This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2021.3055860, IEEE Transactions on Big Data

Discovering and Understanding Geographical Video Viewing Patterns in Urban Neighborhoods

Jiaqiang Liu, Huan Yan, Yong Li, *Member, IEEE,* Dmytro Karamshuk, Nishanth Sastry, Di Wu, *Member, IEEE* and Depeng Jin, *Member, IEEE*

Abstract—Video accounts for a large proportion of traffic on the Internet. Understanding its geographical viewing patterns is extremely valuable for the design of Internet ecosystems for content delivery, recommendation and ads. While previous works have addressed this problem at coarse-grain scales (e.g., national), the urban-scale geographical patterns of video access have never been revealed. To this end, this paper aims to investigate the problem that whether there exists distinct viewing patterns among the neighborhoods of a large-scale city. To achieve this, we need to address several challenges including unknown of patterns profiles, complicate urban neighborhoods, and comprehensive viewing features. The contributions of this paper include two aspects. First, we design a framework to automatically identify geographical video viewing patterns in urban neighborhoods. Second, by using a dataset of two months real video requests in Shanghai collected from one major ISP of China, we make a rigorous analysis of video viewing patterns in Shanghai. Our study reveals the following important observations. First, there exists four prevalent and distinct patterns of video access behavior in urban neighborhoods, which are corresponding to four different geographical contexts: downtown residential, office, suburb residential and hybrid regions. Second, there exists significant features that distinguish different patterns, e.g., the probabilities of viewing TV plays at midnight, and viewing cartoons at weekends can distinguish the two viewing patterns corresponding to downtown and suburb regions.

Index Terms—LDA, geographical behavior, video access, topic modeling.

1 Introduction

Watching videos has become one of the most popular user activities on the Internet. According to the recent reports, video traffic accounts for 64% of all consumer internet traffic in 2014 [1] with 89.6% of Internet users in China watching videos online and each third user doing so on a daily basis [2]. The rapid development of mobile Internet connectivity and proliferation of handheld devices over the recent years have triggered a similar surge in mobile video usage, where a 11-fold increase is expected from 2015 to 2020. Consequently, the audience of online video services is progressively covering more *diverse demographic* groups across wider *geographic areas*, and thus affecting the design of various components of the Internet ecosystem: from geographic-aware content caching techniques to geographically targeted ads and content recommendation systems.

In such a context, it is important to study the geographic access patterns of online video services. Several recent *large-scale geospatial studies* of major user-generated video providers [4], catch-up TV systems [13] and peer-to-peer video streaming services [14] have pioneered this research area by looking at the patterns emerging on coarse-grain geographical scales (e.g., countries and counties), and showing its implications for designing nation- or planet-wide content delivery infrastructures.

However, how video access patterns exhibit on *finer geographic scales*, e.g., urban neighborhoods, is still less explored. The analysis is challenged by several obstacle factors. Firstly, urban neighborhoods are often vaguely defined, which may vary depending on the considered scenario [18] [5]. Secondly, there is lack of a stereotype knowledge on the access characteristics for different types of neighborhoods. Thirdly, the scope of parameters impacting specific patterns of video accesses, including dynamics of *demographic specifics*, *external events*, etc., is so much broad that challenges the whole definition of what should be considered as a repeating geographic pattern.

In this paper, we present the study of geographic patterns for video streaming viewing in urban neighborhoods. We tackle the aforementioned challenges by introducing a comprehensive *geocomputation framework* to identify the video access patterns observed in different urban neighborhoods. Our framework does not require any prior knowledge about neighborhoods in a city and operates by analyzing the *road networks* extracted from digital maps (such as Baidu Maps). We apply a wide range of signals coming from video access logs, including diurnal and weekly access patterns, content and device preferences, etc., to achieve this. We finally, provide a thorough analysis and understanding

- *J. Liu, H. Yan, Y. Li and D. Jin are with Department of Electronic Engineering, Tsinghua University, Beijing, China.*
D. Karamshuk and N. Sastry are with Centre for Telecommunications Research, King's College London.
D. Wu is with School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China.
Email: liyong07@tsinghua.edu.cn.
This work is supported by National Basic Research Program of China (973 Program) (No. 2013CB329105), National Nature Science Foundation of China (No. 61301080, No. 61171065 and No. 61273214, No. 91338203, and No. 91338102).

1. Cisco Visual Networking Index - http://bit.ly/LVhmuL
2. Statistics and Market Data on Video & Entertainment, Statista.com - http://bit.ly/1Vaao4Z

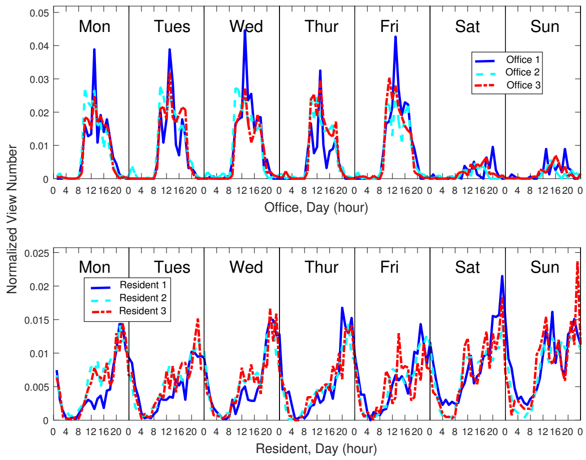


Fig. 1. Temporal video access in office and residential regions in one week.

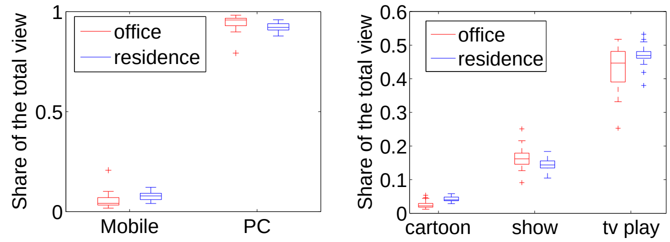
for the discovered patterns including correlation analysis with the categories of land use. Our contributions can be summarized as follows:

- We design a framework for automatic discovering of video viewing patterns in urban neighborhoods. The framework operates, by first dividing the city into non-overlapping regions (named urban neighborhoods in this paper) constrained by road networks. Then, it uses the *latent dirichlet allocation* to model viewing requests over a wide range of dimensions including time of access, device and content types. Finally, it clusters the region-topic distribution to identify distinct geographical viewing patterns.
- We conduct a comprehensive measurement study of *viewing patterns in Shanghai*, one of the largest cities in China, exploiting the two-month dataset collected from a major Internet Service Provider. By employing the designed framework, we discover four distinct patterns and analyze their specific characteristics. Furthermore, we demonstrate that the discovered patterns are distinguishable from each other by a few significant features.

The rest of this paper is organized as follows. In Section 2 we introduce the dataset exploited in our study. Next, we present the framework for discovering geographic video viewing patterns and the experiment results using our dataset in Section 3. Finally, we study the characteristics of each discovered pattern in Section 4, followed by the discussion of related work in Section 5 and conclusion in Section 6.

2 Dataset

In this paper we analyze the dataset collected at the gateways of a major ISP network in Shanghai, one of the largest cities in China, through *deep packet inspection* (DPI). DPI appliances record user requests for video content by parsing the application layer protocol of the packets, e.g., HTTP. Each log record contains details of a single user request including the *anonymized user ID*, *model of the device*, *URL of video* and the *time of the request*. By crawling the URLs of the video content items, we additionally collected the name and the genre of the videos (i.e., cartoon, movie, show, tv play or other) [3], their unique identifiers and content providers. Finally, the location of the users were obtained by resolving their addresses - contained within the ISPs databases - to geographic coordinates using the geolocation API provided by Baidu Map service.



(a) Different types of devices (b) Different types of videos

Fig. 2. The share of views (a) from different types of devices and (b) for different types of videos.

Overall, the analysis in this paper is based on the dataset of 84 million requests collected between Nov. 1st and Dec. 31st 2014 from 1.4 million users for 1 million video content items. The dataset covers top six most popular video content providers in China, i.e., Youku (YK), Iqiyi (AQY), Sohu (SH), Xunlei (XL), Leshi (LS), and Tecent (TX), which are comprehensive content providers that support video on demand service and contain all five types of videos we study. All geographic neighborhoods and demographic groups are well represented in our dataset.

Ethics The goal of our study is to have a good understanding of geographical video viewing patterns in urban neighborhoods. Our dataset is obtained via the collaboration with ISP who has anonymized the user information. Meanwhile, this work has been approved from our university.

3 Discovering Viewing Patterns

In this section, we first present some basic analysis and observations that motivated our study and then introduce the methodology for discovering geographical viewing patterns.

3.1 Intuition and Observation

A key intuition behind our study is that the video *access behavior of users may vary across urban neighborhoods*, e.g., viewing patterns in office locations are different from those in residential places. To verify this intuition, we manually inspect the viewing requests of the users in randomly selected 30 office and residential neighborhoods in Shanghai. Fig. 1 shows the volumes of requests in each week hour, normalized over the total volume of requests in each week, for three arbitrary neighborhoods in each category. The plots suggest considerable differences in the two categories. The view requests in office regions are concentrated around daytime; whereas the peaks of user activity in residential regions is around 8pm at night. Additionally, there are few video views at weekends in office regions.

3. A TV play includes a series of pre-shot videos. For example, Big Bang is a TV play. The category of show includes on-site videos of some talks and performance. For example, the videos of Running Man are in this category. All videos not in the cartoon, movie, show and tv play categories are classified to the category of other, which includes short news, user generated contents, and etc.

Further, we analyze the share of views from mobile gadgets and PCs, and the share of views for different video genres. Fig. 2 shows the box-plot of the results over 30 selected office and residential neighborhoods. It is clear from the plots that residential regions have more views from mobile devices in comparison with the office regions. In fact, for selected residential regions, the median value of the share of views from mobile gadgets is 0.12, which is nearly three times higher that for office regions. This is possibly due to users have more direct access to PCs in their work positions. Additionally, from Fig. 2(b), we find that users prefer to watch cartoons and TV plays at home and watch shows in office. Specifically, the median share of views for cartoons in selected residential regions is two times of that in office regions, while the median share of views for shows in selected residential regions is 11.1% less than that in office regions. The possible reason is that TV plays and cartoons usually include series of videos and users would like to watch them continuously when they have more free time at home.

Overall, these preliminary observations confirm that the video viewing behavior varies substantially between the neighborhoods with different urban functions. However, it is still unknown whether there are some specific video viewing patterns among all neighborhoods in a city. Motivated by the intuitions and to answer this question, we design a framework to automatically extract the potential geographical video viewing patterns, by combining the features of temporal, device type and video genre.

3.2 Methodology

We aim to identify the video view patterns among urban neighborhoods. Since we have no knowledge of viewing patterns, the general idea is to map the viewing requests to the neighborhoods according to the location where they are generated. Then, we exploit viewing requests as features and use unsupervised learning technique to classify the neighborhoods into different types with objective that the viewing requests involved in the neighborhoods of the same type have similar characteristics, which are regarded as "patterns" in this paper.

We use latent dirichlet allocation (LDA) to model the video viewing requests in each region and then use the estimated parameters to cluster regions. Latent dirichlet allocation is a topic model that widely used in document modeling and classification. It regards a document as a bag of words and assumes each word is generated by some latent topics that have specific word distribution. In general, we find there are some similarity between a document and the viewing requests in a region. Particularly, we can treat each viewing request as a word, and all requests in a region are generated by some "topics", e.g., users with similar interests, daily routines and preference of devices. These similarities motivate us to model the view requests with LDA.

Based on above considerations, we design a geocomputation framework to automatically extract geographical viewing patterns from the dataset of viewing video requests. As shown in Fig. 3, the extraction process consists of two main steps. In the first step, the city is divided into non-overlapping neighborhoods according to the road network.

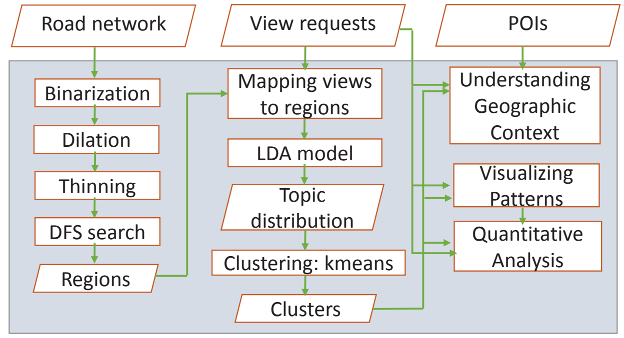


Fig. 3. The framework to automatically discover video viewing patterns.

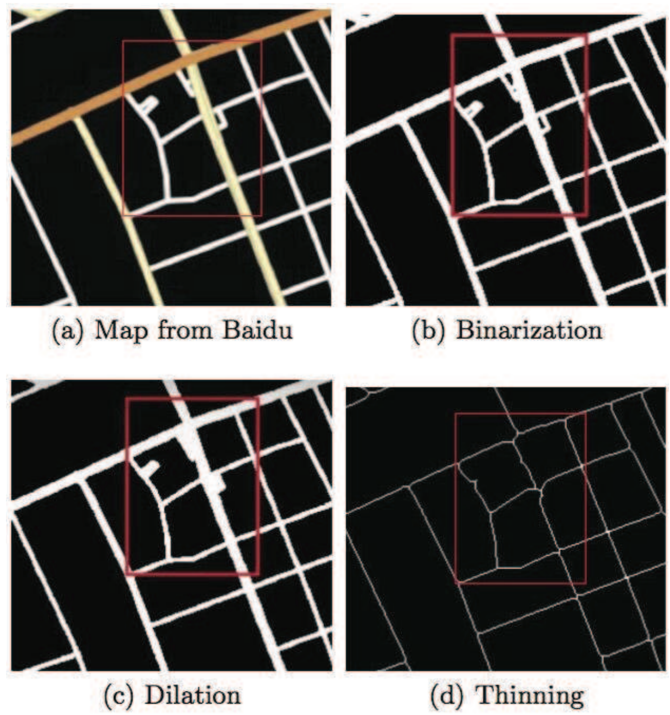


Fig. 4. Illustration of region division process.

Then, neighborhoods are clustered into a multiple of groups based on the observed video viewing requests. Specifically, we first use LDA to model the video viewing requests, discover the topics of users common viewing behavior, and then exploit the region-topic probability distributions as features to cluster regions. Next, the geographical contexts are learned according to the point of interests (POIs) information, and we study how the geographical contexts have impact on the viewing patterns (discussed in Section 4). In the following subsections, we firstly introduce the methodology for dividing the city using the road information and further discuss how we apply LDA to identify video access patterns.

3.2.1 *Region Division*

Following an established practice in land economics [18] [19], we use the regions formed by the *road network* as basic spatial units of neighborhoods to study geographical patterns in access patterns. Instead of using inaccurate public road network data such as OpenStreet [21], we extract the road network directly from digital images provided by credible online map service. Specifically, we exploit the approach introduced by [10] to obtain the boundaries of regions represented by the sequences of lattitude and longitude coordinates. The algorithm proceeds in three steps. Firstly, we obtain an image of road networks from Baidu Maps. As Fig. 4(a) shows, there are three levels of roads, represented by orange (level 1), yellow (level 2), and white (level 3), respectively. We then perform a binarization operation on the road map. As a result, we get a binary image, with 0-s representing roads and 1-s representing blank spaces (Fig. 4(b)). Secondly, we conduct a dilation operation [15] to remove very small regions caused by lanes and overpasses (Fig. 4 (c)). As there is no lanes in level 3 roads, to avoid merging of some small regions in the city center, this operation is only applied on level 1 and level 2 roads. Third, we perform a thinning operation over all roads to reduce the space occupied by them, as shown in Fig. 4(d). Finally, we obtain the boundaries of each region using Moore's neighborhood tracing algorithm [22], as shown in Algorithm 1. The basic idea of this algorithm is to extract the trace of a boundary by iteratively searching the boundary pixels until the first boundary pixel is visited for a second time. Here a pixel represents 0 or 1 in a binary image.

Algorithm 1: Moore's neighborhood tracing algorithm [22]

```
Input: A binary image T
Output: A list B of boundary pixels
Initialization
B ← Φ;
s ← the first black pixel is found when the cells of T
    is scanned from bottom to top and left to right;
B ← append s into B;
p = s; // the current boundary pixel
b ← the backtrack of p;
M(p) ← Moore neighborhood of pixel p;
c ← the next clockwise pixel (from b) in M(p);
while c ≠ s do
    if c is black then
        B ← append c into B;
        b = p;
        p = c;
        c ∈ M(p) : the next clockwise pixel (from b)
          in M(p);
    else
        b = c;
        c ← the next clockwise pixel (from b) in
          M(p);
return B;
```

3.2.2 Pattern Identification

Before diving into the details of pattern identification methodology, we firstly introduce latent dirichlet allocation (LDA) – the model we use to characterize the viewing requests in each region. LDA is one of the topic models that are widely used to automatically extract hidden themes from a large corpus of documents [3]. In LDA, topics are distinguished by the probability distributions over words, and documents are represented by the probability distributions over topics. Assume there are K topics and let β_k represents the topic-word distribution of k-th topic over all V words. Further, let θ_d represent the document-topic distribution of d-th document, where $\theta_{d,k}$ is the proportion of k-th topic in document d, and let Z_d represent the sequence of topics to generate d-th document, where $Z_{d,n}$ is the selected topic for the n-th word. With these notations, each document is generated as follows:

1) Generate $\beta_k \sim Dir(\eta)$;
2) For each document d, generate $\theta_d \sim Dir(\alpha)$;
3) For nth word in document d, generate topic $Z_{d,n} \sim Multi(\theta_d)$, then generate the word $w_{d,n} \sim Multi(Z_{d,n})$.

Here, α and η are two parameters of LDA, which represent the prior document-topic distribution and topic-word distribution, respectively. Dir represents the dirichlet distribution, and $Multi$ represents the multinomial distribution. More details of the most widely used *EM algorithm* and *Gibbs sampling* method to obtain the parameters in LDA model can be found in [3].

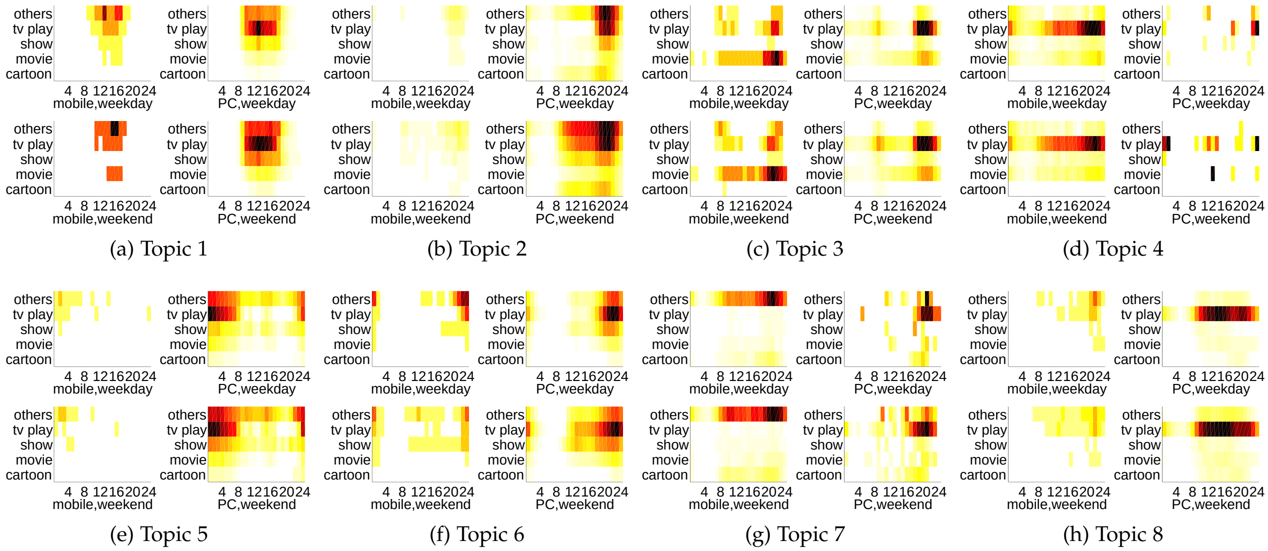
We use a four-tuple $V = (L_v, D_v, T_v, C_v)$ to represent the video request. For each request v, $v \in V$, L_v represents the location of the subscriber, D_v represents the utilized device, T_v represents the time of this request, and C_v represents the requested content. To apply LDA model, we discretize each dimension to form "documents" and "word vocabulary". Specifically, each request is mapped to a region, according to the subscriber's location. The devices are categorized into two groups: *mobile* includes mobile phones and tablets, *PC* includes personal computers. The videos are mapped to five types: cartoon, movie, show, TV play and others, directly according to the type value in the crawled video meta data. The time is discretized into one hour time bins. Time in a week is discretized into different bins in order to capture weekly patterns. By combining the device, content and time information, a cuboid is formed. Each cubic lattice is regarded as a word, where cubic lattice (d, c, t) represents the view requests using $d-$th type of device, watching $c-$th type of videos during $t-$th time bin.

By applying Gibbs sampling and EM algorithm, we obtain the word probability distribution $\beta_k(d, c, t)$ of each topic k and topic probability distribution θ_r of each region r. $\beta_k(d, c, t)$ denotes the probability that a request is generated at time bin t, using device d, and for videos in the category of c. This probability distribution quantitatively characterize some common user behavior of watching videos in time, device and video dimension. Since the distributions vary across topics, together they can be exploited to analyze the viewing behavior in a geographic region, as topic distribution θ_r suggests. Specifically, $\theta_r(k)$ denotes the probability that a viewing request in region r is generated according to the "word" distribution of topic k. Therefore, the regions with similar topic distribution would have similar video viewing behavior. To this end, we can use topic distributions as features to cluster regions and further extract video viewing patterns.

3.3 Experiment Results

We acquire the road map of Shanghai from Baidu Maps, and obtain the regions of neighborhoods using the approach in-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2021.3055860, IEEE Transactions on Big Data



(a) Topic 1 (b) Topic 2 (c) Topic 3 (d) Topic 4

(e) Topic 5 (f) Topic 6 (g) Topic 7 (h) Topic 8

Fig. 5. Visualizing the video viewing behavior of each topic.

troduced in previous section. The region division results in 9416 regions with more than 83 million total view requests.

According to the similarity of topic-word distributions between different topics, we obtain eight topics. Based on the introduced methodology, we regard each region as a document and train the LDA model using the selected 9146 regions.

Fig. 5 visualizes the probability distribution of view requests in each topic via heat map. Some common watching behaviors can be observed from this figure. Specifically, in Topic 1, there is a high probability of watching TV plays and shows on PC, and watching movies and TV plays on mobile device in day time, which reflects the behavior of watching videos at office. In contrast, the view requests in Topic 2 and Topic 6 are usually generated at weekends or at night of weekdays, which reflects the behavior of watching videos when go back home after work. Topic 2 are different with Topic 6 in that Topic 6 has a higher probability of watching videos on mobile devices at midnight (0am). Further, Topic 4 captures the behavior of watching most of TV plays on mobile devices, which is possibly generated by users with no direct access to PCs. Topic 3 and Topic 7 captures the behavior of watching movies and other videos (usually news reports and user generated videos) on mobile devices. But different to Topic 4, there is also a high probability of watching TV plays at night in these topics. Topic 5 reflects the behavior of watching videos on PCs at late night (0am-4am). Finally, in Topic 8, the probability of watching TV plays on PC is high through all wake time, possibly reflecting the watching behavior of users who have a lot of spare time.

We then cluster the regions according to the topic distribution of them using K-means. By referring to the average *silhouette value* and observing the actual clustering results, we set the number of clusters to four. Table 1 summarizes the basic statistics of four discovered clusters. Cluster 1 (C1) is largest cluster that contains 31.4% of regions, which in their turn, cover 58.5% of the user base and 47.9% of view requests. However, the average number of view requests per subscriber is only 0.64, the lowest among all clusters, possibly implying that users in C1 spend less time on watching videos.

TABLE 1
The basic statistics of each cluster.

cluster ID	C1	C2	C3	C4
# regions	2958	1954	2061	2443
# subscribers ($\times 10^4$)	62.03	8.45	17.85	17.77
# view requests ($\times 10^6$)	39.79	8.90	15.89	18.52

TABLE 2
The average cluster-topic distribution.

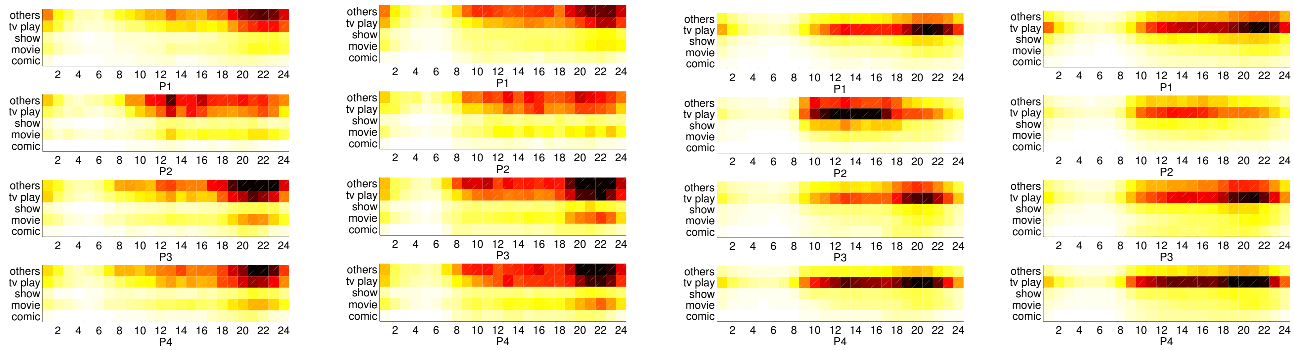
Topics	1	2	3	4	5	6	7	8
C1	0.13	0.13	0.05	0.03	0.08	0.33	0.03	0.23
C2	0.55	0.07	0.08	0.03	0.05	0.07	0.03	0.11
C3	0.15	0.34	0.12	0.04	0.06	0.09	0.03	0.17
C4	0.14	0.13	0.12	0.04	0.05	0.10	0.03	0.39

Table 2 presents the average region topic distribution of each cluster. Each cluster has a dominant topic with the highest probability among all topics - an indicator which helps to understand the differences in viewing behavior for each cluster. For example, Topic 1 is the dominant topic of Cluster 2, therefore, we can expect that in the regions of Cluster 2 (C2), most of viewing requests are generated between 10am and 4pm. Besides, the probability of Topics 4, 5 and 7 are less than 10% in all clusters. By analyzing the results in Fig. 5, we find that this is caused by the difference of viewing behavior they characterize. For instance, for Topic 5, 55.6% of video requests are generated in the night hours (12am-7am) potentially indicating accesses from home locations; whereas 97.8 % and 87.8% of accesses in Topic 4 and 7, respectively, (in contrast to the average of 8.44%) are coming from mobile devices - indicating a prominent mobile access pattern.

4 Understanding Viewing Patterns

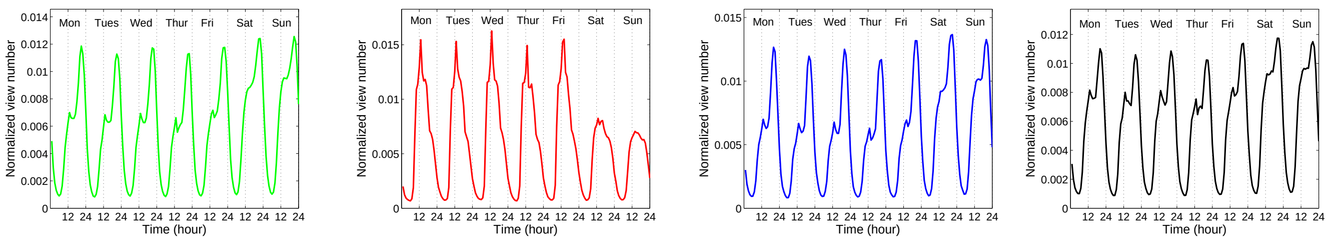
In this section, we study and understand the discovered viewing patterns from multiple aspects. First, we visualize the viewing patterns to directly illustrate their characteristics. Then, we quantitatively analyze the viewing patterns from the aspect of time, device and type dimensions. After that, we investigate the question of whether there are some

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2021.3055860, IEEE Transactions on Big Data



(a) Mobile, weekday (b) Mobile, weekend (c) PC, weekday (d) PC, weekend

Fig. 6. Visualizing the viewing patterns.



(a) P1 (b) P2 (c) P3 (d) P4

Fig. 7. The average normalized view number against time of four discovered patterns.

geographical contexts corresponding to the discovered patterns by studying the distribution of point of interests. Finally, we use Kullback-Leibler distance to identify the most significant features of viewing requests that differentiates the four patterns.

4.1 Visualizing Viewing Patterns

We define a viewing pattern as the average of normalized number of view requests over regions of a cluster. Therefore, we can obtain four patterns from the experiment results, which are denoted as P1, P2, P3 and P4 in following analysis. Fig. 6 presents the heat map of hourly viewing requests for 5 different types of videos, where a darker color represents a higher number of views. Since the share of views from mobile devices is low in general, to visualize its characteristics, we use one max value in the normalization of (a) and (b), and another max value in the normalization of (c) and (d). Besides, as the pattern is similar in the weekdays and during weekends, we use the result of Monday and Saturday to represent them respectively.

We can directly obtain some observations from Fig. 6. As Fig. 6(c) shows, in the weekday, P2 has more views during daytime from 10am to 5pm, while P1 and P3 has more views at night. In P4, the hourly number of views also achieves highest at night, but different to P1 and P3, P4 also attracts a lot of views at daytime. Further, by comparing Fig. 6(c) and Fig. 6(d), we see that P2 has less views at weekend. Moreover, Fig. 6(c) and Fig. 6(d) shows that most PC generated view requests are for videos of TV plays. Turn to Fig. 6(a) and Fig. 6(b) which illustrates the views from mobile devices, we observe the similar temporal trend for P1, P3 and P4, while we see no obvious temporal hot spot for P2. Besides, instead of TV play, the videos of others account for the highest ratio of views among the five studied genres.

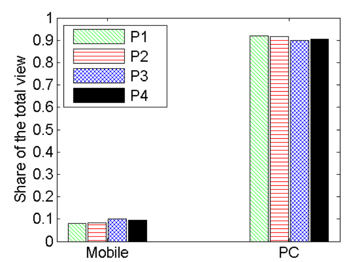


Fig. 8. The share of views from different types of devices.

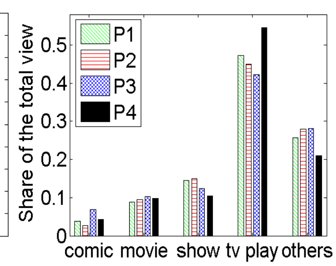


Fig. 9. The share of views for different video genres.

What is more, the share of movie is higher for the views from mobile devices than that from PC devices.

4.2 Quantitative Analysis of Viewing Patterns

The features of a view request includes three dimensions: time, device and video genre. To help understand the discovered patterns, we start by analyzing the distribution of view requests along each dimension. Fig. 7 shows the results in time dimension, i.e., hourly normalized view requests in a week of each pattern. Obviously, P2 has one prominent daily peak, whereas all three other patterns have two daily peaks. Besides, P2 has fewer views at weekends than weekdays. Moreover, while P1, P3, and P4 featuring similar daily trends, they diverge in the ratio of two peaks, which are 0.564, 0.483 and 0.671 respectively. Further more, there are also some differences between P1 and P3. For example, the peak hour in P1 is around 9pm, which is 1 hour later than that of P3. In addition, the relative view number in 0am to the peak hour is 0.61 in P1, while it is only 0.37 in P3.

Fig. 8 and Fig. 9 shows the share of views using different types of devices and the share of views for different types of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2021.3055860, IEEE Transactions on Big Data

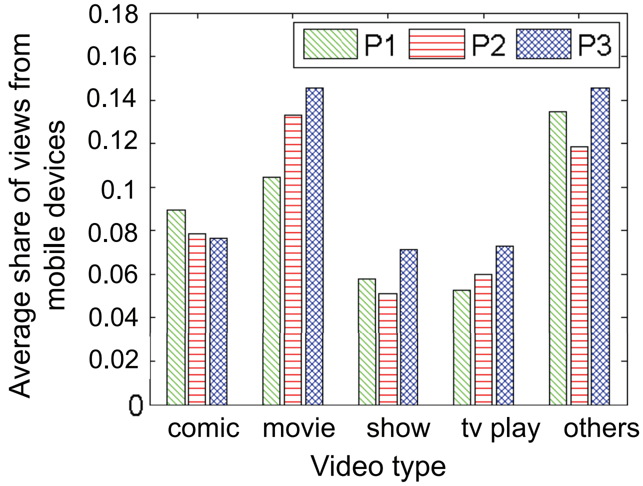


Fig. 10. The share of mobile views for different video genres.

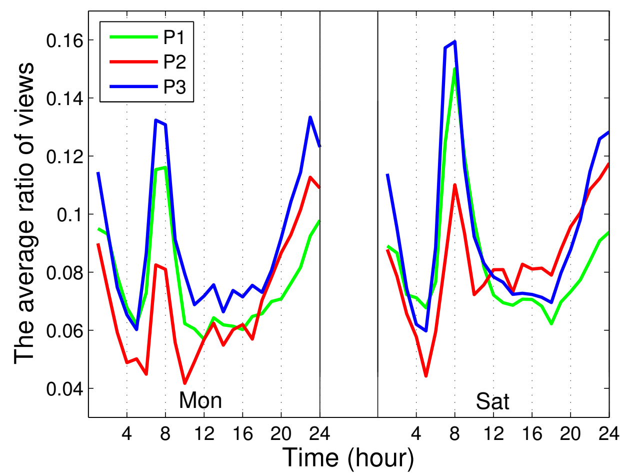


Fig. 11. The hourly share of views from mobile devices.

videos respectively. We observe that the share of views from mobile devices are nearly equally among the four patterns, while that of P3 and P4 are a little higher. For different video genres, P1 and P2 have more views for shows than P3 and P4; while P3 have more views of cartoon, and P4 have more views of TV Play.

As we have shown in Fig. 6, TV play accounts for most of view requests generated from PC, while others accounts for most of views generated from mobile devices. To quantitatively analyze this characteristic, we calculate the share of views from mobile devices in terms of video genre and plot it in Fig. 10. It shows that generally the share of mobile devices is low, less than 15%, despite of video genres considered. However, on average, the videos of others have almost two times of shares of mobile views than show and TV plays. Besides, we observe that the specific share of views still varies across different patterns. For example, in P3, movie accounts for 14.5% of view requests from mobile devices, while this value is only 10.5% in P1.

Further, since the time of a day would influence the usage of mobile devices. To characterize it, we calculate the average share of views from mobile devices per hour in each pattern and show it in Fig. 11. We observe that there is a significant change of the shares from mobile devices in one day. Specifically, it has two peaks: in the morning around 8am and at midnight around 11pm. This implies that users are more likely to use mobile devices to watch videos in the early morning and late night, corresponding to the common time of getting up and going to bed. Moreover, we can observe that in P1 and P3, the share at 8am is higher than that at 11pm, while it is opposite in P2. Besides, the peak share in P1 and P3 at 8am is obviously higher than that of P2.

4.3 Geographic Contexts of Viewing Patterns

An interesting question is where are the neighborhoods corresponding to each pattern located. To answer this question, we show the regions of the four clusters on the map in Fig. 12 [4]. The map suggests a vivid spatial distribution:

4. White regions indicate the neighborhoods with less than 500 requests, therefore, excluded from our analysis.

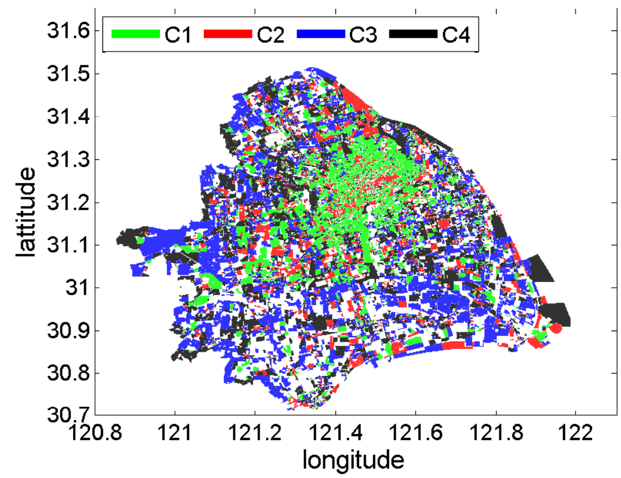


Fig. 12. The spatial distribution of regions in different clusters.

most of the neighborhoods of P1 are located at the city center, the neighborhoods of P3 are mainly concentrated at the surrounding suburbs, and neighborhoods of P2 and P4 are distributed all over the city. By comparing this result with the temporal patterns from Fig. 7, where P2 have fewer views at weekends, P1 and P3 have more views at evening and at weekend, we can infer that P2 are mainly generated by views from office locations, while P1 and P3 are generated by views from residential regions at city center and suburb areas respectively.

To learn more about the geographical context, we study the distribution of POIs corresponding to each pattern. A POI is a point location with a specific function such as restaurant and shopping mall. We collect 21 categories of POIs from Baidu Map, ranging from food to village (the specific types are shown in Table 3). Since the region areas may lead to great differences of the number of POIs with a specified category in different regions, it is unreasonable to compare the POI categories with regions. Thus, we use the density of each POI category to eliminate the influence of different region areas. Formally, for region r, the density of

TABLE 3
Average density of POI categories in four discovered clusters. (The density is marked according to its rank among four clusters, with darker color indicating higher rank.)

category	P1	P2	P3	P4
food	1.24	1.70	0.39	0.72
hotel	1.34	1.42	0.42	0.80
shopping	1.09	1.69	0.46	0.87
entertainment	1.20	1.80	0.39	0.69
sports	1.14	2.17	0.23	0.60
school	1.57	1.25	0.47	0.63
scenery spot	1.00	2.08	0.40	0.71
tourism zone	1.79	0.38	0.00	1.46
financial area	0.87	2.62	0.32	0.50
office building	0.75	3.32	0.09	0.26
company	0.86	2.43	0.37	0.62
trade area	1.56	1.00	0.13	1.12
factory	0.75	1.46	0.87	1.13
industrial park	0.95	2.24	0.50	0.57
science park	1.13	2.41	0.30	0.37
economic zone	0.51	3.18	0.11	0.66
high-tech zone	0.72	3.39	0.00	0.33
residential district	2.15	0.52	0.39	0.57
life service	1.30	1.58	0.47	0.69
township	0.99	0.88	1.32	0.93
village	0.77	0.68	1.42	1.27

j-th POI category, denoted by $d_{r,j}$, is defined as follows:

$$d_{r,j} = \frac{\text{The number of j-th POI category in region r}}{\text{The area of region r (measured in } km^2\text{)}}.$$

Moreover, as the total number of POIs of different categories varies significantly due to their functional difference. To eliminate the influence, we take out a linear normalization that makes the average density of each POI category end up to one. Specifically, the normalized density of j-th category of POI for region r, denoted by $\overline{d}_{r,j}$, is defined as:

$$\overline{d}_{r,j} = \frac{\sum_{r\in R} d_{r,j}}{|R|}.$$

Then we average the normalized densities over regions of each cluster, which are summarized in Table 3. The data is marked according to its rank among four clusters, with darker color indicating higher rank.

From Table 3, we find that the normalized density of residential districts is highest among all categories of POIs for P1. It is also much higher than that of other patterns. For P2, the POIs of high-tech zone, office building, economy zone and financial area have highest densities, which suggest that the neighborhoods corresponding to P2 are mainly office regions. The categories of POI with highest densities for P3 are village and township, suggesting that most neighborhoods of P3 are located in suburb areas, which agrees well with the geographical distribution in Fig. 12. The top four densest POI categories for P4 are tourism development zone, village, factory and trade area. As factory and trade area are work places, and village is residential place, it suggests that a neighborhood corresponding to P4 may include both work and residential areas. In conclusion, there is a strong correlation between the four discovered patterns and geographical downtown residential, office, suburb residential and hybrid regions.

4.4 Geographical Viewing Features

In the studying of viewing patterns, we notice that the features, i.e., the word vocabulary cuboid built on device type, video genre and time bin, are not equally important in terms of distinguishing the four patterns. For example, it is hard to differentiate the four patterns by referring to view requests during 2am $\sim$ 5:00am (see Fig. 6). We now formally investigate the importance of different features by measuring the distance between their distributions. Specifically, for a region r and a feature $f = (d, c, t)$, let p_r^f represents the probability that users in r watches c-th type of video at time bin t via d-th type of device. Let $Q_i^f(x)$ represents the probability distribution function of p_r^f for all $r \in C_i$. Then, a higher difference between $Q_i^f(x)$ and $Q_j^f(x)$ indicates a higher importance of f in distinguishing pattern i and pattern j. Since the Kullback-Leibler distance is widely used to measure how different two probability distributions are, we use it to define the importance of f in distinguishing pattern i and pattern j, which is formally formulated as follows:

$$I_{i,j}^f = \int_0^1 ln\frac{Q_i^f(x)}{Q_j^f(x)}Q_i^f(x)dx + \int_0^1 ln\frac{Q_j^f(x)}{Q_i^f(x)}Q_j^f(x)dx.$$

The overall importance of f, represented by I^f, is then defined as the sum of $I_{i,j}^f$ between all $(i, j), i \neq j$ pairs.

We use the appearance frequency of f as the approximation of p_r^f, and then estimate $Q_i^f(x)$ using the histogram of $p_r^f, \forall r \in C_i$. Table 4 presents top 20 features to distinguish P1, P2 and P3 (We only consider weekday and weekend, and thus the number of total features is $2 \times 5 \times 2 \times 24 = 480$). To ease the understanding, we aggregate continuous time bins with same device and video genre (despite of their rank). The results illustrate that only PC is involved in all top 20 features, which implies that watching videos on PC is more related to geographical locations than watching videos on mobile devices. With regard to video genre, TV play, show and others are more correlated. Particularly, the high probability of watching these videos on the weekday morning is a strong signal that distinguishes P2 from other

TABLE 4
The top 20 features that distinguish different patterns(aggregated in time).

Time	Device	Video genre
weekday, 9am $\sim$ 12am	PC	Show
weekday, 9am $\sim$ 12am	PC	Others
weekday, 9am $\sim$ 1pm	PC	TV play
weekday, 1pm $\sim$ 2pm	PC	Show
weekday, 2pm $\sim$ 4pm	PC	TV play
weekday, 4pm $\sim$ 5pm	PC	Show
weekday, 8pm $\sim$ 10pm	PC	TV play
weekend, 6pm $\sim$ 7pm	PC	TV play
weekend, 8pm $\sim$ 10pm	PC	TV play
weekend, 11pm $\sim$ 12pm	PC	TV play

TABLE 5
The top 5 features that distinguish P1 (mainly downtown residential regions) and P3 (mainly suburb residential regions).

Time	Device	Video genre
weekday,11pm $\sim$ 12pm	PC	TV play
weekend,11pm $\sim$ 12pm	PC	TV play
weekday,10pm $\sim$ 11pm	PC	TV play
weekend,4pm $\sim$ 5pm	PC	Cartoon
weekend,0am $\sim$ 1am	PC	TV play

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2021.3055860, IEEE Transactions on Big Data

patterns. For example, for more than 60% of regions of P2, which are mainly office regions, the probability of watching shows is higher than 0.004; while less than 20% of regions of other patterns can achieve this value.

Moreover, Table 5 shows the top 5 features that differentiates P1 and P3. We observe that these features mainly correspond to views from 11pm to 1am (next day), which is consistent with the aforementioned phenomenon we have identified, i.e., the relative view at 0am is higher in P1 (usually downtown residential regions) than in P3 (usually suburb residential regions). A surprising finding is that viewing cartoon during 4pm $\sim$ 5pm at weekend ranks the fourth significant feature. Specifically, the probability of this feature is higher than 0.0014 for more than 60% of regions of P3, while only 35% of regions of P1 has the value higher than 0.0014. In fact, we find that viewing cartoons at weekend in other time periods, such as 10am $\sim$ 12am, are included in top 20 features, implying that viewing cartoons at weekend is also a significant signal to distinguish P1 and P3.

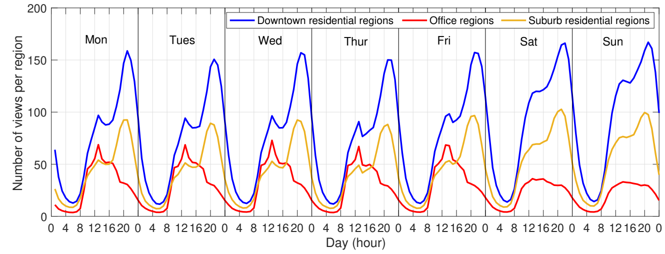
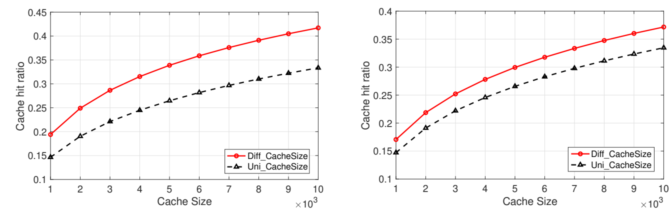


Fig. 13. The temporal distribution of views per region in different kinds of regions.



(a) Downtown residential regions (b) All three kinds of regions

Fig. 14. Cache hit ratio during the night when using different caching methods.

4.5 Summary

We summarize and highlight the characteristics of the identified four viewing patterns as follows:

(1) P1: In this pattern, there are more views in the evening, with a peak at 9pm, and at the weekends. Besides, the number of view requests at midnight (around 0am) is relatively higher. The share of mobile views achieves to the maximum at 8am. Furthermore, this pattern has a strong correlation with downtown residential regions.

(2) P2: In this pattern, most of views are concentrated at daytime, with a peak at 1pm in weekdays. The share of mobile views arrives the maximum at 0am. In comparison to other patterns, the probability of watching videos at morning is higher. The share of views for show is also higher. This pattern has a strong correlation with office regions.

(3) P3: In this pattern, there are more views in the evening, with a peak at 8pm, one hour earlier than P1. In comparison with other patterns, the probabilities of watching cartoons at weekends, and watching movies via mobile devices are higher. This pattern has a strong correlation with suburb residential regions.

(4) P4: In this pattern, there are more views in the evening, but the ratio between two daily peaks is higher than that of P1 and P2. Besides, the probability of watching TV play is higher. This pattern has a strong correlation with hybrid regions that contain both office locations and residential areas.

4.6 Implications

We discuss some important implications based on our findings. Our analysis reveals that the viewing behaviors have a positive correlation with the geographical contexts, which can be applied to video content delivery and recommendation.

For example, we observe that in the downtown residential regions most views appear at night, as shown in Fig. 13. In this case, network providers should push more popular videos to caches deployed near the users in advance to alleviate the heavy network loads. To illustrate it, we design two methods for cache content selection. The first is *Uni_CacheSize* method that allocates the same size of cache to each kind of regions, while the second method called *Diff_CacheSize* allocates different cache sizes according to the view demand in corresponding regions. To be specific, given the total cache size D, the allocated cache size d_i is proportional to the view demand s_i in i-th kind of regions, which is expressed as $d_i = Ds_i / \sum_i s_i$. Since we focus on the potential benefits of each method, it is assumed that the view demand is regarded as a prior. In practical systems, the view demand can be predicted from historical viewing records, which is not our focus in this paper. With these two methods, we first calculate the required cache size and select top popular videos for the caches deployed in regions during 7pm $\sim$ 12pm. Then, we compare the cache performance in terms of cache hit ratio when using different methods. The results are shown in Fig. 14, and it is observed that for *Diff_CacheSize* method, its cache performance achieves the great improvement in downtown residential regions. When 8000 videos are cached, its hit ratio is 26% higher than the *Uni_CacheSize* method. Moreover, for all 3 kinds of regions, *Diff_CacheSize* method also obtains the good cache performance. These results imply that it is beneficial to efficiently allocate cache resource based on the view demand in different kinds of regions. Further, we analyze the potential benefits of caching different types of videos during 7pm $\sim$ 12pm in downtown residential regions. Ideally, when all the videos of a specified type viewed during such time period are stored in the cache, the optimal cache performance can be obtained. We compute its cache hit ratio, and plot the result in Fig. 15. We find that TV play achieves the highest hit ratio. Although it is difficult to predict and store all the videos viewed in future in real systems, it still provides an important insight for network providers to take effective measures to approach the optimal cache performance, *e.g.*, network providers can cooperate with the content providers

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2021.3055860, IEEE Transactions on Big Data

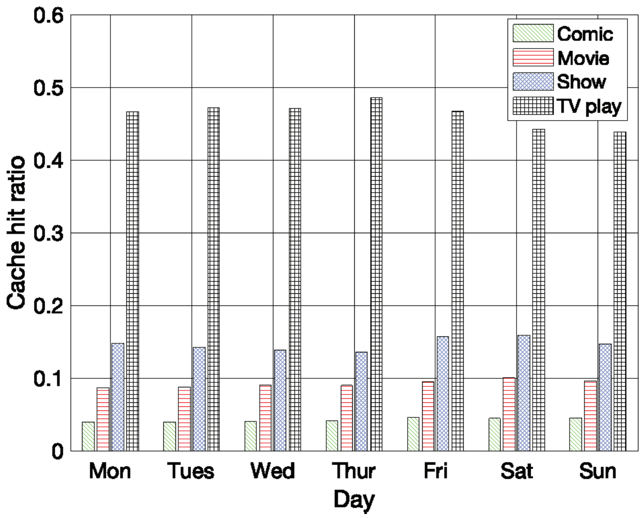


Fig. 15. Cache hit ratio during the night in downtown residential regions.

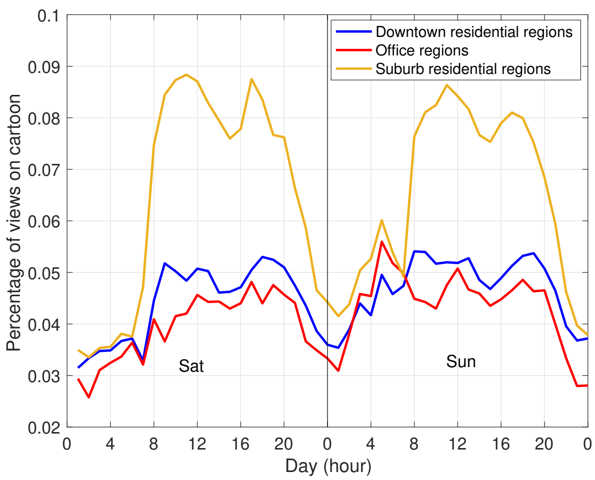


Fig. 16. The percentage of hourly views on cartoon in different kinds of regions.

to proactively cache their uploaded popular videos.

In addition, it is observed that users in suburb regions prefer watching cartoons at weekends, as shown in Fig. 16. Thus, for network providers, it is better to place more popular cartoons in the caching appliances in this period of time. If all the cartoons viewed at the weekend are cached in advance, the hit ratio in suburb residential regions obtains at least 50.7% higher than that in other kinds of regions. For content providers, when optimizing their recommender systems, they should actively recommend more popular cartoons to users who visit these regions.

4.7 Limitations

Our findings are obtained based on the dataset from a large city of Shanghai, China. It has some limitations on the generalization of the findings. While, our analysis metrics and methods can be applicable to the cases of other cities. As our future work, when the relevant data are available, we will further explore the viewing patterns in other cities by our designed framework, and make the comparisons of findings between different cities to validate the generalization.

5 Related Work

Some works have studied the geographical popularity of videos. Brodersen *et al.* [4] analyze the views of Youtube and illustrate that most views of a video are arising from a confined area, e.g., a country, rather than from a global one. They also analyze how social sharing affects the geographical properties of videos. Li *et al.* [16] investigate the geographic interests on videos provided by PPTV, a VoD service provider in China. Different provinces are used as the basic geographical locations in their work. Crepaldi *et al.* [23] study the location selection of server for video on demand (VoD) service. Lv *et al.* [24] analyze the geospatial features of users' online browsing behavior on different content services especially video services. Inspired by these works, we study the geographical patterns of video viewing behavior inside a city. In our previous work [17], we have studied the spatial popularity and similarity in video consumption in a large city. Unlike it, we use a different angle that focus on how the viewing requests in a geographical region are distributed over various dimensions instead of the video-centric angle.

Some other works [1], [2], [20], [25], [26], [27] study the user behavior and cacheability of videos by analyzing data collected from different video systems. Abrahamsson and Nordmark in [1] analyze the access patterns of a TV-on-demand system and show the overall cacheability is high. Cha *et al.* [12] analyze the statistics of requests and its relationship with the videos' characteristics of Youtube. Ben *et al.* [2] investigate the cacheability of youtube videos in cellular network. Avrachenkov *et al.* [26] propose a distributed cooperative caching algorithm for VoD system in a cellular network. Li *et al.* [14] study user behavior of a mobile video-on-demand system named PPTV. Zhang *et al.* [20] investigate the performance of a hybrid CDN-P2P based VoD system. Different to these works that focus on the behavior of users and properties of videos of one content provider, our work aims at studying the overall geographical video viewing behavior covering the top six most popular video content providers in China using the data collected in an ISP network.

We also note some works address the problem of automatic land usage identification using data related to human daily activities [5], [6], [7], [8], [9], [10], [11]. Most of these works use mobile phone data [7], [8], [11]. Other works use the taxi data [10] or online web data [5], [9]. In this paper, we observe that there is correlation between video viewing behavior and land usage. However, instead of identifying land use, we mainly focus on discovering and understanding geographical video viewing patterns.

6 Conclusion

In this paper, we studied the geographical video viewing patterns in urban neighborhoods. We designed a framework that combines the time, device, and content domain features to automatically characterize and discover the viewing patterns in urban neighborhoods. Using this framework, we conducted a comprehensive measurement study and discovered four distinct viewing patterns, by analyzing a dataset of 84 million viewing video requests collected from a major ISP of China. We find that each pattern has a very specific geographic context.

Furthermore, we analyzed the time, device, and content domain characteristics of each pattern and revealed various interesting observations. For instance, users are more likely to watch shows in office locations than at home, and users in suburb areas are more likely to watch cartoons at weekends than users in downtown areas. Our analysis provides a comprehensive understanding of geographic patterns in video viewing behavior, which is helpful to optimize video content discovery, delivery, and recommendation.

References

[1] H. Abrahamsson and M. Nordmark. Program popularity and viewer behaviour in a large tv-on-demand system. In *Proc. of IMC*, pages 199–210. ACM, 2012.

[2] F. Ben Abdesslem and A. Lindgren. Cacheability of youtube videos in cellular networks. In *Proc. of workshop on All things cellular: operations, applications, & challenges*, pages 53–58. ACM, 2014.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.

[4] A. Brodersen, S. Scellato, and M. Wattenhofer. Youtube around the world: geographic popularity of videos. In *Proc. of WWW*, pages 241–250. ACM, 2012.

[5] J. Cranshaw, R. Schwartz, J. I. Hong, and N. Sadeh. The livehoods project: Utilizing social media to understand the dynamics of a city. In *International AAAI Conference on Weblogs and Social Media*, page 58, 2012.

[6] Z. Cao, S. Wang, G. Forestier, A. Puissant, and C. F. Eick. Analyzing the composition of cities using spatial clustering. In *Proc. of SIGKDD International Workshop on Urban Computing*, page 14. ACM, 2013.

[7] V. Soto and E. Frías-Martínez. Automated land use identification using cell-phone records. In *Proc. of workshop on MobiArch*, pages 17–22. ACM, 2011.

[8] T. Pei, S. Sobolevsky, C. Ratti, S.-L. Shaw, T. Li, and C. Zhou. A new insight into land use classification based on aggregated mobile phone data. *International Journal of Geographical Information Science*, 28(9):1988–2007, 2014.

[9] C. K. Vaca, D. Quercia, F. Bonchi, and P. Fraternali. Taxonomy-based discovery and annotation of functional areas in the city. In *Ninth International AAAI Conference on Web and Social Media*, 2015.

[10] J. Yuan, Y. Zheng, and X. Xie. Discovering regions of different functions in a city using human mobility and pois. In *Proc. of SIGKDD*, pages 186–194. ACM, 2012.

[11] J. L. Toole, M. Ulm, M. C. González, and D. Bauer. Inferring land use from mobile phone activity. In *Proc. of SIGKDD international workshop on urban computing*, pages 1–8. ACM, 2012.

[12] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In *Proc. of IMC*, pages 1–14. ACM, 2007.

[13] D. Karamshuk, N. Sastry, A. Secker, and J. Chandaria. On factors affecting the usage and adoption of a nation-wide tv streaming service. In *Proc. of INFOCOM*, pages 756–764, IEEE, 2015.

[14] Z. Li, G. Xie, M. A. Kaafar, and K. Salamatian. User behavior characterization of a large-scale mobile live streaming system. In *Proc. of WWW*, pages 307–313, ACM, 2015.

[15] Yasuaki Oishi, A region-dividing approach to robust semidefinite programming and its error bound. In *American Control Conference*, pages 7–pp, IEEE, 2006.

[16] Z. Li, G. Xie, J. Lin, Y. Jin, M.-A. Kaafar, et al. On the geographic patterns of a large-scale mobile video-on-demand system. In *Proc. of INFOCOM*, pages 397–405. IEEE, 2014.

[17] H. Yan, J. Liu, Y. Li, D. Jin and S. Chen. Spatial Popularity and Similarity of Watching Videos in a Large City. In *Proc. of Globecom*, pages 1–6. IEEE, 2016.

[18] M. Mehaffy, S. Porta, Y. Rofe, and N. Salingaros. Urban nuclei and the geometry of streets: The emergent neighborhoods model. *Urban Design International*, 15(1):22–46, 2010.

[19] S. Porta, P. Crucitti, and V. Latora. The network analysis of urban streets: a dual approach. *Physica A: Statistical Mechanics and its Applications*, 369(2):853–866, 2006.

[20] G. Zhang, W. Liu, X. Hei, and W. Cheng. Unreeling xunlei kankan: understanding hybrid cdn-p2p video-on-demand streaming. *IEEE Transactions on Multimedia*, 17(2):229–242, 2015.

[21] "OpenStreetMap." https://www.openstreetmap.org/ ,Accessedon2016.9.28.

[22] "Moore Neighborhood Tracing Algorithm." https://en.wikipedia.org/wiki/Moore_neighborhood,Accessedon2016.9.28.

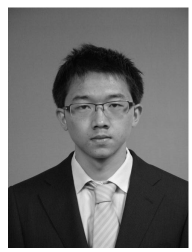
[23] T. Crepaldi, N. L. S. d. Fonseca and E. C. Xavier. Selection of Servers for Video on Demand Service over Hybrid Cloud. In*Proc. of IEEE ICC*, pp. 1–7, 2018.

[24] Lv, Q., Qiao, Y., Zhang, Y. et al. Measuring Geospatial Properties: Relating Online Content Browsing Behaviors to Users' Points of Interest. In*Wireless Personal Communication*, pp. 1469–1498, 2018.

[25] M. Irshad, A. Anand and M. Bisht. Modelling Popularity Dynamics Based on YouTube Viewers and Subscribers. In *International Journal of Mathematical, Engineering and Management Sciences*, vol. 4, no. 6, pp. 1508–1521, 2019.

[26] K. Avrachenkov, J. Goseling and B. Serbetci. Distributed Cooperative Caching for VoD with Geographic Constraints. In *Proc. of IEEE WiOPT*, pp. 1–8, 2019.

[27] R. Zhang, S. Jia, Y. Ma and C. Xu. Social-Aware D2D Video Delivery Method Based on Mobility Similarity Measurement in 5G Ultra-Dense Network. In *IEEE Access*, vol. 8, pp. 52413–52427, 2020.



Jiaqiang Liu received his B.S. degree in Electronic Engineering from Tsinghua University, Beijing, China, in 2012, and now he is a PhD student in Tsinghua University. His research interests include software defined network, data center network, network function virtualization and cloud computing, etc.

Huan Yan received the B.S. degree in electronic information engineering from Nanchang University, Nanchang, China, in 2009, and M.S. degree in communication engineering from Hangzhou Dianzi University, Hangzhou, China, in 2012. He is currently working toward the Ph.D. degree in the department of electronic engineering of Tsinghua University, Beijing, China. His current research interests include network big data and software-defined network.

Yong Li (M'09–SM'16) received the B.S. degree in electronics and information engineering from Huazhong University of Science and Technology, Wuhan, China, in 2007 and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2012. He is currently a Faculty Member of the Department of Electronic Engineering, Tsinghua University.

Dr. Li has served as General Chair, TPC Chair, TPC Member for several international workshops and conferences, and he is on the editorial board of two IEEE journals. His papers have total citations more than 3600 (six papers exceed 100 citations, Google Scholar). Among them, eight are ESI Highly Cited Papers in Computer Science, and four receive conference Best Paper (run-up) Awards. He received IEEE 2016 ComSoc Asia-Pacific Outstanding Young Researchers and Young Talent Program of China Association for Science and Technology.

Dmytro Karamshuk is a Postdoctoral Research Associate with King's College London, London, U.K. His research is focused on studying behavior of Internet users to assist design of future generation communication networks.

Nishanth Sastry (S'09–M'11) received the Bachelor's degree from Bangalore University, Bangalore, India, in 1999, the Master's degree from The University of Texas at Austin, Austin, TX, USA, in 2001, and the Ph.D. degree from the University of Cambridge, Cambridge, U.K., in 2011, all in computer science. He is a Senior Lecturer with King's College London, London, U.K. He has over 6 years of experience in the industry (Cisco Systems, Bangalore, India, and IBM Software Group, Westford, MA, USA) and industrial research labs (IBM T. J. Watson Research Center, Cambridge, MA, USA). His work has ranged over several layers of the network stack, and he is currently involved in building better networked systems by harnessing social network information. Dr. Sastry's honors include a Best Undergraduate Project Award, a Best Paper Award from the Computer Society of India, a Yunus Innovation Challenge Award at the Massachusetts Institute of Technology IDEAS Competition, a Benefactor's Scholarship from St. John's College, Cambridge, U.K., a Cambridge Philosophical Society Research Studentship, a Cisco Achievement Program Award, and several awards for his work at IBM.

Di Wu is an associate professor and associate department head in the Department of Computer Science, Sun Yat-sen University, Guangzhou, China. He received the B.S. degree from the University of Science and Technology of China in 2000, the M.S. degree from the Institute of Computing Technology, Chinese Academy of Sciences, in 2003, and the Ph.D. degree in computer science and engineering from the Chinese University of Hong Kong in 2007. From 2007 to 2009 he worked as a postdoctoral researcher in the Department of Computer Science and Engineering, Polytechnic Institute of NYU, advised by Prof. Keith W. Ross. He is the co-recipient of the IEEE INFO- COM 2009 Best Paper Award.

Depeng Jin received the B.S. and Ph.D. degrees from Tsinghua University, Beijing, China, in 1995 and 1999 respectively both in electronics engineering. He is a professor at Tsinghua University and the chair of Department of Electronic Engineering. Dr. Jin was awarded National Scientific and Technological Innovation Prize (Second Class) in 2002. His research fields include telecommunications, high-speed networks, ASIC design and future Internet architecture.