Enabling high throughput for delay sensitive streams using epidemic replication in clustered MP2P devices

Constandinos X. Mavromoustakis
Department of Computer Science,
University of Nicosia
46 Makedonitissas Avenue, P.O.Box 24005
1700 Nicosia, Cyprus
+357-22-841730
mavromoustakis.c@unic.ac.cy

Helen D. Karatza
Department of Informatics,
Aristotle University of Thessaloniki
54124 Thessaloniki, Greece
(+30) 2310 997974
karatza@csd.auth.gr

ABSTRACT
The forwarding replicated file-chunk strategy and dissemination scheme for enabling availability of the requested resources by other mobile users is considered as a crucial parameter. In this paper, epidemically replicated file chunks using a transition-based approach of a chained model of an infectious disease with susceptible, infected, recovered and death states, enforces the mobile nodes whether to host a file chunk or not, or when no longer a chunk is needed to purge it. In the examination of the proposed framework the effective throughput E_{ff} as a function of the packet loss parameter and the available Bandwidth is considered to examine the system's behavior under continuous changes. Additionally this work examines the throughput of the system and the enforcement degree of the proposed epidemic-based file chunk replication policy.

Categories and Subject Descriptors
C.2.1 Network Architecture and Design.

General Terms
Verification, Reliability, Measurement, Algorithms, Performance.

Keywords
High throughput replication model, Epidemic Replicas, File Reliability, Mobile Peer-to-Peer Devices.

1. INTRODUCTION

Dealing with the asymmetry characteristics in wireless devices and the scarceness in wireless resource sharing availability MP2P networks should form the topology dynamically and in way that the network formation will survive to host the end-to-end resource availability between hosts. Multimedia streaming in such environments is facing reliability problems where usually resources are wasted in the end-to-end connectivity guarantee and network formation. Additionally delay is not optimal due to the instability of the availability of the node involved in the path for the provision of the "*on the run*" file and requested resources availability. As a result of all the above there is a severe Throughput instability showing often the inadequacy in hosting delay sensitive applications. In a Peer-to-Peer network, some factors aggregate to performance reduction. These factors are based on connection quality or node and resource availability or maintenance between peers on the network. Peers are prone to failures and failures against performance are then aggravated by short connections times or sudden disconnections (with chained unpredictable disconnections due to range and battery failures) and small network formation factor. As a result redundant information is disseminated to all neighboring nodes in order to contribute to network reformation. But this substantially aggravates the overall performance since it generates redundant information that is destined only for a small group of mobile users. Moreover it utilizes the network resources (capacity and channels) causing additional delay which aggravates the overall throughput of the system [1]. In this work a new model for disseminating information is presented where the Epidemic model is used and applied into a MP2P system. Along with the designed model and framework a throughput model approach is proposed and encapsulated in the epidemic model for primarily enabling the stability in the offered resource reliability and increase the throughput of the system by increasing the throughput per source-destination pair scale with the number of nodes n. The proposed model enables the involved nodes in the path to contribute into the diffusion process pathetically and increases the multicasting diffusion throughput response significantly.

2. RELATED WORK
File sharing is probably the most widespread P2P application. It is estimated that as much as 70% of the network traffic in the Internet can be attributed to the exchange of files, in particular music files [12] (more than one billion downloads of music files can be listed each week [13]). When a mobile node makes an explicit request for a resource, and the whole network is flooded with a single query, as is the case with many mobile ad-hoc route discovery algorithms [3, 4], similar to file discovery by query flooding in P2P networks like Gnutella and unlike the proposed scheme in [5] which enables efficient and consistent access to data without redundant message generated communication overhead. Flooding in a wireless network is in fact efficient as compared to in wired networks because of wireless multicast advantage [6]. Improvements of the basic flooding approach using advertisements and geographic information have also been recently studied [7, 8]. Event dissemination protocols use gossip to carry out multicasts.

Digital Object Identifier: 10.4108/ICST.MOBIMEDIA2009.7449
http://dx.doi.org/10.4108/ICST.MOBIMEDIA2009.7449

Different mechanisms to broadcast information in the complete network where proposed in the recent past exploring the advantages in disseminating data in a specific geographic area (Geocast) [9]) as well as the caching approaches used, for enabling the requested data content to be available and discoverable [10, 11] at any time such that content can be discovered in a peer-to-peer manner without having network partitioning problems. Additionally id requested data is available the timing end-to-end availability should use the selective or unselective dissemination process to forward the requested packets to destination [1, 10, 11]. However selective and criteria-based dissemination procedure based on mobile nodes' content requirement and the reactive multicast group establishment is still a relatively unexplored area. Proactive multicasting can be disastrous and prohibitive, particularly given the fact that membership varies over time due in dynamically changing topology (frequent and unexpected changes in topology/managing group membership is difficult), content requirement can be diverse (need for multiple multicast groups), and the asymmetry in nodes' resources which significantly affects the stability (hosting many capacity and energy constraints). In [14] authors identify several interesting effects at the link-layer, notably the highly irregular packet reception contours, the likeliness of asymmetric links, and the complex propagation dynamics of simple protocols. An epidemic algorithm in [15] is proposed based on strictly local interactions for managing replicated databases in a robust way for unpredictable communication failures. An extension of this method is studied in [16] where the "repopulation" process is applied for facing node failures. In this work the decentralization and autonomy of control of all the nodes in a cluster which has at least one active transmission (delay sensitive data/multimedia packets requested from a node to another) is proposed and studied using the using an Epidemic-based replication approach in clustered MP2P devices. In the proposed model there is no central coordinating authority for the organization of the network (setup aspect) or the use of resources and communication between the peers in the network (sequence aspect). Examination through simulation is performed for the offered reliability by the epidemic collaborative replication scheme and combined throughput optimization scheme, showing the significant increase in grade of robustness in file sharing reliability among mobile peers and a remarkable throughput response.

3. EPIDEMIC COLLABORATIVE REPLICATION FOR MAINTAINING FILE SHARING CONSISTENCY

3.1 Cooperation scheme between clustered Mobile Peer-to-Peer devices

The exploitation of the time-variation of the user's channel due to mobility in [2] creates a new concept by splitting the packets of each source node to as many nodes as possible. Therefore, strategies of this type incur additional delay, because packets have to be buffered until the channel becomes sufficiently strong for transmission(s). In this work in order to avoid any redundant transmissions and retransmissions we propose the clustered-based mobility configuration scenario which is set in figure 1. Clusters enable the connectivity between nodes and the local (within a cluster) control of a specified area. On the contrary with [1, 18] in this work a different mobility scenario is examined where the node controlled area is not specified (unless a cluster can not be formed) -like the Landscape in [1] and the mobility scenario is entirely different having more flexible registering *in* and *out* of a cluster. Cluster network formation works as follows: each cluster is responsible to host newly added nodes and measures (Cluster Head (CH) responsibility) whether these nodes can host new file chunks. If the new node entered the cluster *i* has available remaining capacity greater than the existing CH, then the Gradual Energy Tree-based (GET) is formed [19] for maintaining connectivity through the node's remaining energy. GET is used to form a tree having as a root, the node with the least remaining energy, and as children the nodes with the higher residual energy. GET configuration is applied only locally, on a 2-hops basis to each node to prevent huge information delivery to nodes which are located far from source node and to maintain 2-hops recursive network topology formation as in the proposed scenario depicted in [19]. No CH has a restriction to directly communicate with any other CH. Additionally the selected CH has as a responsibility to drive the transfers (between nodes) and restrict transfers which may be inadequate in terms of resources (coverage, connectivity, lack of relay nodes etc). Devices which are moving into different locations where the nodes' density is not adequate to serve the hosted users are not hosted neither as CH nor as a relay node in a specified cluster. Thus in our model we have enabled the *probation slot* parameter which basically evaluates the time the node which has entered the cluster and after T_s *probation slot* the node can be either CH candidate or a member of a cluster to share resources.

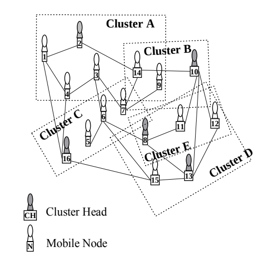

Figure 1. The initial network formation in a Mobile Peer-to-Peer system according to clustering formation and node self-awareness.

Users' request patterns may vary because some users may require transmitting delay sensitive data which in turn require low latency while others may require data being transmitted with low communication overhead. Devices are purely independent in terms of mobility patterns as well as in terms of capacity having spatiotemporal asymmetry. Connectivity can change network state also when a user moves to a different location and data need to be delivered from a source user to another then the relay mechanism can be interrupted and user experiences data losses. Assume that node 1 has a download request from node 15 in different cluster (fig.1). The CH that node 1 is set, gets informed about this transfer and supervises the correct delivery and correct connectivity between peer in the cluster A. The same exists for node 15 as well for all nodes which are not set in one of these clusters (via a third cluster/i.e. cluster C). In this way the connectivity can be adequately formed and nodes can transfer data and file chunks by temporarily cache them onto intermediate nodes as explored in the following section. The proposed scheme could also enable scalable broadcast among mobile nodes while reduces the acknowledged delivery of information in topologically changing environments.

Digital Object Identifier: 10.4108/ICST.MOBIMEDIA2009.7449
http://dx.doi.org/10.4108/ICST.MOBIMEDIA2009.7449

3.2 Epidemic object replication scheme for reliable file sharing

Epidemic-like algorithms [17] which follow a nature paradigm could improve the cooperative response of these dynamically changing topologies. By applying simple epidemiological rules to efficiently spread information by only having a local view of the environment, the outcome could in essence a heuristic optimization approach to the reliability problem [10]. Reliable file sharing can be determined by relying on epidemic algorithms, a breed of distributed algorithms that find inspiration in the theory of epidemics. Epidemic (or gossip) algorithms constitute a scalable, lightweight, and robust way of reliable disseminating information to a recipient or group of recipients, by providing guarantees in probabilistic terms. Based on certain characteristics, epidemic algorithms are amenable to the highly dynamic scenarios. In this work a promiscuous caching [1] is used which means that data can be cached "anywhere, anytime". However this enables trade offs in consistency for availability which is faced with cooperative cluster-based connectivity and the caching used in the clusters described earlier. We assume that a common lookup service is followed by all devices in the network cooperating via a shared platform. In a dynamic MP2P network each user might desire to share or download a file or files with other users (peers). Many conditions must be satisfied for reliable communication between mobile peers. On one hand users due to their mobility might draw away from the user (peer) when a file sharing communication takes place. On the other hand a sudden network partitioning or network split could occur because of network's dynamic topology which is continuously changing. Thus a proactive dissemination scheme must be determined in order to prevent the cutoff in file sharing communication. On the contrary with [18] this work assumes an isolated system comprising of a variable number of mobile nodes ([18] assumes a fixed number in a Landscape and pre-known mobility pattern like in VANets) confined in a predefined geographic region. These nodes are mobile, and communicate with each other in a wireless (radio) ad-hoc manner. The limited connectivity coverage that MP2P systems offer, results in significant delay in downloading a message or file (group of packets). In the proposed scenario each node carries some unique data items and the number of nodes which have been requested to deliver file chunks also varies.

A chain model was chosen to determine the file sharing termination criteria as follows: each mobile host m_k has a predetermined capacity *M*. At any time in the network each m_k has a state. There are three different states that m_k can be characterized: the susceptible state *S(t)* represents the number of hosts in the system which are "susceptible," infected state *I(t)* represents the number of "infected" hosts, and *R(t)* represents the "recovered" hosts. A host is in susceptible state *S(t)* if the device does not share any information with any other host. In turn host/node 1 is in infected state *I(t)* if a file(s) share occurs. Finally node 1 is in "recovered" state *R(t)* if any shared file(s) are no longer pending. A chain model of an infectious disease with susceptible, infected, recovered and death states is used shown in figure 2.

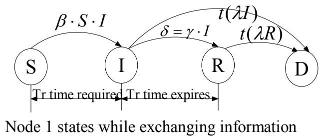

Figure 2. The chained model of an infectious disease with *susceptible*, *infected*, *recovered* and *death* states.

Adopting the framework from an infectious disease model [17], a host is set as "infected" if a file sharing (or a group of data packets) are pending. Suppose there are *k* hosts in the system, then a host is sharing a resource with *β(k−1)* other hosts per unit time. *S (k−1)* do not have yet the disease. Therefore, the transition rate from state S to state I becomes

$$Filesharing = (nu\_infected) \cdot (dld\_Rate) \cdot (nu\_NOT\_share)$$

$$Filesharing = I[\beta(k-1)] \cdot \left[\frac{S}{k-1}\right] \quad (1.1)$$

where β is the contact rate for *k* hosts.

Then the downloaded (no longer pending) rate is :

$$\delta = \gamma \cdot I \quad (1.2)$$

where γ is the download rate and I is the number of infected devices. The file sharing process should be performed in time limit [0, t) and the infection (download) has upper limit time duration otherwise the infection stops and the node is transferred into the recovered state as follows:

$I_{TTL} = \tau \cdot \beta \cdot S \cdot I$ where τ is the time limit that a delay sensitive transfer/download can be processed without losses.

In the proposed scheme –on the contrary with [1]- we have introduced the death state which comprises of marking the spread of the deletion of a data item to any nodes which have hosted for time t the data item. Deleted items get the death certificates after a specified timestamp and nodes diffuse this deletion like the updates in any other state (refer to figure 2 chain model where an item can not remain after a certain time on a specified node). For this state two thresholds *t1*, *t2* – most sites delete it after *t1* and retention sites delete after *t2*

Assume that a device P has such a certificate for data item x then if by any chance an obsolete update for x reaches P; P will react by simply spreading the death certificate for x again to all nearby nodes in the cluster (informing also the CH). This is to clean up with redundant files the devices' storage units and memories.

$$d_{TTL} = t(\lambda I) + t(\lambda R) = t(\lambda I + \lambda R) = t\lambda(I + R) \quad (2.1)$$

Buffer Purging Enforcement Degree (BPED) as a function of the death rate considered in figure 2 above, as follows:

$$\mathrm{BPED} = \sum_{t=0}^{t_\tau} f(\lambda t) \cdot R \quad (2.2)$$

The $\beta \cdot S \cdot I$ is called π coefficient which indicates the enforcement degree of the diffusion process and $\pi \cdot \tau$ the enforcement at a bounded time delay. π has the dimension of $\left[\frac{1}{Time}\right]$. Previous examinations of the behavior of small scale systems [19] showed that relatively small populations could be faced with a stochastic model. Thus taking into account that π depends on the measures of *S(t)* and *I(t)* and the probability of transmitting the information, we can derive *S(t)* as follows:

Digital Object Identifier: 10.4108/ICST.MOBIMEDIA2009.7449
http://dx.doi.org/10.4108/ICST.MOBIMEDIA2009.7449

$$\frac{dS}{dt} = -\beta \cdot S \cdot I = -\pi \tag{3.1}$$
$$\frac{dI}{dt} = \beta \cdot S \cdot I = \beta(N-I) \cdot I = \beta NI - \beta I^2$$

By solving the first order differential equation the outcome is:

$$I(t) = \frac{N}{1 + e^{-\beta \cdot N \cdot t}(N-1)} \tag{3.2}$$

According to the definition of spreading ratio equation (3.2) becomes:

$$I'(t) = \frac{I(t)}{N} = \frac{N}{N \cdot (1 + e^{-\beta \cdot N \cdot t}(N-1))} \tag{3.3}$$
$$I'(t) = \frac{1}{1 + e^{-\beta \cdot N \cdot t}(N-1)}$$

Equation 3.3 is referred as the Cumulative Distribution Function (CDF). In the implementations of the scheme the locations will be updated and measured according to the following:

$$L(t) = L(t-1) + S_t \cdot \vec{d} \tag{4}$$

Where *L(t)* is the new location *L(t-1)* the previous location at step time *(t-1)*, S_t is the speed of each device and $\vec{d}$ is the directed unit vector [17]. Additionally we estimate the average delay experienced by all the peers in downloading a multi-part file as follows:

$$\bar{d}^{(m)} \approx \frac{\tau_0}{m} \log_2 n \tag{5}$$

where *m* is the number of identical sized chunks that the file is divided, *n* is the number of peers and τ_0 is the amount of time taken to download the whole file, if downloaded from a single peer.

3.3 Enabling high throughput using Epidemic replication

In order to evaluate the modeled approach we have measured the effective throughput E_{ff} as a function of the packet loss parameter and lost packets as follows:

$$EffectiveThroughput = E_{ff} = 1 - (Packet_{loss}) \cdot (\frac{PacketTrasferredSize}{PacketTrasferredTime}) \cdot (\frac{1}{Bandwidth})$$

(6)

4. SIMULATION EXPERIMENTS AND DISCUSSION

To demonstrate the methodology discussed in this paper, we performed exhaustive discrete time simulations of the proposed scenario under several different conditions. In the implementation-simulation of this work we used our own libraries implemented in C/Objective C programming language. We assume a system consisting of several mobile nodes, e.g., mobile users equipped with notebooks or PDAs and wireless network interfaces. All mobile nodes collaborate via a shared application that uses a distributed lookup service. Radio coverage is small compared to the area covered by all nodes, so that most nodes cannot contact each other directly. Additionally, we assume IEEE 802.11x as the underlying radio technology.

Figure 3 shows the mean number of file chunks transfer sessions between clusters, with the number of the successfully received packets of completed downloads. This figure shows that even in the present of high mean number of file chunks transfer sessions between clusters the packets that are successfully received are kept at high levels. In Figure 4 the end to end packet delay (ms) between different clusters with the mean number of file chunks transfer sessions is presented where it is obvious that the epidemic collaboration enables significantly less end to end packet delay, even if the number of file chunks transfer sessions increases. The death rate with the mean number of file chunks recovery requests is presented in figure 5. If the death rate will not be present this can be in a sense considered as a delay factor [1] (searching delay onto nodes) and as unneeded/redundant information hosted onto nodes. Figure 6 shows the total transfer delay with the mean number of file chunks simultaneous transmissions. Figure’s 6 outcome is considered very interesting extracting that even in the presence of multiple simultaneous transmissions, the total transfer delay is kept relatively at low levels.

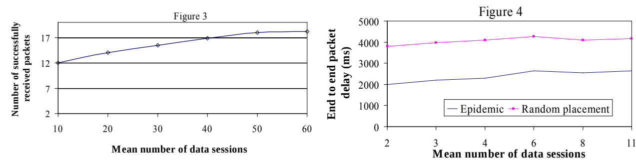

Figure 3. Mean number of file chunks transfer sessions between clusters, with the number of the successfully received packets(completed downloads), Figure 4. End to end packet delay (ms) between different clusters, with the mean number of file chunks transfer sessions.

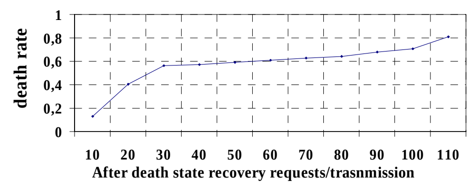

Figure 5. Death rate with the mean number of file chunks recovery requests.

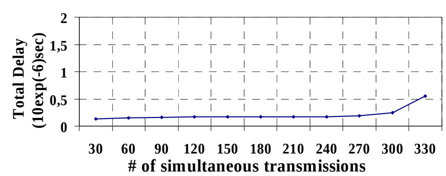

Figure 6. Total transfer Delay with the mean number of file chunks simultaneous transmissions.

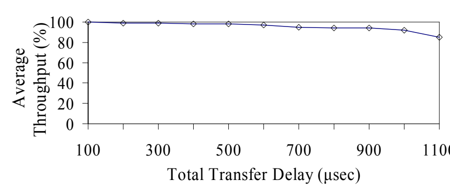

Figure 7. Throughput evaluation with the Mean Total Transfer Delay (mean of all transfers for the total delay).

Figures 7 and 8 show the throughput evaluations under the Mean Total Transfer Delay (mean of all transfers for the total delay) and the Mean number of nodes that are participating in the cluster in the transferring (or propagating file chunks respectively. Both figures show the robustness of the throughput response under

Digital Object Identifier: 10.4108/ICST.MOBIMEDIA2009.7449
http://dx.doi.org/10.4108/ICST.MOBIMEDIA2009.7449

these measures. When the Total Transfer Delay is increasing, the scheme is shown to be robust in the overall throughput offered throughout the simulation.

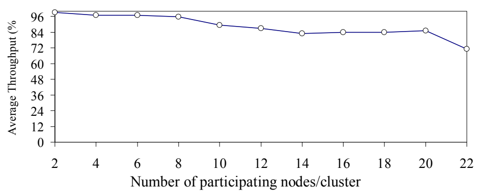

Figure 8. Throughput evaluation with the Mean number of nodes that are participating in the cluster in the transferring (or propagating file chunks.

5. CONCLUSIONS

In this paper we have proposed an approach for reliable file chunks sharing among mobile peers using a stateful metaphor from epidemics. The modified epidemic approach uses a delay protocol which creates replication of the file chunks in order to enable P2P reliable file sharing. In order to clean up the nodes/users hosting no longer needed and redundant files, we have introduced the Buffer Purging Enforcement Degree (BPED) as a function of the death rate. The proposed scheme seems to behave satisfactory, allowing high SDR for completed files. A comparison with other schemes for file chunks or generic object's placement is one of our upcoming research targets. By comparing the placement strategy will enable future researches and implementations to use this method in large lookup databases, in which a common multiclient application is used.

6. REFERENCES

[1] C. X. Mavromoustakis, H. D. Karatza, "Under storage constraints of epidemic backup node selection using HyMIS architecture for data replication in mobile peer to peer networks" Journal of Systems and Software, Elsevier Volume 81, Issue 1, January 2008, Pages 100-112.

[2] M. Grossglauser, D. Tse, "Mobility Increases the Capacity of Ad Hoc Wireless Networks", Proceedings of IEEE Infocom 2001, pp 312-319.

[3] X. Hong, K. Xu, and M. Gerla. Scalable routing protocols for mobile ad hoc networks. IEEE Network Magazine, 16(4), Jul-Aug 2002, Pages 11–21.

[4] E. Royer and C. Toh. A review of current routing protocols for ad-hoc mobile wireless networks. IEEE Personal Communications, Apr. 1999 vol. 6 (2), Pages 46–55. doi:10.1109/98.760423.

[5] M. K. Aguilera, A. Merchant, M. Shah, A. C. Veitch, and C. T. Karamanolis, "Sinfonia: a new paradigm for building scalable distributed systems". In 21st ACM Symposium on Operating Systems Principles (SOSP), 2007SOSP'07 ACM Symposium on Operating Systems Principles No21, Stevenson, Washington , ETATS-UNIS (14/10/2007), Pages 159-174.

[6] J.E. Wieselthier, G.D. Nguyen, and A. Ephremides. Energy-efficient broadcast and multicast trees in wireless networks, Mobile Networks and Applications archive, Volume 7, Issue 6 (December 2002), Pages 481-492.

[7] J. Tchakarov and N. Vaidya, Efficient content location in wireless ad hoc networks. In IEEE International Conference on Mobile Data Management (MDM), 2004, Pages 74- 85.

[8] C. X. Mavromoustakis, H. D. Karatza, "Dispersed information diffusion with level and schema-based coordination in mobile peer to peer networks", Cluster Computing (Computer Communications & Networks), Springer Netherlands, Volume 10, Issue 1, (March 2007), Pages 33-45.

[9] Y. B. Ko and N. H. Vaidya. Flooding-based geocasting protocols for mobile ad hoc networks. Mobile Networks and Applications, 7(6), 2002, Pages 471–480.

[10] T. Hara. Effective replica allocation in ad hoc networks for improving data accessibility. In Proceedings of IEEE INFOCOM, pages 1568–1576. IEEE Computer Society, 2001.

[11] N. Davies, K. Cheverst, K. Mitchell, and A. Friday. Caches in the air: Disseminating tourist information in the guide system. In In Proceedings of Second IEEE Workshop on Mobile Computing Systems and Applications, 25-26 Feb 1999, Pages 11-19.

[12] M. Stump, Peer-to-peer tracking can save cash: Ellacoya. Retrieved November 15, 2008, from http://www.ellacoya.com/news/pdf/10_07_02_mcn.pdf.

[13] F.Oberholzer, K. Strumpf, The Effect of File Sharing on Record Sales - An Empirical Analysis. Retrieved November 15, 2004, from: http://www.unc.edu/~cigar/papers/FileSharing_March2004.pdf.

[14] D. Ganesan, B. Krishnamachari, A. Woo, D. Culler, D. Estrin, and S. Wicker, "Complex behavior at scale: An experimental study of low-power wireless sensor networks". UCLA, 2002, Technical Report UCLA/CSD-TR 02-0013.

[15] A. Demers, D. Greene, C. Hauser, W. Irish, and J. Larson, "Epidemic algorithms for replicated database maintenance". Proceedings of the 6th Annual ACM Symposium on Principles of Distributed Computing, ACM Press, 1987, pp 1-12.

[16] J. Kulik, W. R. Heinzelman, and H. Balakrishnan, "Negotiation-based protocols for disseminating information in wireless sensor networks". Wireless Networks, 2002, vol 8(2-3), pp 169-185.

[17] F. Brauer and C. Ch/avez, "Mathematical Models in Population Biology and Epidemiology" Springer-Verlag New York, Inc., 2001.

[18] C. Mavromoustakis, H. Karatza, "'Reliable File Sharing Scheme for Mobile Peer-to-Peer Users Using Epidemic Selective Caching'". Proceedings of IEEE International Conference on Pervasive Services (ICPS), Santorini, Greece, July 2005, pp.169-177.

[19] C. Mavromoustakis, H. Karatza, "Dispersed information diffusion with level and schema-based coordination in mobile peer to peer networks", Cluster Computing (Computer Communications & Networks), Springer Netherlands, Volume 10 , Issue 1, (March 2007), pp 33-45.

Digital Object Identifier: 10.4108/ICST.MOBIMEDIA2009.7449
http://dx.doi.org/10.4108/ICST.MOBIMEDIA2009.7449