

Spatial Popularity and Similarity of Watching Videos in Large-Scale Urban Environment

Huan Yan, Jiaqiang Liu, Yong Li¹, Senior Member, IEEE, Depeng Jin, and Sheng Chen, Fellow, IEEE

Abstract—With the popularity of watching mobile videos, a major form of multimedia contents, many works focus on the geographic features of user viewing behaviors, but few study them in the context of an entire metropolitan city. Different regions of a large city have different intensity of economy activities with respect to their different distances to the downtown, and how this will influence video popularity and similarity is still unclear. To quantitatively study the spatial popularity and similarity of watching videos in a large urban environment, we collect a dataset with two-month video view requests from the largest network provider in Shanghai, containing the top six content providers, and study the spatial features of video access in regions of different scales. We find that: 1) video popularity and similarity exist at different scales of city division; 2) the concentration of video popularity becomes higher as the region is closer to downtown; and 3) when comparing the regions of same scale, the similarity of popular videos becomes lower as the region is farther away from the downtown. Finally, we correlate our findings with cache deployment, advertising, and video recommendation to illustrate the implications.

Index Terms—Video spatial popularity, video spatial similarity, urban computing, city division.

I. INTRODUCTION

IN RECENT years, the advance in network technologies and proliferation of mobile devices boosted the popularity of multimedia contents. Online videos, as a major form of multimedia contents, have attracted a significant number of views [1], [2]. Content providers (CPs) such as YouTube, Netflix that offer online videos are willing to have a good understanding of video watching behaviors, which gives an important guideline on how to improve the quality of video services and maximum the revenue. To deal with it, the analysis based on real-world dataset, as an important direction to understand human behaviors in multimedia systems, can be also applied to characterize the patterns of video views. Many recent works collect viewing data from CPs to study the time-varying popularity of videos [12], [13], and some of them

utilize the temporal feature to predict the video popularity [4]. Different from them, our study focuses on the spatial characteristics of watching videos, which is greatly important for networking. For example, most CPs use content distribution networks to improve user experience of video watching. Such systems depend on the global distributed set of cache servers that store most popular video contents, and thus significantly reduce the transmission delay. However, since the viewing patterns may vary across geographic regions, how the cache should be deployed in different regions becomes a challenging question. Therefore, it is important to study geographical consumption patterns of videos.

Under such context, understanding the spatial popularity and similarity of watching videos with the evolution of mobile networks is particularly important. Raman *et al.* [33] identified the similarity patterns of TV streaming consumption across the regions of varying demographics, and designed Wi-Stitch for efficient content sharing. Many other works [5], [6] focus on the geographic popularity of video views in the world and country scale. In this paper, we seek to explore how the popularity and similarity of watching videos are at different scales of urban environment. Obtaining the relationship between different regions and user viewing behaviors helps explore whether the activities of economy (*i.e.*, most economically developed regions locate in the downtown) will influence the popularity and similarity of user views, which also benefits content providers, enabling them to provide better video service for example.

Thus, associating with different scales and different regions in a metropolitan area, we focus on two key problems that need to be investigated:

Problem I: How is the video popularity distributed? At different scales, due to diverse user interests, the concentration of video popularity could be different. Thus, the influence of the region size needs to be investigated. On the other hand, between the downtown and the suburb, whether there exist the differences of the concentration is also a key consideration.

Problem II: How is video request similarity distributed? Under a given scale, how do we characterize the difference of video requests between different regions? Meanwhile, what is the difference of the similarity of user views between the downtown and suburb? These are the issues that must be addressed.

These two problems are also critical in multimedia systems since various applications such as online advertising and multimedia content recommendation can benefit greatly from

Manuscript received August 11, 2017; revised January 6, 2018, March 12, 2018, and March 14, 2018; accepted March 15, 2018. Date of publication March 20, 2018; date of current version June 8, 2018. The associate editor coordinating the review of this paper and approving it for publication was M. Mellia. (Corresponding author: Yong Li.)

H. Yan, J. Liu, Y. Li, and D. Jin are with the State Key Laboratory on Microwave and Digital Communications, Beijing National Research Center for Information Science and Technology, Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China (e-mail: liyong07@tsinghua.edu.cn).

S. Chen is with the School of ECS, University of Southampton, Southampton SO17 1BJ, U.K.

Digital Object Identifier 10.1109/TNSM.2018.2817549

answering them. In order to systematically study these two problems, we analyze a dataset containing two-month real-world video requests. The data is collected at the gateway of a major Internet service provider (ISP) in a large city in China, Shanghai. More than 200 million view requests for videos recorded come from top six most popular content providers. We divide the city into non-overlapping regions of four scales by using cellular network infrastructure of the same ISP. At the same time, we define topic as the set of videos of episodes/clips belonging to the same program, which aggregates videos from different content providers. Next, we discuss spatial features of watching videos related to the above introduced two problems by defining metrics of popular topic number, view concentration and popular topic similarity to characterize them. We have an in-depth analysis based on these metrics, and summarize major findings corresponding to the two problems as follows:

- Well-known Pareto principle of video popularity is validated to exist at any size of scales in the city. Also, we observe that the concentration of video popularity becomes higher as the size of scale increases, and it becomes higher near the downtown than in the suburb. In particular, associating with different video types such as movies, cartoons, show and TV series, we find that in the downtown the concentration on TV series and shows is higher while opposite is true for the other two types.
- We observe that the similarity between the sets of popular videos of any two regions of small scale is low, but the similarity increases with the scale. Also, the regions nearer the downtown have much more similar popular video requests than the region in the downtown. Furthermore, between the regions in the downtown and in the suburb, the similarity of movies is lowest while the similarity of shows is highest.

Finally, we exploit our findings for potential applications on cache deployment strategies, advertisement and video recommendation. Taking the cache deployment as an example, according to our results, we need to cache different contents with the consideration of local video popularity.

The rest of this paper is organized as follows. After describing the datasets and basic data processing in Section II, we introduce the methodology and metrics in Section III. Then, in Section IV, we analyze the concentration and similarity of video popularity, and then discuss the obtained results. After that, we discuss the implications of our findings in Section V. After the survey of related works in Section VI, we finally conclude the paper in Section VIII.

This paper is extended from our previous work published in [29] to conduct a more detailed and insightful analysis. First, we add the analysis of the necessity of video aggregation, and explore whether view requests are concentrated in a small fraction of regions under each scale. Then, we add the analysis of relationship between popular topics with different types and their views in the city. To illustrate the spatial similarity of popular topics, we add the analysis of the geographic distribution of most popular topics in regions under small scale. Importantly, based on our findings in the studying of the spatial popularity and similarity of watching videos, we

add and classify three important implications of exemplified applications.

II. DATA COLLECTION AND PROCESSING

A. Data Collection

In this paper, we analyze the logs of deep packet inspection appliances deployed at the service gateways of a major ISP network in Shanghai, one of the largest cities in China. The logs consists of individual entries, where each entry represents a view requests for a specific video, and records the information of user ID, location, and URL for the request. We further obtain the context of the requested video, including its content provider, ID, type, and description of name, by crawling the URL from Web. From the perspective of ISP, this dataset is comprehensive to study the spatial popularity and similarity of watching videos in a city, because all view requests, despite of the content providers, need to go through the ISP gateway. Overall, our dataset is collected during Nov. 1 and Dec. 31, 2014, and it covers 200 million view requests, 9 million users, and 7 million videos from top six most popular content providers in China.¹

In addition, as we aim to investigate the influence of the region size on the concentration and similarity of video popularity, we further collect information of mobile network infrastructure in Shanghai by the same ISP, including the ID, location and associated district of each base station (BS), base station control (BSC) and mobile switch center (MSC). Combining these two datasets of logs of view requests and mobile network structure enables us to characterize the user interests on specific video topics in a specific region and study the spatial features of video popularity and similarity under different scales.

Ethics: Our study aims at providing a good understanding of spatial popularity and similarity on video consumption in a city. We obtain the dataset by collaborating with ISP who has anonymized user ID and processed the URLs. At the same time, this work has received the approval from our university.

B. Aggregating Videos by Topic

Videos in our dataset come from six content providers, and some of them could be duplicated or aliased, which should be aggregated to eliminate their inference on popularity analysis. Furthermore, instead of a single video, *e.g.*, one episode of a TV series, we are more interested in the spatial popularity of a set of videos with the same topic, *e.g.*, videos of the same TV series. To this end, we aggregate videos by topic, defined as a set of videos or episodes and/or clips in the same TV series, movie, show or cartoon program. Specifically, we crawl the program lists of TV series, cartoon, show and movie from the websites of six content providers. We then use the title of each program as the keyword of one topic, and classify the videos into topics by matching these keywords with the

¹We release some anonymized dataset under our collaborator's approval, including the access logs with the timestamp, anonymized user id, location in terms of longitude and latitude, website id and video id. They are available in <https://drive.google.com/drive/folders/1TM2XcnDwndGg6DnXpY0k01-x0EvfVOI-?usp=sharing>.

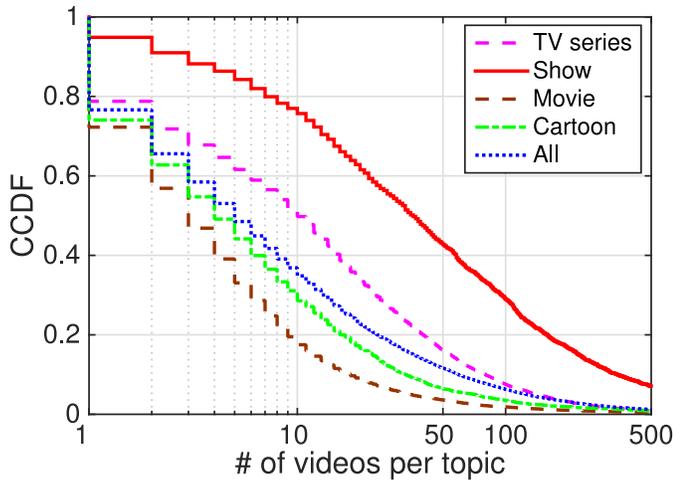


Fig. 1. The CCDF of the number of videos per topic.

TABLE I
THE STATISTICS OF VIEW REQUESTS DATA

	TV series	Show	Movie	Cartoon
# of topics	8533	2007	11193	2451
# of videos (10^6)	0.73	0.37	0.22	0.08
# of views (10^6)	127.9	37.2	24.2	8.7

name strings. Specially, in terms of movies, we use the keyword to group the feature, the clips and trailers of the same movie.

Finally, we obtain 24,184 topics, which covers 77% of overall videos among the video types including TV series, show, movie, and cartoon.² Table I summarizes some basic statistics of the topics with these types. To illustrate that the topic provides an appropriate granularity to aggregate videos, we plot the complementary cumulative distribution function (CCDF) of the number of videos in each topic in Fig. 1. We can observe that 90% of topics have less than 50 videos, which suggests that the topic provides a moderate granularity for the aggregation. Specifically, the topic belonging to show usually has more videos due to a large number of clips; while the topic of movie contains the least number of videos. Further, we plot the CCDF of the number of videos per topic in terms of providers in Fig. 2. Among different CPs, we observe that at least 78% of topics have less than 50 videos. Although there are some differences in the distributions, they are not very significant.

We further study the involvement of content providers in each topic to illustrate the necessity of video aggregation. Specifically, we first classify the topics into six classes according to the number of providers they involve, *e.g.*, class two CPs mean that the videos in this topic come from two content providers, and then calculate the fraction of topics, videos, and views that each class accounts for. As shown in Fig. 4, although the number of topics covering one provider occupies high percentage in each type, the topics with at least two providers account for much more videos and views. For example, considering TV series in Fig. 4 (a), there are nearly 80%

²The rest of videos cannot be classified because their names are usually about the content of the episodes and/or clips and do not contain the title of the corresponding program. These videos are removed in our following analysis.

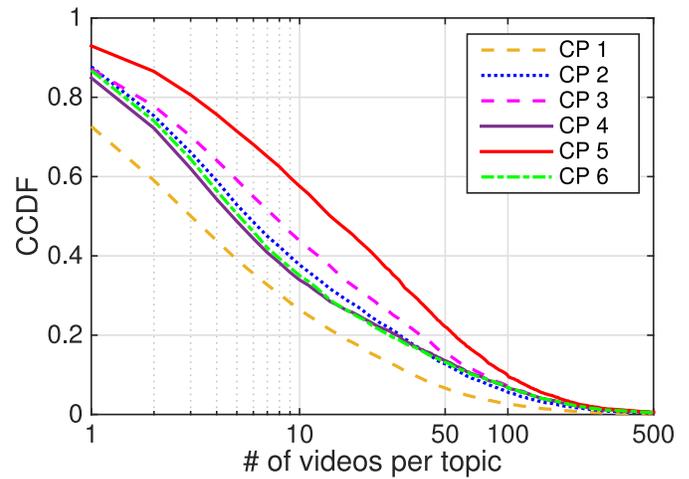


Fig. 2. The CCDF of the number of videos per topic in terms of CPs.

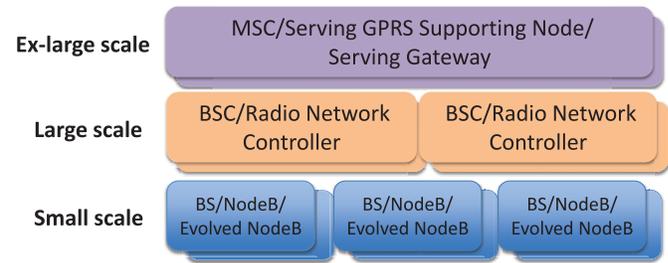


Fig. 3. The communication entities of the mobile network corresponding to the defined scales.

TABLE II
THE CONTENT PROVIDERS (CPs) IN TOP 5 VIEWED TOPICS, WHERE SMALLER RANK NUMBER REPRESENTS THE TOPIC WITH HIGHER VIEWS

Rank	CP1	CP2	CP3	CP4	CP5	CP6
1		✓				
2	✓				✓	✓
3	✓	✓	✓		✓	
4	✓		✓			✓
5	✓		✓	✓	✓	✓

of views from the topics with 6 providers, which indicates a large overlap between the providers. However, in Fig. 4 (d), we observe that the fraction of topics belonging to cartoons is least in 6 providers, which exhibits the largest diversification in the catalog of the providers. Overall, about 35% of topics involving over one provider obtain 90% of videos and 95% of views. This suggests that from the perspective of ISP, the videos offered by any single provider are insufficient to capture user behaviors on a specific topic. This observation can be further confirmed by analyzing the content providers involved in the requests for the top viewed topics: as shown in Table II, there are four most viewed topics containing at least three providers, which cover all six providers. In summary, these results show that it is appropriate to aggregate videos from multiple providers into topics, which helps to study video popularity and similarity.

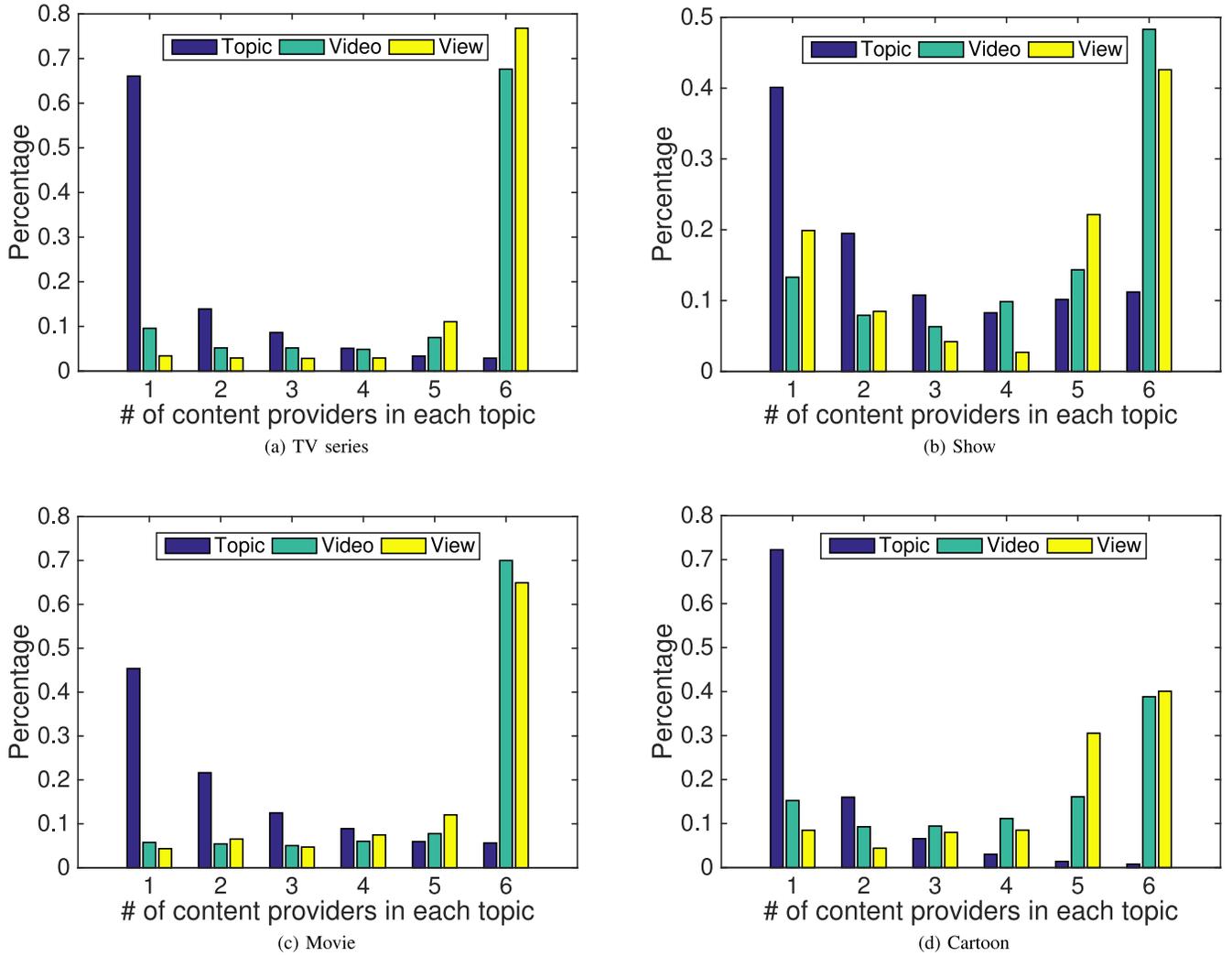


Fig. 4. The fraction of topics, videos, and views occupied by six classes of topics belonging to four types involving different numbers of CPs.

C. City Division

To study the spatial popularity and similarity of watching videos at different scales, we divide the city into multiple non-overlapping regions with different sizes. In general, there are two major approaches to obtain the fine-grained regions: using the city structure, and relying on the network infrastructure. Further, regardless of which approach, the formed regions can be easily labeled as suburb and downtown in an artificial manner. Because of potential benefits on cache deployment, advertisement and video recommendation (details in Section V), we choose the second approach to study spatial viewing patterns in this paper. Specifically, as the view requests are first sent to the BSs or access points (APs), we obtain the coverage area of each BS/AP and regard them as basic regions. Then we build larger regions by aggregating adjacent basic regions according to the structure of the ISP's network infrastructure.

As for each base station we only know its location, we use Voronoi diagram to obtain their coverage areas. Formally, given the locations of N BSs', denoted by $\{l_1, l_2, \dots, l_N\}$, the Voronoi diagram gets the coverage areas for the BSs as $\{R(l_1), R(l_2), \dots, R(l_N)\}$, where $R(l_i)$ denotes the coverage

area of the i -th BS, and any location p inside $R(l_i)$ satisfies that the Euclidean distance between p and l_i is smaller than that between p and l_j for any $l_j \neq l_i$.

To obtain larger regions, we aggregate coverage areas of adjacent base stations by referring to the cellular network infrastructure. As shown in Fig. 3, the entity in a larger scale controls multiple entities in the smaller scale. For example, a MSC controls multiple BSCs, while a BSC controls multiple BSs. Based on these relationships, we aggregate BSs belonging to the same entity to form larger regions. Besides, to make our study more comprehensive, we also construct larger regions by aggregating BSs belonging to the same geographical district. Overall, we divide the city by four scales: the division using BS's coverage area as the small scale, and the division by aggregating adjacent BSs in the same district as medium scale, the division by aggregating BSs controlled by the same BSC and MSC as the large and extra-large scales respectively. Finally, there are 4137, 93, 18 and 7 regions in the small, medium, large and extra-large scales, respectively. Fig. 5 shows their geographic distribution, where a smallest region represents the coverage area of a BS, and the adjacent smallest regions with the same

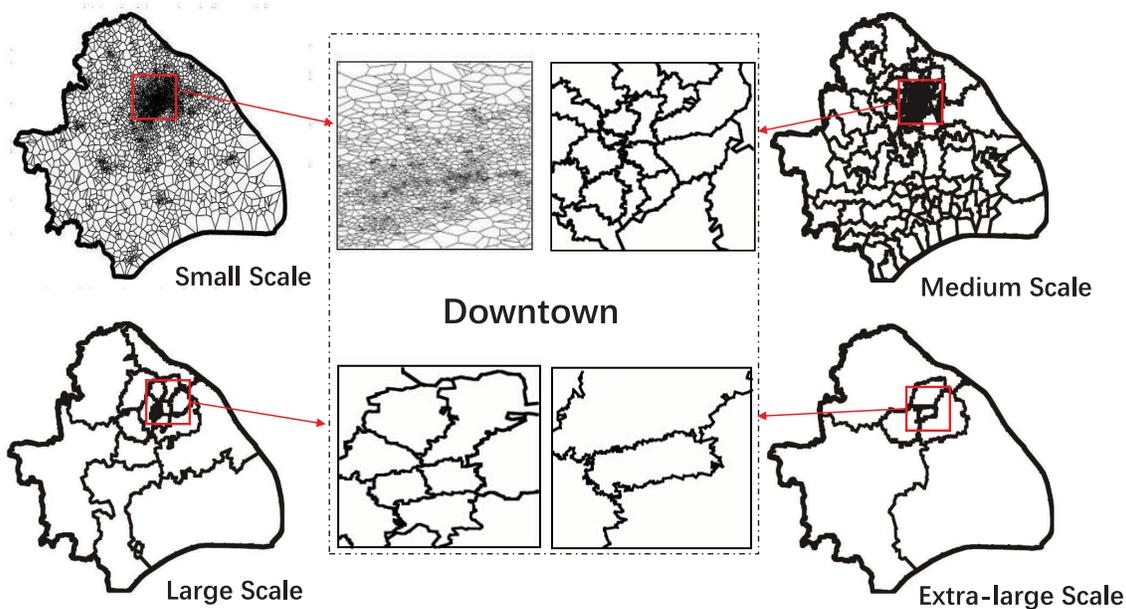


Fig. 5. The geographic distribution under different scales.

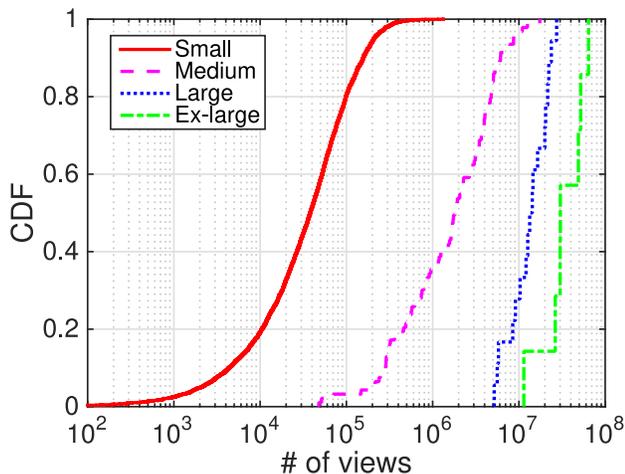


Fig. 6. The number of views in regions of different scales.

color represents that they are aggregated into the same larger region.

We can divide the city by four scales of different sizes using above approach. To validate the rationality of such division, we focus on two questions: whether the number of the aggregated views in regions of one scale is significantly different to another scale; and whether view requests are concentrated in a small fraction of regions under each scale. To answer above questions, we plot the CDF of views under four scales in Fig. 6. We observe that the number of views in regions of the same scale exhibits the concentration that corresponds to a relatively small range. For example, 20% and 80% of views are within one order of magnitude under small scale, which are 10488 and 98795 respectively. On the other hand, the median number of views has different orders of magnitude that is from tens of thousands to tens of millions.

To quantitatively analyze view distribution over regions of the same scale, we calculate the fraction of views accounted

by different fractions of regions that attracts highest views. We show the results in Fig. 7. From the results, we observe that the percentage of video views in each region has little difference under a given scale, which is inconsistent with the Zipf's law that denotes the standard geographical concentration [20]. This implies that the views do not exhibit geographical concentration under different scales. This observation also stands when we count the number of views for different video types separately. Therefore, we conclude that it does not exhibit the geographical concentration under different scales regardless of video types.

Associating with the above two aspects, we conclude that dividing the city at the defined four scales provides an alternative method to study the spatial popularity and similarity of watching videos.

III. METHODOLOGY AND METRICS

In order to quantitatively answer the two problems regarding the spatial characterizations of video popularity and similarity, we investigate the view patterns related to them. Before introducing the methodology and related metrics, we first describe the notations utilized below.

Let $\{p_1, p_2, \dots, p_M\}$ denote the topic set, where M is the number of topics. The number of topics that receive at least one view in region i is represented by n_i . For each region, we sort the topics in descending order by the number of views, and $u_{i,k}$ denotes the number of users that watch the k th sorted topic whose view number is denoted as $v_{i,k}$. $v_i = \sum_{j=1}^M v_{i,j}$ represents the total number of views of region i . We are now ready to describe the methodology used in our study.

A. Problem I: Video Popularity Concentration

Users in some regions may be interested in only a few popular topics, while for other regions video popularity may cover

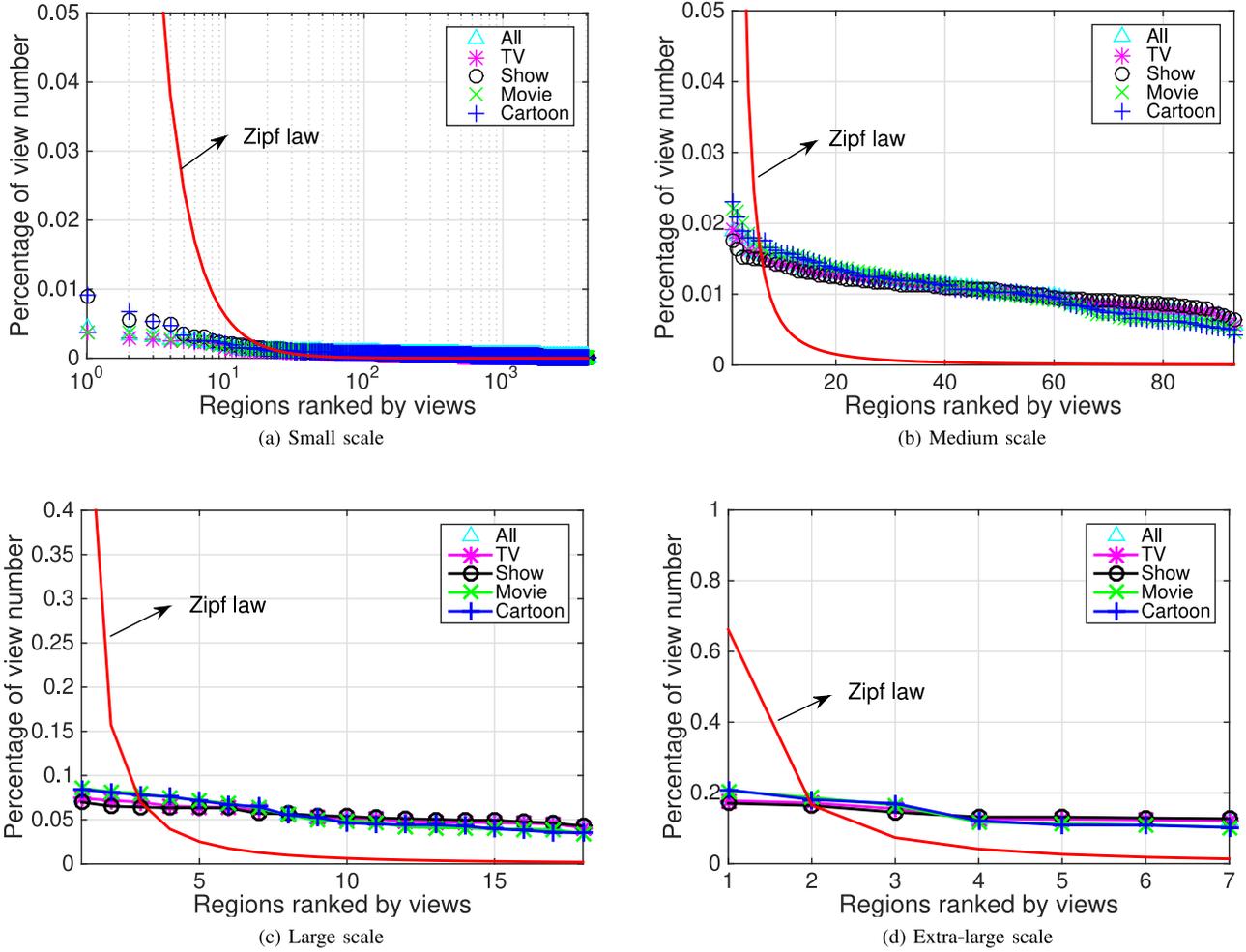


Fig. 7. The percentage of view number by each user as a function of top viewed regions.

a large number of topics. Therefore, in order to study the concentration of video popularity, we investigate the distribution of the number of popular topics over regions.

Intuitively, the number of views can be used as a popularity metric, so we define the topics with more than a given number of views as popular ones. Specifically, given a desired view number n , we define **popular topic number** of region i , denoted by $P_i(n)$, as follows:

$$P_i(n) = \sum_{j=1}^M I(v_{i,j} - n), \quad (1)$$

where $I(x) = 1$ if $x \geq 0$ and otherwise $I(x) = 0$.

On the other hand, to eliminate the influence of the specified view number, we use the fraction of the top viewed topic to define popular topic. Specifically, given f ($0 < f < 1$), we define the most viewed topics with at least f of the total topics as popular topics. Thus, we denote the **popular topic number** of region i as $Q_i(f)$, computed by:

$$Q_i(f) = \lfloor f \times n_i \rfloor. \quad (2)$$

A higher value of $P_i(n)$ or $Q_i(f)$ indicates that region i has more popular topics.

In general, if a large fraction of views concentrates on a few popular topics, it represents the high concentration of

video popularity. On the other hand, if the views are uniformly distributed to many different topics, the concentration is low. Therefore, in order to study the spatial characterization of video popularity, we investigate the concentration of views on popular topics under each scale. To quantify it, we define **view concentration** of popular topics of region i as follows:

$$\begin{aligned} C_i^P(n) &= \frac{1}{v_i} \sum_{j=1}^{P_i(n)} v_{i,j}, \\ C_i^Q(f) &= \frac{1}{v_i} \sum_{j=1}^{Q_i(f)} v_{i,j}. \end{aligned} \quad (3)$$

In the above definition, a higher value of $C_i^P(n)$ or $C_i^Q(f)$ indicates a higher view concentration and, consequently, a higher concentration of video popularity.

To eliminate the influence of user number on the number of views in a region, we define average views of popular topics by each user in region i as follows:

$$A_i(f) = \frac{1}{Q_i(f)} \sum_{j=1}^{Q_i(f)} v_{i,j}/u_{i,j}. \quad (4)$$

According to the above definition, a higher value of $A_i(f)$ indicates the higher view concentration of video popularity by each individual user.

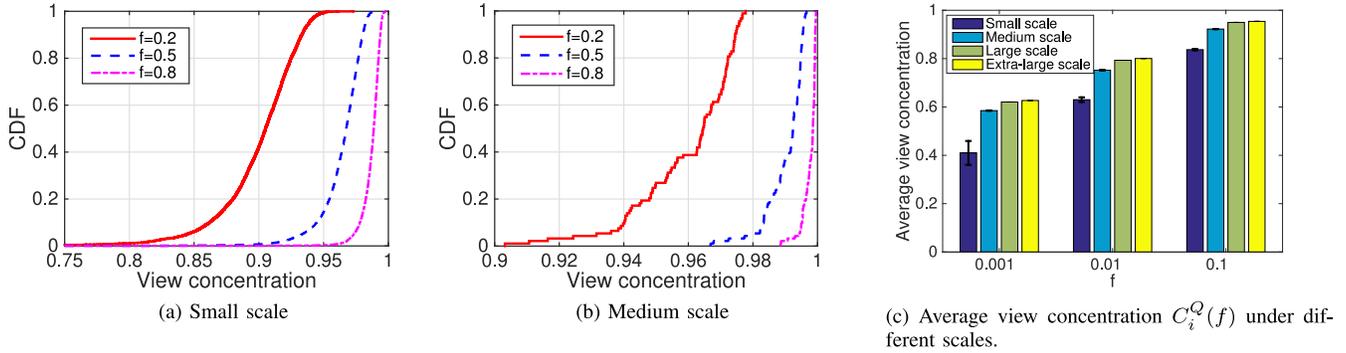


Fig. 8. View concentration $C_i^Q(f)$ under different scales.

B. Problem II: Video Request Similarity

Video requests in a region concentrate on the most popular topics. If different regions have similar set of popular topics, there is no significant difference with video popularity. Therefore, in order to answer the question of how similarity of popular video requests is, we study the similarity between the popular topics of different regions. Note that Jaccard similarity coefficient is a well defined measure to characterize the similarity between two sets, specifically, it equals to 1 if the two sets are equal, and it is smaller than 0.5 if they have low similarity. We also use Jaccard similarity coefficient to quantify the similarity between different sets of popular topics. Let Φ_i^k denote the set of top k ranked popular topics of region i . Then, we define **popular topic similarity** of region i and region j , where $i \neq j$, as follows:

$$J_{i,j}^k = \frac{|\Phi_i^k \cap \Phi_j^k|}{|\Phi_i^k \cup \Phi_j^k|}. \quad (5)$$

According to the definition, users in regions i and j have different popular topic requests if $J_{i,j}^k$ is small.

IV. SPATIAL POPULARITY AND SIMILARITY ANALYSIS

In this section, we study the view patterns by utilizing the introduced methodologies and defined metrics with the purpose to answer the two key problems about spatial popularity and similarity of watching videos.

A. Concentration of Video Popularity

In order to study Problem I of how concentration of video popularity is in regions of four scales, we investigate the popularity distribution of topics and user view concentration on video topics in the city.

As shown in Fig. 10(a), we observe that there is a strong log-log linear relationship among the topics sorted by views. Specifically, the number of views on the topic with its rank k is proportional to $\beta k^{-\alpha}$ that indicates the power law distribution of topic's popularity. This reveals that most views concentrate on the most popular video topics. To further demonstrate it, we plot the CCDF of views by sorting the topics with their views in descend order in Fig. 10(b), where we find that top 10% topics attract over 80% of views regardless of TV series, show,

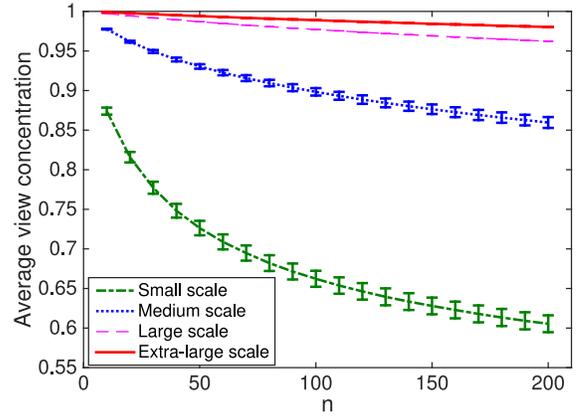


Fig. 9. Average view concentration $C_i^P(n)$ under different scales.

movie or cartoon. Furthermore, the fraction of overall views on top 10% topics is larger than 90%. We can conclude that the concentration of video popularity follows the well-known Pareto principle of video popularity in the city.

To study the concentration of video popularity under four different scales, we analyze the view concentration under these scales according to (3). As shown in Fig. 8(a), view concentration in each region of small scale is over 0.75 when $f = 0.2$. The corresponding value is higher under the medium scale seen in Fig. 8(b). When considering different scale in Fig. 8(c), we observe that top 10% of popular topics attract more than 80% of views, while in the large and extra-large scales, top 10% of popular topics have more than 90% of views. This shows that the results conform to the Pareto principle under different scales. Further, we fit the distribution of view concentration $C_i^Q(f)$ under different scales, and then obtain the shape parameters in Pareto distribution as 0.581, 0.7057, 0.7894 and 0.8066 in small, medium, large, extra-large scales respectively. Thus, *the larger scale implicates the higher concentration of video popularity*. Similarly, in Fig. 9, we observe that in the medium scale, popular topics whose view number exceeds 200 attract more than 70% of views, while in the small scale, popular topics whose view number exceeds 50 have more than 80% of views. Therefore, we validate that *the Pareto principle of video popularity applies to different scales*.

In order to explore their spatial distribution of popular topics, we count the popular topic number according

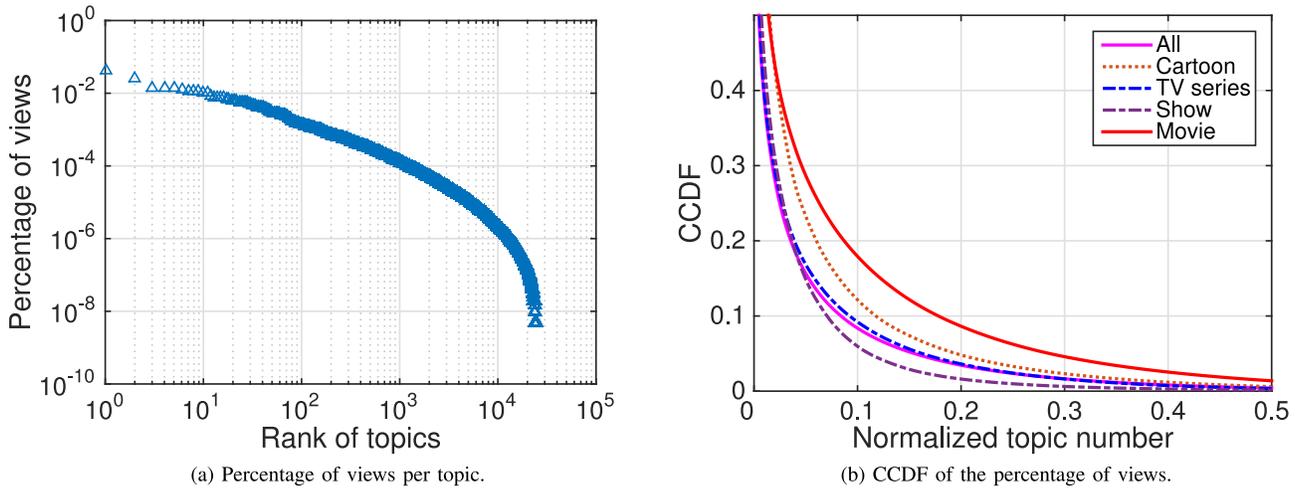


Fig. 10. The distribution of the number of views per topic.

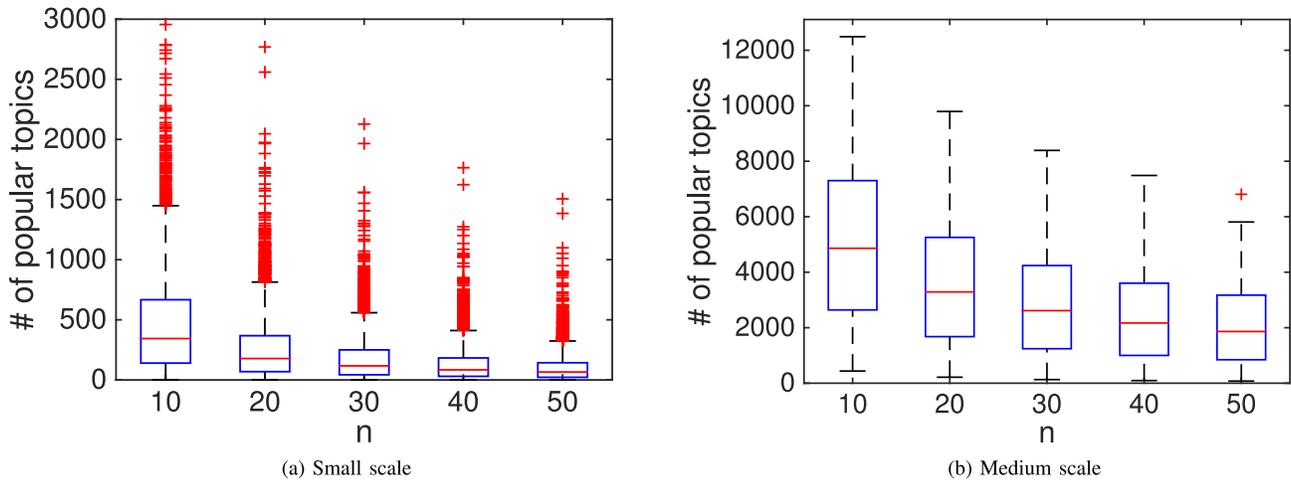


Fig. 11. The popular topic number P_i as a function of the number of views.

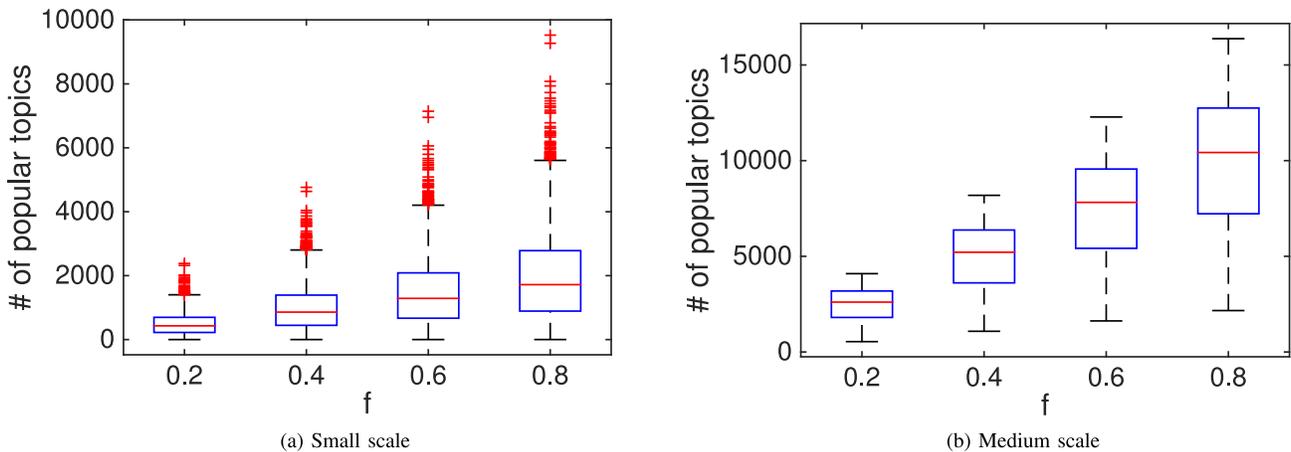


Fig. 12. The popular topic number Q_i as a function of f of topics.

to (1) and (2), and plot the distribution of popular topics in the regions of small and medium scales. In Fig. 11, we observe that there is a wide gap between the smallest and largest number of popular topics in regions of both small and medium scales. Similarly, In Fig. 12, we

find that the number of popular topics exists great differences among the regions of same scale. This shows that each region has different numbers of popular topics and the concentration of video spatial popularity differs significantly.

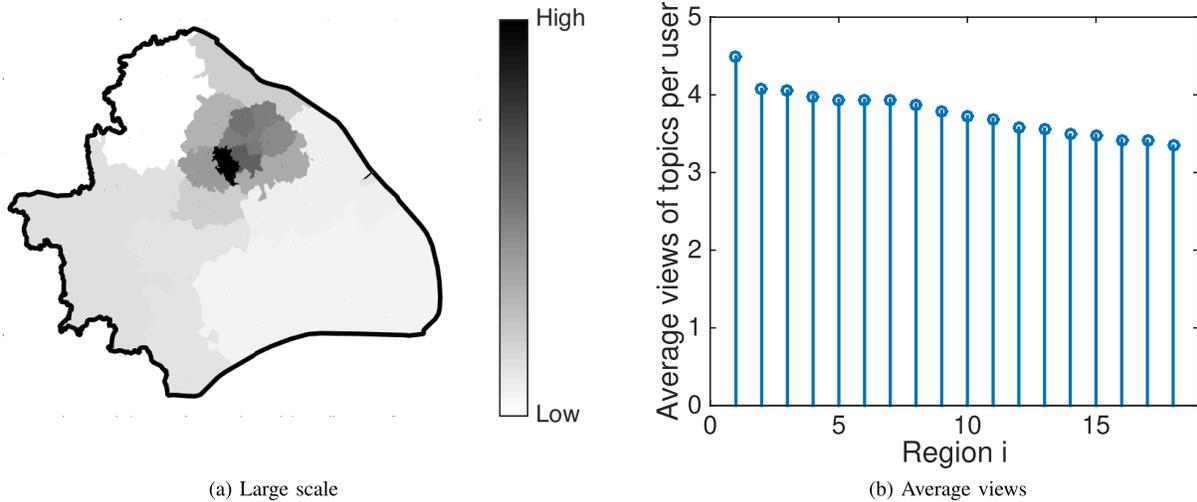


Fig. 13. The spatial distribution of views of top 30% topics by each user in large regions.

To understand the concentration of video popularity in the spatial dimension of the city, we calculate the average number of views of top 30% topics by each user according to (4) in the large scale. We use heatmap to intuitively depict the results in Fig. 13(a), where the darker color represents the larger number of views. We find that the average views increase from the suburb to the downtown, and reach the peak in the downtown. This shows that video popularity exhibits higher concentration near the downtown. Fig. 13(b) plots the average views per topic by each user in each region, which reflects the concentration of video popularity for an individual user.

Further, we investigate the difference of video popularity between the suburb and downtown. Since we seek to analyze the difference of video popularity between the suburb and downtown, it requires to select the specific regions. Meanwhile, previously we analyze the popularity and similarity of watching video across regions of different scales, where the number of user has been averaged. Therefore, this factor has limited impact on the results. When we select two regions for the analysis, we cannot neglect it because the great inequality of user number would significantly influence the conclusion. Thus, we select two regions corresponding to these two locations where the numbers of users are approximately the same. Next, we calculate the number of popular topics whose minimum number of views is range from 50 to 1000 with four types. In Fig. 14, the terms S_C, S_T, S_S and S_M denote the topics with cartoon, TV series, show and movie in the suburb, respectively, while D_C, D_T, D_S and D_M represent the topics with cartoon, TV series, show and movie in the downtown, respectively. From the results, we observe that the number of topics with TV series and shows in the downtown is larger than that in the suburb, while the topics with cartoon and movie attract higher concentrations in the suburb than that in the downtown. One of the reasons is that the more diversity of TV series and shows can attract more preferences across all ages, especially in the downtown. As a result, there are great differences of video popularity in terms of the topic types between the suburb and downtown in the

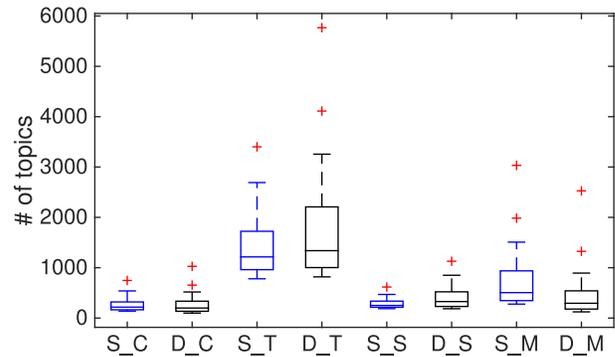


Fig. 14. The number of popular topics whose minimum number of views is range from 50 to 1000 with different types between suburb and downtown.

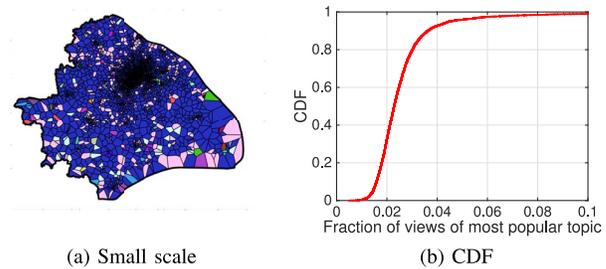


Fig. 15. Illustrating the geographic feature of most viewed topics in small regions.

city, and hence implicates that different regions corresponding to different intensity of economy activities have a great impact on video popularity.

B. Similarity of Video Requests

Now, we focus on Problem II of how similarity of video requests is. To address it, we leverage the metric of popular topic similarity to analyze whether the popular topics are the same in different regions.

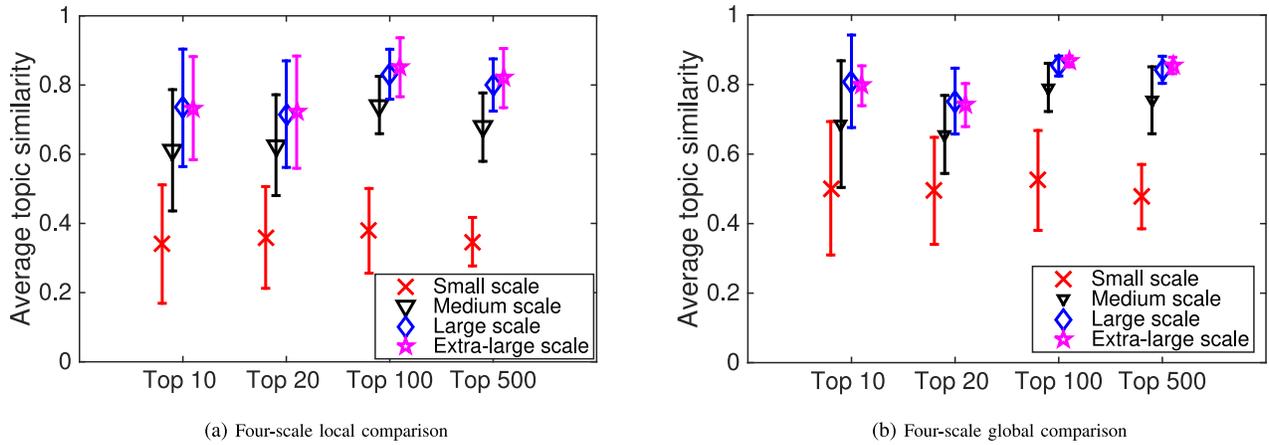


Fig. 16. The popular topic similarity (a) compared with that of locally popular topics under four scales; (b) compared with globally popular topics under four scales.

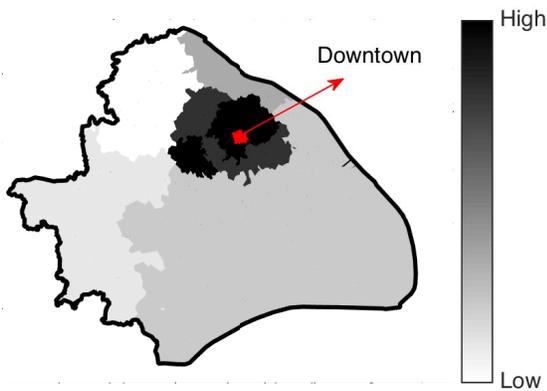


Fig. 17. The spatial distribution of similarity of top 15 topics in large regions compared with that in the downtown.

We select the topics with the most views in each region under small scale, and plot their geographic distributions in Fig. 15 (a), where different color represents different most popular topics. We observe that there are no significant differences in colors of geographically adjacent regions, which indicates that the most popular topics in most of regions are the same. However, from Fig. 15 (b), we see that the percentage of views of the most popular topic in these regions is relatively low, which shows that we cannot neglect the influence of other popular topics that reflect the diversity of video popularity in these regions.

To quantitatively study the diversity of popular topics between regions in different scales, we choose different number of popular topics in each region to calculate their similarities according to (5). Fig. 16 (a) depicts the average similarity with different numbers of popular topics, where we observe that higher similarity exists in larger scales. For example, the similarity of top 10 topics in the extra-large scale is 0.75, which is about 2 times of that in the small scale. This is because the popular topics exhibit the characteristic of concentration in larger regions. To further illustrate this characteristic, we select the most viewed topics in the city and compute their similarity with most popular topics in each region. We show the results of similarity in Fig. 16 (b). From Fig. 16 (b), we

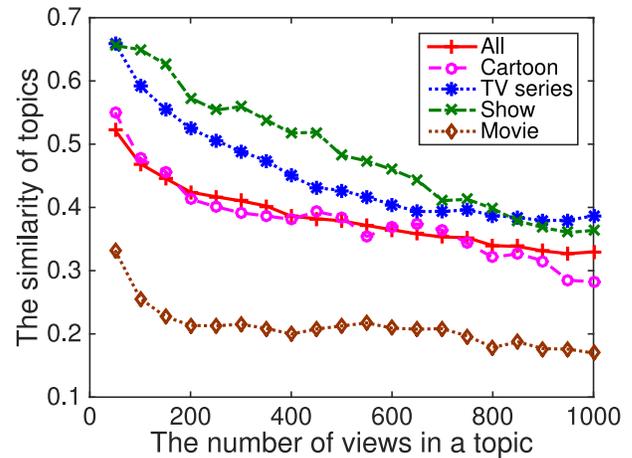


Fig. 18. The similarity of popular topics with different types between suburb and downtown.

observe different similarities under different scales, where the larger the scale is, the higher the similarity can be observed, *e.g.*, the average similarity of top 10 topics under the larger scale is close to 0.8, while it reaches 0.5 under the small scale.

The above analysis focuses on the similarity of popular topics in regions of different scales. To investigate the similarity of popular topics between the downtown and suburb, we choose the concentration of the downtown, the region filled with red color in Fig. 17, as the baseline and calculate the similarity of top 15 topics between this region and other regions in the large scale. As shown in Fig. 17, where the darker color represents higher similarity, we find that the similarity of popular videos becomes low from the downtown to the suburb. Specifically, we choose a region in the suburb and calculate the similarity of popular topics that have more than a given number of views in terms of video types. As shown in Fig. 18, the similarity of topics belonging to movies is lower than that belonging to cartoons, shows and TV series, *e.g.*, the similarity of the popular topics that belong to movies and have more than 200 views is only about 0.2, which is the lowest compared with the similarity of topics belonging to other three types. Further, the topics belonging to shows account for the highest similarity when the number of views in a topic is smaller than 800,

which reveals users in these two regions favour similar topics belonging to shows.

In summary, we observe that great difference of popular topics exist at different scales. Especially, the similarity of popular videos is low in the small scale, while it increases with the increase of the scale. Furthermore, the nearer the region is to the downtown, the higher the similarity is. At the same time, users request the similar popular topics belonging to the show and different topics related to the movie between the downtown and suburb.

C. Summary of Our Findings

By analyzing the spatial characterization of video popularity and similarity in the city, we have obtained some interesting and important findings, which are summarized as follows:

1) Under the studied four scales, most views are concentrated on a few popular topics, and the concentration of video popularity becomes higher in larger scale. Under a given scale, the concentration of video popularity exists great difference in terms of the number of popular topics in the regions.

2) Considering region locations' influence on the concentration of video popularity, we find that higher concentration exists in the downtown. Video popularity on TV series and shows have higher concentration in the downtown, while the higher concentration of video popularity on cartoons and movies exists in the suburb.

3) The popular topic similarity between different small-scale regions is low, and it increases with the scales. Considering region locations' influence on the similarity of popular video requests, we observe that the similarity becomes higher when closer to the downtown. Furthermore, comparing similarity according to topic types between the downtown and the suburb, we find that shows and movies exhibit the highest and lowest similarity, respectively.

V. IMPLICATIONS

We have studied the spatial popularity and similarity of watching videos. Based on our findings, we can infer some important implications on three exemplified applications, i.e., cache deployment, advertisement and video recommendation.

A. Cache Deployment

Since current mobile networks use the all-IP-based infrastructure, cache appliances can be easily deployed in them. Especially, the appearance of network function virtualization brings larger flexibility for cache deployment. Thus, the demand for efficient video delivery as well as the support of flexible deployment makes cache become an important element of future mobile network. Especially, in the city scale, ISP can deploy multiple cache appliances, with each one serving users of the deployed region. However, fundamental understandings of cache for mobile network, such as how cache performs with a large number of viewing requests for large-scale video contents and effective cache deploying methods, are still lacking. Although Least Recently Used (LRU) and Least Frequency Used (LFU) are two widely used caching strategies, it is still difficult to deduce which scale of regions

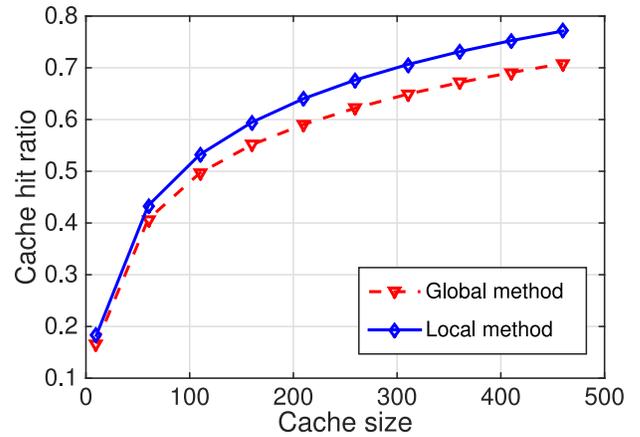


Fig. 19. Cache hit ratio under global and local cache strategies.

cache can be effectively deployed and how much size of cache is in different regions. In such context, the understanding of the spatial popularity and similarity of watching videos based on the real-world data would benefit the solving of above mentioned cache problems.

Our findings provide a valuable guideline on strategies of cache deployment, *e.g.*, video popularity exhibits the concentration even in small-scale region and thus cache can be deployed with a flexible scale. Further, we observe that different regions have different similarities of video popularity, which suggests that different regions should cache different topics.

Further, we quantitatively analyze the important “what-if” questions in the cache deployment. Such questions include: what would be the hit cache ratio if caching different number of topics? If the cache strategy is designed agnostic to the locality of regions (*e.g.*, downtown or suburb regions), how much the cache performance would be worse compared to if it considers this feature?

We use two simple cache strategies named global and local methods to investigate these problems. In the global method, each region caches the global top n topics. On the other hand, local method requires to cache the local top n topics in each region. Assume that the demand of watching topics is known as a priori. We compute the cache hit ratio of two strategies against the cache size, and plot the results in Fig. 19. Note that the cache size is defined as the number of topics. From the results, we observe that the hit ratio increases with the cache size, and local method achieves the highest hit ratio. Especially, when cache size reaches 460, the hit ratio obtained by using local method is 7% higher than global method. This implies that the cache performance can be improved by considering local video popularity. For example, when we need to design an effective cache strategy, it is better to be aware of proximity to the downtown and suburb.

B. Online Advertisement

Online advertisement, as a valuable revenue for content providers, is often placed before the playback of videos. Rather than simply placing the advertisements on arbitrary videos,

the advertisers are eager to target more right users by creating effective advertising. To maximize the advertising effects in a cost-effective way, how much the advertisements should be put in different geographic locations and what videos are suitable and valuable for more advertisements should be taken account into, especially in an area of large scale urban environment. These requirements motivate us to thoroughly investigate the spatial properties of watching videos.

Through the comprehensive analysis of large-scale viewing request data with user locations, we now show that our spatial video popularity characterization can assist deciding the strategies of advertisement. For example, the concentration of video popularity is higher in the regions of downtown, which suggests more advertisements should aim at these specified regions. Further, we should take advantage of the difference of the concentration on the different types between the downtown and suburb to provide the meaningful reference for the advertisers, *e.g.*, putting more advertisements on the TV series in the regions of downtown. In addition, there are other important factors (*e.g.*, what time the advertising is more effective) that worth to be studied to extend the results. As future work, we plan to introduce the viewing time to discuss the video popularity and similarity. We believe the spatial-temporal understanding of video watching can provide more valuable guidelines for the advertising.

C. Video Recommendation

Advertisement can bring substantial profits for content providers only when more views can be attracted. Thus, it is necessary to design an effective video recommendation to increase the hitting ratio of videos. However, many recommendation systems only focus on the popularity of videos without considering the differences between user interests and video contents. In other words, it is critical to recommend appropriate videos that are suitable for region-specified users, in order to bring higher investment of advertisements. To implement it, we conduct trace-driven analysis to observe the geographic features of watching videos.

Our findings aid the design of the video recommendation system. For example, higher concentration on TV series in the downtown (seen in Fig. 14) indicates that the corresponding videos should be recommended to the users in these regions. Furthermore, other spatial characteristics of watching videos such as the similarity under different scales can also be considered to efficiently implement the excellent video recommendation.

Similar with the cache deployment, we also investigate the key “what-if” questions in recommendation system. For instance, what the recommendation performance would be if recommending different number of topics? If the recommendation algorithm is designed agnostic to the locality of regions, how much the recommendation performance would be worse compared to if it considers this feature? To explore them, we use two similar algorithms: global and local algorithms, where global algorithm recommends the global top viewed topics to users, and the local popular topics are recommended in local algorithm. To evaluate the performance of recommendation

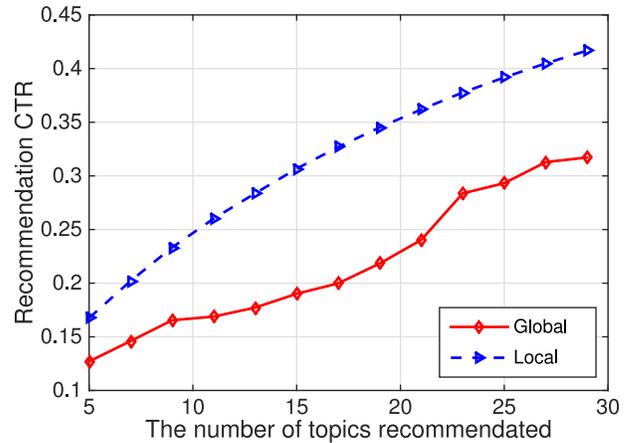


Fig. 20. Recommendation CTR under global and local recommendation algorithms.

system, we use recommendation click through rate (CTR), which is computed as the number of clicks of recommended topics divided by the number of times the recommended topics are showed in a region. Fig. 20 shows the CTR of different algorithms as a function of the number of recommended topics. From the results, we find that CTR of the local algorithm is higher than global one, and the CTR differences between two algorithms is greater than 5%. This demonstrates our analysis that it is beneficial to recommend targeted videos for users in specified regions.

VI. RELATED WORK

Two major directions of research that are correlated with our work are concerned with the investigation of user viewing behaviors on videos and the analysis of geographic features of watching videos respectively.

A. User Viewing Behaviors

User viewing behaviors in Internet video system have been studied in many works [7]–[13], [30]–[32]. Lai and Wang [1] show that video popularity is greatly affected by external links of video sharing sites by the analysis on Youku and YouTube. Yu *et al.* [16] provide the analysis about user behavior, video popularity and their impacts on recommendation. Zhou *et al.* [12] focus on the video popularity dynamics with time for different types of videos and its applications on caching strategies. Crane and Sornette [26] focus on dynamics and peak fraction of viewing behaviors on YouTube to make the prediction model. Reference [3] study the prediction of social video popularity towards cross-domain learning. Figueiredo *et al.* [27] investigate how the popularity growth of YouTube videos evolves with time. Cheng *et al.* [2] make a deep analysis on the characteristics of YouTube videos from statistics and social networking aspects. Abrahamsson and Nordmark [14] study the access patterns and program popularity in a TV-on-demand system. Similarly, [18] investigate the viewing behavior and user activity pattern in the PPTV live streaming system. Acharya *et al.* [28] concentrate on the characterizing videos on world wide Web by classifying

how users access these videos. Li *et al.* [10] characterize different user behaviors of watching mobile videos from IPTV and VoD systems in comparing the 3G and WiFi access methods. Chen *et al.* [24] build a user behavior model to study video browsing, switching and early departure behaviors. In addition, Lin *et al.* [9] investigate peer-assisted video delivery in WiFi networks and conduct the analysis of viewing time, user population and user locality. Shafiq *et al.* [23] present the feature of mobile video streaming performance and model its influence on user engagement from network operators' perspective. Unlike them, we study the spatial features of video popularity and similarity from the perspective of ISP since our data contains view requests of videos from multiple content providers.

B. Geographic Features of Watching Videos

There have been some studies on geographic features of watching online videos [15], [17], [19], [22], [33]. Brodersen *et al.* [5] study whether YouTube videos exhibit geographic locality of YouTube videos over the world and show that most of views come from the videos in a single country. Platt *et al.* [21] study videos trending across several nations and make the contribution to understand the international cultural impact and potential of videos. Li *et al.* [6] provide an in-depth study on the geographic patterns of mobile video consumption of PPTV among different provinces within a country, and obtain the distinct geographic popularity features on popular and non-popular videos. Also, [17] analyze the relationship between the user interest on video categories and locations of mainland China, overseas and Hongkong-Macao-Taiwan. Scale *et al.* [22] exploit the geographic information from social cascades to predict the patterns of video views in future. Huguenin *et al.* [25] find that the content locality and geographic locality are correlated based on the analysis of YouTube dataset. Zink *et al.* [34] explore the correlation between global popularity obtained from YouTube website and local popularity from local network trace. Our work differs with these works in that we aim at exploring the spatial features of video popularity and similarity based on different granularities of the divisions under the city scale.

VII. DISCUSSION

At that time of data collection, it is not popular to watch videos through cellular network due to high expense of cellular traffic. Therefore, ISP offers us the logs of video access over wired and WiFi networks, but without the cellular networks. This has limited effects on our results.

In addition, mobility of users may also affect the spatial popularity, which can provide additional insights on temporal access in downtown and suburb. However, our dataset cannot distinguish mobile users as the relevant work is now in process by ISP. In the future work, we plan to study them when related dataset is available.

VIII. CONCLUSION

In this paper, we aim at understanding the spatial characterization of popularity and similarity of watching videos in

the city scale. We carry out a comprehensive analysis of the concentration and similarity of video popularity in the spatial dimension over a large-scale video viewing logs that covers six most popular video content providers in China. Through the analysis, we obtain some interesting but important findings, which are summarized as follows:

(i) With regarding video popularity when considering different scales and region locations in a city, we find that under all four studied scales, video view requests conform with Pareto principle, and thus the concentration of video popularity exists at all these scales. Furthermore, under the same scale, the differences of concentration exist in different regions and the regions in the downtown exhibit higher concentration than ones in the suburb.

(ii) With regarding video request similarity when considering different scales and region locations in a city, we find that the similarity of video popularity in a region varies with each other, which is especially significant under the small scale. The nearer the region is to downtown, the higher similarity there is when compared with a chosen region in the downtown.

These findings offer useful guidelines for cache deployment, advertisement and video recommendation.

REFERENCES

- [1] K. Lai and D. Wang, "Understanding the external links of video sharing sites: Measurement and analysis," *IEEE Trans. Multimedia*, vol. 15, no. 1, pp. 224–235, Jan. 2013.
- [2] X. Cheng, J. Liu, and C. Dale, "Understanding the characteristics of Internet short video sharing: A YouTube-based measurement study," *IEEE Trans. Multimedia*, vol. 15, no. 5, pp. 1184–1194, Aug. 2013.
- [3] S. D. Roy, T. Mei, W. Zeng, and S. Li, "Towards cross-domain learning for social video popularity prediction," *IEEE Trans. Multimedia*, vol. 15, no. 6, pp. 1255–1267, Oct. 2013.
- [4] Z. Tan, Y. Wang, Y. Zhang, and J. Zhou, "A novel time series approach for predicting the long-term popularity of online videos," *IEEE Trans. Broadcast.*, vol. 62, no. 2, pp. 436–445, Jun. 2016.
- [5] A. Brodersen, S. Scellato, and M. Wattenhofer, "YouTube around the world: Geographic popularity of videos," in *Proc. WWW*, Lyon, France, 2012, pp. 241–250.
- [6] Z. Li *et al.*, "On the geographic patterns of a large-scale mobile video-on-demand system," in *Proc. IEEE INFOCOM*, Toronto, ON, Canada, 2014, pp. 397–405.
- [7] A. Finamore, M. Mellia, M. Munafò, R. Torres, and S. G. Rao, "YouTube everywhere: Impact of device and infrastructure synergies on user experience," in *Proc. ACM IMC*, Berlin, Germany, 2011, pp. 345–360.
- [8] 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. ACM IMC*, San Diego, CA, USA, 2007, pp. 1–14.
- [9] J. Lin *et al.*, "Mobile video popularity distributions and the potential of peer-assisted video delivery," *IEEE Commun. Mag.*, vol. 51, no. 11, pp. 120–126, Nov. 2013.
- [10] Y. Li, Y. Zhang, and R. Yuan, "Measurement and analysis of a large scale commercial mobile Internet TV system," in *Proc. ACM IMC*, Berlin, Germany, 2011, pp. 209–224.
- [11] Y. Borghol, S. Ardon, N. Carlsson, D. Eager, and A. Mahanti, "The untold story of the clones: Content-agnostic factors that impact YouTube video popularity," in *Proc. KDD*, Beijing, China, 2012, pp. 1186–1194.
- [12] Y. Zhou, L. Chen, C. Yang, and D. M. Chiu, "Video popularity dynamics and its implication for replication," *IEEE Trans. Multimedia*, vol. 17, no. 8, pp. 1273–1285, Aug. 2015.
- [13] J. Wu, Y. Zhou, D. M. Chiu, and Z. Zhu, "Modeling dynamics of online video popularity," *IEEE Trans. Multimedia*, vol. 18, no. 9, pp. 1882–1895, Sep. 2016.
- [14] H. Abrahamsson and M. Nordmark, "Program popularity and viewer behaviour in a large TV-on-demand system," in *Proc. ACM IMC*, Boston, MA, USA, 2012, pp. 199–210.

- [15] H. Yin *et al.*, "Inside the bird's nest: Measurements of large-scale live VoD from the 2008 olympics," in *Proc. ACM IMC*, Chicago, IL, USA, 2009, pp. 442–455.
- [16] H. Yu, D. Zheng, B. Y. Zhao, and W. Zheng, "Understanding user behavior in large-scale video-on-demand systems," in *Proc. EuroSys*, Leuven, Belgium, 2006, pp. 333–344.
- [17] Z. Li *et al.*, "Watching videos from everywhere: A study of the PPTV mobile VoD system," in *Proc. ACM IMC*, Boston, MA, USA, 2012, pp. 185–198.
- [18] Z. Li, G. Xie, M. A. Kaafar, and K. Salamatian, "User behavior characterization of a large-scale mobile live streaming system," in *Proc. WWW*, Florence, Italy, 2015, pp. 307–313.
- [19] M. Cha, P. Rodriguez, J. Crowcroft, S. Moon, and X. Amatriain, "Watching television over an IP network," in *Proc. ACM SIGCOMM*, Vouliagmeni, Greece, 2008, pp. 71–84.
- [20] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law," *Contemp. Phys.*, vol. 46, no. 5, pp. 323–351, 2005.
- [21] E. L. Platt, R. Bhargava, and E. Zuckerman, "The international affiliation network of YouTube trends," in *Proc. AAAI ICWSM*, Oxford, U.K., 2015, pp. 318–326.
- [22] S. Scellato, C. Mascolo, M. Musolesi, and J. Crowcroft, "Track globally, deliver locally: Improving content delivery networks by tracking geographic social cascades," in *Proc. WWW*, Hyderabad, India, 2011, pp. 457–466.
- [23] M. Z. Shafiq *et al.*, "Understanding the impact of network dynamics on mobile video user engagement," in *Proc. ACM SIGMETRICS*, Austin, TX, USA, 2014, pp. 367–379.
- [24] L. Chen, Y. Zhou, and D. M. Chiu, "Video browsing—A study of user behavior in online VoD services," in *Proc. ICCCN*, Nassau, The Bahamas, 2013, pp. 1–7.
- [25] K. Huguenin, A. Kermarrec, K. Kloudas, and F. Taïani, "Content and geographical locality in user-generated content sharing systems," in *Proc. NOSSDAV*, Toronto, ON, Canada, 2012, pp. 77–82.
- [26] R. Crane and D. Sornette, "Robust dynamic classes revealed by measuring the response function of a social system," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 41, pp. 15649–15653, 2008.
- [27] F. Figueiredo, F. Benevenuto, and J. M. Almeida, "The tube over time: Characterizing popularity growth of YouTube videos," in *Proc. WSDM*, Hong Kong, 2011, pp. 745–754.
- [28] S. Acharya, B. Smith, and P. Parnes, "Characterizing user access to videos on the World Wide Web," in *Proc. SPIE/ACM MMCN*, 2000, pp. 130–141.
- [29] H. Yan, J. Liu, Y. Li, D. Jin, and S. Chen, "Spatial popularity and similarity of watching videos in a large city," in *Proc. IEEE GLOBECOM*, Washington, DC, USA, 2016, pp. 1–6.
- [30] B. Wang, X. Zhang, G. Wang, H. Zheng, and B. Y. Zhao, "Anatomy of a personalized livestreaming system," in *Proc. ACM IMC*, Santa Monica, CA, USA, 2016, pp. 485–498.
- [31] H. Yan *et al.*, "A first look at user switching behaviors over multiple video content providers," in *Proc. AAAI ICWSM*, Montreal, QC, Canada, 2017, pp. 700–703.
- [32] Z. Li, J. Lin, K. Salamatian, and G. Xie, "Social connections in user-generated content video systems: Analysis and recommendation," *IEEE Trans. Netw. Service Manag.*, vol. 10, no. 1, pp. 70–83, Mar. 2013.
- [33] A. Raman, N. Sastry, A. Sathiseelan, J. Chandaria, and A. Secker, "Wistitch: Content delivery in converged edge networks," in *Proc. Workshop Mobile Edge Commun.*, Los Angeles, CA, USA, 2017, pp. 13–18.
- [34] M. Zink, K. Suh, Y. Gu, and J. Kurose, "Watch global, cache local: YouTube network traffic at a campus network: Measurements and implications," in *Proc. Multimedia Comput. Netw.*, vol. 6818, 2008, Art. no. 681805.



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



Jiaqiang Liu received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2012, where he is currently pursuing the Ph.D. degree. His research interests include software-defined network, data center network, network function virtualization, and cloud computing.



Yong Li (M'09–SM'16) received the B.S. degree in electronics and information engineering from the 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, where he is currently a Faculty Member with the Department of Electronic Engineering.

He has served as General Chair, TPC Chair, and TPC Member for several international workshops and conferences. He is on the editorial board of two IEEE journals. He has over 4000 citations. Among them, ten are ESI Highly Cited Papers in Computer Science. He was a recipient of four Conference Best Paper (Run-Up) Awards and the IEEE 2016 ComSoc Asia-Pacific Outstanding Young Researchers and Young Talent Program of China Association for Science and Technology.



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



Sheng Chen received the B.E. degree in control engineering from East China Petroleum Institute, Dongying, China, in 1982, the Ph.D. degree in control engineering from City University, London, U.K., in 1986, and the higher Doctoral and Doctor of Science degrees from the University of Southampton, Southampton, U.K., in 2005. From 1986 to 1999, he held a research and academic appointments with the University of Sheffield, Sheffield, U.K., the University of Edinburgh, Edinburgh, U.K., and the University of Portsmouth, Portsmouth, U.K. Since 1999, he has been with the Department of Electronics and Computer Science, University of Southampton, where he is currently a Professor of Intelligent Systems and Signal Processing. He has published over 550 research papers. He is a Distinguished Adjunct Professor with King Abdulaziz University, Jeddah, Saudi Arabia. His current research interests include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, intelligent control system design, and evolutionary computation methods and optimization. He was a recipient of the ISI Highly Cited Researcher Award in engineering in 2004. He is a fellow of IET. He was elected as a fellow of the U.K. Royal Academy of Engineering in 2014.