Proactive Selective Neighbor Caching for Enhancing Mobility Support in Information-Centric Networks

Xenofon Vasilakos, Vasilios A. Siris, George C. Polyzos, and Marios Pomonis
Mobile Multimedia Laboratory
Department of Informatics
Athens University of Economics and Business
76 Patission Str., Athens, Greece GR10434
{xvas, vsiris, polyzos}@aueb.gr, mariospomonis@gmail.com

ABSTRACT

We present a Selective Neighbor Caching (SNC) approach for enhancing seamless mobility in ICN architectures. The approach is based on proactively caching information requests and the corresponding items to a subset of proxies that are one hop away from the proxy a mobile is currently connected to. A key contribution of this paper is the definition of a target cost function that captures the tradeoff between delay and cache cost, and a simple procedure for selecting the appropriate subset of neighbors which considers the mobility behavior of users. We present investigations for the steady-state and transient performance of the proposed scheme which identify and quantify its gains compared to proactively caching in all neighbor proxies and to the case where no caching is performed. Moreover, our investigations show how these gains are affected by the delay and cache cost, and the mobility behavior.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*distributed networks, network communications.*

General Terms

Design, Performance

Keywords

Mobility, Information-Centric Networking, Publish-Subscribe

1. INTRODUCTION

Information-Centric Networks (ICNs) have two key differences compared to IP networks that have implications to mobility: First, they employ a receiver-driven model where receivers request content by its name, asynchronously, from the publishers. Second, the transport of content from publishers to receivers is performed in a connectionless (stateless) manner, in contrast to TCP's connection-oriented (statefull) end-to-end control involving location-dependent IP addresses. Both features allow mobile receivers to re-send requests for content they did not receive while they were at their previous attachment point or could not receive while they were disconnected, without requiring re-establishing a connection or cumbersome and costly overlay solutions such as mobile IP.

Despite the aforementioned support for mobility, increased delay for receiving data can be incurred when a mobile sends a request for content but disconnects or moves to a new attachment point before receiving the requested data. This delay can be significant for applications with strict delay requirements. Examples of such applications include real-time emergency notification services, teleconferencing, and online gaming, which are sensitive to the end-to-end delay. Another example is streaming multimedia services which are sensitive to delay jitter.

The key premise of this paper is that this delay can be reduced by exploiting knowledge from information requests (or subscriptions), which is available due to the receiver-driven model of ICN architectures, and knowledge from the users' mobility behavior; this knowledge can be utilized by *proxies* that enhance mobility support. Proxies can be viewed as special caches with additional functionality to handle information requests on behalf of mobiles, and pre-fetch and cache items that match a mobile's requests while the mobile is in a handover phase or is disconnected from the network. Such an application of caching is different from its typical application of serving requests for the *same content* from *different users.* Understanding the advantages of caching for enhancing mobility support, we believe, is important for understanding the overall benefits from in-network caching in ICN architectures.

There are three types of caching solutions identified in literature based upon *when* and *where* subscriptions and information items are cached: i) reactive approaches [2, 5, 12, 13], ii) durable subscriptions [4], and iii) proactive approaches [1, 7]. In *reactive approaches*, when a mobile disconnects from its current proxy, the latter keeps caching items that match the mobile's subscriptions. When it reconnects, the mobile informs the new proxy of the old proxy's identity, and the new proxy requests from the old proxy all items that have been cached during the disconnection period. This reactive procedure has the disadvantage of increased delay for the

ICN'12, August 17, 2012, Helsinki, Finland.

new proxy to start forwarding items to the mobile, since the proxy must receive all the items cached in the old proxy before it can start forwarding. This delay can be avoided with both durable subscriptions and proactive caching solutions.

With *durable subscriptions*, proxies maintain a mobile's subscriptions and cache items matching these subscriptions, independent of whether the mobile is connected to the proxy or not. Without any additional mechanisms, in order to avoid loosing information, all proxies within a domain that a mobile can possibly connect to would need to maintain subscriptions and cache matching items; this incurs significant memory costs.

Finally, *proactive caching* solutions enhance mobility by proactively caching items matching a mobile's subscriptions in *single hop distance "neighboring" proxies*, before the mobile disconnects. The selection of proxies to proactively cache items can be based on prediction [1], or knowledge of all neighboring proxies lying one hop ahead [7]. The work of [1] does not propose a specific prediction algorithm, whereas with the approach of [7], when a mobile disconnects *all neighboring proxies* start caching items matching the mobile's subscriptions. Hence, when the mobile connects to one of these proxies, it can quickly receive items that were transmitted during its disconnection. Proactive approaches trade-off buffer space for reduced delay in forwarding items to subscribers.

A selective neighbor caching approach has been proposed for reducing the handover delay in WLANs [9]. The motivation is that, even when the number of neighboring APs is large, at most 3 or 4 access points are targets of the handoffs [9]; it is likely that a similar motivation applies for ICNs, since the movement of users is the cause for mobility in both cases. However, the application of such an idea for supporting mobility in ICNs is different due to the different nature of the problem: in ICNs the objective is to forward to mobiles items that match their subscriptions, whereas in [9] the objective is to proactively send a mobile's context to neighboring access points in order to reduce association delay. As a result, the model and corresponding procedure for selecting the subset of neighbors is fundamentally different. We also note that mobility prediction has been used to improve QoS support during handovers in cellular networks, e.g., see [3, 11].

In this paper we propose a Selective Neighbor Caching (SNC) approach which *selects an appropriate subset of neighboring proxies* that minimizes the mobility costs in terms of expected average delay and caching costs. The preliminary ideas of the target cost function used in SNC were presented in an extended abstract [10]; in the current paper we present the procedure for selecting neighbor proxies to proactively cache information items and discuss its implementation. Moreover, we present and discuss performance results that illustrate the tradeoff between the delay and cache cost, and we identify and quantify SNC's gains compared to full proactive caching, i.e., caching at all neighbor proxies, and when caching is not used.

The remainder of the paper is structured as follows: In Section 2 we present the target cost function and decision procedure for the proposed SNC approach, and in Section 3 we discuss its application over two representative ICN architectures: *Publish-Subscribe Internetworking* (PSI) and *Content-Centric Networking* (CCN). In Section 4 we evaluate the steady-state and the transient performance of SNC, comparing it with the case of full proactive caching and when no caching is performed. Finally, in Section 5 we conclude the paper identifying ongoing research directions.

2. SELECTIVE NEIGHBOR CACHING

Consider a mobile that is initially connected to proxy i, Fig. 1. Assume that the probabilities p_{ij} of the mobile connecting to proxy $j \in J$ (set of i's neighbor proxies) after it disconnects from i are known. Note that the term *neighbor* does not correspond to geographic proximity, but is related to the mobile's sequence of network attachment points; for example, when a mobile is connected to a WLAN in a subway station, the neighboring proxies can be those located in other subway stations.

The key idea of the SNC approach proposed in this paper is to select an optimum subset of neighbor proxies to which i will send the mobile's subscriptions, and which will proactively cache information items matching these subscriptions. If a mobile connects to one of the proxies in this subset, then the mobile can immediately receive information items it did not receive due to its disconnection from the previous proxy. Let the probability that the mobile connects to a proxy in the subset $S \subseteq J$ be $P_{\text{hit}}(S) = \sum_{j \in S} p_{ij}$. The optimum subset of neighbor proxies S^* is the set that minimizes the following *target cost function*:

$$P_{\text{hit}}(S) \cdot C_{\text{hit}} + (1 - P_{\text{hit}}(S)) \cdot C_{\text{miss}} + N(S) \cdot C_{\text{cache}}\,, \quad (1)$$

where $N(S)$ is the number of proxies in set S. C_{hit} is the delay experienced by the mobile in order to receive its requested information items from a proxy j that has proactively cached items matching its subscriptions; the probability for the mobile to move to such a proxy is $P_{\text{hit}}(S)$. $C_{\text{miss}} > C_{\text{hit}}$ is the delay cost for a mobile to receive items from their publishers, which can occur with probability $1 - P_{\text{hit}}(S)$. Finally, C_{cache} is the cost for a proxy to cache a mobile's subscriptions and the matching items. If we assume that the memory requirements for storing subscriptions is small relative to information items, then C_{cache} depends linearly on the memory requirements for storing a single item.

The proposed target cost function captures the tradeoff between the average delay for a mobile to receive an information item, which is given by the first and second terms of (1), and the corresponding cache cost, which is given by the third term of (1). Increasing the set S of neighbors that proactively cache items increases $P_{\text{hit}}(S)$, hence on one hand reduces the average delay for obtaining information items since $C_{\text{miss}} > C_{\text{hit}}$, but on the other hand increases the buffers that are required to cache the items. A key issue is how to quantify the cache cost in order to add it to the delay. In Section 2.2 we discuss how the cache cost can be quantified in the practical case where proxies have a fixed cache size.

When a mobile disconnects[1] from its current proxy i, then i notifies its neighbor proxies in the set S^* to start caching items that match the mobile's subscriptions, starting from the item after the last item forwarded to the mobile. When the mobile connects to a new proxy, then the latter informs the old proxy that the mobile has reconnected, and subse-

---

[1] Depending on the disconnection period's duration, an alternative is that neighbor proxies start to proactively cache prior to a mobile's disconnection. The target cost function (1) still remains the same.

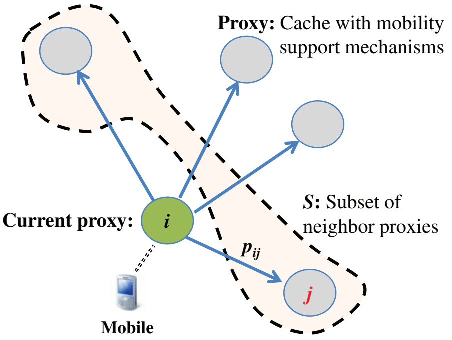


Figure 1: A mobile is currently connected to proxy i. Based on the probabilities for the mobile to connect to proxy i's one hop neighbors, SNC selects the subset of neighbor proxies to proactively transmit the mobile's subscriptions in order to minimize the target cost function (1).

quently the old proxy informs all neighbor proxies in S^* to stop caching items for that specific mobile. In Section 3 we discuss how the above communication between proxies can be implemented in two representative ICN architectures.

2.1 Finding S^*

Next we discuss how to find the set S^* that minimizes the target cost function (1). Consider a subset $S' \subseteq J$. If items matching a mobile's subscriptions are pre-fetched by all proxies in S', the total cost would be

$$P_{\text{hit}}(S')\cdot C_{\text{hit}} + \left(1 - P_{\text{hit}}(S')\right)\cdot C_{\text{miss}} + N(S')\cdot C_{\text{cache}}\,. \quad (2)$$

Assume now that we want to decide whether to include a proxy j in the subset of proxies that pre-fetch information items. The total cost when items are pre-fetched by proxies in the set $S' \cup \{j\}$ is

$$P_{\text{hit}}(S'\cup\{j\})\cdot C_{\text{hit}} + \left(1 - P_{\text{hit}}(S'\cup\{j\})\right)\cdot C_{\text{miss}} + (N(S)+1)\cdot C_{\text{cache}}\,. \quad (3)$$

Pre-fetching items in proxy j would be beneficial if and only if cost (3) $\leq$ (2), which is equivalent to

$$p_{i,j} \geq \frac{C_{\text{cache}}}{C_{\text{miss}} - C_{\text{hit}}} = \frac{\frac{C_{\text{cache}}}{C_{\text{hit}}}}{\frac{C_{\text{miss}}}{C_{\text{hit}}} - 1}\,. \quad (4)$$

The numerator in (4) denotes the cost for caching, while the denominator denotes the gain with caching, in terms of reduced delay. Hence, finding the set S^* that minimizes the cost function (1) involves deciding individually for each neighboring proxy whether it should proactively cache information items matching a mobile's subscriptions, using (4). In addition to involving a simple comparison, (4) can be applied centrally at the original proxy i, or decentralized at every neighbor $j \in J$ to decide whether to proactively cache items or not.

In equations (1) and (4) we have assumed that the delay cost C_{miss} for obtaining items from their publishers is the same for all publishers and for all items; the latter is equivalent to assuming that all items have the same size. In practice, there can be a different delay cost for receiving items from different publishers and for different item sizes. The delay costs C_{hit} and C_{miss} can be estimated from actual measurements of the time for the mobile to receive items from a local proxy or from the item's publisher, respectively, after issuing the request for that item. Note that these delay costs refer to the neighbor proxies J, which the mobile can connect to. Hence, the estimation of the costs C_{hit} and C_{miss} should be performed at each neighbor proxy $j \in J$. Moreover, the cache cost C_{cache} can depend on the buffer utilization, as we discuss in Section 2.3, hence different neighbor proxies can have different cache costs. From the above, (4) can be written more generally as

$$p_{i,j} \geq \frac{\frac{C^j_{\text{cache}}}{C^j_{\text{hit}}}}{\frac{C^j_{\text{miss}}}{C^j_{\text{hit}}} - 1}\,. \quad (5)$$

An advantage of implementing (5) in the neighbor proxies $j \in J$, rather than the original proxy i, is that this would avoid the need for communicating the ratios $C^j_{\text{cache}}/C^j_{\text{hit}}$ and $C^j_{\text{miss}}/C^j_{\text{hit}}$ from all the neighbor proxies to proxy i. On the other hand, the transition probability $p_{i,j}$ can be estimated by the ratio of the number of transitions from proxy i to proxy j, which is known to j, over the total number of transitions from proxy i; the latter needs to be communicated from i to all its neighbor proxies $j \in J$. Estimation of the transition probabilities can be enhanced with information such as location, orientation, and road/path topology, which have been used for mobility prediction in cellular networks [3, 11].

2.2 Cache cost

The target cost function (1) and (5) involves the delay cost for obtaining an information item and a cost for proactively caching an information item. In a practical application of the proposed SNC approach to a system where proxies have a fixed cache size, a goal would be to achieve a high cache utilization. This suggests that when the cache utilization is low, then the cache cost should also be low to allow more proactive caching in neighbor proxies. On the other hand, when the cache utilization is high then the cache cost should also be high in order to reduce the amount of proactive caching. Based on the above, the cache cost can be defined as

$$C^j_{\text{cache}} = \frac{a}{1 - \rho^j_{\text{util}}}\,, \quad (6)$$

where ρ^j_{util} is the cache utilization at proxy j. The cache cost coefficient a can be adjusted to achieve a minimum cache utilization.

2.3 Extensions

The transition probability on the left-hand side of (5) captures the average behavior of all mobiles that connect to proxy i. If there are different types of mobiles, with mobiles of the same type having a similar mobility behavior, whereas mobiles of a different type have a distinct mobility behavior, then one can apply (5) considering the specific proxy transition probability for each type of mobile. A mobile's type can be identified through some characteristic or information made available by the mobile.

Equation (5) assumes that the cost for proactively caching an information item is incurred by one mobile, that requested the specific information item. However, when more than one mobile request the same information item, then the caching cost should be shared by the mobiles requesting that item. If the average number of requests for an information item k is n_k, then the decision inequality (5) can be written as:

$$p_{i,j} \geq \frac{\frac{C^j_{\text{cache}}}{C^j_{\text{hit}}}}{n_k \cdot \left(\frac{C^j_{\text{miss}}}{C^j_{\text{hit}}} - 1\right)} .$$

3. SNC OVER PSI AND CCN

SNC can be implemented in software modules running in the proxies, without modifying the underlying ICN architecture. The implementation of SNC involves the following: i) estimation of the cost ratios $C^j_{\text{miss}}/C^j_{\text{hit}}$ and $C^j_{\text{cache}}/C^j_{\text{hit}}$ and the transition probabilities p_{ij} that appear in (5), ii) selection of the neighbor proxies that will perform proactive caching, iii) transmission, from the original proxy to its neighbor proxies, of a mobile's requests and notifications to start/stop proactive caching, and iv) notification containing the mobile's new attachment point which is sent by the mobile's new proxy to its old proxy.

Recall from Section 2.1 that costs and transition probabilities can be estimated locally at each neighbor proxy $j \in J$. The only information that needs to be communicated in order to estimate the transition probabilities is the total number of transitions from the old proxy i to any of its neighbors; this communication involves a one-to-many dissemination of the same information, from proxy i to all its neighbors. Similarly, the transmission of a mobile's requests and notifications to start/stop proactive caching from the original proxy to neighbor proxies also involves a one-to-many dissemination of the same information. If the decision of whether to proactively cache information items is taken at the neighbor proxies, then the aforementioned one-to-many communications can be performed in a receiver-driven (pull-based) fashion. Hence, they can be appropriately implemented using the receiver-driven (pull-based) communication primitive in the Content-Centric Networking (CCN) architecture [8], exploiting CCN's ability to effectively disseminate that same information to multiple interested receivers. Similarly, it can be implemented in the Publish-Subscribe Internetworking (PSI) [6] architecture, which supports a receiver-driven model at the rendezvous (or resolution) layer. Moreover, one-to-many dissemination can exploit both CCN and PSI's native multicast capabilities.

Finally, the last communication mentioned above (in iv)) involves the old proxy being notified of the mobile's new proxy. This can be implemented in both PSI and CCN by having the old proxy issue a subscription message (in PSI) or an Interest (in CCN) that corresponds to an information item containing the mobile's connection status. Once the mobile connects to a new proxy, the latter issues a publication announcement (PSI) or a Data packet (CCN) that matches the aforementioned subscription or Interest, respectively. Note that this communication exploits the ability to asynchronously issue subscriptions or Interests prior to publication announcements or Data packets in the PSI or CCN architecture, respectively.

4. PERFORMANCE EVALUATION

In this section we present analytical investigations for the steady-state and simulation investigations for the transient performance of SNC. The analytical investigations illustrate the tradeoff between the average delay, the cache cost, their influence on the total cost, and how the tradeoff and gains of SNC depend on the cost ratios $C_{\text{miss}}/C_{\text{hit}}$, $C_{\text{cache}}/C_{\text{hit}}$ and on the mobility pattern. The simulation investigations show the transient delay gains of SNC when the cache size is fixed and when the proxy transition probabilities, based on which the selection of neighbors to proactively cache items is made, are estimated on-line rather than a priori known as in the analytical investigations.

For the analytical investigations we consider scenarios with one proxy and 4 neighboring proxies, whereas in the transient investigations we consider scenarios with 5 and 8 neighbor proxies. The analytical investigations show the average delay, the cache cost, and the total cost, which are defined below:

$$\text{Total cost} = \underbrace{P_{\text{hit}} \cdot C_{\text{hit}} + (1 - P_{\text{hit}}) \cdot C_{\text{miss}}}_{\text{Average delay}} + \underbrace{N(S) \cdot C_{\text{cache}}}_{\text{Cache cost}} . \tag{7}$$

The analytical results show SNC's gain, defined as the reduction of the total cost relative to full proactive caching, for which $S = J$ and $P_{\text{hit}} = 1$, and when no caching is used, for which $S = \emptyset$ and $P_{\text{hit}} = 0$. The simulation results show SNC's gains in terms of average delay, compared to full and no caching.

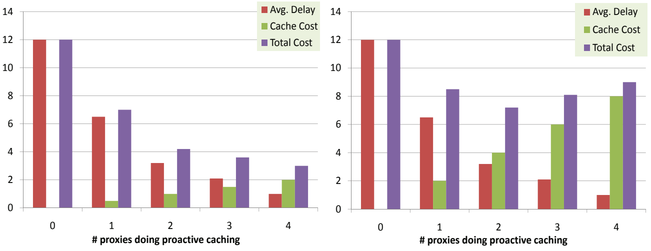
4.1 Analytical investigations

In this subsection we present analytical investigations that illustrate the tradeoff between the average delay and caching cost, and how the cost ratios $C_{\text{miss}}/C_{\text{hit}}$ and $C_{\text{cache}}/C_{\text{hit}}$, and the proxy transition probabilities influence SNC's gains compared to the case where no proactive caching is performed and when all neighbors proactively cache a mobile's subscriptions; the latter corresponds to the proactive approach of [7].

Average delay and cache cost tradeoff: Fig. 2 shows the tradeoff between the average delay and the caching cost. As expected, the average delay decreases when more neighbor proxies proactively cache a mobile's subscriptions. On the other hand, the caching cost increases when more neighbor proxies proactively cache subscriptions. SNC captures this tradeoff and selects the subset of neighbors that gives the lowest total cost. Different cost ratios $C_{\text{miss}}/C_{\text{hit}}$ and $C_{\text{cache}}/C_{\text{hit}}$ result in a different selection of neighbors that proactively cache information items: SNC selects 4 neighbor proxies ($N^* = 4$) in Fig. 2(a), 2 neighbors in Figs. 2(b) and 2(d), and zero neighbors in Fig. 2(c). Note that the two neighbors SNC selects in Figs. 2(b) and 2(d) are the two with the highest transition probability (0.5 and 0.3).

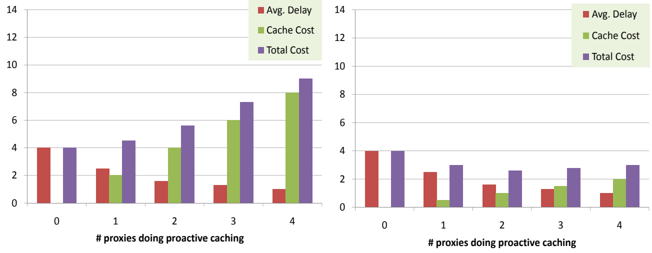
Influence of C_{cache}, C_{miss}, and proxy transition probabilities: Fig. 3 shows that as the ratio $C_{\text{miss}}/C_{\text{hit}}$ increases SNC's gains compared to full proactive caching are lower; this occurs because the optimal number of proxies for proactive caching increases, Fig. 3(b). On the other hand, SNC's gains compared to the case where no caching is performed increase with $C_{\text{miss}}/C_{\text{hit}}$.

Fig. 4 shows the influence of the ratio $C_{\text{cache}}/C_{\text{hit}}$ on the gains and the optimal number of neighbors that perform proactive caching, for two proxy transition probabilities. Observe that increasing the ratio $C_{\text{cache}}/C_{\text{hit}}$ reduces SNC's gains compared to the case where proactive caching is not performed, whereas it increases SNC's gains compared to the case of proactive caching in all neighbors. Fig. 4(a) shows that when mobiles move to neighboring proxies with



(a) $C_{\text{miss}}/C_{\text{hit}}=12$, $C_{\text{cache}}/C_{\text{hit}}=0.5, N^*=4$

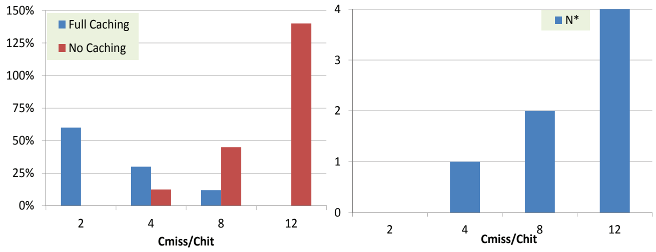
(b) $C_{\text{miss}}/C_{\text{hit}}=12$, $C_{\text{cache}}/C_{\text{hit}}=2, N^*=2$



(c) $C_{\text{miss}}/C_{\text{hit}}=4$, $C_{\text{cache}}/C_{\text{hit}}=2, N^*=0$

(d) $C_{\text{miss}}/C_{\text{hit}}=4$, $C_{\text{cache}}/C_{\text{hit}}=0.5, N^*=2$

Figure 2: Tradeoff between average delay and caching cost. SNC considers this tradeoff and selects the subset of neighbors that gives the lowest total cost.



(a) Gain

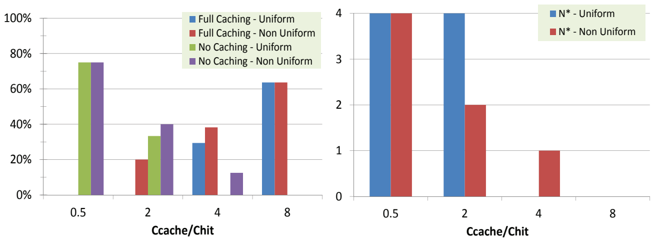
(b) N^*

Figure 3: Influence of $C_{\text{miss}}/C_{\text{hit}}$ on gains and optimal number of proxies that perform pre-fetching $N^* = N(S^*)$. $C_{\text{cache}}/C_{\text{hit}} = 1$, $\{p_{ij}\} = \{0.5, 0.3, 0.1, 0.1\}$.

a uniform probability, then SNC's gain compared to full caching and no caching are smaller in comparison to the case of non-uniform probabilities for intermediate cost ratios $C_{\text{miss}}/C_{\text{hit}}$ (2 and 4 in Fig. 4). Fig. 4(b) shows that when the proxy transition probabilities are uniform, the optimal number of proxies to perform pre-fetching exhibits a bimodal behavior: either it is optimal to not pre-fetch in any proxy or it is optimal to pre-fetch in all neighboring proxies.

4.2 Simulation investigations

Next we evaluate SNC using the OMNeT++ simulation framework. Due to space limitations, we present the results from a small set of simulations. We consider 100 mobiles which are uniformly distributed to 5 proxies. From its initial proxy, a mobile can move to one of 5 other proxies with probabilities 80%, 10%, 5%, 2.5%, 2.5%. We assume that the destination proxy with probability 80% is different for different initial proxies, hence the number of mobiles that connect to the destination proxies from any initial proxy is uniform. After a mobile moves to a new proxy, a new mobile connects to one of the initial 5 proxies, hence the total number of mobiles in the system remains 100, and each proxy has on average 20 mobiles. We also consider a scenario with 8+8 proxies, where the total number of mobiles is 160, uniformly distributed to the different proxies. Information items requested by mobiles have the same size, and each proxy can cache up to 20 items. For the above scenario, we have found that an appropriate value for the cache cost coefficient in (6) is $a = 2$. If we consider a minimum target cache utilization (e.g., 80%), the cache cost coefficient can be estimated from equations (5) & (6), and the cache size, distribution of mobiles, and proxy transition probabilities. Finally, note that the transition probabilities are estimated online, rather than a priori known as in the analytical investigations.



(a) Gain

(b) N^*

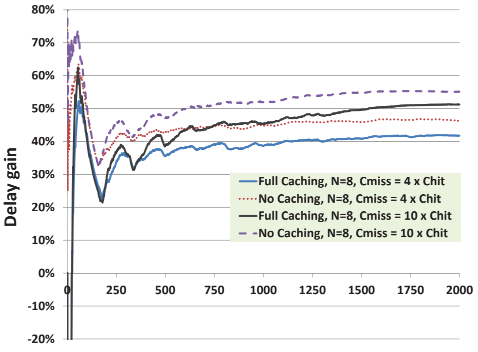
Figure 4: Influence of $C_{\text{cache}}/C_{\text{hit}}$ on gains and optimal number of proxies that perform pre-fetching. $C_{\text{miss}}/C_{\text{hit}} = 12$, $\{p_{ij}\} = \{0.5, 0.3, 0.1, 0.1\}$ (non-uniform), $\{0.25, 0.25, 0.25, 0.25\}$ (uniform).

Fig. 5(a) shows that SNC can achieve lower delay compared to both full proactive caching and no caching. As expected, the delay improvement is higher for a higher ratio $C_{\text{miss}}/C_{\text{hit}}$. The convergence time is necessary for the cache utilization and the estimated proxy transitions probabilities to converge. Fig. 5(b) shows SNC's delay gain for a different number of proxies. The above results show that the intelligent selection of neighbor proxies to cache information can efficiently utilize caches, significantly improving delay.

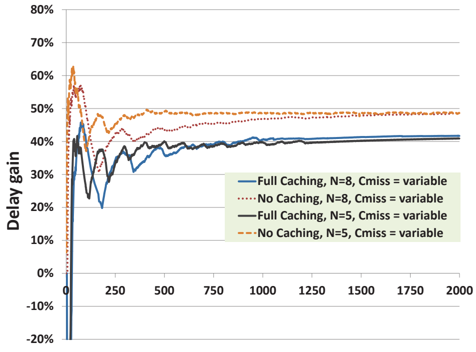
5. CONCLUSIONS AND FUTURE WORK

We have presented a Selective Neighbor Caching (SNC) approach for enhancing seamless mobility support in ICN architectures. SNC selects the set of neighbor proxies to proactively cache information requests based on the tradeoff between cache cost and delay. We have presented analytical and simulation investigations that quantify SNC's gains compared to full proactive caching and when no caching is performed.

Ongoing work includes investigating how the network topology, traffic demand, information item size, disconnection period duration, and in-network caching influence SNC's performance, and how the cache cost can be adapted on line for scenarios with a different number of mobiles and proxies, proxy transition probabilities, and cache sizes. Also, we are investigating the case of multiple levels of proxies. When there are proxies at different levels, the delay cost for a proxy at a higher level is higher. On the other hand, a proxy at a higher level can potentially serve a larger number of mobile attachment points, hence using proxies at a higher level can



(a) $C_{\text{miss}}/C_{\text{hit}} = 4, 10$



(b) 5 and 8 destination proxies

Figure 5: Delay gains based on simulation. In the left figure, $C_{\text{miss}}/C_{\text{hit}}$ is randomly selected in the range $[4, 5]$.

reduce the number of proxies that need to proactively cache information, hence the total cache cost.

6. ACKNOWLEDGMENT

The work in this paper was supported by the FP7 ICT project PURSUIT, under contract ICT-2010-257217.

7. REFERENCES

[1] I. Burcea, H. Jacobsen, E. Lara, V. Muthusamy, and M. Petrovic. Disconnected Operation in Publish/Subscribe Middleware. In *Proc. of IEEE Int'l Conference on Mobile Data Management (MDM)*, 2004.

[2] M. Caporuscio, A. Carzaniga, and A. L. Wolf. Design and Evaluation of a Support Service for Mobile Publish/Subscribe Applications. *IEEE Transactions on Software Engineering*, 29(12):1059–1071, December 2003.

[3] M.-H. Chiu and M. A. Bassiouni. Predictive Schemes for Handoff Prioritization in Cellular Networks Based on Mobile Positioning. *IEEE Journal on Selected Areas in Communications*, 18(3):510–522, March 2000.

[4] U. Farooq, E. Parsons, and S. Majumdar. Performance of Publish/Subscribe Middleware in Mobile Wireless Networks. In *Proc. of International Workshop on Software and Performance (WOSP)*, 2004.

[5] L. Fiege, F. C. Gartner, O. Kasten, and A. Zeidler. Supporting mobility in content-based publish/subscribe middleware. In *Proc. of Middleware 2003, LNCS, volume 2672*, 2003.

[6] N. Fotiou, P. Nikander, D. Trossen, and G. Polyzos. Developing information networking further: From psirp to pursuit. In *Proc. 7th ICST Int'l Conf. on Broadband Commun., Networks, and Systems*, 2010.

[7] A. Gaddah and T. Kunz. Extending mobility to publish/subscribe systems using a pro-active caching approach. *Mobile Information Systems*, 6(4):293–324, 2010.

[8] V. Jacobson, D. K. Smetters, J. D. Thornton, M. Plass, N. Briggs, and R. Braynard. Networking Named Content. In *Proc. of CoNEXT*, 2009.

[9] S. Pack, H. Jung, T. Kwon, and Y. Choi. SNC: A Selective Neighbor Caching Scheme for Fast Handoff in IEEE 802.11 Wireless Networks. *ACM Mobile Computing and Communications Review*, 9(4):39–49, October 2005.

[10] V. Siris, X. Vasilakos, and G. Polyzos. A Selective Neighbor Caching Approach for Supporting Mobility in Publish/Subscribe Networks. In *Proc. of ERCIM Workshop on eMobility. Held in conjunction with WWIC*, 2011.

[11] W.-S. Soh and H. S. Kim. QoS provisioning in cellular networks based on mobility prediction techniques. *IEEE Communications Magazine*, 41(1):86–92, January 2003.

[12] V. Sourlas, G. S. Paschos, P. Flegkas, and L. Tassiulas. Mobility support through caching in content-based publish/subscribe networks. In *Proc. of IEEE/ACM Int'l Conf. on Cluster, Cloud and Grid Computing*, 2010.

[13] J. Wang, J. Cao, J. Li, and J. Wu. MHH: A Novel Protocol for Mobility Management in Publish/Subscribe Systems. In *Proc. of IEEE Int'l Conference on Parallel Processing (ICPP)*, 2007.