UNIVERSIDADE LUSÓFONA
| Escola de Comunicação, Artes e Tecnologias da Informação

siti
Informatics Systems and Technologies Research Unit
Unidade de I&D em Sistemas e Tecnologias Informáticas

Survey on Opportunistic Routing for Delay/Disruption Tolerant Networks

SITI Technical Report SITI-TR-11-02

2010

Authors/Editors

Author	Affiliation
Waldir Moreira	SITI, Lusófona University
Paulo Mendes	SITI, Lusófona University

Survey on Opportunistic Routing for Delay/Disruption Tolerant Networks

Waldir Moreira, and Paulo Mendes
SITI, Lusofona University, Portugal.
{waldir.junior, paulo.mendes}@ulusofona.pt

Abstract

Complex Dynamic Networks are characterized by a stochastic nature where the behaviour of nodes is random and unknown. From the routing viewpoint, data exchange is rather challenging in such networks since paths between any pair of nodes may never exist or delay may be too long to be accepted by current data transport protocols. However, opportunistic routing has proven to be able to cope with the stochastic behaviour present in networks that display topologies with patterns of connection between nodes that are neither purely regular nor random. With opportunistic routing, forwarding decisions are made using locally collected knowledge about node behaviour to predict which nodes are likely to deliver a content or bring it closer to the destination. Currently, there are a significant number of proposals to opportunistically route data based on time-variant graphs used by Delay/Disruption Tolerant Networks (DTN), each with a different goal and based on different evaluation criteria. Hence, this technical report aims to survey and analyse literature with the goal of devising classification and evaluation methods to characterize, analyse and compare existing and future proposals to route over complex networks such as DTNs.

Keywords

Complex Networks; Delay/Disruption Tolerant Networks, Opportunistic Routing, Taxonomy, Evaluation Framework

I. Introduction

Life style in modern society is creating an increasing demand for users to have constant capability to exchange information. However, providing pervasive connectivity to users that have a dynamic behaviour is challenging since most social and technological networks display topologies with patterns of connection between nodes that are neither purely regular nor purely random. Examples of such networks are online social networks and Delay/disruption Tolerant Networks (DTN). DTNs are networks of self-organizing wireless nodes connected by multiple time-varying links, and where end-to-end connectivity is intermittent. Even in urban scenarios, it is possible to face intermittent connectivity due to dark areas, such as inside buildings and metropolitan systems, as well as public areas with closed access points or even places overcrowded with wireless access points. Unavailability of wireless connectivity can be also a result of power-constrained nodes that frequently shut down their wireless cards to save energy [1].

Due to intermittent connectivity, routing protocols based on the knowledge of end-to-end paths perform poorly, and numerous opportunistic routing algorithms have been proposed instead. Some opportunistic routing protocols use replicas of the same message to combat the inherent uncertainty of future communication opportunities between nodes. In order to carefully use the available resources and reach short delays, many protocols perform forwarding decisions using locally collected knowledge about node behaviour to predict which nodes are likely to deliver a content or bring it closer to the destination (cf. Fig. 1). Nodes must have enough processing power and storage to keep data until another good intermediate carrier node or the destination is found [2], following a store-carry-and-forward (SCF) paradigm.

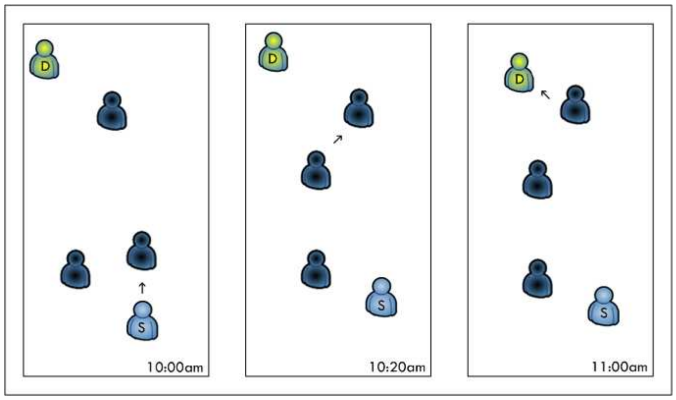

Figure 1. Example of opportunistic routing.

In DTNs, communication opportunities are not entirely random, which means that communication patterns (e.g., such as time of last encounter between two nodes) can be used to attempt inferring future contact opportunities as well as to exploit node mobility in order to achieve a fast delivery of data by flooding the network (e.g., *Epidemic* [3]). Similar results can be achieved by controlling flooding based on: i) delivery probability optimizations (e.g., *Spray and Wait* [4]); ii) history of encounters (e.g., *PROPHET* [5]); priority (e.g., *MaxProp* [6]); future prediction (e.g., *EBR* [7]). A recent approach of predicting future contact opportunities is to aggregate previous contacts in a social graph and to use metrics from complex network analysis (e.g., centrality and similarity) to assess the utility of a node to carry a piece of content. Such approaches apply the concept of communities formed by nodes based upon their social relationships (e.g., *Bubble Rap* [8]), by considering user mobility and interests (e.g., *SocialCast* [9]), and node popularity (e.g., *PeopleRank* [10]).

Although routing proposals tried to present some improvement to cope with link intermittency, there is no clear understanding about what the most suitable solution may be. The major reason for this poor understanding about existing routing proposals is the lack of a taxonomy to identify common properties, and the lack of an objective framework to evaluate existing approaches. In this context, the goal of this technical report is three fold: first, to provide an objective comparison of the most representative proposals. Second, based on the comparative analysis, to propose an objective classification and evaluation framework of existing and future opportunistic routing proposals. Last but not least, by surveying the literature, we aim to identify a consolidated set of open research issues.

There are currently several approaches aiming to devise different opportunistic routing solutions. This technical report aims to devise classification and evaluation methods to characterize, analyse and compare existing and future proposals to route over complex networks such as DTNs. The technical report is structured as follows: To better introduce the problem space, section II provides a model of DTNs, as an example of a complex network. In section III, we provide an overview of current opportunistic routing solutions, while section IV gives an insight into existing taxonomies for opportunistic routing as well as our own proposal. Section V describes our proposal to devise an universal evaluation framework for opportunistic routing. In section VI, we provide some insights about potential open research issues identified based on an objective comparison of existing proposals. We conclude in section VII.

II. A DTN MODEL OF COMPLEX DYNAMIC NETWORKS

This section introduces a generic model for DTNs that is important to sustain a better understanding about opportunistic routing. We started by describing a DTN graph model in terms of the type of contacts and network behaviour. Then, we discuss how network density is affected by social graphs.

A. Generic Model

From a conceptual point of view, a DTN consists of a node meeting schedule and workload. A node meeting schedule is a directed multigraph $G = (V, E)$, where V and E represent the set of nodes and edges, respectively. Each direct edge E between two nodes (V_1, V_2) represents a meeting opportunity between them, and it is annotated with a tuple $E = (t_s, t_e, s, T)$, where t_s is the starting time of the meeting, t_e is the ending time of the meeting, if known, s is the size of the transfer opportunity (i.e., contact capacity), and T is the contact type. The workload is a set of messages $M = \{(u_1, v_1, s_1, t_1, P), (u_2, v_2, s_2, t_2, P),...\}$, where the i^{th} tuple represents the source, destination, size, time of creation at the source, and priority, respectively, of message i.

As mentioned before, a DTN model encompasses the notion of type of contact or connectivity (T). In current networks, the connectivity of a link or path is generally given as a binary state (i.e., connected or disconnected). In DTNs, a richer set of connectivity options is required to make efficient routing decisions. Most importantly, links (and paths, by extension) may provide a scheduled, predicted or opportunistic communication. *Scheduled contacts* imply some a priori knowledge about adjacent nodes regarding future availability of links for message forwarding. Scheduled links are the most typical cases for today's Internet and satellite networks.

Predicted contacts correspond to communication opportunities wherein the probability of knowing whether a link will be available at a future point in time is strictly above zero and below one. Such links are the result of observed behaviour (e.g., a person may use its home Internet connection with significant probability for any given time period) being characterized using statistical estimation. Predicted links are only becoming of serious concern recently, namely in ad-hoc wireless networks where node mobility may be significant.

However, in more challenging environments, such as mission-critical networks for instance, the future location of communicating entities may not be neither known nor predictable. These types of contacts are known as opportunistic. Such *opportunistic contacts* [1] are defined as a chance to forward messages towards a specific destination or a group of destinations. In such unpredictable scenario, it is important to take into account the time that a node must wait until it meets another node again (i.e., *inter-contact time*), the duration of these contacts (i.e., *contact duration*), and the quality of the contact in terms of the set of information that can be transferred (i.e., *contact volume*).

Independently of the type of connectivity, a link in a DTN is direction-specific. For example, a dial-up connection originating at a customer's home to an Internet Service Provider (ISP) may be scheduled from the point of view of the customer but unscheduled from the point of view of the ISP.

Another concept that must be introduced is that of *network behaviour*. As networks can be formed on-the-fly, their behaviour can either be deterministic or stochastic [11], depending on the type of used links. In this technical report, we focus on dynamic scenarios where the behaviour of the network is described in stochastic terms, based on users' mobility and social behaviour. In a dynamic scenario, users move around carrying their personal devices, which opportunistically come into contact with each other, resulting in topology changes. There are three major cases, based on the level of mobility in the network. First, it is possible that only a few set of entities are mobile. In this case, contacts appear and disappear based solely on the quality of the communication channel between them. Second, it is possible that some, but not all, nodes in the network are mobile. These nodes may be exploited to guarantee some level of network connectivity. Since they are the primary source of transitive communication between two non-neighboring nodes in the network, an important routing question is how to properly distribute data among these moving nodes. Third, it is possible that the vast majority, if not all, of the nodes in the network are mobile. In this case, a routing protocol will most likely have more overlapping forwarding options during contact opportunities. A realistic model of the mobility behaviour of nodes (e.g., community-based mobility model [12], time-variant community mobility model [13]) is very useful to devise efficient opportunistic routing protocols.

Although forwarding may be done opportunistically based on the mobility behaviour of humans, the efficiency of routing proposals may depend on the accuracy of mobility prediction. Considering that human behaviour is based on social relationships, a social-aware approach to opportunistic routing may present higher efficiency, since human relationships are less volatile than human mobility. Hence, an open research area is the utilization of social models into opportunistic routing. Such social models may be based on information retrieved, when possible, from online social networks, as well as from the "social" behaviour of networking nodes. Social behaviour of networking nodes may be based on inter-meeting times (i.e., showing common mobility behaviour), contact duration (i.e., showing proximity) as well as contact volume (i.e., showing common interests and trust). Such social parameters define *community* behaviour (i.e., represent the social ties with other nodes), as well as *friends* (i.e., where social ties are determined based on interest and trust) [14]. However, a social model of DTNs should also take into account *familiar strangers* (i.e., nodes with no social ties, but that have frequent meetings due to common habits) as well as *strangers* (i.e., nodes with no social ties, but that meet sporadically).

Other characteristic of network behaviour is the capacity of nodes (i.e., buffering capability and contact volumes). Many nodes, such as mobile phones, are limited in terms of storage space, transmission rate, and battery life. Others, such as buses, may not be as limited. Routing protocols can utilize this information to best determine how messages should be scheduled, transmitted, and stored in order to avoid the starvation of nodes. Only recently has the scientific community started taking resource management for DTNs into consideration, and this is still an active area of research.

B. Network Density

The first DTNs were developed to scenarios where communication suffers with very long delay links (e.g., space communications) or where communication is often disrupted, meaning that contacts do not occur very frequently (e.g., rural areas). The challenges imposed by these sparse networks are more related to transport protocols than routing. For instance, in space communications the problem is to find a transport protocol able to send packets over wireless links with very long delays. Another example applies to rural areas, where the lack of telecommunications infrastructure leads to the adoption of ferry type of solutions [15], in which the source passes the information to a moving node that carries it directly to the destination(s). Besides space communications and message ferrying, another example of a solution suitable to be applied to sparse networks are Data mules [16], in which specific nodes move around collecting information from several nodes (e.g., sensors) carrying such information to a well-known destination.

These and other routing solutions to sparse networks were already characterized by Zhang (2006) [11]. We aim to analyse the utilization of opportunistic routing in daily life scenarios, based on the social and mobility behaviour of people in urban areas with dark places (i.e., areas either full of closed APs or with many open APs but with very high interference) or places without coverage at all (i.e., subway, public areas).

A social graph offers a natural representation of the resulting node contact set over time. In such a graph, an edge could mean that two nodes see each other frequently, because they have a social connection (i.e., friends), or because they are frequently in the same place without actually knowing each other (i.e., familiar strangers). Hence, a graph edge is intended to have predictive value for future contacts: Time Expanded Graphs [17] have been proposed to capture timing information of contacts into a dynamic graph.

However, considerable scalability issues arise, as one would essentially need to store a graph for every time instant. On the one hand, one could create an edge if at least one contact has occurred in the past between the two nodes, but this would result in a significant dense graph, after a certain network lifetime. On the other hand, a past contact could represent an edge only if occurred during a given time interval. However, if the time interval is too small, the resulting graph might be too sparse.

In both cases, meaningful differentiation between nodes using complex network analysis may be rendered impossible. It is thus important to carefully design this mapping in order for edges in the graph to maintain their predictive value. In this case, the routing challenges refer to the ability of forwarding messages towards destinations with high delivery probability, within a time frame that is useful for the lifetime of the message, and with good usage of network resources.

III. Opportunistic Routing Approaches

This section aims to provide an analysis of the most relevant proposals to perform opportunistic routing, following the concepts and scenarios described in section II.

According to Jain et al. [18], deterministic routing approaches are excellent from a performance point of view, since the more information a node can get from the network, the wiser its forwarding decision will be. However, the needed extra knowledge brings more complexity to the solution and even makes it impossible to be implemented due to the dynamic nature of DTNs. It was already shown that, although the optimal solution will need to have a broad knowledge about the network behaviour and traffic demands [18], even the most simple oracle, called *Contact Oracle*, which contains information about contacts between any two nodes, is unrealistic as it is equivalent to knowing the time-varying characteristics of such networks. Other oracles were defined, including a *Queuing Oracle* (i.e., information about instantaneous buffer occupancies), and a *Traffic Demand Oracle* (i.e., information about the present or future traffic demand), but their assumptions about the network behaviour is even more severe. The existence of such oracles would support deterministic routing algorithms able to compute an end-to-end path (possibly time dependent) before messages were actually transmitted. For example, if a *Contact Oracle* is available, modified Dijkstra with time-varying cost function based on waiting time would be enough to find the route. However, in DTNs the most realistic assumption is that network topology is not known ahead of time.

Based on the analysis of deterministic routing approaches, it is clear that the most suitable solution would be the one based on a local probabilistic decision, aiming to forward messages based on the opportunities raised by any contact within range. More elaborated solutions may also take other information into account to increase the efficiency of message progression towards a destination. Examples of such extra information are: history data about encounters, mobility patterns, priority of information, and social ties.

This section aims to contribute to a broad understanding of existing opportunistic routing approaches, which are of great interest to this technical report. These routing proposals are grouped into three categories: single-copy, aiming to improve the usage of network resources; epidemic, aiming to increase delivery probability; and, probabilistic-based, aiming to find an optimal balancing between both previous categories.

A. Single-copy Routing

In resource-constrained networks, which can occur in urban areas with high spectrum interference, opportunistic routing may lead to waste of resources when trying to deliver messages to a destination or set of destinations. Such waste of resources is mainly due to the utilization of message replication aiming to increase delivery probability.

Aiming to optimize the usage of network resources, some approaches avoid replication of messages, forwarding messages to single next-hops based on available connectivity and some form of mobility prediction. This means that these proposals perform single-copy forwarding, i.e., only one copy of each message traverses the network towards the final destination. Such copy can be forwarded if the node carrying it decides (i.e., randomly, or based on a utility function) that another encountered node presents a higher probability to deliver the message.

Minimum Estimated Expected Delay (MEED) [19] is an example of a single-copy forwarding approach that uses contact history (i.e., connection and disconnection times of contacts) to aid forwarding. Contact history is a metric that estimates the time a message will wait until it is forwarded. A per-contact routing scheme is used to "override" regular link-state routing decision. That is, instead of waiting for nodes enclosed in the path with the shortest cost (based on *MEED* values), it simply uses any other contact opportunity (i.e., node) that arises prior to what is expected to forward the message. To be able to do this, *MEED* must recompute routing tables each time a contact arrives and also broadcast this information to every other node in the network.

Spyropoulos et al. (2008) [20] present six examples where this type of forwarding is considered: i) *direct transmission*, where messages are forwarded only to the destination; ii) *randomized routing*, where messages are only forwarded to encountered nodes that have forwarding probability p, where $0 < p \leqslant 1$; iii) *utility-based routing with 1-hop diffusion*, where forwarding takes place if the utility function - based on encounter timers - of an encountered node is higher than the one of the current carrier; iv) *utility-based routing with* transitivity, in which an utility function towards a destination is updated also considering intermediate nodes that have high utility to such destination; v) *seek and focus routing*, where *randomized routing* is used to find the best starting point towards the destination and henceforth a utility-based approach is used to find the destination; and vi) *oracle-based optimal algorithm*, with which the future movement is known beforehand allowing optimal forwarding decisions to be taken aiming to delivery messages in a short amount of time.

From a resource consumption viewpoint, single-copy forwarding approaches are quite interesting, since they keep the usage of network (e.g., bandwidth) and node (e.g., energy, storage) resources at a low level. However, they suffer from high delay rates which, consequently, may result in a low delivery ratio. Another issue is related to the amount of knowledge that needs to be exchanged/available in order to aid forwarding, which in some scenarios generates too much overhead and may be impossible to implement.

To mitigate such problem, the following category of opportunistic routing approaches relies only on current contacts to increase delivery probability.

B. Epidemic Routing

Ubiquitous communication is a feature that has been present in our everyday life and requires messages to be delivered with high probability to their destinations, with the help of intermediary nodes able to implement the SCF paradigm.

Since networks are created by people moving around, opportunistic contacts are considered to increase delivery probability in networks with the aforementioned characteristics in a proposal called *Epidemic* [3]. With an *Epidemic* routing approach, every node in the network gets at least a copy of each message. Such a full replication strategy leads to an increase of the delivery rate. Replication of messages is done by means of summary vectors that are exchanged between any two encounters. Such summary vectors contain the list of messages each node is carrying, allowing nodes to exchange all messages that the other node is lacking. The proposal indeed increases delivery rate, since every potential forwarder has, with high probability, a copy of the message, assuming contacts with significant duration and sufficient buffer space in each node. However, to avoid waste of resources, each host sets a maximum buffer size that it is willing to allocate for epidemic message distribution. In general, nodes drop older messages in favor of newer ones upon reaching their buffer's capacity. This means that the efficiency of the delivery process depends upon the configured buffer space, which may be substantially improved with the usage of Bloom filters [21]. In order to avoid messages to be replicated indefinitely, a hop count field determines the maximum number of epidemic exchanges that a particular message is subject to, being packets dropped based on the locally available buffer space. Since the number of hops towards the destination is not known in advance setting a hop count may decrease the delivery probability. A stale-data removal mechanism could be more efficient by removing messages that were already delivered.

In an attempt to avoid waste of network resources, other proposals emerged based on a controlled replication approach. That is, the number of nodes which get a message copy is reduced by probabilistically choosing next nodes or by using a utility function.

C. *Probabilistic-based Routing*

Probabilistic approaches are based on the estimation/prediction of what is the next best set of carriers for each message based on some probability metric aiming to maximize delivery probability. Probabilistic forwarding protocols require node mobility patterns that exhibits long-term regularities such that some nodes consistently meet more frequently than others over time: the mean inter-meeting time between two nodes in the past will be close to that in the future with high probability.

Within this category, proposals attempt, first of all, to optimize the delivery probability while avoiding full replication of messages. Besides this core concern, there are proposals that take into account the capabilities of nodes, the priority of messages and the availability of resources aiming to achieve a high delivery ratio with low delay and resource consumption. Another concern of some probabilistic-based approaches is to use as few meta-data as possible aiming to decrease concerns with respect to energy, processing and bandwidth saving. Having considerable control overhead increases contention in the network resulting in message discarding and retransmissions.

Since 2003, different delivery probability metrics have been proposed including frequency encounters [5], [6], [22], [7], aging encounters [23], [24], [25], aging messages [4], [26], resource allocation [27], [28], and social similarity [29], [30], [8], [9], [10]. These are the probabilistic-based routing proposals that are analysed in this technical report.

1) Frequency Encounters: Proposals based on these metrics have in common the fact that they rely on the knowledge about how many times nodes meet in a network. The proposals considered in this sub-category are the *Probabilistic ROuting Protocol using History of Encounters and Transitivity* (PROPHET), *MaxProp, Prediction*, and *Encounter-Based Routing* (EBR).

One of the most well known approaches, being currently considered by the *DTN Research Group* (DTNRG) of the *Internet Research Task Force* (IRTF), is the *Probabilistic ROuting Protocol using History of Encounters and Transitivity* (PROPHET) [5]. *PROPHET* uses a probabilistic metric called delivery predictability, which indicates what is the likelihood of a node to deliver a message to a destination based on its past contacts with such destination. If a pair of nodes does not encounter each other for a while, they are less likely to be good forwarders of messages to each other, thus their delivery predictability is reduced in the process. Transitivity is a property of this predictability where if a node A regularly meets a node B and node B regularly meets node C, this implies that node C is also a good node to forward messages to node A. Delivery predictability helps to decide whether the node should replicate or not a given message. Upon a contact, nodes exchange summary vectors that also have information regarding delivery predictability. This information is used to update their own delivery predictability vectors, which are used to make decisions about message forwarding. *PROPHET* delivers messages only to nodes which are better (in terms of delivery predictability) than the current carrier node, resulting in a reduction of the consumption of network resources and a high probability of messages delivery.

However, flooding can still occur with *PROPHET* if the message is originated in a node with low delivery predictability (i.e., node with low mobility) towards the destination, and this node only encounters nodes with higher delivery predictability. Another potential drawback is the usage of unrealistic mobility models such as random waypoint, or models based on the community concept, but where the probability of moving between communities are statically defined and do not vary with time.

MaxProp [6] is another probabilistic-based approach that uses a metric called delivery likelihood of messages, by having each node keeping track of a probability of meeting any other peer. Using an incremental averaging method, nodes that are seen frequently obtain higher delivery likelihood values over time. Each time two nodes meet, they exchange their delivery likelihood probabilities towards other nodes. Based on the delivery likelihood values computed by other nodes and by itself, a carrier of a message computes a cost for each possible path to the destination, up to n hops. The cost for a path is the sum of the probability of each contact on the path not occurring. This cost estimation, along with the hop count, are then used to order messages for scheduling and for dropping. In addition, *MaxProp* assigns a higher priority to new messages (i.e., low hop count) to increase their chance of reaching the destination faster, and tries to prevent reception of the same message twice by including a hop list in each message, and uses acknowledgments to notify all nodes about message delivery. Upon contact, two nodes exchange messages in a specific priority order: first, messages that have these nodes as final destinations; second, information for estimating delivery likelihood; third, acknowledgments to remove stale messages; fourth, messages that have not traversed far in the network; and, fifth, send messages with highest priority.

By combining the estimation of message delivery likelihood with message priority and acknowledgments, *MaxProp* is able to reach good performance regarding message delivery probability and message latency with transfer opportunities limited in duration and bandwidth. However, it is considered that nodes have unlimited storage for their own messages and limited storage for the messages coming from other nodes, which has an impact on the overall performance.

It is important to mention that the delivery likelihood metric used in *MaxProp* is different from the delivery predictability metric employed in *PROPHET* in the sense that the former depends solely on the probability of nodes to meet each other while in *PROPHET* it depends on the probability to meet the destination itself, which means that *PROPHET* requires more state. However, *MaxProp* requires nodes to compute possible paths to the destination by concatenation of delivery probability between nodes while *PROPHET* does not require further computation, since messages are forwarded if a node has higher delivery predictability towards the destination than the carrier.

Song and Kotz proposed a prediction-based approach (hereafter referred to as *Prediction*) [22] that makes use of contact information to estimate the probability of meeting other nodes in the future. As happens with *PROPHET* and *MaxProp*, *Prediction* also uses historical contact information to estimate the probability of meeting other nodes in the future. However, unlikely previous approaches, *Prediction* estimates the contact probability within a period of time, based on a metric called timely-contact probability, which is used to compute the contact frequency between two nodes, as follows: the contact history between two nodes i and j is divided into a sequence of n periods of ΔT starting from the start time (t_0) of the first contact in history to the current time. If node i had any contact with node j during a given period m, which is $[t_0 + m\Delta T,\, t_0 + (m+1)\Delta T]$, the contact status of the interval I_m is set to 1. The probability of node i meeting node j in the next ΔT can be estimated as the average of the contact status in prior intervals. In *Prediction*, whenever two nodes meet, they exchange the indexes of all their messages. If the destination of a message is not the node in contact, the probability to deliver such message through that node is computed. If the probability of delivering the message via the contacted node (based on the past average number of encounters with the destination) within a defined period of time is greater than or equal to a certain threshold, the message is passed to the node in contact. This proposal presents two methods for choosing the next node: the decision can be taken by the node that is sending the message or by the one which is receiving it. In the former case, a meta-data message is necessary to determine if the receiver is a good next hop. As for the latter, the receiver decides whether or not to keep the copy of the message considering its own probability of coming into contact with the destination.

Prediction achieves good performance regarding message delivery with low number of message transmission and duplication. However, its performance and storage usage is directly proportional to the *Time-To-Live* (TTL) allowed for messages which makes it more suitable for networks tolerant to very long delays (i.e., sparse networks).

The proposal *Encounter-Based Routing* (EBR) [7] also considers the number of times nodes meet in order to predict the rate levels of future encounters. It simply counts the number of contacts a node has with other nodes (Current Window Counter) and determines the Encounter Value (EV) that represents the node's past rate of encounters. The higher EV is, the higher the probability of successful message delivery. This also determines the number of replicas of a message that the relay node will get in each contact. Nodes maintain their past rate of encounters to predict their rate of future encounters. When nodes meet, they first update their EV values and estimate the EVs ratio, which is used to determine the number of tokens of a message replica that will be passed to each neighbor. This is a kind of time-to-live parameter also used by other approaches such as *Spray and Focus*. For security reasons, prior to EV update, nodes exchange information that will aid them to determine if their EV values are correct.

The prediction of future encounters used by *EBR* allows the improvement of latency and message delivery by reducing traffic overhead (i.e., unwanted copies). Messages are only exchanged with nodes that have high encounter rate, which avoids routes that may not result in delivery, and minimizes network resource usage. In addition, the proposal implements a security measure that avoids black hole denial-of-service attacks from malicious nodes pretending to be part of the network and announcing fake EV values. However, this approach presents some drawbacks in scenarios with multiple communities that have low rate of inter-contact times. In these scenarios, packets may be forwarded to nodes that indeed have higher EV values, but such packets will get stuck within the source community. Also, this security measure incurs in wasting contact opportunities for determining the reliability of the encountered node, and the proposal may have its performance degraded in scenarios where nodes have short contact times.

2) Aging Encounters: In this sub-category, the age of encounters are taken into consideration, and the proposals that fall into it are the *Exponential Age SEarch* (EASE), *FResher Encounter SearcH* (FRESH), and *Spray and Focus.*

The *Exponential Age SEarch* (EASE) proposal [24], first presented in IEEE INFOCOM 2003, was one of the first proposals to consider history of encounters with a specific destination to support opportunistic forwarding. In addition to that, *EASE* also considers geographic position of nodes where a node would make routing decisions based on the time and location of its last encounter with every other node in the network. With *EASE,* every node maintains local information about the time and location of its last encounter with other nodes in the network. To be considered a good next hop, a node must either have a more recent encounter with the destination than the current holder of the packet or the node must be physically closer to the destination. Once the new hop is found, any position-based algorithm (e.g., DREAM [31]) can be employed to route the packet towards it. These two phases, namely next hop search and packet routing, are repeated until the best next hop is the destination itself. A second version of the proposal (*Greedy EASE* [24]) is able to change the chosen next hop during the routing phase as it checks the age of the last encounter with the destination at every hop.

EASE performs quite well in scenarios with nodes presenting random walk mobility patterns. The performance results show that the routes towards the destination have the same length (sometimes smaller) as the optimal case even for very large distances between the communicating nodes. However, this proposal is highly dependent on the mobility pattern and destination speed. Its performance is easily degraded if the destination moves fast turning the solution more costly (i.e., in terms of route length). Added to that, if mobile nodes shutdown their wireless interface (for energy-saving purposes), estimates considering location are of no use.

The *FResher Encounter SearcH* (FRESH) [23] proposal is an example of a "blind" routing protocol, since it has no notion of coordinates. Each node keeps track of the time elapsed since the last encounter with every other node, and uses this information to choose the next hop for message forwarding. When a sender wishes to initiate data forwarding, it must first search for the next hop that is determined by the time elapsed since the last encounter between this potential next hop and the destination. This search is omni-directional and is done in concentric rings with increasing radius until the next hop is found. It is necessary that nodes keep track of their one-hop neighbors to maintain encounter tables updated.

FRESH takes advantage of the time-distance correlation where the distance traveled during a time interval of duration t is positively correlated with t (e.g., a node met a few minutes ago is closer than a node met two hours before). With that, it is able to improve the performance of route discovery with no need for global knowledge of the network. Instead, the proposal is based on a distributed implementation where next hop search is defined in terms of local information (e.g., encounter tables). Like with *EASE*, performance depends on nodes mobility processes as the time-distance correlation becomes noisier with heterogeneous speeds. That means that a node that has just encountered the destination may not be close to it if the destination is moving too fast. Another issue is that *FRESH* may suffer with loops in routing if source and destination are not part of a connected subset of nodes. This is indeed a problem especially in scenarios where isolated cluster of nodes may be formed. There is also an overhead related to the need of having one-hop neighbor encounter table updated.

The more recent *Spray and Focus* [25] approach proposes a scheme were a fixed number of copies are spread initially exactly as in *Spray and Wait* [4] (with the subtle difference of using only $\frac{1}{3}$ or $\frac{1}{2}$ of the L messages normally used in *Spray and Wait*), but then each copy is routed independently according to single-copy utility-based scheme with transitivity [20]. In the spraying phase, ideally it would be good to be able to choose as relays the L nodes that most frequently encounter the destination. However, waiting for a “better” relay may mean that opportunities to spread extra copies are forfeited. Hence, the *Spray and Focus* scheme uses a greedy spraying phase by implementing a binary spraying algorithm to minimize the amount of time it takes to spray all L copies, moving the problem of looking for a possibly better relay to the focus phase. In the focus phase, each potential router maintains a timer for every other node in the network, recording the time elapsed since the two nodes came within transmission range for the last time. These timers are similar to the age of last encounter [23] and contain indirect location information. For a large number of mobility models, it can be shown that a smaller timer value on average implies a smaller distance from the node in question.

The *Spray and Focus* proposal outperforms flooding-based (i.e., *Epidemic*) and single-copy schemes (i.e., *Randomized Flooding*, and *Utility-based Flooding*) as well as the other spraying algorithm (i.e., *Spray and Wait*) under realistic mobility scenarios (e.g., modeling human behaviour), by forwarding messages to nodes which have a “closer” relationship (determined by the encounter timers) with the destination. Also, *Spray and Focus* presents good performance in scenarios with heterogeneous mobility using an algorithm that is able to diffuse timer information much faster than regular last encounter based schemes. However, since its performance is highly dependent on the use of encounter timers, it can be easily degraded in scenarios where nodes are highly mobile as timers quickly become obsolete.

3) Aging Messages: These proposals have in common the fact that they aim to avoid messages to be kept being forwarded in the network by creating metrics that define the age of message copies. *Spray and Wait*, and *Optimal Probabilistic Forwarding* (OPF) are the proposals comprised by this sub-category.

One of the first proposals was *Spray and Wait* [4], which decouples the number of transmissions per message from the total number of nodes, generating a small number of transmissions in a large range of scenarios. Initially, copies of a message are spread quickly in a manner similar to epidemic routing. However, *Spray and Wait* stops when enough copies have been sprayed in order to avoid flooding, while guaranteeing that at least one copy will reach the destination with high probability. By exploiting the mobility of nodes, *Spray and Wait* operates in two steps: first, the source determines a certain number of adjacent nodes that are going to get a copy of the message (spraying phase). In the second step, the nodes that got a copy of the message deliver it directly to the destination when it gets within range (waiting phase). Each generated message can have L copies of it distributed in the network. This number of copies can be determined in two ways: i) based on the number of nodes M and size of the network N; ii) by estimating M when both M and N are unknown. Once L is known, the source can spread the copies of the message by passing only one copy of the message (*Source Spray and Wait*), or $L/2$ copies (*Binary Spray and Wait*) to each encountered node. In the latter case, the receiver of $L/2$ copies will spread the obtained copies in the same way. If the destination is not found in the spraying phase, the nodes holding one copy of the message will forward it directly to the final destination.

Spray and Wait exploits node mobility being able to limit the total number of copies and transmissions per message resulting in an energy-efficient solution with low delivery delay, although the achieved delay is inversely proportional to the number of copies. If nodes move quickly enough around the network, *Spray and Wait* shows that only a small number of copies can create enough diversity to achieve close-to-optimal delays. However, there is no acknowledgment mechanism to get rid of copies of already delivered messages and no mechanism to select the best set of forwarders in the spraying phase. In what concerns computational effort, determining L is not an easy task, since it is necessary to know M beforehand and it depends on nodes performing independent random walks. This can easily result in an inaccurate measure of L, which degrades the algorithm's performance. This problem is even worse in large dense networks with frequent disconnections and nodes following different mobility patterns.

The *Optimal Probabilistic Forwarding* (OPF) [26] protocol replicates a message upon node encounter if, by doing so, such action increases the overall delivery probability of such message. That is, if this action maximizes the joint expected delivery probability of the copies to be placed in system (i.e., in the sender and receiver nodes of the message). *OPF* aims to maximize the delivery probability based on a particular knowledge about the network, relying on the assumptions that node mobility exhibits long-term regularity (enabling the estimation of mean inter-meeting times) and that each node knows the mean inter-meeting time of all pairs of nodes in the network. *OPF* metric reflects not only the direct delivery probability of a message, such as in *PROPHET*, but also the indirect delivery probability when the node can forward the message to other intermediate nodes, as in *MaxProp*. However, unlike *MaxProp*, *OPF* metric reflects a hop-count-limited forwarding scheme, based on a function of two important states of a message: remaining hop-count and residual lifetime. Such utility function may estimate the effect that packet replication may have on the expected delivery rate while satisfying the constant on the number of forwardings per message (*OPF* has a performance awareness as happens with *RAPID*, for instance).

With *OPF* every message has a residual time-to-live (T_r) that also denotes a given meeting time slot, and nodes know the mean inter-meeting time ($I_{i,j}$) between any two nodes i and j in the network. This is used to determine the meeting probability ($M_{i,j}$) among any two nodes and the delivery probability (P_{i,j,K,T_r}) between nodes i and j of a message with K remaining hop-count and T_r residual time to live. The delivery probability is simply given by $P_{i,d,0,T_r}$ if the message cannot use anymore hops to reach the destination. However, when the message is at K hops from the destination, forwarding will take place if the combined probability of the two new copies of the message at the next time-slot T_{r-1} (i.e., $1-(1-P_{i,d,K-1,T_{r-1}}) \times (1-P_{j,d,K-1,T_{r-1}})$) is greater or equals to the probability of not forwarding it at all (i.e., $P_{i,d,K,T_{r-1}}$). So, when node i meets node j, whether i should forward the copy to j depends on whether replacing the copy in i with two logically new copies (i.e., in i and j) increases the overall delivery probability. *OPF* also comes in another version where K is then substituted by the number of logical tickets (L, as in *Spray and Wait*) which are going to be distributed between the two replacing copies in a message forwarding.

OPF is able to achieve good overall delivery rate with a subtle increase in delay since it only forwards messages to really good relay nodes. Its performance is better if the relationship between nodes is greater and the hop count (K) allowed for each message is chosen wisely. However, this good performance comes with a cost, since it is really dependent on the amount of routing information available. In networks where only local information is available due to the dinamicity of nodes, *OPF* will have its performance degraded since it needs the mean inter-meeting time of all nodes in the network and such information may be difficult to obtain. Added to that, since the mobility model considered follows an exponential inter-meeting time, the measurements may not represent human behaviour as it is known that a power law distribution better represents such behaviour [14].

4) Resource Allocation: In what concerns approaches that are aware of available resources, we have the *Resource Allocation Protocol for Intentional DTN* (RAPID), and *PRioritized EPidemic* (PREP).

The proposal dubbed *Resource Allocation Protocol for Intentional DTN* (RAPID) [28] opportunistically replicates messages based upon a utility function that estimates the effect that packet replication may have on a predefined performance metric in a network with resource constraints. When nodes meet, they exchange meta-data about packets to be delivered and acknowledgements about already delivered packets, along with packets destined to each other. With the meta-data, nodes are able to determine the marginal utility (which must have the highest increase based on the performance metric considered) of replicating packets between them. *RAPID* calculates the effect of replication considering resources constraints by exchanging meta-data through an in-band control channel that allows it to have a global state of the network resources (e.g., length of past transfers, expected meeting times, list of delivered packets, delivery delay estimate for buffered packets, changed information on packets since last exchange).

RAPID has a small cost in the usage of contact opportunities due to the utilization of an in-band control channel. Moreover, the information exchanged in such channel may not always be updated due to node mobility, delivery delay, and unacknowledged packets. This cost may be very high in bandwidth-constrained scenarios with short-lived contact opportunities. Besides that, there are no overall performance guarantees since heuristics are based on sub-optimal solutions supported by one metric at a time. In addition, the performance is also related to the used mobility pattern (e.g., predictable vehicular movements) and can be degraded in scenarios with unpredictable mobility patterns.

In what concerns metric-based approaches, the *PRioritized EPidemic* (PREP) [27] proposal is also based on message prioritization and the idea of prediction. *PREP* introduces the average availability (AA) metric that measures the average fraction of time a link will be available in the future (i.e., the inter-node cost) and defines drop and transmit priorities (in which lower values indicate high priority) for each message. By using a regular discovery algorithm, each node finds out about its links towards neighboring nodes. According to available information about the past state of the link (i.e., up/down), a node can determine the availability of the links for future use. Then, costs are assigned to links based on their AA values and epidemically broadcasted in the network. To find the lowest cost path to a destination (or its whereabouts), the Dijkstra's shortest path algorithm is employed. Any changes to AA values will trigger link costs updates, which are again broadcasted through *Link State Advertisements*.

PREP is able to generate a gradient of message replication density that is inversely proportional to the distance towards the destination. That is, messages' copies are kept as close as possible to their destinations. *PREP* sets a drop priority where messages shall be discarded, upon a full buffer, according to a cost determined by the distance between the message's holder and destination. The greater the distance, the higher the drop priority. In addition, a transmission priority is also set to messages considering their expiry time and the cost previously mentioned. This allows a wiser usage of resources (e.g., storage and bandwidth) having a good effect on message delivery. However, in dynamic scenarios, *PREP* has its performance degraded since its delivery capability is inversely proportional to the level of disruption happening in the scenario.

5) Social Similarity: Within the previous solutions for opportunistic routing, one of the processes that may lead to significant consumption of energy, processing capability and bandwidth is the prediction of mobility patterns, which is quite common to all probabilistic routing approaches. One alternative may be to devise probabilistic solutions that exploit not only mobility of nodes but also their social similarities. The reason is that mobility patterns change faster (causing the appearance of unwanted traffic due to out-of-date information) than social relationships between people within a society. The fact that social relationships are less volatile than mobility behaviour has been proven [8] to be rather useful in forwarding decisions. Hence, since 2007 several approaches have been investigating the exploitation of social relationships as well as individual interests in order to improve the delivery rate while decreasing the consumption of network resources. The proposals that are part of this sub-category are *Label*, *SimBet*, *Bubble Rap*, *SocialCast*, and *PeopleRank*.

The *Label* [29] approach was one of the first proposals to employ social characteristics into opportunistic routing. Experiments were conducted in IEEE INFOCOM 2006 where nodes were labeled telling others about their affiliation/group, and this allowed nodes to forward messages directly to destinations, or to next hops belonging to the same group (i.e., same label) as the destinations. The proposal looked not only to inter-contact time distribution for all the nodes inside a group but also to the inter-contact time distribution between two groups (i.e., friendship ties) with results showing that nodes from one group may be good forwarders for nodes in the corresponding friendship group. The *Label* proposal provided the first indication that exploiting social similarities improves delivery ratio and especially delivery cost (i.e., total number of messages and duplicates transmitted). Another observation is that friendship between different communities (i.e., unusual connections among nodes of both communities) can be used to slightly improve delivery ratio.

Label's performance is directly related to the allowed message TTL and the mixing rate of nodes in the scenario. Delivery ratio is very low in the case of messages with short TTL, and it is easily degraded if nodes do not mix well. The reason is that *Label* performs only one-hop delivery and just to nodes belonging to the same community as the destination, which means that delivery may fail if the sender never encounters members of the same community as the destination.

Some network nodes may have such a behaviour that make the usage of encounters inefficient to forward messages, due to their sporadic meeting rates. For instance, a node may be involved in a highly clustered network in which none of the nodes have directly or indirectly met the destination node. However, paths between clusters may be insured by nodes that form bridges based on weak acquaintance ties. In this context, the *SimBet* [30] approach proposes to forward data based on the identification of these bridges and the identification of nodes that reside within the same cluster as the destination node. The major contribution is a new forwarding metric based on ego network analysis to locally determine betweenness centrality of nodes (i.e., importance of a node in the system, defined as the number of connections between nodes belonging to different communities that cross the referred node) and social similarity (i.e., probability of future collaboration between nodes in the same community). These two social parameters are determined from an adjacency matrix that each node keeps to track contacts (direct and through neighbor nodes) with other nodes in the network. When a node i meets a node j, they exchange messages they have to each other and request each other's list of contacts. With this list, they can update their own contact list along with their betweenness (Bet) and similarity (Sim) values. Then, they exchange summary vectors that contain the destinations to which they are carrying messages along with their updated Bet and Sim values. For each destination in the summary vector, they determine their *SimBet* utility. With this, a vector of destinations is created containing all the destinations to which the node has highest *SimBet* utility. This vector of destinations is exchanged and nodes exchange messages to destinations present in their own vector and remove such messages from their buffers.

SimBet is able to forward messages even if the destination node is unknown to the sending node or its contacts. In this case, the message is routed to a structurally more central node where the potential of finding a suitable carrier is much higher. Moreover, *SimBet* makes no assumptions about the control of node movements or knowledge of node future movements. Finally, with *SimBet* messages are forwarded solely based on locally obtained information. It works based on forwarding a single copy of each message in the network, which makes it able to reduce resource consumption, mainly buffer space and energy. It has good overall performance regarding message delivery, which is close to *Epidemic*'s but with highly improved delivery cost (i.e., very low number of required forwards to reach destination). However, this proposal may suffer with high delay since the level of contact (i.e., how often and with whom nodes meet) between nodes is a key aspect regarding dissemination of information. That is, if this level of contact between nodes is low, information (e.g., Sim and Bet values, contact lists) will take longer to be updated and diffused. Another issue regarding performance is the contact time, which can have a strong influence especially in scenarios where contacts are short lived.

Another proposal known as *Bubble Rap* [8] also uses node centrality along with the concept of community structure to perform forwarding. With *Bubble Rap,* nodes are grouped based on social parameters (i.e., number of contacts and contact duration) and have a local/global popularity index (obtained from betweenness centrality). Messages within the same community are forwarded using local popularity whereas messages traversing different communities use a combination of local/global popularity to reach the final destination. In the second case, whenever the message is forwarded to a member of the destination's community, the current carrier deletes it from the buffer to prevent further dissemination. The algorithm employed in this proposal is rather simple. If a source node wishes to send a message, all it needs to do is to check if the community of the destination node is the same as its own. If so, for every encountered intermediate node, it compares their local ranks and generates a copy to the encountered node that has higher rank value. Otherwise, it passes this message copy to an encountered node belonging to the same community as the destination or having higher global rank value.

Bubble Rap considers that nodes belong to different size communities and that such nodes have different levels of popularity (i.e., rank). With this, it can mimic human relationship allowing it to achieve good overall performance regarding delivery success rate with acceptable delivery cost. The performance is even better as the number of different communities and message TTL increase showing its capability to deal with human social behaviour. However, reaching a destination in a different community is quite exhaustive, especially if the source node has the lowest rank and its community has a high number of nodes. This will incur undesirable replication within the source community. Moreover, centrality can result in overloading (e.g., processing, energy, buffer) popular nodes since they are outnumbered compared to the number of global nodes. In addition, it does not always mean that a high centrality node has the best contact probability with the destination community.

Differently from previous proposals, *SocialCast* [9] shows that forwarding can be achieved not only based on the social ties and mobility patterns, but also considering the interests of destinations. The proposed routing protocol determines a utility function based on the predicted node's co-location (i.e., probability of nodes being co-located with others sharing the same interest) and change in degree of connectivity (i.e., related to mobility and representing changes in neighbor sets), which is used to calculate how good data carrier a node can be. This proposal is based on the publish-subscribe paradigm, that is, nodes publish content on the network that is received by nodes according to their subscribed interests. Nodes get copies if they have higher utility regarding a given interest than the node currently carrying messages with content matching such interest. The proposal comprises three phases: i) *interest dissemination*, in which each node broadcasts, to its first-hop neighbors, its list of interests along with its updated utilities regarding its interests as well as the last received messages; ii) *carrier selection*, in which if the utility function of a neighbor node regarding a given interest is higher than the current carrier node, this neighbor is selected as the new carrier; and, iii) *message dissemination*, in which messages are replicated to interested nodes and/or passed to the new carrier.

SocialCast allows messages to reach their destinations with a very low number of replications (i.e., reduced resource consumption) and stable latency. The result is good delivery ratio with low TTL values as messages are delivered within few hops. However, the co-location assumption (i.e., nodes with same interests spend quite some time together) may not always be true [10]. Since such assumption is of great importance, the proposal is compromised in scenarios where it does not always apply.

Also considering node mobility and social interaction, *PeopleRank* [10] makes use of stable social information between nodes to decide on forwarding. As its own name suggests, *PeopleRank* sets ranks to nodes according to their social interaction, and use this ranking to decide on the next hop for data exchange as it is known that socially well-connected nodes become the best forwarders for message delivery. This ranking process is analogous to Google's page rank system in which the relative importance of a Web page is determined according to its links to/from a set of pages.

Thus, nodes are ranked according to their position in the social graph, i.e., considering their linkage to other important nodes in the network. In order to determine its rank, a node needs to be acquainted to its socially connected neighbors and their respective ranks. So, when two nodes meet, they exchange their ranks and neighbor sets, update their own ranks, and exchange messages according to the new determined ranking. The node with the highest rank gets the messages.

The more social information is available to nodes, the better is the overall performance regarding delivery success rate. *PeopleRank* is also able to keep the cost associated to message delivery very low and with short delays. However, it is proven that considering only socially connected nodes is not enough to guarantee a good performance level, since socially disconnected nodes are also able to forward messages and could be considered to improve performance.

In summary, proposals that take into consideration the idea of community perform quite well [14] when compared to algorithms simply based on history of encounters, encounter prediction, and message prioritization. However, socially-aware proposals do not consider the age of contacts when computing the centrality levels of nodes (i.e., centrality change according to the node interaction in society) and the communities are statically defined (i.e., communities change according to the social behaviour of nodes). Also, most of these proposals may still lead to a flooded network, in the sense that at the end of the testing period every node in the network has at least a copy of all messages that were generated. In addition, Hossmann et al. (2010) [32] show that, if the contact aggregation considered for determining the social graph is based on time window, some social metrics (e.g., betweenness centrality, similarity) can lead to node homogeneity regarding the given metric. That is, as network lifetime increases, these nodes may have the same characteristic (i.e., popularity) which will result in great impact on the performance of forwarding algorithms. This suggests that social-based algorithms must be carefully designed in order to not end up becoming a mere random-based solution.

In the next section, we present the existing taxonomies regarding DTN routing, and provide a taxonomy aiming to facilitate the identification of existing and yet-to-come proposals.

IV. Taxonomy for Opportunistic Routing

The analysis of opportunistic routing approaches shows the existence of different trends based on distinct goals. On the one hand, we have single-copy forwarding approaches aiming to optimize the utilization of network resources. On the other hand, replication-based approaches aim to optimize delivery probability. Forwarding has the advantage of using network resources properly, but may end up taking too long to delivery the messages. Replication-based approaches present a message delivery probability that is, in almost every solution, very close to optimal, but with a high cost. Aiming to achieve a good balance between a high delivery probability and a low utilization of network resources, several proposals try to avoid flooding the network, by exploiting mobility of nodes, history of encounters, and social parameters.

Upon what is available in terms of routing solutions for DTNs, there is the need to properly classify them within a suitable taxonomy in order to support an objective evaluation analysis. There is a set of prior work that attempted to categorize opportunistic routing algorithms. However, they consider aspects (e.g., level of knowledge) that led to an unbalanced classification, assigning most solutions to a few set of categories, or to very specific classification branches (e.g., by considering for instance information coding or methods to control movement of nodes). Added to that, the lack of a taxonomy able to reflect the overall performance of the network led to group proposals without a suitable way of validate them. The result is that authors are faced with no guidelines to select the most suitable set of proposals that could be used to realistically evaluate the performance of their own algorithm. Consequently, there are solutions being evaluated based on generic benchmarks (e.g., always considering epidemic routing) that are not the most suitable ones to prove, without margin for doubts, that such proposals are the ones with better performance (i.e., good balance between probability of delivery and resource utilization) among the proposals that have similar performance goals.

Hence, this section aims to analyse existing routing taxonomies regarding opportunistic routing, as well as to propose a routing taxonomy suitable for an easy classification and evaluation of current and future proposals.

A. *Existing taxonomies*

The first attempt to create a taxonomy for opportunistic routing was proposed by Jain et al. (2004) [18] based on three types of classification. The most important one is the first type of classification, which divides opportunistic routing according to the knowledge about the network that nodes need to have to perform packet forwarding. The second and third classifications follow a trend already used to classify other type of routing: the second approach classifies routing as proactive (i.e., route computation happens prior to traffic arrival) and reactive (i.e., route computation takes place upon the need for sending data); the third approach classifies routing as source-based (i.e., the complete route is determined by the source), or hop-based (i.e., the next hop is determined in every traversed hop).

In what concerns the knowledge-based taxonomy, knowledge about the network is provided by centralized *oracles*. Four different oracles are proposed: i) *Contact Summary Oracle,* which provides summarized information about contacts (e.g., average waiting time until next contact, average number of contacts); ii) *Contacts Oracle*, which provides more detailed information related to contacts between nodes at any point in time (e.g., number of contacts in a given period); iii) *Queuing Oracle*, which provides information about buffer utilization at any time; iv) *Traffic Demand Oracle,* which provides information about present or future traffic demand. According to the authors, the more knowledge a routing solution can get, the better its performance will be, which leads to increasing unrealistic approaches since such ubiquitous knowledge is impossible to get in a dynamic network.

The knowledge levels used by each oracle can be: zero, where solutions use no information about the network to perform routing; partial, where solutions can route by using only the *Contact Summary* oracle, or one/both of the *Contacts* and *Queuing* oracles; and complete, where all oracles (*Contacts*, *Queuing*, and *Traffic Demand*) are considered. Fig. 2 illustrates all the different levels of knowledge used in this taxonomy as well as the existing oracles, and the relationship between performance and knowledge.

In realistic scenarios existing routing solutions are, however, classified as having zero or partial knowledge. This is due to the nature of DTNs where network topology is not known beforehand, which makes it difficult to have a central entity (e.g., oracle) providing information to aid routing decisions. Not to mention the delay incurred to gather/process such information, which can become quite unfeasible in scenarios with short-lived contacts between nodes. Regarding the time and place for routing decisions, most of the solutions can be seen as hop-by-hop reactive approaches.

In our opinion, a taxonomy based on network knowledge, as well as proactive/reactive and source/hop characteristics is not the most suitable one to provide a balanced classification of opportunistic routing approaches.

The most complete taxonomy for opportunistic routing, to our knowledge, was provided by Zhang (2006) [11]. This taxonomy follows the work presented by Jain et al., classifying protocols according to the information they require from network as well as relevant routing strategies (e.g., pure forwarding,

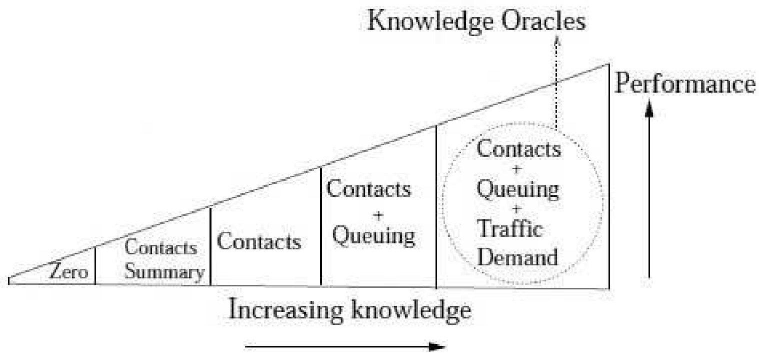

Figure 2. Knowledge oracles (Jain et al. [18]).

estimation of forwarding probability). The result is a taxonomy based on two categories, namely deterministic and stochastic routing. In the former case, node movement and future connections are known beforehand (i.e., nodes are completely aware about topology), which follows the oracle-based taxonomy proposed by Jain et al. In the case of stochastic approaches, the behaviour of nodes and network is random and unknown. Hence, routing decisions depend upon local conditions leading to simple solutions where messages are replicated on every contact up to more complex solutions where the use of history of encounters, node mobility patterns, and message coding are considered for routing decisions. Zhang also classifies proposals regarding whether (or not) node movement can be controlled.

We can say that Zhang's proposal complements the proposal of Jain et al. by including a set of more realistic (stochastic) approaches. However, Zhang's proposal emphasizes aspects that are orthogonal to different routing categories (i.e., coding methods) and puts emphasis on categories that are specific to deterministic networks (e.g., node movement control), which do not represent a scenario where opportunistic routing may have greater impact in our daily life, as explained in section II-B.

Balasubramanian et al. (2007) [28] propose a taxonomy based on two types of classification criteria regarding the routing strategy and the effect on performance metrics. Based on the first criterion, solutions are divided into two routing strategies: replication-based, where packets are replicated and then transferred to the next hop; and, ii) forwarding-based, where only one copy of the message traverses the network. The second classification criterion is used to divide solutions based on the effect routing decisions have on performance metrics. There are two types of effect: i) incidental, where the effect of decisions do not take into consideration resource constraints; and, ii) intentional, which determines the effect of a routing strategy on metrics considering such constraints.

We believe that classifying routing solutions according to their routing strategy is the most suitable approach. However, there is more to it than what is presented by Balasubramanian et al. since there are other performance metrics, besides incidental/intentional usage of resources, that can be used to distinguish among solutions that follow the same generic strategy (i.e., replication).

The same idea of classifying solutions based on their routing strategy (i.e., forwarding/replication) is also followed by Song and Kotz (2007) [22] and Nelson et al. (2009) [7]. In relation to the work presented by Balasubramanian et al., the classification presented by Song and Kotz is able to further divide replication-based proposals taking into consideration, not only the effect that they have on network resource consumption, but also on delivery probability. Still in what concerns the classification of replication-based approaches, Nelson et al. propose to divide them into flooding-based and quota-based. This classification is quite important since its shows that it is not the fact that messages are duplicated that may lead to a flooding situation, as the one that occurs with epidemic routing, but the employed routing metric for deciding on replication. Nelson et al. shows that there are probabilistic-based replication strategies that actually end up flooding the network (flooding-based), since at the end of the experimental period every node has at least one replica of each message, while others (quota-based) have higher success in controlling the number of replicas in the network. However, this taxonomy is incomplete in the sense that it does not consider routing categories based on metrics such as encounter number, resource usage, or social similarities.

In another survey [33], D'Souza and Jose (2010) classify routing solutions into three major categories: i) flooding-based, where nodes flood the network to increase delivery probability or apply some measures to control such flooding by bounding the number of messages copies to be distributed in the network and by embedding additional information to messages blocks to reduce flooding effects; ii) history-based, in which the history of encounters between nodes is taken into account to improve routing decisions; and, iii) special devices-based, where stationary or mobile devices are used to improve communication among the communicating nodes. Such special devices can also consider social interaction among nodes to perform routing decisions. Like Zhang, D'Souza and Jose consider aspects that are orthogonal (e.g., network and erasure coding) and could be easily applied to other categories. Although this taxonomy proposal succeeds in including the new trend (i.e., social aspects) observed in the last three years, it includes this trend under a category that does not comply with the regular behaviour found in DTNs. The reason for that is that distinguishing proposals according to whether or not they use special stationary/mobile devices to improve data exchange is not realistic as the network/nodes will have to present a deterministic behaviour in order to correctly place these devices in the system.

The most recent classification proposed by Spyropoulos et al. (2010) [34] groups opportunistic routing proposals according to message exchange scheme they employ: forwarding (only one message copy traverse the network), replication (message is replicated in different levels ranging from every node getting a copy up to more elaborate solutions based on utility functions), and coding (where message can be coded and processed at the source or as it travels throughout the network). The authors also identify the different types of utility functions that can be applied to either forwarding or replication message schemes. Such functions are categorized according to their dependency on the destination (i.e., destination dependent/independent). Additionally, they classify DTNs according to characteristics that have major impact on routing such as connectivity, mobility, node resources, and application requirements. The authors succeed in mapping the routing solutions to the different types of DTNs. Still, the proposed opportunistic routing classification considers categories that can be orthogonal (i.e., coding) and does not include the social similarity trend that emerged in the last years. They simply refer to the social aspects as a mere destination-dependent function in which we believe comprises a new research direction that includes social relationships, interests, and popularity to achieve suitable delivery probability with shorter delay and cost.

Fig. 3 summarizes the analysed taxonomies. Such figure shows an evolution towards stochastic approaches that do not require any knowledge about the global network topology. Deterministic approaches, based on some kind of centralized oracles, are not realistic and do not contribute to a taxonomy aiming to support an objective evaluation of current and future proposals. Within stochastic approaches, and since 2007, there is a clear trend to classify routing strategies based upon their success in achieving a good balance between delivery probability (e.g., message replication) and usage of network resources (e.g., forwarding).

Keeping in mind the ultimate goal of balancing performance and resource usage, current taxonomies are not capable of illustrating the different families of routing metrics that have been used to devise replication-based approaches able to avoid network flooding. The next section presents a proposal to extend the taxonomy (presented by Nelson et al.) that most reflects the behaviour of DTNs with a set of categories that represent recent trends in stochastic opportunistic routing.

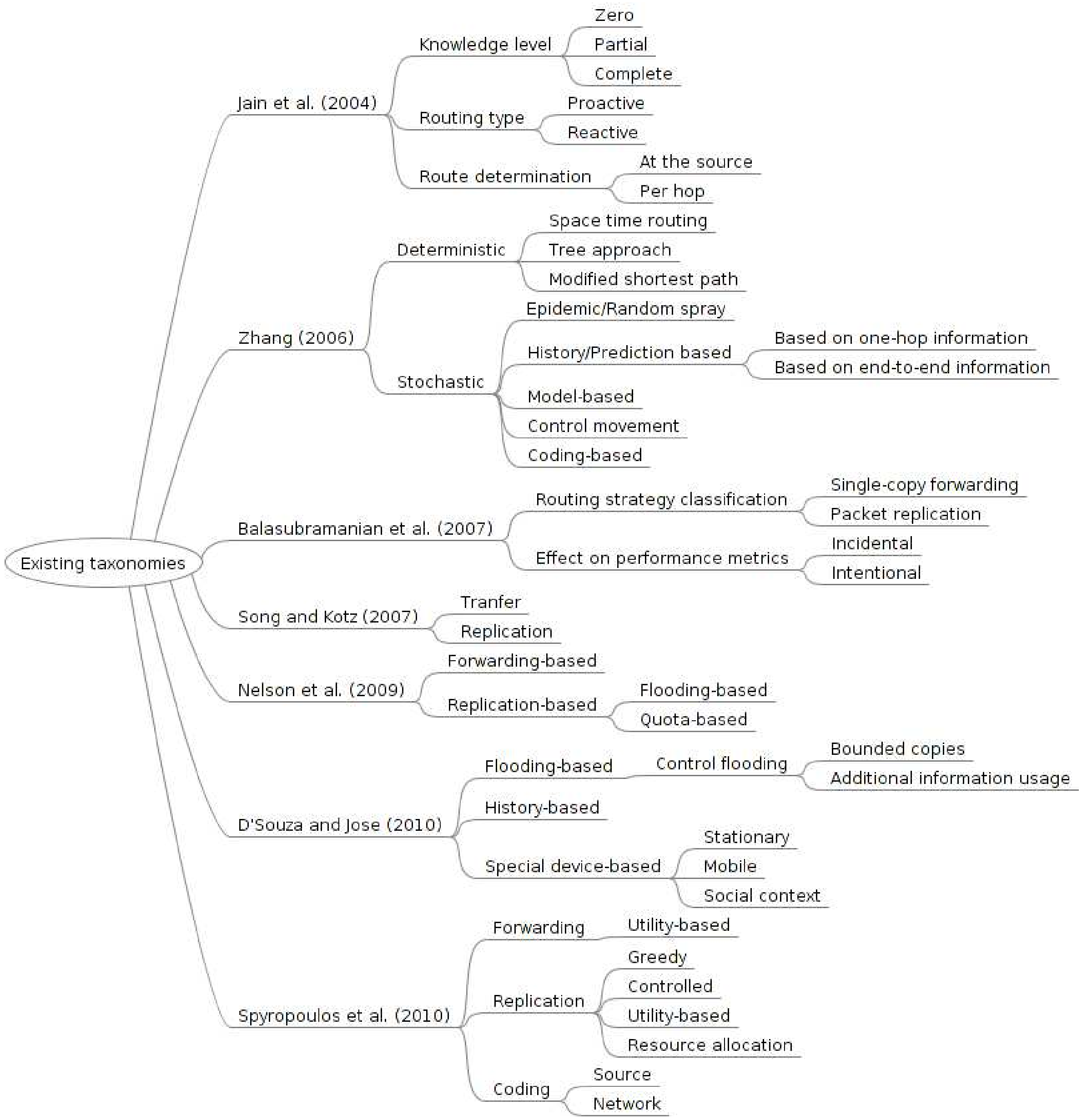

Figure 3. Existing taxonomies.

B. Proposed Taxonomy

One can conclude that the existing taxonomies generally focus on the analysis of opportunistic routing proposals based on their efficiency (e.g., level of knowledge employed to achieve higher delivery rates [18], forwarding schemes that result in different performance levels [11], [28], [22], limiting the number of messages copies in the network to spare resources [7], [33]), but they lack an analysis of the characteristics of the graph structure used.

We believe that focusing on an analysis of the topological features (e.g., contact frequency and age, resource utilization, community formation, common interests, node popularity) assumed by each proposal may lead to a more stable taxonomy useful to study real-world networks such as computer networks operated based on social behaviour.

Thus, in this section, we propose a taxonomy based upon the analysis of recent opportunistic routing proposals (cf. section III). The goal is to devise a homogeneous taxonomy able to classify available

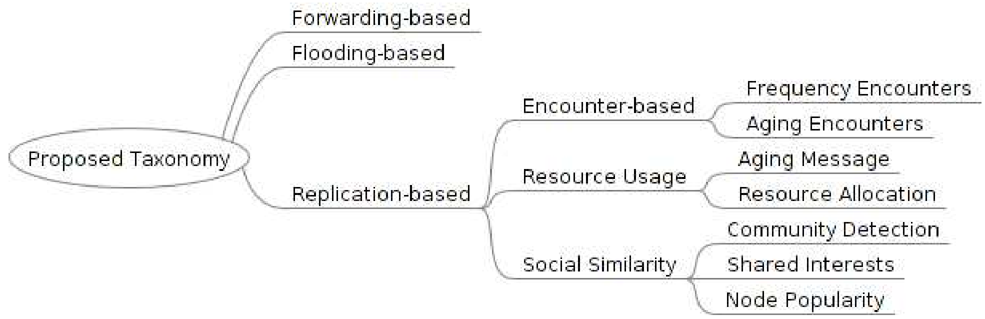

Figure 4. Taxonomy for opportunistic routing in DTNs.

proposals and identify future trends in what concerns opportunistic routing. Moreover, it is expected that the proposed taxonomy may support an objective evaluation of future routing proposals, by providing a realistic representation of the similarity between different routing families, aiming this way to avoid comparison between proposals that have a completely different routing strategy and set of metrics. Social similarity is an example of recent metrics that have clearly created a new trend in the investigation of opportunistic routing [29], [30], [9], [8], [10]. The reason for this is that social behaviour takes into account human relationship characteristics such as contacts with other people, time spent with these people, the level of relationship between people, among others. And, since computing devices are carried by humans, social-based forwarding decisions can consider people's socially meaningful relationships, where the relevant information come from aspects such as human mobility, interaction and social structures. This information can be used to perform forwarding, because the topology created from human social behaviour varies less than the one based on mobility and thus such solutions deserve being categorized.

Fig. 4 illustrates the proposed taxonomy that complements the most recent trend (replication vs. forwarding) with the analysis done of nineteen proposals published between 2000 and 2010.

It is important to mention that all studied opportunistic routing proposals take advantage of node mobility to forward data ahead, where some of them (e.g., *Epidemic* [3], *Direct Transmission* [20]) are rather simple and use the resulting contacts to reach the destination, while others are more elaborate and consider social aspects in order to find the destination (e.g., *LABEL* [29], *PeopleRank* [10]). This is the reason why we do not devote a specific category (as *Model-* or *Control Movement-based* in Zhang [11] and *Mobile Device-based* in D'Souza and Jose [33]) since this is an inherent feature of opportunistic proposals and we look for features that help us understand their graph structure.

Thus, the proposed taxonomy is based on an initial classification of all proposals as forwarding-, flooding-, or replication-based. The forwarding-based category is also known as single-copy forwarding (e.g., *MEED*, and approaches in Spyropoulos et al. [20]) since all approaches propose that only one copy of each message traverses the network towards the destination. From the resource consumption viewpoint, this category of approaches is quite interesting since it keeps network (e.g., bandwidth) and node (e.g., buffer space) resources usage at a low level; however, all approaches suffers in general from high delay rates that, consequently, results in a low delivery rate.

Nelson et al. propose, depending on the level of duplication, that algorithms within the replication-based approach be divided into flooding-based and quota-based. The flooding-based algorithms are able to increase delivery rate to a very high level, whereas the quota-based algorithms, in general, allow a more wise usage of resources, resulting in low delay and reduced flooding overhead since they tend to spread less copies of messages in the network.

We do consider these different levels of replication but unlike Nelson et al., we propose that flooding-based algorithms be classified out of the replication branch. First, because we only consider proposals that allow every node to spread a copy of each message to every other node that they meet (e.g., *Epidemic*). And, also due to the fact that having (or not) the quota-based feature (i.e., where the number of created copies does not depend on the number of network nodes) can be found in the different algorithms identified throughout our analysis (cf. section III-C).

Despite being an aggressive approach, the flooding-based strategy is able to increase delivery rate, but at the same time leads to a high consumption of resources. Such waste of resources can be avoided with algorithms that try to somehow control flooding. This control starts by limiting the number of copies injected in the network, if it is able to avoid nodes ending up with a replica of every created message, the algorithm has the quota-based feature.

So, the replication-based approaches have as common goal an attempt to increase the delivery rate by sending several copies of the initial message through different nodes to quickly reach the destination before message expiration time. Since these approaches consider different routing algorithms and metrics, we propose the following sub-categories.

The first sub-category is the encounter-based, where nodes choose next hops based either on frequency encounters (e.g., *PROPHET*, *MaxProp*, and *Prediction,* and *EBR*), or aging encounters (e.g., *FRESH*, *EASE*, and *Spray and Focus*). In the former, proposals consider the history of encounters with a specific destination to support opportunistic forwarding of messages or the frequency nodes met in the past to predict future encounters. As for the latter, proposals consider the time elapsed since the last encounter with the destination to decide about next hops.

Resource usage is the second sub-category, in which decisions are made considering the age of messages (e.g., *Spray and Wait*, and *OPF*) or knowledge about local resources (e.g., *PREP*, and *RAPID*). Aging messages proposals have in common the fact that they aim to avoid messages to be kept being forwarded in the network by creating metrics that define the age of message copies. As for the resource allocation proposals, they take forwarding decisions that wisely use available resources.

The last sub-category is related to social similarity, where proposals start to follow more complex algorithms aiming first at avoiding flooding with high probability, and exploiting social behaviour. We propose to divide social similarity algorithms into: community detection, shared interest, and node popularity.

Community detection approaches (e.g., *SimBet*, *Label, BubbleRap*) rely on the creation of node communities taking into consideration people social relationships translated to contact numbers and duration of contact among nodes. These approaches suffer with the overhead of community formation. The shared interest approach (e.g., *SocialCast*) relies on the assumption that nodes with the same interest as the destination of the message are good forwarders since they have high probability to meet. But this assumption may not always be true [10], since a node with similar interest to a given group may not even come in contact with this group of nodes. Still within the social similarity category, there are approaches that are based only on a process of ranking people in terms of their popularity without a straight dependency upon neither the computation of communities nor the synchronization of interests. Node popularity approaches (e.g., *PeopleRank*) make use of social information to generate ranks to nodes based on their position on a social graph, using such ranking to decide on the next hop for data exchange. Although social similarity algorithms provide stable graphs, it is proven that relying on socially connected nodes may not be enough to guarantee a good performance, which can be improved with the inclusion of some degree of randomness in the forwarding decision [10].

It is easily observed that all categories of our taxonomy present advantages and disadvantages. However, our goal was not to identify the winner category, but to: i) devise a homogeneous taxonomy able to classify available proposals and identify future trends within opportunistic routing context; and ii) support an objective evaluation of future routing proposals, by providing a clear identification of the similarity between different routing families.

V. Universal Evaluation Framework

Even with a suitable taxonomy, there is the need to identify a *Universal Evaluation Framework* (UEF) to provide a way to fairly evaluate opportunistic routing proposals based upon realistic scenarios (i.e., able to represent the real requirements of DTN applications) and based on common metrics such as delivery rate and cost, delay, and energy efficiency. The motivation to devise a UEF is that current opportunistic routing proposals, no matter if they are based on forwarding or replication, or if they consider levels of social relationship or not, do not always consider neither a similar set of performance metrics nor comparable experimental scenarios.

We start by providing a brief overview of the most recent evaluation framework available in the literature. Then, to devise an efficient UEF, we analyse in this section sixteen opportunistic routing proposals plus the *Epidemic* one in what concerns evaluation benchmarks, metrics and scenarios. The goal of having a UEF is to avoid the creation of future proposals comprising irrelevant performance metrics, evaluated in specific scenarios and without proper benchmarks. In other words, we present an evaluation framework that spans: i) network type, size, and resources; ii) number of nodes, the relationship between them, their interest, available resources, and willingness to participate in communication; and iii) performance metrics, such as delivery rate, delivery cost, delivery delay, goodput, message expiration, and energy efficiency.

A. Existing Evaluation Framework

Regarding evaluation frameworks, we highlight the most recent proposal based on *Evolving Graph* (EG) theory to design/evaluate least cost routing protocols. EG provides a formal abstraction for dynamic networks and reflects the different connectivity graphs in the time domain by considering node mobility. The result is that connectivity of links are transcribed into subgraphs for different instant in time.

Thus, Ferreira et al. (2010) [35] take into consideration one of the formalized EG criteria (i.e., foremost) to determine journeys (i.e., future temporary connections between nodes that can form a path over time) in which data can quickly reach its destination.

This evaluation framework provides designers with an algorithm that is able to reach good performance in scenarios where connectivity patterns are known beforehand. Additionally, the algorithm can be used as lowerbound reference to compare opportunistic routing solutions.

The work of Spyropoulos et al. (2010) [34] also provides principles to help developers in designing routing solutions based on their classification of opportunistic routing, identified utility functions, and DTNs characteristics. The authors show that by knowing the application characteristics and requirements, the choice/design of routing solutions is eased.

Still, both works lack a guideline of how performance metrics and experimental setups can be used. Thus, our goal is to provide a set of experimental setups to aid designers in fairly assessing the performance of existing and their yet-to-come opportunistic routing solutions in comparison studies.

B. Benchmarks

To start with, Fig. 5 shows the relationship among the seventeen routing proposals (in rectangulars) analysed in what concerns their evaluation benchmarks. With this first analysis, we aim to identify evaluation similarities among these proposals. The arrows leaving rectangular, representing a proposal, indicate other proposals that were used as benchmark to evaluate it. In Fig. 5, we can see that *EBR* and *RAPID* are the proposals that were evaluated against the highest number of related work, five (e.g., *EBR* used *Epidemic*, *MaxProp*, *Spray and Wait*, *Spray and Focus*, and *PROPHET*) and three (*RAPID* used *MaxProp*, *Spray and Wait*, and *PROPHET*) proposals, respectively. The number of incoming arrows reflects the importance that a proposal has as a benchmark to other related proposals. It is clear

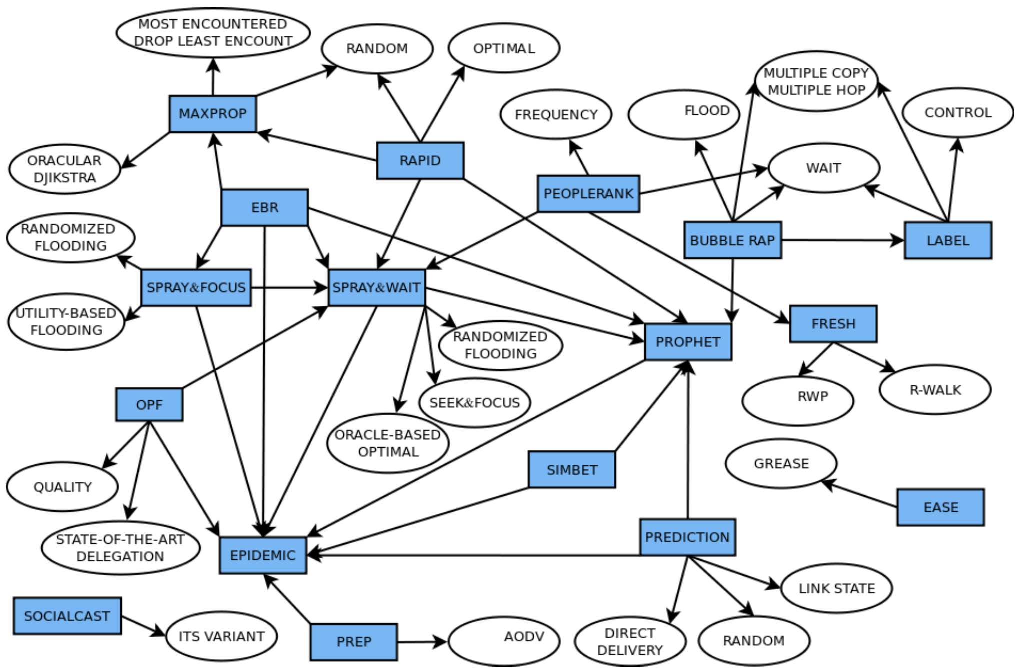

Figure 5. Analysis of most common benchmarks.

that *Epidemic* (with eight incoming arrows), *PROPHET* (with six incoming arrows) and *Spray and Wait* (with five incoming arrows) are the proposals that have been used most often as benchmark.

What is important to extract from Fig. 5 is that there are no rules for comparing proposals whereas authors always rely on proposals that have clearly worse performance in relation to their own, given the considered conditions. This is the first motivation to propose UEF, since we believe it is imperative to be able to assess the performance through a fair comparison among related work.

Fig. 5 also shows several other methods (in ellipses) that are used to evaluate most of the probabilistic proposals. These methods, which were not analysed in this technical report, include extensions of the proposed algorithms (e.g., *GREASE* in the *EASE* proposal). From the evaluation methods, the most used ones follow a random approach where either a message copy is created based on some forwarding probability p [4], [25], [22] or messages are randomly distributed upon an opportunity [6], [28].

It is worth mentioning that from all the analysed proposals, only *MaxProp*, *Label*, and *FRESH* are not compared against any other related work. Instead, they are either evaluated against other methods, as it is the case of *MaxProp* and *Label*, or based on scenarios with different mobility models, namely Random Waypoint and Random Walk (i.e., *FRESH*). Other isolated cases are *SocialCast* and *EASE*, which are compared only to a variant of themselves.

There are several proposals that rely on some kind of optimization algorithms for their evaluation: *Spray and Wait* (*Oracle-based Optimal*), *MaxProp* (*Oracular Dijkstra*), and *RAPID* (*Optimal*). These optimization methods differ in what concerns the used control information and techniques. The former case includes future movements (*Oracle-based Optimal*) and future meetings (*Oracular Dijkstra*/*Optimal*), while the techniques include a central entity that holds required knowledge (*Oracle-based Optimal*), adaptation of the shortest-path Dijkstra (*Oracular Dijkstra*), or devising the approach as Integer Linear Program optimization problem (*Optimal*).

Tables I, II, and III summarize Fig. 5 in what concerns the proposals that belong to the most relevant branches of the proposed taxonomy (all except forwarding- and flooding-based). The goal is to analyse how solid is the evaluation of a proposal in what concerns the most relevant related work (their own category) and the most used benchmark. For the sake of simplicity, Tables I, II, and III do not show the proposals that did not consider related work for their evaluation (i.e., *FRESH, SocialCast, EASE, MaxProp,* and *Label*). The analysed proposals are illustrated in each column, together with the indication of their taxonomy category.

Table I
TAXONOMY-BENCHMARK RELATIONSHIP FOR ENCOUNTER-BASED APPROACHES.

	Encounter-based			
	Frequency Encounters			Aging Encounters
	PROPHET	PREDICTION	EBR	SPRAY&FOCUS
EPIDEMIC	X	X	X	X
FRESH				
PROPHET		X	X	
SPRAY&WAIT			X	X
MAXPROP			X	
LABEL				
SPRAY&FOCUS			X	

Table II
TAXONOMY-BENCHMARK RELATIONSHIP FOR RESOURCE USAGE APPROACHES.

	Resource Usage			
	Aging Messages		Resource Allocation	
	SPRAY&WAIT	OPF	PREP	RAPID
EPIDEMIC	X	X	X	
FRESH				
PROPHET	X			X
SPRAY&WAIT		X		X
MAXPROP				X
LABEL				
SPRAY&FOCUS				

Table III
TAXONOMY-BENCHMARK RELATIONSHIP FOR SOCIAL SIMILARITY APPROACHES.

	Social Similarity		
	Community Detection		Node Popularity
	SIMBET	BUBBLE RAP	PEOPLERANK
EPIDEMIC	X		
FRESH			X
PROPHET	X	X	
SPRAY&WAIT			X
MAXPROP			
LABEL		X	
SPRAY&FOCUS			

From Tables I and II, we can see that the proposals that were compared to the highest number of related work were *EBR* and *RAPID*, compared to five and three other proposals, respectively. For instance, *EBR* is an example of a proposal with a solid evaluation benchmark (cf. Table I), since it was compared to two out of three other related work in its category (*PROPHET,* and *MaxProp*), and to two of the three most used benchmarks (*Epidemic,* and *Spray and Wait*).

From the proposed taxonomy and considering Tables I, II, and III, we can observe that proposals belonging to a specific category (e.g., Flooding-based: *Epidemic*) are used for comparison purposes with proposals belonging to more "intelligent" categories. It is important to point out that some proposals (e.g., *PREP*) consider *Epidemic* just as an upperbound for representing the best performance values in terms of delivery rate. However, in most cases (e.g., *Prediction*, *EBR*), such comparison is highly unfair since *Epidemic* ends up being evaluated under different assumptions and conditions (e.g., different number of nodes, message size, buffer size) from those used in its design, and based on different performance metrics (e.g., composite metrics in *EBR*).

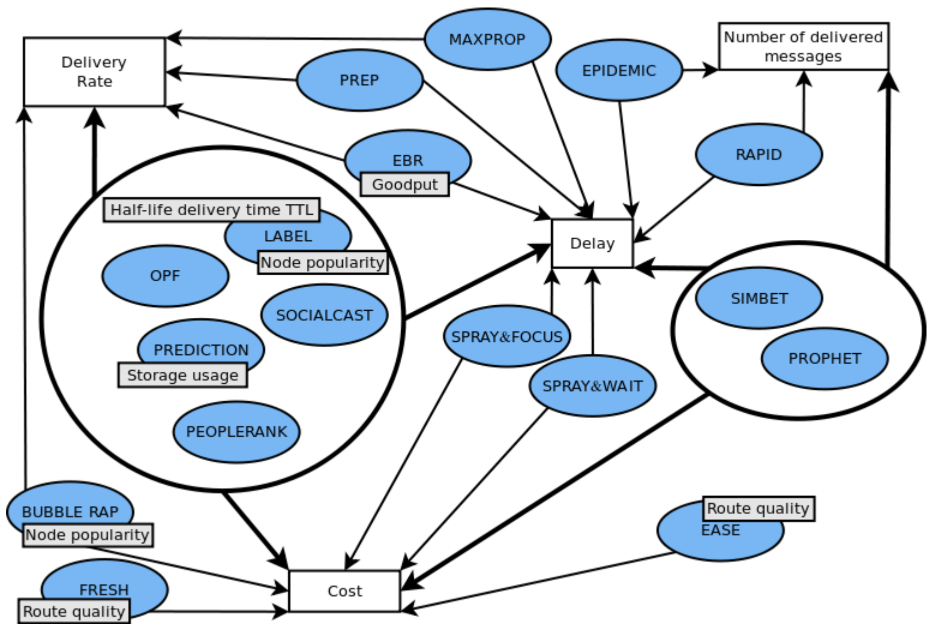

Figure 6. Analysis of performance metrics.

C. Performance Metrics

The analysis of the benchmark strategies used to evaluate the different routing proposals reveals a lack of consistency and fairness in the usage of evaluation scenarios and metrics. This evidence leads us to our second motivation for proposing a UEF: proposals must be evaluated considering similar performance metrics and the same evaluation conditions in order to fairly assess their performance.

Hence, we analyse the performance metrics that were considered for the evaluation of the classified proposals (cf. Fig. 6). For the sake of simplicity, we use general terms (i.e., delivery rate, cost, delay, number of delivered messages) to identify the most used performance evaluation metrics (wide rectangulars) since terminology varies according to the proposal. Our analysis also identifies other metrics that are relevant for specific proposals (narrow rectangulars). The identified metrics are generic in the sense that different proposals may use slightly different instantiation of them. For instance, in what concerns the delay metric, *RAPID* works with the worst-case delay while *Epidemic* routing works with message delivery latency. It is important to mention that to simplify Fig. 6, some proposals have been grouped (e.g., *SimBet*, and *PROPHET*) as they consider the same set of metrics.

Based on our analysis, we can conclude that the most important metrics for opportunistic routing are: delivery rate, cost, delay and number of message delivered. In what concerns the delivery rate, we define it as the *number of messages that have been delivered per unit of time, out of the total number of messages created*. This metric is used in nine of the seventeen proposals under different terminology such as: number of messages that have been delivered out of the total of messages created (*Label*), the fraction of sourced bundles that are delivered to the destination (*PREP*), the ratio of the number of messages delivered to the number of total messages generated (*Prediction*), the proportion of messages that have been delivered out of the total unique messages created (*Bubble Rap*), the ratio between the actual number of messages delivered to the interested subscribers and the ideal one (*SocialCast*), the message delivery ratio (*EBR*), simply delivery rate (*MaxProp*, *OPF*), or the success rate of the algorithm normalized by the success rate of flooding within a delay period (*PeopleRank*).

The delivery rate is a very important metric to be considered since it represents the effectiveness of the proposal. Normally, proposals are classified as having good performance if they achieve high delivery rates, i.e., can delivery as much messages as possible to destinations in a useful time frame.

However, the effectiveness of a proposal has to be weighted against the cost associated to the message delivering process. In our analysis, we define the cost of opportunistic routing as the *the number of replicas per delivered message*. Each byte can correspond to data (including message copies) or control information, and its transmission has an energy and processing cost. This metric is used in twelve of the seventeen analysed proposals under different instantiations, such as: the number of forwarded messages (*PROPHET*, *SocialCast*), the number of transmissions per delivered message (*Spray and Wait*), the number of transmissions (*Spray and Focus*), the total number of messages transmitted across the air (*Label*, *Bubble Rap*), the total number of messages transmitted during the simulation across all nodes (*Prediction*), the total number of meta-data units transmitted during the simulation across all nodes (*Prediction*), the number of times a message copy occurred due to replication (*Prediction*), the total number of forwards (*SimBet*), the number of forwardings (*OPF*), or the number of retransmissions (*PeopleRank*).

Depending on the approach, the cost metric can also be associated to the overhead necessary to build a route from a source to a destination (*search cost* in *FRESH*), the total number of transmissions required to forward/search in order to deliver a message (*relative cost of routes* in *EASE*), as well as the distribution of the number of hops needed for all the deliveries (*hop distribution* in *Label*, *Bubble Rap*), and the average number of hops per message (*SimBet*).

At a first glimpse looking at delays on scenarios that are susceptible to high delays may sound contradictory. Nevertheless, each message has a time-to-live that is correlated to its utility. Moreover, it is important to remove already delivered messages as soon as possible from the network in order to avoid waste of resources. Thus, making sure that messages reach their destination within a useful time frame is also important from the viewpoint of the performance. Based on Fig. 6, it is clear that most of the routing proposals consider delay as a performance metric, more precisely fourteen out of seventeen proposals. In our analysis, we define the delay factor of opportunistic routing as the *time require to deliver all the bytes encompassing a message*. Delay is also seen differently depending on the proposal, and it can be considered as the message delivery latency (*Epidemic*), the message delivery delay (*PROPHET*, *Spray and Wait*), the latency of delivered packets (*MaxProp*), the delay distribution (*Label*, *PeopleRank*), the delivery delay (*Spray and Focus*), the fraction of bundles that are delivered within a given delay bound (*delay CDF* in *PREP*), the average delay and worst-case delay (*RAPID*), the duration between the generation time and delivery time of a message (*Prediction*), the average end-to-end delay (*SimBet*), the average latency (*SocialCast*), the average delay (*EBR*), or simply delay (*OPF*).

In summary, we can conclude that guaranteeing a low cost to maximize delivery rate with low delay seems to be highly desired for opportunistic routing solutions, since this guarantees that end-users have access to a significant amount of useful information with a good usage of network resources. This is reflected in the number of proposals that consider delivery rate (nine), cost (twelve) and delay (fourteen) as performance metrics, as seen in Fig. 7.

Besides these three metrics, a small number of proposals (four) also consider the *number of delivered messages*, as a performance metric under different terminology such as: percentage of delivered messages (*Epidemic*); number of received messages (*PROPHET*); number of packets delivered before a deadline (*RAPID*); total number of messages delivered (*SimBet*). However, it is our opinion that the number of delivered messages is directly related to the delivery rate and so should not be considered as major evaluation metric.

Also in Fig. 7, we can observe five other metrics that are only used to evaluate specific proposals: route quality used by *FRESH* and *EASE*; node popularity used by *Label* and *Bubble Rap*; half-life delivery TTL used by *Label*; storage usage used by *Prediction*; and, goodput used by *EBR*. The route quality metric is defined as the difference between the route found by *FRESH/EASE* and the route with shortest number of hops. The popularity metric is the number of contacts between a node and the others in the network. The half-life delivery TTL metric is the TTL value that would allow half of

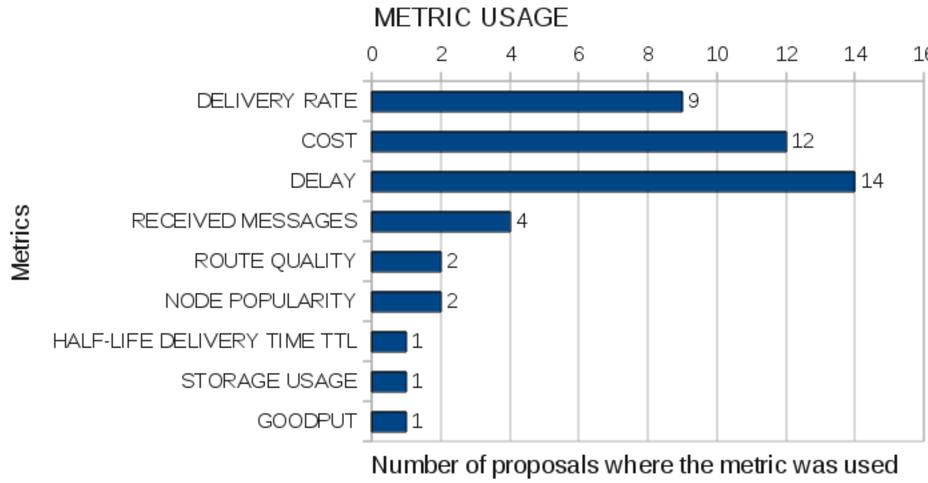

Figure 7. Metrics usage for performance evaluation.

the messages to be delivered. The storage usage metric is the maximum storage (in bytes) used across all nodes, and the goodput metric is the number of messages delivered divided by the total number of messages transferred (including those transfers that did not result in a delivery).

From the five metrics that are specific to some proposals, we believe that *node popularity*, *half-life delivery TTL*, and *goodput* are the most interesting ones since they can identify nodes connecting different node clusters, and guarantee message delivery within a useful time frame.

As a last remark, we can say that it is easily observed the lack of a convention regarding the terminology used for performance evaluation metrics and also that proposals are seen from different perspectives regarding such metrics. This leads to a difficult evaluation process and provides no fairness when comparing proposals from different categories.

D. Evaluation Scenario

The lack of convention observed from the analysis of performance metrics is also evident in what concerns the used evaluation scenarios, which brings us to our third motivation to propose the creation of an UEF: the definition of a set of guidelines to be considered when creating experimental scenarios.

Tables IV and V summarize what we could observe from the proposals under study regarding the evaluation conditions. It highlights the most common aspects that were taken into account by each of the proposals for performance assessment, such as the number of nodes, the number of source and destination pairs, meeting time (i.e., contact time) and time between such meetings (i.e., inter-meeting time), area size (also referred to as network density, and expressed in number of nodes in the surface area), message size, network load (i.e., number of generated messages), message TTL (expressed either in hops or time units), size of node buffer, mobility model and node speed, transmission range, and beacon usage.

Tables IV and V show that proposals are very different in terms of the experimental scenario considered for performance assessment. It is also easily observed that some of the proposals that are mostly used as benchmark (e.g., *Epidemic*) provide detailed information about the evaluation setup, which is quite useful. But most important, although the aspects pointed out in this table are of great importance, the lack of consistency require some analysis to allow us to derive a default scenario set up that should be used by all proposals.

From the analysis concerning the evaluation scenarios, we identify two classes of parameters of great importance. The first one is related to the density of the network , including network area, number of

Table IV
NETWORK DENSITY PARAMETERS CONSIDERED BY THE DIFFERENT PROPOSALS.

PROPOSALS	AREA DENSITY (Km² / number of nodes)	MOBILITY MODEL /SPEED (m/s)	NODE MEETING/ INTER-MEETING TIME	TRANSMISSION RANGE (m)	BEACON CONTROL
EPIDEMIC	0.45/50	Node chooses a point and walks there (0 – 20)	Pairs of hosts come in contact periodically and randomly	10 to 250	Internet MANET Encapsulation Protocol
PROPHET	0.45/50 (RWP), and 4.5/50 (CM)	Random Waypoint Mobility (0 – 20), and Community (10 – 30)	Pairs of hosts come in contact periodically and randomly	50 and 100	
SPRAY&WAIT	0.25/100 (Traffic load) 0.04/200 (Connectivity)	Random Walk Mobility	Exponentially distributed meeting time	5 to 35	
SPRAY&FOCUS	0.04/100	Random Walk, Random Waypoint, and Community based Mobility		5 to 35	Nodes periodically transmit beacons to recognize each other's presence
PREP	9/25	Random Waypoint Mobility (5 – 15)		250	Hello protocol
RAPID	388/20 and 40		Meeting time distributions are exponentially distributed / Inter-meeting time between nodes follow either an exponential or a power law distribution		Scans for other buses 100 times a second
PREDICTION	0.81/5142	Mobility traces from CRAWDAD		Nodes could discover and connect each other instantly when they were associated with a same AP	
SIMBET	Lab and College area/100	Human traces		Bluetooth	
BUBBLE RAP	Infocom05/41 Hong-Kong/37 Cambridge/54 Infocom06/98 Reality/97	Human traces	Inter-meeting time follows a power-law distribution	Bluetooth	
EBR	15/26, 51, and 101 (VMM) 9/26, 51, and 101 (REDMM, and RWP)	Vehicular-based Map-driven (2.7 – 13.9) Role-based, Event-driven Disaster Mobility (1 – 20) Random WayPoint (0.5 – 1.5)		250 m	
OPF	Unknown/300 and 40	Human and Vehicular			
PEOPLERANK	MobiClique/27 SecondLife/150 Infocom/65, 47, and 62 Hope/414		Median Contact time: 90, 150, 180, and 240s Median Intercontact time (10, 15, 25, and 30mn)	Bluetooth	

nodes, mobility model, node meeting and inter-meeting times, transmission range and beacon control. The second class is related to traffic and encompasses distribution of sources and destinations, load generation, message size, message TTL, and buffer size.

In what concerns network density (cf. Table IV), this class allows designers to learn how good their proposals behave when nodes interact in sparse and dense scenarios. Scenarios with sparse networks are good to analyse if proposals are able to use sporadic node contact to guarantee the require probability of delivery, since delays are expected to be high in such scenarios. Dense networks are important to evaluate the capacity of using a large number of forwarding opportunities including the smart usage of randomness (Mtibaa et al. [10] have proven that some level of randomness by choosing different nodes other than the ones considered by the algorithm, is useful to achieve better performance results).

Network density may be configured by using three parameters: network area, number of nodes, and mobility models. Regarding the number of nodes, we can observe from Table IV that the trend is to consider a roughly average number between 100 and 150 nodes (excluding extreme cases such as *FRESH*, *EASE*, and *Prediction*). This number of nodes may be enough to configure dense or sparse scenarios depending on the considered network area, and node mobility model. The network area can span a conference building as well as a city area, so designers must consider these two extreme cases in order to better assess the quality of their proposals. By considering different areas, along with different mobility models, the challenge faced by algorithms increases since different levels of sparseness will emerge as simulations run. Since a normal assumption to have is that most of the mobile nodes are carried by people, experimental scenarios should consider realistic human mobility models. Random models may not be suitable, even though humans are very unpredictable when it comes to movements. Moreover, people are part of communities [14] that represent their social relationship with others, considering interests and tastes [9], [10]. Hence, considering the relation between mobility models and social interaction within a society seems to be a good method for assessing the performance of a given algorithm.

We believe that mobility model should also include variations in node speed and pause time in order to describe realistic movement situations. These parameters certainly influence the contact and inter-contact times of nodes. Different proposals [4], [29], [28], [8] have correctly assumed that contacts and inter-contact times are best described as exponentially and power law distributions. However, other proposals [6], [10] simply obtain the contact and inter-contact times from the considered datasets.

Another parameter that may have an impact on the density of the network is transmission coverage. Quite a few devices are already equipped with both Bluetooth and Wi-Fi cards, so transmission range must be looked upon. Based on the analysed proposals, the transmission range should be set to an average value of 100 m, assuming that devices work with Wi-Fi. While Wi-Fi direct is still not built into devices, we suggest a minimum transmission range of 10 m, which represents Bluetooth contact between mobile nodes, and a maximum transmission range of 250 m in cases where fixed entities (i.e., access point) may be used for information relay.

Lastly, we observed how proposals use beacons, which are useful to find out more information about potential new neighbors. However, beacons may not be always present since nodes are mobile, and so may go to sleep mode quite often in order to spare battery. This has some pros and cons. Some of the pros relate to the fact that battery lifetime can last longer. However, this may result in the loss of good contact opportunities. According to the studied proposals, considering beaconing every 100 ms should be enough to achieve a good balance between battery lifetime and network knowledge. However, further investigation is required to validate this assumption.

In what concerns traffic (cf. Table V), one of the first consideration is about the number of data sources and destinations. From the considered proposals, we can observe that in almost all cases, these node pairs are randomly chosen, and comprise a subset of the total number of nodes. We believe that source/destination pairs must remain the same throughout the comparison process, as it is highly expected to offer the same conditions to all the proposals under evaluation. In the analysed proposals, sources are configured with a load generation algorithm, which is the setup aspect that brings more variation to experiments. There are proposals that generate a message per second [3], [5], a number of packets per hour [6], a number of messages uniformly distributed [29], [8], a given number of bytes/messages per second/minute per source [27], [7], as well as proposals that give little [25] or no information about the load used in the network. Load is indeed a setup parameter that must be carefully addressed, which is not the case with current approaches. For instance, the used load distribution may be related to the mobility model.

Another setup parameter related to traffic is message size, which plays an important role when measuring the consumption of network (e.g., bandwidth) and node (e.g., buffer, and power) resources since the size of messages may be different (depending on the applications generating them). Also, depending on the average contact duration (especially in highly mobile scenarios) and message size, data exchange may not even happen in each contact. Especially if the proposal requires exchange of metadata information [28], [27] prior to exchanging real data, in which case nodes may waste a portion of a potentially short contact time.

Just a few proposals explicitly mention message size, which can be of 1 KB [3], [27], [28], [22] or

Table V
TRAFFIC PARAMETERS CONSIDERED BY THE DIFFERENT PROPOSALS.

PROPOSALS	SOURCE/DEST DISTRIBUTION	# GENERATED MESSAGES (LOAD)	MESSAGE SIZE (KB)	MESSAGE TTL	BUFFER SIZE
EPIDEMIC	45 / 44	one message per second	1	1, 2, 3, 4, and 8 hops	10, 20, 50, 100, 200, 500, 1000, and 2000 messages
PROPHET	45 / 44 (Random) and 2 / 1 (Comm.)	one message/sec (RWP), and 20mgs/sec – 2mgs/5sec (CM)		3 and 11 hops	200 messages
SPRAY&WAIT	1 / 1	Node generate a new message with an inter-arrival time distribution uniform.		4000 – 6000 unit times	
SPRAY&FOCUS	1 / 1	Moderate number of CBR traffic sessions		1000 – 10000 time units	
PREP	1 / 1	40 to 200 bytes/sec/node	1		1 to 6MB
RAPID	1 / All	4 packets/hour/node	1		100KB (Power law) 40GB (Trace driven)
PREDICTION	1 / 1	After each contact event in the contact trace, a message is generated with a given probability	0.08 to 1	Unlimited	Unlimited
SIMBET	1 / 1	A single message is generated between each node included in the subset			
BUBBLE RAP	1 / 1	1000 messages			
EBR		1, 2, and 4 messages/minute/ Source	25		1MB
OPF	1 / 20 (NUS) and 1 / 1 (UmassDieselNet)			3 and 1 to 5	
PEOPLERANK					

vary between 10 to 100 KB [6], [7]. Based on the considered proposals, we suggest a variation of the message size from 1 to 100 KB in order to provide a more realistic evaluation. For instance, people may write short messages/emails while on the move, so the variation of the size of messages easily represents the applications and time employed when using portable devices on the go.

Another aspect that affects traffic levels in the network is message TTL, which can be represented in number of hops [3], [5], [26] or time units [4], [25]. If messages have high values of TTL, they may end up increasing network/node resource consumption. On the other hand, if messages have a small TTL, they may not even reach their destination. From the considered proposals, we observe that TTL normally varies between 3 to 10 hops in average. The minimum value is justified by a study made by Hui et al. [8] showing that only 5% of nodes have some level of relationship with the destination in the first hop. However, the interaction values improve (around 35%) for more than 3 hops.

Buffer usage can also influence the performance of a given proposal. Considering unlimited buffer space is not realistic at all since we cannot assume that users are willing to share all the storage room with other users. Of course, this may be different in scenarios where nodes are only there to serve others [28]. But, in general, buffer space should be limited (varying according to the device) to a size of around 200 messages (considering 10 KB messages), based on the analysed proposals.

Based on the analysis provided in this section, it is easily seen that not only are solutions evaluated against others that belong to different categories but also consider different performance evaluation metrics as well as different conditions to decide on performance. This means that the way proposals are evaluated has an important impact on how they are classified as having or not satisfactory performance.

VI. Open Research directions in Opportunistic routing

From the analysed proposals, it is clear that one recent trend for opportunistic routing solutions is going towards the awareness about social parameters, related to inter-meeting times or related to more complex social similarities. However, we can see that many aspects must still be investigated to achieve maximized delivery rate with reduced delivery cost and delay.

We have seen that social ties can devise different communities [14] and these communities can indeed aid in data forwarding. Forwarding can reach even better results if interests and tastes [9] of users are taken into account and if the social relationship of nodes is explicit [10]. However, how should such communities be formed? In the analysed proposals, either a simple node gathering algorithm is applied (e.g., *SocialCast*) or a more detailed approach is done by observing the number of contacts and their duration to determine the social ties between nodes with the goal of creating communities (e.g., *Bubble Rap*). These are good approaches, but since we are dealing with humans, the structure of these communities can change over time, and so far routing proposals work under the strong assumption that communities are static. In addition, it is still not clear what is the best number of contacts and their duration in order for a node to be part of a community. Another question about social-aware proposals is if shall we simply consider common interest and use this commonality to group nodes together aiming to find good forwarders?

Besides the trend about social parameters, it was proven that, even if social connection is fully available, performance does not approach optimal values [10]. This means that nodes with which we do not have much contact can still be a good relay of information. Hence, there is an open issue to set up a good threshold between randomized and social-based forwardings.

Another open question has to do with the infrastructure. Currently, all proposals consider isolated networks, but in several environments, such as a city, it would bring extra benefits to devise routing algorithms able to make use of infrastructure when it is available, in order to increase the probability to reach a good performance level.

Regarding nodes themselves, should we consider that they are always willing to share their constrained resources on behalf of others? A scalable solution may be to consider that willingness of users may be implicitly obtained by the communities they are part of. However, this approach may end up not respecting users' bounds and wishes.

Can physical medium features be taken into account? We know that wireless medium is highly unstable and many factors (e.g., physical obstacles, weather) can have significant impact. So, instead of only considering the interests and social ties, we could also consider the link characteristics between nodes and this way avoid data being sent over low quality links. This would certainly improve performance in terms of delivery delay.

From what we have learned about the existing proposals, it will be very relevant the exploitation of users' movements that are tightly related with their social relationships, interests, as well as link quality characteristics. Such relationships can be used to identify communities and define trust level among users. Community formation must happen on-the-fly and must change as the users interact. This community approach will aid in forwarding data to nodes belonging to a community with similar interests or that of a specific destination. Trust can be applied especially to identify malicious nodes nearby (e.g., DoS attack) and avoid such nodes for data exchange.

Fig. 8 shows an example of what we think an opportunistic routing should be able to do, namely avoid insecure areas as well as nodes with low resources, and to exploit mobility of nodes taking into account their willingness to forward information, intrinsic trust levels and social similarities.

Finally, it is clear from the analysed proposals that most of the solutions aim to deliver data from a source to a destination. However, this point-to-point model is not the most useful one in a DTN. It is our opinion that solutions for opportunistic routing should allow the dissemination of information from diverse sources to a set of interested people, similar to what is the context of *SocialCast* [9].

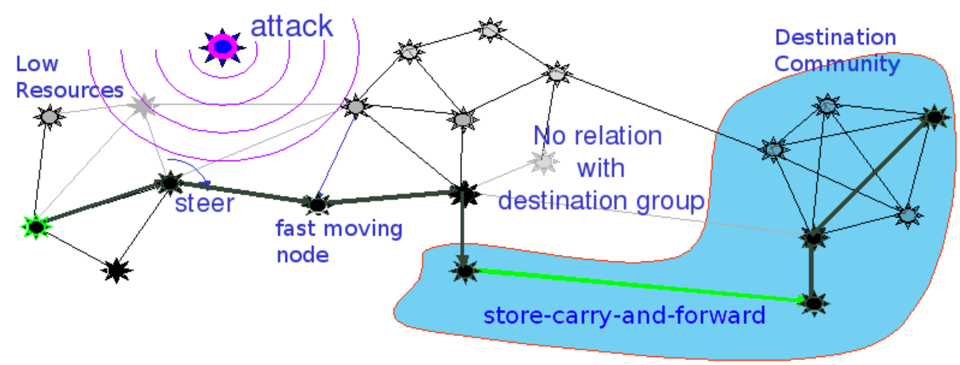

Figure 8. Generic characteristics of opportunistic routing.

VII. Summary and Conclusions

A lot of research has been done in order to provide suitable routing to networks with link and node intermittent behaviour. The result was the investigation of a large number of solutions aiming to solve routing problems in specific scenarios. Nevertheless, we observe a rich collection of mechanisms that still leave some performance gaps (e.g., community adaptability) open. So, our goal is to provide a survey of the most relevant solutions devised in the last ten years.

We start this technical report by presenting the two possible main scenarios which opportunistic routing proposals may face, namely sparse and dense networks. According to their characteristics, each scenario imposes a level of challenge regarding routing solutions being the dense scenario the most challenging one (and the focus of this technical report) since nodes present a very dynamic behaviour that may result in infrequent connections and additionally the quality of such connections may be endangered by high wireless interference levels. After identifying the main scenarios, we analysed nineteen proposals, and classified them according to the way they deal with message replication. We divide them into three major groups: the single-copy routing, where no replication is taken into account and only one copy of the message traverses the network, thus sparing network and node resources; the epidemic routing, which replication is performed upon each contact in order to increase delivery rate; and, the probabilistic-based routing, in which proposals perform a wiser replication in attempt to balance resource consumption and delivery rate.

The trend observed in proposals that span the last ten years is towards a probabilistic approach where mobility is exploited taking advantage of stochastic encounters between nodes. In addition, social parameters (e.g., relationships, interests) have been used in the last three years to aid in data forwarding since topologies based on social aspects tend to be less volatile (i.e., vary less) than the ones based purely on mobility patterns.

It is clear that these proposals are somehow related, thus we analysed the existing taxonomies aiming to find commonalities that would allow us to devise an objective comparison. We found out that most of the existing taxonomies simply classify proposals either as forwarding- or replication-based giving emphasis to network capacity optimization, and flooding avoidance. Furthermore, features (e.g., coding methods) that we believe are orthogonal (i.e., can be applied in different categories), were considered as categories. Additionally, new proposals were not comprised by these taxonomies (e.g., *Bubble Rap*, *SocialCast*, *PeopleRank*). So, we propose a new taxonomy based on nineteen proposals, spanning the last ten years, including proposals in identified new trends (i.e., the employment of social parameters).

Another issue we notice regards the way these proposals were being evaluated. Performance evaluation considered different aspects (e.g., node number, mobility models) preventing a more reliable evaluation. This is even more evident in those proposals that were compared to “better” ones in scenarios totally different from the ones they have been designed for. With that, we concluded that a Universal Evaluation Framework (UEF) must be created to allow a more consistent evaluation process, especially when comparisons between older and new approaches take place. In what concerns a UEF, our contribution includes a close analysis of many experimental aspects of the proposed solutions according to two classes of parameters that have a great impact in the performance of the overall system, namely network density and traffic. The stochastic analysis of such parameters used by different proposals helped us to identify a set of network (e.g., area, nodes, mobility) and traffic (e.g., load, message size, buffer) parameters that must be considered by default for future performance evaluation. Besides identifying such parameters, our contribution also includes suggestions of values for such parameters.

To conclude, we provide some considerations of what we think are the most important research issues to be addressed in the future when it comes to opportunistic routing. These considerations especially highlight topics related to the most recent proposals including the exploitation of social similarities.

ACKNOWLEDGMENT

Thanks are due to FCT for financial support and for PhD grant (SFRH/BD/62761/2009) to Waldir Moreira.

REFERENCES

[1] F. Warthman, Delay-tolerant networks (dtns): A tutorial, Warthman Associates. Version 1.1, May, 2003.
[2] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, H. Weiss, Delay tolerant network architecture, IETF Network Working Group. RFC 4838, 2007.
[3] A. Vahdat, D. Becker, Epidemic routing for partially connected ad hoc networks, Tech. Rep. CS-200006, Duke University (2000).
[4] T. Spyropoulos, K. Psounis, C. S. Raghavendra, Spray and wait: an efficient routing scheme for intermittently connected mobile networks, in: Proceedings of ACM SIGCOMM WDTN, Philadelphia, USA, August, 2005.
[5] A. Lindgren, A. Doria, O. Schelén, Probabilistic routing in intermittently connected networks, SIGMOBILE Mob. Comput. Commun. Rev. 7 (3) (2003) 19–20.
[6] J. Burgess, B. Gallagher, D. Jensen, B. N. Levine, Maxprop: Routing for vehicle-based disruption-tolerant networks, in: Proceedings of INFOCOM, Barcelona, Spain, April, 2006.
[7] S. Nelson, M. Bakht, R. Kravets, Encounter-based routing in DTNs, in: Proceedings of INFOCOM, Rio de Janeiro, Brazil, April, 2009.
[8] P. Hui, J. Crowcroft, E. Yoneki, Bubble rap: social-based forwarding in delay tolerant networks, in: Proceedings of ACM MobiHoc, Hong Kong, China, May, 2008.
[9] P. Costa, C. Mascolo, M. Musolesi, G. P. Picco, Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks, Selected Areas in Communications, IEEE Journal on 26 (5) (2008) 748–760.
[10] A. Mtibaa, M. May, M. Ammar, C. Diot, Peoplerank: Combining social and contact information for opportunistic forwarding, in: Proceedings of INFOCOM, San Diego, USA, March, 2010.
[11] Z. Zhang, Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: Overview and challenges, IEEE Communications Surveys and Tutorials 8 (1-4) (2006) 24–37.
[12] M. Musolesi, C. Mascolo, A community based mobility model for ad hoc network research, in: Proceedings of ACM MobiHoc REALMAN, Florence, Italy, May, 2006.
[13] W. J. Hsu, T. Spyropoulos, K. Psounis, A. Helmy, Modeling spatial and temporal dependencies of user mobility in wireless mobile networks, IEEE/ACM Trans. Netw. 17 (2009) 1564–1577.

[14] P. Hui, People are the network: experimental design and evaluation of social-based forwarding algorithms, Tech. Rep. UCAM-CL-TR-713, University of Cambridge, Computer Laboratory (2008).
[15] 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, May, 2004.
[16] R. C. Shah, S. Roy, S. Jain, W. Brunette, Data mules: Modeling a three-tier architecture for sparse sensor networks, in: Proceedings of IEEE ICC SNPA, Anchorage, USA, May, 2003.
[17] L. R. Ford, Jr., D. R. Fulkerson, Flows in Networks, Princeton University Press, 1962.
[18] S. Jain, K. Fall, R. Patra, Routing in a delay tolerant network, in: Proceedings of the ACM SIGCOMM, Portland, USA, August, 2004.
[19] E. P. C. Jones, L. Li, P. A. S. Ward, Practical routing in delay-tolerant networks, in: Proceedings of the ACM SIGCOMM WDTN, Philadelphia, USA, August, 2005.
[20] T. Spyropoulos, K. Psounis, C. S. Raghavendra, Efficient routing in intermittently connected mobile networks: the single-copy case, IEEE/ACM Trans. Netw. 16 (1) (2008) 63–76.
[21] B. H. Bloom, Space/time trade-offs in hash coding with allowable errors, Commun. ACM 13 (7) (1970) 422–426.
[22] L. Song, D. F. Kotz, Evaluating opportunistic routing protocols with large realistic contact traces, in: Proceedings of ACM MobiCom CHANTS, Montreal, Canada, September, 2007.
[23] H. Dubois-Ferriere, M. Grossglauser, M. Vetterli, Age matters: efficient route discovery in mobile ad hoc networks using encounter ages, in: Proceedings of ACM MobiHoc, Annapolis, USA, June, 2003.
[24] M. Grossglauser, M. Vetterli, Locating Mobile Nodes with EASE: Learning Efficient Routes from Encounter Histories Alone, IEEE/ACM Transactions on Networking 14 (3) (2006) 457–469.
[25] T. Spyropoulos, K. Psounis, C. S. Raghavendra, Spray and focus: Efficient mobility-assisted routing for heterogeneous and correlated mobility, in: Proceedings of IEEE PERCOM Workshops, White Plains, USA, March, 2007.
[26] C. Liu, J. Wu, An optimal probabilistic forwarding protocolin delay tolerant networks, in: Proceedings of ACM MobiHoc, New Orleans, USA, May, 2009.
[27] R. Ramanathan, R. Hansen, P. Basu, R. Rosales-Hain, R. Krishnan, Prioritized epidemic routing for opportunistic networks, in: Proceedings of ACM MobiSys MobiOpp, San Juan, Puerto Rico, June, 2007.
[28] A. Balasubramanian, B. Levine, A. Venkataramani, Dtn routing as a resource allocation problem, in: Proceedings of ACM SIGCOMM, Kyoto, Japan, August, 2007.
[29] P. Hui, J. Crowcroft, How small labels create big improvements, in: Proceedings of IEEE PERCOM Workshops, White Plains, USA, March, 2007.
[30] E. M. Daly, M. Haahr, Social network analysis for routing in disconnected delay-tolerant manets, in: Proceedings of ACM MobiHoc, Montreal, Canada, September, 2007.
[31] S. Basagni, I. Chlamtac, V. R. Syrotiuk, B. A. Woodward, A distance routing effect algorithm for mobility (dream), in: Proceedings of ACM/IEEE MobiCom, Dallas, USA, October, 1998.
[32] T. Hossmann, T. Spyropoulos, F. Legendre, Know thy neighbor: Towards optimal mapping of contacts to social graphs for dtn routing, in: Proceedings of IEEE INFOCOM, San Diego, USA, March, 2010.
[33] R. J. D'Souza, J. Jose, Routing approaches in delay tolerant networks: A survey, International Journal of Computer Applications 1 (17) (2010) 8–14, foundation of Computer Science.
[34] T. Spyropoulos, R. N. Rais, T. Turletti, K. Obraczka, A. Vasilakos, Routing for disruption tolerant networks: taxonomy and design, Wirel. Netw. 16 (8) (2010) 2349–2370.
[35] A. Ferreira, A. Goldman, J. Monteiro, Performance evaluation of routing protocols for manets with known connectivity patterns using evolving graphs, Wirel. Netw. 16 (2010) 627–640.