Routing Algorithm of Smart Grid Data Collection based on Data Balance Measurement Model

Xiaochun Jia, Xingyu Chen, Sujie Shao, Feng Qi
State Key Laboratory of Networking and Switching Technology,
Beijing University of Posts and Telecommunications,
Beijing 100876, China
jxch@bupt.edu.cn

***Abstract*—Data collection communications system for electric power communication network is an important entity network in electric power system. It's significant to provide a reasonable routing mechanism for data collection system. The major risk of original routing mechanism is the burst congestion. In the data collection network of smart grid, that some nodes may suffer from more congestion than others even become the bottleneck of network become the new source of risk. In order to overcome these problems, this paper proposes a routing mechanism to realize data balance. This mechanism firstly abstracts a multi-gate network model of data collection system. Then, it proposes a routing measurement model to balance data traffic named data balance measurement model (DBMM). Subsequently it proposes a routing algorithm based on DBMM, named RA-DBMM. The RA-DBMM handles the data traffic to achieve data balance by modifying the routing decision measurement function. So the queue length of queuing data and buffer capacity of every nodes are taken into consideration. Simulation illustrates that the proposed RA-DBMM has biggish enhancement in minishing the congestion, reducing loss ratio, overcoming bottleneck and increasing throughput of network.**

Keywords—Smart grid; Data collection; Routing algorithm; Data balance; Available capacity

I. INTRODUCTION

Electric power communication network plays an important role to exchange information in electric power system [1]. Smart grid has some peculiar characteristics compared to traditional telecomm network, such as the larger network scale, higher demand for timeliness [2], lower mobility and relatively fixed network topology. The data collected by smart meters is delivered to the gateways by multi-hop. The meters closer to gateways will maybe transmit more data, thus these meters easily lead to long queue length and congestion. Therefore, a reasonable data collection routing mechanism is an important issue to be addressed in electric power communication network.

There have been many studies about improving routing mechanisms for data collection. The destination-sequenced distance-vector routing mechanism (DSDV) is introduced in [3].The DSDV routing mechanism that each node maintains a routing table, is based on Bellman-Ford algorithm. However, it cannot effectively balance data traffic and avoid congestion in some key nodes. The concept of the Max-Weight algorithm for a multi-hop network was introduced in [4]. It will schedule any packet through a specific route according to the queue-length difference of each single-direction single hop link to achieve the purpose of delay balance. However, due to a lack of contribution from the hop-count and queue length to the routing decision measurement, Max-Weight algorithm cannot guarantee that all data gets through the optimal route. The authors of [5] take the data traffic balance method based on ring routing, to avoid congestion of critical node in the central area. And in [6], a probability-based data traffic balance method is put forward, by way of using multiple gateway nodes as forwarding nodes, to avoid the congestion of gateway. It cannot consider data balance of the whole network.

In this paper, it takes the DSDV routing mechanism as reference. Firstly, the network model and mathematical model of smart grid data collection networks are established. Then this paper puts forward the data balance measurement model (DBMM), and routing algorithm based on DBMM (RA-DBMM) by modifying the routing decision measurement function to take the queue length and buffer capacity of each node into account. Next, the MATLAB is adopted to simulate the transmission failure rate and delay performance of RA-DBMM and Bellman-Ford algorithm used by original DSDV. According to the simulation results, it is concluded that the characteristics of the RA-DBMM compared with Bellman-Ford algorithm used by original DSDV can effectively reduce network congestion, decrease the transmission failure rate, improve network reliability and throughput, and shorten the delay under heavy load conditions.

II. NETWORK MEDLE AND MATHEMATICAL MEDLE

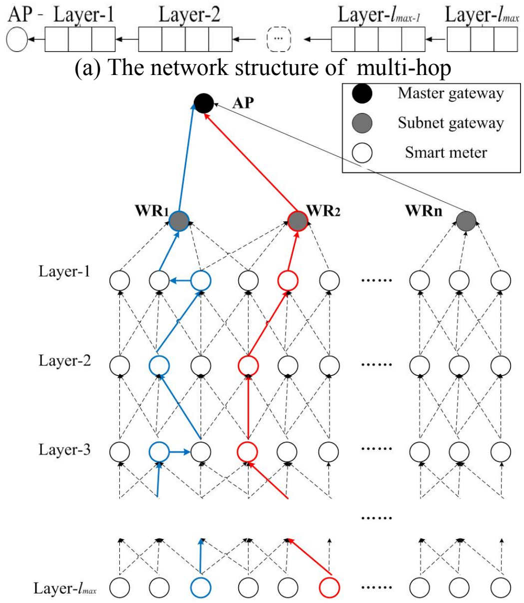
A. Data collection routing mechanism

Routing mechanism for data collection refers to that conduct routing discovery and maintenance processes through routing algorithm efficiently and reliably. Smart grid in data collection system mainly consists of smart meters and gateways. Gateway can be divided into master gateway, that is, the intelligent access point (AP) and subnet gateway, that is, the wireless router (WR). Smart meter is denoted as a node in the network structure. In WMNs, as long as the distance is within a certain range, the communication can be realized between two nodes.

All the nodes are incorporated into a unified and connected network, so that all sub-networks are combined into a large

This work was supported by the National High Technology Research and Development Program of China (2012AA050801)

unified network with multi-gateway. In this unified network, smart meter of any sub-network can access to any reachable gateway of multiple gateways based on local data traffic variation [7]. Further, since the communication distance between nodes is limited, therefore the multi-hop communication method should be taken [4]. So smart meters can directly communicate with the available gateway, and can also transponder data to gateway through other smart meter acting as a router.



(a) The network structure of multi-hop

(b) The routing process of network

Fig.1. An example of a network structure

In this paper, data collection network of smart grid consists of at least one AP node, several WRs subnet nodes connected to at least one AP and a number of smart meters nodes. In order to ensure that any smart meter has at least one routing path to the AP, it is assumed that all the smart meter nodes are distributed within a region that has $l_{\max}$ distance from any WR node, and $l_{\max}$ means the hop count of the furthest meter node away from WR node.

This paper classifies smart meter nodes based on the minimum number of hop count from any one WR node. All the smart meters directly connected to a WR are the layer-1 nodes, that is, the smart meter from any WRs in one hop. All the smart meters connected to the layer-1 node are the layer-2 nodes, that is, the smart meter from any WRs in two hops. In this way, we define the smart meters away from any WRs node in *n* hops as layer-*n* nodes. Each node maintains its initial data queue, that is, historical data. The network structure is shown in Fig.1.

B. Mathematical model

In this paper, the network structure is expressed as (N, L). The symbol N is all the communication devices in the entire network. The L represents all the available single-direction communication routes. If the nodes $u, v \in N$, (u, v) represents a link from node *u* to node *v*. Although all the destination nodes of data packets in the data collection network are the master gateway AP, the AP is only connected to all subnet gateways, WR, thus data packets must be forwarded to the destination node via WR. So N is divided into two sets of nodes, one is WR, the subnet gateway nodes collection as destination nodes to forward data and the other is represented as M, the smart meters that can both send and forward data. Each node maintains a historical record of itself, mainly including queue length and available capacity, denoted respectively as $L(v)$ and $A(v)$. When the $A(v)$=0, the node *v* will become the bottleneck of the whole network, and continuing to send data to node *v* will lead to failed transmission.

In addition, a transmission cycle is taken as the smallest unit of time in this paper. Assuming in the initial state Times=0, which means it has N transmission cycles when Times=N. Data packets being transmitted can only in single hop in each transmission cycle. Assuming that in a transmission cycle, any node can send only one packet. The process of sending and receiving is independent of each other. The symbol *F* is the collection of the data traffics. Any one data traffic is denoted as *f*. *S(f)* represents the source node sending the data packet traffic. $Data_f[t](f \in F)$ represents the number of packets entering the network carried by *f* at Times=*t*.

DSDV using the Bellman-Ford algorithm as routing algorithm, and the number of hop count as measure, therefore when Bellman-Ford algorithm is applied to the wireless routing algorithm, weight denoted as *w* decided only by the reachability between two nodes. That is *w(u,v)=1,* when node *u* connected with node *v*.

$$w(u,v) = hop(u,v) = \begin{cases} 1, & \text{connected} \\ \infty, & \text{not connected} \end{cases} \tag{1}$$

The shortest routing path from the source node *s* to the destination node *d* is the path comprised by the minimum number of hop count. So metric between two nodes is defined as follows.

$$\text{Metric(u,v)} = H_{u \to v}^{\min} \tag{2}$$

III. ROUTING ALGORITHMR BASED ON DATA BALANCE

A. Routing measurement model

In this paper, the concept of available capacity is introduced to help DSDV choose route that can overcome congestion and avoid overflow, in order to ensure the reliability of the entire network. The definition of available capacity is the capacity have not been occupied that still can be used to cache and forward data coming. The total capacity of node *v* is denoted as *C (v)*, and the queue length that the capacity occupied as *Q (v)*, thus it can be inferred that the calculation of available capacity is *A (v)*.

$$A(v) = C(v) - Q(v) \tag{3}$$

That is, the available capacity of a node is the total capacity minus the queue length of queuing data.

We adopt the way modifying the routing decision measurement function. It should not consider only the number of hop count as a metric, but also the queue length of queuing data and buffer capacity. In this model, the available capacity *A* and hop count are integrated as the routing metric for selecting appropriate routes.

This model defines a formula determining the weight between two nodes. When the queue length of a node reaches or exceeds its buffer capacity, if allowed to continue to transmit data to the node, it will form a data overflow, even leading to transmission failure. In order to select the next hop node with larger available capacity, the weight between two nodes should be inversely proportional to the effective capacity of the next hop node. Thus the weight can be quantified as:

$$w(u,v) = hop(u,v) \times 1/\mathrm{A}(v) \tag{4}$$

Meanwhile, in order to make the weight formula meaningful, some appropriate corrections of defining the available capacity are made as follows:

$$\mathrm{A}(v) = \begin{cases} \mathrm{C}(v) - \mathrm{Q}(v), \mathrm{Q}(v) < \mathrm{C}(v) \\ 0.1, \mathrm{Q}(v) = \mathrm{C}(v) \end{cases} \tag{5}$$

After redefining the weight, measurement model between any two nodes can be quantified as:

$$Metric(S,D) = min \sum_{u \to v}^{S \to D} w(u,v) \tag{6}$$

According to formula (6), when the queue length of next-hop node grows, this measurement model can balance data traffic effectively. This measurement model is named as data balance measurement model, that is, DBMM.

B. *Routing process*

In order to make DBMM better used in DSDV routing protocols, the structure of routing table should be modified. The hop fields of routing table maintained by each node will be replaced by the synthesized metric used to determine the next hop node. Furthermore, in order to complete the routing process, the available capacity should be updated appropriately.

The available capacity can never be a constant value. Because every transmission will change the capacity and this change will alter this node and link weight between it and the previous node of transmitting data, thus updating information is extremely important for a routing algorithm. The updating scheme for RA-DBMM based on updating scheme of DSDV is founded in this paper.

When selecting route for the transmission of data packets, the process of updating the capacity of the node triggered in the routing table updates is first updated. That is, each time the data transfer is completed, the metric field in routing tables of related nodes is first modified. Subsequently, the triggered routing tables are updated, and the determination of the next hop node is based on the updated metric field in order to ensure that the latest information of the available capacity.

IV. RERFORMANCE EVALUTION

In this section, to verify the RA-DBMM can effectively overcome the network bottleneck existed in data collection network of smart grid, the performance of RA-DBMM and Bellman-Ford algorithm used by original DSDV in data collection routing process with matlab are simulated. The simulation is done in a 100m × 100m square network topology environment. In this network, the effective communication range of a node is 20m. Topology consists of a master gateway, four subnet gateways, 100 smart meters and all communication links between any two nodes, as shown in Fig.2.

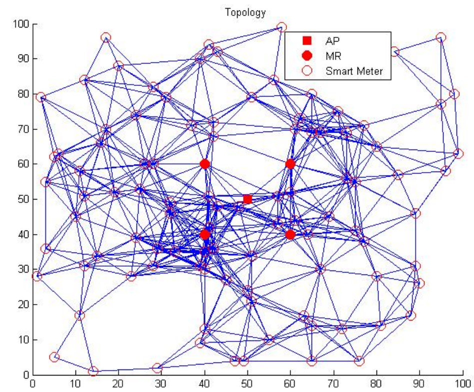


Fig.2. Topology of data collection network

A. *Transmission failure ratio*

Transmission failure ratio can be defined as the ratio of lost packets to all the sent packets. When congestion occurs in some key nodes, the length of queuing data is greater than the buffer capacity of the node, transmission failure will be generated. Transmission failure ratio directly reflects the reliability of the entire network, thus the main purpose of this algorithm is to reduce the transmission failure ratio. Meanwhile, when the network topology and the whole data traffic are the same, the case that transmission failure ratio reduces means packet delivery ratio increases, therefore the network throughput also grows.

Fig.3 shows the transmission failure ratio of Bellman-Ford algorithm and RA-DBMM with the changes in network load. The improvement of transmission failure ratio made by the RA-DBMM according to the Bellman-Ford is shown in Table I. In the beginning of the simulation, the network load is light, the transmission failure happens significantly later in RA-DBMM compared to Bellman-Ford algorithm. In case the network load is below a certain threshold, RA-DBMM can significantly improve the problem of network bottlenecks. As shown in Table I, the improvement rate of transmission failure ratio is as high as 80.28%, and almost overcome bottlenecks. With the smart meter continuously generating data packets, the network load increases, the curve of transmission failure ratio becomes stabilized when the network load exceeds the capacity threshold of network. And the transmission failure ratio of RA-DBMM proposed in this paper is still significantly lower than the Bellman-Ford algorithm. As shown in Table I, the improvement rate of transmission failure ratio gradually stabilizes at 16.75%. In the same network load, the transmission failure ratio curve of Bellman-Ford algorithm stabilizes earlier than RA-DBMM, therefore RA-DBMM is better able to maintain a reliable network , while provides a greater data throughput.

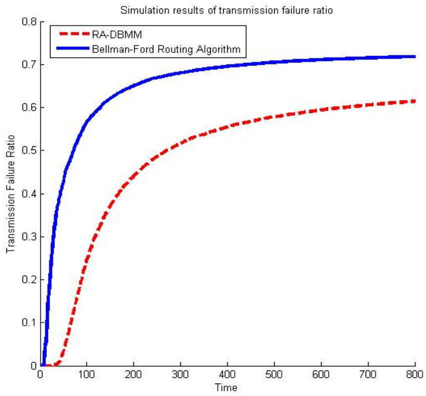


Fig.3. Simulation results of transmission failure ratio

B. Average delay

The total delay consists of transmission delay, propagation delay, nodal processing delay and queuing delay. Since propagation delay and nodal processing delay are relatively fixed and have little effect on the total delay, thus in this paper the impact of propagation and nodal processing delay is not considered. The transmission delay and queuing delay have become the two most important factors that affect the performance of the network average delay. In this paper, transmission delay is mainly associated with the hop count, and queuing delay mainly consists of the length of queuing data. Assuming that the total delay from a node to the main gateway node is the data flow D_f, the total delay is defined as follows:

$$D_f = \sum_{(u,v)\in f} Q(v) + hop(u,v) \tag{7}$$

Fig.4 shows that the average delay of the RA-DBMM is less than that of the original Bellman-Ford algorithm. While at the beginning of the lighter load, the average delay of RA-DBMM is slightly more than Bellman-Ford routing algorithm. As shown in Table I, the average delay of RA-DBMM compared to Bellman-Ford routing algorithm increases 13.34 %. But under light load conditions, the overall delay is small, thus the absolute delay value of RA-DBMM compared to Bellman-Ford routing algorithm is not so much, so the overall performance of data collection network will not have much impact. When the load increases, the average delay of RA-DBMM gradually becomes less than the Bellman-Ford routing algorithm. As shown in Table I, the average delay of RA-DBMM compared to Bellman-Ford routing algorithm reduces 25.97%. And the average delay curve of RA-DBMM is more stable. The average delay from each layer hierarchical node to the AP node is stable. Thus, delay performance can be planned by choosing reasonable location of MRs and AP to avoid large and uncontrollable delay fluctuations due to congestion in some key nodes in Bellman-Ford algorithm

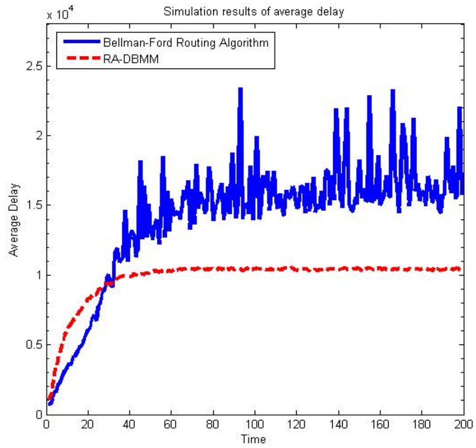


Fig.4. Simulation results of average delay

TABLE I. PERFORMANCE IMPROVEMENT OF RA-DBMM COMPARED TO BELLMAN-FORD ROUTING ALGORITHM

Load Situation	Transmission failure ratio improvement ratio	Average delay reduced rate
Light Load	80.28%	-13.34%
Heavy Load	16.75%	25.97%

REFERENCES

[1] Fan Z, Kulkarni P, Gormus S, et al. Smart grid communications: Overview of research challenges, solutions, and standardization activities[J]. Communications Surveys & Tutorials, IEEE, 2013, 15(1): 21-38.

[2] Wang W, Xu Y, Khanna M. A survey on the communication architectures in smart grid[J]. Computer Networks, 2011, 55(15): 3604-3629.

[3] C. E. Perkins and P. Bhagwat, "Highly Dynamic Destination Sequenced Distance Vector Routing (DSDV) for Mobile Computers, AGM SIGCOMM'94, pp. 234-244.

[4] Gharavi H, Xu C. Traffic Scheduling Technique for Smart Grid Advanced Metering Applications[J]. Communications, IEEE Transactions on, 2012, 60(6): 1646-1658.

[5] Bhaya G, Manoj B S, Murthy C S R. Ring based routing schemes for load distribution and throughput improvement in multihop cellular, ad hoc, and mesh networks[M]//High Performance Computing-HiPC 2003. Springer Berlin Heidelberg, 2003: 152-161.

[6] Hegarty D F, Kechadi M T. Topology preserving dynamic load balancing for parallel molecular simulations[C]//Supercomputing, ACM/IEEE 1997 Conference. IEEE, 1997: 36-36.

[7] Gharavi H, Hu B. Multigate communication network for smart grid[J]. Proceedings of the IEEE, 2011, 99(6): 1028-