




Please cite the Published Version

Vasilakos, Xenofon, Al-Khalidi, Mohammed , Siris, Vasilios A, Reed, Martin J , Thomos, Nikolaos  and Polyzos, George C (2017) Mobility-based proactive multicast for seamless mobility support in cellular network environments. In: 1st Workshop on Mobile Edge Communications (MECOMM), 21 August 2017, Los Angeles, CA.

DOI: <https://doi.org/10.1145/3098208.3098213>

Publisher: ASSOC COMPUTING MACHINERY

Version: Accepted Version

Downloaded from: <https://e-space.mmu.ac.uk/633518/>

Usage rights:  In Copyright

Additional Information: © ACM 2017. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in MECOMM '17: Proceedings of the Workshop on Mobile Edge Communications, <http://dx.doi.org/10.1145/3098208.3098213>

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

Mobility-based Proactive Multicast for Seamless Mobility Support in Cellular Network Environments*

Xenofon Vasilakos

Department of Informatics, School of Information Sciences and Technology, Athens University of Economics and Business
76 Patission Str
Athens, Greece GR 10434
xvas@aueb.gr

Martin J. Reed

School of Computer Science and Electronic Engineering, University of Essex
Wivenhoe Park
Colchester, UK CO4 3SQ
mjreed@essex.ac.uk

Mohammed Al-Khalidi

School of Computer Science and Electronic Engineering, University of Essex
Wivenhoe Park
Colchester, UK CO4 3SQ
mshawk@essex.ac.uk

Nikolaos Thomos

School of Computer Science and Electronic Engineering, University of Essex
Wivenhoe Park
Colchester, UK CO4 3SQ
nthomos@essex.ac.uk

Vasilios A. Siris

Department of Informatics, School of Information Sciences and Technology, Athens University of Economics and Business
76 Patission Str
Athens, Greece GR 10434
vsiris@aueb.gr

George C. Polyzos

Department of Informatics, School of Information Sciences and Technology, Athens University of Economics and Business
76 Patission Str
Athens, Greece GR 10434
polyzos@aueb.gr

ABSTRACT

Information-Centric Networking (ICN) is receiver driven, asynchronous and location-independent, hence it natively supports client-mobility. However, post-handover delay is a problem for delay-sensitive mobile applications, as they need to (re-)submit their subscriptions and wait for them to get resolved and (probably re-) transmitted before receiving the demanded data. To avoid this problem and optimize performance, this paper proposes a Mobility-based Proactive Multicast (MPM) scheme. Unlike reactive or blind multicast solutions proposed in the past, MPM takes autonomous decisions locally at various network access points (cells) prior to the movement of mobile clients, using a semi-Markov mobility prediction model that predicts next-cell transitions, along with anticipating the duration between the transitions for an arbitrary user in a cellular network. Since cellular backhaul links are typically a bottleneck, MPM trades-off effectively part of the capacity of the (congested) backhaul link for a decreased delay experienced by users after handovers thanks to a congestion pricing scheme used for backhaul capacity allocation. Our preliminary performance evaluation results show that MPM captures well the temporal locality of mobile requests due to the semi-Markov mobility prediction

*This work was carried out within the project POINT, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643990.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MECOMM '17, August 21, 2017, Los Angeles, CA, USA

© 2017 Association for Computing Machinery.

ACM ISBN 978-1-4503-5052-5/17/08...\$15.00

<https://doi.org/10.1145/3098208.3098213>

model, hence it achieves a better performance compared to both a (i) blind/naïve multicast and a (ii) content popularity-based proactive multicast.

CCS CONCEPTS

• **Networks** → **Wireless access points, base stations and infrastructure**; *Cross-layer protocols*;

KEYWORDS

proactive multicast; markov; mobility prediction

ACM Reference format:

Xenofon Vasilakos, Mohammed Al-Khalidi, Vasilios A. Siris, Martin J. Reed, Nikolaos Thomos, and George C. Polyzos. 2017. Mobility-based Proactive Multicast for Seamless Mobility Support in Cellular Network Environments. In *Proceedings of MECOMM '17, Los Angeles, CA, USA, August 21, 2017*, 6 pages.

<https://doi.org/10.1145/3098208.3098213>

1 INTRODUCTION

Information-centric Networks (ICNs) [20] can seamlessly support subscriber host mobility. Mobile receivers merely need to (re-)issue their subscriptions for Information Objects (IOs) after attaching to a new network location, and then have their subscriptions resolved by the network before the actual data transfer starts, utilizing anycast from some source to the current network location of the requesting mobile users. Nevertheless, increased delay for receiving data in ICN networks can still occur after a mobile completes a handover. Specifically, users need to *wait* for their subscriptions to be resolved before starting to receive their desired IOs. Alongside propagation delay, this can cause Quality-of-Service (QoS) degradation for *delay* sensitive applications which demand a tight delay bound, upon handovers, e.g. video streaming applications.

Past efforts [2, 7, 22] in the literature adapt *multicast* for handling large mobile populations. This idea can be the basis for supporting seamless mobility in heterogeneous wireless networking environments comprised by different types of micro, small, pico or femto cells, as well as Wi-Fi hotspots. Towards this aim, the Publish-Subscribe communication in ICNs is *intrinsically multicast*, which natively allows a *wide* distribution of the requested IO data around the network “areas” of the mobile subscribers. However, blind/naïve multicast in all possible network regions where the mobile resides, or can possibly move to, is resource-wasting. This is important as, although wireless technologies continuously increase their capacity, the corresponding capacity of the *backhaul* links which connect the wireless cells to the backbone network (typically ADSL or VDSL) are constrained relative to demand, particularly in scenarios that involve *large mobile populations*.

Inspired by proactive solutions designed for enhancing mobility support in cellular and/or Pub-Sub mobile environments via caching such as [3, 4, 10, 11], and particularly by work which focuses on leveraging *mobility prediction* information such as [19, 21], this paper presents an efficient *Mobility-based Proactive Multicast (MPM)* model tailored for delay-sensitive mobile applications. The solution exploits information on (i) users’ individual mobility via a *semi-Markov mobility prediction* model and knowledge of users’ requests, and applies a (ii) *congestion pricing* scheme for utilising the congested backhaul capacity of cells in an efficient manner, so as to take proactive multicast actions that reduce the expected cost of delay for users who move and attach to another cell. The proposed solution models the mobility behaviour of users in a cellular network as a semi-Markov process. It is applied for predicting the next-cell transitions, along with anticipating the duration between the transitions for an arbitrary user in a cellular network. The parameters used in the model can be derived from the cellular data that can be readily obtained from traffic logs. The basic MPM model follows a simple, lightweight and fully decentralised process of autonomous proactive multicast decisions at cells, and can be extended with bandwidth allocation replacement actions (MPM-R) that may further increase delay cost gains. Moreover, the solution can best fit within the network layer of ICN architectures which allow to have a full knowledge of users’ requests and to leverage the native multicast abilities of Pub/Sub communication. Nonetheless, the solution can also be applied as an application-layer solution over standard TCP/IP network architectures, in order to reduce the delay of establishing a new TCP connection between a mobile receiver and a remote server.

In what follows, Section 2 discusses the semi-Markov mobility prediction model that is used for taking the efficient proactive multicast decisions. Next, Section 3 presents the MPM model. Section 4 illustrates the performance evaluation of the proposed scheme. Section 5 describes the related work before we summarise our conclusions and future plans in Section 6.

2 SEMI-MARKOV PREDICTION MODEL FOR MOBILITY

Cellular networks are usually divided into cells, each served by at least one base station. Every cell in the network is identified by a unique cell ID which is also used to track and identify users’

location. The location of a user can also be identified by his/her geographic coordinates. In particular, users’ coordinates are directly mapped to the cell ID in which a user is located. Since our mobility prediction facilitates network resources management, it is therefore assumed that a user’s location identification via cell ID is sufficient from network’s perspective. Furthermore, the mobility history of a mobile user is also described by the list of successive visited cell IDs during users trip (i.e., the sequence of cells that the user was connected to throughout his journey). Thus, users’ mobility history patterns can be periodically recorded using the cell-ID representations. In this context, each cell that is building a mobility pattern profile, records the number of handovers made to neighbouring cells, as well as the residence time (time spent in the current cell i before transition to the next cell l occurs). This allows for the computation of the cell-transition probabilities $p_{i,l}$ and the distribution of cell residence times at each location from the set of recorded residence times. We assume that the network, at each of its cells, keeps a record of session residence time and the cell ID of the next-cell transition [1].

In this work, user mobility is modelled by a semi-Markov process [13], which allows for arbitrary distributed residence times and can be viewed as a process with an embedded Markov chain. Here, the embedded points are the time instants when a user attaches to a new cell. In a semi-Markov process, the successive state occupancy’s (visited cells represented by cell IDs) are governed by the transition probabilities $p_{i,l}$ of a Markov process. The residence time in any state depends on both the current-state and the next-state where the user will move. The semi-Markov kernel for a time-homogeneous process is given by $\Phi_{i,l}(t)$ [15], which denotes the probability that immediately after making the transition into state i , the process makes a transition to state l within t time units. $\Phi_{i,l}(t)$ is defined as

$$\Phi_{i,l}(t) = Pr\left\{X_{n+1} = l, T_{n+1} - T_n \leq t | X_n = i\right\} \quad (1)$$

where X_n and X_{n+1} represent the state of the system after n and $(n + 1)$ transitions, respectively. T_n and T_{n+1} are the times at which the n th and $(n + 1)$ th transitions occur, respectively. We can further express the kernel as $\Phi_{i,l}(t) = P_{i,l}\Psi_{i,l}(t)$, where

$$\Psi_{i,l}(t) = Pr\left\{T_{n+1} - T_n \leq t | X_{n+1} = l, X_n = i\right\} \quad (2)$$

$\Psi_{i,l}(t)$ represents the conditional probability that a transition will take place within t time units, given that the process has just entered state i and next will move to state l . The residence times in such a process can follow any arbitrary distribution. This allows for a convenient departure from the common assumption of exponentially distributed residence times, thus permitting a more accurate representation of the temporal behaviours [6].

Since it is known that as time goes to infinity, $\Psi_{i,l}$ tends to one, we have $P_{i,l} = \lim_{t \rightarrow \infty} \Phi_{i,l}(t)$. We also define the kernel $\phi_{i,l}(t) = p_{i,l}\psi_{i,l}(t)$, where $\psi_{i,l}(t)$ is the corresponding residence time and $P = [P_{ij}]$, $\forall i, j \in [1, n]$ is the state transition probability matrix

defined as:

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & p_{2,3} & \cdots & p_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n,1} & p_{n,2} & p_{n,3} & \cdots & p_{n,n} \end{bmatrix} \quad (3)$$

where

$$\phi_{i,l}(t) = Pr\{X_{n+1} = l, T_{n+1} - T_n = t | X_n = i\} \quad (4)$$

and

$$\psi_{i,l}(t) = Pr\{T_{n+1} - T_n = t | X_{n+1} = l, X_n = i\} \quad (5)$$

Utilizing the past handover history of a user which is represented as a tuple \langle attachment time, cell ID \rangle , the state transition probability matrix P and the residence time distribution matrix Ψ are initialized as follows:

$$p_{i,l} = \frac{|H_{i,l}|}{|H_i|} \quad (6)$$

$$\psi_{i,l}(\tau) = \frac{|H_{i,l,\tau}|}{|H_{i,l}|} \quad (7)$$

where $H_{i,l}$ is the number of handovers from cell i to l , and H_i is the total number of user handovers from cell i regardless of the destination. $H_{i,l,\tau}$ is the number of handovers of user's from cell i to l with a residence time within the defined τ time interval. Whenever there is a handover from cell i to l , $p_{i,l}$ and $\psi_{i,l}(\tau)$ and $\phi_{i,l}(\tau)$ are updated. The cell with the highest $\phi_{i,l}(\tau)$ probability is chosen as the predicted future destination when the time spent in cell i falls within time interval τ . This prediction process can be used for both offline learning, where a training period is needed before actively applying the learned prediction matrices, or it can also be used for online learning where the probabilities are continuously updated based on the experienced transitions. Therefore, any changes in the mobility behaviour are directly reflected on the prediction probabilities. In this work, we focus on the offline prediction model as operators are aware of their transition statistics. We use the resulting (residence time dependent) probability matrix to predict mobile nodes next destination cells.

3 MOBILITY-BASED PROACTIVE MULTICAST MODELS

In this section we present our proactive multicast allocation decision model. Our objective is to minimise the average delay across all requested IOs, subject to the cell backhaul bandwidth constraints. Note that the proactive actions taken in one cell are always independent from the corresponding decisions taken by other cells, hence our model is fully distributed. Let q_s^l denote the probability that the mobile requesting IO s moves to cell l and B_l denote the maximum capacity at the backhaul link of l . For the sake of simplicity Probability q_s^l is defined by the prediction model discussed in Sec. 2, where $q_s^l \equiv \phi_{i,l}(\tau)$ when the time spent in cell i falls within time interval τ . Also, let S_l be the set of IOs that are currently requested by mobiles that have non-zero probability to

move to l and L be the set of all cells. We define the following optimization problem:

$$\min_{b_s^l} \sum_{s \in S_l} \mathcal{D}_s \quad (8)$$

$$\text{subject to } \sum_{s \in S_l} o \cdot b_s^l \leq B_l, \quad \forall l \in L, \quad (9)$$

where $\mathcal{D}_s = \sum_{l \in L} \mathcal{D}_s^l$ is the average delay for obtaining s after completing a handover, b_s^l is equal to one if s is proactively multicast by l and zero otherwise. \mathcal{D}_s^l is equal to $q_s^l D_{miss}$ if s is *not* proactively multicast by l ($b_s^l = 0$), in which case the user has to suffer a delay for re-submitting its request and wait for it to be resolved in order to start getting the data of s . Alternatively, if s is proactively multicast by l , then the expected delay cost for obtaining s by the mobile nodes which attach to l is $q_s^l D_{hit}$ ($b_s^l = 1$). The latter cost is smaller than $q_s^l D_{miss}$, as the data are immediately available to the users that attach to l and D_{hit} represents only a minimum delay for completing the attachment process of mobiles to l in the MAC layer. The above optimization problem involves *autonomous* proactive decisions by each cell $l \in L$ for each IO s requested by any mobile that can possibly move to l , based on the mobile transition probabilities of the mobiles requesting s to l .

3.1 Basic model

In order to efficiently utilize the available backhaul link of l , we introduce a *congestion price* p_l which is adapted with each newly coming request to l based on the current aggregate demand for multicast $\sum_{s \in S_l} o \cdot b_s^l$ at l and the available bandwidth of the link B_l :

$$p_l = \left[p_l' + \gamma \left(\sum_{s \in S_l} o \cdot b_s^l - B_l \right) \right]^+, \quad (10)$$

where p_l' is the previous price value before receiving the latest request and that γ is a price update factor determining how quickly the price adapts to changes in the demand for proactive multicast. The multicast decision for s at cell l is based on rule

$$b_s^l = \begin{cases} 1 & \text{if } Q_s^l (D_{miss} - D_{hit}) \geq p_l, \\ 0 & \text{Otherwise,} \end{cases} \quad (11)$$

where Q_s^l aggregates the individual transition probabilities q_s^l of every mobile with an active request for s , i.e. $Q_s^l = \sum q_s^l$. Note that Q_s^l expresses local IO popularity with respect to (w.r.t.) the cell l , which as shown in Sec. 4 can capture temporal locality or user requests. Note that when no mobile currently being served with data from s remains connected to l as well as when there are no other proactive multicast mobile requests for s (i.e., when $Q_s^l = 0$), the allocated part of the backhaul capacity is freed and multicast stops at l . Evidently, the decision rule in (11) provides a fully decentralized *Mobility-based Proactive Multicast (MPM)* model approach for taking autonomous actions for each IO s in each cell l . Adjusting the congestion price with (10) causes decisions to make an efficient use of available backhaul capacity and to achieve the optimization target set with (8): when the backhaul is i) underutilized, the price decreases and allows for more IOs to be proactively multicast;

however, if ii) the multicast demand is greater than the backhaul capacity, then the price increases and causes to reduce the number of IOs being multicast. Last, for the case that the price is such that iii) the proactive multicast demand equals the backhaul capacity, then the proactively multicast IOs correspond to the highest values of $Q_s^l(D_{miss} - D_{hit})$, which achieves the optimization target of (8).

Regarding who takes the decision, one option is for l to use the cache congestion price p_l and to apply the decision rule (11). Note that D_{hit} as well as D_{miss} can be easily computed based on a local history from past mobile attachments to l . Alternatively, the multicast decision can be taken at the mobile or some proxy on its behalf. In this case, the mobile needs to learn the current price p_l from all $l \in L$ as well as D_{hit}, D_{miss} .

3.2 Model extensions

We extend our basic MPM model to integrate *replacement* decisions for backhaul reservation, hence yielding a *Mobility-based Proactive Multicast model with Replacements (MPM-R)*. The goal of the adapted extensions is to improve the model's effort for capturing temporal locality via updating proactive decisions with each newly incoming (resp. leaving) request. The former have an impact on Q_s^l , thus MPM-R makes an effort to keep the capacity in the backhaul proactively allocated only for the IOs that more likely to be served to multiple users, based on fresh information about mobile requests.

Adapting replacements to our model changes the process of proactive decisions in twoways: (i) requests are not immediately satisfied nor rejected with a decision rule like (11) and (ii) actions are *not* subject to a congestion price for the requested capacity part of backhaul of the cell. Instead, an object s is immediately multicast if its implied bandwidth need o_s fits in the unallocated backhaul capacity; otherwise MPM-R explores the possibility of evicting one or more IOs e that are currently being multicast. To decide which IOs to evict, MPM-R follows a procedure according to which IOs e with size o_e are polled for eviction in order of increasing $G(e)/o_e$ until there is enough free capacity in the congested backhaul of the cell for s to fit and provided that

$$\sum_{\forall e} G(e) / \sum o_e < G(s) / o_s, \quad (12)$$

$$G(e) = Q_s^l \cdot (D_{miss} - D_{hit}). \quad (13)$$

If evicting the next e violates (12) during this process, then the multicast request for s is dismissed. The purpose of (12) is to optimize the *total gain per utilized backhaul capacity unit* as a heuristic for tackling the knapsack combinatorial optimisation problem; *i.e.* maximising the total gain of the multicast IOs given their different individual gain values $G(s)$, different capacity reservation demand o_s and the limited capacity of the backhaul link. Evidently, (12) can be omitted if all requests refer to the same $o_s \equiv o$.

3.3 Model computational complexity

The basic MPM model is very lightweight, implying only an $O(1)$ computational complexity per multicast request decision. On the contrary, capacity reservation replacements add a significant burden on proactive actions, as the MPM-R model needs to maintain a *gain-based ordering* for all its pending multicast requests. This can be commonly done with a *binary heap*. The originally proposed heapsort insertion algorithm implies (in worst case) an $O(n \cdot \log(n))$

complexity cost to maintain the ordering after taking n multicast decisions. Even if a faster insertion method [9] is used, it will still require $(2 + o(1))n$ actions. Though significantly lowered with [9], MPM-R decisions imply an $O(n)$ complexity that is worse than $O(1)$ in the case of MPM.

4 EVALUATION

This section presents a preliminary performance evaluation with the use of a custom Java simulator¹.

4.1 Simulation setup

We use a realistic car traffic mobility trace and a corresponding cell distribution trace for the greater urban area of the city of Cologne². Due to practical computational resource restrictions, we use six randomly sampled regions of the trace, each of which corresponds to $2.5 \text{ km} \times 2.5 \text{ km}$ sub-areas. The sampled areas are used in two stages: first, for extracting the semi-Markov mobility prediction model and, second, for running the mobile simulations for evaluating the performance of the proactive multicast decisions with MDM and MDM-R.

User demand for proactive multicast: The number of mobiles during a simulation run depends on the sampled area and the registered time in the mobility trace. We adapt a series of five synthetic traces of video file requests produced with the GlobeTraff [12] workload generator. Each trace corresponds to a catalog of 440 videos and requests follow a Zipfian distribution with parameter $s = 0.9$. Each user issues ten requests at any time. Such a demand model can represent modern mobile applications like Facebook, which auto-play multiple videos while users scroll down their "timeline". Also, recall that mobiles correspond to cars (or buses), *i.e.* requests can be assumed to originate from multiple mobile devices in each traced vehicle.

Available multicast resources: Regarding the total "supply" of multicast resources that are subject to a restricted backhaul capacity, the number of cells in each simulation run depends also on the assumed sampled area. Since the original trace contains information only about the location of the cells, we assume microcells with a coverage range of approximately 500 m and a backhaul link capacity that is equal to 20 Mbps.

Gain metric (90% upon multicast hits): We assume that the delay D_{hit} (resp. D_{miss}) experienced for obtaining an IO is the same for all mobiles in case of a multicast hit (resp. miss), *i.e.* when the requested IO is (resp. is *not*) proactively multicast by the destination cell. Upon multicast hits the implied *delay cost gain* is equal to 90% relative to the case of a multicast miss. This corresponds to the delay induced by the submission of a user's IO subscriptions and the in-network resolution & initial path setup process before the actual data transfer starts. The remaining 10% of delay cost corresponds to the MAC-layer connection, which is orthogonal to our problem and solution.

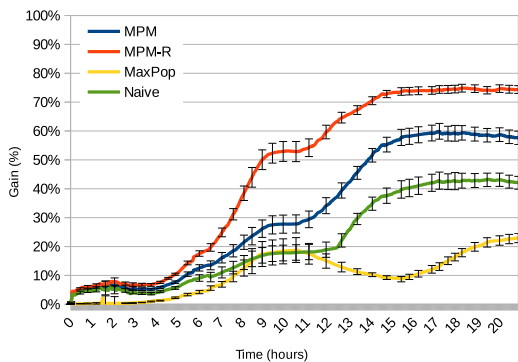


Figure 1: Transient performance gains over a period of approximately 21 hours.

4.2 Results

Fig. 1 presents transient performance gains along with 95% confidence intervals over a period of approximately 21 hours taken from the mobility trace. Apart from MPM and MPM-R, the graph includes the gains of two benchmark models: (i) MaxPop and (ii) Naive. *MaxPop* multicasts only the *topmost popular* IOs that can fit the backhaul capacity of cells based on their long-term requests frequency, while *Naive* blindly preallocates the available cell backhaul capacity by *neglecting congestion* conditions and users' *mobility* information.

Notice the fluctuations in all gain results. They are due to the dynamic demand conditions which change with time on the X axis of the graph. When traffic increases (resp., decreases) during the simulation, so does demand. A higher (resp., lower) demand causes MaxPop gains to increase (resp., decrease) and at the same time MPM and (particularly) Naive gains to stabilize (resp., increase) up until demand reduces (resp., increases) again. This is because with higher (resp., lower) traffic/demand, the chances of requesting a top-most popular IO are higher (resp., lower). Still, the performance of *MPM* remains higher than the one of *MaxPop*. This is because MaxPop decisions can *not capture temporal locality*, unlike our model which exploits up-to-date information on users' requests and mobility conditions via the recorded cell residence times used as input to the underlying semi-Markov mobility prediction model. As expected, the graph shows a better performance of MPM compared to Naive as well. Mobility prediction and congestion pricing help to utilize the congested backhaul link more efficiently and, hence, to achieve greater gains with proactive multicast.

Finally, the results show an improved performance of our model with capacity reservation replacements, i.e. MPM-R reaches closer to an optimal allocation of the congested backhaul resources than the basic MPM model. This is because MPM-R continuously adapts its decisions to the dynamic demand conditions reflected on Q_s^t in (13). However, it is also important to note that *MPM* has a *relatively good performance compared to MPM-R*, as MPM is significantly more lightweight than MPM-R (see 3.3).

5 RELATED WORK

Proactive solutions have been previously proposed for the purpose of fast handovers [14] and for vehicular Wi-Fi access [8]. The work of [16] proposes a proactive model for selective neighbour caching that reduces the handover delay in wireless LANs. The underlying motivation in the above works lies upon the fact that at most 3 or 4 APs are possible handover targets, even when the total number of neighbouring APs is large. In addition, the approach of [5] improves QoS support during handovers in Wi-Fi and cellular networks based on mobility information to undertake proactive cache actions that reduce *handover delays*. Applying proactive actions similarly to the aforementioned, only based on multicast rather than caching, is also possible for reducing resolution & path setup (resp. connection setup) delay in ICN (resp. over TCP/IP): for instance, the objective in [16] is to proactively send a mobile's context to neighbouring APs to *reduce association delay*.

Finally, many recent [3, 11, 21] efforts focus on improving the service of *popular* IOs with proactive solutions designed for heterogeneous wireless networking environments composed of femto or pico cells and Wi-Fi hotspots as a remedy for *backhaul bottlenecks*. Besides these popularity-based solutions, there are proposals that focus on *mobility* prediction [17, 19] or combine mobility prediction to IO *popularity* [18, 21]. The most interesting conclusion out of these efforts is the importance of mobility information for capturing temporal locality. As it turns out, mobility information can yield better performance results compared to mere IO popularity information in several scenarios w.r.t. certain demand conditions or IO catalog characteristics like popularity skewness, catalog and individual IO size, etc. Particularly in urban environments, the works of [18, 21] stress the significance of up-to-date mobility information on the accuracy of proactive actions due to the higher road network complexity, traffic congestion and the variety of mobility habits and routes.

6 CONCLUSIONS

This paper presents a Mobility-based Proactive Multicast model for cellular networks which uses a semi-Markov mobility prediction scheme. The solution tries to reduce the delay experienced by mobile users when they attach to a new cell. Within an ICN context, the model is designed with Pub/Sub-implied delay in mind, but it can be also applied over regular TCP/IP network architectures to reduce the time needed for establishing a new connection between a requesting mobile and a remote IO server. Our preliminary performance results based on a realistic mobility scenario indicate that gains from mobility-based proactive multicast actions are relatively higher compared to the cases of IO popularity-based and blind proactive multicast. Our model decisions use fresh information on user demand and mobility, hence allowing to adapt multicast decisions to changes in temporal locality. This outcome largely coincides with the conclusions of work [18, 21] in the literature on proactive caching models, which relate the accuracy of proactive actions in cellular networks to the latest mobility information in urban environments.

¹<https://github.com/xvasilakos/EPMS-Sim>

²Available here: <http://kolntrace.project.citi-lab.fr/>

REFERENCES

- [1] Haitham Abu-Ghazaleh and Attahiru Sule Alfa. 2010. Application of Mobility Prediction in Wireless Networks Using Markov Renewal Theory. *IEEE Trans. Vehicular Technology* 59, 2 (2010).
- [2] Suman Banerjee, Bobby Bhattacharjee, and Christopher Kommareddy. USA, August, 2002. Scalable application layer multicast. In *Proc. of the ACM SIGCOMM Conf. on Applications, Technologies, Architectures, and Protocols for Computer Comm.*
- [3] Ejder Bastug, Mehdi Bennis, and M erouane Debbah. 2014. Living on the edge: The role of proactive caching in 5G wireless networks. *IEEE Comm. Magazine* 52, 8 (2014).
- [4] Ioana Burcea, Hans-Arno Jacobsen, Eyal de Lara, Vinod Muthusamy, and Milenko Petrovic. USA, January, 2004. Disconnected Operation in Publish/Subscribe Middleware. In *IEEE Int. Conf. on Mobile Data Management*.
- [5] M.-H. Chiu and M. A. Bassiouni. 2000. Predictive Schemes for Handoff Prioritization in Cellular Networks Based on Mobile Positioning. *IEEE Journal on Selected Areas in Comm.* 18, 3 (March 2000).
- [6] R emi Choquet, Arnaud B echet, and Yann Gu edon. 2014. Applications of hidden hybrid Markov/semi-Markov models: from stopover duration to breeding success dynamics. *Ecology and Evolution* 4, 6 (2014).
- [7] Yang-hua Chu, Sanjay G Rao, and Hui Zhang. 2000. A case for end system multicast. In *ACM SIGMETRICS Performance Evaluation Review*, Vol. 28. ACM.
- [8] Pralhad Deshpande, Anand Kashyap, Chul Sung, and Samir R Das. Poland, June, 2009. Predictive methods for improved vehicular Wi-Fi access. In *Proc. of the int. conf. on Mobile systems, applications, and services*.
- [9] Robert W. Floyd. 1964. Algorithm 245: Treesort. *Comm. ACM* 7, 12 (1964).
- [10] Abdulbaset Gaddah and Thomas Kunz. 2010. Extending mobility to publish/subscribe systems using a pro-active caching approach. *Mobile Information Systems* 6, 4 (2010).
- [11] Negin Golrezaei, Karthikeyan Shanmugam, Alexandros G. Dimakis, Andreas F. Molisch, and Giuseppe Caire. USA, March, 2012. FemtoCaching: Wireless video content delivery through distributed caching helpers. In *Proc. of the IEEE INFOCOM*.
- [12] Konstantinos V. Katsaros, George Xylomenos, and George C. Polyzos. Turkey, May, 2012. GlobeTraff: A Traffic Workload Generator for the Performance Evaluation of Future Internet Architectures. In *Int. Conf. on New Technologies, Mobility and Security*.
- [13] Jong-Kwon Lee and Jennifer C Hou. Italy, May, 2006. Modeling steady-state and transient behaviors of user mobility: formulation, analysis, and application. In *Proc. of the ACM int. symposium on Mobile ad hoc networking and computing*.
- [14] Arunesh Mishra, Min-ho Shin, and WA Arbaush. Hong Kong, March, 2004. Context caching using neighbor graphs for fast handoffs in a wireless network. In *Proc. of the Annual Joint Conf. of the IEEE Computer and Comm. Societies, INFOCOM*.
- [15] Apollinaire Nadembega, Abdelhakim Hafid, and Tarik Taleb. 2015. A Destination and Mobility Path Prediction Scheme for Mobile Networks. *IEEE Trans. Vehicular Technology* 64, 6 (2015).
- [16] S. Pack, H. Jung, T. Kwon, and Y. Choi. 2005. SNC: A Selective Neighbor Caching Scheme for Fast Handoff in IEEE 802.11 Wireless Networks. *ACM Mobile Computing and Comm. Review* 9, 4 (October 2005).
- [17] Vasilios A. Siris, Xenofon Vasilakos, and George C. Polyzos. Australia, June, 2014. Efficient proactive caching for supporting seamless mobility. In *Proc. of IEEE Int. Symposium on a World of Wireless, Mobile and Multimedia Networks*.
- [18] Xenofon Vasilakos, Vasilios A. Siris, and George C. Polyzos. 2016. Addressing niche demand based on joint mobility prediction and content popularity caching. *Computer Networks* 110 (2016).
- [19] Xenofon Vasilakos, Vasilios A. Siris, George C. Polyzos, and Marios Pomonis. Finland, August, 2012. Proactive selective neighbor caching for enhancing mobility support in information-centric networks. In *ACM Proc. of the ICN Workshop*.
- [20] George Xylomenos, Christopher N. Ververidis, Vasilios A. Siris, Nikos Fotiou, Christos Tsilopoulos, Xenofon Vasilakos, Konstantinos V. Katsaros, and George C. Polyzos. 2014. A Survey of Information-Centric Networking Research. *IEEE Comm. Surveys and Tutorials* 16, 2 (2014).
- [21] Feixiong Zhang, Chenren Xu, Yanyong Zhang, K. K. Ramakrishnan, Shreyasee Mukherjee, Roy D. Yates, and Thu D. Nguyen. USA, June, 2015. EdgeBuffer: Caching and prefetching content at the edge in the MobilityFirst future Internet architecture. In *IEEE Int. Symposium on A World of Wireless, Mobile and Multimedia Networks*.
- [22] Ben Yanbin Zhao, John Kubiatowicz, Anthony D Joseph, et al. 2001. Tapestry: An infrastructure for fault-tolerant wide-area location and routing. (2001).