to increase the probability of the message delivery. This paper evaluates mechanisms to transform the social encounters into congestion aware metrics so that high-ranking central nodes are downgraded when they encounter congestion. The network transformations are aimed at aggregating the connectivity patterns of the nodes to implicitly accumulate the network information to be utilized by centrality measures for routing purposes. We have analyzed the performance of the metrics' computed centrality measures using routing simulation on three real-world network traces. The results revealed that betweenness centrality along with the congestion aware network metrics holds the potential to deliver a competitive number of messages. Additionally, the proposed congestion aware metrics significantly balance the routing load among the central nodes.

## 1. Introduction

Communication in opportunistic networks is dependent on the mobility patterns of the wireless nodes that constitute the network. An opportunity to exchange data among the network is created whenever two nodes come across the communication range of each other. The wireless nodes exhibit a store-carry-forward mechanism wherein they hold the characteristics of both data-mule and router [1]. This intermittent nature of the opportunistic network causes the delivery time of a message to vary from a few seconds to several days depending on the node characteristics (mobility model), network characteristics (network density and network diame-

ter), and message characteristics (message size and the distance between source and destination) [2].

An efficient opportunistic network routing protocol accurately predicts connectivity pattern using available node-node link characteristics such as contact count, contact duration intercontact time (time elapsed between two consecutive contacts) [1], frequency, and duration of the contacts that are created as a result of the social interaction of the users of the nodes. Due to this versatile nature of opportunistic networks, extracting an accurate routing metric for different kinds of opportunistic networks becomes a strenuous process. However, in any network, a few nodes are considered central nodes (also known as Hub) since they are

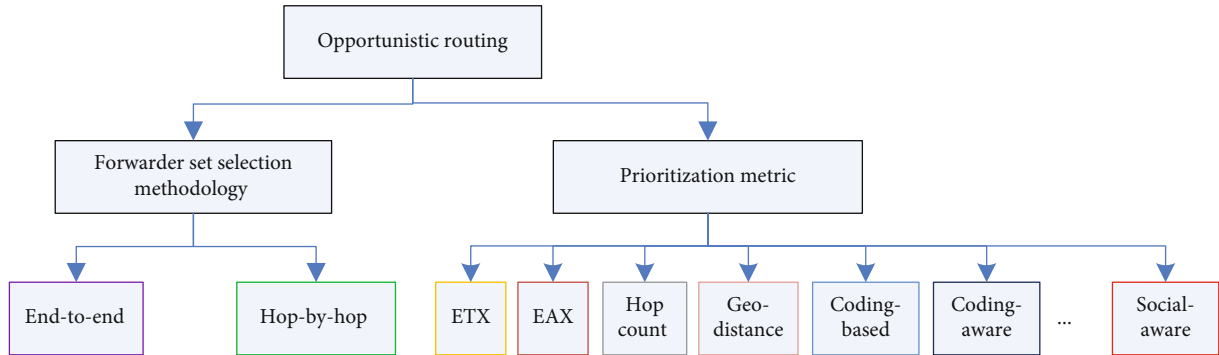


FIGURE 1: Components of opportunistic routing.

playing a pivotal role in routing, e.g., these central nodes contribute more in the process of message forwarding than the other nodes [3].

The fundamental notion of opportunistic routing consists of two main components, i.e., forwarding methodology and prioritization as shown in Figure 1. Forwarding methodology can be based on hop-by-hop selection of forwarder/central nodes or end-to-end selection of forwarder node set. On the other hand, the prioritization is based on a metric (i.e., node sociality, Expected Transmission Count (ETX), geo-distance, hop count, etc.) dependent on the particular nature of a wireless network or wireless network application and it ultimately suppresses undue packet forwarding.

Network centrality is considered a vital tool for network analysis to identify because such central nodes can play a key role in disseminating information in the network. Most of the centrality measures are defined for static networks (contact information among nodes does not change with time). This work focuses on evaluating the routing performance of the network metrics that are adapted to implicitly harness the congestion-related information of the links between mobile nodes using centrality measures. The congestion aware nature of the investigated metrics manipulates the centrality ranking in such a way that a high-rank mobile node will be ranked lower than its counterparts if it carries higher traffic volume.

The remainder of the paper is organized as follows: Section 2 presents existing state-of-the-art works that use centrality measures for routing. Section 3 presents the metrics that are used to transform the opportunistic networks. Details of the network traces along with the simulation setup parameters are discussed in Section 4. Section 5 presents the results and discusses the performance of simulated network metrics. Conclusion and future work are presented in Section 6.

## 2. Related Work

The process of central node identification has been utilized in multiple works to make routing decisions in different kinds of communication networks. Literatures acknowledge the synergy among social network tools like centrality measures and ad hoc networks as a fertile research area that has the potential for designing network routing protocols for oppor-

tunistic networks [4]. Researchers have the choice among various centrality measures available in network theory [5]. The centrality of a node in a network can be determined by using various centrality measures such as degree centrality [6], betweenness centrality [5], eigenvector centrality [7], page-rank [8], hub centrality [9], and centrality [10]. The computation process of most of the centrality measures is centralized and required complete network information. However, in the context of opportunistic networks, it is not practical for a node to have the complete information of the network and to calculate various centrality measures ranking a node concerning its suitability for routing [10]. Moreover, due to the dynamic nature of opportunistic networks, the contemporary centrality measure cannot be helpful to calculate node centrality [11].

The existing centrality measures have been applied to address multiple issues of ad hoc network such as routing [12], congestion [13, 14], and energy conservation [1]. Wang et al. proposed a scheme for superior forwarding node selection that is based on the concept of value strength that relies on social network structure [15].

The authors simulated the protocol on real-world traces extracted from Flickr to show high information converge ratio. Zhu et al. [16] have presented a routing protocol “ZOOM” for opportunistic forwarding in vehicular networks that uses network centrality metrics in the absence of primary routing information, i.e., intercontact time. The concept of centrality-based community is employed for opportunistic routing where a message is pushed towards a central community with the understanding that nodes in such a community have a high probability to get connected with the destination of the message [17]. It has also been argued that nodes with a high value of betweenness centrality are susceptible to face traffic congestion as well as energy depletion because of their high probability of participation in the routing process [18]. Miralda et al. [19] have employed a variant of betweenness centrality based on fuzzy logic with the aim of energy conservation that is crucial in Io nodes in opportunistic networks. They argue that nodes with high local betweenness centrality are prone to consume more energy and face buffer occupancy problems during the routing process, and consequently, distributing the routing process with low local betweenness centrality can help in conserving the energy.

The effectiveness of closeness and betweenness centrality to identify the influential nodes for routing in opportunistic networks is demonstrated in [20]. Sivalingam and Chellappan [21] have used the concept of entropy for routing purposes in tactical wireless networks. Entropy is defined in terms of the downstream degree of a vertex, which is the number of eligible vertices for forwarding. The investigation also includes the correlation among centralities as a function of network connectivity and network mobility showing that the closeness centrality (relevant to the shortest path) has obtained a higher correlation with degree-based centrality measure as compared to betweenness centrality. The research community agrees that network-theoretic concepts can be useful for the identification of influential nodes for routing in infrastructure-less environments. However, this brings about the challenge of how centrality measures for static networks can be transformed for making accurate routing decisions in opportunistic networks using metrics that preserve the link characteristics between the nodes for congestion handling. The list of link characteristics includes node connectivity count, link error rate, link duration, and hidden node problems. The focus of this work is to investigate the congestion aware metrics using node contact patterns to enhance the routing performance in the opportunistic network.

*Gap analysis:* we have discussed several recent studies that have focused on improving routing performance and reducing energy consumption using centrality metrics. The focus of these studies is mostly computing centrality measures with minimum network overhead that requires global network knowledge. However, none of the recent works have investigated the implicit congestion avoidance capabilities of the network metrics that can be sensed and shared among the network devices with little or no overhead. This paper investigates three such network metrics by integrating them with multiple centrality measures to gauge their congestion awareness.

### 3. Opportunistic Network Metrics

The metrics are aimed at facilitating the computation of centrality measures while maintaining the temporal characteristics of the contact among the network nodes. The aforementioned centrality measures have been described concerning the static networks. However, opportunistic networks are dynamic where nodes may join and leave the network. We can transform the dynamic link behavior of the network nodes using the following metrics to analyze their performance for opportunistic network routing. We have divided the presented metrics into two classes, i.e., congestion oblivious and congestion aware.

*3.1. Congestion Oblivious Metric.* Most of the existing literature relies on the congestion oblivious metrics that are not affected by the current traffic load of the routing nodes.

*Aggregate network:* it is the simplest metric for the centrality computation for dynamic networks. An opportunistic network can be seen as a sequence of static graphs at a particular point in time as shown in Figure 2(a). A simple network

is created aggregating all edges that existed at any point in time [22] as shown in Figure 2(b). An aggregated graph for a dynamic network can be represented as an  $n \times n$  adjacency matrix  $A$ , where each element  $a_{ij}$  is defined as follows:

$$\alpha_{ij} = \begin{cases} 1 & \text{nodes } i \& j \text{ made a contact,} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

The primary advantage of this method is simplicity. However, aggregating all edges may result in losing temporal information that is vital for routing decisions in opportunistic networks. A connection between any two nodes that once came in contact with each other will be represented as a permanent connection.

*Contact count:* it represents a weighted network with edges representing the number of contacts that occurred between two nodes. This metric will relate two frequently contacting (two nodes contact each other when they are in their wireless range) nodes stronger than those that have contacted each other less frequently. Contact count graph can be represented as a weighted adjacency matrix  $\text{ContCnt}$  where each element  $\eta_{ij}$  is defined as the total contact count between nodes  $i$  and  $j$ . This metric allocates weight to contacts that occur frequently; however, it does not favor contacts occurring at regular intervals. Two nodes that connect infrequently but on regular basis (daily) over a longer span will be given less priority as compared to the nodes that get connected very frequently over a shorter period.

*3.2. Congestion Aware Metrics.* Duration-based metrics have not been investigated to handle congestion. The focus of this work is to evaluate the congestion aware metrics that affected implicitly the traffic load of a node, and they can be used to lower the routing suitability of the nodes that are facing higher routing load [23].

*Average contact duration:* it represents a weighted network with edges representing the average duration of the contacts between any two nodes. The longer the average contact duration between two nodes, the stronger is the relationship between them and vice versa. The contract duration of two nodes will be reduced if they participate in large volume transmissions. Thus, their centrality rank will be lowered due to congestion. The average duration network can be represented as a weighted adjacency matrix  $\text{Dur}$  where each element  $\lambda_{ij}$  is defined as follows:

$$\lambda_{ij} = \frac{\sum_{k=1}^{\eta_{ij}} \text{contactduration}_k}{\eta_{ij}}. \quad (2)$$

*Intercontact time:* it represents a weighted network with edges representing the mean time elapsed between two consecutive contacts of two particular nodes. The smaller the intercontact time between two nodes, the stronger is the relationship between them and vice versa. This is a duration-based metric that will affect the relationship between two nodes if they encounter network congestion. The duration of the contact time of the nodes will be reduced, and

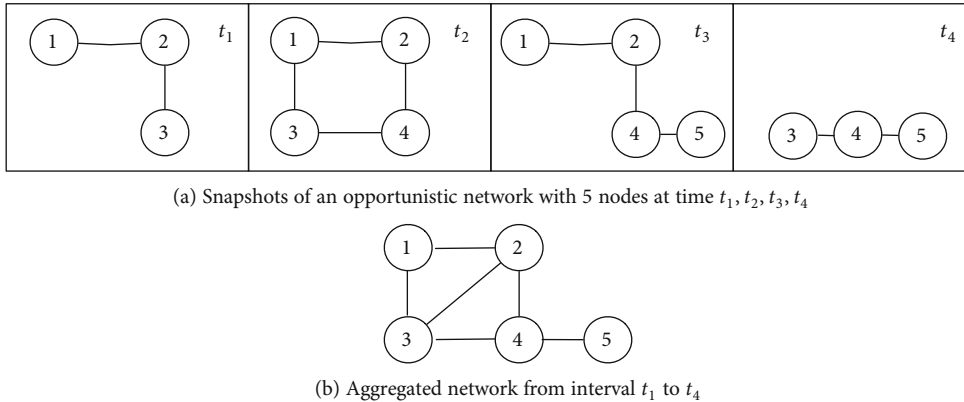


FIGURE 2: Opportunistic network representation as an aggregated network.

correspondingly, intercontact time will increase. Intercontact time network of nodes  $i$  and  $j$  with  $\eta_{ij}$  contact count can be represented as weighted adjacency matrix  $ICDur$  where each element  $\mu_{ij}$  is defined as follows:

$$\mu_{ij} = \frac{\sum_{k=1}^{\eta_{ij}} \text{intercontactDur}_k}{\eta_{ij}}. \quad (3)$$

The metric considers the mean of all the intercontact time duration elapsed among all contacts of two nodes. The behavior of this metric is somewhat similar to the contact count. This metric will downgrade the relationship between two nodes that either stop contacting each other or start to incur longer delays.

All of the metrics except aggregate graph are dynamic in nature, and two metrics, i.e., average contact duration and intercontact time are sensitive to congestion faced by the involved nodes. Considering dynamic metric, the node ranking is not static and the ranking of a central node bounds to degrade with passing time when the node encounters congestion while congestion sensitive metrics are employed.

#### 4. Experimental Setup

We have considered three different kinds of datasets, namely, MIT cell tower, MIT Bluetooth, and IBM access points, and all of these have been obtained from Community Resource for Archiving Wireless Data at Dartmouth (<http://www.crawdad.org/>). The selected dataset promises to represent the real-life mobility pattern of users while grabbing the basic social contact behavior. The motivation behind choosing these three traces is to analyze the range of spectrum between dense and sparse networks. Two of the data traces have been synthesized from reality mining project [24] from MIT spans 19 months, i.e., February 2004 to August 2005, whereas the third data trace consists of the SNMP logs for one month from an IBM campus [25]. Since the contact duration of MIT reality mining is longer than IBM trace, we have filtered the MIT data to match the time span of IBM traces.

The sparse network is extracted from Bluetooth logs of MIT traces (MITBT) where each node scans for active Bluetooth neighbors in the interval of every five minutes and

stores the duration of contact times. For the sake of comparison with other traces and simplicity, we have limited our experiments to one month of connectivity trace, where any visible Bluetooth device was considered a candidate connection. Reduction of the trace time span has been considered on the basis of connectivity times, i.e., one month, where nodes have maximum connectivity in terms of time duration. The highest connectivity period, i.e., November 2004, showed 1858 Bluetooth nodes suggesting a huge number of undesignated nodes as compared to the designated 81 nodes that were designated to gather the data. It is noteworthy that a few undesignated devices had more connectivity and interaction with the network than the designated nodes.

In the case of IBM Access Point trace, Simple Network Management Protocol (SNMP) is used to poll access points (AP) every 5 minutes, from July 20, 2002, through August 17, 2002 [25]. The total of 1366 devices has been polled over 172 different access points during approximately 4 weeks. We have extracted the traces of 928 devices after discovering the existence of 3 clusters in this network and then choosing the biggest cluster with respect to node count. The biggest cluster has been identified by analyzing the connectivity pattern among devices. The 3 extracted clusters represent the devices belonging to 3 buildings, and the biggest cluster is considered for the simulations. Since the authors of the dataset [25] have polled the access point for connected devices every 5 minutes, we assume that the snapshot data remains constant for the next 5 minutes to turn these samples into continuous data. In the rare cases where this would cause an overlap with another snapshot from another access point, we assume that the transition happens halfway between the two snapshots. It is also assumed that two nodes that are connected to one access point during the overlapping time period are connected to each other. Thus, key features of such network are low mobility and medium transmission range.

The second trace, the MIT cell tower, is used according to the similar principle as that of the IBM trace. The only difference is that instead of functioning as access points, cell towers are used to gather the contact times of the nodes with each other; thus, the resulting network can be characterized as a very dense network due to the high range of the cell tower. The MIT cell tower provides continuous data; therefore, it consists of a large number of contacts with small duration

(less than 10 seconds). Due to several lapses in data gathering, mentioned by the creators of the data, only 89 of 100 devices have been included that visit 32768 different cell towers. Similar to Bluetooth traces, Nov 2004 turned out to be the maximum activity month with 81 devices.

It is imperative to mention that the assumption that two devices connected to one base station (access point or cell tower) introduces inaccuracies [26]. On one hand, it is overly optimistic, since two devices attached to the same access point may still be out of range of each other. On the other hand, the data might omit connection opportunities, since two nodes may pass each other at a place where there is no base station, and this contact would not be logged. Another issue with these datasets is that the devices are not necessarily colocated with their owner at all times (i.e., they do not always characterize human mobility). Despite these inaccuracies, such traces are a valuable source of data, since they span many months and include thousands of devices. Additionally, the datasets used in this study promise to represent the real-life mobility patterns and social networking behaviors of users because the traces are extracted using the mobile devices [27]. Authors in [28] analyzed IBM traces and extracted the usage (session duration and traffic volume) and mobility patterns (number of associated users) of WLAN users. Bhaumik and Batabyal utilized graph tools including centrality and clustering coefficient to propose message dissemination protocols for delay tolerant networks using MIT traces [29]. Further details of the routing simulation mechanism are available in [2].

*Centrality computation:* as stated earlier, the first two weeks are used for bootstrapping the centrality measures. This process is continued in the latter part of the simulation, and the centrality of each device is updated at a 10-minute interval. This interval is extended to 360 minutes when the devices report their last activity for one day and the next activity occurs on the next morning. The results of the computations are assumed to be transmitted to all devices instantaneously.

*Link sharing:* each device can maintain the communication session with one other device at any point in time. There are enough independent channels available that any number of node pairs can communicate at the same time with full bandwidth, independent of their proximity to another pair. This aspect will play a key role in analyzing the effect of traffic congestion on devices when *average contact duration* and *intercontact time* are used for centrality measure computation.

*Congestion awareness:* wireless devices rely on nonsharing channel allocation for data transfer. This aspect is focused to adapt congestion aware metrics based on the amount of data being forwarded through a device. Whenever two devices come into transmission range, messages may be exchanged depending on the current ranking of the devices obtained using one of the centrality measures. When two devices start exchanging messages, then these devices will be invisible to the rest of the surrounding devices. Thus, the ranking obtained using the duration dependent metrics, i.e., *average contact duration* and *intercontact time*, will deteriorate for the devices that attempt to transmit large data.

TABLE 1: Simulation parameters.

Message count	100
Message size	1.6E3...1.6E7 B
Message lifetime	7 days
Centrality computation interval	10-360 min
Message size distribution	Power law
Bandwidth (low)	100 kB/s
Bandwidth (high)	10,000 kB/s

The peripheral simulation parameters are summarized in Table 1. 100 messages of varying sizes ranging from 1600 B to 1.6E7 B are simulated. The size distribution followed a power law, i.e., a few messages having a large size and many small size messages. We have performed experiments with three centrality measures using the transformed networks. A centralized version of centrality metrics is considered for the sake of comparison. One may consider that the accuracy of metrics will decrease when egocentric variants of the centrality measures based on local information will be available to the individual nodes. A summary of the metrics used with various kinds of centrality is presented in Table 2.

The simulated protocol follows a hop-based routing where every node forwards a message replica to the next node if the centrality measure value of the next node is higher than the current node. In other words, if the receiving node is relatively central to the current node then the current node forwards a replica of one message to another node. The message is replicated during this process and is delivered to the destination if any of the nodes currently in possession of the message replica make a contact with the destination. If the source node of the message has a lower centrality, then the message will be replicated more as compared to the message whose source has a higher centrality in the network.

Considering the example presented in Figure 3, two mobile devices D1 and D2 are shown within transmission range before the exchange of messages. Size and destination of each message are shown along each message. It is assumed that the centrality measure of D2 is higher than that of D1. When the transmission phase is over due to the termination of the contact, we can see that two messages destined for each device have been transmitted first followed by the messages that are required to be forwarded. D2 has one more message labelled with blue color that is destined for device D2. However, device D2 that is assumed to have higher centrality value receives two new messages from device D1. One of the messages is destined for device D2, and the other is replicated for forwarding purpose. Two messages destined for D5 and D7 have not been exchanged due to the termination of the contact.

## 5. Results and Discussions

To establish the correlation between centrality measures and routing importance of nodes, we have simulated epidemic routing to obtain the routing importance of network nodes and then analyzed its relation with the values obtained using centrality measures. A node may have participated in the

TABLE 2: Transformed metrics and centrality measure combinations.

Transformed metric	Betweenness (B)	Closeness (C)	Degree (D)
Aggregated network (Agr)	X	X	X
Contact count (ContCnt)	X	X	
Contact duration (Dur)	X		
Intercontact duration (ICDur)	X	X	

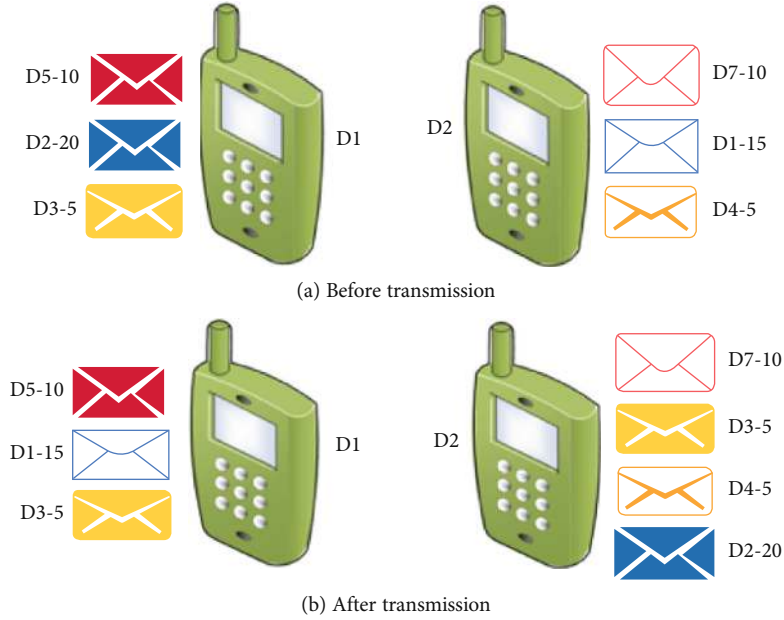


FIGURE 3: Example showing the working of the message forwarding protocol.

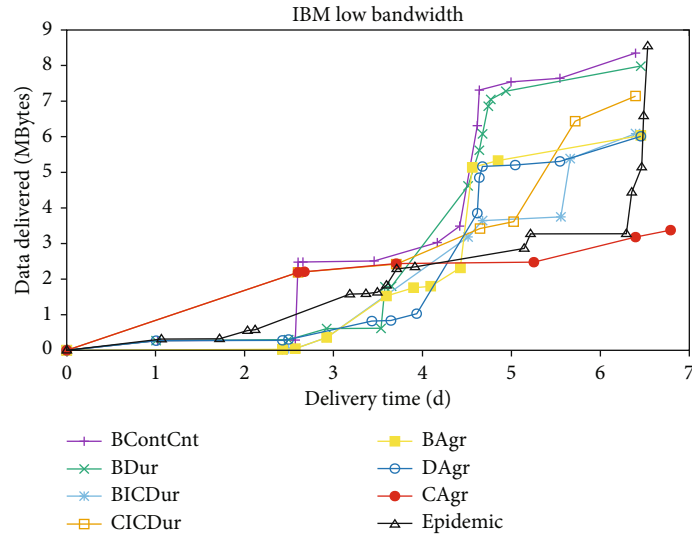
	BContCnt	BDur	BICDur	CContCnt	CICDur	BAgr	DAgr	CAgr	MTCnt
BContCnt	1								
BDur	0.45	1							
BICDur	0.29	0.56	1						
CContCnt	-0.3	-0.42	-0.4	1					
CICDur	-0.32	-0.48	-0.47	0.96	1				
BAgr	0.43	0.72	0.58	-0.47	-0.49	1			
DAgr	0.51	0.74	0.6	-0.74	-0.76	0.81	1		
CAgr	-0.3	-0.41	-0.39	0.99	0.96	-0.46	-0.73	1	
MTCnt	0.49	0.51	-0.42	-0.51	-0.53	-0.61	0.72	-0.51	1

(a) Correlation grid for IBM trace

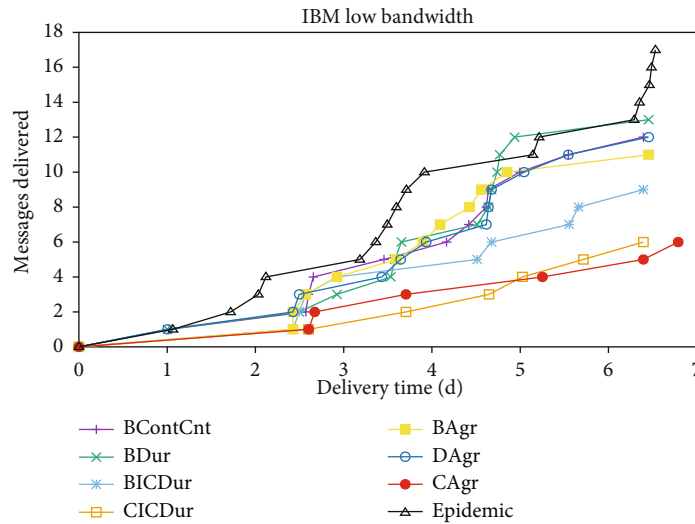
	BContCnt	BDur	BICDur	CContCnt	CICDur	BAgr	DAgr	CAgr	MTCnt
BContCnt	1								
BDur	0.68	1							
BICDur	0.45	0.63	1						
CContCnt	-0.24	-0.27	-0.26	1					
CICDur	-0.46	-0.49	-0.47	0.67	1				
BAgr	0.75	0.77	0.66	-0.26	-0.47	1			
DAgr	0.7	0.78	0.67	-0.51	-0.72	0.75	1		
CAgr	-0.24	-0.27	-0.26	0.99	0.67	-0.26	-0.51	1	
MTCnt	0.69	0.69	-0.66	-0.27	-0.48	-0.78	0.74	-0.26	1

(b) Correlation grid for MITBT trace

FIGURE 4: Correlation grid for three traces.



(a) Amount of data delivered



(b) Number of messages delivered

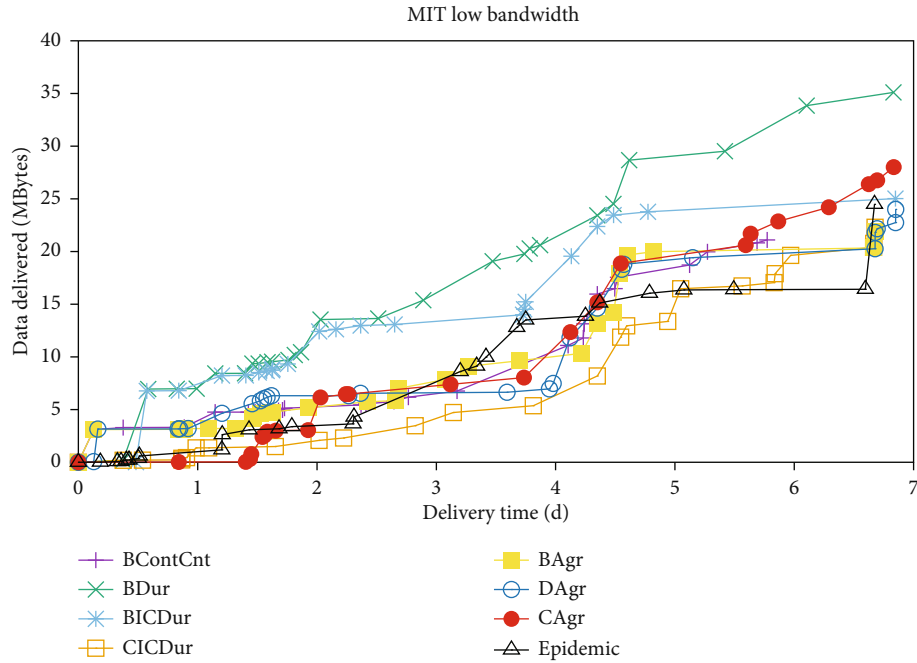
FIGURE 5: Routing performance of centrality metric IBM network.

transmission of multiple messages, and we can associate the number of messages, and a node has transmitted to the next-hop with its routing importance. A node that is centrally positioned in the network is expected to participate in the forwarding of a larger number of messages as compared to other nodes that do not have the central position. Correlation is represented among centralities with varying transformed metrics in the form of a grid in Figure 4. The green color represents the positive correlation, and the yellow color shows the negative correlation. The first letter of each label in Figure 4 represents the type of centrality followed by the metric used for centrality computation, i.e., BContCnt represents contact count betweenness. The scheme of abbreviations used in Figure 4 is described in Table 2.

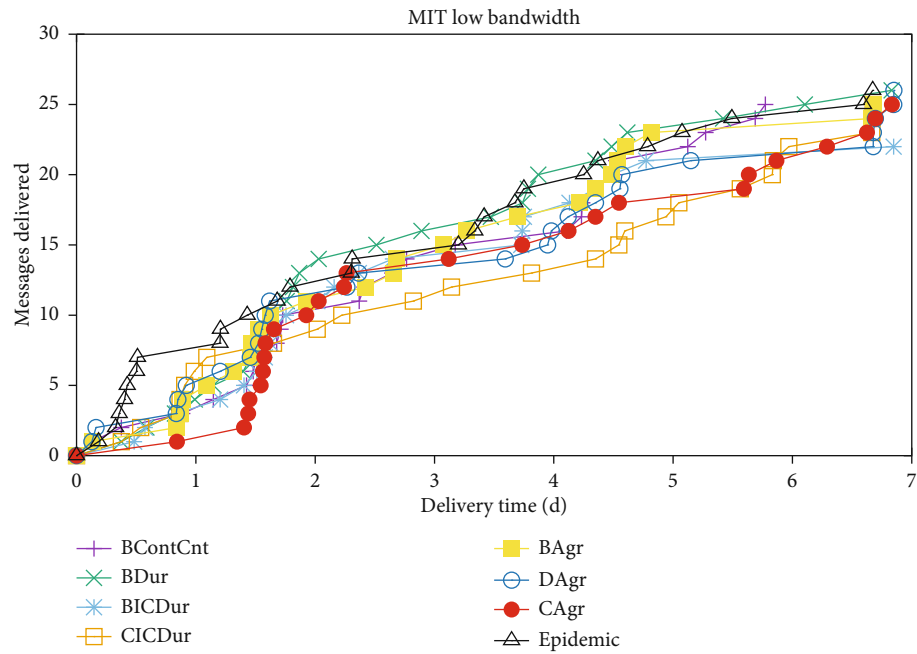
Correlation among investigated centralities is higher in IBM and MITBT traces as compared to the MIT trace. The reason is that MIT consists of a large number of very small

duration contacts. As the MIT trace is gathered with the help of cell towers and in many cases, the connection between nodes and cell towers breaks frequently particularly for those nodes that have to select among multiple cell towers due to their location in the overlap area of these cell towers. These aspects result in a high contact count of the MIT trace very high without significantly affecting the duration features of the contacts.

Another aspect observed in Figure 4 is that the same pair of centrality measures for all metrics generally shows a higher correlation with the exception of the MIT trace. Degree centrality shows a consistently negative correlation to the closeness centrality in all the traces. Closeness centrality with respect to intercontact time and contact count shows a negative correlation to messages transmitted. Betweenness centrality with respect to contact count and contact duration shows a positive correlation to the messages transmitted in all traces.



(a) Amount of data delivered



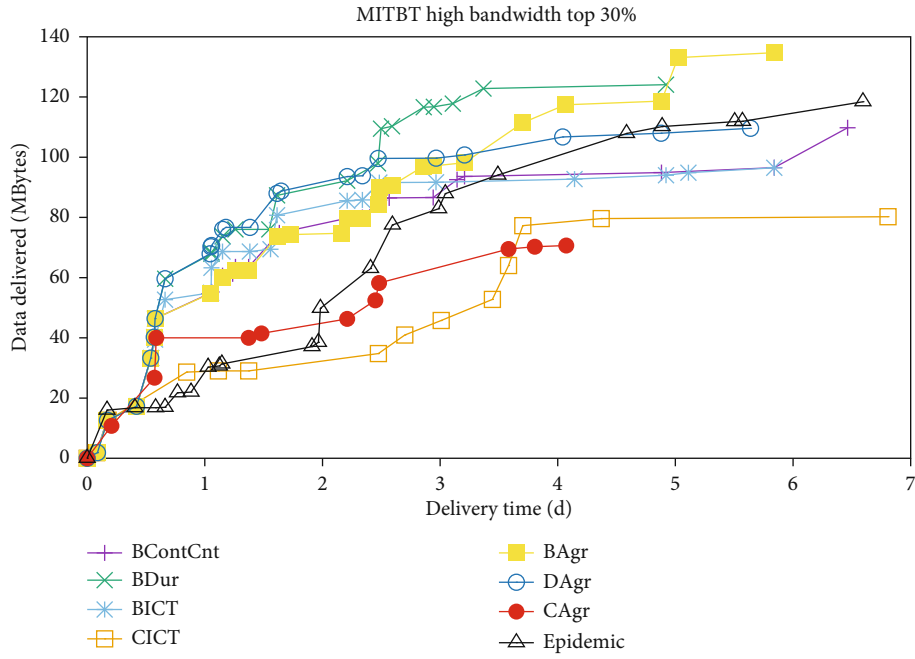
(b) Number of messages delivered

FIGURE 6: Routing performance of centrality metric MIT network.

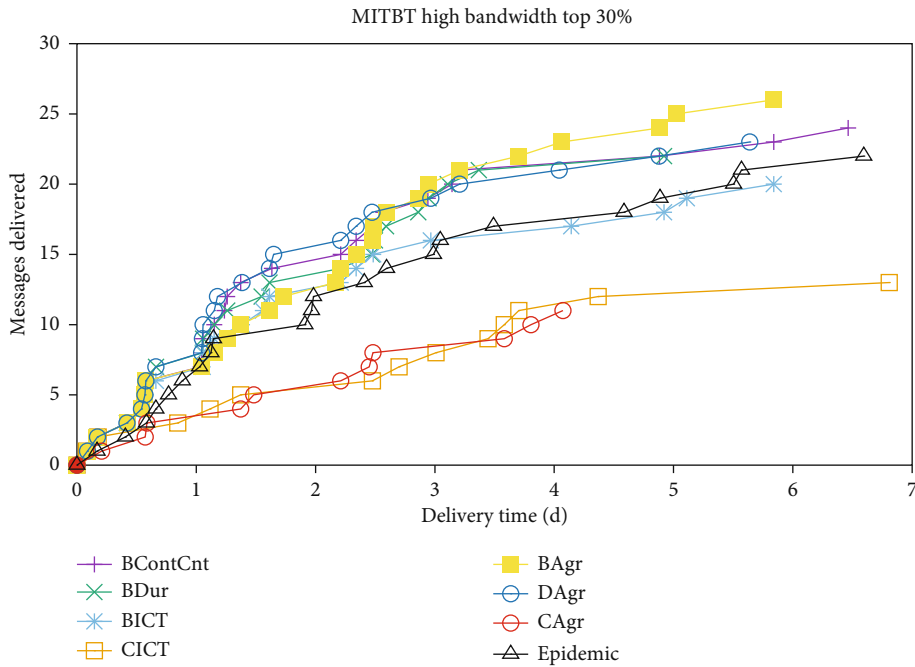
The correlation analysis has been used to reduce the scope of the experiments. The simulations are conducted using those combination of centrality measures and metrics that have low correlation with other combinations. From the above discussion, we conclude that the contact count and aggregate network betweenness can be considered as reliable routing metrics for opportunistic network routing for all three traces. Moreover, the correlation shown in Figure 4 is based on the results for the whole trace period, i.e., 1 month.

The results discussed in this section represent several messages and the amount of data delivered during the allocated 7 days of a time span to each message. Each pair of plots consists of the amount of data delivered (a) and messages (b) in each pair of Figures 5–7. For the sake of comparison, we have included epidemic protocol [13] where nodes try to replicate all the messages to the nodes that come in contact with it. An interesting aspect observed in the results of all traces (Figures 5–7) is that majority of centrality metrics have delivered somewhat similar performance to





(a) Amount of data delivered

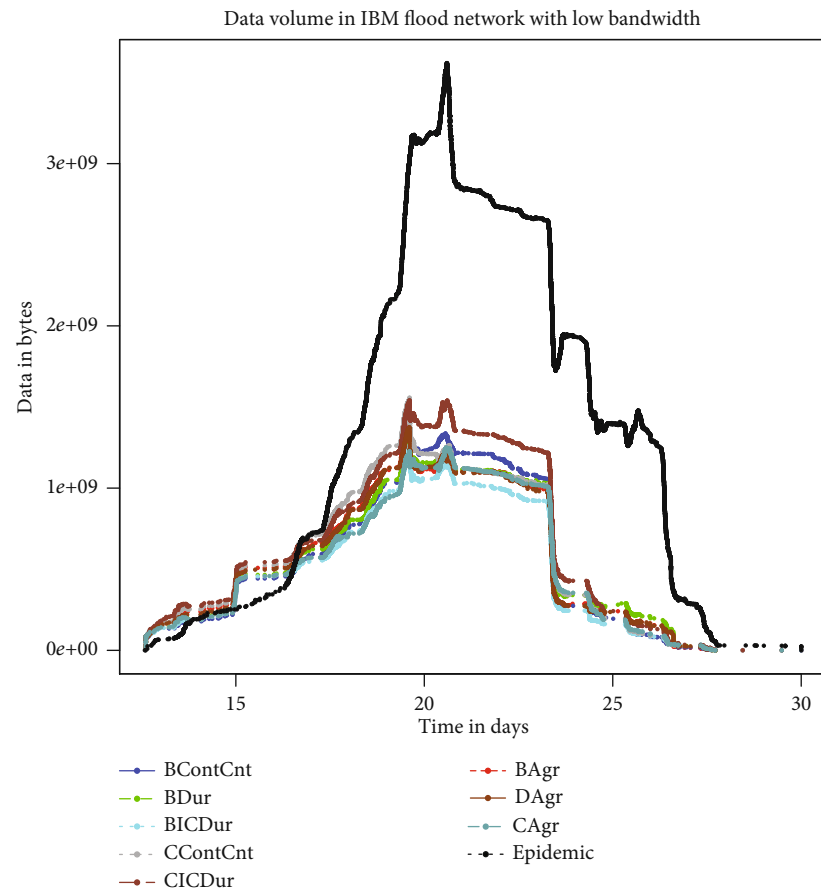


(b) Number of messages delivered

FIGURE 7: Routing performance of centrality metric MIT network.

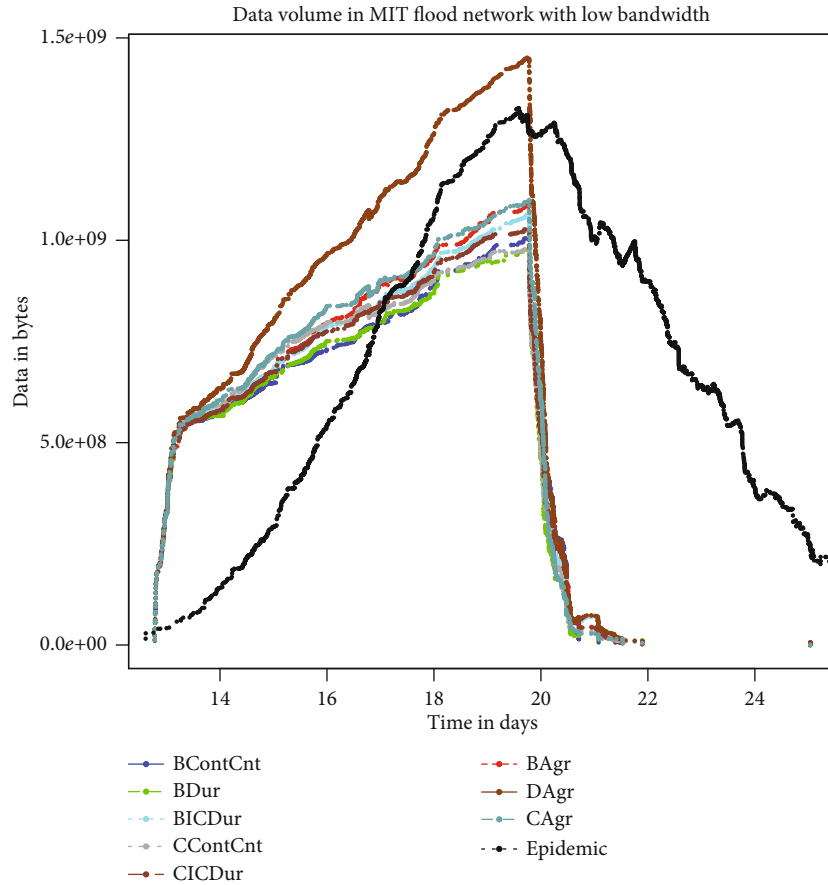
epidemic routing. It is imperative to mention that epidemic routing has not been able to attain the best performance in low bandwidth scenarios because devices are not able to forward and successfully replicate messages to other devices due to traffic congestion. A device is not able to forward a message unless it has received a complete replica of the message. The variation in the performance behavior of the centrality metrics in the three traces is due to the variation in contact patterns of the three traces. In the case of MIT trace (small duration frequent contacts) low bandwidth, several of

the centrality metrics have performed better than epidemic routing because of the overhead suffered as shown in Figure 6. Small contact durations made epidemic routing performance partially vulnerable as nodes have failed to replicate their messages despite consuming scarce bandwidth, and duration betweenness and contact count betweenness are among the metrics that have delivered the maximum amount of the bytes whereas degree closeness and intercontact time closeness are among the metrics with the low performance.



(a) Data volume for IBM trace

FIGURE 8: Continued.



(b) Data volume for MIT trace

FIGURE 8: Network memory utilization for IBM and MIT with low bandwidth.

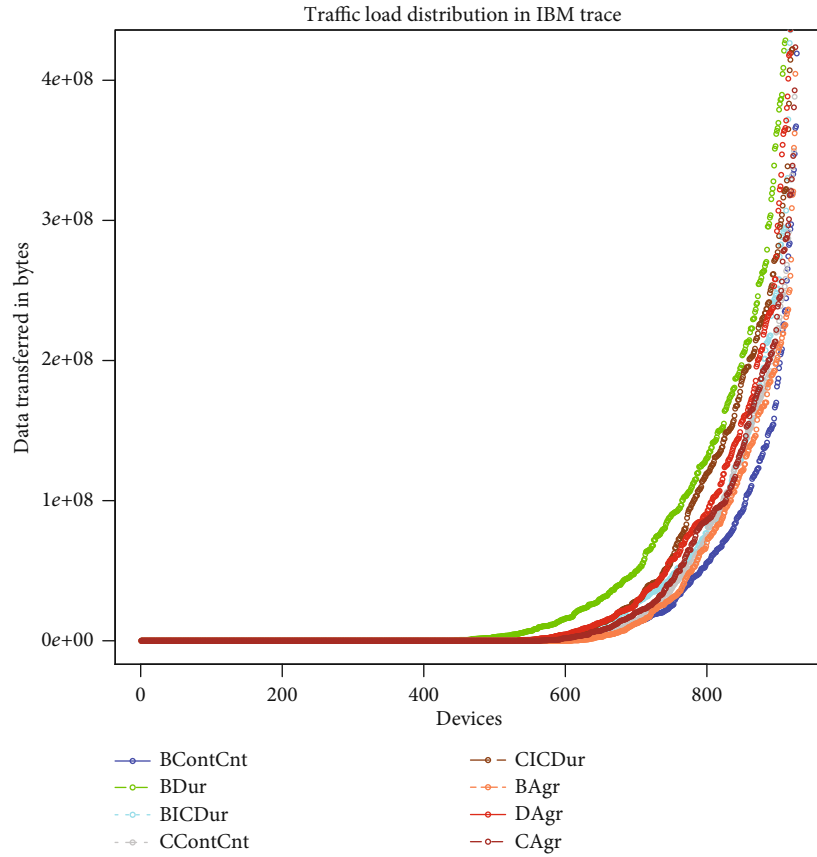
Epidemic routing delivered the maximum number of messages for IBM trace that comprises of contacts that occur with relatively low frequency as shown in Figure 5. It is followed by degree betweenness, contact count betweenness, and contact duration betweenness. In the case of bytes delivered, the list of top performers includes contact count betweenness and contact duration betweenness that have delivered approximately the same number of bytes as epidemic routing, however, in a shorter period. The delivery ratio is the minimum for MITBT trace because it is the sparsest dataset (very low frequency of contacts) among the three that have been utilized for experimentation as shown in Figure 7. The maximum amount of data is delivered by epidemic routing; however, degree betweenness and contact count betweenness have outperformed epidemic routing in several cases. The results are somewhat consistent with IBM and MIT as betweenness-based forwarding mechanism is among the best performance for MITBT as well.

Figure 8 shows the amount of memory consumed in the network, i.e., the consumed network storage increases as the nodes replicate the messages in their possession and transmit to the other nodes by increasing the storage utilization. Once the message is delivered or the lifetime of the message expires, the nodes remove the replica of the message by releasing the local storage. We have assumed the unlimited amount of local storage for each device so that protocols

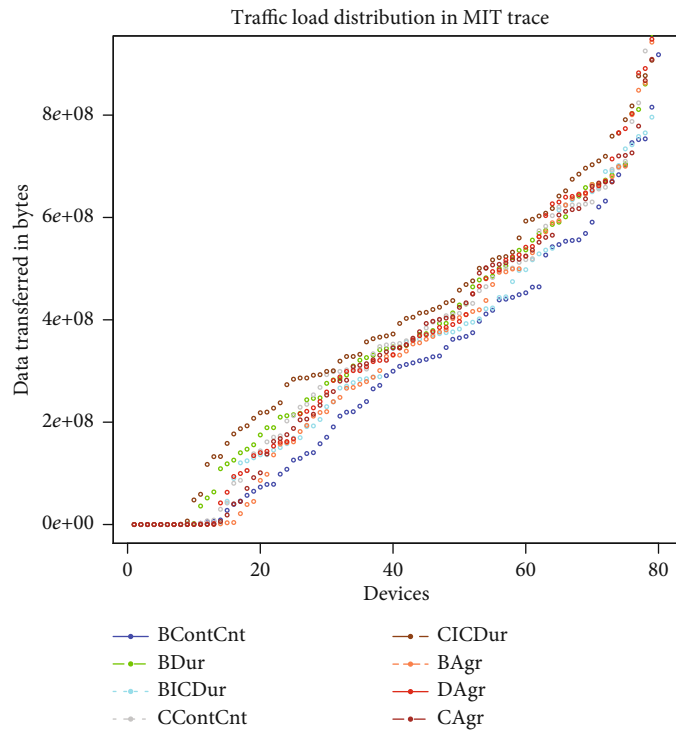
may exploit maximum storage to show maximum performance potential. Taking a closer look at the overheads involved during the centrality metric-based routing protocols as shown in Figure 6, it is noteworthy that protocols using congestion sensitive metrics with betweenness centrality have delivered the competitive number of messages. Also, the respective traffic volume is lower as compared to epidemic routing that shows the utilization of fewer resources.

**5.1. Congestion Awareness.** When we analyze the routing performance results of Figures 5–7 from the viewpoint of congestion management, the IBM trace shows that contact count betweenness is among the top three centrality forwarding strategies concerning being message count and data volume. As discussed earlier, IBM trace has of relatively lower contact frequency with long durations. The contact-count-based centrality measures have a relatively strong correlation with their aggregate-based counterparts. Aggregate-based centrality measures reported a strong correlation with the message transmission count of each device showing that these measures are prone to be affected by congestion.

Figure 9 shows the load shared by each device during the routing process of all the metrics in IBM and MIT traces. The  $x$ -axis represents the devices in sorted order concerning bytes transferred, and  $y$ -axis represents the corresponding number of bytes. Both traces (IBM Figure 9(a) and MIT Figure 9(b))



(a) The traffic load on devices in IBM trace



(b) The traffic load on devices in MIT trace

FIGURE 9: Bytes transferred through individual devices with low bandwidth.

show that average duration and intercontact time have utilized more devices in a relatively balanced way as compared to aggregate and contact count metrics. BDur in IBM trace used more than 450 devices; however, BContCnt used approximately 300 devices, which shows BContCnt burdened a smaller group of devices creating higher congestion than BDur. In the case of MIT trace, both BDur and CICDur have utilized devices in a more balanced way as compared to BContCnt.

## 6. Conclusion

In this study, several adapted centrality-based routing metrics are evaluated for opportunistic network routing. The centrality measures have been computed using three metrics that preserve the link characteristics among nodes. Influential nodes concerning each centrality measures are identified to analyze the performance of centrality measures. The results show that betweenness centrality-based metrics are twice as good as the closeness centrality metrics. Moreover, the performance of betweenness metrics has been comparable to the epidemic routing. The overhead of all centrality-based routing mechanism is significantly lower than that of epidemic routing.

All transformations can be calculated locally (no global network knowledge required). However, the centrality measure computation has to be adapted to allow any node to estimate its centrality along with its neighbors to make efficient forwarding decisions. In the future, we intend to devise a mechanism to estimate local centrality measures so that the individual nodes can make routing decisions with the help of the information available in their immediate neighborhood [30].

## Data Availability

The data used to support the findings of this study have been deposited in the reference [24] repository.

## Additional Points

**Highlights.** The routing simulations in this paper are performed by adapting the real-life social networks that are extracted using the traces of wireless devices. The results presented in this paper are extended from [31] which indicates that the message delivery ratio of the congestion aware metrics is observed compatible with the message delivery ratio for epidemic routing. The highlights of the article are as follows: evaluation of routing performance of three opportunistic network metrics, i.e., *contact count*, *intercontact time*, and *average contact duration*, to harness link congestion information; a comprehensive congestion analysis concerning both complete network and individual nodes of the network metrics by simulation using three real-life social networks.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] A. Dhungana and E. Bulut, "Energy balancing in mobile opportunistic networks with wireless charging: single and multi-hop approaches," *Ad Hoc Networks*, vol. 111, article 102342, 2021.
- [2] M. Islam and M. Waldvogel, "Optimizing message delivery in mobile-opportunistic networks," in *Baltic Congress on Future Internet Communications*, pp. 134–141, Riga, Latvia, 2011.
- [3] A. Jain, "Betweenness centrality based connectivity aware routing algorithm for prolonging network lifetime in wireless sensor networks," *Wireless Networks*, vol. 22, no. 5, pp. 1605–1624, 2016.
- [4] Y. Xiao and J. Wu, "Data transmission and management based on node communication in opportunistic social networks," *Symmetry*, vol. 12, no. 8, p. 1288, 2020.
- [5] L. C. Freeman, "Centrality in social networks: conceptual clarification," *Social Networks*, vol. 1, pp. 215–239, 1979.
- [6] G. Sabidussi, "The centrality index of a graph," *Psychometrika*, vol. 31, no. 4, pp. 581–603, 1966.
- [7] P. Bonacich, "Power and centrality: a family of measures," *American Journal of Sociology*, vol. 92, no. 5, pp. 1170–1182, 1987.
- [8] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer networks and ISDN systems*, vol. 30, no. 1-7, pp. 107–117, 1998.
- [9] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of ACM*, vol. 46, no. 5, pp. 604–632, 1999.
- [10] C. Miralda, E. Donald, M. Ikeda, K. Mutso, L. Barolli, and M. Takizawa, "Effect of node centrality for IoT device selection in opportunistic networks: a comparison study," *Concurrency and Computation: Practice and Experience*, vol. 30, no. 21, article E4790, 2018.
- [11] A. Rasheed, S. Amjal, and A. Qayyum, "Adaptive routing update approach for VANET using local neighbourhood change information," *Malaysian Journal of Computer Science*, vol. 27, pp. 307–327, 2014.
- [12] P. Yuan, P. Liu, and S. Tang, "Exploiting partial centrality of nodes for data forwarding in mobile opportunistic networks," in *IEEE 17th International Conference on Computational Science and Engineering (CSE)*, Chengdu, China, 2014.
- [13] G. Goudar and S. Batabyal, "Point of congestion in large buffer mobile opportunistic network," *IEEE Communications Letters*, vol. 24, no. 7, pp. 1586–1590, 2020.
- [14] H. B. Liaqat, F. Xia, Q. Yang, L. Liu, Z. Chen, and T. Qiu, "Reliable TCP for popular data in socially-aware ad-hoc networks," in *Proceedings of the 15th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, New York, NY, USA, 2014.
- [15] J. Wang, C. Jiang, T. Q. S. Quek, X. Wang, and Y. Ren, "The value strength aided information diffusion in socially-aware mobile networks," *IEEE Access*, vol. 4, pp. 3907–3919, 2016.
- [16] H. Zhu, M. Dong, S. Chang, Y. Zhu, M. Li, and X. Shen, "Zoom: scaling the mobility for fast opportunistic forwarding in vehicular networks," in *IEEE INFOCOM Proceedings*, pp. 2832–2840, Turin, Italy, 2013.
- [17] P. Santi, *Mobility Models for Next Generation Wireless Networks*, John Wiley & Sons Ltd, 2012.
- [18] M. Tulu, M. Mkiramweni, R. Hou, S. Feisso, and T. Younas, "Influential nodes selection to enhance data dissemination in

- mobile social networks: a survey,” *Journal of Network and Computer Applications*, vol. 169, article 102768, 2020.
- [19] C. Miralda, E. Donald, M. Ikeda, K. Mutso, L. Barolli, and M. Takizawa, “A decision-making system based on fuzzy logic for IoT node selection in opportunistic networks considering node betweenness centrality as a new parameter,” in *International Conference on Intelligent Networking and Collaborative Systems*, Cham, 2020.
- [20] C. M. Kim, I. S. Kang, Y. H. Han, and Y. S. Jeong, “An efficient routing scheme based on social relations in delay-tolerant networks,” in *Lecture Notes in Electrical Engineering Heidelberg*, pp. 533–540, Springer, Berlin, 2014.
- [21] K. M. Sivalingam and V. Chellappan, “Application of entropy of centrality measures to routing in tactical wireless networks,” in *2013 19th IEEE Workshop on Local Metropolitan Area Networks (LANMAN)*, Brussels, Belgium, 2013.
- [22] T. Amah, M. Kamat, K. Bakar, W. Moreira, A. Oliveira, and M. Batistita, “Preparing opportunistic networks for smart cities: collecting sensed data with minimal knowledge,” *Journal of Parallel and Distributed Computing*, vol. 135, pp. 21–55, 2020.
- [23] R. Akbar, F. Safaiea, and E. Khodadad, “A novel adaptive congestion-aware and load-balanced routing algorithm in networks-on-chip,” *Computers and Electrical Engineering*, vol. 71, pp. 60–76, 2018.
- [24] N. P. A. Eagle, “Reality mining: sensing complex social systems,” *Journal Personal and Ubiquitous Computing*, vol. 10, no. 4, pp. 255–268, 2006.
- [25] P. Cao, G. Li, A. Champion, D. Xuan, S. Romig, and W. Steve, “On human mobility predictability via WLAN logs,” in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, Atlanta, GA, USA, 2017.
- [26] A. Trivedi, J. Gummesson, and P. Shenoy, “Empirical characterization of mobility of multi-device internet users,” 2020, <https://arxiv.org/abs/2003.08512>.
- [27] G. Poucin, B. Farooq, and Z. Patterson, “Activity patterns mining in Wi-Fi access point logs,” *Computers, Environment and Urban Systems*, vol. 67, pp. 55–67, 2018.
- [28] F. Ganji, Ł. Budzisz, F. G. Debele et al., “Greening campus WLANs: energy-relevant usage and mobility patterns,” *Computer Networks*, vol. 78, pp. 164–181, 2015.
- [29] P. B. Suvadip Batabyal, “Analysing social behaviour and message dissemination in human based delay tolerant network,” *Wireless Networks*, vol. 21, no. 2, pp. 513–529, 2015.
- [30] Z. Gao, Y. Shi, and S. Chen, “Identifying influential nodes for efficient routing in opportunistic networks,” *Journal of Communications*, vol. 10, no. 1, pp. 48–54, 2015.
- [31] M. A. Islam, M. A. Iqbal, M. Aleem, and Z. Halim, “Analysing connectivity patterns and centrality metrics for opportunistic networks,” in *International Conference on Communication, Computing and Digital Systems (C-CODE)*, Islamabad, 2017.