Congestion Aware and Adaptive Service Discovery for MANET

1S. Pushpalatha and 2P. Jaganathan

1Associate Professor, PSNA College of Engineering & Technology, Department of Computer Science and Engineering, Dindigul, India
2Professor, PSNA College of Engineering & Technology, Department of MCA, Dindigul - 624622, India

ABSTRACT

Background: The essential requirement of Service discovery protocol in Mobile Ad hoc network is to provide quality of service guarantee to the service clients. The nodes involved in the Group based service Routing Protocol in peer-to-peer network have the function of service requesting, service providing, service forwarding and caching the service information. Each node has the responsibility to balance their load while transmitting more information. The overloaded service provider and the overloaded network lead to congestion. Objective: The local queue length information is used to select the non-congested neighbor and service providers. Results: The resources such as bandwidth and energy are efficiently utilized to improve throughput and delay. Extensive simulations are carried out using NS2 for proving higher throughput compared with the existing system. Conclusion: The proposed cross layer congestion aware service discovery normalizes the network traffic and evenly distributing the service requests to the service providers. Also, it reduces the energy consumption of service providers and intermediate nodes.

INTRODUCTION

Mobile ad-hoc network is a dynamic topology based network in which services, either hardware or software, are likely to be shared automatically. Many service discovery protocols have been designed with various routing protocols to find the route to the relevant service. As there is a need for sharing real time multimedia applications, Quality of Service (QoS) based efficient service discovery protocols are required in dense and scalable network. The delay sensitive applications such as video on demand, interactive games would require minimum service accessing delay to satisfy the end users. The QoS is directly connected with Quality of Experience (QoE) of end user who demands the improved throughput.

In order to provide QoS and QoE based invocation, the service provider and the service accessing route should not be overloaded. As many nodes are sharing the network and each node has more than one role, it is also mandatory to control the load of the nodes in the network. Pure service discovery protocols could not work well when sudden burst of data occurs. The burst traffic will cause congestion. Because of congestion the problems like long delay, high overhead and low throughput occurred. Therefore, our aim is to build the congestion free service discovery protocol.

Group based service discovery (GSD) uses DARPA markup language to describe the services on nodes in the network. Services are classified into several groups in GSD. The service group information of the nodes is cached in its neighbours. Multiple instances of the same service may exist on different nodes. So, each node in the network maintains a cache called as service Information Cache (SIC). This protocol identifies the functionally similar service for the required service request. When a request packet is to be forwarded, the node intelligently select forwarding nodes based on the information from SIC. The service requests can be forwarded to an instance of the services rather than to a particular node address.

Group based service routing is a cross layer Group service discovery. It is integrated with Ad hoc On demand Distance Vector Routing Protocol (AODV) called as Group based Service Routing (GSR) to find the services in the peer nodes. The integration of route discovery, route reply and route maintenance of AODV with discovery mechanism, results in reduced overhead. The popular services are always accessed by many client nodes. But, the loads among various service providers are not evenly distributed. The heavily loaded nodes cause congestion which lead to packet loss, packet retransmission, more energy consumption and
control overhead. If the next hop neighbor node could not be identified, CSREQ is broadcasted in the network. This kind of broadcasting floods CSREQ packets in the network. This may cause congestion in the network. Hence, there is a need for congestion adaptive mechanism in GSR to improve QoS. The objective of integrating congestion control in service discovery is to:

(i) Reduce service invocation delay
(ii) Reduce energy Consumption by avoiding retransmission of lost packets
(iii) Extend the lifetime of mobile nodes.

Related work:
It is getting more and more important for mobile users to discover their required services (e.g. printers, fax machines, multimedia applications and software services) without any reconfiguration. Many works have been done by integrating the service discovery with network layer using cross layer approach. They address the architecture of service discovery, but they do not concentrate on quality of service invocation.

Dipanjan Chakraborthy et al (2006a) have proposed a semantic service discovery scheme, Group based Service Discovery (GSD), in which the nodes intelligently forward the requests based on its service description semantics. In peer-to-peer network, Service request packets are selectively forwarded to accurately locate the service with the help of Service Information Cache (SIC). It also reduces traffic by limiting the announcements to minimum number of hops. The number of service request packets are further reduced by the optimized methods such as PCPGSD (PFCN, CRN and PRN enhanced GSD protocol), CNPGSDP (Candidate Node Pruning Group based Service discovery Protocol), FNMGSN (Forward Node Minimization enhanced Group-based Service Discovery Protocol) mentioned in the articles of Gao, Z.G et al (2006a), Gao, Z.G et al (2006b), Zhenguo Gao et al (2011). Dipanjan Chakraborthy et al (2006b) have proposed Group based Service Routing (GSR), a new routing and session management protocol which integrates service discovery and routing. GSR also combines transport layer features and provides end-to-end session management that detects disconnections, link and node failures and enables session redirection to handle failures. The broadcast storm problem in the above mentioned methods causes congestion. In the article of Zhenguo Gao et al (2006c), a cross layer service discovery protocol, Forward Node Selection based Cross-layer Service Discovery Protocol (FNSCSDDP) is proposed in which the hello packets are used for backbone construction to facilitate service discovery tasks. The forwarding nodes are selected based on the information collected from local topology.

KinWah Kwong et al (2006) have proposed a congestion aware search protocol in which congestion control and object discovery are integrated using Congestion Aware Forwarding, Random Early Stop and Emergency Signalling for unstructured P2P networks. This congestion based approach reduces hit delay and increases hit rate. It does not concentrate on the load of the service provider and does not ensure successful service invocation. Kaouthier Abrougui et al (2012) have proposed a service discovery protocol integrated with routing protocols for vehicular ad hoc networks to guarantee load balancing on routing paths between service providers and service requesters. The QoS aware location-based service discovery protocol shows that the improvement in response time and bandwidth usage while satisfying client requests for various QoS requirements.

Christopher et al have (2013) proposed a cluster based QoS-aware service discovery architecture using swarm intelligence. In this architecture, the client sends a service request with its required QoS parameters such as power, distance and CPU speed to the source cluster head. Intra and inter cluster shortest path routing is implemented using Swarm intelligence. Each cluster head maintains the QoS constraints in the service table and server table. These tables are used to select the QoS aware server. Hassan Artaf et al (2014) have proposed a cluster based QoS aware service discovery in which the cluster heads maintain the directory of services with QoS scores of service providers from service requestor nodes. The service requestor requests the data with QoS data of the required service.

Puri et al (2009) have proposed congestion avoidance and load balancing approach. Congestion is one of the reasons for a route break due to excessive load on the network and leads to failure of nodes and route change in the ad-hoc network. Selection of less congested routes improves the overall network performance. Ad hoc On-Demand Multipath Distance Vector (AOMDV) selects a path with a lower hop count and discards routes with higher hop count. The proposed AODV-Multipath preserves the higher hop count routes in the routing table and utilizes it as alternate path when link failure occurs. Queue length indicates congestion level of a node in the network. Queue length and Hop count value are used together to select a route to avoid congestion and to do load balancing. When the shortest path is congested, then the alternate longer path can be selected. The broadcasting of route request packets is rejected by intermediate nodes if the routes are already congested. This protocol avoids congestion and balances the load in the network. This work improves the packet delivery ratio, increases throughput by avoiding link failure. But, it increases average delay.

S. Santhosh baboo and B.Narasimhan (2009) have proposed a hop by hop congestion control mechanism by combining data-rate, Buffer queuing delay, Link Quality and MAC Overhead as Routing
metric. SenthilKumaran T and Sankaranarayanan (2011a) have proposed an Early detection congestion and self cure routing mechanism to find non congested path between source and destination. This path is the primary route. Each node has two routing tables, primary table and neighbors table. The primary table is used to direct packets on the primary route. A primary path of a node predicts the congestion status and periodically broadcasts a Congestion Status Packet (CSP) to its neighbors with Time To Live (TTL) = 1.

Senthil Kumaran T and Sankaranarayanan (2011b) have presented an Early congestion Detection and Adaptive Routing (EDAPR) protocol which detects congestion at a node level by calculating queue status value and finding congestion status. It constructs Non-Congested Neighbors set which contains both one hop and two hop neighbors and initiates the route discovery procedure by using NHN set to find a non-congested path to a destination. After the route discovery, data packet is transmitted to the destination. EDAPR reduces overhead to find congestion free path and decreases the flooding packets. It requires additional space in the nodes to store NHN table. Bisengar Ahmed et al (2012), have proposed Adaptive Mobility aware AODV (AMAODV) for heterogeneous Mobile Ad-hoc Networks (H-MANETs). H-MANETs are composed of nodes with different transmission range. AMAODV is based on the routing metrics such as distance, relative velocity, queue length and hop count. This protocol considers distance, relative velocity, queue length and hop count between each node and one hop neighbor.

So far, little work has been carried out on congestion aware service discovery. In scalable and dense network, the network traffic leads to packet loss which initiates packet transmission. So, the load of the node is considered for forwarding the packets to avoid packet loss and retransmission. This will improve the performance of the service discovery architecture. The proposed work is analyzed using GSR protocol.

**Early congestion detection technique:**
The average queue length is used to predict the congestion well in advance. The average queue length is calculated as

\[ \text{avgqnew} = (1 - w_q) \times \text{avgqueold} + w_q \times \text{curr\_que} \]  

In eq (1), \( w_q \) queue weight, is a constant (\( w_q = 0.002 \) from RED queue experimental results (Floyd, 1997))

Minimum and maximum threshold values for the queue length are used to control the traffic.

Minimum threshold =35% of packet size
Maximum threshold = 2 * Minimum threshold
Warn Line (WL) = queue length / 2

Transient congestion need not result in a significant increase in the average queue length. As a result, the average queue length (\( \text{avgqnew} \)) changes much slower than the current queue length (\( \text{curr\_que} \)).

**Proposed work:**
The QoS parameters such as throughput, invocation delay and energy utilization are essential metrics for accessing real time multimedia applications in MANET. In addition to that, it is mandatory to extend the lifetime of the Service Provider (SP) by consuming less energy during the service discovery. In order to reduce energy consumption, the number of retransmissions has to be avoided. Balancing the network traffic among the nodes reduces energy consumption and extends the lifetime of the nodes. Congestion-Aware and Adaptive Group based Service Routing (CAA-GSR) protocol is to balance the network traffic load only depending on the local queue length information. Congestion aware Service Request (CSREQ) packet forwarding approach would identify a lightly loaded Service Provider (SP) with the information available in the Service Information Cache (SIC) and average queue length of the forwarding nodes. In this work, it is assumed that the path to the popular SPs is almost congested. When the accurate required service is not available, the equivalent service in the congestion free path is preferred. A slight modification is added in GSD to accept a similar service in case the accurate service is not available in the congestion free path. Instead of forwarding the packet to a particular service, the CSREQ is forwarded to all services in the group. Any one of the relevant service is selected in the congestion free path. When non congested path is selected for forwarding CSREQ packet, the CSREQ loss is reduced and overloading the path to SP nodes is also avoided. It avoids retransmission of CSREQ packets and reduces control packets overhead.

**Methodology:**
(i) Average Queue length aware packet forwarding approach
(ii) Service Request Queue Maintenance.

It consists of four phases.
(a) Congestion aware Service Request (CSREQ) packet forwarding
(b) Admission Control Mechanism at Service Provider
(c) Congestion aware Service Reply (CSREP) packet forwarding
(d) Congestion aware Service Invocation (CSINV)

**a. Congestion aware Service Request Packet Forwarding:**
When a Service Client (SC) requires a service (\( S_i \)), it checks its SIC to get the list of neighbour nodes having the information about \( S_i \). The information may be a service Group (\( G_i \)) or a service
The SC node generates a CSREQ packet and forwards to the selected neighbours (N_j^i). When the neighbor node receives CSREQ, it calculates average queue length (q_{av}(t)) of the node in order to know about the congestion status. The average queue length information provides the total flow information about the nodes in the service request path towards the Service Provider (SP). If q_{av}(t) is less than the minimum threshold (T_{min}) and current queue length is less than the minimum threshold (T_{min}), the neighbor list is identified from the SIC for the required service. Then the CSREQ is forwarded to the selected neighbours and it continues until a Service Provider (SP) node is identified. The reason for selecting minimum congested path is to avoid packet loss during service invocation phase. The SP also checks the average queue length and current queue length to store the CSREQ in Service Request Buffer (SRB).

### Table 1: Variables used in this paper.

| q_{av}(t) | Queue length of node j at time ‘t’ |
| N_j^i | Average queue length of node j at time ‘t’ |
| \( N_j^i \) | Set of next hop neighbours destined for service ‘s’ |
| A_i | External arrival packets to node j |
| T_{max} | Maximum threshold value of interface queue |
| T_{min} | Minimum threshold value of interface queue |
| \( \alpha_j \) | Fraction of time that link j spends serving node j |

The load of link at each node at time ‘t’

\[
L_i = \frac{q_{av}(t)}{BW} \quad (2)
\]

In equation (2), BW is bandwidth of the shared medium. The network is stable when \( L_i < 1 \).

**b. Admission control Mechanism at Service Provider:**

When CSREQ is received by a SP, it checks its current queue status. If \( q_{av}(t) < T_{min} \) and \( q_{av}(t) < T_{max} \), CSREQ is placed in the SRB. The CSREP is sent to SC in the reverse route of service request. The number of CSREQ in SRB is limited based on the transmission rate and bandwidth availability of SP.

Number of CSREQ in SRB at time ‘t’ is

\[
N_{csreq}(t) = \text{Bandwidth available /transmission rate of the service} \quad (3)
\]

The static bandwidth allocation method is applied in this protocol in which each node gets fixed bandwidth and the requests are accepted based on the bandwidth availability.

**c. Congestion aware Service Reply (CSREP) packet forwarding:**

It is assumed that CSREQ and CSREP use the same path. The intermediate nodes, which receive CSREP, create an entry in the Forwarding Table (FT) for the new session. FT consists of the following fields.

[FT-entry]:<Service Client-id, Service Provider id, Service-Request-id, Predecessor-id, session-id >>

Service Client-id is sender of service request; Service Provider id is the node providing the service, Service-Request-id is requested service, CSREQ path is included in CSREP packet and Predecessor-id is extracted from CSREP packet. Predecessor-id is node from which CSREQ is received and is used to forward the service reply in the reverse route. session-id is to identify the session, a unique identifier.

**d. Congestion aware Service Invocation (CSINV):**

When the SC receives more than one CSREP in the congestion free path, it selects any one of the relevant service. The service is selected in the congestion free path by considering the service invocation delay. The selected path is the QoS path which provides minimum packet loss and minimum delay.

**Illustrative Example:**

In figure(1), there are 14 nodes and 3 group of services such as a, b and c. The group ‘a’ has three services a_1, a_2 and a_3. The group ‘b’ has three services b_1, b_2 and b_3. The group ‘c’ has three services c_1, c_2 and c_3. When node A requests service a_1, it sends CSREQ to the selected neighbors from its SIC. Then the neighbor nodes B,C,D receive CSREQ. They check their average queue length. If \( q_{av}(t) < T_{min} \) \&\& \( T_{min} < q_{av}(t) < T_{max} \), send CSREP to SC in the reverse route. When A receives CSREP, it compares the acknowledged service with the required service. The service which has high resemblance with the required service is accepted and SINv packet is sent to the suitable node. If the required service is not possible to access in the non-congested path, the equivalent service is accessed considering the end to end service invocation delay.
Packet Forwarding Algorithm
1: Receive Packet p;
2: Switch (packet type)
3: {
4:     Case CSREQ:
5:         if (receiving node is SP)
6:             if (ql_{ave}(t) < T_{min}^{S} & ql_{j}(t) < T_{min}^{S})
7:                 Store Sr in buffer
8:                 Send CSREP
9:             else
10:                 drop CSREQ;
11:         }
12:     if (receiving node is not SP)
13:         If (ql_{ave}(t) < T_{min}^{Q} & ql_{j}(t) < WL)
14:             get N_j from SIC;
15:             Forward CSREQ to the set of neighbours in N_j;
16:         else
17:             drop CSREQ;
18:     Case CSREP:
19:         forward ‘p’ to the predecessor in the reverse route.
20:     Case SINV:
21:         if (session-id is available in FT)
22:             Forward data packet towards SC.
23:         }
24: }

Fig. 1: Example – Congestion Aware and Adaptive GSR.

Simulation results:
Our proposed work is simulated using NS2[19]. In this simulation, the channel capacity of mobile hosts is set to 2 Mbps. The distributed coordination function (DCF) of IEEE 802.11 for wireless LANs is used as the MAC layer protocol. The simulation is done in 800 meter x 800 meter in 100 seconds for sets of 25 mobile nodes. Initial locations and movements of the nodes are obtained using the random waypoint (RWP) model of NS2. Each node moves independently with the same average speed. All nodes have the same transmission range of 250 meters. In our simulation, the minimal speed is 3 m/s and maximal speed is 5 m/s. The pause time is 5 seconds. The simulated traffic is Constant Bit Rate (CBR) of packet size 512 bytes. The simulation is carried out for 100 seconds. There are 3 different service groups (a,b,c) and each group has three different services. 60% of the total nodes provide services and others are user agents. For cross layer service routing AODV protocol is used. For each scenario, the simulation is repeated 5 times and the average is calculated. The advertisement interval time is 10s and the expiry time of the service in SIC is 5s. TTL of CSREQ packet is 10 in this scenario. The size of the service is 100 KB.
Fig. 2: Service Hit Rate of GSR and CAA-GSR

Fig. 3: Service Hit Rate of CAA-GSR with limited acceptance rate.

a. Service Hit Rate:

The number of service requests generated and the number of accepted requests are shown in figure (2). 30, 45, and 60 CSREQs are generated in three different scenarios. The hit rate is less than GSR. Figure (3) shows the number of CSREQ accepted and rejected during the simulation time. When the number of CSREQ is restricted per unit time, the number of packet drop also reduces. The number of service requests accepted by each SP is 4 in 20 seconds duration. As the nodes are sharing the medium, the packets transmitted from the neighbour nodes will be affected by collision during the service invocation. If the number of sessions is limited, the collision is avoided.

b. Packet Delivery Ratio:

There are 3 scenarios for generating 30, 45 and 60 service requests in 100 seconds. The service size of each SP is 100 KB. The packet delivery ratio is shown in figure 4. The packet delivery ratio of CAA-GSR is greater than GSR by considering the queue length.

c. End-to-end delay:

Average end to end delay for CAA-GSR is less than GSR (figure 5). When the number of CSREQ is limited to each SP, end to end delay is less than normal CAA-GSR (figure 6).

Fig. 4: Packet Delivery Ratio

Fig. 5: End to end delay.

Fig. 6: End to end delay (limited).

d. Energy Consumption:

The simulation is done for 100 seconds with 25 nodes. The residual energy is calculated 5 times and the average is taken. The service requests are generated periodically and the average of energy consumption is calculated. This is shown in figure 7 and 8. The initial energy of all the nodes is 100 joules. The 500 KB service is transmitted by the SPs.
The service provider nodes are continuously servicing the clients. In each scenario some service providers are randomly chosen to check the residual energy. From figure 7 and 8, it is shown that the energy is almost evenly consumed at all SPs and the lifetime of the nodes is approximately same except very few nodes. It is assumed that the applications would provide better performance when the service discovery integrated with lower level information. The 500 KB service is transmitted for analyzing the energy consumption.

**Fig. 7:** Remaining energy for 30 requests.

**Fig. 8:** Remaining energy for 60 requests.

**Conclusion:**

The proposed cross layer congestion aware service discovery increases the efficiency of GSR protocol. The service requests are accepted in the non congested path to reduce packet loss. The CAA-GSR normalizes the network traffic and evenly distributing the service requests to the service providers. Also, it reduces the energy consumption of service providers and intermediate nodes. The broadcast storm problem of GSR protocol is also be eliminated by this approach. The average lifetime of the service providers are extended by integrating the service discovery with congestion awareness. Also, the QoS parameters such as successful service invocation and service invocation delay are improved. The residual energy has to be combined with average queue length information to find the service provider for successful service completion as a future enhancement.

**REFERENCES**


International Journal of Computer Science and Information Security, Vol. 3, No.1