DOI: 10.3969/j. issn. 1673-5188. 2015. 04. 007
http://www.cnki.net/kcms/detail/34.1294.TN.20151118.0733.002.html, published online November 18, 2015

Gateway Selection in MANET Based Integrated System: A Survey

Ye Miao, Zhili Sun, and Ning Wang
(Institute for Communication Systems (ICS), University of Surrey, Guildford, GU27XH, UK)

Abstract

Taking advantage of spontaneous and infrastructure-less behavior, a mobile ad hoc network (MANET) can be integrated with various networks to extend communication for different types of network services. In the integrated system, to provide inter-connection between different networks and provide data aggregation, the design of the gateway is vital. In some integrated networks with multiple gateways, proper gateway selection guarantees desirable QoS and optimization of network resource utilization. However, how to select gateway efficiently is still challenging in the integrated MANET systems with distributed behavior terminals and limited network resources. In this paper, we examine gateway selection problem from different aspects including information discovery behavior, selection criteria and decision-making entity. The benefits and drawbacks for each method are illustrated and compared. Based on the discussion, points of considerations are highlighted for future studies.

Keywords

gateway selection; MANET; integrated network; routing algorithms; QoS

1 Introduction

As mobile technology develops, mobile devices have become more powerful with high transmission speed, low power consumption and low cost. Hence, the paradigm of multihop ad hoc networking has become popular. An established mobile ad hoc network (MANET) can cooperate with various network systems to extend communication areas for applications and ensure better user experience. The application domains, e.g. public safety [1], intelligent transportation system, mesh community and mobile data off-loading [2], can benefit from such integration. One of the major challenges of the integrated networks is to serve large amount of traffic data from end devices with limited resources. Thus, in order to satisfy the Quality of Service (QoS) of the different applications and devices, proper and efficient mechanisms should be developed for suitable allocation of available network resources.

In an integrated system, a MANET node can communicate with external networks via gateways. A gateway node, with multiple interfaces and protocols, acts as interworking unit responsible for forwarding traffic between different types of network. To provide interoperability and end-to-end network connectivity for end devices, gateways need to be implemented with various networking functions that regulate data and control traffic. To optimize network resource utilization and achieve satisfied QoS, the data transmissions must be distributed and make best use of allocated resources accordingly which can be achieved by proper gateway selection mechanisms. However, due to the distributed behavior of mobile nodes and limited network resources, there are still some issues to be addresses when designing gateway selection mechanisms.

In this paper, we investigate the gateway-selection problem. The main contribution of this paper is to provide a survey of gateway selection mechanisms in a MANET based integrated networks. The previous studies are categorized in terms of different criteria, such as information discovery method, selection metric, and decision-making entity. We present explanations and comparison of the studies in each category, and specify some load-balancing based algorithms. With the development of hardware techniques, networks such as machine-to-machine network (M2M) and Internet of things (IoT) can also be integrated with MANETs for data aggregation. With a large number of mobile devices and enormous amounts of data, load-balancing is essential for alleviating traffic congestion and even traffic distribution in order to improve network performance. We also point out future study directions based on the surveyed papers and analysis.

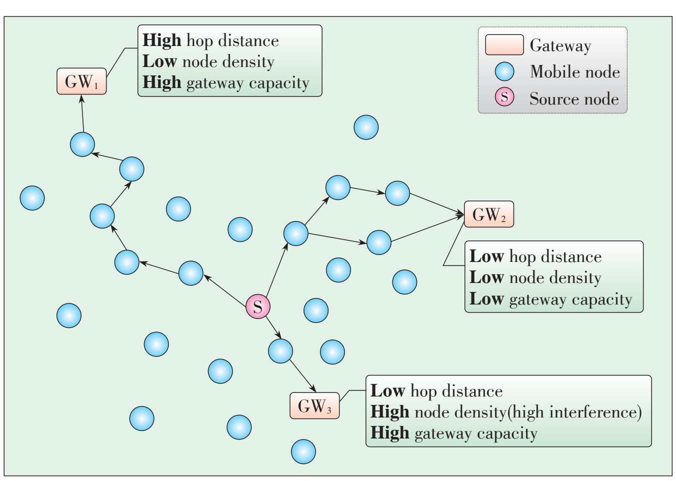
The rest of the paper is organized as follows: In section 2, we discuss and state properties of gateway selection problem. In section 3, the current gateway selection mechanisms in terms of different categories are reviewed. In section 4 we highlight the open research challenges and issues for future research. Finally, in section 5, we conclude the paper.

2 Gateway Selection Problem Statement and Discussion

In order to communicate with external networks, mobile nodes should target one or more trusted gateways to direct their traffic to. During this procedure, several aspects need to be considered. Firstly, the gateways are limited by processor capacity and outgoing transmission link capacity. Thus, the amount of traffic that goes through a gateway must not exceed the capacity of the gateways; otherwise, the data packets will be dropped. This may lead to more packets having to be retransmitted, longer delay, and poorer throughput. In addition, gateways may have different capacities of connections that are not known by the mobile nodes. This may cause unevenly utilization of gateways' resources. Furthermore, mobile nodes select gateways spontaneously without considering other nodes, it is possible for them to have the same preferable paths at the same time, which may cause congestions on wireless channel both in-network and around gateways.

In order to make best use of the network resources and avoid problems, it is essential to consider network performance during gateway and path selection, when there are multiple paths and gateways available in network. Gateway selection problems include 1) which gateway to select for each data packet flow, 2) how to route the flows through multi-hop network paths established between mobile nodes and selected gateways, and 3) increase network throughput and improve QoS in integrated system. Gateway selection problem can be broken into three steps: measurement, decision and execution [3]. The first step is path/gateway information discovery and collection. Information accuracy is guaranteed in this step. The second step is decision phase that concentrates on the algorithms and decision metrics. Based on the information collected in the measurement step and algorithm designed, a certain path/gateway is decided in this step. Execution, the last step, illustrates which terminal devices make the decisions. The decisions can be made in nodes locally, or they can be made centrally and distributed to nodes.

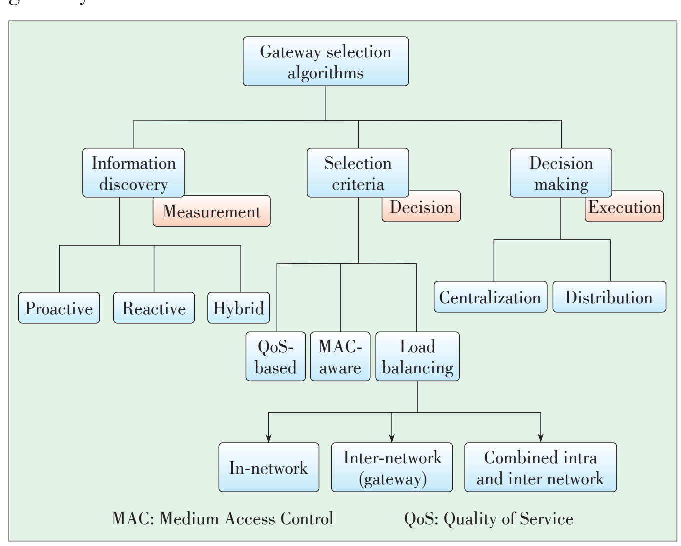
Fig. 1 illustrates the topology under consideration. The considered topology is formed from a MANET with large number of mobile nodes that can communicate with each other within communication range. Any mobile node can be the source node which requires external communication. Multiple gateways are available to provide external links and multiple paths are possible to route to gateways. In Fig. 1, for the source node s, there are currently three gateways available. They have different properties: gateway 1 and gateway 3 are able to handle more traffic and have higher capacity compared with gateway 2. And clearly, the node density around gateway 3 is higher than that around gateway 1 and gateway 2, which means it is highly possible that interference needs to be considered when it comes to gateway 3. In terms of location position, gateway 1 is the most far away one, which means hop distance delay is the largest theoretically. For gateway 2, there are two paths established with different hop distance and node density (transmission interference). Should the decisions made in node s or in other designed entity? How to make sure the validation and accuracy of these information? With all these information, which gateway and path should be selected to the node s to achieve the best network performance? Gateway selection aims at solving these questions.



▲Figure 1. Target MANET topology with multiple gateways and available paths.

3 State-of-the-Art Gateway Selection in an Integrated MANET

Various approaches have been proposed and in this section we present a literature review of the studies in gateway selection problem. Different gateway selection mechanisms are designed to achieve various functions and objectives in different systems. The studies are categorized in terms of various aspects as shown in **Fig. 2**. A basic category criterion is the way a gateway is discovered and how the gateway/path information is collected. For this purpose, proactive, reactive and hybrid mechanisms are proposed. Each mechanism has its own advantages. Another criterion is the means of decision-making for gateway selection in network level. The decision can be made



▲Figure 2. Category illustration of gateway selection algorithms.

in a central point (central controller) and a centralized mechanism is needed in this case. On the contrary, the decision can be made at local nodes that require the mechanisms to be distributed. The decision making we discuss here mainly focus on the terminal devices that make the decisions and the means of execution. Hence, the decision making is related to the execution process. Based on the objective of algorithm design, the studies of gateway selection can be categorized into QoS - aware, Medium Access Control (MAC)-aware and load balancing based group. **Table 1** summarizes the studies briefly in terms of different criteria. In the following part, these studies will be illustrated with more design details.

3.1 Decision Making

Based on the way of decision-making, the studies can be grouped as centralized solutions and distributed solutions. The most common methodology proposed to address gateway selection problem is to have a centralized controller that collects network information and runs an algorithm to determine gateways for mobile nodes which are grouped as centralized solutions. Gateways can also be responsible for making selection decisions, and we group these as gateway centralized studies. On the contrary, decision-making can occur in local nodes and based on this, the distributed solutions are designed for node-centric behavior. Based on the collected gateway and path information, mobile nodes make their own decisions for choosing which gateway to deliver traffic flows to.

The work in [4] considers 1) eliminating the gateway flapping, 2) avoiding harmful effect caused by gateway selection interference, and 3) balancing traffic at TCP flow level and improving the performance and fairness of flows. It is assumed that there is routing protocol executed to establish routes and perform routing within network. The network topology graph and routes are assumed to be obtained from the routing protocol. Gateways are connected to a central controller by wired network. When a set of flows passes through the gateway, the topology and path information is passed to the controller. Based on this, the controller executes a fast gateway selection algorithm and informs the gateway new association.

▼Table 1. Studies of gateway selection

Information discovery	Ref.	Selection criteria	Decision making
Proactive	[7]	Expected throughput; MAC link interference	Distributed
	[18]	Expected link quality; Interference ratio; Gateway load	
	[15]	Gateway capacity; Hop distance	Centralized
	[8]	Contention level; Congestion level; Hop distance	Gateway centralized
	[9]	Hop distance; Traffic volume	
	[14]	Gateway traffic load; Hop distance threshold	
Reactive	[4]	No. of Register nodes; Expected interference; Hop distance	Centralized
	[10]	Traffic volume	Distributed
	[19]	Gateway traffic load; Hop distance	
Hybrid	[5]	Path available period; Path capacity; Path latency	Distributed

The selection algorithm is periodically executed to refresh the associations and all the flows of a node have the same gateway association. The node-gateway association is decided according to two ranked lists. The unlocked nodes (nodes with no gateway association) are ordered according to the number of valid paths (gateways). The valid paths are ordered in ascending order based on the gateway load (number of nodes the gateway serves) and path cost (expected interference and hop distance). Unlocked nodes are assigned one by one to the gateway with minimum load and cost. Even though the aim is balancing the load at the flow level, this paper only explains node-gateway association.

Based on the network-wide information, centralized solutions are likely to determine the most suitable gateway for a certain node and make optimized decisions for all the considering nodes. However, they suffer from the fact that all the information has to be collected at a single point. In the topology-dynamically-changing network, to capture the latest overall network condition is challenging.

3.2 Selection Criteria

As mentioned, there are questions in gateway selection procedure: 1) which gateway to select for each traffic flow, 2) how to route data flows through multi-hop network paths established between selected gateways and mobile nodes. Various approaches to addressing these issues significantly affect the performance of the entire system. They can be designed to increase network throughput, improve QoS, and increase the efficiency of network resource usage. Considering the objective and emphasis of selection criteria, the studies can be classified into different categories: QoS-based, MAC-aware and Load balancing based algorithms. Selection criteria, which determine decision metrics, should be categorized to the decision process.

3.2.1 QoS Based Studies

To improve QoS, some performance metrics are designed to evaluate the quality of path/gateway based on network and gateway parameters. QoS based gateway selection mechanisms favours some network parameters, such as hop distance, end-to-end delay, link connectivity, and residual load capacity of gateways. Some early works were designed based on a single metric, however, single metric based algorithm sometimes needs to sacrifice another network performance. Hence, trade-off between network parameters should be taken into account and algorithms with multiple metrics are proposed.

The authors in [5] consider multiple QoS path parameters such as path availability period, available capacity and latency. In order to achieve considerations of both in-network and gateways, the parameter of the end-to-end path between a mobile node and a gateway node are computed. The path availability period is estimated based on mobility model and epoch length caused by its movement is considered. To avoid in-network bottlenecks, the residual load capacity of a path is formulated as the minimum available load capacity at any node along the path. Each node makes an independent decision on which gateway to choose based on the metric value to each gateway. The metric is based on a linear combination of the three terms.

The authors in [6] aim at balancing the inter/intra-MANET traffic load over multiple IGWs (Internet gateways). The paper provides a metric consisting three components: shortest Euclidean distance (hop count), inter-MANET traffic load and intra-MANET traffic load. The metric is a linear combination of these three components. The inter-MANET traffic load is represented by the number of registered MANET nodes (both local and visiting) at the IGW. The intra-MANET traffic load within the network is related to the optimal node density to delivery traffic successfully. However, the definitions of inter and intra-MANET traffic given in paper are partly redundant and not validated well.

3.2.2 MAC Based Studies

The selection decision over different routing paths depends on link performance indicators, which are influenced by the MAC parameters. Hence, the appropriate routing metric should adapt to the corresponding MAC parameters for efficient joint functioning. In [7], authors develop load balancing algorithm on top of expected throughput (ETP) routing metric. ETP is a MAC-aware routing metric which takes into account the capacity reduction of a link due to its interaction with other links in its contention domain. ETP takes into account all the existing flows on the paths at a node. It aims to determine a delay-optimal routing forest. Unlike the algorithms mentioned previously, this study does not explicitly use gateway advertisements or requests to gateways from nodes. All the information is achieved by neighboring messaged exchange. Control message overhead is sufficiently decreased. In this paper, the bit rate and the packet success probability of a link to all its potential parent nodes is available at current node. Based on these, next hop with minimum expected delay experienced by a bit is chosen which achieves local optimization. Benefiting from the distributed manner, it is expected to minimize the expected per bit delay over the entire network.

3.2.3 Load Balancing Based Studies

Due to the distributed behavior of mobile nodes, nodes in certain locations could be more vulnerable to congestion because they may have a higher number of neighbors or they are responsible for transmitting the majority of the packets. Gateway nodes can also be vulnerable to congestion because they aggregate all the network traffic. The congestion is caused by unevenly distribution of the traffic. Load balancing schemes are essential to distribute the traffic evenly and alleviate congestion. Furthermore, the load balancing schemes can be used to decrease the potential delay caused by congestions and improve the resource utilization. Therefore, load balancing is an important consideration for improving network performance. There are a number of studies that have been performed to focus on balancing network traffic and they differ in the metric used for representing path/gateway traffic load. These studies can be sub-categorized into: in-network load-balancing, gateway load-balancing and combine network and gateway load-balancing.

The in-network load-balancing mechanisms emphasize that the traffic between source nodes and gateways is distributed between a set of alternative paths in order to avoid congestions and improve network performance. Most of them focus on design of path performance metric. In [8], the contention level, congestion level, and hop distance are combined as the a selection metric to avoid areas which have high data traffic or many nodes contending for channel access. The authors in [9] propose a potential-field-based routing scheme to reflect both distance and traffic volume at each node to multiple gateways. Based on different path quality metrics, these studies distribute data traffic across the network. However, compared to mobile nodes, the gateway nodes are more vulnerable to congestions and easier to become bottlenecks because they aggregate all the network traffic.

The gateway load-balancing mechanisms aim at distributing traffic load among multiple gateways in order to reduce gateway load imbalance and to maximize the total network throughput. The related approaches mostly consider gateway load for selection and the gateway load is evaluated with different gateway parameters, such as actual amount of traffic [10], residual capacity [11], the number of active flows, average queue length [12], the number of registered nodes [13], and so on. The authors in [14] consider the gateway load and channel contention in cost function for gateway selection. Each node calculates its own cost contribution to a gateway by multiplying the hop distance by the reciprocal of the number of nodes whose traffic is forwarded. Each gateway receives route requests from all mobile nodes, calculates total costs (summation of the reciprocals of the number of nodes for which traffic is forwarded) and then includes the related information in its periodic advertisement message to mobile nodes. There is also a hop count threshold introduced, and the gateways that are more than a certain hop distance are not considered as a candidate. However, the cost function is not explained very well and is not justified as a good metric.

In [15], the authors design a practical, robust framework for load-balanced routing and gateway selection under variable traffic loads. In this architecture, end-to-end QoS provisioning and network resource management are deployed on a centralized controller called the Network Manager (NM). The NM is responsible for collecting information on the network topology and the available capacities of all links, maintaining a list of active flows. Two schemes are proposed: (i) to assign a given flow to the gateway with maximum residual capacity (MRC), (ii) to assign the flow to the gateway with maximum normalized capacity (MNC), which is the ratio of residual gateway capacity to hop distance. To always select gateways with a lighter load, these solutions may evenly distribute data packets among multiple gateways in average during a period. However, they do not take into account the effects of route-switching, which may cause gateway flapping and also degrade in-network performance [16].

There are only a few studies about combining load balancing within the network and among gateways. In [17], gateway load, route interference and expected link quality are combined as metric for nodes to choose the best gateway. The gateway load is defined as the average queue length at the network interface of accessing gateway. To estimate the gateway load accurately, an exponential weighted moving average is used to smooth out variations. Route interference is defined as the sum of the link interferences along the route. Link interference is the maximum of the node interference, where node interference is defined as the percentage of the time nodes sense wireless activities in the channel. Finally, the expected link quality is defined as the forward link packet delivery ratio. Using the combined metric (of gateway load, route interference, and expected link quality), the aim is to evenly distribute traffic across the network and equalize the traffic load on gateways. However, in real scenarios, gateways may have different capacities. Even though the data traffic is uniformly distributed between multiple gateways, a few gateways may still be overloaded. Hence, the utilization, which is defined as ratio of traffic load over maximum capacity of a gateway, should be considered in future work. Moreover, this study uses proactive approach for information discovery, where gateways periodically broadcast advertisement packets including gateway load and mobile nodes periodically broadcast probe packets including node interference factor and packet delivery ratio. A proactive approach may keep routes up to date, but it increases interference due to large number of routing packets.

In [18], a source based gateway selection scheme is proposed and it combines path metric and gateway load. Path metric is a function of link metric, which is defined as a combination of expected link quality and interference ratio. Expected link quality is the success rate of transmitted probe packets. The interference ratio is the ratio of the sum of interference power from all interfering nodes over the maximum tolerable interference at the receiver. The gateway load is based on the average interface queue length and represented as gateway capacity. The gateway capacity is the sum of the capacities of three interfaces in each gateway, where the capacity of each interface is defined as subtraction of maximum interface queue capacity and current interface queue length. This study also proposes a waiting time before switching to a better path because an immediate switch may cause congestion and oscillations when several source nodes switch to a better path at the same time. However, it does not consider that the better path may change during this waiting time period. Even though the gateway load is part of selection metric, this study focuses on decreasing packet loss/retransmission, instead of balancing network load. Another issue of this study is that only source nodes are designed and implemented with gateway/path selection algorithm. In practical scenarios, source nodes are normally far away from gateways and they may not be sensitive enough to catch the latest network conditions which dynamically change. Hence, the intermediate nodes should also be able to make decisions based on the latest path information to improve selection accuracy.

3.3 Gateway and Information Discovery

The gateway/path discovery is about information collection and the accuracy of path parameters is an important design objective. The functionality mechanisms (QoS aware, MAC-aware and load balancing) are usually built on top of discovery algorithm. Things that need to be considered with the different methods for gateway/path discovery are network control overhead packets, computation complication of algorithm, and initial delay of path setup.

In [6] and [14], gateways proactively broadcast their current information of the number of registered nodes and network topology periodically. For proactive behavior, there will be large control packets transmitting in network, which consumes network resources. Besides, the path/gateway information may not be updated frequently enough to keep up to date. The advantage is the routing information will be always available in mobile nodes and packets can be redirected immediately when it arrives at the node.

The authors in [19] use reactive behavior to collect the path/gateway information, which means only when the route to a certain destination is required, the mobile node will initiate a route request. The current path/gateway condition will be integrated into the route reply packets. Control packets in this case include route request and route reply packets which will be much smaller than those in the proactive algorithm. On the other hand, the route initiation takes time, which increases the end-to-end delay of packet transmission.

In [5], the authors use a hybrid gateway discovery algorithm. The proactive and reactive zones are divided by the number of hops from mobile nodes to gateway node. Each gateway periodically broadcast the gateway advertisement message which contains its parameters within a proactive region of k hops. The nodes in reactive zone discover the gateway node and path information by sending gateway discovery message. In this case, the number of control messages will be reduced compared with the proactive scheme. Also, the nodes within k-hop distance have less route establishment time compared to the reactive scheme. However, it requires more complicated computation work in this scheme and how to define the optimal proactive region is critical in this scheme.

4 Open Issues and Challenges for Future Studies

Integration of MANET and other networks has been well studied, but when it comes to gateway selection, there are still some additional issues that must be taken into account and addressed. In this section, the challenges and issues are explained and discussed. **Table 2** summaries the challenges and future directions.

▼**Table 2. Challenges and directions for gateway selection**

Challenge	Problem Statement	Direction
Route flapping	Synchronization rerouting; Traffic unevenly distribution	To evaluate route change effects; Traffic control
Performance trade-off	Multiple objective; Resource and cost conflict	To evaluate costs (individual and combine objectives)
Path information update	High accuracy cost, energy and network resource consuming	To evaluate information update cost; Maximize frequency and limit resource consuming

4.1 Routing Flapping

The most distinguished feature of ad hoc networks is that the self-organization which makes it impossible to achieve accurate, instant-react routing management in-network. Most of the protocols are designed in a distributed manner, and there will be latency for nodes' reaction to changes in the network topology. It is likely that at one time all the nodes find out the current default gateway is overloaded, so they turn to register in another gateway. This leads to overloading on the targeted gateway. The overloading cannot be alleviated until the time that the overloaded information of this gateway is known by the majority of mobile nodes. At this time, the other gateway starts to overload again. This synchronized rerouting problem causes routing flapping (also referred to as the ping pong effect). It is hard to maintain an absolutely balance in load distribution, however, optimal schemes can be applied to alleviate this flapping effect. There must be coordination between gateways and mobile nodes to intelligently select the gateway and alleviate flapping.

Furthermore, a route change in a node may affect others, leading to affect overall network performance. The more aggressive of the gateway switching/route change is, the more likely to be poor network performance is. The trade-off between route decision and network performance also need to be considered. Thus, how to evaluate the effect of route changes and how to control the changes to achieve the best performance gain must be further studied.

4.2 Performance Trade-Off

With respect to the network view, there is a cost if additional improvements are made beyond the natural network behavior. To achieve realistic situation, multiple-objectives are normally required, for example, to achieve load-balancing and multi-QoS at the same time. However, when a certain objective is achieved, normally some other may not be accomplished. In this case, how to coordinate the behavior and performance of different metrics is crucial and the trade-off of different achievements needs to be considered.

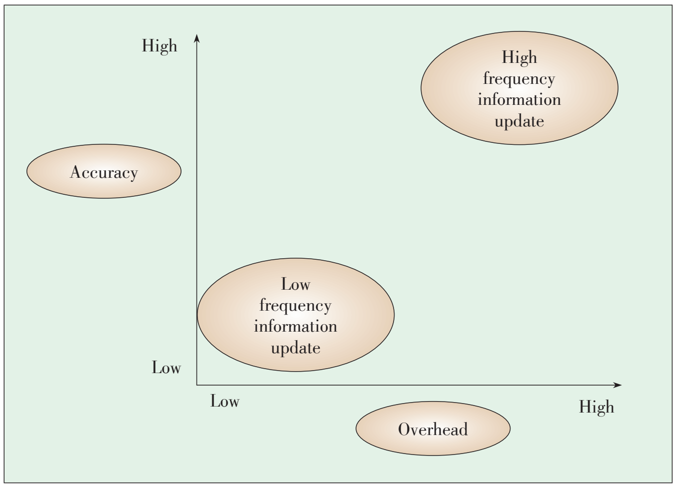
For example, when path reliability is of high priority, a more reliable, longer distance path may be chosen for higher reliability. In this case, the overall data delivery ratio will be increased but also the end-to-end delay. Due to large amount of data aggregated, it is likely to be congested when the nodes have a large number of neighbor nodes or the nodes are on the shortest path. To alleviate the congestion, a less traffic path with longer hop distance can be selected. For example, gateway 1 will be chosen instead of gateway 2 and gateway 3 in Fig. 1. Load-balancing is achieved but delay is increased as the cost.

Different QoS and priorities need to be considered for various applications and scenarios/topologies. In terms of applications, the voice message values the transmission delay and jitter. On the other hand, file transmission prefers high delivery ratio other than delay. When it comes to application scenarios, emergency response integrated system requires both delay and delivery ratio to be high level performance. Most of the mechanisms are designed to be effective in a specific scenario. The challenging issue in this case is the design of coefficients that should be able to not only achieve promising performance in theory but affect the realistic situation. Besides, the goal is to make best use of network resource and sacrifice the least.

Furthermore, the most challenging part is how to design an adaptive mechanism that is able to cope with different cases through its own learning/adaptive function and achieve the related goals.

4.3 Information Update Frequency (Trade-off Between Accuracy and Cost)

The route/gateway selection decision depends on current path conditions. Parameters of path conditions are collected from nodes by different information discovery methods. The frequency of path parameter updates determines how accurately a node can adapt to changes. There is a trade-off between acquired accuracy and resource efficiency including overhead packets and energy consumption (**Fig. 3**). More frequent updates may achieve more accuracy when adapting to network dynamics, but it may also cause higher control messages consume and higher energy consumption. If the update is not frequent enough, the changes of state of gateways/paths may not



▲**Figure 3. Trade-off between control overhead and information accuracy.**

be informed to mobile nodes in time. The information stored in nodes could be out of date. Hence, how to limit the update frequency and at the same time achieve the desirable accuracy needs to be investigated. In addition, route changes may affect the diffusion of path information. Hence, to make sure the information is disseminated efficiently, the route change effect should also be considered as one aspect.

In multihop Wireless Sensor Networks (WSNs) with static gateways, the nodes close to the gateways are more likely to deplete their battery supplies before the far-away nodes due to the intersection of multihop routes and concentration of data traffic towards the gateways. To achieve uniformity of energy consumption, the usage of mobile gateways is proposed and explored recently [20], [21]. As the gateway moves, the hotspot nodes around the gateway change and energy increaseingly drains through the network. Furthermore, to harvest multiple-input multiple-output (MIMO) gains in WSNs, several nodes which are equipped with one or more antennas emulate a multi-antenna node and independent paths between the transmit and receive sides can be realized by having multiple spatially separated nodes [22], [23].

5 Conclusions

To provide inter-connection between intrinsically different networks, data aggregation function and network performance guarantee, the gateway in an integrated system is vital to satisfy requirements. The gateway selection problem can be addressed as a centralized optimization problem, where performance metrics (e.g. end-to-end delay, overall network throughput, etc.) should be maximized. To achieve the objectives, the network topology and traffic flows need be assumed to be static and known. However, either network topology or traffic flows are uncertain and highly dynamic in the system. The gateway selection problem, on the other hand, can be addressed as a network performance improvement problem in terms of different routing metrics (e.g., traffic load, path delay, multiple QoS metrics, etc.). The information is gathered in a distributed learning way that may decreases timeliness. Besides, in order to achieve local optimization, certain paths may be sabotaged by related nodes.

Algorithms for future designs should be elaborated to cope with realistic scenarios and deal with the challenges stated in the last section. Update of path/gateway quality parameters must be designed carefully and efficiently, in which case it guarantees the information accuracy and achieves performance gain with the least network cost. Due to dynamic topology and real-time traffic condition, the route decisions should be dynamic and based on current information. It is desirable to design the algorithm manageable and adjustable to different network conditions. Furthermore, the problem can be further extended to gateway selection in multi-domain wireless networks [24], [25], where the inter-domain costs need to be considered.

References

[1] T. Taleb, Y. H. Aoul, and T. Ahmed, "Challenges, opportunities, and solutions for converged satellite and terrestrial networks," *IEEE Wireless Communications*, vol. 18, no. 1, pp. 46–52, Feb. 2011. doi: 10.1109/MWC.2011.5714025.

[2] A. Aijaz, H. Aghvami, and M. Amani, "A survey on mobile data offloading: technical and business perspectives," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 104–112, Apr. 2013. doi: 10.1109/MWC.2013.6507401.

[3] A. Oliveira, Z. Sun, M. Monier, *et al.*, "On optimizing hybrid ad-hoc and satellite networks—the MONET approach," in *Future Network and Mobile Summit*, Florence, Italy, 2010, pp. 1–8.

[4] J. J. Galvez, P. M. Ruiz, and A. F. G. Skarmeta, "Responsive on-line gateway load-balancing for wireless mesh networks," *Ad Hoc Networks*, vol. 10, no. 1, pp. 46–61, Jan. 2012.

[5] S. H. Bouk, I. Sasase, S. H. Ahmed, and N. Javaid, "Gateway discovery algorithm based on multiple qos path parameters between mobile node and gateway node," *Journal of Communications and Networks*, vol. 14, no. 4, pp. 434–442, 2012.

[6] Q. Le-Trung, P. E. Engelstad, T. Skeie, and A. Taherkordi, "Load-balance of intra/inter-MANET traffic over multiple internet gateways," in *MoMM' 08*, Linz, Austria, 2008, pp. 50–57. doi: 10.1145/1497185.1497200.

[7] V. Mhatre, H. Lundgren, F. Baccelli, and C. Diot, "Joint MAC-aware routing and load balancing in mesh networks," *ACM CoNEXT' 07*, New York, USA, 2007, article no. 19. doi: 10.1145/1364654.1364679.

[8] Y. Fu, K.-M. Chan, K. S. Tan, and B. S. Yeo, "Multi-metric gateway discovery for MANET," in *IEEE 63rd Vehicular Technology Conference*, Melbourne, Austrialia, 2006, pp. 777–781. doi: 10.1109/VETECF.2006.1682930.

[9] S. Jung, M. Kserawi, D. Lee, and J.-K. K. Rhee, "Distributed potential field based routing and autonomous load balancing for wireless mesh networks," *IEEE Communications Letters*, vol. 13, no. 6, pp. 429–431, Jun. 2009. doi: 10.1109/LCOMM.2009.090110.

[10] Y. Miao, Z. Sun, F. Yao, *et al.*, "Study on research challenges and optimization for internetworking of hybrid manet and satellite networks," in *PSATS*, Toulouse, France, 2013, pp. 90–101. doi: 10.1007/978-3-319-02762-3_8.

[11] C.-F. Huang, H.-W. Lee, and Y.-C. Tseng, "A two-tier heterogeneous mobile ad hoc network architecture and its load-balance routing problem," *Mobile Networks and Applications*, vol. 9, no. 4, pp. 379–391, Aug. 2004. doi: 10.1023/B:MONE.0000031605.84447.1d.

[12] D. S. Nandiraju, L. Santhanam, N. Nandiraju, and D. P. Agrawal, "Achieving load balancing in wireless mesh networks through multiple gateways," in *IEEE MASS*, Vancouver, Canada, 2006, pp. 807–812. doi: 10.1109/MOBHOC. 2006. 278655.

[13] J. Shin, H. Lee, J. Na, *et al.*, "Load balancing among internet gateways in ad hoc networks," in *IEEE 62nd Vehicular Technology Conference*, Dallas, USA, 2005, pp. 1677–1680. doi: 10.1109/VETECF.2005.1558227.

[14] K. Tada and M. Yamamoto, "Load-balancing gateway selection method in multi-hop wireless networks," in *IEEE GLOBECOM*, Honolulu, USA, 2009, pp. 1–6. doi: 10.1109/GLOCOM.2009.5425325.

[15] E. Ancillotti, R. Bruno, and M. Conti, "Load-balanced routing and gateway selection in wireless mesh networks: design, implementation and experimentation," in *IEEE WoWMoM*, Montreal, Canada, 2010, pp. 1-7. doi: 10.1109/WOWMOM.2010.5534991.
[16] V. Pham, E. Larsen, P. E. Engelstad, and Ø. Kure, "Performance analysis of gateway load balancing in ad hoc networks with random topologies," in *MobiWAC' 09*, Tenerife, Spain, Oct. 2009, pp. 66-74. doi: 10.1145/1641776.1641787.
[17] U. Ashraf, S. Abdellatif, and G. Juanole, "Gateway selection in backbone wireless mesh networks," in *IEEE WCNC*, Budapest, Hungary, 2009, pp. 2548-2553. doi: 10.1109/WCNC.2009.4917735.
[18] M. Boushaba, A. Hafid, and M. Gendreau, "Source-based routing in wireless mesh networks," *IEEE Systems Journal*, vol. PP, no. 99, pp. 1-9, 2014. doi: 10.1109/JSYST.2014.2317453.
[19] Y. Chaba, R. B. Patel, and R. Gargi, "Efficient multiple DYMO routing protocol with gateway selection for hybrid manets," *International Journal of Computer Theory and Engineering*, vol. 4, no. 1, Aug. 2012. doi: 10.7763/IJCTE.2012.V4.541.
[20] S. Yu, B. Zhang, C. Li, and H. Mouftah, "Routing protocols for wireless sensor networks with mobile sinks: a survey," *IEEE Communications Magazine*, vol. 52, no. 7, pp. 150-157, Jul. 2014. doi: 10.1109/MCOM.2014.6852097.
[21] C. Tunca, S. Isik, M. Donmez, and C. Ersoy, "Distributed mobile sink routing for wireless sensor networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 16, no. 2, pp. 877-897, 2014. doi: 10.1109/SURV.2013.10013.00293.
[22] D. N. Nguyen and M. Krunz, "A cooperative MIMO framework for wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 10, no. 3, May 2014.
[23] Q. Gao, Y. Zuo, J. Zhang, and X.-H. Peng, "Improving energy effciency in a wireless sensor network by combining cooperative MIMO with data aggregation," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 8, pp. 3956-3965, Oct. 2010. doi: 10.1109/TVT.2010.2063719.
[24] Y. Song, S. H. Y. Wong, and K.-W. Lee, "Optimal gateway selection in multi-domain wireless networks: a potential game perspective," in *17th Annual International Conference on Mobile Computing and Networking*, Las Vegas, USA, Sept. 2011, pp. 325-336. doi: 10.1145/2030613.2030650.
[25] S. Zhong and Y. Zhang, "How to select optimal gateway in multidomain wireless networks: alternative solutions without learning," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5620-5630, Nov. 2013. doi: 10.1109/TWC.2013.100113.121828.

Manuscript received: 2015-09-02

Biographies

Ye Miao (y.miao@surrey.ac.uk) received her BSc in Electronic Information Science and Technology from China Agricultural University, China, in 2010 and MSc in Mobile and Satellite Communication from University of Surrey, UK, in 2011. She is currently a PhD student at the 5G Innovation Centre, Institute for Communication Systems (ICS), University of Surrey, UK. Her research interests are QoS and routing solutions in MANETs, integrated MANET-Satellite networks and WSNs.

Zhili Sun (z.sun@surrey.ac.uk) is the chair of Communication Networking and a Professor at the 5G Innovation Centre, ICS, University of Surrey, UK. He has been with University of Surrey since 1993. He got his PhD degree in Computer Science from Lancaster University, UK, in 1991. He worked as a postdoctoral research fellow with Queen Mary University of London from 1989 to 1993. He has been principle investigator and technical co-coordinator in many projects within the EU framework programs, ESA, EPSRC, and industries, and published over 125 papers in international journals, book chapters, and conferences. He is the sole author of the book titled *Satellite Networking—Principles and Protocols* by Wiley in 2005, a contributing editor of the book *IP Networking Over Next Generation Satellite Systems* published by Springer in 2008, and a contributing editor of the 5th edition of the text book *Satellite Communications Systems—Systems, Techniques and Technology* published by Wiley in 2009. His research interests include wireless and sensor networks, satellite communications, mobile operating systems, traffic engineering, Internet protocols and architecture, QoS, multicast and security.

Ning Wang (n.wang@surrey.ac.uk) received his B.Eng. (honours) degree from Changchun University of Science and Technology, China, in 1996, his M.Eng. degree from Nanyang University, Singapore, in 2000, and his PhD degree from University of Surrey, UK, in 2004. He is a senior lecturer at the 5G Innovation Centre, ICS, University of Surrey, UK. His research interests mainly include information centric networking (ICN), content and data caching management, network optimisation techniques and QoS.