A Tunable K-Hop Mobility-Based Outsourcing Replication Model for Optimizing End-to-End Reliability in MP2P Systems Using Community-Oriented Neighboring Feedback

Constandinos X. Mavromoustakis

Department of Computer Science, University of Nicosia
46 Makedonitissas Avenue, P.O. Box 24005, 1700 Nicosia, Cyprus
mavromoustakis.c@unic.ac.cy

Abstract. This work proposes a scheme which takes into account the k-hop mobility-based outsourcing strategy where the mobile devices are using a sequential scheme for caching requested high ranked resources onto neighboring nodes in the k-hop path in order to be available during a request. The scheme takes into consideration the relay epoch and the mobility aspects of each node in the k-hop path in order to disseminate effectively and within a specified duration any requested file chunks. The combined community oriented model enables the involved nodes in the path to contribute into the diffusion process pathetically according to the k-hop replication diffusion scenario using a feedback oriented mechanism which increases the multicasting diffusion throughput response significantly as simulation results show.

Keywords: mobility and adaptation methodology, MP2P systems, caching and outsourcing data, reliability technique, community-oriented neighboring feedback.

1 Introduction

The present work is motivated by the design of protocols in Mobile Peer-to-Peer (MP2P) networks that seek for nodes in order to assign to them specific roles such as forwarding capability as mediator nodes in cooperative caching, as well as for message ferrying nodes in Delay Tolerant Networks [1], rebroadcasting nodes in vehicular networks [2] et.c. Peers are prone to failures and aggravate the end-to-end performance whereas short connections times or sudden disconnections (with chained unpredictable disconnections due to range and battery failures) reduce the overall resource availability of the MP2P system. The present work proposes a new technique called as k-hop mobility-based cache replication strategy where the mobile devices are using a sequential scheme for caching requested high ranked resources onto other neighboring nodes in the k-hop path in order to be available during a request. The scheme takes into consideration the relay epoch and the mobility aspects of each node in the k-hop path in order to disseminate effectively and within a specified duration

F. Alvarez and C. Costa (Eds.): UCMEDIA 2010, LNICST 60, pp. 1–9, 2012.

any requested packets. The proposed scheme utilizes the systems resources and comprises of a new model for disseminating information in a MP2P system. The proposed model enables greater stability in the offered resource availability on-the-go, and end-to-end reliability, whereas it increases 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 according to the k-hop dissemination scenario using structured topologies and increases the multicasting diffusion throughput response significantly as simulation results show.

2 Related Work

Determination of the appropriate values of fundamental device parameters (e.g., the optimal caching parameter in accordance with other tuned measures) is a difficult task. 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] the scheme for information dissemination should be efficient enough, to enable recipients of the information to receive the requested info, and at the same to time reduce in the least possible degree the redundant duplications. 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, the proposed scheme considers the flooding occasion which was found [6, 7 ,8] that reduces dramatically the end-to-end performance of the network.

Different 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. Additionally if all nodes are moving using a pattern or a path (as in vehicular networks) the requested data should be available within a specified interval and the selective or unselective dissemination process should forward the requested packets to destination [1, 2 10, 11] in a bounded time delay since devices are moving and the topology changes in time. Thus cooperative caching can speedup the streaming process, since the processing and delivery of multimedia content are not independent. Thus, cooperative caching results in lower latency, energy dissipation and packet loss [12].

The asymmetry in nodes' resources which significantly affects the stability of the sharing process (hosting many capacity and energy constraints) should be balanced using a scheme fulfilling all these requirements. A previous effort on a part of this problem is studied in [13] where the "repopulation" process is applied for facing node failures. The proposed model uses decentralized control in the manipulation of the requested resources and the communication between the peers in the network (sequence aspect). This applies in particular to the fact that no node has central control over the other. In this respect, communication between peers takes place directly. The outsourcing concept for the k-hop scheme is proposed in the next section which attempts to fill the trade-offs between user's mobility, reliable file sharing and limited throughput in exchanging delay sensitive data streams in mobile peer to peer environments. Examination through simulation is performed for the offered reliability by the collaborative replication k-hop scheme showing the increase in the grade of robustness in sharing resources among mobile peers.

3 K-Hop Cooperation Scheme and Mobility Model in Clustered Mobile Peer-to-Peer Devices

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 consisting of the single lane approach and multiple lanes approach. Clusters enable the connectivity between nodes and the local (within a cluster) control of a specified area. On the contrary with [8] in this work a different mobility scenario is examined where the node controlled area is not specified (unless a cluster cannot be formed) - like the Landscape in [8, 10]. The mobility scenario is based on a real time following a probabilistic path like in the real vehicular pathways.

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 this node becomes a CH and maintains the connectivity. 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). 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. 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.

In order to enable recoverability a replication scheme is designed using the k-hop replication algorithm as follows: each node creates, using the common look-up table for the requests, a ranking-based criteria selection of the requests of any files and/or packets by the nearby nodes. Since each node-based on the mobility scenario examined in the next section- will remain for a specified amount of time in the same path then the node *i*, which gets the requests forward copies of the packets to k-nodes in the path in the two directions (as in figure 1). Then each node which receives the requests do the same ranking procedure for their requests and forward the requested file chunks to k-nodes in the path in both traversing directions. The ranking process for the *i-resource* of the *N-node* ($Rank(i_N)$) that is set in the cluster $C_{i.}$ takes place only in the formed cluster C. In this way there is a spine-based tree of replicated objects consisting of several nodes in the clustered path, while nodes with high ranked requests for specified file chunks continually create replicas for other nodes. This model enables high resource availability while maintaining the connectivity among peers that are requesting common resources.

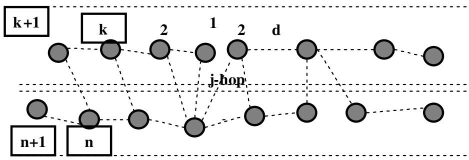

Fig. 1. Inter-cluster outsourcing of file chunks in order to be available according to ranking requests. j-hop comparisons with j-1 neighboring nodes are taking place for avoiding saturation of the replication scenario.

The tree of replicated objects is created-as in [10], considering the mobility model and the likelihood in accessing the certain path using social-model interactions, the update rate of the requests, the cost of the n-hop replications as well as the replay costs between m-different relay regions. Considering that the above scenario is used in real-time like in a vehicular raw-lane network, where requests of the j-hop may have a different direction, then it stands that for $Rank(i_N)$:

$$Min[Rank(i_N)]\forall N \notin C_i \quad (1.0)$$

is minimized for the node that the resource was downloaded at least once or when the distance d is over a certain threshold D_{thress} from the k-hop peer- which means that the requested resource(s) set on this node has been redirected to any other path. Equation 1.0 sets the rank of the node containing the requested resource to minimum for the nodes that are not member of the cluster where the resource was requested iff d is over a certain threshold D_{thress}. If the D_{thress} becomes big enough then the resource is isolated and it no longer belongs to the C_i. This enables the prevention of huge duplicated information delivery, whereas it considers the nodes which are located far from source node and to maintain only the j-hops duplications-after the performed comparisons- avoiding redundant transmissions.

Considering the k-hop scenario of figure 1, the evaluated duration of the requested file chunks is evaluated as follows:

$$C_d = k \cdot E_i \quad (1.1)$$

where C is the caching duration that is allowed for node i and E_i is the relay epoch according to the number of hops permitted. Therefore it stands that the greater the number of hops, then the greater the time duration that is allowed to be achieved. The delay epoch duration is modelled according to equation 1.5, taking into account the hop-count path, ping delays and the total delays from the end-to-end perspective.

3.1 Using Community Oriented Relay Regions

When streaming in a region and the packets sent are considered as prioritized, then these packets have a bounded time delay τ to reach any specified destination. The streaming parameter is based on the number of hops from a source to a destination and on the relay region and enclosure graph [9]. The relay region of a certain transmitter–relay node pair (u,w) identifies the set of points in the plane (node locations) for which communication through the relay node is more reliable than direct communication. Formally a *Relay region is considered to be as in [9]:*

$$RR_{u\to w} = \{(x, y) \in \Re^2 : P_{u\to w\to(x,y)} > P_{u\to(x,y)}\}. \tag{1.2}$$

where $P_{u\to w\to(x,y)}$ is the probability for a certain node *u* to transfer the file chunks from a source node *x* to destination node *y* via the *w* based on the connectivity and the social interactions described in the next section. Thus in an end-to-end path $\forall u \in P_n$ the minimized ping delays between the nodes in the end-to-end path the minimized evaluated delay is according to the:

$$d_p = Min\sum_{i=1}^{n} D_i \tag{1.3}$$

where D_i is the delay from a node *i* to node *j*, and d_p is the end-to-end available path. Therefore the delay epoch $E_{i(t)}$ of each node is defined as a function of the number of created replicas on the j-hosts as follows:

$$E_{i(t)} = d_{r_{i\to j}} \cdot r_{i\to j} / Total\_d_{r_{i\to j}} \tag{1.4}$$

where *D* is the delay via the ping assigned durations, $r_{i\to j}$ is the number of replicas form node i to j in the j-hop path and $Total\_d_{r_{i\to j}}$ is the total duration that all the requested replicas can be downloaded from the j-hop path.

3.2 Considering Mobility Models for Enabling Minimum Latency Outsourcing for Collaborative Neighboring Caching

Assume that we consider a Brownian-like motion with semi-random fields of characterization. In Brownian-like motion, the position of a node at a given time step depends (in a certain probabilistic way) on the node position at the previous step. In particular, no explicit modeling of movement direction and velocity is used in this model. An example of Brownian-like motion is the model used in [15]. Mobility is modeled using three parameters: p_{stat}, p_{move} and m. The first parameter p_{stat}, represents the probability that a node remains stationary for the entire simulation time. Parameter p_{move} is the probability that a node moves at a given time step. Parameter m models, to some extent, velocity: if a node is moving at step i, its position at step i + 1 is chosen uniformly at random in the square or side 2m centered at the current node position. Since in our case study we have examined the scenario of the movements of the nodes where nodes are moving in real-time pathways (roads,streets, corridors et.c.), it is important to denote that the probabilistic Brownian two-dimensional motion [10] can be an emulation of the real-time movements of the users in a certain pathway.

Definition 1.3 (Likelihood of the multiple selection connectivity path): The multiple selection connectivity path of a certain transmitter–relay node pair *(u,w)* traversing the formed *n* paths/clusters for time t, can be maintained within this time t if the nodes follow the same trajectory movements with likelihood as follows:

$$C_{u\to w} = \{(x, y, u, w) \in \Re^2 : P_{C_1\in(u,x,y,w)} > P_{C_2\in(u,x,y,w)} \forall u, w \in C_n\}. \quad (2)$$

Thus in an end-to-end path $\forall u \in P_n$, the likelihood of following the path P_1 instead of P_2 for node u and w in the cluster C_1 is following the p_{move} of the node which is the probability that a node moves at a given time step into a certain direction using the unit vector m.

3.2.1 Intercommunity Streaming and Download Frequency of the File Chunks

The metrics modeled are community-oriented and are considering the number of created clusters $C_N(t)$ in a specified Relay region of a certain transmitter–and a number of receivers (1, N] under the relay node pair (u,w) -as a modified definition of [9]- as follows the:

$$C_N(t) = \frac{2|h_N(t)|}{|I_{C(N)}(t)|\cdot(|I_{C(N)}(t)|-1)}, \text{ iff } P_{u\to w\to(x,y)} > W_N(t) \quad (3.1)$$

where *W* is the Community streaming factor and is defined as the number of existing communities in the intercluster communicational links at a given time instant. The $h_N(t)$ is the number of hops in the existing clusters and the $I_{C(N)}(t)$ is the number of interconnected nodes N in the cluster $C_N(t)$.

A community is defined as a dense sub-graph where the number of intracommunity edges is larger than the number of intercommunity edges [9]. *W* can be defined according to the download frequency of the file chunks in the intercommunity as follows:

$$W_N(t) = \frac{DldRate\cdot\#\,sharingChunks}{Total\#\,dlds(t)\cdot\#\,inactiveChunks} \quad (3.2)$$

where in (1.4) the download rate is considered in contrast with the number of chunks being shared in a specified instant time *t*.

3.2.2 Neighboring Feedback for File Chunk Indices

Neighbor $N_j(t)$ at a certain instant of time, informs the *k*-neighbor receivers for the existence of the file chunk onto this node, according to the following:

$$R_{j \to 1..k} = \{ \lim_{N \to K} C_{n(t)} \in W_N(t) : h_N(t) > \frac{N(N-1)}{2} \} \quad (3.3)$$

This means that for a specified amount of time the neighbors collaboratively can provide any node which exists in the Community with a streaming factor W with the feedback and can be locally informed about any requested file chunk at the specified time t. Provided that all the assumptions were made under the k-hop replication scheme described in the previous section in order to enable reliable file chunks sharing among mobile peers using a neighboring cluster-based feedback outsourcing.

4 Performance Analysis through Simulation and Discussion

The implementation-simulation of the proposed scenario was performed in C/Objective programming language libraries were used as in [8] and as a routing approach we have used a combination of the Zone Routing Protocol (ZRP) [16] with the Cluster-based Routing Protocol (CRP). Figure 2 shows the number of hops with the delay/latency for the nodes that requested replicas were created. This figure shows that in the case of not having a feedback response, the number of hops that replicas were created is significantly increased compared with the number of replicas created

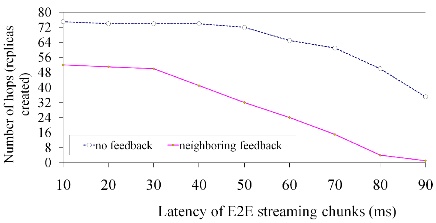

Fig. 2. Number of hops with the delay/latency for the nodes that requested replicas were created

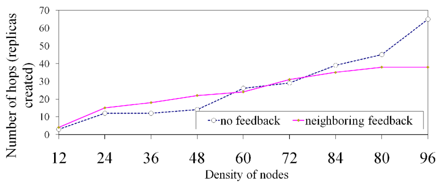

Fig. 3. Number of hops with the density of the nodes in the cooperation scenario

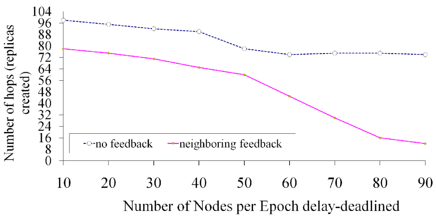

Fig. 4. Number of hops with the density of the nodes per epoch delay deadline used in the cooperation scenario

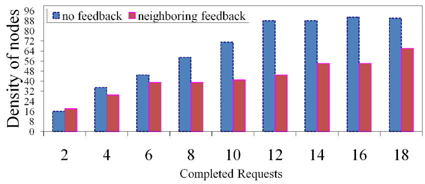

Fig. 5. Density of nodes with the completed requests

when feedback takes place. Additionally the latency in the first case is significantly higher and as shown reduces the throughput response of the system significantly. Figure 3 shows the number of hops with the density of the nodes in the cooperation scenario whereas figure 4 shows the number of hops with the density of the nodes per epoch delay deadline used in the cooperation scenario. In figure 5 the density of nodes with the completed requests is shown whereas in figure 6 depicts that neighboring feedback can significantly increase the streaming stability in a multistreaming end-to-end path.

Figures 7 shows the throughput evaluations under the Mean Total Transfer Delay (mean of all transfers for the total delay). The robustness of the throughput response under these measures is depicted when the Total Transfer Delay is increased, where the scheme is shown to be adequately behaving in the overall throughput offered, throughout the simulation.

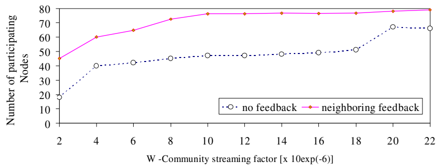

Fig. 6. Mean number of participating nodes in the intercluster feedback provision with the Community Streaming factor-*W*

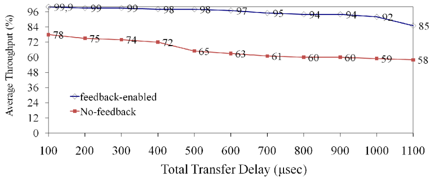

Fig. 7. Throughput evaluation with the Mean Total Transfer Delay (mean of all transfers for the total delay)

5 Conclusions and Further Research

This work considers the modeled social interactions in the intercommunity domain enabling the created clusters to interact via social-oriented feedback as well as the existing individual movements of each node. The streaming is achieved through the collaborative outsourcing file chunk policy, where, in cooperation with the k-hop replication, the scheme avoids problems like file chunk redundancy and the increased utilization of storage resources. Experimental results show that the proposed scheme behaves satisfactory, allowing high SDR for completed files.

Current and future research directions include the modeling of the mobility pattern of the peers by using approaches like the fractional Brownian motion as well as other mobility schemes taking into account the global requests and different network partitioning parameters.

References

[1] Hui, P., Crowcroft, J., Yoneki, E.: BUBBLE Rap: Socialbased forwarding in Delay Tolerant Networks. In: Proceedings of the ACM MOBIHOC Conf., pp. 241–250 (2007)

[2] Zhang, M., Wolff, R.S.: Routing protocols for vehicular ad hoc networks in rural areas. IEEE Communications Magazine 46(11), 126–131 (2008)

[3] Hong, X., Xu, K., Gerla, M.: Scalable routing protocols for mobile ad hoc networks. IEEE Network Magazine 16(4), 11–21 (2002)
[4] Picconi, F., Massoulie, L.: Is There a Future for Mesh-Based live Video Streaming? In: Peer-to-Peer Computing 2008, pp. 289–298 (2008)
[5] Aguilera, M.K., Merchant, A., Shah, M., Veitch, A.C., Karamanolis, C.T.: Sinfonia: a new paradigm for building scalable distributed systems. In: 21st ACM Symposium on Operating Systems Principles (SOSP), SOSP 2007 ACM Symposium on Operating Systems Principles No 21, Stevenson, Washington, pp. 159–174. ETATS-UNIS (2007)
[6] Ko, Y.B., Vaidya, N.H.: Flooding-based geocasting protocols for mobile ad hoc networks. Mobile Networks and Applications 7(6), 471–480 (2002)
[7] Hara, T.: Effective replica allocation in ad hoc networks for improving data accessibility. In: Proceedings of IEEE INFOCOM, pp. 1568–1576. IEEE Computer Society (2001)
[8] Mavromoustakis, C., Karatza, H.: Reliable File Sharing Scheme for Mobile Peer-to-Peer Users Using Epidemic Selective Caching. In: Proceedings of IEEE International Conference on Pervasive Services (ICPS), Santorini, Greece, pp. 169–177 (July 2005)
[9] Mavromoustakis, C.X., Karatza, H.D.: Community oriented neighboring feedback for efficient end-to-end delay-sensitive MP2P streaming via temporal relay-based replication. Journal of Algorithms & Computational Technology 5(2), 177–198 (2011)
[10] Mavromoustakis, C.X., Karatza, H.D.: Dispersed information diffusion with level and schema-based coordination in mobile peer to peer networks. Cluster Computing (Computer Communications & Networks) 10(1), 33–45 (2007)
[11] Ko, Y.B., Vaidya, N.H.: Flooding-based geocasting protocols for mobile ad hoc networks. Mobile Networks and Applications 7(6), 471–480 (2002)
[12] Blough, D., Resta, G., Santi, P.: A Statistical Analysis of the Long-Run Node Spatial Distribution in Mobile Ad Hoc Networks. In: Wireless Networks vol. 10, pp. 543–554. Kluwer Academic Publishers (2004)
[13] Kulik, J., Heinzelman, W.R., Balakrishnan, H.: Negotiation-based protocols for disseminating information in wireless sensor networks. Wireless Networks 8(2-3), 169–185 (2002)
[14] Dembo, A., Peres, Y., Rosen, J., Zeitouni, O.: Thick points for spatial Brownian motion: multifractal analysis of occupation measure. Ann. Probab. 28(1), 1–35 (2000)
[15] Santi, P., Blough, D.: The critical transmitting range for connectivity in sparse wireless ad hoc networks. IEEE Transactions on Mobile Computing 2(1), 25–39 (2003)
[16] Haas, Z., Pearlman, M.: The performance of query control schemes for the zone routing protocol. ACM/IEEE Transactions on Networking 9(4), 427–438 (2001)
[17] Pemantle, R., Peres, Y.: What is the probability of intersecting the set of Brownian double points? Ann. Probab. 35, 2044–2062 (2007)