

Session Based Admission Control for Congestion Control and Load Balancing in Group Based Service Routing

S. Pushpalatha and P. Jaganathan
PSNA College of Engineering and Technology, Dindigul, Tamil Nadu, India

Abstract: Service discovery is the technology automatically finds services which is crucial in MANET. The proposed research is designed to avoid congestion in cross layer based service routing and to balance the load among the nodes as well as service providers. The session level congestion control scheme accepts a new service request based on the active session count which is complement to packet-level congestion control schemes. The session control is done at the admission level for the UDP sessions based on the remaining bandwidth in the network. It extends the lifetime of service providers and improves quality of service in service routing protocol. This research is an extension of group based service routing protocol by including session based congestion control scheme by extracting the active session information from service information cache, session table and session estimation table. It provides lesser packet loss and higher session completion rate than queue length based congestion control.

Key words: Group based service routing, manet, congestion control, session control, load balance

INTRODUCTION

Mobile ad hoc network, an infrastructure less dynamic network is mandatory for emergency environment. Flexibility and adaptability are obligatory for the service discovery of ubiquitous and pervasive computing applications in MANET. Service Discovery Protocol (SDP) allows mobile nodes to advertise their services and locates the services automatically without user intervention. So far many service discovery architectures have been proposed to propagate the service advertisement in the network and to identify the relevant service (Ververidis *et al.*, 2008). Various service discovery architectures like Distributed Service Discovery Model (DSDM) (Artail *et al.*, 2008), Service Location Protocol (SLP) (Liu and Issarny, 2004) are proposed for MANET.

Group based Service Discovery (GSD) is one such architecture and many versions of GSD protocols such as PCPGSD (PFCN, CRN and PRN enhanced GSD protocol), CNPGSDP (Candidate Node Pruning enhanced Group-based Service Discovery Protocol), FNMGSDDP (Forward Node Minimization enhanced Group-based Service Discovery Protocol) have been proposed (Chakraborty *et al.*, 2006a, b; Gao *et al.*, 2006a, b, 2011).

These approaches improve the efficiency of service request packet forwarding by including little additional information in cache. However, the control packets are further reduced by utilizing the cached information exhaustively. The existing GSD versions neither prevent a

large number of sessions from contending with each other nor block new session from joining the network during congestion. These protocols do not concentrate on improving the quality of service invocation by considering the resources of mobile nodes and wireless network.

MANET has some practical design issues such as limited bandwidth, dynamic nature of topology and decentralized coordination. When many service requests are accepted simultaneously and only a small amount of resources are utilized for each session and the session will go for long time. This affects the flow completion time of service discovery applications. In the mobile environment, session break will occur due to mobility when the session goes for long time. Moreover, congestion occurs due to non-availability of resources for simultaneous service sessions. Congestion caused by a large number of sessions leads to rejection of few numbers of competing sessions. Congestion caused by lesser number of sessions with high transmission rate lead to reduction in data rate. Admission control mechanism can be augmented in service discovery to admit new session in order to avoid congestion. Hence the session control is adapted in the future work to increase successful session completion rate and reduce session completion time. In this research, congestion control based service invocation is done with the help of service information cache and session table. The session-based congestion control is to prevent the network and service provider from becoming overloaded and to ensure that longer sessions can be completed. Service provider

augmented with the session control mechanism is able to ensure a fair guarantee of service completion for any accepted session, independent of a session length. To improve end user satisfaction in future next generation service-oriented networks, the session aware congestion control framework is designed for improving useful throughput in terms of the number of sessions completed successfully per unit of time.

Literature review: Some of the service discovery protocols and congestion aware protocols are explained here. Jinshan Liu and Valerie Issarny, have proposed a framework for QoS-aware service location in ad hoc networks. The framework consists of QoS specification to capture most significant QoS-related properties with minimal computation cost associated with QoS management and a benefit function for evaluating user's perspective and resource consumption. Chakraborty *et al.* (2006a, b) have proposed a semantic service discovery scheme, Group based Service Discovery (GSD) in which the nodes forward the requests using service description semantics. Service request packets are selectively forwarded in the peer-to-peer network to identify and locate the accurate service with the information available in Service Information Cache (SIC).

FNMGSDP (Forward Node Minimization enhanced Group-based Service Discovery Protocol) is to reduce number of next hop nodes by utilising SIC effectively (Gao *et al.*, 2011). Chakraborty *et al.* (2006a, b) have proposed Group based Service Routing (GSR), a new cross layer service discovery protocol in which routing, service discovery and session management are integrated. 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. In the above mentioned methods, there is a chance for parallel service invocation sessions which leads to congestion.

KinWah Kwong and Danny Tsang have proposed a congestion aware search protocol in which congestion control is integrated with object discovery for unstructured P2P networks. It consists of three mechanisms such as congestion aware forwarding, random early stop and emergency signalling. In congestion adaptive group based service discovery, queue length is obtained from MAC layer in order to forward service requests in the network to search for an appropriate service provider. This ensures that the service accessing path is a congestion free path (Pushpalatha and Jaganathan, 2015). The novel search-based session-level congestion control approach is proposed for the Internet, to complement packet-level congestion control scheme. Generally, congestion is caused

by a large number of interacting TCP flows or lesser number of flows with large amounts of data. A combination of Golden Section Search Gradient Ascent algorithm (GSS+GA) and Close-Probe (CP) algorithm is complement to existing packet level mechanisms. Ayari *et al.* (2010) have investigated a novel session-aware admission-control for improving service providers profitability. It consists of session identifier module, admission control module and load management module. A measure of the load includes the usage of the server's resources, such as CPU, memory, network buffers, I/O and application server backlog queue length. It improves the success rate compared to standard request-aware admission-control.

Nafea *et al.* (2014) have proposed the performance and load management of web Servers. To avoid overloading the web server, an admission control scheme is proposed which is based on the class-based priority scheme that classifies customer's requests into different classes. This is implemented using π calculus. In reactive bandwidth-aware node-disjoint multipath routing protocol, routes are determined based on the specified bandwidth and a Session Admission Control (SAC) process permits a session to enter into the network based on the current availability of network bandwidth (Lal *et al.*, 2015). Pham and Nguyen (2015) have proposed a joint route discovery and distributed session control strategy to setup multimedia sessions for adhoc networks.

So far, little work has been carried out on congestion aware service discovery for MANET. In scalable and dense network, the network traffic causes packet loss which leads to packet retransmission. 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 research integrates routing and session based admission control in GSR protocol.

MATERIALS AND METHODS

In this research, Quality of Service (QoS) of service invocation is improved in GSR Protocol. The service invocation delay and service completion rate are the two important QoS parameters in service discovery applications to satisfy the end users. These parameters are improved by controlling traffic in the network. The implicit bandwidth reservation in terms of session control offers stable network or to keep the network traffic in control. The number of service requests is controlled at the intermediate nodes and Service Provider nodes to control the service invocation sessions. The number of sessions accepted at each node depends on the available bandwidth in the transmission

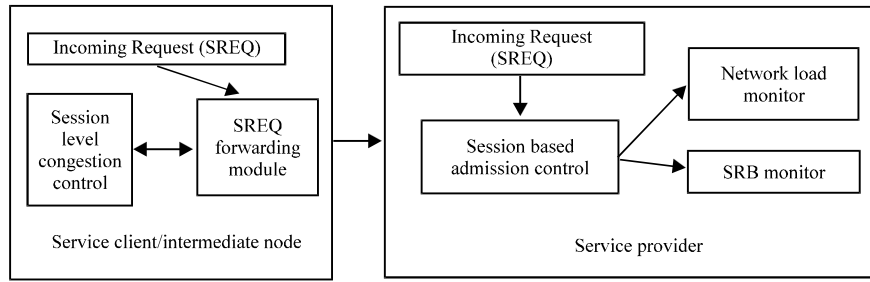


Fig. 1: Session based admission control

Table 1: Notations and explanations

Symbol	Explanation	Threshold values
N_j^s	Set of next hop neighbours destined for service 's'	-
$th(N_{node}^{SS})$	Number of sessions handled by each node	N
$th(N_{tr}^{SS})$	Number of sessions handled in the transmission range	K
S_{ij}	Session from source i to destination j	-
TR_i	Transmission rate of node 'I'	-
BW_{total}	Total bandwidth	-
BW_{data}	Bandwidth used for data session	80% of BW_{total}
$BW_{control}$	Bandwidth used for control packets	20% of BW_{total}

Table 2: Session estimation table

Nodes in transmission range	Number of sessions
A	2
B	1
C	3
D	1

range which avoids congestion. The estimation of the upper limit of number of sessions handled by the nodes based on the resources available at SP node and network resources is done in advance. The service provider nodes and the intermediate nodes of the service path handle session control mechanism.

Session level congestion control: The network traffic is predicted through session table. In order to get the session details of the neighbours of a node 'i', the session table is exchanged among neighbours. From the session table number of sessions used by one hop neighbours are calculated and stored in a Session Estimation Table (SET) of the node. The notation and their explanation are given i.e., Table 1. The fields in SET are:

<< neighbour id, number of sessions >>

Service discovery and session table establishment: Initially, when a Service Client (SC) requires a service (S_i), it generates a Service Request (SREQ) packet and forward the SREQ packet to a set of neighbours (N_j^s) selected from Service Information Cache (SIC) and Session Estimation Table (SET). The neighbor nodes which receive SREQ, forward the SREQ packet if they are not the requested SPs. The intermediate nodes include their identity in SREQ before forwarding it. Once the SREQ reaches the SP, it applies Session based Admission Control (SAC) procedure and constructs Service Reply (SREP) packet. The SREP packet is forwarded to SC in the reverse route of SREQ. The complete route is copied from SREQ into SREP and the intermediate nodes update Session Table (ST) using this information. The fields in the session table are:

For each and every node in SET, mark the node as 'safe node' if $N_{node}^{SS} < N$. The 'safe nodes' list form the Congestion Free Neighbour (CFN) set. When the SREQ packet is received, if the total number of sessions of 1-hop neighbours is $< th(N_{tr}^{SS})$ and the number of sessions of the node is $< (N_{node}^{SS})$, the SREQ packet is forwarded. Each session is assumed to have a fixed transmission rate. The number of bits transmitted should be less than the total bandwidth to avoid congestion (Table 2).

<< session id, SP, SC , Intermediate nodes >>

For fixed transmission rate, The bandwidth required in time 't' depends on the number of sessions at time 't' and transmission rate of a session:

$$Bw_{req} = k \times TR \tag{1}$$

For variable transmission rate:

$$Bw_{req} = k_{i=1}^k TR_i \tag{2}$$

In Fig. 1, the architecture of session based admission control is provided. It shows the session level admission control for service clients, intermediate nodes and service Providers.

In Eq. 1 and 2, $k = 1$ is number of sessions in the transmission range. In this study, only fixed transmission rate is considered. When the number of sessions handled

by the neighbour nodes is known in advance, the arrival rate of SREQ is controlled. The maximum number of sessions in the transmission range is determined as:

$$K = BW_{total}/TR \quad (3)$$

In Eq. 3, K is maximum number of sessions in the transmission range of a node:

$$N = K/n \quad (4)$$

In Eq. 4, N is the maximum number of sessions in each node in the range when the total number of nodes in the range is 'n'. As the number of sessions is fixed in advance, the bandwidth is properly utilized.

Algorithm 1: Session Control based SREQ Forwarding

- At node 'i', Collect session tables from neighbours
- Construct SET
- Receive SREQ packet
- If $N_{tr}^{ss} \geq K$ and $N_{node}^{ss} \geq N$, reject SREQ.
- Else
- Find one hop congestion free neighbour set(CFN)
- $\{CFN-SIC\} = \{\text{neighbours from SIC}\} \cap \{CFN\}$
- Forward SREQ to selected one hop neighbours in set CFN-SIC

In order to forward a SREQ packet to a destination, the SC node selects congestion free one hop neighbour nodes from SET and forwards SREQ to the selected non congested forwarding node set of A. Algorithm 1 explains the SREQ forwarding mechanism. This process is repeated until a SP is identified. From the SP, SREP is forwarded in the reverse route of SREQ. A major advantage of this approach is to identify the non-congested route with minimum number of control packets.

Session based Admission Control (SAC) at service providers:

The session based SREQ forwarding mechanism identifies a lightly loaded SP. The SP node does not accept load beyond its capacity. The Session Request Buffer (SRB) has space to accommodate newly arrived session, SREQ identifier is inserted in SRB. The waiting time in SRB for a service request (S_r) is double the time required for transmitting the service. The SRB monitor module monitors the SRB length to control incoming requests. In addition to that network monitor module monitors the available bandwidth to accept or reject service requests. The waiting time of a request in SRB is represented by $WaitTime_{sreq}(SRB)$ and the time required to transmit a session is represented by $Time_{ss}^{tr}$:

$$WaitTime_{sreq}(SRB) = Time_{ss}^{tr} \times 2 \quad (5)$$

The total number of sessions handled (SH_{count}) by the SP is forwarded in SREP. The SP nodes with minimum

SH_{count} is used to balance the load among the servers. When the SC receives SREP packets from more than one SP through its neighbours, it selects the SP with minimum SH_{count} . Likewise, the load balancing is done to extend the lifetime of the node. Then, the SINV packet sent by the SC, confirms the session and the SP node which does not receive SINV deletes SREQ from SRB:

$$Length(SRB) = SS_{max} \times 2 \quad (6)$$

Number of SREQ packets in SRB is denoted by length (SRB) and the maximum number of parallel sessions transmitted by the SP node is denoted by SS_{max} .

The proposed research is illustrated in Fig. 2. When the node A requires a service 'a_2', the number of sessions handled by that node (N_{node}^{ss}) and the neighbour nodes (N_{tr}^{ss}) are gathered from SET. Then, the CFN set is identified. SIC and ST tables are shown in the Fig. 2. The ST has source, destination and intermediate nodes. Let us assume $N = 4$ and $K = 20$ in this example. N_{tr}^{ss} is of node 'A' is 7 which is <threshold value 'K'. The congestion status of the neighbours is gathered. Then, the congestion free nodes from SIC (CFN-SIC) is determined.

$$\begin{aligned} \{CFN-SIC\} &= \{\text{neighbours from SIC}\} \cap \{CFN\} \\ &= \{B,C,D\} \cap \{B,D\} \\ &= \{B,D\} \end{aligned}$$

The nodes B and D are identified as congestion free nodes and SREQ is forwarded to B and D. When the nodes B and D receive the same SREQ, if they are not SPs, the same procedure is repeated to forward the SREQ. The nodes M and H receive SREQ from B and D respectively. When SP receives SREQ, forward SREP if it is not overloaded. The SC selects appropriate SP, when more than one SREP is received. SH_{count} is taken into account when more than one SP provides the same service. As M and H are service providers of service group 'a', they forward SREP to A. A finds that node M provides the service a_2 and sends Service Invocation (SINV) packet to M. The selected path from node A to M is a congestion free path which ensures the successful service invocation. Also, packet loss is reduced during service invocation in GSR-SC. In order to improve the quality of User Datagram Protocol (UDP), the retransmission has to be handled when packet loss occurs. But, the session control avoids packet loss. The interference of neighbor nodes is not considered in this approach.

Where:

- ' γ ' = Ratio of neighbours which handle minimum number of session
- ' τ ' = Number of hops in the routing path

- Service reply forwarded in the network

$$SREP_{tot} = \sum_{i=1}^{mSp} mSREP \times \tau \quad (15)$$

where, 'mSP' is number of service providers. Total number of control packets in 't' simulation period:

$$Control_{alt} = SADV_{tot} + SREQ_{tot} + SREP_{tot} \quad (16)$$

Additional control packets during alternate route discovery:

$$control_{alt} = \sum_{i=1}^{nSF} nSREQ \times n / \gamma \times \tau \quad (17)$$

where, nSF is number of session failure predicted.

RESULTS AND DISCUSSION

The proposed research is simulated using NS2. In this simulation, the bandwidth of the medium 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 carried out in 1500×1500 m for the sets of 25 and 50 mobile nodes. Initial locations and movements of the nodes are obtained using the Random Way Point (RWP) model. Each node moves independently with the same average speed. All the nodes have the same transmission range of 250 meters. In the simulation, the minimum speed is 3 m/s and maximum 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 run for 500 seconds. 60% of the total nodes provide services and others are the service clients. AODV protocol is modified for service routing. For each scenario, the simulation is repeated 5 times and the average is calculated. Time-To-Live (TTL) of CSREQ packet is 4 in this scenario. The size of the service is 250 KB. The transmission rate of the service at each SP is from 8 packets/sec to 32 packets/sec. The sessions allowed in each node is 6 and in the network 48 sessions. In this research, we focus on session control for UDP flows. Initial energy is 100 J. Service related information is as Table 3.

Session admission ratio and session completion ratio: The simulation is carried out for 100, 150, 250, 300 and 400 sec in

Table 3: Simulation parameters for services

Parameter name	Value
Serevice group	a, b, c
Services	a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2, c_3
Service groups	3
Service providers	60%
Advertisement interval	10s
Expiry time of the service in SIC	5s
TTL of CSREQ	10

Table 4: Session admission ratio and session comple

No. of SREQ admit rate	Session completion rate	Session completion time(s)	Average throughput
4	99.5	51.15	98.6
5	99.5	51.36	98.5
6	99.5	51.58	98.5
7	99.4	51.64	98
8	99.4	51.66	98

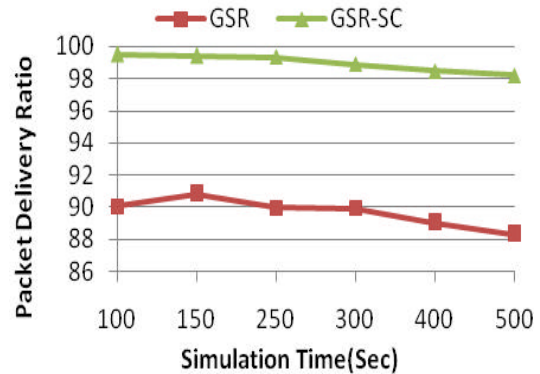


Fig. 3: Packet delivery ratio

different scenarios. Every 5 sec a service request is generated by any one of the nodes. The transmission rate of the service provider is 8 packets per second. The throughput is approximately 98% up to 6 parallel sessions. The remaining 2% loss is due to the mobility of the nodes. The dropped SREQ packets are retransmitted after 5 seconds. The session completion rate is optimized to 99.5%. The remaining 0.5% is due to the mobility of the node which breaks the session. The shorter session completion time is expected in mobile networks (Table 4).

Packet delivery ratio (PDR):

$$PDR = P_r / P_s \quad (18)$$

Where:

P_r = The number of packets received

P_s = The number of packets sent

The simulation is run for 100, 150, 250, 300 and 400 seconds. Every 5 seconds a service request is generated from any one of the nodes. Figure 3 shows the average PDR

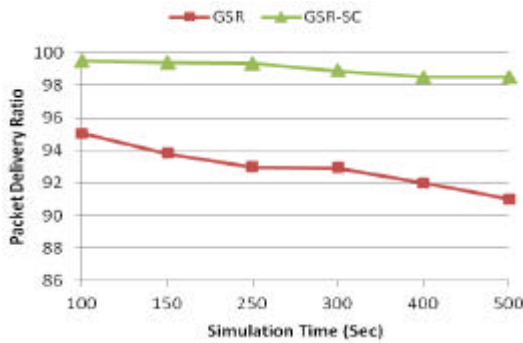


Fig. 4: Packet delivery ratio

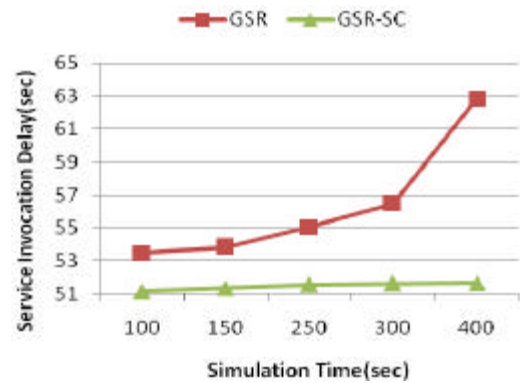


Fig. 6: Service invocation delay

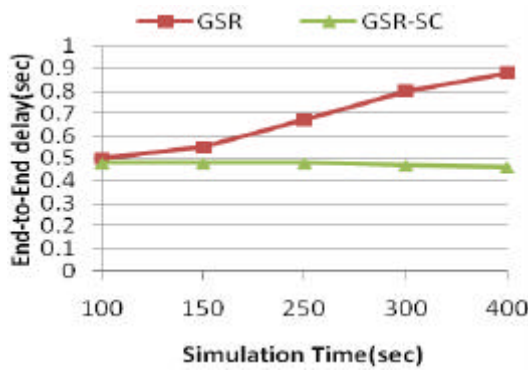


Fig. 5: End-to end delay

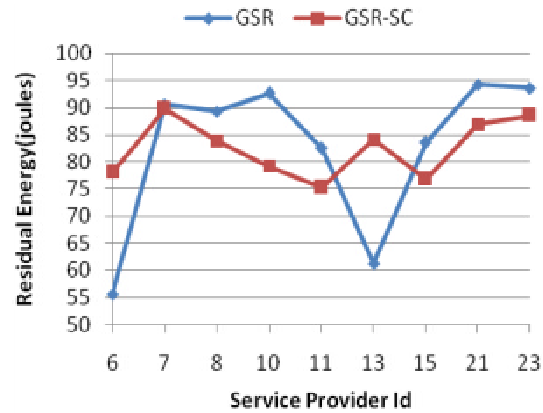


Fig. 7: Energy consumption

of the service sessions. The transmission rate of the service provider is 8 packets per second. GSR-SC has the improvement over GSR up to 48 sessions in the same transmission range. The >48 sessions the packet loss occurs due to unavailability of bandwidth. In order to share the bandwidth among the neighbor nodes, the maximum number of sessions is controlled to 6 per node. In addition to that, packet loss occurs due to the interference of neighbour nodes and mobility of the nodes. In GSR-SC, the number of sessions in the transmission range is controlled to avoid packet loss due to interference. GSR-SC has 10-15% of the improvement in the PDR than GSR. If the transmission rate is increased for the same number of sessions, packet loss increases.

Figure 4 shows the average PDR for transmitting 32 packets per second. When increasing the transmission rate from 8 packets to 32 packets, packet loss increases for the same number of parallel sessions. While increasing the transmission rate, the number of parallel sessions is reduced to avoid packet loss. GSR-SC has 5-10% improvement in the PDR than GSR. A constant number of sessions in the transmission range are followed during all simulation scenarios. The hop by hop session control for service request packet forwarding balances the load of the service providers.

End to end packet delay and service invocation delay: The end to end packet delay is calculated under various timings 100, 150, 250, 300 and 400 seconds in different scenarios and is shown in Fig. 5. Figure 5 and 6 shows the end-to-end delay and service invocation delay respectively. The end-to-end delay for GSR-SC is <GSR. The service invocation delay is directly proportional to end-to-end delay. Since the packets are transmitted without the need for retransmission, service invocation delay of GSR-SC is <GSR. If the quality of the service is negligible, retransmissions are not required.

Energy consumption: The residual energy is calculated for all the nodes and the residual energy of some of the peer nodes are shown in Fig. 7. The simulation is done for 400 seconds. The transmission rate is 8 packets per second. In GSR protocol, some nodes have drained more energy than other nodes. The nodes 6 and 13 have consumed more energy than other nodes. In GSR protocol, the service providers are not selected based on the load handled. In GSR, there is a chance of selecting the same SP repeatedly since there is no constraint in selecting SP. In GSR-SC, the

nodes have consumed energy evenly. The round robin approach and SH_{count} of the node are used to select a SP and this stabilizes the load among the SPs. This shows that the session based admission control leads to load balancing among the SPs.

CONCLUSION

The session level congestion control is used to avoid congestion in GSR protocol. The hop by hop session based admission control utilizes the bandwidth efficiently. The proposed strategy controls the number of competing sessions to optimize for the successful session completion rate and the session completion time. The remaining energy is determined to verify the participation of service providers. The residual energy shows that the load is shared among all service providers. Detailed evaluations of session control algorithms under various traffic scenarios and load changes showed that session control algorithms have benefits over no session control during session congestion. For calculating the number of sessions in the transmission range, only 1-hop neighbour's sessions are considered. The nodes in the interference are not considered. The packet loss can be completely avoided by considering interference nodes.

REFERENCES

- Artail, H., K.W. Mershad and H. Hamze, 2008. DSDM: A distributed service discovery model for manets. *Parallel Distrib. Syst. IEEE. Trans.*, 19: 1224-1236.
- Ayari, N., L. Lefevre and D. Barbaron, 2010. Design and evaluation of a session-aware admission control framework for improving service providers profitability. *J. Internet Eng.*, 4: 260-268.
- Chakraborty, D., A. Joshi, Y. Yesha and T. Finin, 2006a. Toward distributed service discovery in pervasive computing environments. *Mobile Comput. IEEE. Trans.*, 5: 97-112.
- Chakraborty, D., A. Joshi and Y. Yesha, 2006b. Integrating service discovery with routing and session management for ad-hoc networks. *Ad Hoc Netwo.*, 4: 204-224.
- Gao, Z., L. Wang, M. Yang and J. Wang, 2011. FNMGS DP: An optimized group-based service discovery protocol for MANETs. *Wirel. Pers. Commun.*, 57: 137-162.
- Gao, Z., L. Wang, X. Yang and D. Wen, 2006a. PCPGSD: An enhanced GSD service discovery protocol for MANETs. *Comput. Commun.*, 29: 2433-2445.
- Gao, Z., L. Wang, M. Yang and X. Yang, 2006b. CNPGSDP: An efficient group-based service discovery protocol for MANETs. *Comput. Netw.*, 50: 3165-3182.
- Lal, C., V. Laxmi, M.S. Gaur and S.B. Ko, 2015. Bandwidth-aware routing and admission control for efficient video streaming over MANETs. *Wirel. Netw.*, 21: 95-114.
- Liu, J. and V. Issarny, 2004. QoS-aware service location in mobile ad hoc networks. *Proceedings of the 2004 IEEE International Conference on Mobile Data Management*, January 19-22, 2004, IEEE, Berkeley, California, USA., ISBN: 0-7695-2070-7, pp: 224-235.
- Nafea, I., M. Younas, R. Holton and I. Awan, 2014. A priority-based admission control scheme for commercial web servers. *Int. J. Parallel Program.*, 42: 776-797.
- Pham, V.T. and T.D. Nguyen, 2015. Joint routing-distributed session control for highly dynamic ad hoc networks. *Proceedings of the Sixth International Symposium on Information and Communication Technology*, December 3-4, 2015, ACM, Hue City, Viet Nam, ISBN: 978-1-4503-3843-1, pp: 195-202.
- Pushpalatha, S. and P. Jaganathan, 2015. Congestion aware and adaptive service discovery for MANET. *Aust. J. Basic Appl. Sci.*, 9: 265-272.
- Ververidis, C.N. and G.C. Polyzos, 2008. Service discovery for mobile ad hoc networks: A survey of issues and techniques. *IEEE. Commun. Surv. Tutorials*, 10: 30-45.