A Reputation Based Framework to Avoid Free-riding in Unstructured Peer-to-Peer network

Ruchir Gupta, Yatindra Nath Singh, Senior Member IEEE

Department of Electrical Engineering, IIT, Kanpur

{rgupta,ynsingh}@iitk.ac.in

Abstract—Free riding is a major problem in peer-to-peer networks. Reputation management systems are generally proposed to overcome this problem. In this paper we have discussed a possible way of resource allocation on the basis of reputation management system i.e. probabilistic allocation based on reputation. This seems to be a better way for allocation of resources because in this case nodes that don't have very good reputation about each other, may also serve each other at least some amount of resource with finite probability. This avoids disconnect between them. Algorithms are presented for optimizing the shared capacity, reputation based probabilistic allocation that is optimal for a node, and formation of interest groups on the basis of similarity between interests of nodes.

I. Introduction

The reputation computation should be such that it gives enough credit to the nodes those are contributing more to the network. If some node is requested for smaller size resource, it should not be at loss. We are proposing a probabilistic resource allocation method in this paper. It uses reputation of users while allocating the resources. This kind of allocation ensures that a node with low reputation gets resources with some small finite probability. This will avoid the disconnect between two nodes that have low reputation for each other. This is important when resources are distributed in the network.

The incentive systems proposed in literature do not consider the interests of peers i.e. the neighbourhood of a peer is not made up on the basis of its interests. Although, gnutella servant keeps few last query replying peers in the cache but it is limited to that only. This system is rather inefficient because two nodes may have many interests in common while not being neighbours. Whereas two neighbour nodes may have very few common interests. Hence, the better option is to make neighbourhood on the basis of similarity of interests and reputation.

In this paper, we have proposed an algorithm for the optimization of shared capacity of a node (section III), a method to compute the reputation (section IV), probabilistic resource allocation based on reputation (section IV) and server selection according to interests of node

and reputation (section V). Finally numerical results are given to verify the hypotheses (section VI).

1

II. RELATED WORK

Various groups have suggested different techniques for resource allocation in peer-to-peer networks. Kung et.al. [1] proposed selection of a peer for allocation of resource according to its contribution to the network and usage of resources. In the same context, Feldman et.al. [2] proposed a new term – generosity of the node. It is estimated as the ratio of the service provided by the node to the service received by the node. Nodes will be served as per their estimated generosity. Banerjee et.al. in [3] proposed that a node will calculate the expected utility function for requesting node and on that basis it will decide if service has to be provided or not. In [4], [5] the resource allocation algorithm for single link limited capacity systems has been proposed. These papers considers network as market and proposes that second price auction leads to optimality. [4] assumes that resource is available everywhere except at the requesting node. Thus, a node is not required to have interaction with many nodes. Ma et.al. [6] proposed progressive water filling algorithm on the basis of marginal utility for allocation of resources among different requesting nodes. The base of bucket for water filling is proposed to be varying according to the contribution of requesting node. Ma et.al. [7] proposed to allocate the resource to requesting node on the basis of their contribution and requirement of bandwidth. Yan et.al. in [8] proposed a ranking based resource allocation scheme. Resource allocation is done according to utility and ranking of requesting peer to ensure max-min fairness.

Social networks are formed on the basis of interests of users. This fact is been capitalised to improve query search as well as recommendation network in peer-to-peer networks [9], [10], [11], [12]. In [9] BitTorrent traces are studied and it is concluded that interest based grouping of peers results in an efficient system. It also proposes a DHT based system to implement this kind of group formation. Wang *et.al.* [11] proposed interest

based online social communities that are headed by super nodes and nodes join the communities according to their interests. These communities will have a trust relationship among its members. In [12] a friend network is proposed on the basis of similarity of interests.

III. CAPACITY SHARING

As every node in the network is rational, hence it will try to share minimum amount of resource to increase its pay-off. If there is a reputation management system implemented in the network, nodes are compelled to share the resources to get the quality of service from the network. Nodes, being rational, will try to optimize the amount of shared resources. We propose a method for nodes to optimize the shared capacity to get the required quality of service from network. In this method, nodes will initially share some amount of resource. This amount of shared resource will be periodically reviewed and adjusted for optimality.

Initially a node will share the capacity as per its perceived download requirement. By perceived download requirement, we mean a rough estimate of its average download requirement. This need not to be accurate as it will be updated later on. But, it should neither be too low to ruin the reputation of a newcomer node nor be too high to cause a cost penalty. If no estimate is available, initially half of total download capacity can be shared.

Node will tweak the value of its shared capacity by periodically increasing and decreasing it by some amount δ to get the optimal point where it will get maximum advantage. While doing so, node will follow the following method.

- If decrease in sharing capacity does not decrease significant average download, it implies that node is sharing more than required resource and hence it should decrease it.
- If increase in sharing capacity increases significant average download, it implies that node is sharing less than required resource and hence it should increase it.
- 3) If decrease in sharing capacity decreases significant average download, it implies that either node was on optimal point (if it was preceded by an increase) so it should get back to that point or is now sharing even lesser than what it should have shared.
- 4) If increase in sharing capacity does not increases significant average download, it implies that either node was on optimal point (if it was preceded by an decrease) or is now sharing even more than what is required.

This process is shown in algorithm 1.

 ϵ is a parameter that is kept for overcoming the effect of demand variation in the network. When node observes high variation of the demand in the network, value of ϵ will be increased.

Algorithm 1 Shared upload capacity adjustment of a node

k = 0; and A(k) = -1 {k is the instant when node reviews its sharing capacity and A(k) is the indicator variable which shows the action taken at a particular repeat $D_k \leftarrow$ average data download for kT to (k+1)T $U_s = U_s + \delta \cdot A(k) \{U_s \text{ is shared capacity of the node}\}\$ if $|D_k - D_{k-1}| \le \epsilon$ then if A(k) = A(k-1) or A(k)=0 then $A(k+1) \leftarrow -1$ else **if** A(k) = 1 && A(k-1) = -1 **then** $A(k+1) \leftarrow 0$ end if end if else if $A(k) = -1 \&\& A(k-1) = 1 \&\& D_k < D_{k-1}$ then $A(k+1) \leftarrow 1$ end if if $D_k > D_{k-1}$ then $A(k+1) \leftarrow 1$ end if end if $k \leftarrow \mod_5(k+1)$

IV. REPUTATION BASED SYSTEM

Network is only meaningful if nodes are interacting with each other and contributing to each others' interest. We have a network of nodes that are rational in nature. Such nodes contribute in the network only when they have some incentive for doing so. To avoid this problem, a reputation based incentive system can be used. In such a system nodes keep the record of behaviour of other nodes observed by itself or on the basis of recommendation of different nodes. This kind of system forces rational nodes to contribute to the network. To implement this kind of a system, we need to formulate a way for estimation of reputation and a way for allocation of resource according to the estimated reputation.

A. Reputation Management System

until Node is in the network

Ideally, reputation should be the measure of cooperative behaviour of a node which is an abstract quantity and it is a private information of a node. So, it is difficult to measure the cooperative behaviour of a node and we can only measure its implications with some degree of uncertainty. However, it can be estimated with certain accuracy on the basis of behaviour observed by a node.

There could be a number of ways to observe the behaviour of a node. One such method may be to use the ratio of received data rate to requested data rate. The advantage of such technique is that if some node is asking for less amount of data, the serving node will not earn a bad reputation. Moreover this kind of reputation remains between 0 and 1 as given in ??, [13], i.e.,

$$t_{ij} = \left(\frac{q_{a,ji}}{\min(q_{ij,ay}, q_{f,ji})}\right)^{1-\eta_i} \times \frac{\hat{q}_{w,ji}}{q_{r,ji}},\tag{1}$$

hence it is easy to handle. Here t_{ij} is the reputation of node j for node i, $q_{r,ij}$, $q_{w,ji}$, $q_{a,ji}$, $q_{f,ji}$ and $q_{ay,ji}$ are the requested, willing, actual, feasible and accepted service rates respectively. The disadvantage of this kind of system is that it does not takes into account the amount of request. It means that if a node is asking for less amount of resource from a node and more amount of resource from another node and both are fulfilling node's demand, both will get similar gain in reputation. However, the node that was requested more resource, had to pay more in comparison to the other one.

This problem can be taken care of by giving different weights to different transactions as per the amount of resource requested by that node. Weights should be such that these should range between 0 to 1 and biggest service request should get maximum weight. Requesting node has a fixed download capacity that is generally the maximum of its download requirement.

A node can calculate the reputation of a node with following formulation,

$$t_{ij} = \left(\frac{q_{a,ji}}{\min(q_{ij,ay}, q_{f,ji})}\right)^{1-\eta_i} \times \frac{\hat{q}_{w,ji}}{q_{r,ji}} \times \frac{q_{r,ji}}{q_{r,i,d}}.$$
 (2)

Here $q_{r,i,d}$ is the download capacity of node i. It may be noted that we have multiplied the factor $\frac{q_{r,i}}{q_{r,i,d}}$ to the t_{ij} as estimated by equation $\ref{eq:total_state}$ in paper $\ref{eq:total_state}$?

Keeping the download capacity $(q_{r,i,d})$ instead of requested resource $(q_{r,ji})$ in denominator, we can overcome the above mentioned problem. However, $q_{r,i,d}$ is quite large compared to $q_{r,ji}$, this will make reputation values very low. Apart from it, every node has a different value for $q_{r,i,d}$ and it will be a problem for aggregation because a node having same kind of behaviour with two nodes of different download capacities will have different value for the t_{ij} . To over come these problems nodes will multiply their reputation table with $(q_{r,i,d}/Q_{r,d})$. Here $Q_{r,d}$ is the universal scaling factor known by all the nodes.

However, this will only work well if a node is only interested in resources that have size of same order. If some node 'A' is rich with resources that are small in size, this node will have a very small value of reputation for a node that is requesting for all size of resources because 'A' would have been asked only for small amount of resource. Consequently, when 'A' asks for small amount of resource, it will get the resource with a very small probability. This issue will be further investigated in next subsection.

B. Probabilistic Resource Allocation

In probabilistic resource allocation, node probabilistically decides if it will provide the resource to requesting node or not. It means, when a node 'A' requests for resource from node 'B', node 'B' checks the reputation table and converts reputation of 'A' to its effective reputation. Here, by effective reputation we mean reputation that is adjusted according to the requested amount of resource. To calculate this value, node multiplies the reputation value with ratio of its download capacity to requested amount of resource, i.e.

$$t_{ij,effective} = t_{ij} \times \frac{q_{r,i,d}}{q_{r,ij}}.$$
 (3)

Now in proposed system, if a node is asking similar amount of resource as it supplied, it will be given same quality of service. If it asks for smaller resource than it supplied, it gets even better quality of service whereas if opposite happens, it gets a poor quality of resource.

Once node i gets the effective reputation of a node j, it selects the node j with probability proportional to its reputation. It means node i generates a random number. If this generated number is smaller than the reputation of requesting node i.e. node j here, multiplied by a constant (v_i) , requesting node is selected to provide the resource. v_i is the constant that ensures the requirement of selected nodes remains around the shared capacity so that it can be optimally utilised. Mathematically,

$$P_{allocation,i,j} = \begin{cases} \acute{P}_{allo} & \text{if } \acute{P}_{allo} < 1. \\ 1, & \text{otherwise.} \end{cases}$$
 (4)

Where,

$$\acute{P}_{allo} = (t_{ij,effective})^{x} \cdot \nu_{i}$$

Here $P_{allocation,i,j}$ is the probability by which j will be allocated the resource. Node will be selected for resource allocation if,

$$rand \le P_{allocation,i,j}$$
 (5)

Here *rand* is the random number generated by the node and x is the reputation exponent. It is used because a low reputation node can only increase its reputation if it serves with x < 1. Its value has been calculated in [15].

Nodes will dynamically and periodically adjust the value of v_i to get the optimality. To do so, node will measure the utilised part of its shared capacity and fulfilment level of demand of selected nodes. If it is not able to utilise its shared capacity regularly, it increases the value of v_i and if demand of selected node is not fulfilled over the time, value of v_i is decreased.

After the selection of nodes, the shared capacity is distributed among selected nodes. If the total demand of requesting nodes is less than the shared capacity of serving node, every node is allocated resource as per their requirement. If this total demand is greater than the shared capacity, node needs to use some kind of allocation algorithm such that, the serving node can get

maximum advantage and nodes can not play game by asking for resources greater than their requirement.

A node will be maximum benefited when it has highest chance of getting selected for resource allocation i.e. by maximizing its $P_{allocation}$ for a future time when it will need some resource, according to equation (2). If serving node is doing this calculation, it can be assumed that $q_{a,ij} = min(q_{ji,ay}, q_{f,ij})$. Hence equation reduces to,

$$t_{ji,effective} = \frac{\hat{q}_{w,ij}}{q_{r,ij}} \times \frac{q_{r,ij}}{q_{r,ji}}.$$
 (6)

 $\hat{q}_{w,ij}$ can be replaced by $q_{o,ij}$ as we are discussing about allocation in a particular round where number of nodes, their demands and shared capacity has already been fixed. Hence equation (6) reduces to,

$$t_{ji,effective} = \frac{q_{o,ij}}{q_{r\,ii}}. (7)$$

 $q_{r,ji}$ can only be predicted statistically. We can observe that if a node 'A' is asking less amount of resource to node 'B' then 'B' can only get less amount of resource because if it will demand for bigger resource, its effective reputation will come down and hence it will not be selected for service by node 'A'. Hence if A is asking lesser resource from B, that implies B is asking lesser resource from A.

For simplicity ν can be taken as 1. Therefore, the optimisation problem a node needs to solve becomes,

$$\max \sum_{i} (t_{ji,effective})^{x} \implies \max \sum_{i} \left(\frac{q_{o,ij}}{q_{r,ij}}\right)^{x}.$$
 (8)

Such that

$$\sum_{j} q_{0,ij} = U_{s,i}
q_{0,ij} \leq q_{r,ij}; \forall j
\sum_{j} q_{r,ji} > U_{s,i}$$

Here *x* is a constant. Its value lies between 0 and 1. For our case its value is 0.75.

This is a difficult optimisation problem. For that, we need to observe the function a^x . Here a varies from 0 to some finite value and x is a constant between 0 and 1 as mentioned above. Two facts are easy to observe, first it is a monotonically increasing function and second its rate of change is monotonically decreasing function. Therefore it is evident that initially our objective function will get the maximum increment if resource is allocated to node that corresponds to smallest $q_{r,ij}$. After some allocation, increase in the value of objective function will decrease and now it will be more for any other node that has requested more data than first one. Now it will be beneficial to allocate data to this second node. After some allocation, any other node may result in more increment and this process continues till the resource allocation is complete.

On the basis of reason mentioned above, we propose an algorithm for allocation. First a node i decides about the minimum unit of allocation. Let us call this Δ_i . On the basis of Δ_i and the amount of total resource shared $(U_{s,i})$, i calculates the total number of allocation units $(U_{su,i})$ such that $U_{su,i} = \frac{U_{s,i}}{\Delta_i}$. Now, i constructs an allocation array U_{sua} that has the dimension of $U_{su,i} \times \acute{N}$. Here \acute{N} is the number of nodes selected for data allocation by equation (5) such that

$$U_{sua}(k,j) = ((k)^x - (k-1)^x) \cdot \left(\frac{\Delta_i}{q_{r,ij}}\right)^x. \tag{9}$$

Here k and j are row and column indices of U_{sua} . Elements of U_{sua} are sorted and indices of top $U_{su,i}$ element are stored in a vector of dimension $1 \times U_{su,i}$. The number of times any particular node comes in this vector will be allocated the same number of units.

In this kind of allocation, nodes asking for less amount of data will be given data first. Hence if a node asks for more data than its requirement, it loses the allocation part. This kind of allocation will also fulfil our second requirement.

As requests will be coming temporally in arbitrary fashion, it is necessary to define a policy followed by a node for provisioning the service. If node will serve the request as and when it comes, node will always remain busy in doing so. Moreover, nodes that has got the bandwidth, will get allocation again and again. If node will service the requests periodically, there is a chance with finite probability that a low reputation node will get the service while a high reputation node may keep on waiting.

Hence, a node should have a dynamic policy about serving instants. It means that when total reputation of requesting nodes crosses a certain threshold, node will serve the accumulated requests. If over a certain period of time total reputation of requesting nodes does not cross the threshold, node will serve the requests accumulated by this time. While summing up the reputation of requesting nodes, it is ensured that high reputation nodes get preferred to the nodes of lower reputation. Whenever, a node serves new requests, node will first do the selection process for newcomer nodes and then it will redistribute the resources among newly selected nodes and already existing nodes.

V. Server Selection

A. Common Interest Groups

In peer-to-peer file sharing network, different users have common interests. For a user, it is beneficial to make neighbours that share interests with him and ready to serve him. Therefore, a node should adopt a strategy to form its neighbourhood according to similar interests with good reputation nodes.

Interest is an abstract notion so classification of nodes on the basis of interest is difficult. Even if it is done, this will be a very large set that will be difficult to handle. Therefore, interest group should be formed on the basis of files, users requested or provided. However, users with different interests may request same file. For example, a song may be liked for different reasons like music, singer or lyrics. But, if two users are requesting for more and more similar files, probably they may have some common interests. As the number of similar files grows, probability of two peers choosing file due to same interest increases while choosing it for different interests decreases.

Therefore, we propose that a node will compute the similarity coefficient of the other nodes in the network. The similarity coefficient (χ_{ij}) of node j will be calculated by node i using

$$\chi_{ij} = \begin{cases} v_{ij} \cdot log_{base_i}(\Omega_{ij} + 1) & \text{if } \Omega_{ij} < base_i. \\ v_{ij}, & \text{otherwise.} \end{cases}$$
 (10)

Here Ω_{ij} is the number of times node i has queried to node j or vice-versa, v_{ij} is the ratio of answered queries to total queries between node i and node j. $base_i$ will be dynamically adjusted periodically as per the accuracy of similarity coefficient of the node. It means if the selected neighbours can not answer sufficient number of requests, value of $base_i$ will be increased.

B. Inclusion of Reputation in Neighbourhood Formation

As discussed earlier, for server selection, a node need to form its neighbourhood using interests and reputation. This can be done by combining reputation (t_{ij}) and similarity coefficient (χ_{ij}) for node j. The combined score can be used to rank the other nodes in the network. This rank can be used to select the server i.e. where to send queries.

The combining can be done as follows.

$$score_{ij} = \alpha \cdot \chi_{ij} + (1 - \alpha) \cdot t_{ij}.$$
 (11)

Here α is a combination coefficient between 0 and 1. Value of α will depend upon the stability of common interest network. If a node has newly joined the network, it has to build the interest network hence α will be taken high. Once it has a stable interest network, value of α will be decreased to have more contribution of reputation in the score.

VI. Numerical Results

We have done performance evaluation of reputation system and resource allocation system for a network of 200 nodes. We have also evaluated interest based group formation algorithm for a network of 1000 nodes. We have considered the discrete time instants for the purpose of measurement and estimation in the simulations. Every slot is termed as an iteration. First 50 iterations have been taken as an acquaintance period i.e. a node will allocate their bandwidth without referring to the reputation table.

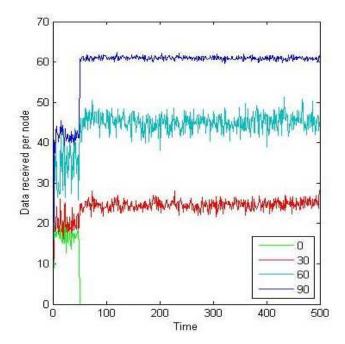


Fig. 1. Average data received per node within the same network. Different graphs show nodes with different shared capacities

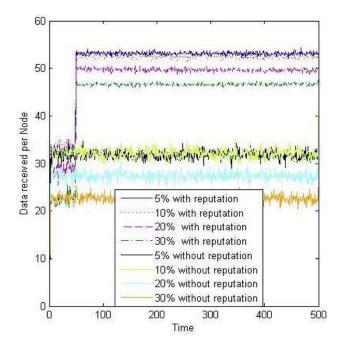


Fig. 2. System performance with different percentage of free riders.

Figure 1 presents the average data received by nodes sharing different amount of resource to the network. Here, it is evident from figure that the node that is sharing more data, is getting better quality of service. Figure 2 shows the performance of system in presence of different percentage of free riders. We can see in figure that from 5% to 10% decay in system performance is almost negligible. After that, system performance decreases by small amount. So we can say that system performance

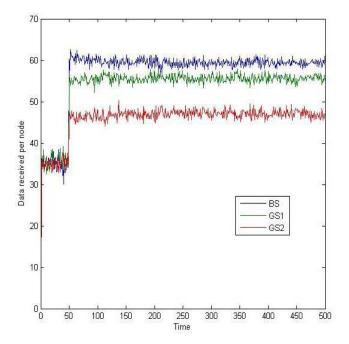


Fig. 3. Data received by nodes with different strategies

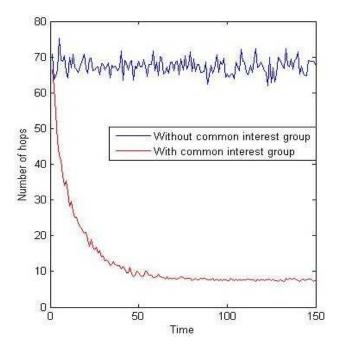


Fig. 4. Average number of nodes queried required for query resolution

does not deteriorate much due to free riders.

Figure 3 shows the data received by peers asking for different amount of data in the network. Here, BS represents the nodes that request for the amount of resource as per its requirement whereas GS1 and GS2 represents the nodes that requests the amount of resource multiple time to their requirement. GS2 requests more times than GS1. Here it can be seen that nodes making request as per their requirement are getting better quality of service whereas nodes that are trying to exploit network by

making requests multiple times are not getting that kind of quality of service. This discourages the tendency of exploitation of by making multiple time requests.

Figure 4 shows the average number of nodes queried required for resolution of query in interest based and non-interest based network. Here, it can be seen that, if node forms interest groups, its query gets resolved in much lesser number of hops than number of hops in other case.

VII. Conclusion

In this paper, we have discussed allocation of resource by node on the basis of reputation. Allocation has been done probabilistically, i.e., requesting node has been offered resource with a probability proportional to its reputation. If total demand of selected nodes is more than offering node's shared capacity, allocation will be done to optimise the gain in reputation of offering node. An algorithm has been proposed for the same. This algorithm also ensures that nodes do not request more than their actual demand. An algorithm for formation of common interest group and shared capacity optimisation has also been proposed. Numerical results show that proposed algorithms work as per the requirement.

References

- H. Kung and C.-H. Wu, "Differentiated admission for peer-topeer systems: Incentivizing peers to contribute their resources," in Proceedings of the 1st Workshop on Economics of Peer-to-Peer Systems.
- [2] M. Feldman, K. Lai, I. Stoica, and J. Chuang, "Robust incentive techniques for peer-to-peer networks," in *Proceedings of the 5th ACM conference on Electronic commerce*, EC '04, (New York, NY, USA), pp. 102–111, ACM, 2004.
- [3] D. Banerjee, S. Saha, S. Sen, and P. Dasgupta, "Reciprocal resource sharing in p2p environments," in *Proceedings of the fourth interna*tional joint conference on Autonomous agents and multiagent systems, AAMAS '05, (New York, NY, USA), pp. 853–859, ACM, 2005.
- [4] A. Satsiou and L. Tassiulas, "Reputation-based resource allocation in p2p systems of rational users," *IEEE Trans. Parallel Distrib. Syst.*, vol. 21, pp. 466–479, Apr. 2010.
- [5] M. Meo and F. Milan, "A rational model for service rate allocation in peer-to-peer networks," in INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings, pp. 1–5, 2006.
- [6] R. T. B. Ma, S. C. M. Lee, J. C. S. Lui, and D. K. Y. Yau, "Incentive and service differentiation in p2p networks: a game theoretic approach," *IEEE/ACM Trans. Netw.*, vol. 14, pp. 978–991, Oct. 2006.
- [7] H. Ma and H. fung Leung, "A demand and contribution based bandwidth allocation mechanism in p2p networks: A gametheoretic analysis," in *Advanced Information Networking and Applications*, 2006. AINA 2006. 20th International Conference on, vol. 1, pp. 1005–1010, 2006.
- [8] Y. Yan, A. El-atawy, and E. Al-shaer, "Ranking-based optimal resource allocation in peer-to-peer networks," in 26th Annual IEEE International Conference on Computer Communications (INFOCOM), 2007.
- [9] G. Liu, H. Shen, and L. Ward, "An efficient and trustworthy p2p and social network integrated file sharing system," in *Peer-to-Peer Computing (P2P)*, 2012 IEEE 12th International Conference on, pp. 203–213, 2012.
- [10] T.-M. Wang, W.-T. Lee, T.-Y. Wu, H.-W. Wei, and Y.-S. Lin, "New p2p sharing incentive mechanism based on social network and game theory," in *Advanced Information Networking and Applications* Workshops (WAINA), 2012 26th International Conference on, pp. 915– 919, 2012.
- [11] Q. Wang, J. Wang, J. Yu, M. Yu, and Y. Zhang, "Trust-aware query routing in p2p social networks," Int. J. Commun. Syst., vol. 25, pp. 1260–1280, Oct. 2012.

- [12] R. Farahbakhsh, N. Crespi, A. Cuevas, S. Adhikari, M. Mani, and T. Sanguankotchakorn, "socp2p: P2p content discovery enhancement by considering social networks characteristics," in *Computers* and Communications (ISCC), 2012 IEEE Symposium on, pp. 000530-000533, 2012.
- 13] R. Gupta and Y. N. Singh, "Trust estimation in peer-to-peer network using blue," *CoRR*, vol. abs/1304.1649, 2013.
 [14] R. Gupta and Y. N. Singh, "Reputation aggregation in peer-to-peer network using differential gossip algorithm," *CoRR*, vol. abs/1210.4301, 2012.
 [15] P. Gupta and Y. N. Singh, "A reputation based framework to
- [15] R. Gupta and Y. N. Singh, "A reputation based framework to avoid free-riding in unstructured peer-to-peer network," CoRR, vol. abs/, 2013.