A Probabilistic and Rebalancing Cache Placement Strategy for ICN in MANETs

Cheng Zhang[1,2(✉)], Chunhe Xia[1,2], and Haiquan Wang[1,3]

[1] Beijing Key Laboratory of Network Technology, Beihang University, Beijing, China
zalacheng@foxmail.com
[2] School of Computer Science and Engineering, Beihang University, Beijing, China
[3] College of Software, Beihang University, Beijing, China

Abstract. Cache placement is a key component of information centric networking (ICN) for optimizing network performance. Previews works on caching placement mainly focus on fixed or semi fixed network. We propose a cache placement/update strategy in a mobile ad hoc network (MANET) environment where content consumer and provider are constantly mobile. Based on hotspot recognition and user requirement modeling, two approaches named *Pcaching* and *Cache Rebalancing* are proposed to settle the cache placement/update problem in one region and between different regions. We evaluate the performance of our approach compared with existing caching scheme and show that *Pcaching* has its advantage in both hit ratio and hop count over previews strategy.

Keywords: Cache placement · ICN · MANET

1 Introduction

Based on named data, Information-Centric Networking (ICN) has become a new way of thinking as traditional Host-Centric Internet communication paradigm lacks efficiency in data access and content distribution measures. ICN uses name-based routing to achieve location independence and enables a series of valuable features such as in-network caching, multicast and mobility support [1]. Therefore, it provides a more efficient means of communication, making it much easier for users to retrieve content. When user requests content, any node receiving the request can act as a server and respond to the request with a copy, thereby reducing the network congestion, access latency, and bandwidth consumption.

As ICN uses name-based routing, it naturally fits the application of mobile peer-to-peer communications where nodes are highly mobile, and in lack of steady network topology. In-network caching mechanism faces the challenge of how to place/distribute contents in a mobile network in order to make the network efficiently working. In this paper, we consider popular content replication and placement problem in a mobile ad hoc network (MANET), so as to optimize network performance as network nodes are constantly moving. We propose a placement/update strategy in a MANET environment

Y. Sun et al. (Eds.): ChineseCSCW 2018, CCIS 917, pp. 228–238, 2019.
https://doi.org/10.1007/978-981-13-3044-5_17

where content consumer and provider are both mobile. The model decides where or if a popular content data should be cached or updated in a MANET environment in order to maintain ICN stability and efficiency. We try to address this issue by bringing hot city geographic locations, or hotspots, into consideration, so that our model can place popular data contents according to different hotspots and user requirements. City taxi behavior and geographic features are also studied to help recognize and divide city hotspot regions because taxi moving behaviors are strongly related to their geographic features.

2 Related Works

Most content replication and placement strategy researches have been focused on fixed networks [2–4], which is natural because content caching and placement problems in Web caching and CDNs need to be addressed by taking network topology and throughput into account. However, in a MANET environment, due to dynamic network topology, high provider or consumer mobility and different throughput rate, ICN researchers have to adapt light-weighted caching placement strategies in order to reduce protocol overhead. Due to its host-to-content routing natural, ICN has two basic roles, content consumer and content provider. Paper [5] proposed a content provider mobility scheme in Named Data Networking (NDN), adding a Locator into NDN Interest package and a mapping system which maps identifier to locator. The mapping system is a DNS-like service for users to find out the newest location of content provider on the move. A content consumer mobility ICN scenario is introduced by [6] where mobile users move between different sites while requesting content from caching servers within every site through WIFI connection. A location aided Content Management Architecture for a content-centric MANET is proposed in [7], in which content consumer and provider are both mobile. By binding data to a geographic location, the approach intends to keep a content copy within pre-setup geographic boundaries based on GPS location information. Through proactively replicating the content when needed, the approach can maintain data availability within the boundaries. Different from existing work which mainly studies only one-sided mobility or uniform-distributed geo-location placement strategy, we aim to consider both consumer and provider mobility, as long as the impact of different user request patterns in real city geo-locations, to design an ICN caching update and placement strategy that can work in a highly dynamic MANET environment.

3 Model Overview and Specification

The goal of our model is to manage content placement/update strategy in a MANET environment where content consumer and provider are both mobile. The model decides where or if a popular content data should be cached or updated in a MANET environment in order to maintain ICN stability and efficiency. We try to address this issue by bringing hot city geographic locations, or hotspots, into consideration, so that our model can place popular data contents according to different hotspots and user requirements. City taxi moving behavior and geographic features are also considered in order to recognize and divide city hotspots as taxi behaviors are strongly related to geographic features. The model has three key technologies as shown in Fig. 1 to be further described below:

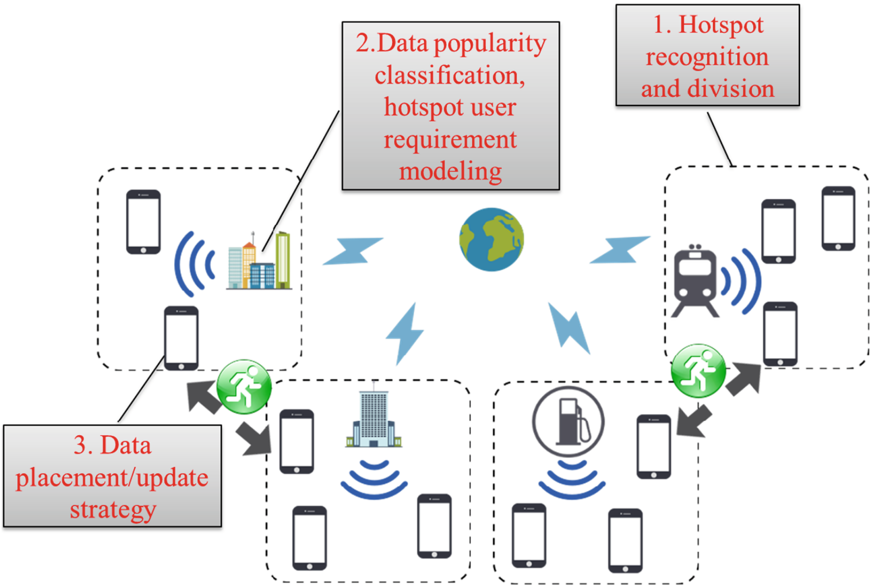

Fig. 1. Model overview and key technologies

Hotspot Recognition and Division

- A city can be divided into two kinds of regions named “load-region” and “drop-region” based on the density of passenger load or drop events extracted from real taxi traces, so we can get load regions and drop regions where the load or drop event happens frequently.

Data Popularity Classification, Hotspot User Requirement Modeling

- Existing caching research [7] tells that the content access rate follows Zipf-like distribution, so we can use this distribution to generally model content popularity.
- In our model, a History Interest Table (HIT) is introduced in order to record historical content requirements of users in a period of time. By using HIT, we can fit a content requirement distribution function (CRF) for a caching node to model user requirement probability distribution.

Data Placement/Update Strategy

- Within one region, users tend to fit with a similar data requirement model. Therefore in order to meet with this requirement, we propose a multi-factor data placement/update strategy in order to cope with HIT, battery status, caching occupation, content time-effectiveness and content popularity.
- With different CRFs, we diverse content requirements in different locations. So we propose an inter-region data update strategy to be adaptable to locations when user moves from one region to another. The strategy makes node periodically updates its own caching in order to fit in with different regions it enters.

3.1 Hotspot Recognition and Division

Through preview researches [8], we found that the taxi moving behaviors or its geographic features are significantly related to the taxi occupation statuses. Instead of dividing the area simply into equally regions, we make the area into two kinds of regions named "load-region" and "drop-region" based on the density of passenger load or drop events extracted from real taxi traces. When a passenger steps into a taxi, the current location is recorded as load-region regions. The destination region, where the passenger steps out, is recorded in the set of drop-region regions. We study the relationship between load-regions and drop-regions, arguing that drop events can better describe passenger moving behaviors, and using drop-regions to form different city hotspots in which data contents are required in respect to user requirement model.

Hotspot Recognition. With real taxi trace data of Beijing city, we analyze the distributions of load-events and drop-events within one week. Through clustering algorithms, we manage to get clustering ranking results, or top hotspots, for both load-events and drop-events, as shown in Fig. 2:

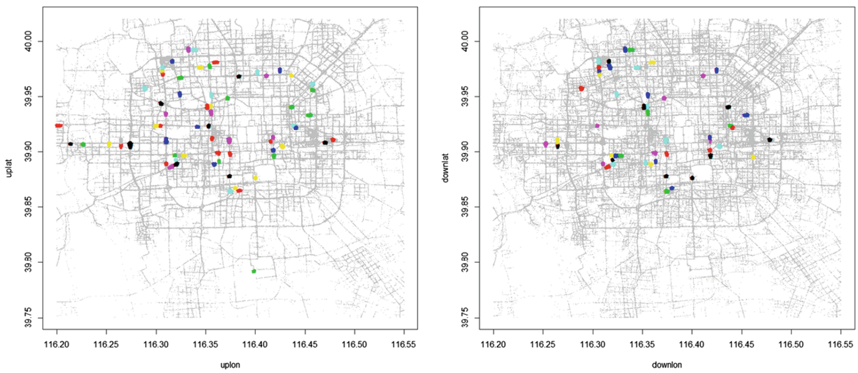

Fig. 2. Hotspot recognition for load and drop events

Hotspot Division. We argue that drop-events are more suitable data to describe people's interest and intention, so in order to simplify calculation we choose top 20 drop-event clusters from Fig. 2 to roughly recognize hotspots. From Fig. 2 we can see that these hotspots include famous scenic spots, large business areas, railway stations and subway stations. After getting clustered hotspots, we have to formally divide those hotspots from the map before we can use them as calculable regions. A Graham's scan method [9] is adopted to find the convex hull of each cluster, and for each convex hull, we calculate its Minimum Bounding Rectangle(MBR) [10] in order to make our division more accurate. With MBR, which is the minimum bounding rectangle for hotspots, we finally get an accurate hotspot region division for MANET.

3.2 Data Popularity Classification and Hotspot User Requirement Modeling

Data Popularity Classification. Inspired by existing researches from Web caching and CDNs, we consider the data access rate by user follows Zipf-like distribution, so we can model content popularity with it. In this distribution, the accessing probability of the r-th ($1 \leq r \leq N_c$) ranking data from a data-set can be denoted as follows:

$$P_r = \frac{1/i^{\alpha}}{\sum_{n=1}^{N_c} \frac{1}{n^{\alpha}}}, r = 1, 2, \ldots, N_c \tag{1}$$

where $\alpha (0 \leq \alpha \leq 1)$ denotes the value of the exponential parameter of the distribution. When $\alpha = 1$, the distribution follows Zipf's law; and when $\alpha = 0$, it follows Uniform Distribution. N_c is the total number of different contents.

Hotspot User Requirement Modeling. In traditional Content-Centric Networking (CCN) solution, which is a realization of ICN, Forwarding Information Base (FIB), Pending Interest Table (PIT) and Content Store (CS) are three key elements to a node in order to participate in the network [11]. FIB contains content advertisements and their routing information in order to forward Interest packages to the content providers that might hold the matching Data. PIT keeps track of Interest packages so the matching Data returned from content providers can follow the reversed path to the data consumers. CS is a storage cache for the nodes in order to temporarily cache the content Data with respect to different caching schemes. When a data requirement is met in a node, the corresponding entry in Pending Interest Table (PIT) is deleted as shown in Fig. 3. However we argue that PIT is in fact a very important basis in order to get user requirement and interest patterns, so the past PIT entries should be saved instead of deleted.
In our model, a History Interest Table (HIT) is introduced in order to record historical content requirements of within a period of time:

HIT <Name, Time, Location>

Thus, the content **Name** that user requests, the requesting **Time** and the **Location** where the request happens are all recorded by HIT. Hence, we can fit a content requirement distribution function CRF for a node in different locations with respect to Zipf's law: $P_{r,j}$, the probability of r-th ranking content being requested in location j.

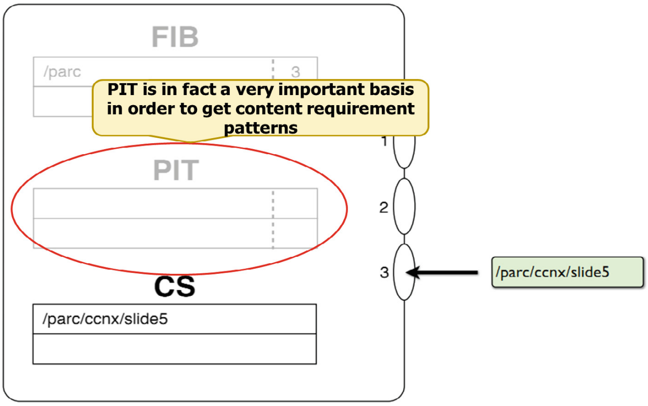

Fig. 3. The importance of PIT

3.3 Data Placement/Update Strategy

The placement/update strategy mainly focuses on optimizing data placement and deciding whether a content data should be replaced by a new one, in order to minimize the access overhead in a dynamic MANET environment. As shown in Fig. 1, we are facing two kinds of problem: placement/update strategy within a single region and between regions. From above user requirement modeling we understand that the requirement pattern tends to be similar within one single region, however differs from different regions. So we consider solving these two problems with ***Pcaching*** and ***Cache Rebalancing***, in both of which HIT plays the key role, as HIT is the most suitable representative to denote user query pattern differences between regions.

Data Placement/Update Strategy Within a Single Region: *Pcaching*. This strategy considers which data should be cached and where to be cached within one region. As the query pattern is similar in a single region, we propose a multi-factor scalable probabilistic cache placement strategy *Pcaching*. *Pcaching* considers HIT, battery status, cache occupation, content time-effectiveness and content popularity to form a Utility Function U that takes the above normalized parameters into account:

$$U = \sum_{i=1}^{N_p} w_i g(x_i) \tag{2}$$

In above function, weights W_i meet $0 \leq W_i \leq 1$ and $\sum_{i=1}^{N} w_i = 1$, and g(x_i) is the normalized parameter from HIT, battery status, and cache occupation, et al. As U has a value within the interval [0:1], it gives a result for caching probability when a node encounters a new data package: if $U \rightarrow 1$, the packet is cached with a high probability; on the contrary, if $U \rightarrow 0$, the packet is cached with a low probability. Therefore, when a node receives a data packet, it can compute its current caching probability for this packet, which is the *Pcaching* process.

Data Placement/Update Strategy between Different Regions: *Cache Rebalancing*. As different regions have different query patterns, or request CRFs, we diverse content requirements in different regions with the help of HIT. Therefore in order to optimize data placement/update, the strategy must be adaptable to locations. Thus we consider a *Cache Rebalancing* approach when node x enters a new region B. The approach asks any node it encounters for the HIT of region B, which means the node can proactively get the user requirement patterns of the new region it enters, and add the new regions CRF to its own one to rebalance its caching table. By periodically *Cache Rebalancing,* the caching network tends to become a steady state. The *Cache Rebalancing* approach is detailed described in below:

When Node x enters a new region B,

- x ask for $\mathrm{HIT_b}$ to any node y it encounters within region B;
- calculate the content request distribution function $\mathrm{CRF_b}$;
- Using $\mathrm{CRF_b}$, x generates interest package for content C_i with a probability of $P(C_i)$ for all the i contents of B, and cache all the responded content;

When a content data C passes Node x,

- x checks its own HIT, calculate its content request distribution function CRF;
- with CRF, x uses *Pcaching* process to computes the caching probability of data C;
- As the request arrival time interval fits Poisson distribution:

$$P_n(t) = \frac{1}{n!}(\lambda t)^n e^{-\lambda t} \tag{3}$$

We assume the data update time interval should also fit upper distribution. By periodically updating its cache, the *Caching Rebalancing* mechanism makes caching exchanges between different regions smoothly and costless. We can further more get the steady state for a cache network in a region:

$$R_{r,j} = P_{r,j} + \sum_{i \neq j} R_{r,i} T_{i,j} \tag{4}$$

$R_{r,j}$ denotes the final probability of the r-th ranking content data being requested in region j, $P_{r,j}$ is the original requesting probability (which can be computed by $\mathrm{HIT_j}$) of the r-th ranking content data being requested in region j, and $\mathrm{T_{i,j}}$ is the transition probability of node x moves from region i to j.

4 Performance Evaluation

We evaluate the performance of our scheme using ndnSIM [12] simulator based on ns-3. We also use real city taxi traces of Beijing to best simulate node mobility. Our goal is to show the benefits of our caching placement strategy by comparing with existing ICN caching placement approaches (Table 1).

Table 1. Simulation parameters

Parameter	Value
Area size	24 km × 24 km
Time of simulation	7200 s
Number of nodes	3000
Radio range	50 m

Hotspot Recognition and Division. In this section, we use DBSCAN clustering algorithm to analyze city taxi traces in order to get the top ranking clusters of load and drop events. Based on the results of clustering, we take a further step to refine our results by using Graham's scan method to compute the convex hull of each cluster, and for each convex hull, we calculate its Minimum Bounding Rectangle (MBR) in order to make the results more accurate, as shown in Fig. 4.

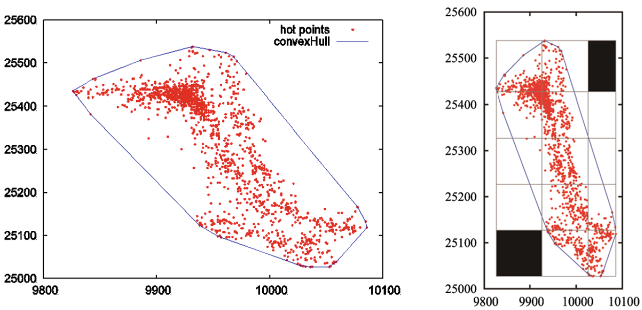

Fig. 4. Minimum bounding rectangle for a hotspot

With MBR, which is the minimum bounding rectangle for hotspots, we finally get an accurate hotspot region division for MANET. In Fig. 5 we can see that these regions include famous scenic spots, large business areas, railway stations and subway stations, which reflex the real social mobility patterns for city taxies.

Data Placement/Update Strategy. In order to demonstrate the benefits of *Pcaching*, we compare *Pcaching* to the default CCN caching strategy LCE (leave copy everywhere) as shown in Fig. 6, which tells the cache hit ratio when Zipf parameter α differs in different cache placement strategies. Large value of Zipf parameter α means a given content gets much higher popularity than others in this region. As can be seen, *Pcaching* shows better cache hit ratio that is nearly two times higher than LCE. This implies that workload of content servers in *Pcaching* is better distributed than LCE. This result is not surprising because LCE strategy caches content chunks aggressively and without any choosing, which could lead to caching redundancy to the whole network, therefore takes down the cache hit ratio of the network. In *Pcaching*, cache hit ratio increases as Zipf parameter increases. This result suggests that the competitive advantage of *Pcaching* comes from those regions that users tend to social and share popular contents through the caching network.

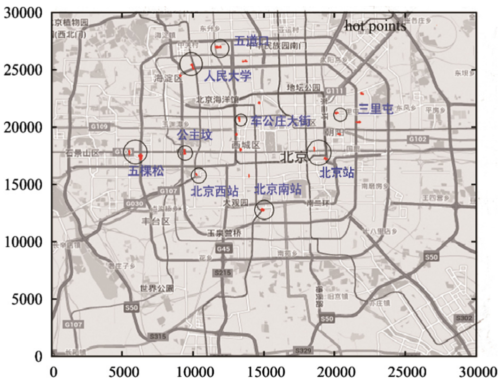

Fig. 5. Top hot regions in a city

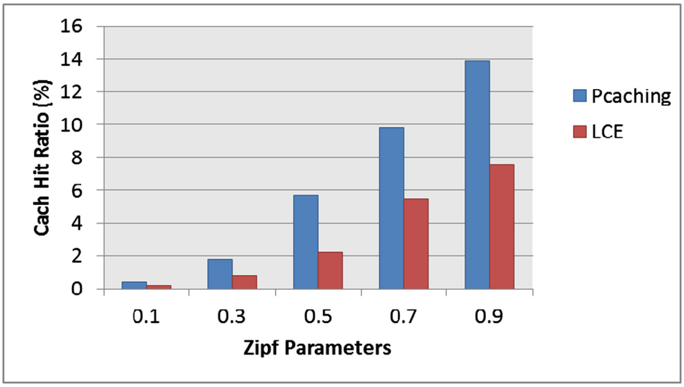

Fig. 6. Cache hit ratio comparison between cache placement strategies

Figure 7 plots the average hop counts by cache size and Zipf parameters. The cache sizes of node ranges from 5 to 25 blocks which represents 5% to 25% of total content number. Compared with the other scheme LCE, we can see that *Pcaching* provides lower average hop counts for different cache sizes. This tells that the *Pcaching* efficiently uses the cache memory to improve the effectiveness of the whole network. We also see that with Zipf parameter increases, average hop counts decreases while *Pcaching* still has advantage to LCE. The results shows hop counts can be reduced by *Pcaching*, and *Pcaching* works better in regions where social communication patterns are frequent.

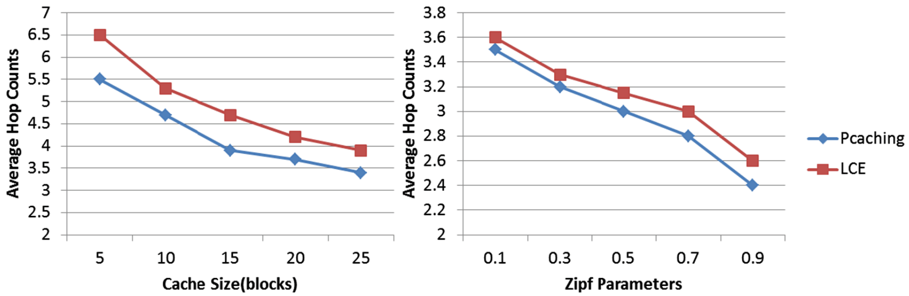

Fig. 7. Hop counts for comparison between cache placement strategies

5 Conclusions

In this article, we study cache placement problem for ICN in a MANET environment. Based on hotspot recognition and user requirement modeling, two approaches named *Pcaching* and *Cache Rebalancing* are proposed to meet the caching placement/update needs for single region and between regions. Through real city taxi mobility trace simulations, we show that our approach outperforms existing CCN traditional caching scheme in both cache hit ratio and hop counts, which indicates that our approach reduces network overhead while maintaining solid performance.

Acknowledgements. This work is supported by these Projects: The National Natural Science Foundation Project of China under Grant No. U1636208; The Co-Funding Project of Beijing Municipal education Commission under Grant No. JD100060630.

References

1. Xylomenos, G., et al.: A survey of information-centric networking research. IEEE Commun. Surv. Tutor. **16**, 1024 (2014)
2. Cohen, E., Shenker, S.: Replication strategies in unstructured peer-to-peer networks. ACM Sigcomm Comput. Commun. Rev. **32**, 177 (2002)
3. Ko, B., Rubenstein, D.: Distributed self-stabilizing placement of replicated resources in emerging networks. IEEE/ACM Trans. Netw. **13**, 476 (2005)
4. Qiu, L., Padmanabhan, N.V., Voelker, M.G.: On the Placement of Web Server Replicas. Cornell University, Ithaca (2000)
5. Jiang, X., Bi, J., Wang, Y., Lin, P., Li, Z.: A content provider mobility solution of named data networking. In: IEEE International Conference on Network Protocols, vol. 1 (2012)
6. Wei, T., Chang, L., Yu, B., Pan, J.: MPCS: A mobility/popularity-based caching strategy for information-centric networks. In: Global Communications Conference, vol. 4629 (2015)
7. Lee, S.B., Wong, S.H.Y., Lee, K.W., Lu, S.: Content management in a mobile ad hoc network: beyond opportunistic strategy. In: IEEE INFOCOM, vol. 266 (2011)
8. Wang, H., Yang, W., Zhang, J., Zhao, J., Wang, Y.: START: status and region aware taxi mobility model for urban vehicular networks. In: Computer Communications Workshops, vol. 594 (2015)

9. Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C.: Introduction to Algorithms, 2nd edn, p. 1297. The MIT Press, Cambridge (2001)
10. Caldwell, D.R.: Unlocking the mysteries of the bounding box. Coordinates Online Journal of the Map & Geography Round Table of the American Library Association (2005)
11. You, W., Mathieu, B., Simon, G.: How to make content-centric networks interwork with CDN networks. In: Network of the Future, vol. 1 (2014)
12. Afanasyev, A., Moiseenko, I., Zhang, L.: ndnSIM: NDN simulator for NS-3 (2012)