

A MULTI ATTRIBUTE RANKING MECHANISM FOR PEER-TO-PEER NETWORKS

¹Swaminathan, B. and ²Sheila Anand

¹Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

²Department of Information Technology, Rajalakshmi Engineering College, Chennai, India

Received 2013-09-13; Revised 2014-01-07; Accepted 2014-03-31

ABSTRACT

P2P systems are envisioned to play a greater role in collaborative applications. P2P environments remove the challenging task of using servers for information sharing. Emerging collaborative P2P systems require discovery and utilization of multi-attribute, distributed and dynamic group of resources to achieve greater tasks beyond conventional file and processor cycle sharing. The process of selection of a peer for collaborative work therefore plays a significant role in accomplishing the task. Collaborative P2P systems use a group of diverse resources like hardware, software, services and data to accomplish the task or application. Hence, ranking of peers based on multiple heterogeneous attributes plays a significant role in enabling the selection of the right peers for collaboration. This study proposes the use of Analytic Hierarchy Process (AHP) for ranking the peers for selection. The relative importance of the attributes has to be decided based on the P2P application that is being collaborated. AHP provides the mathematical technique for decision making for ranking the peers for collaborative activity. The application of AHP for peer ranking has been illustrated with the use of examples. The system has been implemented and tested using Planet Lab dataset. The selection of right peer using this method has improved the process of multi attribute decision making and an optimal decision has been obtained by mapping the requirements to the available resources. The number of criteria used for P2P collaboration has been varied and the results observed shows that the decision making time increases proportional to the number of criteria.

Keywords: Multi Attribute Ranking, Collaboration, P2P, AHP

1. INTRODUCTION

Peer-to-Peer (P2P) network is a type of decentralized and distributed network architecture in which individual nodes in the network act as peers. A peer in the network can act as both server and client, in contrast to the centralized client-server model where client nodes request access to resources provided by central servers. There is no central coordination or authority. In a pure P2P environment, peers can join and leave the network dynamically. In a P2P network, tasks such as searching for files or streaming audio/video are shared amongst multiple interconnected peers who make a portion of their

resources such as processing power, disk storage or network bandwidth directly available to other network participants, without the need for centralized coordination by servers.

P2P networks generally use some form of virtual overlay network to logically connect a subset of the nodes in the physical network (Lua *et al.*, 2005). At the application level, the peers communicate directly through the overlay. Data exchange is carried out over the underlying TCP/IP network. Overlays are used for indexing and peer discovery and make the P2P system independent of the physical network topology. There are two classes of P2P overlay networks: Structured and Unstructured.

Corresponding Author: Swaminathan, B., Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India

In Structured overlay, network topology is tightly controlled and content are placed at specified locations instead of random peers. Chord, Pastry, Tapestry, Kademia, HyperCup, P-Grid, Koorde, CAN are some of the P2P systems based on structured overlays. Structured P2P systems generally use the Distributed Hash Table (DHT) in which (value, key) pairs indicating the peer location where the data object is located. In Unstructured networks, the peers join the network without any prior knowledge of the topology. Unstructured overlay designs of P2P systems include Freenet, Gnutella, FastTrack, Fast Freenet, Local Minima Search (LMS).

P2P systems have many advantages of resource sharing, functionality sharing, ad-hoc collaboration, improved reliability and scalability (Chawathe *et al.*, 2003). However, the decentralized, open and anonymous nature of P2P systems can raise serious concerns for a peer. Peers can join and leave the network dynamically and many of the peers in the network may not have interacted with each other earlier. There is no control or accountability on the content or resources that a peer can share on the network. Hence, the process of selecting the right peer for collaboration and its decision making process becomes difficult.

P2P networks are currently being used for a variety of applications such as file sharing, digital library, video and voice calls and video streaming. File sharing is perhaps one of the most commonly used applications of P2P networks. Some of the popular P2P file sharing applications include Napster, Gnutella, KaZaa, BitTorrent (Pouwelse, 2004), eDonkey, eMule, Limewire, FastTrack, Freenet, OverNet. Digital Library applications have been developed for searching the relevant content on P2P networks. P2P-4-DL (Clarke *et al.*, 2001) is one of the widely used Digital Library applications that are widely used. Skype is a popular and widely used Voice-over P2P (VoP2P) application.

Skype provides services like P2P voice and video calls, voice calls to PSTN endpoints, file transfer, instant messaging and video conferencing (Walkerdine and Rayson, 2004). P2PTV is application software which is used for the redistribution of video streams in real time on P2P network. The distributed video streams are typically TV channels from all over the world. Some of the P2PTV applications include TVUPlayer, PPLive, QQLive, PPStream, Abroadcasting, Zattoo, Octoshape, LiveStation, Joost and Babelgum.

P2P systems are envisioned to play a greater role in collaborative applications. P2P architecture offers some interesting benefits and challenges to collaboration. P2P

environments remove the challenging task of using server for information sharing. Emerging collaborative P2P systems require discovery and utilization of diverse, multi-attribute, distributed and dynamic group of resources to achieve greater tasks beyond conventional file and processor cycle sharing.

This study proposes a multi-attribute ranking mechanism for selection of peers for collaboration in P2P networks. The rest of the paper is organized as follows. The review of related work is presented in section 2. The proposed work is given in section 3. Implementation and results are given in section 4 and section 5 concludes the paper.

2. RELATED WORK

Collaborative P2P system can be defined as P2P system that aggregates a group of diverse resources like hardware, software, services and data to accomplish a greater task (Bandara and Jayasumana, 2013). The process of selection of a peer for collaborative work plays a significant role in accomplishing the task.

A typology for P2P environments has been developed by Fattah (2002). He divides uses of P2P architecture into four applications: User collaboration, application interaction, resource utilization and supercomputing. The first two are grouped under active applications and the later two under passive applications. Active applications are ones in which users or systems do things with P2P to accomplish a task. For example, instant messaging applications, such as ICQ or MSN Messenger and file-sharing systems, such as Napster, would be active applications. Passive applications are one in which idle resources are roped in uses other than their primary function.

Resource utilization applications allow combining of resources to produce large database out of the documents stored on computers scattered around the world. McAfee corporation has produced software that helps manage the distribution of software among users using this idea. Supercomputing applications harness the computing power of PCs on the network or on the internet to aggregate their power and produce virtual supercomputers. SETI@home is a prime example of such an application.

The process of resource collaboration comprises of seven phases-Advertisement, Discovery, Selection, Matching, Binding, Use and Release (Bandara and Jayasumana, 2013). In the Advertisement phase, each peer node advertises its resources and their capabilities. Peers may discover for the resources by generating

request messages to other peers. In the selection phase, the group of resources that satisfies application requirements is selected. In the matching phase, the appropriate combination of resources that are capable of working together has to be determined. Once the required resources are identified, the peers have to communicate and make sure that the selected resources are available for use during the binding phase. The next phase is to utilize the best subset of available resources that satisfy the application requirements. Once the task is completed or binding expires, then release the resources.

A resource is characterized by a set of attributes. When compared to *single-attribute* P2P systems such as file sharing systems, formal characterization of real world, *multi-attribute* resources and queries has received attention recently. For example, BOINC is a volunteer computing platform that is used to remotely execute jobs using idle computing resources. BOINC schedules jobs based on *static attributes* (e.g., CPU speed, total memory, presence of hardware accelerators) of nodes as the jobs are expected to run for several hours and the system is optimized for throughput. In contrast, performance, Quality of Service (QoS) and Quality of Experience (QoE) of latency sensitive applications such as Collaborative Adaptive Sensing of the Atmosphere (CASA) and community cloud computing depend on *dynamic attributes* (e.g., CPU utilization, free memory and bandwidth).

CASA, GENI, UniWiki and aggregated P2P systems depend on some form of resource collaboration. These systems share a variety of resources such as processor cycles, storage capacity, network bandwidth, sensors, special hardware, middleware, scientific algorithms, application software, services and data. These resources are characterized by multiple static and dynamic attributes. For example, CPU speed, free CPU capacity, memory, bandwidth, operating system and a list of installed applications/middleware and their versions may characterize a processing node. These multi-attribute resources need to be combined in a timely manner to meet the performance and QoS requirements of collaborative P2P applications.

Collaborative P2P systems are applicable in a wide variety of contexts such as Distributed Collaborative Adaptive Sensing (DCAS), grid/cloud computing, opportunistic computing, Internet and social networks www.cnrl.colostate.edu/Projects/CP2P/. CASA, the Center for Collaborative Adaptive Sensing of the Atmosphere, creates a distributed collaborative adaptive sensor network that sample and monitor the atmospheric phenomena. The primary objective of CASA is to

improve observations and forecasting of weather by developing new paradigms for sensing the atmosphere. P2P computing fits naturally to this new era of user-driven, distributed applications utilizing resource-rich edge devices. Thus, there is a tremendous opportunity to create value by combining societal trends with P2P systems.

CASA is an emerging heterogeneous network of weather radars, processing nodes and data fusion algorithms (e.g., tornado tracking and precipitation estimating algorithms) that operates collaboratively to detect hazardous atmospheric conditions. Collaborative P2P data fusion provides an attractive implementation choice for CASA real-time radar data fusion, weather monitoring and hazard prediction because data is constantly being generated, processed, pushed and pulled among radars, storage and processing nodes. CASA depends on efficient discovery and utilization of heterogeneous, dynamic and distributed resources that are characterized by multiple attributes.

UniWiki is a collaborative peer to peer system for distributing and managing dynamic content that combines two widely studied technologies: Distributed Hash Tables (DHT) and optimistic replication (Oster *et al.*, 2009). The architecture of UniWiki is built on top of a reliable, inexpensive and consistent DHT-based storage, any number of front-ends can be added, ensuring both read and write scalability, as well as suitability for large-scale scenarios.

Global Environment for Network Innovations (GENI) is being actively used for network research and education. It enables academic and industrial researchers perform new classes of experiments that tackle critically important issues in global communications networks. GENI also enables promoting innovations in network science, security, technologies, services and applications; and provide collaborative and exploratory environments for academia, industry and the public to catalyze discoveries and innovation.

The P2P collaboration applications like SETI@home, BOINC uses static attributes (Heien *et al.*, 2011). The designs of multi-attribute systems have relied on assumptions such as independent and identically distributed attributes. Characteristics of multi-attribute resources and queries from PlanetLab and SETI@home were analyzed to determine the values and behavior of model parameters.

The process of identifying the right peers for collaboration requires ranking of the peers. Peers having higher probability in resource sharing can be selected based on the ranking of peers. This study formalizes the

application of the ranking algorithm for peer to peer collaborative applications to determine the right peer for collaboration based on multiple attributes.

3. PROPOSED MULTI-ATTRIBUTE RANKING (MAR) FOR P2P NETWORKS

Peer collaboration begins with identifying a peer for hosting and co-ordination of the collaboration activity. A request to host a session is sent and the host peer agrees to manage the collaboration activity. The host sends a collaboration invitation to the peers in the P2P network. All peers who want to collaborate may add the host to their contact managers. Most of the peers will send invitation responses, whether accepted or declined, back to the host peer in a timely fashion. The host peer processes all invitation responses to determine who has accepted, who has declined and who has not answered. It may cancel invitations to those who have not answered, or perform some other activity.

Collaboration session can now happen between the registered peers. P2P applications can use the collaboration infrastructure to coordinate communications and complete their tasks. The process of ranking of peers is significant for selecting the peers for collaboration for a particular task or application. The collaborative work of P2P systems require heterogeneous resources to be used. Hence, ranking mechanism based on multiple attributes is required to enable the participating peer to select the right peers for collaboration.

Many multiple attribute decision making methods are used for ranking in the manufacturing environment. Analytic Hierarchy Process (AHP) is one of the most popular analytical techniques for ranking. Designed to reflect the way people actually think, AHP continues to be the most highly regarded and widely used decision making method. AHP can efficiently deal with objective as well as subjective attributes. AHP has been extensively used in integrated manufacturing (Putrus, 1990), in the evaluation of technology investment decisions (Boucher and McStravic, 1991), in flexible manufacturing systems (Wabalickis, 1988), layout design (Cambron and Evans, 1991) and also in other engineering problems (Wang and Raz, 1991). To the scope of our knowledge, AHP has not been applied in P2P collaboration. This study proposes the use of AHP for ranking peers for collaboration in P2P networks.

The application of AHP technique for multi-attribute ranking for P2P collaboration is explained and illustrated with examples. The calculation techniques of AHP begin

with the construction of a matrix expressing the relative values of a set of attributes. For example, the relative importance of the resources like number of processors, CPU speed and memory size is determined with respect to the cost of resource sharing and ease of operation. The relative importance of the attributes has to be decided based on the P2P application that is being collaborated. Usually, the decision-maker has to choose his answer among 1-9 discrete choices. Each choice is a linguistic phrase. Some examples of such linguistic phrases are: "A is more important than B", or "A is of the same importance as B", or "A is a little more important than B", as tabulated by the rating scale table of (Saaty, 1980).

The importance can be scaled as equal importance, weak importance of one over another and strong importance. For example, the rating scale of Saaty is used for assigning a number for relative importance of the attributes used for illustration, 1 for equal importance, 3 for weak importance of one over another and 5 for strong importance. A basic assumption is that if attribute A is absolutely more important than attribute B and is rated at 5, then B must be absolutely less important than A and is valued as 1/5.

The initial matrix for the pairwise comparison is given with the principal diagonal containing entries of 1 as each factor is as important as itself. The sample attributes are named as number of core Processors (P), CPU speed (C) and Memory size (M). The initial matrix can now be given as:

Attributes	C	P	M
C	1		
P		1	
M			1

The criteria can be fixed based on the importance of the attributes relative to the objective used for peer ranking with respect to cost of resource sharing and ease of operation. The matrix obtained is named as *Judgement Matrix*. The Judgement matrix for sample attributes like Number of core processors, CPU Speed and Memory size are fixed based on the observed data and is computed as follows.

The number of core processors used for computation would be single, dual or double dual processors. This is represented as numerical values 1, 2 and 4 for illustration. The dual processors are assumed to be suitable for most of the computations. Hence, more importance with the value 5 is given for dual processors. The single processors are given the next level of importance comparing with dual processors. The double

dual processors are assumed to be of weak importance with the value representation of 1. The judgement matrix for number of core processors is represented as follows:

Number of processors	1	2	4
1	1	1/5	3
2	5	1	5
4	1/3	1/5	1

The relative importance for the attribute CPU speed is computed by determining the discrete values for a range represented as GHz. Here, the range 2.25-2.5 GHz is considered with the relative importance value to be 5, the ranges 2.00- 2.25 GHz and 2.25-2.5 are with the values to be 3. The judgement matrix for the attribute CPU Speed is given as:

CPU speed	1.75-2.00	2.00-2.25	2.25-2.50
1.75-2.00	1	1/3	1/5
2.00-2.25	3	1	1/3
2.25-2.5	5	3	1

The memory size required is categorized as the ranges 1-3 GB, 3-4 GB and 4-32 GB. The relative importance value is computed to be 5 for 3-4 GB, 1-3 GB and 4-32 GB. Hence, the judgement matrix for the attribute memory size is given as:

Memory size	1-3	3-4	4-32
1-3	1	1/5	3
3-4	5	1	5
4-32	1/3	1/5	1

The interdependency among the attributes is determined and is represented using the matrix termed as overall preference matrix. The pairwise comparisons are carried out for all factors to be considered and the matrix is completed. There is no standard way to make the pairwise comparison. Considering a P2P collaborative application, let it be supposed that the attribute Number of Processors among the attributes considered is slightly more important than Memory size. In the matrix it is rated as 5 in the cell P, M and 1/5 in M, P. Also, memory size is considered to be more important than CPU Speed and is rated as 3 in the cell M, C and 1/3 in the cell C, M. The overall preference matrix obtained as:

Attributes	C	P	M
C	1	1/5	1/3
P	5	1	1/3
M	3	3	1

Table 1. Priority vector of the attributes

Attributes		
C	P	M
0.2	0.1	0.2
0.7	0.26	0.7
0.1	0.64	0.1

The eigenvector called as the Priority Vector for the attributes is calculated for every attribute. The elements in each row of the matrix are multiplied with each other and then the n^{th} root is taken for computing the priority vector values. Since, the number of attributes is 3, cubic root value is calculated and is summarized **Table 1**.

The judgement can be verified by calculating a Consistency Ratio (CR) to measure how consistent the judgements have been relative to large samples of random judgements. If the CR is greater than 0.1, the judgements are untrustworthy because they are too close to randomness, pairwise comparisons have to be re-evaluated and must be repeated. The computed CR values are 0.117, 0.03 and 0.117.

Additionally, the priority vector for considering the interdependency among the attributes is calculated to be (0.11, 0.32 and 0.57). The next step is to normalize the relative values by dividing the values with their sums which is termed as final priority vector. The normalized values for the judgement matrix are computed and the final priority vector for the considered sample is (0.349, 0.547 and 0.104). The value 0.547 shows that the attribute Number of Processors is given more importance; 0.349 shows that the attributes CPU speed and the Memory size is given less importance compared to Number of processors.

4. IMPLEMENTATION AND RESULTS

P2P collaboration involves the support of the resources from various peers. It is assumed that every peer advertises its available resources that could be dedicated for collaborative tasks. The requesting peer gives its specification and ranking of peers is done on the basis of the attributes that would satisfy the requirements. This system was implemented in Java and tested with the dataset generated by www.planet-lab.org from computer networking research laboratory. The system was tested with 640 peers by varying the number of attributes to be 3, 5 and 7. The attributes that were considered in this system are given in **Table 2**.

Each attribute has a value that belongs to a given domain. The domain is typically bounded and may represent a continuous or discrete value. For example,

free memory is continuous, number of processors is discrete. Attribute values are further classified as static (e.g., CPU speed, operating system) and dynamic (e.g., free memory and transaction rate).

The performance analysis of the system is compared for the attributes with criteria to be 3, 5 and 7. The graph shown in Fig. 1 shows the performance analysis of the system for 640 peers.

The graph shows that initially, the time taken for decision making is independent of the criteria selection. However, the time increases proportionally to the number of criteria. The process of decision making is fast when the number of peers is less

irrespective of the number of criteria and is shown using the graph in Fig. 2.

Table 2. Attribute representation

Attribute	Notation used	Units	Data type
CPU speed	CSp	GHz	Float
No. of core processors	NCore	-	Integer
CPU Free space	CFree	%	Float
Memory size	MSize	GB	Float
Memory free space	MFree	%	Float
Disk size	DSize	GB	Float
Disk free space	DFree	GB	Float
Data receiver rate	Rx	Bps	Float
Data transmission rate	Tx	Bps	Float

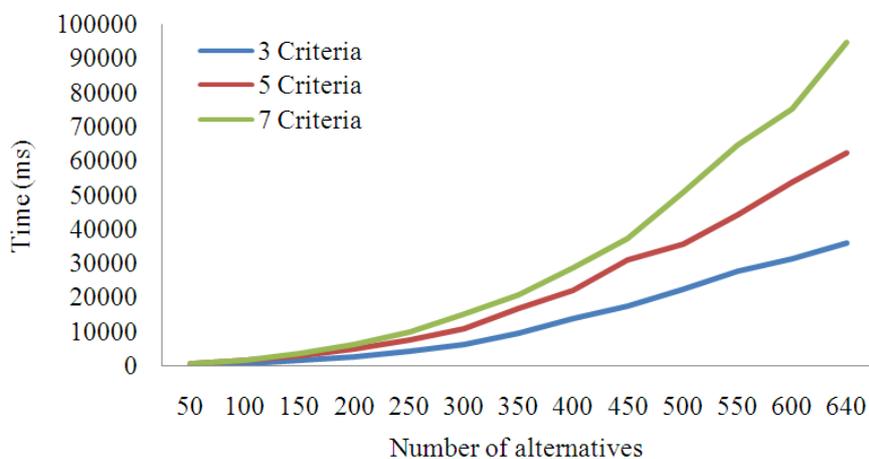


Fig. 1. Performance analysis of the system for 50-640 peers

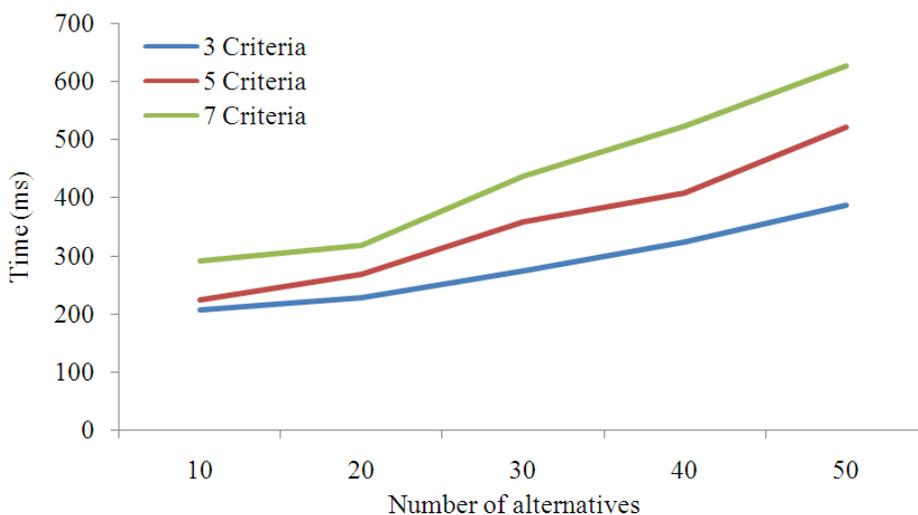


Fig. 2. Performance analysis of the system for 10-50 peers

5. DISCUSSION

Prior work in P2P collaboration has focused on allowing teams to work together over a network. For instance, groove networks, allows secure digital collaboration. The team members can instant message, share files, browse the web together. Other P2P collaboration efforts are more application centric. Basic Support for Cooperative Work (BSCW) is a collaborative workspace over the web that supports document upload, event notification and group management. CASA uses a distributed sensor network to monitor the atmospheric for weather forecasting. GENI is a collaborative platform that allows users to aggregate diverse resource, for example, data pooling and sharing. UniWiki is a collaborative peer to peer system for distributing and managing dynamic content.

Emerging collaborative P2P systems look to accomplish different tasks by aggregating a group of diverse resource. For collaborative P2P work, in the peer selection phase, the group of resources that satisfy application requirements have to be selected. Peer selection has to be based on defined criteria and on multiple attributes. AHP is a well-known technique that has extensively applied in the manufacturing environment and to the best of our knowledge, has not been applied for P2P collaboration.

In the proposed work, the decision making has been performed based on criteria, which are defined to meet the user/application requirements. The criteria for peer selection are based on multiple attributes like, processors, CPU speed and Memory size. The results show that AHP helps to arrive at an optimal decision regarding peer selection by mapping the requirements to the available resources. The application of AHP for peer selection has been tested for varying number of criteria and the performance comparison has been presented.

6. CONCLUSION

P2P-based collaboration is a paradigm that meets the requirements of users, but requires enabling technologies and techniques to be researched and implemented. The proposed approach would assist in peer ranking and selection for P2P collaboration. This system considers multiple heterogeneous attribute and the interdependency of the attributes. The relative importance of the attributes has to be decided based on the P2P application that is being collaborated. AHP provides the mathematical technique for decision

making for ranking the peers for collaborative activity. This study can be further enhanced by applying and testing the peer ranking approach for a larger dataset and for real time environment.

7. REFERENCES

- Boucher, T.O. and E.L. MacStravic, 1991. multiattribute evaluation within a present value framework and its relation to the analytic hierarchy process. *Eng. Economist*, 37: 55-71. DOI: 10.1080/00137919108903055
- Cambron, K.E. and G.W. Evans, 1991. Layout design using the analytic hierarchy process. *Comput. Indust. Eng.*, 20: 221-229. DOI: 10.1016/0360-8352(91)90026-3
- Heien, E., D. Kondo and D.P. Anderson, 2011. Correlated resource models of internet end hosts. *Proceedings of the 31st International Conference on Distributed Computing Systems*, Jun. 20-24, IEEE Xplore Press, Minneapolis, MN., pp: 278-287. DOI: 10.1109/ICDCS.2011.13
- Lua, E.K., J. Crowcroft and M. Pias, 2005. A survey and comparison of peer-to-peer overlay network schemes. *IEEE Commun. Surveys Tutorials*, 7: 72-93. DOI: 10.1109/COMST.2005.1610546
- Fattah, H.M., 2002. P2p: How peer-to-peer technology is revolutionizing the way we do business. Dearborn Trade Publishing.
- Oster, G., P. Molli, S. Dumitriu and R. Mond'egar, 2009. UniWiki: A collaborative P2P system for distributed wiki applications. *Proceedings of 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises*, Jun. 29-Jul. 1, IEEE Xplore Press, Groningen, pp: 87-92. DOI: 10.1109/WETICE.2009.42
- Bandara, H.M.N.D. and A.P. Jayasumana, 2013. Collaborative applications over peer-to-peer systems-challenges and solutions. *Peer-to-Peer Netw. Applic.*, 6: 257-276 DOI: 10.1007/s12083-012-0157-3
- Walkerdine, J. and P. Rayson, 2004. P2P-4-DL: Digital library over peer-to-peer. *Proceedings of the 4th International Conference on Peer-to-Peer Computing*, Aug. 25-27, IEEE Xplore Press, pp: 264-265. DOI: 10.1109/PTP.2004.1334957
- Pouwelse, J., 2004. The BitTorrent P2P file-sharing system, Detailed Measurement Study.

- Clarke, I., O. Sandberg, B. Wiley and T.W. Hong, 2001. Freenet: A distributed anonymous information storage and retrieval system. proceedings of the International Workshop on Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability, Jul. 25-26, Springer, USA., pp: 46-66. DOI: 10.1007/3-540-44702-4_4
- Putrus, P., 1990. Accounting for intangibles in integrated manufacturing (nonfinancial justification based on the Analytical Hierarchy Process). Inform. Strategy, 6: 25-30.
- Saaty, T.L., 1980. The Analytic Hierarchy Process. 1st Edn., McGraw Hill International. New York.
- Chawathe, Y., S. Ratnasamy, L. Breslau, N. Lanham and S. Shenker, 2003. Making gnutella-like P2P systems scalable. Proceedings of the Conference on Applications, Technologies, Architectures and Protocols for Computer Communications, Aug. 25-29, ACM., New York, pp: 407-418. DOI: 10.1145/863955.864000
- Wabalickis, R.N., 1988. Justification of FMS with the analytic hierarchy process. J. Manufactur. Syst., 17: 175-182. DOI: 10.1108/17410380510626204
- Wang, L. and T. Raz, 1991. Analytic hierarchy process based on data flow problem. Comput. IE, 20: 355-365. DOI: 10.1016/0360-8352(91)90007-S