Impact of Human Behavior on Social Opportunistic Forwarding

Waldir Moreira[a,*], Paulo Mendes[a]

[a]*COPELABS, University Lusofona, Lisbon, Portugal*

Abstract

The current Internet design is not capable to support communications in environments characterized by very long delays and frequent network partitions. To allow devices to communicate in such environments, delay-tolerant networking solutions have been proposed by exploiting opportunistic message forwarding, with limited expectations of end-to-end connectivity and node resources. Such solutions envision non-traditional communication scenarios, such as disaster areas and development regions. Several forwarding algorithms have been investigated, aiming to offer the best trade-off between cost (number of message replicas) and rate of successful message delivery. Among such proposals, there has been an effort to employ social similarity inferred from user mobility patterns in opportunistic routing solutions to improve forwarding. However, these research effort presents two major limitations: first, it is focused on distribution of the intercontact time over the complete network structure, ignoring the impact that human behavior has on the dynamics of the network; and second, most of the proposed solutions look at challenging networking environments where networks have low density, ignoring the potential use of delay-tolerant networking to support low cost communications in networks with higher density, such as urban scenarios. This paper presents a study of the impact that human behavior has

---

on opportunistic forwarding. Our goal is twofold: i) to show that performance in low and high density networks can be improved by taking the dynamics of the network into account; and ii) to show that the delay-tolerant networking can be used to reduce communication costs in networks with higher density by taking the behavior of the user into account.

1. Introduction

Wireless devices have become more portable and with increased capabilities (e.g., processing, storage), which is creating the foundations for the deployment of pervasive wireless networks, encompassing personal devices (e.g. smartphones and tablets). Additionally, wireless technology has been extended to allow direct communication: vehicle-to-vehicle - for safety information exchange; device-to-device - aiming at 3G offloading; Wi-Fi direct - overcome the need for infrastructure entities (i.e., access points).

The combination of pervasive wireless devices and direct wireless communication solutions can be used to support the deployment of two major type of applications: end-to-end communication in development regions, since today's Internet routing protocols may operate poorly in such environments, characterized by very long delay paths and frequent network partitions; and low cost communication, namely data sharing, in urban scenarios, to bypass expensive data mobile communications and the unreliable presence of open Wi-Fi access points.

These networking scenarios (from development regional to large urban scenarios) are characterized by network graphs with different densities, which pose different challenges in terms of data forwarding. The challenge that we aim to tackle in this paper is to investigate the impact of human behavior on oppor-

2

tunistic forwarding, namely the awareness about users' social and data similarities.

Most of the prior art has been studying data transfer opportunities between wireless devices carried by humans, by looking at the distribution of the intercontact time, which is the time gap separating two contacts between the same pair of devices [1]. In challenging networking environments, opportunistic contacts among mobile devices may improve communications among peers as well as content dissemination, mitigating the effects of network disruption. This gave rise to the investigation of opportunistic networks, of which Delay-Tolerant Networks (DTN) are an example, encompassing different forwarding proposals to quickly send data from one point to another even in the absence of an end-to-end path between them. Such proposals range from flooding content [2] in the network up to solutions that take into account the social interactions among users [3, 4, 5, 6, 7, 8]. In the latter case, wireless contacts are aggregated into a social graph, and a variety of metrics (e.g., centrality and similarity) or algorithms (e.g., community detection) have been proposed to assess the utility of a node to deliver a content or bring it closer to the destination. Nevertheless, the structure of such graphs is rather dynamic, since users' social behavior and interactions vary throughout their daily routines. This brings us to our first assumption: forwarding algorithms should be able to exploit social graphs that reflect people's dynamic behavior. Prior art have studied forwarding algorithms that consider only the global network structure, without taking people's behavior into account [9]. In this paper, we show that forwarding algorithms that exploit social graphs reflecting the variations in people's daily routines are able to improve the performance of social-aware opportunistic networking.

Our second challenge was to analyze how to expand the deployment of DTN technology, which is normally seen only as useful to allow communications in challenging environments, such as development regions. For this study, we focus on data sharing since this should be the most interesting application to take advantage of low cost communication in dense networks, such as in an urban scenario. In this case, we studied two hypothesis to ensure good performance

3

when the density of the network increases: i) forwarding based on social graphs, where aggregation is based only on social similarities; and ii) forwarding based on behavior graphs, where aggregation is done by combining different aspects of human behavior, such as social similarities and data similarities (derived from the interests that users demonstrate in specific type of data).

Hence, in this paper we aim to investigate the possibility of developing an opportunistic forwarding system able to support low-cost services in dense networking scenarios as well as basic services in extreme networking conditions, by exploiting social as well as data similarities among users. Our work shows which type of opportunistic forwarding scheme is more suitable for delay-tolerant applications, based on the density of the network in scenarios spanning from developing regions to urban environments. Our findings lead to a new research challenge aiming to expand the impact of DTNs: the investigation of self-awareness mechanisms able to adapt their forwarding schemes based on the context of the user, namely the density of the network where he/she is currently.

The remainder of the paper is structured as follows. Section 2 aims to motivate our work, namely in what concerns the goal to study methods to expand the deployment of DTNs, and the impact that a better understanding of human behavior can have in the development of efficient forwarding solutions. In Section 3 we present our definition of network density based on the deployment scenarios that we look at to pursuit our study and experiments. Section 4 presents a set of forwarding algorithms that are considered in our study, including our proposals. In Section 5, we show the performance results of opportunistic forwarding over different network densities. Section 6 concludes our work, and identifies future research challenges to expand the impact of DTNs.

2. Motivation

The growing number of mobile devices equipped with a wireless interface and the end-user trend to shift toward wireless technology are opening new possibilities for networking. In particular, opportunistic communication embodies

4

a feasible solution for environments with scarce or costly infrastructure-based connectivity. A lot of attention has been given to the development of opportunistic forwarding solutions for networks with scarce connectivity, which are considered a natural fit for DTN technology. However, it is our belief that opportunistic forwarding can also be applied to more dense networks, were Internet communications are expensive, or applications aim to take advantage of direct communications among people.

In what concerns challenging networks, the most common approach has been to make use of social similarities to improve performance over the overall network. In this case, our work is motivated by the fact that such approaches ignore the behavior patterns that people present in their daily routines [10], which may lead to further performance improvements.

In what concerns the application of DTN and opportunistic forwarding to dense networks, most of the prior art aims to implement a store-carry-and-forward communication model that exploits specific devices found in urban scenarios, such as buses [11] and cars [12]. It is our understanding that the development of opportunistic forwarding solutions should not depend upon specific equipments only found is some scenarios, since this mitigates the deployment expansion of such proposals. It is our belief that the success of the DTN technology depends on its deployment range, which can only be ensured if such technology is based on pervasive wireless devices, such as smartphones: these devices are present in development regions as well as urban scenarios. In the latter case, communications between smartphones can also exploit mobility patterns of different vehicles ridden by people.

In order to design useful applications, it is vital to have a good understanding of the target environment and its users. Different types of user behavior may result in different network conditions and shall have a huge impact on whether or not a particular application is of interest to the user. A fair amount of work has been done on studying human mobility traces in order to gain understanding of real life mobility patterns and how those affect the properties of the opportunistic networks that are possible in that environment [13, 1, 14].

5

Although mobility patterns are important properties of the network, it is also important to understand the impact the human behavior, such as data interests, have on these networks. Hence, our work aims to tackle this new research trend, expanding social awareness to human behavior awareness. Among the different human behavior metrics that can be considered, we focus our attention on data similarities since data sharing is the most common application in the Internet. The study of data similarities depends on which applications are in place in the network and how the users use them. Usage patterns also depend on the users' context, so the same data patterns do not apply to all users. Approximations of some use cases might be possible to derive from the way cellular networks are used, but that will most likely not be applicable to all types of applications.

Looking at data similarities may improve the performance of opportunistic forwarding [3, 4]. However it is not clear if the improvement is higher than exploiting social interactions and structure (i.e., communities [5], as well as levels of social interaction [6, 7]). Thus, combining social and data similarities shall bring benefits (i.e., faster, better content reachability) to opportunistic forwarding. Hence, in this work we aim to show when the exploitation of social similarities results in a good performance, and when such performance can be augmented by combining them with data similarity metrics.

3. Network Scenario Characterization

One can observe that a networking scenario may vary according to its density. Sparse scenarios are characterized by very long delays (e.g., space communications [15]) and communication suffers with frequent disruption mostly due to the lack of infrastructure and geographic location (e.g., rural areas [16], riverside communities [17, 18]). It is common to see solutions relying on message ferries [19] or data mules [20] as to overcome the missing infrastructure. This is a classic scenario whose challenges are more related to transport protocols (i.e., dealing with extremely high delays) than routing itself.

6

Table 1: Network densities of the considered scenarios

Scenario	Cambridge	MIT	Synthetic
Identified density	26.83	47.01	148.80

In what concerns dense scenarios, communication can take place through both infrastructured (e.g., access points, cell towers) and infrastructureless (e.g., WiFi direct, bluetooth) means. Still, disruption remains a problem, but now seen from a different perspective: the dynamic bebavior of users (e.g., high mobility), different sources of interference (e.g., overlapping spectrum), poor coverage (e.g., areas full of closed access points) are factors that may contribute to link intermittency, despite all the available surrounding infrastructure. With the advances in the industry for portable devices and wireless technologies, this type of scenario is easily observed nowadays in urban settings.

In this work, we see dense (urban) scenarios as imposing new research challenges for opportunistic networks given the aforementioned characteristics. Due to the popularity and capabilities of mobile devices, users want to be able to sent and retrieve data anytime and anywhere. In other words, a free Internet scenario with low cost denominator networking imposed to users is a reality.

With this in mind, we define the density of the network according to the surroundings of the users and independently of the existence of infrastructure: what matters is i) that nodes can communicate directly; and ii) that the average node degree reflects the number of contact opportunities a node may have in such specific scenarios. We studied the characteristics of the three scenarios which are considered in the experiments in Sec. 5: the CRAWDAD traces of Cambridge [21] that corresponds to contacts of 36 students during their daily activities, the MIT [22] traces comprising 97 Nokia 6600 smart phones distributed among the students and staff of this institution, and the synthetic mobility scenario that encompasses 150 walking people, following the *Shortest Path Map Based Movement* model (i.e., nodes randomly choose destinations and use the shortest path to reach them). With the Gephi v0.8.2 [23] analysis tool, we accounted for the network densities of these scenarios summarized in Table 1.

7

Table 1 displays the scenarios in increasing order of density. Thus, it is expected that the routing solutions have an increasing performance behavior as network density increases. This is due to the different contact opportunities that a node may have, which increase (and therefore can be beneficial) for routing purposes.

4. Opportunistic Forwarding in Wireless Networks

This section presents the most relevant and latest opportunistic forwarding proposals, considering whether they make use of social and/or data similarity metrics. Similarity metrics are used to build graphs over which such forwarding proposals operate [24]. That is, instead of considering the number and frequency of contacts due to the mobility of hosts, such approaches take into account more stable social (e.g., common social groups and communities, node popularity, levels of centrality, social relationships and interactions, user profiles) and/or data (e.g., shared interests, interest of users in the content traversing network, content availability, type of content) aspects, aiming to reduce the cost of opportunistic forwarding. Moreover, opportunistic forwarding proposals may take into account the dynamics of user behavior, i.e., the resulting social graphs may consider what happens in terms of social interactions throughout the daily routine of the users.

Table 2 summarizes the type of similarity (i.e., social and/or data) considered by the opportunistic forwarding proposals and whether (or not) they consider the observed user behavior to build dynamic social graphs.

Bubble Rap [5], *CiPRO* [7], *SocialCast* [3], and *ContentPlace* [4] belong to the category that considers social similarity and/or data similarity metrics, but does not suitably reflect the dynamism of user behavior in the underlying social graph.

Bubble Rap combines node centrality with the notion of community to make forwarding decisions. The centrality metric identifies hub nodes inside (i.e., local) or outside (i.e., global) communities. Messages are replicated based on

8

Table 2: Opportunistic Forwarding Proposals

Proposals	Social similarity metrics	Data similarity metrics	Dynamic graphs
Bubble Rap	Communities and centrality		
CiPRO	User profile		
SocialCast		Shared interests	
ContentPlace	Social relationship and communities	Interest on the content	
dLife	Social weight and node importance		✓
SCORP	Social weight	Content type and interest on the content	✓

global centrality until they reach the community of the destination host (i.e., a node belonging to the same community). Then, it uses the local centrality to reach the destination inside the community.

CiPRO considers the time and place nodes meet throughout their routines. *CiPRO* holds knowledge of nodes (e.g., carrier's name, address, nationality, ...) expressed by means of profiles that are used to compute the encounter probability among nodes in specific time periods. Nodes that meet occasionally get a copy of the message only if they have higher encounter probability towards its destination. If nodes meet frequently, history of encounters is used to predict encounter probabilities for efficient broadcasting of control packets and messages.

SocialCast considers the interest shared among nodes. It devises a utility function that captures the future co-location of the node (with others sharing the same interest) and the change in its connectivity degree. Thus, the utility function measures how good message carrier a node can be regarding a given interest. *SocialCast* functions are based on the publish-subscribe paradigm, where users broadcast their interests, and content is disseminated to interested parties and/or to new carriers with high utility.

ContentPlace considers information about the users' social relationships to improve content availability. It computes a utility function for each data object

considering: i) the access probability to each object and the involved cost in accessing it; ii) the social strength of the user towards the different communities which he/she belongs to and/or has interacted with. The idea is having the users to fetch data objects that maximize the utility function with respect to local cache limitations, and choosing those objects that are of interest to users and can be further disseminated in the communities they have strong social ties.

The next category considers solely social similarity metrics and take into account the dynamism of user behavior while building the underlying social graph. *dLife* [6] is in this category and it takes into account the dynamism of users' behavior found in their daily life routines to aid forwarding. The goal is to keep track of the different levels of social interactions (in terms of contact duration) nodes have throughout their daily activities in order to infer how well socially connected they are in different periods of the day. Forwarding takes place by considering either the social strength (i.e., weight) among users or their importance in specific time periods.

Finally, the last category comprises both social and data similarity metrics and the dynamism observed in the behavior of users. *SCORP* [8] belongs to this category. It considers the type of content and the social relationship between the parties interested in such content type. *SCORP* nodes are expected to receive and store messages considering their own interests as well as interests of other nodes with whom they have interacted before. Data forwarding takes place by considering the social weight of the encountered node towards nodes interested in the message that is about to be replicated.

For the remainder of this paper, we consider one representative from each of the described categories: *Bubble Rap*, for being solely based on social similarity metrics; *dLife* and *SCORP*, for considering social and data similarity metrics and for being the proposals which satisfactory capture the dynamic user behavior in the resulting social graphs. These proposals are enough to help us illustrate how the performance of opportunistic routing proposals in networks with different densities can be further improved by considering the user dynamic behavior.

10

5. Evaluation of Opportunistic Forwarding over Different Network Density Scenarios

In this section, we analyze the performance behavior of *Bubble Rap, dLife,* and *SCORP* over different network density scenarios. With the experiments in this section, we want i) to show that performance in low (Sec. 5.2) and high (Sec. 5.3) density networks can be improved by taking the dynamics of the network into account; and ii) to show that the delay-tolerant networking can be used to reduce communication costs in networks with higher density by taking the behavior of the user into account (Sec. 5.4).

5.1. Methodology and Simulation Settings

Simulations are carried out in the Opportunistic Network Environment (ONE) simulator [25]. Results are presented with a 95% confidence interval and in terms of averaged delivery probability (i.e., ratio between the number of delivered messages and the number of messages that should have been delivered), cost (i.e., number of replicas per delivered message), and latency (i.e., time elapsed between message creation and delivery).

The trace scenarios comprise 36 (Cambridge) and 97 (MIT) nodes carrying devices during their daily activities. The synthetic mobility scenario simulates 3 groups (A, M, and B) of 50 people each, who carry nodes equipped with 250-Kbps Bluetooth interfaces, and moving with speed up to 1.4 m/s. The reason for considering traces and synthetic mobility scenario relates to the fact that: i) with the former, we have a representation of real user behavior; and ii) with the latter, we are able to have a network density much higher from the perspective of the user, as defined in Sec. 3. Most analysis have been based on datasets with low density, collected in a constrained setting, which is not representative for realistic use cases of the networks being studied. If one is interested in the properties of a large scale urban environment, it is probably not meaningful to study traces collected from 36 or 97 user at a conference or university campus.

Across all experiments, proposals experience the same load and number of messages that must reach the destinations. In the Cambridge trace (cf. Sec.

11

5.2), the *Bubble Rap/dLife* source sends 1, 5, 10, 20 and 35 different messages to each of the 35 destinations, while the *SCORP* source creates 35 messages with unique content types, and the receivers are configured with 1, 5, 10, 20, and 35 randomly assigned interests. Thus, we have a total of 35, 175, 350, 700, and 1225 generated messages. The *msg/int* notation represents the number of messages sent by *Bubble Rap* and *dLife* sources, or the number of interests of each of the *SCORP* receivers. Since *Bubble Rap/dLife* sources generate more messages, in this scenario node 0 (the source) has no buffer restriction and message generation varies with the load: 35 messages/day rate (load of 1, 5, and 10 messages), and 70 and 140 messages/day rates (load of 20 and 35 messages, respectively).

As for the synthetic mobility scenario (cf. Sec. 5.3), 200 messages are generated. With *Bubble Rap* and *dLife*, node 0 (group *A*) generates 100 messages to nodes in groups *B* and *M*, and node 100 (group *B*) generates 100 messages to nodes in groups *A* and *M*. For *SCORP*, each group has different interests: group *A* (*reading*), group *B* (*games*), and group *M* (*reading* and *games*). The source nodes, 0 and 100, generate only one message for each content type, *game* and *reading*. This guarantees the same number of messages expected to be received, i.e., 200. Also, by varying the node pause times between 100 and 100000 seconds, we have different levels of mobility (varying from 3456 to 3.4 movements in the simulation). In this scenario, all source nodes have restricted buffer, but rate is of 25 messages every 12 hours. This is done so that *Bubble Rap/dLife* do not discard messages prior to even trying exchange/deliver them given the buffer constraint.

Finally, in Sec. 5.4, the load generated is equivalent to 6000, 78000, and 200 messages to be delivered across all experiments for Cambridge, MIT, and synthetic mobility scenarios, respectively.

Regarding message TTL, we set it to be unlimited in order to observe the performance behavior (i.e., buffer consumption, number of replicas) of the forwarding proposals in networks with high traffic load. Message size ranges from 1 to 100 kB. Despite nodes may have plenty of storage, we consider nodes having different capabilities (i.e., smartphones). Thus, nodes have buffers limited

12

to 2 MB as we consider that nodes may not be willing to share all their storage space. The performance evaluation follows the guidelines of a Universal Evaluation Framework (UEF) [26?] to guarantee fairness in the assessment.

As for proposals, *Bubble Rap* uses the K-Clique and cumulative window algorithms for community formation and centrality computation as in [5]. As for *dLife* and *SCORP*, both consider 24 daily samples (i.e., each of one hour) as mentioned in [8].

5.2. Performance over Low Density Network

This section presents the performance of opportunistic forwarding proposals over a low density network scenario. Fig. 1 presents the average delivery probability with different messages and interests being generated.
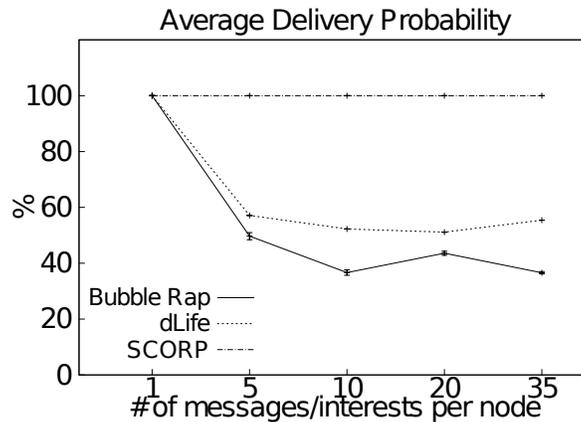

Figure 1: Delivery under different network loads

In the 1 msg/int configuration, formed communities comprise almost all nodes. This means that each node has high probability to meet any other node, which is advantageous for *Bubble Rap* since most of its deliveries happen to nodes sharing communities. Due to the dense properties of the network, *dLife* and *SCORP* take advantage of direct delivery: 57% and 51% of messages, respectively, are delivered directly to destinations.

As load increases, *Bubble Rap* has an 50% decrease in delivery performance. This occurs since it relies on communities to perform forwardings, and conse-

quently buffer space becomes an issue. To support this claim, we estimate buffer usage for the 5 msg/int configuration: there is an average of 80340.7 forwardings, and if this number is divided by the number of days (12^1) and by the number of nodes (35, source not included), we get an average of 191.28 replications per node. Multiplied by the average message size (52kB), the buffer occupancy is roughly 9.94 MB in each node, which exceeds the 2MB allowed (cf. Sec. 4.1).

This estimation is for a worst case scenario, where *Bubble Rap* spreads copies to every encountered node. Since this cannot happen, as *Bubble Rap* also relies on local centrality to reduce replication, buffer exhaustion is really an issue given that messages are replicated to fewer nodes and not to all as in our estimation. As more messages are generated, replication increases: this causes the spread of messages that potentially take over forwarding opportunities from other messages, reducing *Bubble Rap*'s delivery capability.

dLife has a 43% performance decrease when network load increases, as it takes time to have an accurate view of the social weights. This leads to forwardings that never reach destinations given the contact sporadicity. For the 10 msg/int configuration, *dLife* also experiences buffer exhaustion: estimated consumption is 2.17 MB per node. Still, by considering social weights or node importance allows *dLife* a more stable behavior than *Bubble Rap*.

Since content is only replicated to nodes that are interested in it, or have a strong social interaction with other nodes interested in such content, the delivery capability of SCORP raises as the ability of nodes to become a good carrier increases (i.e., the more interests a node has, the better it is to deliver content to others, since they potentially share interests). The maximum estimated buffer consumption of *SCORP* is of 0.16 MB (35 msg/int).

Fig. 2 presents the average cost behavior. In the 1 msg/int configuration, all proposals create very few replicas to perform a successful delivery, 7.95 (*Bubble Rap*), 14.32 (*dLife*), and 23.46 (*SCORP*), as they rely mostly on shared communities and/or direct deliveries. We also observe that *SCORP* produces more

---

[1] In simulation it is worth ˜12 days of communications.

14

replicas than *dLife* due to a particularity in its implementation: *SCORP* nodes with interest in a specific content of a message not only process it, but also replicate it to other interested nodes, thus creating extra replicas.

For the 5, 10, 20 and 35 msg/int configurations, replication is directly pro-portional to the load. Thus, cost is expected to increase as load increases, as seen with *dLife*. The same performance behavior was expected for *Bubble Rap*. However, the observed cost peaks relate to the message creation time and contact sporadicity: when a message is created in a period of high number of contacts, which results in much more replications. This is more evident with *Bubble Rap* at the 5 msg/int configuration as it relies on shared communities to forward: as mentioned earlier, most of the communities comprise almost all nodes, which increases its replication rate.

Despite their efforts, these replications do not improve their delivery prob-abilities, contributing only to the associated cost for performing successful de-liveries.
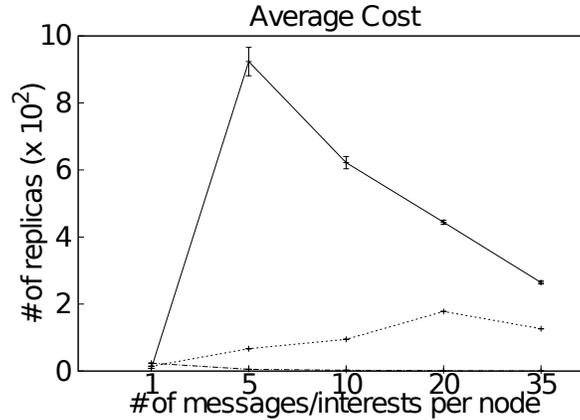

Figure 2: Cost under different network loads

With more interests, a *SCORP* node can serve as a carrier for a larger number of nodes. Consequently, the observed extra replicas make the proposal rather efficient: *SCORP* creates an average of 6.39 replicas across all msg/int

configurations, while *Bubble Rap* and *dLife* produce an average 452.41 and 96
replicas, respectively.

Fig. 3 shows the average latency that messages experience. The latency peak in the 1 msg/int configuration refers to the message generation time: some messages are created during periods where very few contacts (and sometimes none) take place followed by long periods (12 to 23 hours) with almost no contact. Consequently, messages are stored longer, contributing to the increase of the overall latency. This effect is mitigated as the load increases with messages being created almost immediately before a high number of contacts take place.
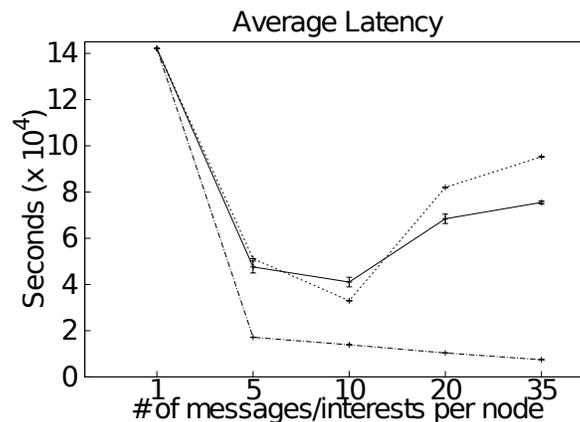

Figure 3: Latency under different network loads

Since latency is in function of the delivered messages, the decrease and variable behavior of *Bubble Rap* and *dLife* is due to their delivery rates decrease and increase, and also to their choices of next forwarders that may take longer to deliver content to destinations. *SCORP* experiences latencies up to approx. 90.2% and 92.2% less than *Bubble Rap* and *dLife*, respectively. The ability of a node to deliver content increases with the number of its interests. Thus, a node can receive more messages when it is interested in their contents, and consequently becomes a better forwarder since the probability of coming into contact with other nodes sharing similar interests is very high, thus reducing latency.

5.3. Performance over High Density Network

This section presents the performance of opportunistic forwarding proposals over a high density network scenario.

Fig. 4 presents the average delivery probability. Given the community formation characteristic of this scenario, *Bubble Rap* relies mostly on the global centrality to deliver content. By looking at centrality [5], we observe very few nodes (out of the 150) with global centrality that can actually aid in forwarding, i.e., 19.33% (29 nodes), 10.67% (16 nodes), 21.33% (32 nodes), and 2% (3 nodes) for 100, 1000, 10000, and 100000 pause time configurations, respectively. So, these nodes become hubs and given buffer constraint and infinite TTL (i.e., messages created earlier take the opportunity of newly created ones), message drop is certain, directly impacting *Bubble Rap*.
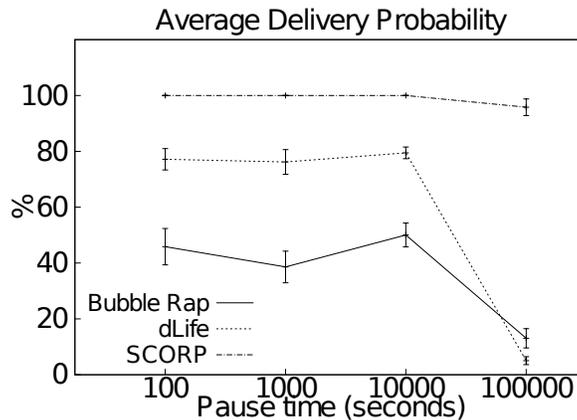

Figure 4: Delivery under varied mobility rates

Given the high number of contacts, the computation of social weight and node importance done by *dLife* takes longer to reflect reality: thus *dLife* replicates more and experiences buffer exhaustion. Indeed, social awareness is advantageous, but still not enough to reach optimal delivery rate in such conditions.

Independent of the number of contacts among nodes, *SCORP* can still identify nodes that are better related to others sharing similar interests, reaching optimal delivery rate for 100, 1000, and 10000 pause time configurations. By

considering nodes' interest in content and their social weights, *SCORP* does not suffer as much with node mobility as *dLife* and *Bubble Rap*.

With 100000 seconds of pause time, the little interaction happening in a sporadic manner (with intervals between 20 and 26 hours) affects *Bubble Rap*, *dLife* and *SCORP* as they depend on such interactions to compute centrality, node importance, and social weights, as well as to exchange/deliver content.

Fig. 5 presents the average cost behavior. As pause time increases, the number of contacts among nodes decreases, providing all solutions with the opportunity to have a stable view of the network in terms of their social metrics with 100, 1000, and 10000 seconds of pause time. This explains the cost reduction experienced by *Bubble Rap* and *dLife*: both are able to identify the best next forwarders, which results in the creation of less replicas to perform a successful delivery.
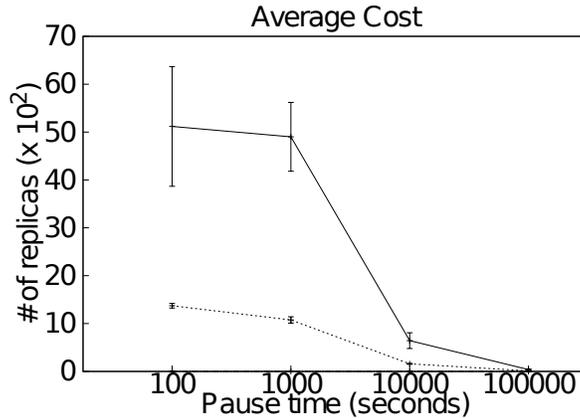

Figure 5: Cost under varied mobility rates

SCORP has a very low replication rate (average of 0.5 replicas) given its choice to replicate based on the interest that nodes have on content and on their social weight towards other nodes interested in such content. When the intermediate node has an increased number of interests (i.e., by having different interests, the node can potentially deliver more content) as observed in Sec. 5.2, replication costs are even lower. Furthermore, *SCORP* suitably uses buffer space with an estimated average occupancy of 0.03 MB per node per day.

With 100000 seconds of pause time, as cost is in function of delivered messages (and deliveries are very low, due to contact sporadicity), proposals have a low cost.
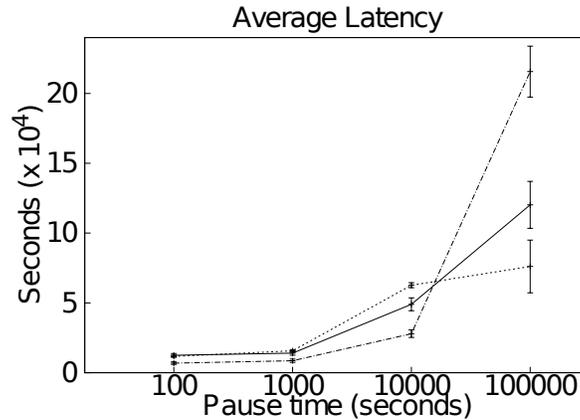

Figure 6: Latency under varied mobility rates

As expected (cf. Fig. 6), latency increases as node mobility decreases: encounters are less frequent, and so content must be stored for longer times. Also, the time that the social metrics take to converge (i.e., a more stable view of the network in terms of centrality, social weight, and node importance) contributes for the increase in the experienced latency. The highest increase in latency with 100000 seconds of pause time is due to contacts happening in a sporadic fashion with intervals between them of up to 26 hours, thus proposals take much longer to perform a delivery.

5.4. Performance over Different Network Densities

This section shows how network density impacts on the performance of opportunistic forwarding proposals. As mentioned before, we want to bring attention to dense scenarios found in urban settings: a panoply of heterogenous devices that could overcome disruption by interacting directly with one another to improve the networking experience of users. Fig. 7 presents the average delivery probability with different network densities.
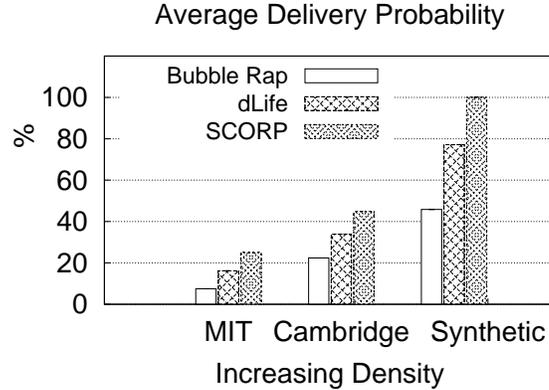
19

Figure 7: Delivery under different network densities

As mentioned in Sec. 3, it was expected that the performance of social-aware opportunistic routing improve with the increase of network density. However, we can observe that content-oblivious *Bubble Rap* and *dLife* experience a decrease in performance in the MIT scenario despite of its identified density (47.01) being almost almost twice as the one identified in Cambridge (26.83). The reason for such behavior lies on the characteristics of each scenario, with MIT nodes covering a much bigger area. Despite of having a higher number of contacts between nodes, the MIT scenario may lead to messages reaching nodes that are not the best forwarders, and these messages, given the unlimited TTL, may end up taking the delivery opportunity of newly created messages. This directly affects the performance of both *Bubble Rap* and *dLife.*

Yet the content-oriented *SCORP* overcomes such features of the MIT scenario since it also considers those nodes that are interested in the content being replicated or that are strongly related to interested parties. Unlike the content-oblivious solutions, *SCORP* has a 6.14% improvement despite the challenging scenario.

Performance behavior for all proposal indeed improve with higher network density (148.8, Synthetic scenario). The reason is tied to the fact that a higher density indicates more contact opportunities for the exchange of messages.

20

Fig. 8 presents the average cost which is expected to increase with network density. This is because the more contact opportunities the scenario has, the more replicas are created by proposals. This can be easily seen with *Bubble Rap*, which creates an average of more than 5000 replicas to perform a successful delivery.
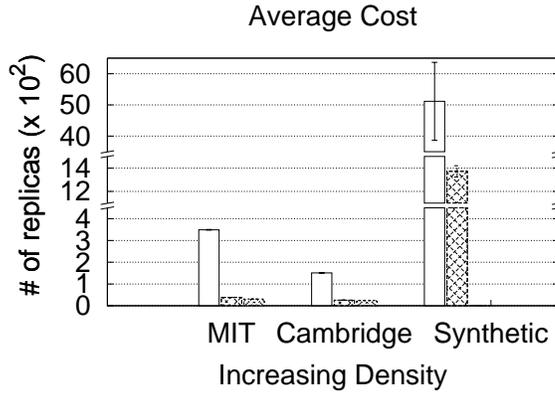
Average Cost

Figure 8: Cost under different network densities

The same cost increasing trend is observed with *dLife*, but in different orders and especially for the synthetic scenario. We believe that the much higher number of copies (up to 98% when compared to other scenarios) is related to the mobility rate. In the Synthetic experiment nodes move a lot, which results in a high number of contacts with many nodes. Consequently, *dLife* takes longer to have a stable view of the network in terms of its social weights and node importances, which leads to the creation of unwanted replicas. *SCORP* has the best cost performance (up to 0.5 replicas to perform a successful delivery), since the more interests a node has, the better forwarder it is. The *M* group (cf. Sec. 5.1) accounts for 50% of the interests existing in the network, which makes it a greater carrier for messages.

Fig. 9 shows the average latency experienced by messages since their creation up to their reception at destination. In the traces experiments, all proposals keep the same trend: average latency increases. This is due to the fact that

nodes in these experiments span different areas and the encounter frequency happens at different rates. This can result in messages being forwarded to nodes who may reach a given destination, but delivery time increases with the area and duration of experiments. SCORP presents a much higher increase in the MIT experiment, as it takes its time to suitably choose the next forwarders (based on their interest on the message's content or social relationship to other interested parties). This added to fact that interactions among nodes happen according to area they move jointly contribute to such latency peak.
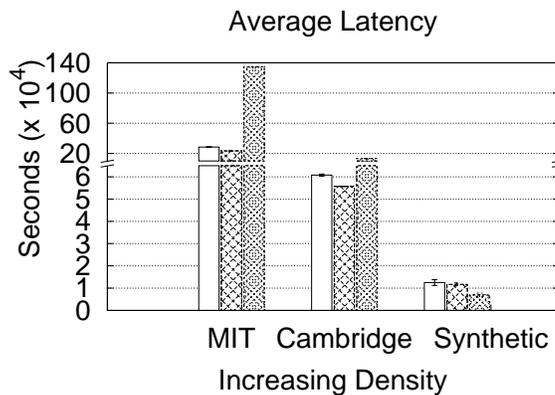

Figure 9: Latency under different network densities

As for the Synthetic mobility experiment, not only the proposals have many different contact opportunities, but also nodes encounter more frequently. This consequently has a positive impact for *Bubble Rap*, *dLife*, and *SCORP* that are able to deliver content in less time (12486s, 11710s , and 6864s, respectively).

6. Conclusions

Opportunistic forwarding can aid communication in two major application scenarios: end-to-end communication in development regions and low cost communication in urban scenarios. Such scenarios do have different network densities which adds more challenges to opportunistic forwarding. The underlying graphs, over which these opportunistic forwarding proposals operate, comprise

22

(e.g., common social groups and communities, node popularity, levels of central-ity, social relationships and interactions, user profiles) and/or data (e.g., shared interests, interest of users in the content traversing network, content availabil-ity, type of content) aspects. Additionally, such graphs may (or not) take into account the dynamics of user behavior.

Thus, in this paper, we exploit the possibility of having an opportunistic forwarding system that can provide support to i) low-cost services in dense net-working scenarios; and ii) basic services in extreme networking conditions (e.g., communications in development regions), considering social and data similar-ities among users as well as the dynamic behavior found in the users' daily routines.

Our results show that opportunistic forwarding, based on social and data similarity metrics and considering the dynamism observed in the users behavior, does answer the communication needs of users in both dense (i.e., urban) and challenged (i.e., development region) scenarios. Performance improvements go up to 54% regarding delivery capability while latency and cost can be reduced by 45% and 99% respectively, when compared to forwarding solely based on data similarity and completely agnostic to user behavior.

These findings point to a new research challenge regarding the impact of DTN application: the investigation of self-awareness mechanisms able to adapt their forwarding schemes based on the context of the user, namely the density of the network where he/she is currently.

References

[1] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, J. Scott, Impact of human mobility on opportunistic forwarding algorithms, IEEE Trans. Mobile Comput. 6 (6) (2007) 606–620.

[2] A. Vahdat, D. Becker, Epidemic routing for partially connected ad hoc networks, Tech. Rep. CS-200006, Duke University (2000).

[3] P. Costa, C. Mascolo, M. Musolesi, G. P. Picco, Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks, IEEE J.Sel. A. Commun. 26 (5) (2008) 748–760.

[4] C. Boldrini, M. Conti, A. Passarella, Design and performance evaluation of contentplace, a social-aware data dissemination system for opportunistic networks, Comput. Netw. 54 (4) (2010) 589–604.

[5] P. Hui, J. Crowcroft, E. Yoneki, Bubble rap: Social-based forwarding in delay-tolerant networks, IEEE Trans. Mobile Comput. 10 (11) (2011) 1576–1589.

[6] W. Moreira, P. Mendes, S. Sargento, Opportunistic routing based on daily routines, in: Proceedings of IEEE WoWMoM, San Francisco, USA, 2012.

[7] H. A. Nguyen, S. Giordano, Context information prediction for social-based routing in opportunistic networks, Ad Hoc Netw. 10 (8) (2012) 1557–1569.

[8] W. Moreira, P. Mendes, S. Sargento, Social-aware opportunistic routing protocol based on user's interactions and interests, in: M. H. Sherif, A. Mellouk, J. Li, P. Bellavista (Eds.), Ad Hoc Networks, Springer International Publishing, 2014, pp. 100–115.

[9] T. Hossmann, T. Spyropoulos, F. Legendre, Know thy neighbor: towards optimal mapping of contacts to social graphs for dtn routing, in: Proceedings of INFOCOM, San Diego, USA, 2010, pp. 866–874.

[10] N. Eagle, A. Pentland, Eigenbehaviors: identifying structure in routine, Behavioral Ecology and Sociobiology 63 (7) (2009) 1057–1066.

[11] I. Komnios, V. Tsaoussidis, Carpool: Extending free internet access over dtn in urban environment, in: Proceedings of ACM MobiCom LCDNet, Miami, USA, 2013.

[12] P.-C. Cheng, K. Lee, M. Gerla, J. Harri, Geodtn+nav: Geographic dtn routing with navigator prediction for urban vehicular environments, Mobile Networks and Applications 15 (1) (2010) 61–82.

24

[13] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, C. Diot, Pocket switched networks and human mobility in conference environments, in: Proceedings of ACM SIGCOMM WDTN, Philadelphia, USA, 2005.

[14] T. Hossmann, T. Spyropoulos, F. Legendre, A complex network analysis of human mobility, in: Proceedings of INFOCOM, Shanghai, China, 2011.

[15] S. Farrell, V. Cahill, Delay- and disruption-tolerant networking, Artech House telecommunications library, Artech House, 2006.

[16] S. Jain, K. Fall, R. Patra, Routing in a delay tolerant network, in: Proceedings of SIGCOMM, Portland, USA, 2004.

[17] W. Moreira, R. Ferreira, D. Cirqueira, P. Mendes, E. Cerqueira, Socialdtn: A dtn implementation for digital and social inclusion, in: Proceedings of ACM MobiCom LCDNet, Miami, USA, 2013.

[18] R. Ferreira, W. Moreira, P. Mendes, M. Gerla, E. Cerqueira, Improving the delivery rate of digital inclusion applications for amazon riverside communities by using an integrated bluetooth dtn architecture, International Journal of Computer Science and Network Security 14 (1) (2014) 17–24.

[19] W. Zhao, M. Ammar, E. Zegura, A message ferrying approach for data delivery in sparse mobile ad hoc networks, in: Proceedings of ACM MobiHoc, Roppongi Hills, Japan, 2004.

[20] R. Shah, S. Roy, S. Jain, W. Brunette, Data mules: modeling a three-tier architecture for sparse sensor networks, in: Proceedings of IEEE Workshop SNPA, Anchorage, USA, 2003.

[21] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, A. Chaintreau, CRAWDAD trace cambridge/haggle/imote/content (v. 2006-09-15), http://crawdad.cs.dartmouth.edu/cambridge/haggle/ imote/content (Sep. 2006).

[22] N. Eagle, A. Pentland, CRAWDAD data set mit/reality (v. 2005-07-01), http://crawdad.cs.dartmouth.edu/mit/reality (Jul. 2005).

[23] M. Bastian, S. Heymann, M. Jacomy, Gephi: An open source software for exploring and manipulating networks, in: Proceedings of AAAI Conference on Weblogs and Social Media, San Jose, California, 2009.

[24] W. Moreira, P. Mendes, Social-aware opportunistic routing: The new trend, in: I. Woungang, S. K. Dhurandher, A. Anpalagan, A. V. Vasilakos (Eds.), Routing in Opportunistic Networks, Springer New York, 2013, pp. 27–68.

[25] A. Keränen, J. Ott, T. Kärkkäinen, The one simulator for dtn protocol evaluation, in: Proceedings of Simutools, Rome, Italy, 2009.

[26] W. Moreira, P. Mendes, S. Sargento, Assessment model for opportunistic routing, in: Proceedings of LATINCOM, Belem, Brazil, 2011.