TMV: Trust-Matrix-Value Based Neighbor Peer Selection for Secure Query Forwarding in P2P Networks

R. Venkadeshan[1](✉) and M. Jegatha[2]

[1] Department of Computer Science and Engineering,
Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham,
Amrita University, Coimbatore, India
r_venkadeshan@ch.amrita.edu

[2] Department of Computer Science and Engineering,
Trichy Engineering College, Tiruchirappalli, Tamilnadu, India
vpragatha@gmail.com

Abstract. Today's internet exploits the P2P network for its unique and special characteristics such as resource sharing, query routing, dynamic topology construction, self-healing in communications and easy and efficient network setup. Moreover, P2P network has undergone many challenges which enable malicious nodes to launch denial of service attacks, allows the adversary peer nodes to deny the query request and even not utilizing its energy for query routing process in the network. These challenges resultant in performance degradation and reduce the query success ratio. In this paper, the authors elaborate the robust trust-based peer communication model that utilizes the scalable topology by building an overlay network with trusted neighbor peers. By having the trustworthiness among the neighbor peer nodes, the privacy, security protection in query searching and success ratio has been increased simultaneously. The main contribution of this paper is to compute Trust-Matrix-Value (TMV) for each peer node, based on the parameters such as query response (QR), resource sharing (RS), information quality (QI) and success ratio (SR). The peer nodes with higher values are considered for further query forwarding and others traffic has been blocked. A comprehensive analysis has been performed and their simulation results proved that the proposed scheme achieves efficient searching with minimal delay, discards by penalizing the malicious peer nodes from the topology, and maximizes the query success ratio in P2P networks.

Keywords: TMV · Trust-Matrix · DoS attack · Security services
Malicious peers · Attacks in P2P · Trust measures

1 Introduction

Peer-to-Peer systems comprise of several distributed applications in which each peer can allow to share their resources with other peers by simple message exchange. The primary goalmouth of the P2P system is to cooperatively sharing the resources among the peers and to aggregates the resources efficiently in order to make available at the

A. V. Deshpande et al. (Eds.): SmartCom 2017, CCIS 876, pp. 82–91, 2018.
https://doi.org/10.1007/978-981-13-1423-0_10

Internet edge. In specific, Gnutella and KaZaA based file sharing of P2P system becomes a most popular model for the interchange of resources among the huge number of internet consumers. The lack of a central controller and non-hierarchical organization of peers make the P2P applications more vulnerable. Without performing the peer authentication, adversary peers can enable various security attacks such as IP spoofing, falsify messages, man-in-the-middle and denial of services (DOS) in the network. A distributed peer-to-peer system [11] should influence resources of all peers such as memory space, bandwidth, and processors to perform query searching and provides better scalability in the network. Besides, in P2P system data has been replicated in multiple peer nodes which provide fault tolerant and there is no single point of failure [13, 14]. Decentralized distributed P2P systems are highly susceptible to Sybil attacks [6], where malicious peer node obtains multiple identities of other peer nodes in the same network (called Sybil node).

To guard against this Sybil attack [4], just monitoring each peer's behavior is not adequate because all the Sybil peers can act smoothly initially, and later unveiling an attack. To overwhelm the above-discussed issues, the work focuses on the functionalities of trustworthiness among the peer nodes and ensures the node privacy while communication. The trust management scheme separates trust peers from the malicious peer nodes, based on the past communication history between them. The formation of trust communities enables each peer node to practice neighborhood of trust for query forwarding which is to defend peer confidentiality in the network.

2 Related Works

In this section, the authors have performed a literature survey based on two primary characteristics: Query searching and secure peer selection mechanisms. In a peer-to-peer network, the resource searching algorithms are hardly divided into two categories: broadcasting method and state knowledge-based searching. By broadcasting approaches like Flooding and Random-walk, the search queries can directly forward to more number of neighbor peers without any constraints. It results in huge overhead in network traffic and does not scale well which incurs variances in performance with minimum success rate and query hits [8]. In flooding based searching technique, querying peer node sends a search query to all its neighbor peers. There is another technique for improving the search efficiency stated in the algorithm [2] is based on query forwarding. Further, in [1, 5], the neighbor peers flood the search query to all of its logical neighbors, excluding the inward bound peers, until the Time-to-Live (TTL) value reaches to zero or the response of the query receives [9].

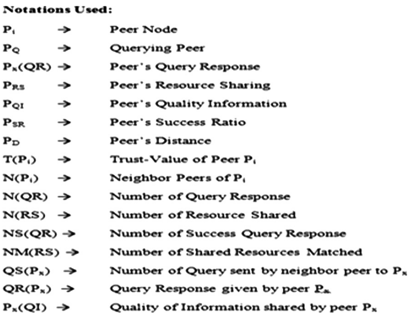
Another approach called Random-k-walker [7], which sends search queries to its *k*-neighbors named "k-walkers", of the querying peer. Every walker selects the next k-neighbor peers randomly then forwards the search query to that walker. This random walker selection mechanism executes until the TTL value expires or receiving the query response. The scope of the search gets an increase as the search mechanism delivered the query request to large peer count, which improves the success ratio [3, 12]. Moreover, the above-stated search mechanism selects the neighbor peer nodes without any selection strategies, so there may be a chance of performance level dissatisfaction. In [16], authors have proposed a searching technique called Query Routing Tree (QRT) in P2P networks. This technique controls the query forwarding hops without loops and reduces the query traffic significantly. Here, the authors have taken the environment as unstructured P2P networks. It is very difficult to compute the cost of the link between communicating peer nodes every time in a dynamic model [2]. In unstructured P2P networks, the peer nodes are highly mobile in nature. So, the cost of the link changes frequently as the topology change. This technique decreases the overall network performance. In [10], authors have proposed a probabilistic prototype to handle trust measures in P2P networks. It performs a local query computation and propagation of trust values of each peer node into modules of other peer nodes. It is well suited for the dynamic P2P networks where each peer has different perspectives towards the other peers with whom it interacts.

Venkadeshan and Chandrasekar [15] have proposed a technique called Peer-ID based authentication scheme in a peer-to-peer network. They have described a mechanism known as Identity-based peer authentication, which authenticates the peer node at the time of entry level into the network. By observing all the above discussion, this paper proposes the Trust-Matrix-Value (TMV) based query routing mechanism which combines the advantages of trusted peer selection and searches performance efficiency in the P2P network.

3 Constructing Trust-Based Topology

In this section, the current work focuses with the construction of a trustworthy environment for communicating peer nodes. In the topology, all the peer nodes perform self-organization of network without any central controller and express their resources with other peers in an autonomous manner. Here, the authors propose a peer selection algorithm called Trust-Matrix-Value (TMV) which constructs trustworthiness among the communicating peers in P2P networks. Each peer node computes trust value of their neighbor peers which used for further query forwarding process. In practical, computing peer's trustworthiness and collecting these trust values in the P2P network is not an easy task. For that, calculate the following characteristics for each peer node: (1) Query Response (QR) frequency, (2) Resource Sharing (RS), (3) Quality of Information (QI), (4) the query Success Ratio (SR), and (5) the transmission peer distance (PD).

Based on these characteristics, neighbor peers with higher TMV value can be targeted for further query forwarding, by this means the network traffic has been routed to trusted zone, the query success ratio is increased and reducing the network traffic overhead significantly. The key factors of this work are, peer selection and query forwarding, each peer should validate the trust value of its neighbor peer nodes before sending or receiving the traffic. The accumulated trust value of the peer nodes is considered and whose value is put down the threshold value is assumed as distrusted and the network traffic through that peer node is blocked. The trust measures are used to evaluate the query traffic from the peer nodes and based on the traffic each peer is updating the trust values of its neighbor peer nodes. In order to increase the efficiency of query search by exploiting the trustworthiness among the peer nodes intelligently,



Notations Used:

P_i	→	Peer Node
P_Q	→	Querying Peer
$P_x(QR)$	→	Peer's Query Response
P_{RS}	→	Peer's Resource Sharing
P_{QI}	→	Peer's Quality Information
P_{SR}	→	Peer's Success Ratio
P_D	→	Peer's Distance
$T(P_i)$	→	Trust-Value of Peer P_i
$N(P_i)$	→	Neighbor Peers of P_i
$N(QR)$	→	Number of Query Response
$N(RS)$	→	Number of Resource Shared
$NS(QR)$	→	Number of Success Query Response
$NM(RS)$	→	Number of Shared Resources Matched
$QS(P_x)$	→	Number of Query sent by neighbor peer to P_x
$QR(P_x)$	→	Query Response given by peer P_x.
$P_x(QI)$	→	Quality of Information shared by peer P_x

Fig. 1. Notations used

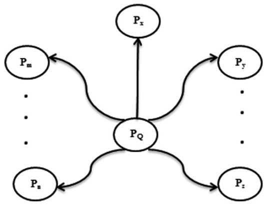


Fig. 2. TMV matrix and topology of querying peer P_Q

the TMV mechanism delivers the high level of security protection for the peer communication. Figure 1 illustrates the different notations used in this trust matrix model.

3.1 TMV Based Query Routing

In Trust-Matrix-Value (TMV) based query routing mechanism, the querying peer node P_Q is handling two different matrixes: Left side-Matrix and Right side-Matrix. The Left side-Matrix called as Bound-Matrix (*BM*), is a $n \times 4$ matrix, where n represents the number of neighbor peer nodes of P_Q and *4* stands for representing the four columns *QR, RS, QI,* and *SR* scores of n neighbor peers of $P_{Q.}$ The four computed scores are placed in the first, second, third and fourth columns of *BM* matrix respectively. The Right side-Matrix called as Distance Matrix (*DM*), is a $n \times 1$ matrix in where n stands for the number of neighbor peers and 1 stands for the entry of neighbor peer nodes distance $D(P_x)$. By having these two *BM* and *DM* matrix, the querying peer P_Q computes the Trust-Matrix-Values (TMV), which results in $n \times 1$ matrix.

Each neighbor peer should assign the computed trust value. The querying peer P_Q uses the T_x resulted in trust value of peer from TMV matrix to choose the neighbors with the top-k scores and forward the search query. Figure 1 listed out the set of notation used and Fig. 2 illustrates the TMV matrix computation process and topology of querying peer P_Q. As depicted in Fig. 2, the neighbor peers of P_Q peer are $\{P_x, P_y, P_z, P_m, and\ P_n\}$.

$$TM(P_Q) = \begin{bmatrix} & & \boldsymbol{BM} & \\ P_x(QR) & P_x(RS) & P_x(QI) & P_x(SR) \\ P_y(QR) & P_y(RS) & P_y(QI) & P_y(SR) \\ P_z(QR) & P_z(RS) & P_z(QI) & P_z(SR) \\ P_m(QR) & P_m(RS) & P_m(QI) & P_m(SR) \\ & & \vdots & \\ & & \vdots & \end{bmatrix} \begin{bmatrix} \boldsymbol{DM} \\ D(P_x) \\ D(P_y) \\ D(P_z) \\ D(P_m) \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} \boldsymbol{TMV} \\ T_x \\ T_y \\ T_z \\ T_m \\ \vdots \\ \vdots \end{bmatrix} \tag{1}$$

3.2 Trust-Matrix-Value (TMV) Calculation

When the search query originated, the querying peer node P_Q calculates the trust value for each of its neighbor peer nodes by using TMV computation. These computed values are treated as the parameter for selecting the appropriate neighbor peers for further query forwarding. The peers with high trust measures are considered as trustworthy peers than other peers. The peer's traffic is either accepted or rejected based on its calculated trust measures by its neighbor peer nodes. Trust-Matrix-Values (TMV) are calculated from the *BM* × *DM* aggregated values, where peers with low trust measures are considered as distrusted and the query traffic for them is consequently reduced. The peers with high trust measures are considered as trustworthy and thereafter the traffic to that peer is fully loaded. This TMV-based query routing mechanism allows the peers to dynamically update the trust values at periodic intervals. Here, the authors use the mean and standard deviation techniques to attain the trustworthiness in a P2P network. Let us consider the data set as {x1, x2, x3 … xn}. Assume that, the mean value *M* of each column in *BM* is computed before calculating the standard deviation of each column of *BM*. It is state

$$\bar{M} = \frac{1}{n} * \sum_{r=1}^{n} BM(r, c). \tag{2}$$

where *n* is the number of neighbor peers of querying peer P_Q, *BM(r, c)* is the r^{th} row and c^{th} column entry of the *BM* matrix. The standard deviation S_D computation as follows:

$$S_D = \int \sqrt{\begin{matrix} \frac{1}{n-1} * \sum_{r=1}^{n} \left(BM(r,c) - \bar{M}\right)^2 & if\, n \geq 2 \\ 0 & if\, n = 1 \end{matrix}} \tag{3}$$

At last, define the *DM* matrix *n* × *1* as

$$DM(c, 1) = \frac{S_D}{\sum_{r=1}^{n} S_r} \tag{4}$$

where *DM(c, 1)* is the weight of the c^{th} column *(c, 1)*-entry.

3.3 Calculation of TMV Parameters

Query Response (QR). In unstructured peer-to-peer networks, several freeloaders are available who utilize the network resources without sharing any of their own resources. This freeloader has been affecting the search performance of P2P network communities. In order to thwart free loaders, *QR* is utilized to make difference among the peer nodes as leech peers and fervent peers. The score of *QR* is computed by each peer node by considering the query response frequency of its neighbor peers. The querying peer

P_Q computes *N(QR)*, the number of query responses for all its neighbor peer nodes. Formally, $P_x(QR)$ is the query response frequency of peer P_x.

$$P_x(OR) = \sum_{x}^{N(P_x)} N(QR) \tag{5}$$

where $N(P_x)$ is the neighbor peers of a querying peer P_Q; that is, $N(P_x)$ are one hop away peers from P_Q. Additionally, querying peer P_Q should compute $QS(P_x)$, is the number of queries sent *(QS)* to the peer nodes that are one step away from P_Q. The number of query response *N(QR)* of the peer P_x is given as,

$$N(QR) = QS(P_x) - QR(P_x) \tag{6}$$

where $QR\ (P_x)$ is the query responses of peer P_x. When *N(QR)* of peer P_x increases, the chance of trust value measures also increases.

Resource Sharing (RS). In P2P networks, an observation has been taken on the aspect of sharing network resources, in which the resource sharing among peer nodes are extremely unbalanced. It has been proven that only limited number of peer nodes is effectively sharing their resources to all other peer nodes and very few percent of peers are properly responding to requested queries. To provide a better success ratio all peers should honestly involve in query forwarding process, but only a small count of volunteer peers are involving which increases the delay in query processing. Besides, each peers' query response abilities vary because of their heterogeneity in resource sharing. In trace analysis has been taken by the authors, shows that only limited peers are sharing their resources effectively. For the reason that, query response involves matching patterns with the searching keywords of all shared resources. Here, the authors improvise the concept as increasing the number of shared files will automatically increase the query success probability. From the above observations, the authors come up with the notion of effective resource sharing *(RS)*, which help us to govern the number of shared resources among the peer nodes in the P2P network. Every peer node should compute $P_x(RS)$ of its neighbor peers, means the number of resources shared by the peers that are one step away from querying peer P_Q.

$$P_x(RS) = \sum\nolimits_{x}^{N(P_x)} N(RS) \tag{7}$$

It is practical that, when a peer shares its maximum resources to other peers, then it is having higher probabilities of query matching than a peer that shares limited resources.

Quality of Information (QI). In P2P network, another observation stated that not all the shared files are used for answering request queries [4]. By considering the number of shared resources with querying peers and the number of resources used to query response has a solid relationship with the responding peers. Having this in mind, exploit the feature, the "Quality of Information" that distinguishes the useful and useless resources. Consider *NM(RS)* be the number of shared information that match with the requested queries. Each querying peer P_Q computes $P_x(QI)$ for all of its neighbor peers. It is stated as,

$$P_x(QI) = \sum\nolimits_x^{N(P_x)} NM(RS) \tag{8}$$

Success Ratio (SR). The next parameter taken for analyses is the query success ratio *SR* of the neighbor peer nodes. Though the possibility of query success ratio may be influenced by the quality and quantity of resource contents, which have been taken the number of success hits provided by the neighbor peer nodes are considered for analyzes the quality of services. Each querying peer P_Q computes *NS(QR)*, is the total amount of success query response of neighbor peers that are one step away from P_Q. It is defined as,

$$P_x(SR) = \sum_x^{N(P_x)} NS(QR) \tag{9}$$

The time is taken for the query success and its efficiency in query forwarding has been improved gradually. Therefore, the queries exchanged between the peers and its traffic is minimized.

4 Experimental Setup and Performance Evaluation

This section, deals with the performance evaluation of proposed TMV-based query routing mechanism with the support of simulator and analysis the results outcome with the existing famous search mechanisms. In the simulation environment, build the topology with peers range from 100–3,000 peer nodes and Table 1 listed the different parameters used in the simulation. By considering the high mobility nature of peers in an unstructured P2P network, the topology design that allows the peer nodes can join (leave) a network at any point.

Table 1. Simulation parameters

Parameters	Values
Nodes	100–3000
Node degree	5
Max TTL	20
Query request per second	1000
Topology area	1000 m * 800 m
Simulation time	600 s

Each peer node can generate the queries with equivalent probability. Here, the comparison among the searching mechanisms like Flooding (FL), *K*-Random-Walk (*k*-RW) and Query Routing Tree (QRT) with the proposed Trust-Matrix-Value (TMV) for the parameters Response Time, Query Success Rate, Searching Hop in Query Hit were

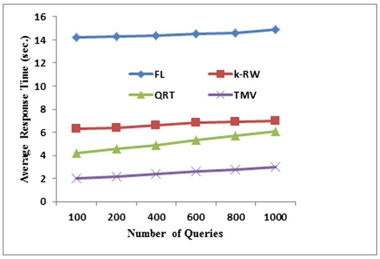


Fig. 3. Average response time of search queries

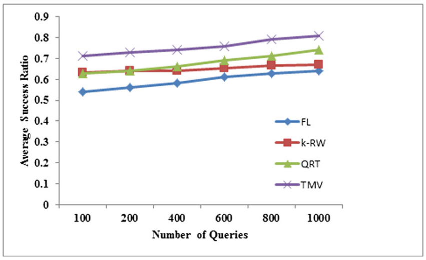


Fig. 4. Average query success ratios for 1000 peers

considered. During simulation, peers in the topology can randomly generate the query message and perform forwarding it to neighbor peers. Figure 3 illustrates the average response time taken to hit the target peer. The outcome of the experiment shows that the proposed TMV achieves the minimum response time than another searching mechanism. In practical, TMV utilizes a minimum amount of network resources to find query matches, it attains the best response time than *k*-RW and QRT. On the other hand, it also acquires low network traffic. Even though the number of query generation increases, TMV performs effectively because it selects trusted neighbor peers each time for further query forwarding. Following it, perform the simulation against the computation of average success rate of searching queries as stated in Fig. 4.

It shows that TMV achieves 70–80% of query responses for the request, whereas FL achieves only 50–60% of query responses due to its high traffic overhead and RW produces only 42–49% of responses because of non-deterministic neighbor selection whereas QRT achieves 61–70%. As an outcome, the proposed TMV reduces the network traffic cost with higher success rate. Figure 5 shows the result of an average number of search nodes with the number of queries. As discussed, TMV has been choosing the neighbor peer nodes based on its trust measures, it uses only a minimum number of hops to obtain the query response. Whereas FL, *k*-RW and QRT searching algorithms consider most of the peer nodes in the process of query searching, the network resources have not been utilized properly. As shown in the result, TMV achieves query response with minimum peer utilization. In the second simulation setup, compare the trusted and distrusted peer-to-peer network by evaluating the following parameters: Success Ratio, and Delay. In Fig. 6, it has been proven that the trusted P2P network achieves high success rate with increasing search query counts than the distrusted environment. In a distrusted network, some of the peer nodes are selfish in nature; it should not use its own resources for query searching and forwarding process.

Figure 7 shows the comparison between query response delay and a number of query nodes. From the outcome, it has been proven that the query delay increases when the number of querying peer increases. The time delay taken for query response is minimal in trusted network compared to distrust network even the querying peer nodes increases.

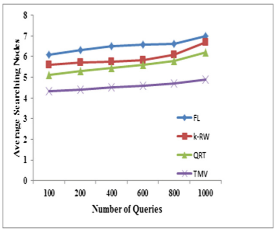


Fig. 5. Average search depths of query for 1000 peers

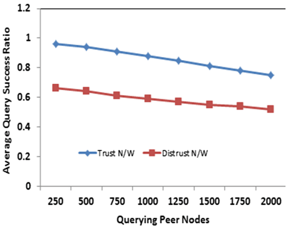


Fig. 6. Average success rate comparisons between trust and distrust netword

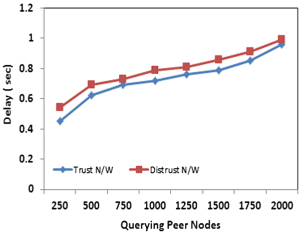


Fig. 7. Average query response delays for 2000 peers

5 Conclusion

In this work, the authors have proposed a Trust-Matrix-Value (TMV) based query searching mechanism for peer-to-peer networks, which is built on the trust management scheme. The key idea of this work is to establish the trustworthiness among the communicating peer nodes before sending or accepting query traffic. Each peer should examine the trust of the neighbor peers before query forwarding. The simulation has been executed to analysis the performance level of TMV algorithm with FL, *k*-RW and QRT. The results exhibit that TMV increases the network performance gradually even under dynamic circumstances. Through the simulation results, TMV is an effective query searching mechanism and optimized algorithm to enhance the query hits in peer-to-peer networks.

References

1. Chavez, E., Graff, M., Navarro, G., Tellez, E.S.: Near neighbor searching with K-nearest references. J. Inf. Syst. **51**, 43–61 (2015). https://doi.org/10.1016/j.is.2015.02.001
2. Chen, K., Shen, H., Zhang, H.: Leveraging social networks for P2P content-based file sharing in disconnected MANETs. IEEE Trans. Mob. Comput. **13**, 235–249 (2015). https://doi.org/10.1109/TMC.2012.239
3. Chiu, Y.M., Eun, D.Y.: On the performance of content delivery under competition in a stochastic unstructured peer-to-peer network. IEEE Trans. Parallel Distrib. Syst. **21**(10), 1487–1500 (2010). https://doi.org/10.1109/TPDS.2010.15
4. An, D., Ha, B., Cho, G.: A robust trust management scheme against the malicious nodes in distributed P2P network. Int. J. Secur. Appl. **7**(3), 317–326 (2013)
5. Gaeta, R., Sereno, M.: Generalized probabilistic flooding in unstructured peer-to-peer networks. IEEE Trans. Parallel Distrib. Syst. **22**(12), 2055–2062 (2011). https://doi.org/10.1109/TPDS.2011.82
6. Gheorghe, G., Lo Cigno, R., Montresor, A.: Security and privacy issues in P2P streaming systems: a survey. Peer-to-Peer Netw. Appl. **4**, 75–91 (2011). https://doi.org/10.1007/s12083-010
7. Hieungmany, P., Souma, T., Shioda, S.: Directional-random-walk - based contents search for unstructured P2P systems. IEICE Trans. Commun. **J98-B**(2), 132–140 (2015)

8. Hsiao, H.-C., Su, H.: On optimizing overlay topologies for search in unstructured peer-to-peer networks. IEEE Trans. Parallel Distrib. Syst. **23**(5), 924–935 (2012). https://doi.org/10.1109/tpds.2011.241
9. Filali, I., Huet, F.: Dynamic TTL-based search in unstructured peer-to-peer networks. Published by IEEE in CCGrid, pp. 438–447 (2010). https://doi.org/10.1109/ccgrid.2010.66
10. Chen, K., Hwang, K., Chen, G.: Heuristic discovery of role-based trust chains in peer-to-peer networks. IEEE Trans. Parallel Distrib. Syst. **20**, 83–96 (2009). https://doi.org/10.1109/tpds.2008.60
11. Lee, H., Nakao, A.: A feasibility study of P2P traffic localization through network delay insertion. IEICE Trans. Commun. **E95-B**(11), 3464–3471 (2012). https://doi.org/10.1587/transxom.e95.b.3464
12. Ghorbani, M.: An adaptive k-random walks method for peer-to-peer networks. Adv. Comput. Sci.: Int. J. **2**(3) (2013)
13. Radhika, N., Thejiya, V.: Trust-based solution for mobile ad-hoc networks. Int. J. Adv. Res. Comput. Sci. Soft. Eng. **4**(5), 73–82 (2014)
14. Shyamala, C.K., Padmanabhan, T.R.: A trust-reputation model offering data retrievability and correctness in distributed storages. Int. J. Comput. Appl. **36**, 56–63 (2015)
15. Venkadeshan, R., Chandrasekar, M.: Effective communication in P2P network by introducing GOSIP-PHE algorithms. Wirel. Pers. Commun. **87**, 923–937 (2016). https://doi.org/10.1007/s11277-015-2625-6
16. Weihua, G., et al.: Enhanced entropy-based resource searching in unstructured P2P networks. Chin. J. Electron. **24**(2), 229–235 (2015). https://doi.org/10.1049/cje.2015.04.002