Wi-Stitch: Content Delivery in Converged Edge Networks

Aravindh Raman   Nishanth Sastry
King's college London
firstname.lastname@kcl.ac.uk

Arjuna Sathiaseelan
University of Cambridge
as2330@cam.ac.uk

Jigna Chandria   Andrew Secker
BBC R&D
firstname.lastname@rd.bbc.co.uk

ABSTRACT

Wi-Fi, the most commonly used access technology at the very edge, supports download speeds that are orders of magnitude faster than the average home broadband or cellular data connection. Furthermore, it is extremely common for users to be within reach of their neighbours' Wi-Fi access points. Given the skewed nature of interest in content items, it is likely that some of these neighbours are interested in the same items as the users. We sketch the design of Wi-Stitch, an architecture that exploits these observations to construct a highly efficient content sharing infrastructure at the very edge and show through analysis of a real workload that it can deliver substantial (up to 70%) savings in network traffic. The Wi-Stitch approach can be used both by clients of fixed-line broadband, as well as mobile devices obtaining indoors access in converged networks.

Keywords

content sharing, edge cooperation, wifi offloading

CCS Concepts

•Networks → Network performance evaluation; Network architectures; *Mesh networks;*

1. INTRODUCTION

Designs for future networks such as 5G are driven by expansionist calls for more capacity [6]. Among other factors, these calls are driven by numerous statistics about the growth of video consumption [22], such as the Cisco Visual Networking Index which forecasts a drastic 82% increase in cellular video streaming. Although several new technologies and novel air interfaces such as mmWave and massive MIMO are being developed to meet these demands in 5G networks, many of these do not work well indoors. This poses a problem because over 80% of of content consumption is predicted to happen indoors [1].

It is generally acknowledged [6, 7] that indoor capacity demands will need to be met using extremely small indoor cells, coupled with offloading to existing fixed-line broadband connections, or investing in expensive new backhaul using fibre-to-the-home. This combination of fixed-line and cellular infrastructure, known as "fixed-mobile convergence",

is expected to be a key of 5G [25, 13, 10]. Indeed, ETSI renamed the MEC working group recently as *Multi-access* Edge Computing, specifically to "address fixed access implementation of the MEC Server (especially WiFi)". (personal communication, 9 Mar, 2017).

Unfortunately, fixed-line networks also have a heavy burden of video traffic – 70% of North American traffic is streaming video and audio, and over 35% of it comes just from one service: Netflix[1]. In the UK, BBC's TV streaming application BBC iPlayer is one of the largest applications on the nation's networks, used by nearly one in three adults, according to the UK communications regulator Ofcom[2]. Such huge loads have led to pitched and highly confrontational battles between content providers such as Netflix and fixed-line ISPs such as Comcast[3]. Although some of these disputes are being resolved with bilateral deals (e.g., Netflix and Comcast[4]) and/or in-network caches (e.g., BBC with major UK ISPs[5]), it takes time to arrange such solutions, and they have to be decided on a case-by-case basis. Furthermore, this additional traffic requires capacity provisioning and may involve expensive commissioning of additional resources, such as fibre to the home, or close to the edge.

Therefore, this paper calls for a complementary approach that makes better use of resources *at the edge*, rather than just offloading mobile edge traffic to fixed-line broadband, or investing in network upgrades. We argue that the capacity needs of video content in converged networks can be gracefully handled by leveraging two fundamental aspects of how consumer end-points are distributed, and how they consume content: First, users are typically organised into relatively dense clusters, in cities and villages (c.f. Fig. 1). This has led to the so-called "chaotic" deployment of WiFi [3], where end users are often within range of a multitude of their neighbours' access points. Thus, *by sharing caches with neighbours, we can stitch together a content delivery network at the edge.* Second, anecdotally, as well as through several studies (e.g., [16]), it is known that interest in content items is highly skewed towards a few extremely popular items. Thus it is entirely possible that the items requested by a user have already been requested by one of her neighbours. Thus, *the caches can be passively populated by the first user*

---

[1] https://www.sandvine.com/pr/2015/12/7/sandvine-over-70-of-north-american-traffic-is-now-streaming-video-and-audio.html
[2] https://www.ofcom.org.uk/__data/assets/pdf_file/0024/26826/cmr_uk_2016.pdf
[3] https://www.wired.com/2014/02/comcast-netflix/
[4] http://www.tubefilter.com/2016/07/06/netflix-comcast-x1-deal/
[5] https://gigaom.com/2007/12/23/419-bbc-to-cache-iplayer-downloads-with-isps-could-soothe-net-neutrality-fe/

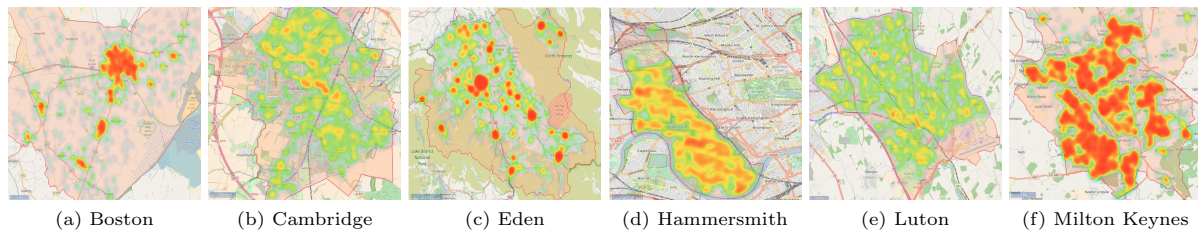| (a) Boston | (b) Cambridge | (c) Eden | (d) Hammersmith | (e) Luton | (f) Milton Keynes |

Figure 1: Heatmap showing the dispersion of BT home WiFi across 6 districts, indicating potential density of caches

who watches that item, and shared with other users who are within wireless range of the first user.

Based on these insights, we present *Wi-Stitch*, an architecture for content delivery in the converged mobile edge. Wi-Stitch caches content at the very edge of the network, *in users' homes* and uses high bandwidth 5G edge technologies such as mmWave [25] or traditional Wi-Fi mesh networks [2] to share content efficiently between neighbours. Wi-Stitch can be operated either by the Fixed-line Internet Service Provider (ISP), the Mobile Network Operator (MNO) that has provisioned indoor small cells, using the Fixed-line ISP as backhaul, or directly by the content provider (CP), without the involvement of either the MNO or the Fixed-line ISP. In each case the entity operating Wi-Stitch deploys and manages caches on their customers' homes, and redirects content requests to neighbours' caches as appropriate. We envision that the cache would be attached to, say, the ISP's modem in the user's home, the small cell station operated by the MNO, or a media streaming device operated by the CP (some, such as versions of Google Chromecast and Apple TV, already come with attached storage, but need to be interconnected with neighbours).

The central contribution of this paper is to evaluate the potential gains possible through our Wi-Stitch architecture using two large real-world datasets: (1) WiGLE[6] that gives us the geographical locations of the caches for six administrative districts in the UK of varying population densities, allowing us to estimate how many cache instances will be within Wi-Fi (or mmWave) range; (2) workload of content consumption records of all 429K users from these locations who watched BBC TV shows on its over-the-top on-demand TV streaming service, BBC iPlayer. Our evaluations show that the Wi-Stitch architecture can deliver substantial (up to 70%) savings in network traffic.

2. RELATED WORK

Given the exponential growth of video and other rich media traffic, a number of proposals have been made to mitigate their impact, using caching. Caching plays a central role in the design of content delivery networks [18], which underpin many of today's video delivery platforms, as well as proposals to offload mobile data by storing temporarily until required by the mobile device [17, 8]. Recent approaches to edge caching and offloading using peer assistance is surveyed in Anjum *et al.* [4]. Caching is also expected to play a large role in future network proposals such as 5G, which have stringent requirements on latency, bandwidth etc, that can only be met by ubiquitous caching [28].

While most approaches tend to apply caching in the context of existing Internet infrastructure, the notion of content-

centric or Information-centric networking (ICN) [11] is considered to be a turning point towards content caching models in future networks, introducing a new network architecture based on ubiquitous caching, and incorporating a naming scheme that makes it easier to cache, avoiding content duplications[26]. Since this Van Jacobson's call-to-arms, numerous ICN architectures have been proposed [29]. Whereas most such ICN proposals talk about caching *in the network*, this paper considers the novel perspective of caching at the very edge of the network, which may be easier to deploy.

This paper follows a line of work looking at traffic savings for BBC content accesses. For instance, [16] looked at factors that affect nationwide take up of the BBC iPlayer streaming application. [15] uses P2P swarms within each ISP to offload traffic from the content provider's server (but not the ISP). [14] preloads content on mobile phones thereby offloading traffic from cellular networks (and minimising the user's consumption of mobile data), but adds traffic to the user's traditional broadband ISP. The closest work, by Nencioni *et al.* [20] uses set-top boxes to speculatively record content for future access, and completely offloads requests for such content from the network. This paper adds to this idea by introducing the notion of sharing between such devices. Moreover, to our knowledge we are the first to analyse the performance of caching strategies across different districts of varying population densities.

3. THE WI-STITCH APPROACH

In this section, we discuss the details of how Wi-Stitch enables a distributed content delivery network at the very edge of the network, through caches in users' homes. We ask and answer three questions: who controls Wi-Stitch, how Wi-Stitch caches are managed and populated, and how does a client retrieve content from Wi-Stitch.

3.1 Who controls Wi-Stitch, and how?

Wi-Stitch may be operated by several kinds of entities: A fixed-line ISP could offer Wi-Stitch as a value added service, for its own customers using streaming video services, as well as for mobile customers of MNOs who are using its infrastructure as bnohaul in small cells. In this case, the cache can be attached to ISP-owned Wi-Fi routers or DSL/cable modems. Alternately, MNOs can deploy Wi-Stitch for their customers, attaching the caches to small cell base stations. Similarly, some Content Providers may have a hardware footprint in the user's home and these typically come equipped with storage (e.g., X-Box and other gaming stations, Apple TV and certain versions of Chromecast) and can deploy Wi-Stitch as well. However, they would need to interconnect with each other, either by deploying a parallel Wi-Fi mesh network, or by reusing the ISP's Wi-Fi router.

In each case, the goal of the operating entity is to present

---

[6]www.wigle.net

Table 1: Content access to BBC iPlayer from British Telecom customers in various locations

District	Area(sq.km)	Pop. density persons/sq. km	#IP address	#Content requests	Mean Content request/IP
Hammersmith and Fulham (HF)	16.40	11 213	19K	630K	33.15
Luton	43.35	4 993	41K	567K	13.82
Cambridge	40.70	3 193	41K	709K	17.29
Milton Keynes (MK)	308.63	847	71K	781K	11.0
Boston	364.90	183	15K	142K	9.46
Eden	2142.00	25	5K	90K	18.0

a unified view of the distributed caches as a single virtualised cache. This may be done using several approaches: One possibility is to use Information-Centric approaches (e.g., [27, 12]) to route users' requests for content to the appropriate caches. A second possibility is to use HTTP- or DNS-level redirections similar to current CDNs [23]. A third possibility is using currently popular SDN and NFV technologies to implement the caches as a virtualised network function [9].

Wi-Stitch is agnostic as to which entity deploys it, or which technology is used to realise the connectivity. Rather, the suitability of Wi-Stitch and the savings realised through this approach depends strongly on how well the caches between neighbours can be shared, which we evaluate in §4.2.

3.2 How are caches managed and populated?

Wi-Stitch is based on the insight that human settlements are clumped together, with a number of homes often located within close proximity of other homes. Thus, Wi-Stitch endeavours to stitch together a number of caches in neighbouring homes into a "cell" which is managed as a single virtualised cache such that the content stored in every home is made available to every other home within the cell. This could be enabled by connecting such caches using a wireless mesh network [2]. For simplicity, we consider cells which are roughly 200m in diameter, which is the typical range for new technologies such as mmWave [25], and also well within the range for Wi-Fi-based mesh networks [5].

In this paper, we consider a straightforward approach where Wi-Stitch caches in user homes are reactively populated when the user watches a content item. Thus, gains are dependent on the overlap between neighbours' content viewing patterns. In certain cases, it may be possible to proactively populate content. For instance, BBC programmes are also broadcast over the air, and this has been exploited for proactive caching [21]. However, we leave the consideration of proactive caches to future work.

3.3 How are cached items retrieved by clients?

Wi-Stitch maintains a mapping of which content item has been delivered to which nodes, and redirects content requests to the appropriate cache. This could be done transparently to the client (e.g., using ICN primitives if adopting an ICN solution, using SDN rewriting at the home router if using SDN (e.g. [2]), or with the content provider using HTTP or DNS redirection). Alternately, the client could actively be involved, and connect to the appropriate access point to directly access the content from there.

4. FEASIBILITY OF WI-STITCH

The benefits of Wi-Stitch depend on two key factors: whether

Table 2: Most common BT SSID prefixes as found in WiGLE as of mid-2016 (does not include the SSIDs customised by the user)

SSID	Count
Auto-BTWiFi	18K
BTWi-fi	42K
BTOpenzone-B	100K
BTOpenzone-H	130K
BT-Fon	231K
BTOpenzone	250K
BTWIFI	570K
BTWifi-X	1M
BTWiFi-with-FON	1.6M

there are sufficient number of neighbouring users to share with, and whether there is sufficient overlap in content consumption patterns, to make use of these neighbours. We explore these using two large-scale datasets. The first, from WiGLE, is a wardriving dataset that gives the geographic location of Wi-Fi access points, allowing us to closely approximate the locations of the caches and their accessibility using Wi-Fi or similar technologies. The second is a dataset of content consumption in a TV streaming application, BBC iPlayer, that is used by nearly 50% of the UK population.

4.1 Datasets

As explained before, Wi-Stitch sharing can happen across all neighbours within range of each other, or just the customers of a single ISP, if Wi-Stitch is being operated by the ISP. Since the latter case results in lesser density of access points, the savings are obviously going to be smaller. Therefore, as a case study, we focus mainly on the sharing possible among customers of one major nationwide ISP, British Telecom (BT). To understand the spectrum of sharing opportunities, we look at six administrative districts[7] of diverse population densities, ranging from Hammersmith and Fulham in London, one of the 10 most densely populated areas in the country with a population density of more than 10,000 people per square kilometre, to Eden, which has the least population density in all of the UK (Table 1). We proceed our study making use of two large datasets:

4.1.1 WiGLE

WiGLE (Wireless Geographic Logging Engine) is an open-sourced platform that uses crowdsourcing to collect the locations of wireless Access Points (APs) across the globe. As

---

[7]https://en.wikipedia.org/wiki/List_of_English_districts_by_population_density

Table 3: Details of the BBC iPlayer dataset

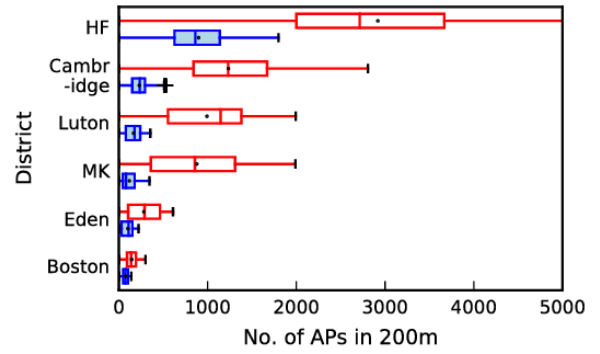|	July
Number of Users	25M
Number of IP addresses	17M
Number of Sessions	215M

Figure 2: Number of users within cells of 200m diameter, considering only BT customers (shaded blue) or all users (red)

of mid-2016 around 250M were reported and $\approx$ 7M from the UK. We use the SSIDs, GPS co-ordinates and timestamps of first and last detection to identify the BT-Access points in a particular location. We identify $\approx$ 4M BT access points based on the patterns of SSID strings assigned by default (e.g., "BTHub3-XWX9"). Due to historic reasons, several different string patterns are observed, with the most common types of SSIDs as shown in Table 2. Our method does not capture SSIDs which have been changed from the default by its users, and therefore can be considered as a lower bound of the sharing possible in each region.

4.1.2 iPlayer

BBC iPlayer is a widely used video streaming application in the UK and is available for both web and mobile platforms. It is a catch-up TV service that makes available for on-demand streaming most of the programmes broadcast on BBC channels across the UK. Content within iPlayer can be available for up to a month depending on licensing terms and other policies. iPlayer is one of the top video sites in the country, second only to YouTube[8]. However unlike YouTube, BBC iPlayer hosts ad-free HD content, and TV shows much longer than the average YouTube video. Our data reported the equivalent of over 40% of the UK's population accessing the iPlayer during July 2014 (Table 3). We focus on accesses from the six administrative districts as shown in Table 1 and from BT customers, a subset of Table 3. Table 1 also gives the population density, as well as the density of accesses (number of requests and number of users making the requests) from each area.

4.2 Exploring Wi-Stitch savings

To understand potential savings, we first use WiGLE to estimate the number of users who would be able to share with each other in Wi-Stitch cells across administrative districts with different population densities[9]. Then we compute possible traffic savings by simulating Wi-Stitch content sharing within cells of different sizes. Finally, we examine the bounds on this savings if cache storage is limited.

4.2.1 Distribution of Access Points

As shown in Fig. 1, users are clumped together in tightly knit clusters, even in the least densely populated areas such as Eden and Boston. Fig. 2 formalises this, considering cells of a fixed 200m diameter, and plotting the distribution of the numbers of users found in such cells in each of the administrative districts. The distributions are plotted as box plots, with a central line showing the median, the ends of the boxes showing the 25th and 75th percentile, and whiskers extending from the 5th to the 95th percentile. We show both

---

[8]http://mediatel.co.uk/newsline/2014/03/28/nielsen-data-report-february-2014/
[9]The WiGLE.net part of the dataset and scripts to reproduce Figs. 1 and 2 are made available at https://bit.ly/wi-stitch. The BBC data contains personal information and cannot be released.

the numbers of BT customers that can be found in cells (in blue), as well as customers of all ISPs (in red), which is clearly a much larger figure. The median number of BT users in a cell ranges from 75 in Boston, a coastal city in Linconshire, England, to over 860 users in Hammersmith and Fulham (HF) in London.

4.2.2 Potential bounds on savings

Next we ask how much traffic savings can be achieved given the distribution of cell sizes we observe. The districts are grouped into cells having a range of populations, and we check what savings can be obtained by simulating the following scenario: Each user who requests a content item first attempts to find it within the cell. If a copy is found, then it can be served locally, contributing to the traffic savings. If the user is the first in the cell to access that content, then it is fetched from BBC's servers and cached for later use by any other user within the cell who requests the same content item. Initially, we assume that there are no storage constraints. We relax this later on and ask how the savings are curtailed by limitations of storage.

Note that our iPlayer data is anonymised, and while location information of the accesses exist (computed by IP geolocation, down to post code or borough level resolution), this does not give the exact latitude and longitude of the access. Thus, we map the location of iPlayer requests to WiGLE access points within that location at random and use this combined dataset for the analysis. We have checked that the results reported here are robust regardless of the exact mapping, by verifying the results are consistent over 50 different random mappings.

Fig. 3 shows how the savings rise as cell size increases when there are no storage constraints, leading up to $\approx$70% savings. The figure also shows that savings evolve in a similar way across cells of various different population densities. To test why this is the case, we fitted power law distributions to the content access requests from different cells, and found that the power law exponent is similar across the regions. The distribution of power law co-efficients across cells was tightly concentrated, with most cells having a power law exponent between 2 and 3. Thus we find similar levels of concentration in accesses towards popular items. The one exception is Hammersmith and Fulham (HF), whose traffic savings stands above the other locations. This could be explained from Table 1, where we see that the number of re-

quests per user is higher for Hammersmith and Fulham than in other locations. In other words, users in this metropolitan London location are heavier users of iPlayer than in other places, leading to a larger and more diverse cache, which in turn leads to a better hit rate and more traffic savings overall.
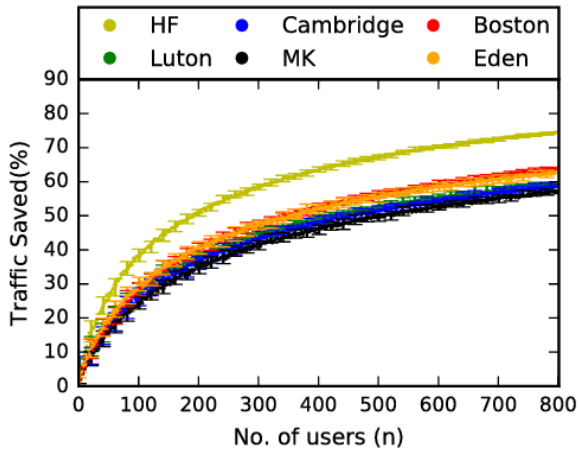

Figure 3: Traffic Savings with 95% confidence interval (CI) across the districts with 'n' people (X-axis) in a cell when the content is opportunistically cached in the local device and made available for the whole cell. To improve readability, the CI bars are shown for every n=10

4.2.3 Effect of storage limits

The savings in Fig. 3 are a result of opportunistic caching assuming an infinite store. To check the effect of storage limits, we compared various common cache replacement strategies, namely Least Recently Used (LRU), Least Frequently Used (LFU), and First in First Out (FIFO). Further, we check the performance of these traditional cache update algorithms against an oracle that can look into the future to determine which items will be accessed and how often, and thus can make the best possible cache replacements.

Fig 4 shows the performance of various cache replacement techniques. To check how the performance varies for different cell sizes, we show the traffic incurred across storage levels for a cell with population $n=120$, 240 and 360 (which covers typical cells in most areas except for Hammersmith and Fulham, where savings are higher than other places). Storage levels are varied from 0% (no cache) to 100% which is the cache size corresponding to the actual data consumed by the users within that cell in that month.

Note that there is a minimum level of traffic that cannot be avoided, since the first user in the cell always has to fetch the item from the origin server. As expected, oracle performs better than the other techniques, and bottoms out to the minimum traffic levels achievable quickly. However, it is interesting to observe that all the other policies also converge to this minimum level – i.e., given reasonable amounts of storage, it is possible to approximate the best possible cache replacement policy (oracle) and pick the "best" items to replace under a particular storage constraint, using various well-known heuristics such as LRU, LFU and FIFO.

LRU is the best performing of these policies and can approximate the Oracle even with only a small amount of stor-

age (≈20% of total data used), followed closely by FIFO. Unexpectedly, we see FIFO performing better than LFU. We conjecture that this may be due to the periodic nature of iPlayer content – as newer episodes of TV shows come online, interest in older episodes wanes, and therefore it becomes less likely that there will be a future access request for older episodes. In other words, interest in content items may naturally follow a rough FIFO order. Although an episode has been frequently accessed in the past, it may be unlikely to be accessed as much as a less frequently used but newer episode, making FIFO a better policy than LFU.

5. CONCLUSION

In this paper we propose Wi-Stitch, a mobile edge architecture for content delivery in converged networks, that stitches together caches of a "cell" of neighbouring homes into a virtualised common cache. Wi-Stitch exploits shared interests in content items across neighbours to derive traffic savings. This architecture can be used to benefit both mobile users connected to indoor or other small cells, as well as users of regular fixed-line broadband. We discuss how each participant in the infrastructure (eg., ISPs, CPs, MNOs) can use Wi-Stitch independently, on their own, or in collaboration with each other. This allows flexibility in operation, traffic savings at the backhaul and better QoS at the user end, with minimal/no additional hardware upgrades except possibly at the very edge of the network, where such upgrades are much easier and cheaper than elsewhere in the network.

We analyse the feasibility of Wi-Stitch and identify the potential savings with two large traces of content accesses and edge cache locations. Using this, we showed that 30–70% of traffic savings could be achieved and also discuss the savings under various cache replacement strategies when curtailed by storage.

Wi-Stitch exploits the current chaotic/clustered deployment of wireless infrastructure as well as locality of interest to enable collaborative cache sharing in converged edge networks. As part of future work, we plan to extend the Wi-Stitch architecture to support collaborative caching strategies using coded content caches [19] as well as developing and evaluating an information/content centric information layer for enabling better content discovery.

6. REFERENCES
[1] White paper: Cisco VNI forecast and methodology, 2015-2020. Technical report, Cisco.
[2] A. Abujoda, D. Dietrich, P. Papadimitriou, and A. Sathiaseelan. Software-defined wireless mesh networks for internet access sharing. *Computer Networks*, 93:359–372, 2015.
[3] A. Akella, G. Judd, S. Seshan, and P. Steenkiste. Self-management in chaotic wireless deployments. *Wireless Networks*, 13(6):737–755, 2007.
[4] N. Anjum, D. Karamshuk, M. Shikh-Bahaei, and N. Sastry. Survey on peer-assisted content delivery networks. *Computer Networks*, 116(C):79–95, Apr. 2017.
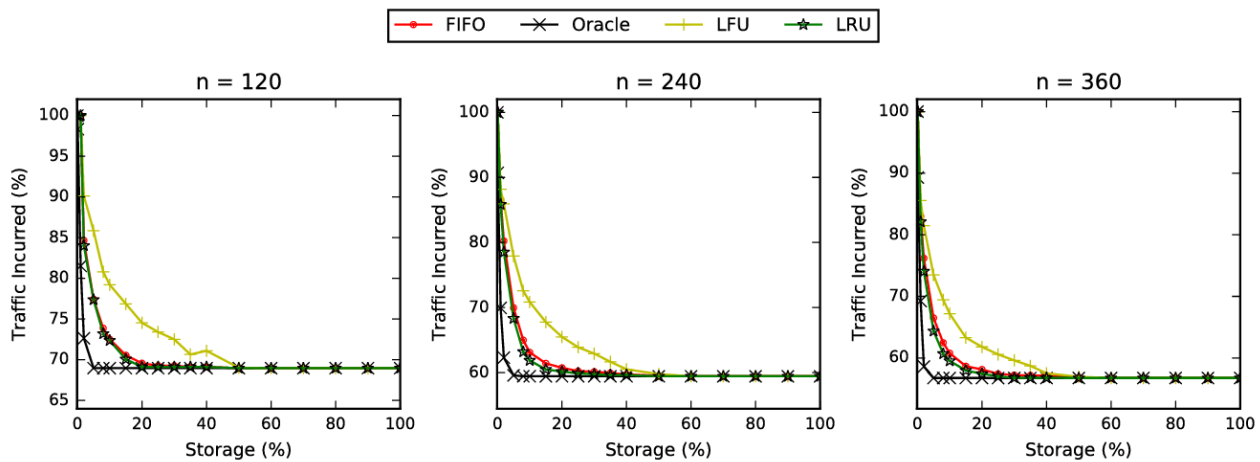
Figure 4: Traffic Incurred for several edge cache store policies, across cells of varying sizes (n=120,240 and 360)

[5] J. Bicket, D. Aguayo, S. Biswas, and R. Morris. Architecture and evaluation of an unplanned 802.11b mesh network. In *Proceedings of the 11th annual international conference on MobiCom*, pages 31–42. ACM, 2005.

[6] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski. Five disruptive technology directions for 5G. *IEEE Communications Magazine*, 52(2):74–80, 2014.

[7] K. Chandra, R. V. Prasad, and I. Nimegeers. An architectural framework for 5G indoor communications. In *Wireless Communications and Mobile Computing Conference*, pages 1144–1149. IEEE, 2015.

[8] S. Dimatteo, P. Hui, B. Han, and V. O. Li. Cellular traffic offloading through wifi networks. In *IEEE 8th International Conference on Mobile Adhoc and Sensor Systems*, 2011.

[9] P. Georgopoulos, M. Broadbent, B. Plattner, and N. Race. Cache as a service: Leveraging SDN to efficiently and transparently support video-on-demand on the last mile. In *IEEE 23rd International Conference on ICCCN*, pages 1–9, 2014.

[10] S. Gosselin, F. Moufida, T. Mamouni, J. A. Torrijos, L. Cucala, D. Breuer, E. Weis, F. Geilhardt, D. V. Hugo, E. Bogenfeld, et al. Fixed and mobile convergence: Needs and solutions. In *European Wireless 2014; 20th European Wireless Conference; Proceedings of*, pages 1–6. VDE, 2014.

[11] V. Jacobson, M. Mosko, D. Smetters, and J. Garcia-Luna-Aceves. Content-centric networking. *Whitepaper, Palo Alto Research Center*, pages 2–4, 2007.

[12] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard. Networking named content. In *Proceedings of the 5th ACM CoNEXT*, 2009.

[13] V. Jungnickel, K. Habel, M. Parker, S. Walker, C. Bock, J. F. Riera, V. Marques, and D. Levi. Software-defined open architecture for front-and backhaul in 5G mobile networks. In *Transparent Optical Networks (ICTON), 2014 16th International Conference on*, pages 1–4. IEEE, 2014.

[14] D. Karamshuk, N. Sastry, M. Al-Bassam, A. Secker, and J. Chandaria. Take-away TV: Recharging work commutes with predictive preloading of catch-up TV content. *IEEE JSAC*, 34(8):2091–2101, Aug 2016.

[15] D. Karamshuk, N. Sastry, A. Secker, and J. Chandaria. ISP-friendly peer-assisted on-demand streaming of long duration content in BBC iPlayer. In *2015 INFOCOM*, pages 289–297, April 2015.

[16] D. Karamshuk, N. Sastry, A. Secker, and J. Chandaria. On factors affecting the usage and adoption of a nation-wide TV streaming service. In *2015 INFOCOM*, April 2015.

[17] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong. Mobile data offloading: How much can wifi deliver? *IEEE/ACM TON*, 21(2):536–550, 2013.

[18] T. Leighton. Improving performance on the internet. *Communications of the ACM*, 52(2):44–51, 2009.

[19] M. A. Maddah-Ali and U. Niesen. Fundamental limits of caching. *IEEE Transactions on Information Theory*, 60(5):2856–2867, 2014.

[20] G. Nencioni, N. Sastry, J. Chandaria, and J. Crowcroft. Understanding and decreasing the network footprint of catch-up TV. In *Proceedings of the 22nd International Conference on WWW*, pages 965–976. International World Wide Web Conferences Steering Committee, 2013.

[21] G. Nencioni, N. Sastry, G. Tyson, V. Badrinarayanan, D. Karamshuk, J. Chandaria, and J. Crowcroft. SCORE: Exploiting global broadcasts to create offline personal channels for on-demand access. *IEEE/ACM TON*, 24(4):2429–2442, Aug 2016.

[22] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, et al. Scenarios for 5G mobile and wireless communications: the vision of the metis project. *IEEE Communications Magazine*, 52(5):26–35, 2014.

[23] G. Pallis and A. Vakali. Insight and perspectives for content delivery networks. *Communications of the ACM*, 49(1):101–106, 2006.

[24] A. Raman, N. Sastry, A. Sathiaseelan, J. Chandaria, and A. Secker. Wi-Stitch: Content delivery in converged edge networks. In *Proceedings of the Workshop on Mobile Edge Communications*, MECOMM '17, pages 13–18. ACM, 2017.

[25] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez. Millimeter wave mobile communications for 5G cellular: It will work! *IEEE access*, 1:335–349, 2013.

[26] D. Rossi and G. Rossini. Caching performance of content centric networks under multi-path routing (and more). *Relatório técnico, Telecom ParisTech*, 2011.

[27] D. Trossen and G. Parisis. Designing and realizing an information-centric internet. *IEEE Communications Magazine*, 50(7), 2012.

[28] X. Wang, M. Chen, T. Taleb, A. Ksentini, and V. C. M. Leung. Cache in the air: exploiting content caching and delivery techniques for 5G systems. *IEEE Communications Magazine*, 52(2):131–139, February 2014.

[29] G. Xylomenos, C. Ververidis, V. Siris, N. Fotiou, C. Tsilopoulos, X. Vasilakos, K. Katsaros, and G. Polyzos. A survey of information-centric networking research. *Communications Surveys Tutorials, IEEE*, 16(2):1024–1049, Second 2014.