Improving data availability via an economic lease model in Mobile-P2P networks

A. Mondal*
Institute of Industrial Science,
University of Tokyo, Japan
E-mail: anirban@tkl.iis.u-tokyo.ac.jp
*Corresponding author

S.K. Madria
Department of Computer Science,
University of Missouri-Rolla, USA
E-mail: madrias@mst.edu

M. Kitsuregawa
Institute of Industrial Science,
University of Tokyo, Japan
E-mail: kitsure@tkl.iis.u-tokyo.ac.jp

Abstract: This work proposes LEASE, a novel Mobile-P2P lease-based economic incentive model, in which data requestors need to pay the price (in virtual currency) of their requested data items to data-providers. In LEASE, data-providing mobile peers lease data items to other mobile peers in lieu of a lease payment. The main contributions of LEASE are three-fold. First, its lease model entices even those users, who have no data to provide, to host data items, thereby improving data availability and MP revenues. Second, its economic model discourages free-riding, which improves connectivity due to higher peer participation. Third, its Vickrey auction-based bidding mechanism for leasing items provides effective leasing incentives to peers for improving data availability. Our performance study shows that LEASE indeed improves query response times and data availability in Mobile-P2P networks.

Keywords: Mobile-P2P, free-riding, economic model, lease, data availability.

Reference to this paper should be made as follows: Mondal, A., Madria, S.K., and Kitsuregawa, M. (2008) 'Improving data availability via an economic lease model in Mobile-P2P networks', International Journal of Computational Science and Engineering (IJCSE), Vol. 1, Nos. 1/2/3, pp. 64–74.

Biographical notes: Anirban Mondal received his Ph.D. degree in 2003 from the National University of Singapore. He is currently a Specially Appointed Researcher at the University of Tokyo, Institute of Industrial Science, JAPAN. His research interests include spatial databases, clusters, spatial indexing, peer-to-peer computing, GRID computing, load-balancing in distributed systems and mobile computing.

Sanjay Kumar Madria is an Associate Professor in the Computer Science department of the University of Missouri-Rolla, USA. His research interests include mobile computing, security, web data warehousing, sensor computing and databases.

Masaru Kitsuregawa is a Professor in the University of Tokyo, Institute of Industrial Science, JAPAN. He is the director of the Center for Information Fusion. His research interests include parallel database processing, data mining, data warehousing, speculative transaction processing, storage systems and digital earth.

1 INTRODUCTION

In a Mobile ad hoc Peer-to-Peer (M-P2P) network, mobile peers (MPs) interact with each other in a peer-to-peer (P2P) fashion. Proliferation of mobile devices (e.g., laptops, PDAs, mobile phones) coupled with the ever-increasing popularity of the P2P paradigm (e.g., Kazaa, Gnutella) strongly motivate M-P2P network applications. Mobile devices with support for wireless device-to-device P2P communication are beginning to be deployed such as Microsoft's Zune.

M-P2P applications facilitate mobile users in sharing information with each other *on-the-fly* in a P2P manner. Some M-P2P application scenarios follow:

- A car user could request other car users for information e.g., locations of nearby available parking slots and restaurants, and traffic reports a few miles ahead.
- A pedestrian could request an available taxi nearby his current location.
- Customers in a shopping mall could share information about the cheapest 'Levis' jeans. They could also swap shopping catalogues with each other.
- Mobile users could exchange songs or video-clips (as in a future mobile eBay market).
- Students in a University campus could exchange information concerning lecture timings and venues. They could also share information about on-campus social and cultural activities.

Such P2P interactions among mobile users are generally not freely supported by existing wireless communication infrastructures. Our target applications mainly concern slow-moving objects such as cars on busy streets, people moving in a market-place or students in a campus.

Data availability in M-P2P networks is typically lower than in fixed networks due to frequent network partitioning arising from user movement and users autonomously switching 'on'/'off' their mobile devices. Additionally, as noted in Kamvar et al. (2003), *free-riding* is rampant in P2P environments. Free-riding implies that a large percentage of peers do not provide any data to other peers. Thus, these free-riders only download data from other peers, but they do not upload any data. According to Adar and Huberman (2000), nearly 90% of the peers in Gnutella were free-riders.

Given that the peers in mobile environments generally have limited resources (e.g., energy, bandwidth), free-riding can be reasonably expected to be even more rampant in M-P2P environments than in static P2P environments. To exacerbate the problem, since MPs generally have limited bandwidth, a data-providing MP can make available only few of its data items to be shared (i.e., the **shared data items**) based on the amount of bandwidth that it would like to share, but it has additional data items (i.e., the **unshared data items**) in the memory. Given the ephemeral nature of M-P2P environments, unshared data items may *expire* before they can be made available to M-P2P users, which further decreases data availability.

M-P2P data availability could be significantly improved if free-riders could be enticed to pool in their bandwidth resources by hosting unshared data items. Hence, we propose LEASE, a novel lease-based economic incentive model for effective collaborative data sharing among MPs with limited resources. In LEASE, data-providing MPs *lease* data items to those who do not have any data items to provide. A data item d (originally owned by MP P) is said to be **leased** by P to MP H when P provides d to H for a pre-specified lease period τ, in lieu of a lease payment (in *virtual currency*). During the period τ, H hosts d, and after τ expires, H deletes the copy of d at itself. Notably, P may lease d simultaneously to multiple MPs. In case any updates are required to the data (e.g., traffic reports in transportation application scenarios), P sends the updates to H. We shall henceforth refer to a data-providing MP P as a **provide-MP**, and the host MP H as a **host-MP**.

Each data item has a *price* (in *virtual currency*). As in Mondal et al. (2006a), data item price depends on access frequency, data quality (e.g., image resolution, audio quality) and the estimated response time for accessing the data item. A query issuing MP pays the *price* of the queried data item to the query-serving MP. Thus, LEASE provides an incentive for free-riding MPs to act as host-MPs so that they can earn **revenue** for issuing their own requests. **Revenue** of an MP is defined as the difference between the amount of virtual currency that it earns (by providing data) and the amount that it spends (by requesting data).

According to Turner and Ross (2004), virtual currency is suitable for P2P environments due to high transaction costs of micro-payments in real currency. Virtual currency has several advantages such as low cost of micro-economic transactions, high level of security due to effective traceability of transactions and convenience of use (i.e., cashless travel). These advantages of virtual currency are likely to be spread by word-of-mouth. We believe that virtual currency would work in practice as larger number of users start using it, thereby making it possible to obtain more services by using the virtual currency.

Interestingly, some efforts in this direction have already been undertaken. A virtual currency, which is designated as Edy (Euro-Dollar-Yen), is already in existence in Japan. Edy can also be in-built into mobile phones, thus it can facilitate micro-economic payments for M-P2P transactions between strangers as well as micro-economic payments for Internet auctions for users, who use mobile Internet services in Japan. Notably, smart cards such as the Suica card used for Japan's train services and the Octopus Card used for Hong Kong's mass transit system may also be viewed as forms of virtual currency. These cards are generally also useful for making micro-payments. Incidentally, Elrufaie and Turner (2004) have discussed secure virtual currency payments.

Leasing benefits both provide-MPs and host-MPs. It facilitates a provide-MP in earning revenue from its unshared data items even without hosting them, especially since unshared data items may expire. It helps a host-MP in earning revenue using other MPs' data items. In the absence of a lease model, MPs without any data to provide cannot earn any revenue, thereby decreasing the overall MP participation. In M-P2P networks, leasing is better than *buying* (permanent ownership transfer) since data items have expiry times, hence their value depreciates significantly over time. Moreover, host-MPs wish to host as many 'hot' data items as possible to maximize their revenues within a short span of time in which they are within the transmission range of other MPs.

The main contributions of LEASE are three-fold:

1. Its lease model entices even those users, who have no data to provide, to host data items, thereby improving data availability and MP revenues.

2. Its economic model discourages free-riding, which improves connectivity due to higher peer participation.

3. Its Vickrey auction-based bidding mechanism for leasing items provides effective leasing incentives to MPs for improving data availability.

Higher peer participation leads to better data availability due to higher available bandwidth and better connectivity. Existing mobile replication schemes (Hara and Madria, 2006) do not combat free-riding, while M-P2P incentive schemes (Wolfson et al., 2004; Xu et al., 2006) do not entice free-riders, which have no data, to provide service.

We have evaluated the performance of LEASE w.r.t. a non-economic model **NL (No_Lease)** since existing M-P2P proposals do not address economic lease-based models. In NL, leasing is not performed and querying is broadcast-based. Our performance evaluation shows that the incentives provided by LEASE entice better MP participation, thus it significantly outperforms NL in terms of query response times and data availability.

At higher workload skew, the performance gap between LEASE and NL increases due to multiple copies of data items created by LEASE in response to load-imbalance conditions. Furthermore, when leasing is performed more frequently, LEASE improves in performance over NL essentially due to the leasing mechanism being able to react more quickly to changes in the access patterns. Our results also indicate that LEASE exhibits good scalability. In essence, our performance evaluation demonstrates that LEASE indeed improves query response times and data availability in M-P2P networks.

To our knowledge, this is the first work to propose a lease-based economic incentive model for M-P2P networks. The remainder of the paper is organized as follows. Section 2 discusses related work, while Section 3 describes the LEASE economic model. Section 4 discusses the Vickrey auction-based leasing strategy used by LEASE. Section 5 describes the algorithms in LEASE. Section 6 reports the performance evaluation. Finally, we conclude in Section 7 with directions for future work.

2 RELATED WORK

This section provides an overview of existing works.

For improving data availability in mobile ad hoc networks (MANETs), Hara and Madria (2006) consider replication of the data. In particular, their proposed **E-DCG+** approach creates groups of MPs that are biconnected components in a MANET, and shares replicas in larger groups of MPs to provide high stability. An RWR (read-write ratio) value in the group of each data item is calculated as a summation of RWR of those data items at each MP in that group. In the order of the RWR values of the group, replicas of items are allocated until memory space of all MPs in the group becomes full. Each replica is allocated at an MP, whose RWR value to the item is the highest among MPs that have free memory space to create it. Notably, Hara and Madria (2006) do not consider any economic scheme to entice peer participation. Moreover, they do not address M-P2P architecture. P2P replication suitable for mobile environments has been incorporated in systems such as ROAM (Ratner et al., 2001), Clique (Richard et al., 2003) and Rumor (Guy et al., 1998). However, these systems do not consider economic models to combat free-riding.

Interestingly, several economic incentive schemes for combating free-riding in MANETs have been proposed (Buttyan and Hubaux, 2003; Chen and Nahrstedt, 2004; Crowcroft et al., 2003; Srinivasan et al., 2003). These schemes essentially provide incentives to peers for relaying messages. However, they do not address M-P2P architecture and they do not entice free-riders to host data.

Xu et al. (2006) provide incentives to MPs for participation in the dissemination of reports about resources in M-P2P networks. Each disseminated report contains information concerning a spatial-temporal resource e.g., availability of a parking slot at a given time and location. Wolfson et al. (2004) consider opportunistic resource information dissemination in transportation application scenarios. An MP transmits its resources to the MPs that it encounters, and obtains resources from them in exchange. Notably, these works primarily address data dissemination with the aim of reaching as many peers as possible, while we consider on-demand services. Moreover, they do not consider incentives for free-riders to host data.

Ensuring secure payments using a virtual currency have been discussed (Elrufaie and Turner, 2004; Zhong et al., 2003). Notably, these secure payment schemes are complementary to our proposal, but they can be used in conjunction with our proposal. Garyfalos and Almeroth (2004) describe Coupons, an incentive scheme that is inspired by the eNcentive framework (Ratsimor et al., 2003), which allows mobile agents to spread digital advertisements with embedded coupons among mobile users in a P2P manner. Straub and Heinemann (2004) propose adPASS, which also deploys coupons for providing incentives.

Kremer et al. (2002) examine non-repudiation systems, which can be incorporated to control the deceiving behaviour of peers. Chakravorty et al. (2005) propose MoB, which is an open market collaborative wide-area wireless data services architecture that can be used by mobile users for opportunistically trading services with each other.

Schemes for combating free-riding in static P2P networks include EigenTrust scores to capture participation criteria (Kamvar et al., 2003). However, these schemes are too static to be deployed in M-P2P networks since they assume peers' availability and fixed topology. Economic schemes for resource allocation in wireless ad hoc networks have also been proposed (Liu and Issarny, 2004; Xue et al., 2005). However, these schemes do not address free-riding. Economic schemes have also been discussed for resource allocation in distributed systems (Ferguson et al., 1993; Kurose and Simha, 1989). However, such schemes do not address M-P2P issues such as node mobility, free-riding, frequent network partitioning and mobile resource constraints.

3 THE LEASE ECONOMIC MODEL

This section discusses the economic model of LEASE. Table 1 summarizes the notations used in this paper.

Symbol	Significance
d	A given data item
M_I	Query-issuing MP for d
M_S	Query-serving MP for d
η	Recent Access frequency of d
δ	Distance between M_I and M_S
DQ	Data quality of d
$size$	Size of d
Ex	Time to Expiry time of d
N_{Copies}	Number of leased copies of d
BA_{M_S}	Bandwidth of the query-serving MP for d
σ_{M_S}	Connectivity of the query-serving MP for d
J_{M_S,t_j}	Job-queue length of the query-serving MP at time t_j

Table 1: Summary of Notations

Data item pricing

In the economic model of LEASE, a query-issuing MP pays the price of its requested data item to the query-serving MP. This provides an incentive for MPs to serve queries, thereby effectively combating free-riding. Now let us see how LEASE computes the price of a given data item.

Using the notations in Table 1, price μ of a data item is computed as follows:

$$\mu = \int_{t_1}^{t_2} \int_{0}^{\delta} (\eta \, dt \times (1/\delta^2) \, d\delta \times DQ \times BA_{M_S})/J_{M_S,t_j} \quad (1)$$

where $[t_2 - t_1]$ represents a given time period. Such time periods are essentially application-dependent. Observe that in Equation 1, we consider the *recent* access frequency η of a given data item d. This takes into consideration the ephemeral nature of M-P2P environments. Notably, each MP M maintains recent access information of data items at itself. M computes the value of η for a given item d by using exponential moving averages of the access frequencies for d for the last N time periods. For our application scenarios, $N = 5$ was found to be reasonable. Notably, for unshared data items, the access frequency η refers to the number of access failures. The price of d increases with increasing value of η to reflect item importance and popularity since items with higher access frequencies are important to M-P2P users.

The price μ of d decreases as the *Euclidean distance* δ between the query issuing MP M_I and the query serving MP M_S (during the time of query issue) increases. This is because larger distance between M_I and M_S implies more delays in the query result reaching M_I. Interestingly, the bandwidth BA_{M_S} allocated by M_S for the download of d from itself also influences query response time. If M_S allocates more bandwidth for the download of d, query response time decreases and vice versa. Thus, as δ decreases and BA_{M_S} increases, μ increases due to faster query response time. Furthermore, as the job queue length J_{M_S,t_j} of M_S increases, μ decreases since M_S's response time for queries on d increases due to higher load. Notably, the dependence of price on query response time is in consonance with the strict deadline requirements of M-P2P queries.

In Equation 1, DQ reflects the quality of data (e.g., image resolution, audio quality) provided by M_S for queries on d. The value of DQ is determined as in our previous works (Mondal et al., 2006a,b,c), where we considered three discrete levels of DQ i.e., *high*, *medium* and *low*, their values being 1, 0.5 and 0.25 respectively. Each MP maintains a table $T_{\epsilon,DQ}$, which contains the following entries: (x%, high), (y%, medium), (z%, low), where x, y, z are error-bounds, whose values are essentially application-dependent and pre-specified by the system at design time. Thus, DQ is computed using the table $T_{\epsilon,DQ}$, which is replicated at each MP and is the same for each MP.

Revenue of an MP

Now let us see how the revenue of a given MP M is computed. Suppose M makes p data items available, and let the price and access frequency of the i^{th} such data item be μ_i and $accs_i$ respectively. Additionally, suppose M leases q of its unshared data items, and let the total sum of lease payments earned by M for the i^{th} leased item over all the time-periods be L_{Get_i}. (We defer the discussion concerning the computation of L_{Get_i} to Section 4.) Observe that M does not earn any currency for its unshared data items that it does not lease, thereby providing further incentive to M to perform leasing. Furthermore, suppose M forwards w messages of other MPs. For forwarding a message, M earns a small constant commission K. Thus, the amount E of currency earned by M is computed as follows:

$$E = \sum_{i=1}^{p}(\mu_i \times accs_i) + \sum_{i=1}^{q} L_{Get_i} + (K \times w) \quad (2)$$

To compute the amount of currency spent by M, suppose the total number of items queried by M is r. Let the price and access frequency (of M) for the i^{th} such queried item be μ_i and $accr_i$ respectively. Additionally, suppose M leases s data items from other MPs, and let the total sum of lease payments that M should make for the i^{th} item that it leases (over all the time-periods) be L_{Pay_i}. (We shall discuss the computation of L_{Pay_i} in Section 4.) Furthermore, suppose v of M's messages are forwarded by other MPs, the relay commission for each such message being K (as in Equation 2). Thus, the amount S of currency spent by M is computed below:

$$S = \sum_{i=1}^{r}(\mu_i \times accr_i) + \sum_{i=1}^{s} L_{Pay_i} + (K \times v) \quad (3)$$

The revenue of an MP is the difference between the amount of currency that it earns and the amount that it spends. Thus, the revenue R of M is computed as follows:

$$R = E - S \quad (4)$$

where the values of E and S are computed in Equations 2 and 3 respectively.

4 VICKREY AUCTION-BASED LEASING

This section discusses the leasing strategy used in LEASE. In particular, LEASE performs leasing by means of Vickrey auctions. Furthermore, peers are provided two different leasing options and they are allowed to choose the option which best suits their requirements.

Role of the provide-MPs and the host-MPs

Each provide-MP maintains recent read-write logs (including timestamps) of its own data items, recent access failures of its unshared items and details (e.g., lease duration, lease payments) of the data items that it leases. This information helps provide-MPs to select their respective shared and unshared data items. Each host-MP maintains recent access information of data items based on queries that pass through itself. Such information facilitates host-MPs in selecting data items that they want to host.

A provide-MP P makes available at itself (i.e., *shares*) data items, with higher revenue-earning potential γ for maximizing its revenue, while leasing out some of its *unshared* data items. Given that $\mu_{i,P}$ is the price of a data item i at P and $acc_{i,P}$ is the recent access frequency of i at P, $\gamma = \mu_{i,P} \times acc_{i,P}$. (For data items that P is currently not making available, $acc_{i,P}$ is the number of times a query failed to obtain the data item at P.) P avoids leasing frequently updated data items due to the high communication overhead (e.g., energy, bandwidth) required for maintaining the consistency of such items.

Vickrey auctions in LEASE

The leasing strategy of LEASE uses a *Vickrey auction* mechanism (Vickrey, 1961). For simplicity, we shall briefly explain the working of Vickrey auctions for a single item, although Vickrey auctions are also applicable to the case for multiple items. Given an item d to be auctioned, multiple prospective buyers send their *sealed* bids to the single seller S of d, and S selects the highest bidder B_{max} as the winner of the auction for d. Then B_{max} pays the price of the second-highest bid to S for obtaining d. This is in contrast with the standard English auction, where B_{max} would have to pay its own bid price to obtain d.

Vickrey auctions are applicable to LEASE since we can view the provide-MPs as the sellers, and the host-MPs as the buyers. Thus, host-MPs would send bids to the respective provide-MPs for the items to be leased. Notably, Vickrey auctions are advantageous for incentive mechanisms since they provide incentives to bidders to bid according to their perception of the true value of an item without considering the bids of other sellers (Liu and Issarny, 2004). Thus, host-MPs have higher incentives to bid for the items on lease.

Periodically, a provide-MP P broadcasts its list of unshared data items, which have low write frequencies, for finding prospective host-MPs to host these items by means of the leasing mechanism. Given its unshared data item d, P receives bids for d from prospective host-MPs and selects the higher-bidding MPs for leasing d to maximize its own revenue. Observe that higher-bidding MPs are likely to be the ones, which are able to provide better *quality of service* for d. This is because only the MPs with better resources for providing good service would bid higher since they need to recoup the investment cost in leasing d so that they can earn some revenue by hosting d. (Recall that in LEASE, better quality of service equates to higher data item prices.)

We define the connectivity of an MP as the number of its one-hop neighbours. Moreover, higher-bidding MPs are also likely to have high connectivity since MPs, which are not well-connected, would not bid high since they would be unlikely to be able to serve adequate number of queries for d within short response times to offset their investment in leasing d. Thus, P prefers to lease d to MPs with higher connectivity to facilitate it in sharing its data items with as many MPs as possible, thereby enabling it to earn more revenue. In essence, P leases d to MPs based on their quality of service and connectivity.

Number of copies of a leased data item

Given an unshared data item d, P decides the number of copies of d to be leased based on the revenue λ that it wishes to obtain from leasing d. P computes λ as follows:

$$\lambda = pc \int_{t_1}^{t_2} (\eta_d \times \mu_d) \quad (5)$$

where $[t_2 - t_1]$ is a given time period, pc is a percentage value, η_d is the number of failed queries on d, and μ_d is the price of d. In Equation 5, the term $(\eta_d \times \mu_d)$ reflects P's estimated lost revenue due to not making d available. Thus, λ is pc% of P's estimated lost revenues. Hence, the estimated revenue from leased data items is shared between the provide-MP and the host-MPs to ensure fairness. Allowing the host-MPs to earn a certain percentage of the revenues from d provides adequate incentive for them to host d since they also incur energy and bandwidth-related costs due to downloads of d. The value of pc is determined by P depending upon the demand of d across the network and the amount of currency that P wishes to obtain by leasing d. Observe that if P sets a high value of pc, relatively fewer prospective host-MPs would be willing to bid for d because they would be likely to earn less amount of currency by hosting d. On the other hand, if the value of pc is too low, P would not gain much currency by leasing d. From our preliminary experiments, we observed that values of pc between 40% and 60% work well for LEASE.

P essentially sums up the bids for d starting from the second-highest bid (due to Vickrey auction properties) until the total value of the bids is greater than or equal to λ. Then, P leases d to the corresponding MP(s) in accordance with the Vickrey auction mechanism. Notably, unlike existing works, LEASE determines the number of copies to be leased based on revenue.

Leasing options

Given a provide-MP P, which leases a data item d for a pre-specified lease period τ, to a host-MP H, LEASE provides two different leasing options. Each option is associated with a *license*, which specifies the freedom which H has in regard to downloads of d from itself.

1. **Bulk lease:** H is provided an *unlimited* license for downloads of d from itself i.e., H can allow any number of downloads of d within the lease period τ. H would be likely to consider this option if it perceives that d has high demand.

2. **Limited lease:** H is provided a *limited* license for downloads of d from itself i.e., H can only allow N downloads of d. Furthermore, P ensures that H is not able to allow more than N downloads of d from itself by means of digital watermarking technologies, which lock d after N downloads.

H decides which leasing option to choose based on the estimated demand for d. Moreover, for the *limited lease* option, the value of N is decided by H depending upon its perceived demand for d. As the value of N increases, the lease payment made by H to P also increases.

Determining the bid price for a host-MP

Host-MPs keep track of queries that pass through themselves to estimate the demand for different data items. This information facilitates them in deciding the data items to bid for, their bid values for these items and the leasing option to choose. A host-MP H bids for data items with higher revenue-earning potential γ for maximizing its revenue. The number of data items for which H bids depends upon its available bandwidth and memory space.

Under the *bulk lease* option, H bids the amount β_{BL} of currency for a given data item d based on d's revenue-earning potential, which depends upon d's popularity, quality, size, estimated expiry time, existing number of copies, amount of bandwidth that it would likely make available for d, and its current job-queue length. (Recall that d's price at H depends upon H's bandwidth and job-queue length.) Using Table 1 (see Section 3), H computes β_{BL} as follows:

$$\beta_{BL} = \int_{t_1}^{t_2} (\eta\, dt \times DQ \times Ex \times BA_{M_S} \times \sigma_{M_S}) / (N_{Copies} \times size \times J_{M_S,t_j}) \qquad (6)$$

where $[t_2 - t_1]$ is a given time period. The access frequency η is based on the queries for d that passed through H. A data item *expires* when its access frequency falls below a certain application-dependent threshold. Data items with higher time-to-expiry facilitate H in earning more revenue by hosting them. Higher bandwidth of H implies better response time for queries on d, while larger job-queue length signifies higher load on H, which suggests increased query response times. Smaller-sized data items help H to maximize its revenue per unit of its limited memory space.

Notably, the bid price of H increases with its increasing connectivity σ_{M_S} because H can afford to bid higher when it has more connectivity. Furthermore, the number η of queries for item d, which passed through H during the previous time-periods, can be interpreted to be the perceived demand for d across the whole network. However, the demand for d at H would depend on the total number of existing copies of d in the network. Hence, the term N_{Copies} occurs in the denominator of Equation 6.

Under the *limited lease* option, H bids the amount β_{LL} of currency for a given data item d based on d's revenue-earning potential, which depends upon d's quality, size, estimated expiry time, amount of bandwidth that H would likely make available for d, its connectivity and its current job-queue length. Using Table 1 (see Section 3), H computes β_{LL} as follows:

$$\beta_{LL} = \int_{t_1}^{t_2} (N\, dt \times DQ \times Ex \times BA_{M_S} \times \sigma_{M_S}) / (size \times J_{M_S,t_j}) \qquad (7)$$

Interestingly, unlike Equation 6, the above equation does not consider the item popularity η because H would not use the *limited lease* option for items with high access frequencies. Furthermore, in contrast with Equation 6, the above equation does not consider the number of copies of the leased item because the demand for the leased item at H is factored into the equation by the inclusion of N, where N is the number of times the data item is allowed to be downloaded from H. H estimates the value of N based on the queries for d which pass through itself.

Notably, the values of L_{Get_i} and L_{Pay_i} in Equations 2 and 3 respectively are determined based on β_{BL} or β_{LL}, depending upon the leasing option under consideration.

Incidentally if a prospective bidder host-MP H wins a bid for an item d, LEASE stipulates that H is required to lease d from the corresponding provide-MP P in accordance with the results of the Vickrey auction i.e., H pays the bid price of the next-highest bidder as lease payment to P. This ensures the avoidance of high communication overheads due to renegotiations and possible reallocations associated with leasing data items. However, LEASE makes an exception under certain special circumstances e.g., when H does not have adequate memory space to host d or when H is running out of energy.

Such special circumstances may occur because when MPs bid for an item, they are not guaranteed to win that item, thus it is highly likely for any given MP to bid for multiple items simultaneously. As a result, sometimes it may happen that an MP wins more leased items than it had previously estimated. Under such circumstances, the bid winner H pays a penalty to the corresponding provide-MP P. The penalty is equal to the difference between the bid price that H was supposed to pay P (according to the rules of the Vickrey auction) and the bid-price of the bidder, which eventually leases the item from P. Observe that the penalty serves as a deterrent for MPs from indulging in significant over-bidding, while ensuring that in the unlikely case of occurrence of over-bidding, the leased items are at least allocated at MPs, which have adequate memory space and/or energy to serve requests for these items with good quality of service.

Prevention of illegitimate leasing behaviours

Data providers periodically broadcast the unique identifiers of host-MPs, to whom they have leased their data items. Thus, MPs can download *updated* copies of data items from the authorized lease-holders, thereby improving the quality of service. (Provide-MPs send updates only to authorized host-MPs.) In case a host-MP H illegitimately hosts a given data item d or if H continues to host d after its lease period of d has expired, other MPs (e.g., relay MPs through which messages for downloads of d would pass) would inform the corresponding provide-MP P, and P would blacklist H. Periodically, provide-MPs broadcast their list of blacklisted MPs. Blacklisted MPs have to pay double the lease payment the next time they want to lease data items from any provide-MP, which acts as a deterrent. Furthermore, digital watermarking technologies, used in conjunction with the above methods, also deter illegitimate leasing behaviours.

Host-MPs make the lease payments to provide-MPs at the time of expiry of the lease so that host-MPs can earn revenue from hosting data items before they pay for the lease. This facilitates seamless integration of newly joined MPs, which may initially be unable to make the lease payment. Host-MPs, which fail to make the lease payment at the end of the lease expiry period, are blacklisted, thereby deterring malicious MPs from abusing the leasing system.

5 ALGORITHMS IN LEASE

Figure 1 depicts the algorithm for a provide-MP P. Lines 1-8 depict how P determines its set of shared and unshared items. Thus, in Line 8, CL is the set of unshared data items of P, and hence, CL constitutes the set of data items for lease. In Line 3, write frequency WF_d of a data item d is computed as (nw_d / τ), where nw_d is the number of writes on d and τ is the lease period. Write frequency threshold TH_{WF} is computed as the average write frequency of all the shared and unshared items in P. In Line 9 of Figure 1, $n = 3$ or $n = 4$ were found to be reasonable values for our application scenarios (as indicated by preliminary experimental results). In Line 9, P's broadcast message contains the unshared data items and their prices to help prospective host-MPs to determine their bid values.

Algorithm *LEASE_Provide_MP*
Spc: Its available memory space
(1) Sort all its data items in in descending order of their
    revenue-earning potential γ into a list L.
(2) for each data item d in L
    /* WF_d is d's write frequency, TH_{WF} is write frequency threshold */
(3)   if ($WF_d < TH_{WF}$)
(4)   if ($size_d \leq$ Spc)   /* $size_d$ is the size of d */
(5)   Fill up its memory space with d
(6)   Spc = Spc - $size_d$
(7)   if (Spc = = 0)   **exit**
(8) Create set CL comprising its **unshared** data items
    /* CL is the set of candidate data items for lease */
(9) Broadcast the set CL to its n-hop neighbours
(10) for each data item d in CL
(11)   Receive bids from prospective host-MPs, which wish to host d
(12)   Arrange the bids in descending order of bid value
(13) $Bid_{Sum} = 0$
(14)   for each bid β from host-MP i
(15) $Bid_{Sum} = Bid_{Sum} + \beta$
(16)   if $Bid_{Sum} \leq \lambda$
(17)   Add i to set $Host_d$
(18)   if set $Host_d$ is non-empty
(19)   Lease d to the MPs in set $Host_d$ with next-highest
             bid values as lease payment /* Vickrey auction mechanism */
(20)   Initialize set $Host_d$ by making it a NULL set
end

Figure 1: LEASE algorithm for provide-MP

Lines 10-20 show the bidding process of LEASE by means of Vickrey auctions. In Lines 11-12, note that the bids could be based either on the *bulk lease* option or on the *limited lease* option, and the sorting of the bids is based on bid value, irrespective of the lease option. Lines 12-17 depict how P determines the number of copies of each leased item based on revenue. In Lines 14-15, the value of β is computed by means of Equations 6 or 7, depending upon whether the bid is for the *bulk lease* or for the *limited lease* option. In Line 16, the value of λ is computed by Equation 5. In Line 19, observe how the *next-highest* bid price is paid by the winning bidders to P. This is in consonance with the mechanism of Vickrey auctions.

Algorithm *LEASE_Host_MP*

CL_i: Candidate data items for lease from provide-MP i
Spc: Its available memory space $\times$ 1.3
(1) for each provide-MP i
(2) Receive broadcast message from i containing items for lease
(3) Add all data items in CL_i to a set *bigCL*
(4) Sort all data items in *bigCL* in descending order of γ
(5) for each data item d in *bigCL*
(6) /* $size_d$ is the size of d */
(7) if ($size_d \leq Spc$)
(8) Add d to a set *BID*
(9) Spc = Spc - $size_d$
(10) if (Spc = = 0) **exit**
(11) for each data item d in set *BID*
(12) Decide whether to bid for the *bulk lease* or the *limited lease* option
(13) Send the bid of β_d to the corresponding provide-MP
(14) if bid is successful
(15) if memory space and energy are adequate
(16) Obtain d from corresponding provide-MP by making the lease payment
(17) else
(18) Withdraw the bid by making a penalty payment to the corresponding provide-MP
end

Figure 2: LEASE algorithm for host-MP

Figure 2 depicts the algorithm executed by a host-MP H. In Lines 1-4, H receives broadcast messages from different provide-MPs concerning data items for leasing and sorts these items in descending order of their revenue-earning potential γ. This is because H wishes to bid for items with relatively higher revenue-earning potential to maximize its revenue. In Lines 5-10, H *simulates* the choice of data items that it wishes to bid for. Observe that *Spc* is 30% more than H's actual available memory space. This is because H may not necessarily be able to obtain a lease for all the data items that it bids for since other MPs may outbid H, hence it is a *simulation*. Thus, H greedily *simulates* the filling up of its memory space by data items with higher revenue-earning potential.

In Line 12, H decides the leasing option to bid for based on the access frequency of the data item for leasing. For items with high access frequencies, H selects the *bulk lease* option, while for other items, it selects the *limited lease* option. In Line 13, the value of β_d is computed by either Equation 6 or 7, depending upon the leasing option selected by H. In Lines 14-18, in case H's bid is successful and it has adequate memory space and energy, it makes the lease payment to the corresponding provide-MP P and obtains the item for lease from P. However, in the case of H's memory space or energy being inadequate, H withdraws its bid by paying a penalty, which we had earlier discussed in Section 4.

6 PERFORMANCE EVALUATION

This section discusses our performance evaluation.

MPs move according to the *Random Waypoint Model* (Broch et al., 1998) within a region of area 1000 metres $\times$1000 metres. The *Random Waypoint Model* is appropriate for our application scenarios, which involve random movement of users. A total of 100 MPs comprise 30 data-providers and 70 free-riders (which provide no data). Each data-provider owns 8 data items comprising 4 *shared* items and 4 *unshared items*. Each query is a request for a single data item. 10 queries/second are issued in the network, the number of queries directed to each MP being determined by a highly skewed Zipf distribution with Zipf factor of 0.9. Communication range of all MPs is a circle of 100 metre radius. Table 2 summarizes our performance study parameters. In Table 2, the lease time-period *LP* refers to the time intervals at which data items are put up for lease by provide-MPs.

Parameter	Default value	Variations
No. of MPs (N_{MP})	100	20,40,60,80
Percentage of data-providers	30%	
Percentage of free-riders	70%	
Zipf factor (ZF)	0.9	0.1,0.3,0.5,0.7
Lease time-period (LP) 10^2 seconds	4	8,12,16
Queries/second	10	
Bandwidth between MPs	28 to 100 Kbps	
Probability of MP availability	50% to 85%	
Size of a data item	50 to 350 Kb	
Memory space of each MP	1 to 1.5 MB	
Speed of an MP	1 to 10 metres/s	
Size of message headers	220 bytes	

Table 2: Performance Study Parameters

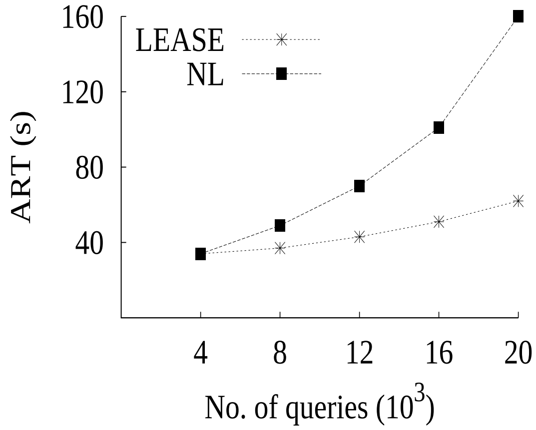
Our performance metrics are **average response time** (**ART**) of a query and **data availability** (**DA**). ART is computed as follows:

$$ART = (1/N_Q)\sum_{j=1}^{N_Q}(T_{f,j} - T_{i,j}) \quad (8)$$

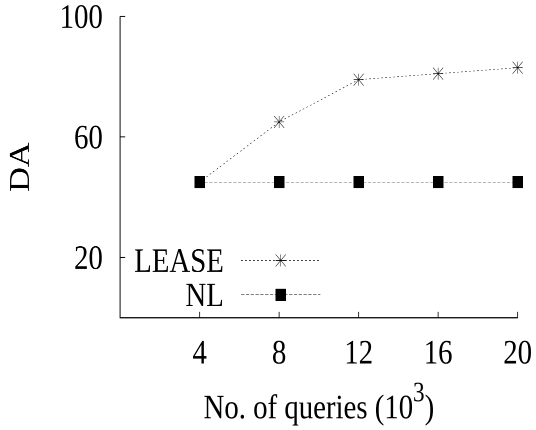
where $T_{i,j}$ is the query issuing time for the j^{th} query, $T_{f,j}$ is the time of the query result reaching the query issuing MP for the j^{th} query, and N_Q is the total number of queries. ART includes data download time, and is computed only for successful queries.

The computation of DA follows:

$$DA = ((N_S/N_Q) \times 100) \quad (9)$$



(a) Average Query Response Time



(b) Data Availability

Figure 3: Performance of LEASE

where N_S is the number of successful queries and N_Q is the total number of queries. Queries can fail due to MPs being switched 'off' or due to network partitioning.

As reference, we adapt a non-economic model **NL (No_Lease)** since existing M-P2P proposals do not address economic lease-based models. In NL, leasing is not performed and querying is broadcast-based. As NL does not provide incentives for free-riders to become host-MPs, only a single copy of any given data item d exists at the owner of d.

Performance of LEASE

Figure 3 depicts the performance of LEASE using default values of the parameters in Table 2. Leasing procedures are initiated only after the first 4000 queries, hence both LEASE and NL initially show comparable performance. The ART of both LEASE and NL increases with time due to the skewed workload ($ZF = 0.9$), which overloads some of the MPs that store 'hot' data items, thereby forcing queries to incur high waiting times and consequently high ART. However, over time, the economic incentives of LEASE entice more MPs to host data items, thereby increasing the resources (e.g., bandwidth, memory space) in the network for creating multiple (leased) copies for the same data item to facilitate load-balancing as well as reduction of the total query hop-counts. LEASE also considers the connectivity of host-MPs, which further decreases its querying hop-counts, thereby decreasing ART. In Figure 3b, DA eventually plateaus for LEASE due to network partitioning and unavailability of some of the MPs.

In contrast, the non-economic nature of NL does not entice the free-riders to host data items via leasing, thus the ART of NL keeps increasing due to overloading of MPs storing 'hot' data items. For NL, DA remains relatively constant since it depends only on the probability of availability of the MPs.

Effect of variations in the workload skew

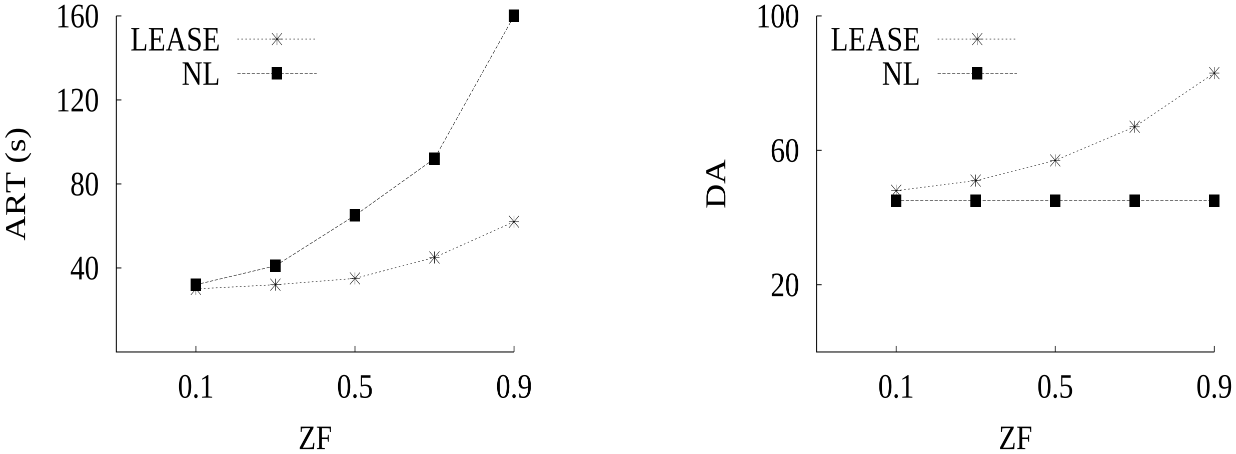
Figure 4 depicts the effect of variations in the workload skew by varying the zipf factor (ZF) associated with queries. As ZF increases (i.e., high skew), ART also increases for both LEASE and NL due to load-imbalance. Observe that the ART-related performance gap between LEASE and NL keeps increasing with increase in ZF. This is because in contrast with NL, LEASE creates multiple copies of data items by means of leasing. However, at low values of ZF (i.e., low skew), the workload skew is too low to exploit the multiple copies created by LEASE. Hence, LEASE and NL perform comparably for lowly skewed workloads.

Incidentally, DA remains relatively constant for NL since it depends only on the probability of availability of the MPs. However, for LEASE, DA improves with increasing value of ZF due to more copies of the data items being created (by means of leasing) in response to the higher workload skew. The explanation for DA eventually reaching a plateau is essentially similar to that of Figure 3b.

Effect of variations in the lease time-period

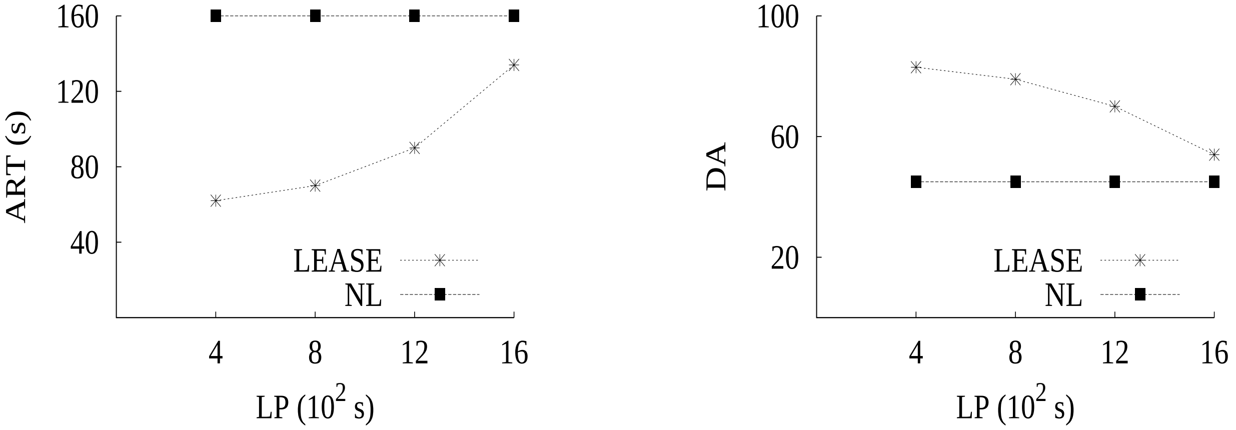
Recall that the lease time-period LP refers to the time intervals at which data items are put up for lease by provide-MPs. Figure 5 depicts the results of varying LP. Since NL does not perform any leasing, it is independent of LP. Thus, the performance of NL remains constant with variations in LP. This experiment was done by issuing 20000 queries at query interarrival rate of 10 queries/s. Thus, for LP values of 400s, 800s, 1200s and 1600s, there were 5, 2, 1 and 1 lease time-periods respectively.

When LP is low, data items are put up for lease more frequently, hence the leasing mechanism is able to react quickly to changing access patterns, hence LEASE exhibits lower ART and higher DA as the value of LP decreases. As LP increases, leasing is performed less frequently, hence the performance of LEASE in terms of both ART and DA



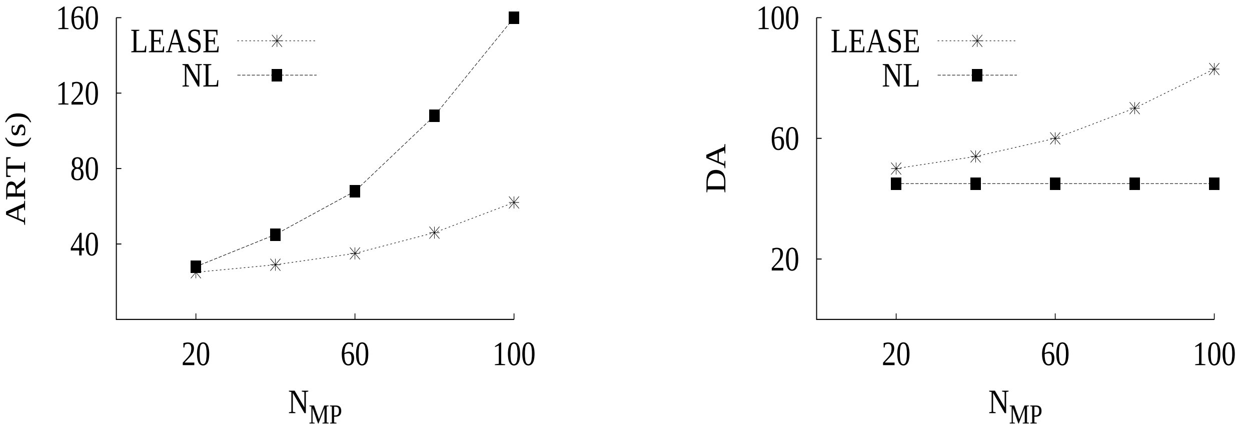
(a) Average Query Response Time (b) Data Availability

Figure 4: Effect of varying the workload skew



(a) Average Query Response Time (b) Data Availability

Figure 5: Effect of varying the lease time-period



(a) Average Query Response Time (b) Data Availability

Figure 6: Effect of varying the number of MPs

degrades. Interestingly, for LP values of 1200 s and 1600 s respectively, there is only one period during which items are put up for leasing. However, the performance varies for these time-points because the leasing occurs earlier for $LP = 1200$ s.

Effect of variations in the number of MPs

To test LEASE's scalability, we varied the number N_{MP} of MPs, while keeping the number of queries proportional to N_{MP}. In each case, 30% of the MPs were data-providers, the rest being free-riders. As the results in Figure 6 indicate, ART increases for both approaches with increasing N_{MP} due to larger network size. At higher values of N_{MP}, LEASE outperforms NL due to the reasons explained for Figure 3. As N_{MP} decreases, the performance gap decreases due to limited leasing opportunities, which results in lesser number of copies for leased data items, thereby making the effect of leasing less prominent.

7 CONCLUSION

We have proposed LEASE, a novel Mobile-P2P lease-based economic incentive model, in which data requestors need to pay the price (in virtual currency) of their requested data items to data-providers. In LEASE, data-providing mobile peers lease data items to other mobile peers in lieu of a lease payment. The main contributions of LEASE can be summarized as follows. First, its lease model entices even those users, who have no data to provide, to host data items, thereby improving data availability and MP revenues. Second, its economic model discourages free-riding, which improves connectivity due to higher peer participation. Third, its Vickrey auction-based bidding mechanism for leasing items provides effective leasing incentives to MPs for improving data availability. In essence, LEASE facilitates the collaborative harnessing of limited mobile peer resources for improving data availability. Our performance study shows that LEASE indeed improves query response times and data availability in Mobile-P2P networks. In the near future, we plan to incorporate game-theoretic ideas into the LEASE model.

REFERENCES

Adar, E. and Huberman, B. A. (2000) 'Free riding on Gnutella', *Proc. First Monday*, Vol. 5, No. 10.

Broch, J. et al. (1998) 'A Performance Comparison of Multi-hop Wireless Ad Hoc Network Routing Protocol', *Proc. MOBICOM.*

Buttyan, L. and Hubaux, J.P. (2003) 'Stimulating Cooperation in Self-Organizing Mobile Ad Hoc Networks', *ACM/Kluwer Mobile Networks and Applications*, Vol. 8, No. 5, pp. 579–592.

Chakravorty, R. et al. (2005) 'MoB: a mobile bazaar for wide-area wireless services', *Proc. MobiCom.*

Chen, K. and Nahrstedt, K. (2004) 'iPass: an Incentive Compatible Auction Scheme to Enable Packet Forwarding Service in MANET', *Proc. ICDCS*, pp. 534–542.

Crowcroft J. et al. (2003) 'Modelling incentives for collaboration in mobile ad hoc networks', *Proc. WiOpt.*

Elrufaie, E. and Turner, D. (2004) 'Bidding in P2P Content Distribution Networks using the Lightweight Currency Paradigm', *Proc. ITCC*, pp. 129–130.

Ferguson, D. F., Nikolaou, C. and Yemini, Y. (1993) 'An economy for managing replicated data in autonomous decentralized systems', *Proc. International Symposium in Autonomous Decentralized Systems*, pp. 367–375.

Garyfalos, A. and Almeroth, K. C. (2004) 'Coupon Based Incentive Systems and the Implications of Equilibrium Theory', *Proc. IEEE International Conference on E-Commerce Technology*, pp. 213–220.

Guy, R. et al. (1998) 'Rumor: Mobile Data Access Through Optimistic Peer-to-Peer Replication', *Proc. ER Workshops*, pp. 254–265.

Hara, T. and Madria, S. K. (2006) 'Data replication for improving data accessibility in ad hoc networks', *IEEE Transactions on Mobile Computing*, September, Vol. 5, No. 11, pp. 1515–1532.

Kamvar S., Schlosser M. T. and Garcia-Molina H. (2003) 'Incentives for combatting free-riding on P2P Networks', *Proc. Euro-Par* , pp. 1273–1279.

Kremer, S., Markowitch, O. and Zhou, J. (2002) 'An intensive survey of non-repudiation protocols', *Technical Report 473, ULB.*

Kurose, J. F. and Simha, R. (1989) 'A Microeconomic Approach to Optimal Resource Allocation in Distributed Computer Systems', *IEEE Trans. Computers*, Vol. 38, No. 5, pp. 705–717.

Liu, J. and Issarny, V. (2004) 'Service Allocation in Selfish Mobile Ad Hoc Networks using Vickrey Auction', *Proc. Current Trends in Database Technology - EDBT Workshops revised papers, LNCS 3268.*

Mondal, A., Madria S. K. and Kitsuregawa M. (2006a) 'CADRE: A Collaborative replica allocation and deallocation approach for Mobile-P2P networks', *Proc. IDEAS*, pp. 21–28.

Mondal, A., Madria S. K. and Kitsuregawa M. (2006b) 'EcoRep: An Economic Model for Efficient Dynamic Replication in Mobile-P2P networks', *Proc. COMAD.*

Mondal, A., Madria S. K. and Kitsuregawa M. (2006c) 'CLEAR: An Efficient Context and Location-Based Dynamic Replication Scheme for Mobile-P2P Networks', *Proc. DEXA*, pp. 399–408.

Ratner, D., Reiher, P.L., Popek, G.J. and Kuenning, G.H. (2001) 'Replication Requirements in Mobile Environments', *Mobile Networks and Applications*, Vol. 6, No. 6, pp. 525–533.

Ratsimor, O. et al. (2003) 'eNcentive: A Framework for Intelligent Marketing in Mobile Peer-to-Peer Environments', *Proc. ICEC*, pp. 87–94.

Richard, B., Nioclais, D. and Chalon, D. (2003) 'Clique: A Transparent, Peer-to-Peer Replicated File System', *Proc. MDM*, pp. 351–355.

Srinivasan, V. et al. (2003) 'Cooperation in wireless ad hoc networks', *Proc. INFOCOM.*

Straub, T. and Heinemann, A. (2004) 'An Anonymous Bonus Point System For Mobile Commerce Based On Word-Of-Mouth Recommendation', *Proc. ACM SAC*, pp. 766–773.

Turner, D.A. and Ross, K.W. (2004) 'A Lightweight Currency Paradigm for the P2P Resource Market', *Proc. Electronic Commerce Research.*

Vickrey, W. (1961) 'Counter speculation, auctions, and competitive sealed tenders', *Journal of Finance*, Vol. 41, pp. 8–37.

Wolfson, O., Xu, B. and Sistla, A. P. (2004) 'An Economic Model for Resource Exchange in Mobile Peer-to-Peer Networks', *Proc. SSDBM*, pp. 235–244.

Xu, B., Wolfson, O. and Rishe, N. (2006) 'Benefit and Pricing of Spatio-temporal Information in Mobile Peer-to-Peer Networks', *Proc. HICSS-39.*

Xue, Y., Li, B. and Nahrstedt, K. (2005) 'Optimal Resource Allocation in Wireless Ad Hoc Networks: A Price-based Approach', *IEEE Transactions on Mobile Computing.*

Zhong, S., Chen, J. and Yang, Y. R. (2003) 'Sprite: A Simple, Cheat-Proof, Credit-Based System for Mobile Ad-Hoc Networks', *Proc. IEEE INFOCOM.*