

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Data Forwarding in Opportunistic Networks

¹Lianggui Liu and ²Zhixin Chen

¹College of Informatics and Electronics, Zhejiang Sci-Tech University,
5 Second Avenue, Xiasha Higher Education Zone, Hangzhou, 310018, China

²College of Information, Central University of Finance and Economics, Beijing 100081, China

Abstract: In this study, state-of-art data forwarding methods in opportunistic networks are profoundly reviewed. Opportunistic networks is one of the most exciting evolutions of the legacy Mobile Ad hoc Networking (MANET) paradigm, in which the assumption of complete paths between data senders and receivers is not required all the time. Opportunistic networks enable users communication in disconnected environments, in which island of connected devices appear, disappear and reconfigure dynamically. As one key and principal problem in any networking, data forwarding and routing play an important role in improving network performance. As present routing protocols are mostly based on one primary assumption that end to end complete route should exist while data need to be transported between source-destination node pair, these protocols do not fit for opportunistic networks. New data forwarding and routing methods and protocols should be proposed for emerging opportunistic networks in order to meet the extensive pervasive networking need from the industry society. Furthermore, in the end of this study, new directions and trends in this area that are worth being studied further are pointed out.

Key words: Opportunistic networks, data forwarding, data routing, overview

INTRODUCTION

With the explosive deployment of mobile wireless devices recently, opportunistic networks (Lilien *et al.*, 2007) is becoming an increasingly popular area in networking research, in which the assumption of having end-to-end paths between the source and the destination is relaxed. Such networks fall into the fields of mobile ad hoc networking (MANET) and Delay-Tolerant Networking (DTN) (Jain *et al.*, 2004). Opportunistic networks enable user communication in an environment where disconnection and reconnection are common and link performance is extremely dynamic. They are very suitable to support the situation where network infrastructure has limited coverage and users have islands of connectivity. By taking advantage of device mobility, information can be stored and forwarded over a wireless link when connection opportunities arise (e.g., an appropriate network contact is met). In this view, traditional Internet connectivity can be considered as a special case of connection opportunity. With numerous emerging applications, opportunistic networks allow a huge number of devices to communicate end-to-end without requiring any pre-existing infrastructure and are very suitable to support pervasive networking scenarios. The network is thus extremely dynamic and is formed by

the evolving contacts among mobile devices and among connected clouds of devices. Opportunistic networks thus encompass the features and methods of delay or Disruption Tolerant Networks (DTN). They are very suitable to support the pervasive networking scenario, in which a huge number of devices carried by users and embedded in the environment communicate wirelessly without requiring any pre-existing infrastructure. By enabling end-to-end communication without requiring complete paths, opportunistic networks are much closer to real pervasive networking scenarios, with respect to the legacy MANET paradigm.

Figure 1 shows an example of opportunistic networks. Here in this network, source node S wants to send data to destination node D. Unfortunately there is no complete route between this source and destination pair, thus the data communication between this pair can not be performed if in a conventional mobile ad hoc network. But in an opportunistic network, the source can first send its data to his neighbor 1, then 1 will carry the data while moving until it meets another node 3 and relays this data to node 3. Node 3 will carry this data until it meets destination D. In the end the source node S can send data to the destination node D successfully even though there is no complete route between them all the time.

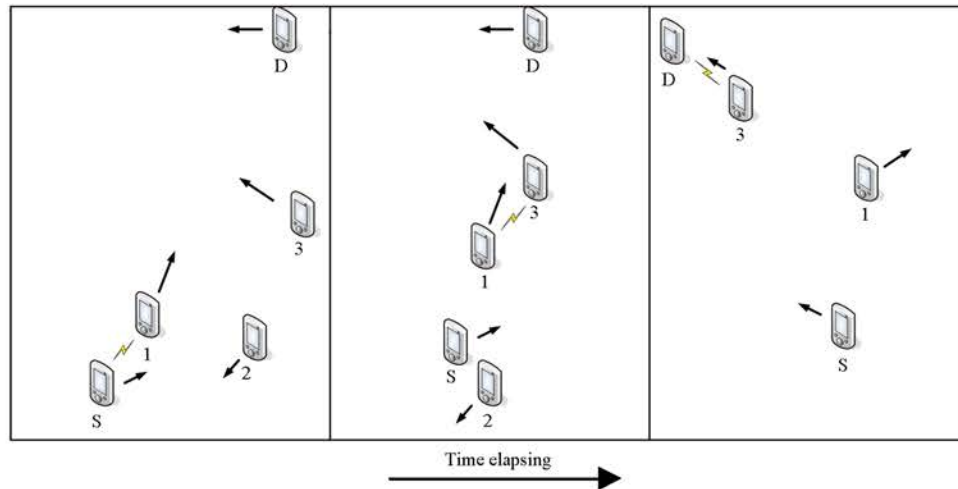


Fig. 1: Time evolving of opportunistic networks

THE RETROGRESSION OF ROUTING IN OPPORTUNISTIC NETWORKS

Data forwarding and routing is one key and principal technology which should be firstly considered while performing networking. Till now, many kinds of routing protocols have been proposed in mobile ad hoc networks, most of which are based on the consistent connectivity between source-destination pair (Liu and Feng, 2005a-c, 2007; Liu and Chen, 2009). Thus, as mentioned in the abstract of this study, most of them are not fit for opportunistic networks because of the intermittent connectivity in opportunistic networks which could result in packet discard. In opportunistic networks, packets will be buffered in present node when the next hop toward the packets' destination node does not exist and wait for the opportunity of network connectivity. Consequently, the selection of the optimal next hop toward the destination node and proper forwarding opportunity will be one of the most important issues which should be ended while designing effective opportunistic networks. Thus, from this viewpoint, routing based on network topology and information of link state in conventional networks will regress to the selection of the optimal next hop forwarding node.

DATA FORWARDING IN OPPORTUNISTIC NETWORKS

In opportunistic networks, network resources are constrained. For example, node depends on the battery power to perform its function, which results in energy constraint; nodes are usually subjected to their limited memory space; the frequency of nodes CPU are limited. On the other hand, nodes can deliver data only

by means of packet copy. If the number of copied packets can not be reduced, severe wireless channel collision will happen and network resources will be exhausted. Thus, the characters of opportunistic networks which include network resource constrained should be considered while designing low-cost data forwarding algorithm for opportunistic networks. Data forwarding is the most compelling challenge in opportunistic networks and the design of efficient data forwarding strategies for opportunistic networks is generally a complicated task due to the absence of knowledge about the network topological evolution. If the network topological evolution is deterministic and known, or at least predictable, data forwarding (when and where to forward data) can be scheduled beforehand so that some optimal objective can be reached. If the network topology evolution is stochastic, that is, the time-evolving topology is stochastic, all the nodes in opportunistic networks can do is to randomly forward data to their neighbors. In this study, we propose a profound review of the state of art on data forwarding algorithms and strategies in opportunistic networks. According to the fact that the information of network behavior can or can not be achieved or predicted ahead of time, we categorize them as follows:

- Deterministic data forwarding
- Randomly evolving data forwarding:
 - (a) Flooding based algorithms
 - (b) History information based algorithms
 - (c) Message-ferrying based algorithms
 - (d) Direct contact based algorithms
 - (e) Network coding based algorithms

Deterministic data forwarding: Here, we review the data forwarding algorithms and strategies where the future movement and connections, that is, network topology evolution can be known by the network node ahead of time. Furthermore, when and which node to forward can be predicted ahead and there should be an end-to-end path between source destination pair before data transmission even though this path will be time-variable.

In recent years, many research have been done in this area (Jain *et al.*, 2004; Handorean *et al.*, 2004; Huang *et al.*, 2007) (Table 1).

For DTN, several routing algorithms are proposed by Jain *et al.* (2004) based on the amount of knowledge about the network topology characteristics and traffic demand. Four knowledge oracles are defined; each oracle represents certain knowledge of the network. The Contacts Summary Oracle contains information about aggregate statistics of the contacts, which could result in time-invariant information. A contact is defined as an opportunity to send data. The Contacts Oracle contains information about contacts between two nodes at any point in time. This is equivalent to knowing the time-varying networks. The Queuing Oracle gives information about instantaneous buffer occupancies at any node at any time. The Traffic Demand Oracle contains information about the present or future traffic demand. Based on the assumption of which oracles are available, corresponding routing algorithms are proposed. For example, if all the oracles are known, a linear programming is formulated to find the best route. If only the Contacts Summary Oracle is available, Dijkstra with time-invariant edge costs based on average waiting time is used to find the best route. If only the Contact Oracle is available, modified Dijkstra with time-varying cost function based on waiting time is used to find the route. All of the algorithms developed (except the zero-knowledge case) mentioned by Jain *et al.* (2004) belong to deterministic ones.

Algorithms proposed by Handorean *et al.* (2004) selects the path of message delivering, depending on the available knowledge about the movement of nodes. Three cases are taken into consideration. In the first case it is assumed that global knowledge of the characteristic profiles with respect to space and time (that is, the

characteristic profiles of the movement and availability of the nodes as functions of time) are completely known by all the nodes. Paths are selected by building a tree first. Such an approach is referred to as the tree approach. Under the tree approach, a tree is built from the source by adding children nodes and the time associated with nodes. Each node records all the previous nodes the message has to travel and the earliest time to reach it. A final path can be selected from the tree by choosing the earliest time (or minimum hop) to reach the desired destination. In the second case, it is supposed that characteristic profiles are initially unknown to nodes. Nodes gain this information through learning the future by letting neighbor nodes exchange the characteristic profiles available between them. Paths are selected based on this partial knowledge. In the third case, to enhance the algorithm in the second case, it also requires nodes to record the past, that is, it stores the sequence of nodes a message has transited within the message itself.

Huang *et al.* (2007) present the characteristics of a vehicular ad hoc network (VANET), which is the Shanghai urban vehicular network (SUVnet). Authors there construct a mobility model using the GPS data collected from more than 4000 taxis in Shanghai. The model is both realistic and large scale. Based on this model, network topology and connectivity of SUVnet are studied. Because of the sparse distribution and dynamic topology of SUVnet, simply utilizing the conventional mobile ad hoc network routing protocols in SUVnet may not achieve a satisfactory performance. Therefore, they apply the delay tolerant network model to SUVnet and evaluate the epidemic routing protocols. They propose a new protocol based on GPS, Distance Aware Epidemic Routing (DAER), to improve the bundle delivery ratio.

As we have mentioned before, in deterministic case, an end-to-end path between source destination pair must exist before data transmission. From this viewpoint, the basic principle of these protocols in deterministic opportunistic networks is the same as that of conventional protocols in ad hoc networks. However, it is

Table 1: Comparison of deterministic data forwarding algorithms in opportunistic networks

Publishing year	Complexity	Applicability	Should node movement pattern be known beforehand?	Algorithm
2004	Integrate with existing different protocols according to different cases	Wide, can be fit for different cases	Yes, the network topology should be known at first to perform relative algorithm	Jain <i>et al.</i> (2004) (except the zero-knowledge case)
2004	Node exchange the characteristic profiles and compute the route	Not wider than case 2, 3	Yes, the global network information is completely known by all the nodes	Handorean <i>et al.</i> (2004) (case 1)
2004	Node exchange the characteristic profiles and compute the route (may be worse than the above)	Wide	No, nodes gain partial information by exchanging the characteristic between neighbor nodes once they meet	Handorean <i>et al.</i> (2004) (case 2, 3)
2007	Simple, integrate with existing epidemic routing protocols	Depend on GPS	Yes, with the help of GPS	Huang <i>et al.</i> (2007)

hard in opportunistic networks to guarantee that end-to-end path between source and destination pair exists all the time.

Randomly evolving data forwarding: Due to the randomly evolving of network topology and link state, the information of network behavior can not be achieved or predicted ahead of time. In this circumstance, data

forwarding algorithms should decide where and when to forward data messages. Representative works can be categorized as follows (Table 2).

Flooding based algorithms (Mundur *et al.*, 2008; Kang and Kim, 2008; Small and Haas, 2003).

Epidemic Routing (Mundur *et al.*, 2008) is one simple method to deliver packets reliably in intermittently connected networks. When each node meets other nodes,

Table 2: Comparison of randomly evolving data forwarding algorithms in opportunistic networks

Publishing year	Complexity	Should node movement or history information be used?	Should forwarding probability be calculated?	Is buffer limited?	Algorithm
2004	Complex	No	No	Yes	(Jain <i>et al.</i> , 2004) (the zero-knowledge case)
2009	Simple	No	No	No	Epidemic routing (Mundur <i>et al.</i> , 2008)
2008	More complex than Mundur <i>et al.</i> (2008)	Yes, nodes calculate their velocity and direction to move (vector) by using their local location information only	Yes	No	(Kang and Kim, 2008)
2003	Similar to Mundur <i>et al.</i> (2008)	No, but some fixed nodes will be needed	No	No	(Small and Haas, 2003)
2002	Simple	Yes and some fixed nodes will be needed	Yes	No	ZebraNet (Juang <i>et al.</i> , 2002)
2004	More complex than Mundur <i>et al.</i> (2008)	Yes, summary vectors and delivery predictability vector should be exchanged	Yes	No	PROPHET (Lindgren <i>et al.</i> , 2004)
2005	More complex than Mundur <i>et al.</i> (2008)	Yes, each node should produce its own delivery probabilities towards each known destination host	Yes	No	CAR (Musoiesi <i>et al.</i> , 2005)
2006	Complex, a high dimensional Euclidean space should be built at first.	Yes, requires the knowledge of all the nodes that are circulating in the space	Yes	Yes	(Leguay <i>et al.</i> , 2006)
2005	Complex, yet control packet overhead and buffer usage are less than Mundur <i>et al.</i> (2008)	Yes, the knowledge of relative velocities of a node and its neighboring nodes is required	Yes	Yes	(Lebrun <i>et al.</i> , 2005)
2005	Complex, similar to the work of Jeremie <i>et al.</i> (2006)	Yes, similar to Leguay <i>et al.</i> (2006)	Yes	Yes	(Leguay <i>et al.</i> , 2006)
2006	Complex	Yes, historical data is required	Yes	Yes	(Burgess <i>et al.</i> , 2006)
2003	Complex	Yes, node estimates the link forwarding probability based on history data	Yes	No	(Tan <i>et al.</i> , 2003)
2004	Complex, message ferry nodes should be elected at first	Yes, ferry nodes should move in two ways	No	No	Message-ferrying approach (Zhao <i>et al.</i> , 2004)
2005	More complex than Zhao <i>et al.</i> (2004), the possibility of interaction between ferries should be considered	Yes	No	No	(Zhao <i>et al.</i> , 2005)
2003	Simple	No	No	Yes, data is copied only once, However, it may incur long delays	(Shah <i>et al.</i> , 2003)
2004	Simple	No	No	Yes, similar to Shah <i>et al.</i> (2003)	(Spyropoulos <i>et al.</i> , 2004)
2005	Simple, integrated with probabilistic routing	No	No	Yes	(Widmer and Jean-Yves, 2006)
2005	Simple, only source node can flood data and relay nodes are designed to forward the data copy only to destination	No	No	Yes	(Wang <i>et al.</i> , 2005)

it infects them with its data. In other words, each node stores packets until it meets the destination node of packets and then it replicates the packets whenever they meet other nodes. It produces a lot of unnecessarily replicated packets, resulting in wasting network resources, which is considered as routing overhead.

A flooding-based Vector Routing protocol was introduced by Kang and Kim (2008) to reduce unnecessary replications without losing delivery ratio. In this approach, nodes calculate their velocity and direction to move (vector) by using their local location information only, not global network information, such as network topology, the location of other nodes and etc. This minimal information is utilized to perform efficient routing.

Small and Haas (2003) proposed the Shared Wireless Infostation Model (SWIM) to monitor the whales' life. Special tags applied to the whales perform periodic data monitoring. Data is replicated and distributed at each pairwise contact between whales and finally arrives at special fixed SWIM stations which are placed on buoys or mobile ones which are placed on seabirds. Hence, both whale-to-whale and whale-to-base-station communications are allowed. From the SWIM stations, data are eventually forwarded onshore for final processing. It is difficult to demonstrate the efficiency of the SWIM system on real whales due to circumstance constraint. However, simulation results are quite realistic since the simulation parameters about both the environment and the whales mobility model have been set according to the observations and studies conducted by biologists on whales real habits. The simulation results show a not negligible delay for arrival of data at the processing base stations. However, improvements are possible by increasing both the number of whales involved and the number of SWIM stations. Finally, mobile SWIM stations have shown better performance than fixed SWIM stations.

History information based algorithms: In ZebraNet (Juang *et al.*, 2002), wireless sensor nodes, namely collars (attached to Zebras), collect location data and opportunistically report their histories when they come in radio range of base stations, or the researchers or data collection objects, which periodically drive through (or fly-over) with receivers to collect data (Juang *et al.*, 2002; Lindgren *et al.*, 2004; Musolesi *et al.*, 2005; Leguay *et al.*, 2006; LeBrun *et al.*, 2005; Burgess *et al.*, 2006; Tan *et al.*, 2003). Collars operate on batteries with/without solar recharge. The goal is to study the animal behaviors through designing a collar and communication protocols that work on Zebras (high data collection rate). The authors study two routing protocols: flood-based routing

protocols and history-based protocols. In the flood-based routing protocol, data is flooded to their neighbors whenever they meet. It is expected that as nodes move extensively and meet a number of neighbors, given enough time, data will eventually arrive at the base station. In the history-based routing protocol, each node is assigned a likelihood of transferring data to the base station based on its past success. A higher value corresponds to a higher probability of eventually being within the range of the base station. Data is forwarded to its neighbor with the highest transferring probability. Experimental results indicate that the flood-based protocol yields higher system throughput if the buffer capacity at each node is large enough, but the energy consumed by the flood-based protocol can be eight times that of the history-based protocol. There is a tradeoff between throughput and energy consumption. Their conclusion is that while flooding makes sense at low-radio-range and low-connectivity points in the design space, it is not a good choice in a high-connectivity regime.

A probabilistic routing protocol, PROPHET (Probabilistic Routing Protocol using History of Encounters and Transitivity), is proposed by Lindgren *et al.* (2004). PROPHET first estimates a probabilistic metric called delivery predictability, $P(a, b)$, at every node a , for each known destination b , which indicates how likely it is that this node will be able to transmit a message to that destination. When two nodes meet, they exchange summary vectors and also a delivery predictability vector containing the delivery predictability information for destinations achieved by the nodes. The information in the summary vector is used to choose which messages to request from the other node. Simulation results show that for the given network, the improvement of packet delivery ratio under PROPHET over the epidemic routing can be up to about 40%.

In the Context-Aware Routing (CAR) protocol (Musolesi *et al.*, 2005), each node in the network is responsible for producing its own delivery probabilities towards each known destination host. Delivery probabilities are exchanged periodically between neighbor nodes so that, eventually, each node can compute the best carrier for each destination node. The best carriers are computed based on the nodes' context. The context attributes needed to elect the best carrier are, for example, the residual battery level, the rate of change of connectivity, the probability of being within reach of the destination and the degree of mobility. When the best carrier receives a message for forwarding, it stores it in a local buffer and eventually forwards it to the destination node when met or, alternatively, to another node with a higher delivery probability. The CAR provides a

framework for computing next hops in opportunistic networks based on the multiattribute utility theory applied to generic context attributes. The simulation results show that CAR is more scalable than epidemic routing while the protocol overhead being approximately constant regardless of the node buffer size.

In the study presented by Leguay *et al.* (2006), information of the nodes' mobility pattern is used for routing. The approach builds up a high dimensional Euclidean space, named MobySpace, where each axis represents a possible contact between a couple of nodes and the distance along an axis measures the probability of that contact to occur. Two nodes that have similar sets of contacts and that experience those contacts with similar frequencies, are close in the MobySpace. The best forwarding node selected for a message is the node that is as close as possible to the destination node in this space. Obviously, in the virtual contact space just described, the knowledge of all the axes of the space also requires the knowledge of all the nodes that are circulating in the space. This full knowledge, however, might not be required for successful routing.

LeBrun *et al.* (2005) proposed a method using the motion vector (MoVe) of mobile nodes to predict their future location. The MoVe scheme uses the knowledge of relative velocities of a node and its neighboring nodes to predict the closest distance between two nodes. After the nodes' future location are calculated, messages are passed nodes that are moving closer to the destination. As compared to epidemic routing, this approach has less control packet overhead and buffer usage.

Leguay *et al.* (2006) presented a strategy that uses a virtual coordinate routing called mobility pattern spaces (MobySpace). The measure of closeness represents the probability that the nodes will come into contact with each other. They showed that this approach consumes less resources than epidemic routing.

Burgess *et al.* (2006) proposed a protocol called MaxProp for effective routing of messages. A node uses MaxProp to schedule packets transmission to its peers and determines which packets should be deleted when buffer space is almost full. Packets are scheduled based on the path likelihoods to peers according to historical data. In addition, several complementary mechanisms, including acknowledgments, a head-start for new packets and lists of previous intermediaries are used in this approach. They showed that their approach performs better than the protocols that have access to an oracle (Jain *et al.*, 2004) that knows the schedule of meetings between peers.

Tan *et al.* (2003) proposed a Shortest Expected Path Routing (SEPR) similar to link-state routing to maintain a

topology map to each other. SEPR first estimates the link forwarding probability based on history data. When two nodes meet, they exchange the link probability update messages called Effective Path Length (EPL). A smaller EPL value suggests a higher probability of delivery. When a node received a smaller EPL, it will update its local EPL value. The EPL is also used in deciding which nodes to forward the messages. Using SEPR protocol, the same message could be forwarded to multiple nodes to increase reliability and to reduce delay.

Message-ferrying based algorithms: In the message-ferrying approach (Zhao *et al.*, 2004), extra mobile nodes are opportunistically used to offer a message relaying service. These nodes are named message ferries and move around in the network where they collect messages from source nodes (Zhao *et al.*, 2004, 2005). Message collection there can be conducted in two ways: (1) Node-initiated message ferrying: the ferry node moves around following a predefined and known path. Each node in the network has knowledge of the paths followed by active ferries and moves to meet ferries when it has data to deliver, (2) Ferry-initiated message ferrying: the ferry node, again, moves around following a predefined, default path. Any source node wishing to deliver messages sends a Service request to the ferry via a long-range radio signal, which also includes its current position. After having received the request from the source node, the ferry changes its trajectory to meet up with the source node.

Zhao *et al.* (2005) study the problem of using multiple ferries to deliver data in networks with stationary nodes and designing ferry routes to minimize average message delay. Multiple ferries offer the advantages of increasing system throughput while reducing message delay and robustness to ferry failures. On the other hand, because the possibility of interaction between ferries should be considered, the route design problem with multiple ferries is more complicated than the single ferry case. The authors present ferry route algorithms for single ferry and multiple ferries cases, respectively. In the single ferry case, solutions for the well-studied traveling salesman problem are adopted. In the multiple ferries case, algorithms assign nodes to specific ferries, synchronize among ferries and assign ferries to specific routes are discussed. Simulation results are utilized to evaluate the performance of route assignment algorithms, especially on the effect of the number of ferries on the average message delay. Numerical results indicate that when the traffic load is low, the improvement in delay due to the increased number of ferries is modest. This is because the delay is dominated by the distance between nodes. However,

when the traffic load is high, an increase in the number of ferries can significantly reduce the delay.

Direct contact based algorithms: In direct contact data forwarding approach, no replication of the original packets is performed and, instead, a source node holds packets until it meets its destination node and delivers the packets directly (Shah *et al.*, 2003; Spyropoulos *et al.*, 2004). Hence, unnecessary replication does not occur. However, it is still not guaranteed that a source node meets the destination node, due to node mobility.

Shah *et al.* (2003) proposed a three-tier architecture that connects spare sensors at the cost of high latency. At the top tier, there are access points or repositories that can be set at convenient places. The middle tier consists of DataMules that are mobile nodes whose mobility pattern is unknown and can communicate with sensors and access points. DataMules have large storage capacity and renewable power. While DataMules moving, they collect data from sensors and forward this data to the access points. The bottom tier consists of sensors that are randomly distributed across a region. In order to save energy, work performed by the sensors is reduced to minimal. DataMules are designed to be capable of short-range wireless communication and can exchange data from a nearby sensor access point they encounter as a result of their movements which are not predictable ahead of time. Thus, DataMules can pick up data from sensors when in close range, buffer it and drop off the data to wired access points when in proximity. The main advantage of the three-tier approach is the potential of large power savings by sensors because communication only takes place over a short range. Simple analytical models are presented to study the scaling of system characteristics as the system parameters, number of sensors, or number of DataMules change. Numerical results provide some relationship between the buffer requirements at the sensors (and at the DataMules) and the number of sensors (and the number of DataMules), respectively. It shows that the change in the buffer capacity on each sensor should be greater than the number of DataMules so that the same success rate can be maintained.

Spyropoulos *et al.* (2004) proposed a simple single-copy routing called direct transmission routing. In this approach, after the source node generates a message, the message is held by the source node until it reaches the destination node. The main advantage of this scheme is that it incurs minimum data transfers for message deliveries. Although this method has minimal overhead, this scheme may incur very long delays for message delivery since the delivery delay for this scheme is unbounded.

Network coding based algorithms: Due to wireless channel loss and other challenges, network coding, a novel technology, has been recently proposed for opportunistic networks. Network coding, coming from information theory, can be utilized in data forwarding to further improve network throughput (Widmer and Jean-Yves, 2005; Wang *et al.*, 2005).

Widmer and Jean-Yves (2005) proposed a communication algorithm that significantly reduces the overhead of probabilistic routing algorithms, making it a suitable building block for a delay-tolerant network architecture. The scheme is based on network coding. Nodes do not simply forward packets they overhear but may send out information that is coded over the contents of several packets they received. Simulation results show that this algorithm achieves the reliability and robustness of flooding at a small fraction of the overhead. In order to describe this method briefly and concisely, we just give a classical example: given a small network consisting of only three nodes A, B and C. A and B can communicate with each other via bi-directional links, so do B and C. Let node A generate the information a and node C generate the information c . Then suppose the information produced needs to be known to all the nodes. Hence, nodes A and C send their information to node B. Then node B, rather than sending two different packets for a and c , respectively, broadcasts a single packet containing $a \oplus c$. Once $a \oplus c$ is received, both nodes A and C can finally infer the missing information (i.e., node A can infer c and node C can infer a). Since network coding based routing is able to deliver the same information with a fewer number of messages injected into the network, it will outperform flooding based routing.

Wang *et al.* (2005) assumed that coded blocks with replication factor r are equally split among the first mr relays (or contacts), for some constant m and those relays must deliver the coded blocks to the destination directly. The original message can be decoded as soon as m contacts deliver their data (that is, as soon as $1/r$ of the coded blocks have been received). This approach sends data dynamically to the first m contacts the node meets, namely, the allocation of the coded blocks is not fixed. It also differs from the history information based approaches discussed earlier in that it does not try to find which contacts have better chances to deliver the data. Instead, it simply forwards to the first m contacts the node meets. Here, all contacts are equally good relays. Both analytic and simulation results show the erasure coding based forwarding in DTNs significantly improves the worst case delay.

CONCLUSION AND FUTURE DIRECTIONS

In recent years, many representative works have been proposed to end the unique problems and to meet the challenges in opportunistic networks. In this study we provide a comprehensive overview of the state of the art on data forwarding in opportunistic networks and classify different data forwarding algorithms into different categories according to the characteristics of the network behaviors and the corresponding approach proposed. To the best of our knowledge, this classification is the newest and most comprehensive one till now. This will assist readers who are new to the field to have a better overall understanding of the frontier of this area and assist them to start conducting research quickly.

Study on data forwarding in opportunistic networks is still in its early stages and there are still many open issues that need to be resolved before the benefits of the opportunistic networks can be fully utilized, which include but are not restricted to: Due to the limited network resource, in order to reduce the overhead produced by data copying and forwarding, the node history information including node velocity and link state should be better used. In this sub-area, there are many interesting works need to be done.

Because network nodes may belong to different society or organization, nodes will be selfish while performing data forwarding for other nodes. Thus, how to stimulate nodes to cooperation during data forwarding process is another one interesting sub-area which is worth to be studied further.

Furthermore, in order to cope with the adverse effect brought by multipath fading in opportunistic networks and get better performance, Physical Layer, MAC Layer and Network Layer based cross layer optimization should be considered while forwarding data in opportunistic networks.

ACKNOWLEDGMENTS

This study was supported by the Scientific Research Foundation of ZSTU (NO. 0704698).

REFERENCES

- Burgess, J., B. Gallagher, D. Jensen and B.N. Levine, 2006. Maxprop: Routing for vehicle-based disruption-tolerant networks. Proceedings of IEEE INFOCOM 2006, Barcelona, Gatalunya, Spain, Apr. 23-29, IEEE Communications Society, pp: 1-11.
- Handorean, R., C. Gill and G.C. Roman, 2004. Accommodating Transient Connectivity in Ad Hoc and Mobile Settings. In: Pervasive Computing, Ferscha, A. and F. Mattern (Eds.). LNCS 3001, Springer-Verlag, Berlin, Heidelberg, ISBN: 978-3-540-21835-7, pp: 305-322.
- Huang, H.Y., P.E. Luo, M. Li, D. Li, X. Li, W. Shu and M.Y. Wu, 2007. Performance evaluation of SUVnet with real-time traffic data. Trans. Vehicular Technol., 56: 3381-3396.
- Jain, S., K. Fall and R. Patra, 2004. Routing in a delay tolerant network. Proceedings of ACM SIGCOMM 2004, Portland, Oregon, USA., Aug. 30-Sept. 3, ACM, New York, pp: 145-158.
- Juang, P., H. Oki, Y. Wang, M. Martonosi, L.S. Peh and D. Dubenstein, 2002. Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with zebranet. Proceedings of ASPLOS'02, San Jose, CA, Oct. 5-9, IEEE Communications Society, pp: 78-93.
- Kang, H. and D. Kim, 2008. Vector routing for delay tolerant networks. Proceedings of IEEE VTC 2008-Fall. Calgary, Canada, Sept. 22-25, IEEE Vehicular Technology Society, pp: 122-135.
- LeBrun, J., C. Chen-Nee, D. Ghosal and M. Zhang, 2005. Knowledge-based opportunistic forwarding in vehicular wireless ad hoc networks. Proceedings of IEEE 61st Vehicular Technology Conference, Stockholm, Sweden, May 30-Jun. 1, IEEE Vehicular Technology Society, pp: 2289-2293.
- Leguay, J., T. Friedman and V. Conan, 2005. Dtn routing in a mobility pattern space. Proceedings of the ACM SIGCOMM Workshop on Delay-Tolerant Networking. Philadelphia, PA, USA., Aug. 22-26, ACM, New York, pp: 276-283.
- Leguay, J., T. Friedman and V. Conan, 2006. Evaluating mobility pattern space routing for DTNs. Proceedings of IEEE INFOCOM 2006, Barcelona, Gatalunya, Spain, Apr. 23-29, IEEE Communications Society, pp: 34-45.
- Lilien, L., A. Gupta and Z. Yang, 2007. Opportunistic networks for emergency applications and their standard implementation framework. Proceedings of IPCCC 2007, New Orleans, Louisiana, USA., Apr. 11-13, IEEE Communications Society, pp: 588-593.
- Lindgren, A., A. Doria and O. Schelén, 2004. Probabilistic Routing in Intermittently Connected Networks. In: Service Assurance with Partial and Intermittent Resources, Dini, P. et al. (Eds.). LNCS., 3126, Springer-Verlag, Berlin, Heidelberg, ISBN: 978-3-540-22567-6, pp: 239-254.

- Liu, L. and G. Feng, 2005a. Swarm intelligence based node-disjoint multi-path routing protocol for mobile ad hoc networks. Proceedings of the 5th International Conference on Information, Communications and Signal Processing, Bangkok, Thailand, Dec. 6-9, IEEE Communications Society, pp: 598-602.
- Liu, L. and G. Feng, 2005b. A Novel Ant Colony Based QoS-Aware Routing Algorithm for MANETs. In: Advances in Natural Computation, Wang, L., K. Chen and Y.S. Ong (Eds.). LNCS. 3612, Springer Verlag, Berlin, Heidelberg, ISBN: 978-3-540-28320-1, pp: 457-466.
- Liu, L. and G. Feng, 2005c. A Novel Heuristic Routing Algorithm Using Simulated Annealing in Ad Hoc Networks. In: Advances in Informatics, Bozanis, P. and E.N. Houstis (Eds.). LNCS. 3746, Springer Verlag, Berlin Heidelberg, ISBN: 978-3-540-29673-7, pp: 849-857.
- Liu, L. and G. Feng, 2007. Simulated annealing based multi-constrained QoS routing in mobile ad hoc networks. *Wireless Personal Commun.*, 41: 393-405.
- Liu, L. and Z. Chen, 2009. Enhanced fisheye state routing using cross-layer optimization. Proceedings of Chinese Control and Decision Conference, Guilin, China, Jun. 17-19, IEEE Industrial Electronics, Singapore, pp: 4673-4676.
- Mundur, P., M. Seligman and G. Lee, 2008. Epidemic routing with immunity in delay tolerant networks. Proceedings of IEEE Military Communications Conference, San Diego, CA, USA., Nov. 17-19, IEEE Communications Society, pp: 1-7.
- Musolesi, M., S. Hailes and C. Mascolo, 2005. Adaptive routing for intermittently connected mobile ad hoc networks. Proceedings of 6th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, Taormina, Giardini Naxos, Jun. 13-16, IEEE Communications Society, pp: 183-189.
- Shah, R.C., S. Roy, S. Jain and W. Brunette, 2003. Data MULEs: Modeling and analysis of a three-tier architecture for sparse sensor networks. *Ad Hoc Networks*, 1: 215-233.
- Small, T. and Z.J. Haas, 2003. The shared wireless Infostation model: A new ad hoc networking paradigm (or Where there is a whale, there is a way). Proceedings of the 4th ACM International Symposium on Mobile ad hoc Networking and Computing, Annapolis, Maryland, USA., Jun. 1-3, ACM, New York, pp: 233-244.
- Spyropoulos, T., K. Psounis and C.S. Raghavendra, 2004. Single-copy routing in intermittently connected mobile networks. Proceedings of 1st Annual IEEE Communications Society Conference on Sensor and Ad hoc Communications and Networks, Santa Clara, California, Oct. 4-7, IEEE Communications Society, pp: 235-244.
- Tan, K., Q. Zhang and W. Zhu, 2003. Shortest path routing in partially connected ad hoc networks. Proceedings of IEEE Globecom 2003, San Francisco, USA., Dec. 1-5, IEEE Communications Society, pp: 1038-1042.
- Wang, Y., S. Jain, M. Martonosi and K. Fall, 2005. Erasure-Coding based routing for opportunistic networks. Proceedings of the ACM SIGCOMM Workshop on Delay-Tolerant Networking, Philadelphia, PA, USA., Aug. 22-26, ACM Press, pp: 229-236.
- Widmer, J. and L.B. Jean-Yves, 2005. Network coding for efficient communication in extreme networks. Proceedings of the ACM SIGCOMM Workshop on Delay-Tolerant Networking, Philadelphia, PA, USA., Aug. 22-26, ACM, New York, pp: 284-291.
- Zhao, W., M. Ammar and E. Zegura, 2004. A message ferrying approach for data delivery in sparse mobile ad hoc networks. Proceedings of 5th ACM International Symposium on Mobile Ad Hoc Networking and Computing, Tokyo, Japan, May. 24-26, ACM Press, pp: 187-198.
- Zhao, W., M. Ammar and E. Zegura, 2005. Controlling the mobility of multiple data transport ferries in a delay-tolerant network. Proceedings of INFOCOM2005, Miami, Florida, USA., Mar. 13-17, ACM, New York, pp: 1407-1418.