Investigating the Impact of Cache Pollution Attacks in Heterogeneous Cellular Networks

Sibendu Paul[1], Anand Seetharam[2], Amitava Mukherjee[3], Mrinal Kanti Naskar[1]
[1]Department of Electronics and Telecommunications Engineering, Jadavpur University, India,
[2]Department of Computer Science, SUNY Binghamton University, USA, [3]IBM India Pvt. Limited
sibendu.paul01@gmail.com, aseethar@binghamton.edu, amitava.mukherjee@in.ibm.com, mrinaletce@gmail.com

***Abstract*—With the growth of Internet-of-Things, mobile data traffic is expected to increase exponentially. To support this rapid growth, heterogeneous cellular networks comprising of femtocells with storage capabilities along with macrocell base stations have been proposed. In this paper, we first investigate the performance impact of a simple randomized cache pollution attack, where the attacker pollutes the cache at the femtocell by requesting unpopular content. We then adopt a principled approach based on the characteristic time of a content in a cache to design an optimized attack strategy. Our experiments show that the proposed attack strategy outperforms the randomized attack with the same attack rate.**

I. Introduction

To alleviate congestion in the cellular network core resulting from the explosive growth of mobile wireless traffic, heterogeneous cellular networks (HCN) comprising of storage-enabled femtocells and macrocell base stations are being deployed. By caching popular content at these femtocells, user performance (e.g., delay, throughput) can be improved by offloading traffic from the macrocell to the femtocells.

Figure 1 shows a HCN consisting of one macrocell and one cache-enabled femtocell with multiple users connected to it. In general, multiple femtocells could be connected to a single macrocell, but we focus on a single femtocell. We assume that some of the users connected to this femtocell are attackers who aim to degrade the performance of legitimate users by polluting the cache in the femtocell by periodically requesting unpopular content.

In this paper, we first investigate the performance impact of a simple randomized cache pollution attack in HCN. We observe via our experiments that the potency of such an attack increases as the attack rate increases. A natural question that arises is, how to design a principled attack strategy that outperforms the randomized cache pollution attack? Therefore, in this paper we leverage the concept of characteristic time of a content in a cache to design an optimized attack strategy. Characteristic time of a content in a cache denotes the expected amount of time a recently accessed content will remain in the cache [1]. We observe that the proposed characteristic time attack outperforms the simple randomized cache pollution attack with the same attack rate.

Prior work [2]–[4] has studied the impact of cache pollution based denial of service (DoS) attacks in information-centric networks consisting of a large number of cache enabled nodes. The authors in [2] demonstrate that such attacks though potent in small networks, quickly lose their potency as network size increases. In contrast to prior work, we study cache pollution attacks in a HCN where there is only a single cache, and propose a new characteristic time based attack.

978-1-5090-6501-1/17/$31.00 ©2017 IEEE

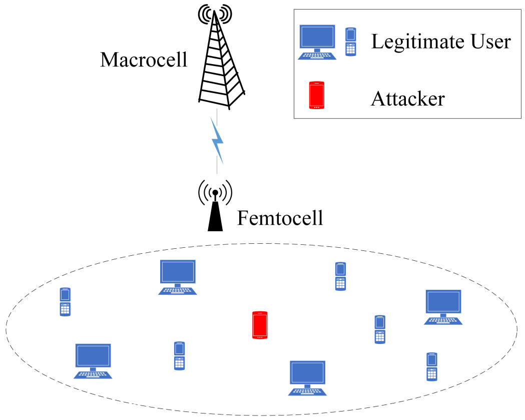


Figure 1: Typical Attack Scenario in HCN

II. Cache Pollution Attacks

Let us consider a HCN as shown in Figure 1 consisting of a cache-enabled femtocell and multiple users, some of which are attackers. We assume that the cache at the femtocell is of size C and that the content universe size is M. We assume that the content popularity varies according to a Zipfian distribution. Each legitimate user requests content at rate λ following the Independent Reference Model (IRM). The attackers leverage the long tail of the Zipfian distribution to inflict a cache pollution attack. We assume that unpopular content set consists of M' least popular pieces of content. By requesting unpopular content, the attackers aim to pollute the cache, thereby degrading the performance of legitimate users.

A. Randomized Cache Pollution Attack

In the randomized attack (with $x\%$), every time an attacker generates a request, it tosses a coin with probability $\frac{x}{100}$ to decide if the request is a malicious request or a normal one. Therefore, in expectation, the number of malicious requests is $\frac{x}{100}\lambda$ and the remaining $\frac{(1-x)}{100}\lambda$ requests are legitimate requests. If x is greater than 100%, the attack rate is greater than λ and all requests are attack requests. In case the attacker decides to place an attack request, it randomly chooses a content from the unpopular content set M'.

B. Characteristic Time Attack

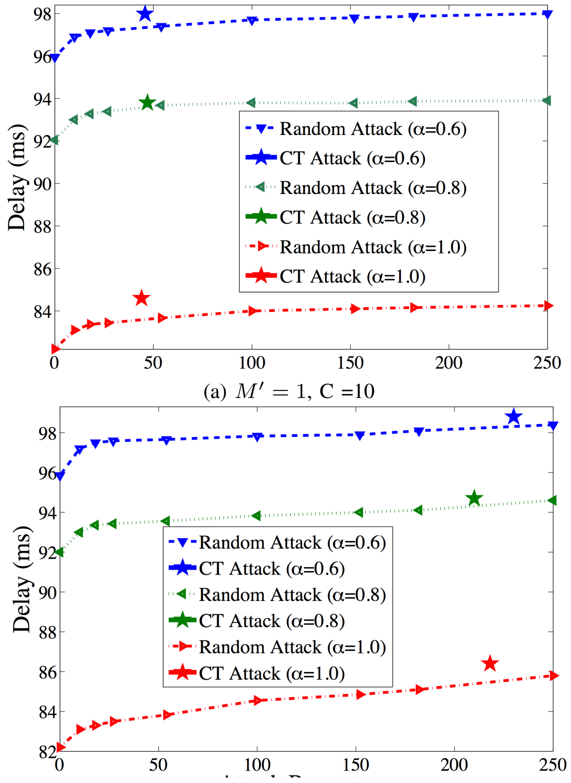
The randomized attack serves as a baseline approach for a cache pollution attack. In this section, we propose an attack strategy based on the concept of characteristic time that optimizes the attack rate and provides superior performance. Characteristic time of a content in a cache indicates the amount of time in future a content is likely to remain in the cache, provided that it has been accessed recently. At the highest level, in this attack, the attacker sends a request for each content in M' after its characteristic time has elapsed. The intuition behind this approach is to ensure that the same unpopular content is inserted back into the cache as soon as it is removed. As unpopular content is being constantly placed in the cache, the performance of legitimate users is decreased.

In order to design an effective attack strategy, we need to estimate the characteristic time accurately for each content in M'. We start with an initial guess of characteristic time based on Che's approximation [1]. We observe from our experiments that the characteristic time estimate obtained via Che's approximation is not an accurate estimate of the characteristic time for unpopular content. We then adopt a binary search algorithm to obtain a better estimate of the characteristic time for each content. We note that the parameters needed to initially estimate the characteristic time using Che's approximation can be determined by using the approach outlined in [5].

III. Performance Evaluation

In this section, we report simulation results demonstrating the performance of the randomized attack as well as the characteristic time attack. For our simulation, we assume that there are nine legitimate users and one attacker. We choose the following parameters, $\lambda = 1$, $M = 1000$, $C = 1$. When a request for a piece of content is served from the femtocell cache, it is a hit, otherwise it is a miss. The values of hit and miss delays are chosen to be 20 ms and 100 ms respectively. We assume that the cache uses the Least Recently Used content replacement policy.

We show results for $M' = 1$ and $M' = 5$ in Figure 2. The continuous lines in Figure 2 show the delay performance of the legitimate users for the randomized attack as the attack percentage increases for different values the Zipfian skewness parameter α. The discrete points in each figure show the performance of the characteristic time attack (shown as CT attack in Figure 2). The results reported are average results obtained over multiple runs of the experiment. The far left or first point on each continuous line shows the performance of the legitimate users in the absence of an attack. As we can observe from the figure, as the attack fraction increases, the effectiveness of the random attack increases. More importantly, the characteristic time attack outperforms the random attack with the same attack rate and the randomized attack requires a higher attack rate to attain the same level of delay degradation.



(a) $M' = 1$, C =10

(b) $M' = 5$, C =10

Figure 2: Latency for varying attack percentage

IV. Conclusion

In this paper, we investigated the effectiveness of cache pollution based degradation of service attack in heterogeneous cellular networks. We studied the performance of a simple randomized attack and also proposed a principled attack strategy that exploits the characteristic time for a content in a cache to improve the attack efficacy. Via simulations, we showed that the effectiveness of the randomized attack increases with attack rate. We also observed that the characteristic time attack outperforms the randomized attack at the same attack rate. As part of our future work, we plan to devise effective strategies to defend against cache pollution attacks.

References

[1] Hao Che, Ye Tung, and Zhijun Wang, "Hierarchical web caching systems: Modeling, design and experimental results", IEEE Journal on Selected Areas in Communications, 20(7):1305–1314, 2002.

[2] Jeffery Gouge, Anand Seetharam, and Swapnoneel Roy, "On the scalability and effectiveness of a cache pollution based DoS attack in information centric networks", in IEEE ICNC, 2016.

[3] Paolo Gasti, Gene Tsudik, Ersin Uzun, and Lixia Zhang, "DoS and DDoS in Named Data Networking", in IEEE ICCCN, 2013.

[4] Matthias Wählisch, Thomas C. Schmidt, and Markus Vahlenkamp, "Backscatter from the data plane: threats to stability and security in information-centric network infrastructure", in Computer Networks, Vol. 57, No.16, pp 3192–3206, 2013

[5] Mostafa Dehghan, Dennis Goeckel, Ting He, and Don Towsley, "Inferring military activity in hybrid networks through cache behavior", in IEEE MILCOM, 2013.