Characterizing the Spatio-Temporal Inhomogeneity of Mobile Traffic in Large-scale Cellular Data Networks

Huandong Wang†, Jingtao Ding†, Yong Li†, Pan Hui‡, Jian Yuan†, Depeng Jin†,
† Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
‡ Department of Computer Science and Engineering, Hong Kong University of Science and Technology
liyong07@tsinghua.edu.cn

ABSTRACT

As the volume of mobile traffic has been growing quickly in recent years, reducing the congestion of mobile networks has become an important problem of networking research. Researchers found out that the inhomogeneity in the spatio-temporal distribution of the data traffic leads to extremely insufficient utilization of network resources. Thus, it is important to fundamentally understand this distribution to help us make better resource planning or introduce new management tools such as time-dependent pricing to reduce the congestion. However, due to the requirement of a large dataset, a detailed, radical and credible network-wide study for the spatio-temporal distribution of mobile traffic is still lacking.

In this work, we conduct such a measurement study. Base on a large-scale data set obtained from 380,000 base stations in Shanghai spanning over one month, we quantitatively characterize the spatio-temporal distribution of mobile traffic and present a detailed visualized analysis. Furthermore, on the basis of quantitative analysis, we find that the mobile traffic loads uniformly follow a trimodal distribution, which is the combination of compound-exponential, power-law and exponential distributions, in terms of both spatial and temporal dimension. Extensive results show that our model is with accuracy over 99%, which provides fundamental and credible guidelines for the practical solutions of the issues in mobile traffic operations.

Categories and Subject Descriptors

C.2.0 [**Computer Communication Networks**]: General–Data communications; C.4 [**Performance of Systems**]: Modeling techniques

Keywords

Mobile Data Traffic; Spatio-Temporal Inhomogeneity; Measurement Study; Trimodal Distribution

HotPOST'15, June 22, 2015, Hangzhou, China.

http://dx.doi.org/10.1145/2757513.2757518.

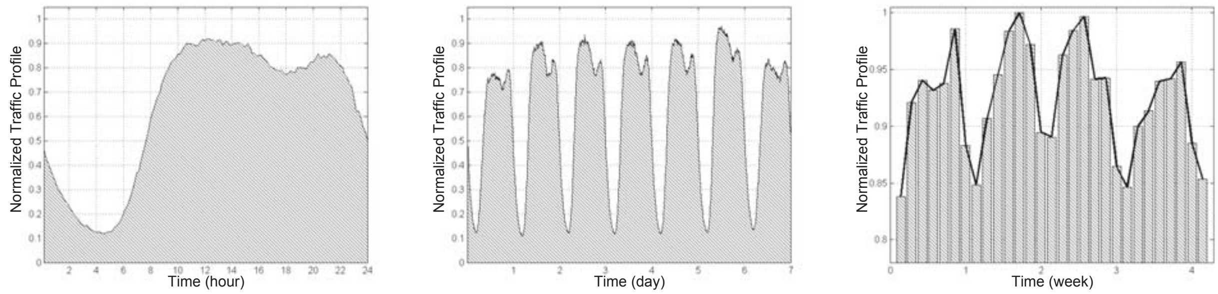
1. INTRODUCTION

With the popularity of smart devices and various bandwidth-hungry applications, the mobile data traffic has been growing rapidly in recent years. Global mobile data traffic is predicted to increase nearly tenfold between 2014 and 2019, reaching 24.3 exabytes per month by 2019. In addition, the dominant traffic in wireless networks has been shifting from voice to mobile video, of which the traffic exceeded 50 percent of total mobile data traffic in 2012, and is predicted to account for nearly three-fourths of the world's total mobile data traffic by 2019 [1]. This explosive growth in bandwidth demand is now forcing network operators to search for solutions to mitigate the network congestion.

Mobile operators have tried every possible solution to keep up with the traffic demand by acquiring additional spectrum, deploying WiFi hotspots, and adopting new technologies such as fourth-generation (4G) Long Term Evolution (LTE) and femtocells, but these approaches are still insufficient for today's challenges of network congestion [5]. Researchers found out that due to highly dynamic behavior of human activity, the traffic usages are inhomogeneous in terms of both spatial and temporal dimension. For example, Michalopoulou *et al.* [4] studied GSM cellular network in Berlin, and showed that Erlang based traffic loads were heavier in the city center and there existed a potential spatial differentiation between business and non-working hours. In the city center or during peak time, the mobile network is seriously congested, while in other space and time, the network resources are insufficiently used. Thus, this inhomogeneous traffic usages are one of the key factors that lead to the serious congestion problem.

In addition, a number of approaches have tried to make better resource planning with the assumption of the inhomogeneity character of mobile data traffic. Sen *et al.* [5] focused on using the time-dependent pricing as a congestion management tool, where the service providers made discounts during off-peak period to incentivize users to time-shift their data demand from the peak to off-peak periods. Peng *et al.* [7] proposed a profile-based approach, which leveraged temporal-spatial traffic diversity and node deployment heterogeneity, and powered off under-utilized base stations under light traffic. All these approaches depend on a fundamental and credible model of the mobile traffic distribution. For this purpose, analyzing and characterizing the properties of the mobile traffic are indeed crucial.

There are few systematic measurement studies of both spatial and temporal distribution of mobile traffic. Lee *et al.* [2] focused on analyzing spatial traffic distributions from the



(a) (b) (c)

Figure 1: The temporal distribution of normalized traffic profile at different time scales (overall traffic).

traffic measurements in commercial cellular networks and proposed a spatial traffic model. However, this model is not accurate enough since the cell traffic was modelled to follow approximately a gamma or Weibull distribution, which cannot character its tail distribution. Laner *et al*. [6] provided an analysis of mean cell load over daytime from a live High Speed Packet Access (HSPA) network. However, their main goal is to construct statistical models for HSDPA data traffic in one single cell. Data traffic in different cells as well as the spatial distributions are ignored. Thus, the spatial and temporal distributions should be also combined to obtain a traffic model with more accuracy and credibility. Moreover, the analysis should be based on a data set of which the scale is large enough.

In this work, we focus on characterizing this inhomogeneity based on a large-scale and long-time data set, which is collected from the commercial cellular networks with 380,000 base stations in Shanghai spanning over a month. By carrying out a thorough analysis on this dataset, we quantitatively characterize the spatio-temporal distribution of mobile traffic and present a detailed visualized analysis. On the basis of quantitative analysis, we show that the total traffic loads per base station at different scales of time as well as the traffic loads per 10 minutes of one base station at different places can be approximated by a trimodal distribution, which is the combination of compound-exponential, power-law and exponential distributions. Extensive results show that our model is highly accurate: the average adjusted R-square statistics between the empirical curves and our model are over 99% at different scales of time and space. Thus, it provides fundamental and credible guidelines for the practical solutions of the issues in mobile traffic operations, such as reducing congestion and improving energy efficiency.

The rest of the paper is organized as follows. In Section 2, we quantitatively characterize the spatio-temporal distribution of mobile traffic, and present a detailed visualized analysis of it. In Section 3, we show that the mobile traffic loads follow a trimodal distribution, which is the combination of compound-exponential, power-law and exponential distributions, in terms of both spatial and temporal dimension. After presenting related works in Section 4, we draw our conclusion in Section 5.

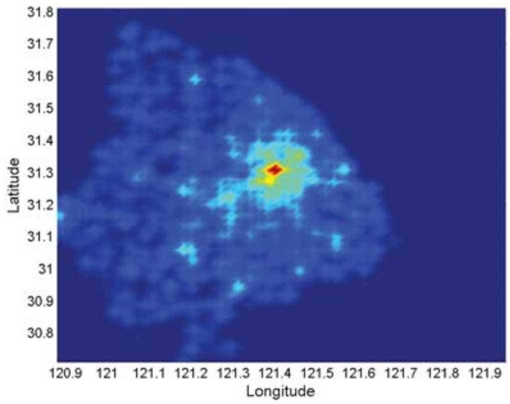
2. VISUALIZED ANALYSIS OF SPATIO-TEMPORAL TRAFFIC DISTRIBUTION

The data set we investigated is collected from the cellular networks of Shanghai by one of the major operators during August 1 to August 31, 2014. Each entry in the data set is corresponding to a continuous data communication, of which the device's ID (anonymized), starting and ending time, base station (BS) ID, BS location and traffic volume are recorded. There are about 1.96 billion entries in our data set, involving over 380,000 BSs.The total volume of traffic recorded in the data set is over 2.8 PB, with over 92 TB per day and 7 GB per BS on average. Thus, this large-scale and fine-grained data set guarantees the credibility of the obtained results.

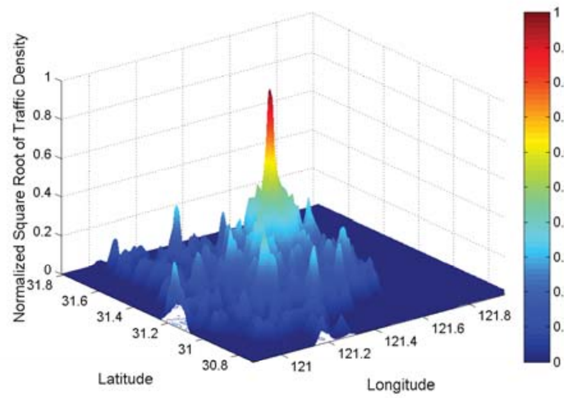
2.1 Temporal Distribution

We start with analyzing the temporal distribution of traffic of each BS in the network. Based on the above introduced massive data set, a preprocessing procedure is conducted to sort the traffic records by time and BS ID, and compute the data traffic according to the usage during a certain period within same BS. Since the duration of one continuous data communication varies from several minutes to several hours, we assume the traffic follows a uniform distribution in the interval and divide each duration into several ten-minute intervals with equivalent traffic, in which intervals shorter than ten minutes are approximately regarded as ten-minute intervals. In addition, we assume the start time of each interval is located at specific moments, which are just 10, 20, 30, 40 or 50 minutes away from the whole point of time. Thus, by this classification and summation, we obtain the traffic load of each BS with the minimum scale of 10 minutes for the whole 31 days.

Fig. 1 shows the aggregate traffic load of the entire network at different time scales, where Fig. 1(a) shows the distribution of the aggregate traffic load for one day period, *i.e.*, Aug. 7th (Thursday), 2015. Fig. 1(b) shows the distribution for seven day period, *i.e.*, from Aug. 3th (Sunday) to Aug. 9th, 2015. The volume of traffic is normalized by the maximum volume of traffic per unit time all over the month. From the results, we can observe that the aggregate network load exhibits a stable periodic behavior tightly coupled with the sleep schedule of humans. Overall, the traffic load is relatively high during the day and is lowest during midnight. There are two peak during the whole day: one is at about 12PM and the other is at about 10PM, which are



(a) Tow-dimensional



(b) Three-dimensional

Figure 2: The two-dimensional and three-dimensional view of the spatial distribution of normalized square root of traffic density.

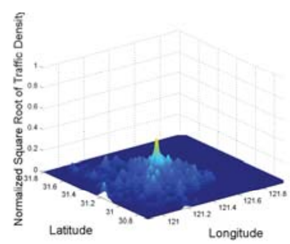
just the lunch time and bedtime of most people, respectively. This indicates that most people trend to consume the data heavily after they have lunch or on bed before they sleep.

Fig. 1(c) shows the daily traffic load of the entire network in the scale of weeks, *i.e.*, from Aug. 3th to Aug. 31th, 2015. The traffic load variations during each day is ignored, while its normalized total volume is plotted. Though not as much obvious as the periodicity of one day shown in Fig. 1(b), the daily traffic load exhibits a week periodic behavior. For example, the traffic load in weekends, such as August 2 and August 3, is obviously less than the traffic load in weekdays, which is resulted from the work schedule of humans. However, traffic loads during weekdays seem more irregular. Overall, there exist inhomogeneous traffic usages at all different scales of time.

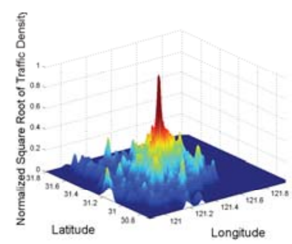
2.2 Spatial Distribution

In this section, we turn our attention to the spatial distribution of the mobile traffic. To characterize the spatial distribution, we first convert geographical names of the BS location to longitudes and latitudes through APIs provided by map service like Baidu. Then, we compute the traffic density, which is defined as the traffic demand per unit area (set to be 1 km^2 by default). We add up the traffic volume of BSs which are located at the same unit area, and obtain the spatial distribution of traffic. Since the peak in the center of the city is too high, traffic density at other place almost cannot be observed in the same figure with linear coordinate. Thus, we use the square root of the traffic density to represent its heaviness.

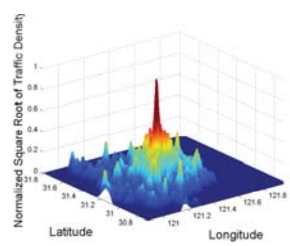
The two-dimensional and three-dimensional view of the spatial distribution of normalized square root of traffic density are shown in Fig. 2(a) and (b), respectively. From the results, it can be observed that the coverage area of mobile traffic just outlines the city of Shanghai. Different color depth means different traffic density of each area, of which the corresponding relation can be obtained in the colorbar of Fig. 2(b). In addition, the peaks of traffic, which in some way show the main concentrated areas of people, such as residential zone or central business district (CBD), spread over the city. There is an outstanding peak in the center of the city with many relatively low peaks all over the city, which indicates the extremely inhomogeneous traffic usages in terms of space.



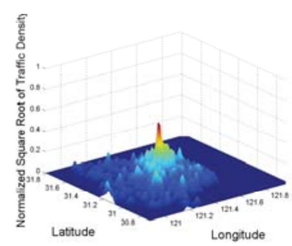
(a) 4AM



(b) 10AM



(c) 4PM



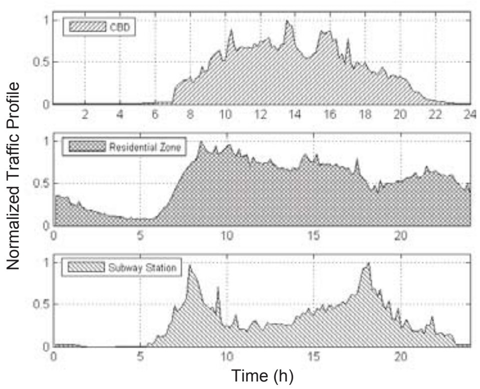
(d) 10PM

Figure 3: The spatial distribution of mobile traffic profile at different time.

2.3 Spatio-Temporal Distribution

In this section, we analyze the characters of mobile traffic by combining the spatial and temporal distribution. Specifically, we first investigate the spatial distribution of mobile traffic profile at different time, and then we focus on the temporal distribution of mobile traffic at different places.

By statistics analysis, we find that the traffic density of the aggregate traffic load of different days do not change obviously. However, the traffic density of different hours varies significantly. Thus, we plot the three-dimensional view of the traffic density of different time in one day in Fig. 3. The corresponding time is 4AM, 10AM, 4PM and 10PM, respectively. From the results, we observe that the highest peak



(a) The temporal distribution of mobile traffic profile at different places.

Base Station	Avg	Var/Avg^2
Subway Station	0.0731	0.9044
CBD	2.3988	0.2445
Residential Zone	1.0000	0.7657

(b) The average and variance of traffic distribution of BSs at different places.

Figure 4: The temporal distribution and related characteristic parameters of mobile traffic profile at different places.

of traffic density in the city center remains unmoved all over the day. When at 4AM, most areas of the city are covered with relatively dark color, indicating that the traffic demand is small as most people are sleeping and their devices are inactive. Meanwhile, at 10PM, which is the moment with the largest traffic demand according to result obtained from Fig. 1(a), most areas of the city are covered with relatively light color, indicating that the traffic demand is large as most people and their devices are active. Then, the lightness of color shows a downtrend at 4PM and 10PM, which again agrees with the results from Fig. 1(a). Overall, there is a huge difference between the traffic density of these different hours, reflecting the extremely inhomogeneous traffic usages in terms of time.

Then, to analyze the characters of traffic demand in areas with different functions, the traffic of BSs in three typical areas (central business district (CBD), residential zone and subway station) is shown in Fig. 4. The types of BS are not limited to those mentioned above, which requires further study and research. In this paper, we only present the traffic profile of the most typical three kinds of BSs. As we can observe from Fig. 4(a), the traffic demand of BS near subway station tends to reach peak point near 8AM and 6PM, which indicate rush hours when people surge to office buildings from their home or in turn, while those BSs near CBD tend to have only one peak near 1PM each day. Meanwhile, the traffic demand of BS near residential zone tends to have only one peak near 8PM, which is the wake-up time of most people, and it shows a sustained downward since peak at 8PM. However, after 6PM, which is the quitting time of most people, the traffic demand of BS near residential zone turns to showing an uptrend, while the traffic demand of BS near CBD shows a sharp downtrend, indicating the population flow from CBD to residential zone. Then, their average and quotient of their variance and the square of the average are presented in Fig. 4(b), where the average traffic volume of BS near residential zone is taken as the baseline. We can observe that the traffic of BS near CBD has the largest mean value and smallest variation, while that near subway station has the smallest mean value and largest variation in contrast. Though the traffic loads of these BSs all exhibit strong inhomogeneity in terms of temporal dimension, their characters in terms of spatial dimension are very different. Thus, it is important to analyze them by combining the spatial and temporal distribution.

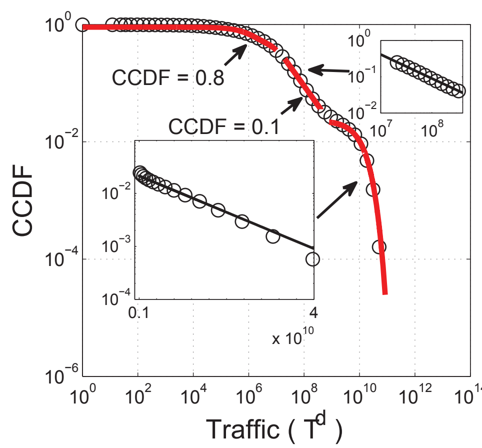
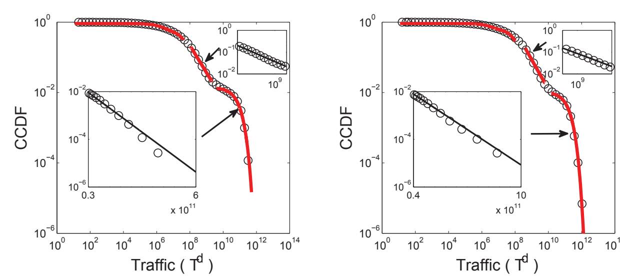


Figure 5: The CCDF of each BS traffic for one day(Aug. 1st, 2014) in a log-log plot and fitting of the model(red) to the empirical data(black), where the red curve is separated into 3 parts by 2 boundary point, $t_0^d = 2 \times 10^7$ and $t_1^d = 7 \times 10^8$. The upper inset displays the power law part of CCDF in log-log scale. The lower inset shows the exponential tails of CCDF in the linear-log plot. The straight lines are the power law and exponential fitting functions, which are correspondingly shown by the second and third red part in the main figure. We also show that the empirical data in $[1, t_0^d]$ can be well approximated by the first function in (1).



(a) (b)

Figure 6: (a) and (b) are the CCDFs of each BS traffic for one week(Aug. 1st to 7th, 2014) and one month(Aug. 2014), respectively, in a log-log plot.

Table 1: The evaluation of trimodal distribution.
(a) The approximated exponents in fitting functions.

exponent	Spatial Distribution (Fig 5)	Temporal Distribution (Fig 7)
α	$(8.67 \pm 0.02) \times 10^{-7}$	$(6.50 \pm 0.67) \times 10^{-1}$
β	$(4.58 \pm 0.01) \times 10^{-8}$	$(2.07 \pm 0.03) \times 10^{-2}$
ε	1.77 ± 0.01	$(7.95 \pm 0.57) \times 10^{-1}$
η	$(6.59 \pm 0.01) \times 10^{-1}$	3.37 ± 0.03
γ	$(8.12 \pm 0.01) \times 10^{-11}$	$(6.38 \pm 0.03) \times 10^{-3}$

(b) The adjusted R-square statistics of the compound, power law and exponential fittings.

Distribution	Spatial			Temporal		
	1day	1week	1month	total	centre	suburb
Compound	99.09%	99.07%	99.00%	99.68%	99.74%	99.45%
Power Law	99.89%	99.68%	99.68%	99.78%	99.98%	99.99%
Exponential	99.78%	99.66%	99.72%	99.76%	99.74%	99.74%

3. MODELING THE SPATIO-TEMPORAL TRAFFIC DISTRIBUTION

In order to accurately characterize the above spatial-temporal distribution of the mobile traffic, we measure the data in the scale of one-day and ten-minute in each BS, denoted as t^d and t^m respectively. Furthermore, we detail these two variables into $t^d_{i,j}$ and $t^m_{k,l}$, where i and k represent the BS ID and time series are marked as j and l. Since one month (Aug. 2014) has 31 days and each day has $24 \times 6 = 144$ time periods of ten minute, ranges of time series j and l are $1 \leq j \leq 31$ and $1 \leq l \leq 4464$. For $t^d_{i,j}$, by fixing temporal variable j, we obtain the information of spatial distribution by the empirical data of BS traffic during the same day. Similarly, we study the temporal distribution by fixing spatial variable k in a set. Considering that 10^4 BSs have 4×10^8 $t^m_{k,l}$, which is difficult to obtain cumulative distribution, we replace the original traffic values with 10^3 quantized values through linear quantization, and denoted them as $\{t^{mq}_{k,l}\}$.

Based on the measurement-driven traffic analysis, we find out that both spatial distribution(i.e., $\{t^d_{i,j}$, j fixed$\}$, see Fig. 5) and temporal distribution(i.e., $\{t^{mq}_{k,l}$, k fixed$\}$, see Fig. 7) can be well described as:

$$P(T > t) = \begin{cases} e^{-\alpha t} + \varepsilon e^{-\beta t}, & t < t_0 \\ t^{-\eta}, & t_0 < t < t_1 \\ e^{-\gamma t}, & t > t_1 \end{cases} \qquad (1)$$

In this trimodal distribution, the first part is a compound curve of the summation of two different exponential functions. As for the second and third part, the exponential tail is connected to the power-law with a hump well above the straight line extrapolation of the power-law, which is significantly different from the usually truncated power-law with the form $t^{-\gamma} e^{-\beta t}$, where the exponential tail is below the straight line of the power law.

Fig. 5 depicts the Complementary Cumulative Distribution Function (CCDF) of BS traffic for one day (Aug. 1st, 2014) in a log-log plot, i.e., $P(T^d_{i,1} > t^d_{i,1})$, as well as fitting of the trimodal distribution model to the empirical data. All the exponents in (1) are shown in Table 1(a). We find out that one-day traffic in 90% of BSs is less than 10^8, and there are 70% of BSs in $[10^6, 10^8]$. Since the highest one-day traffic is about 10^{11}, this result means that traffic of most

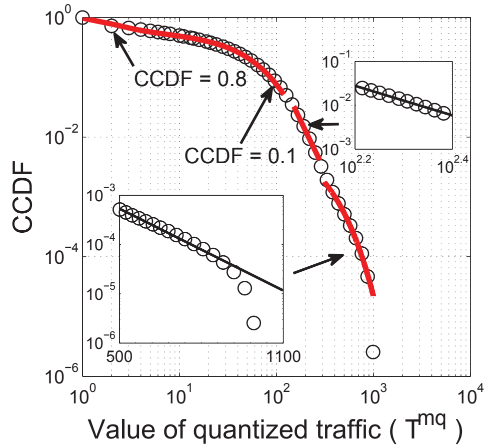
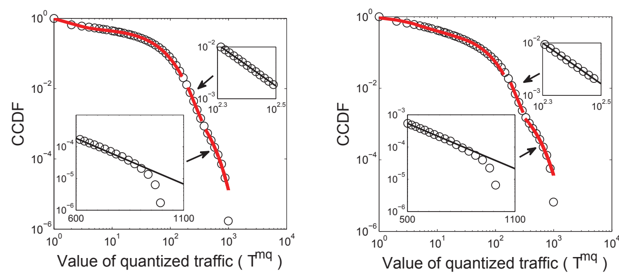


Figure 7: The CCDF of ten-minute quantized traffic in BSs located throughout the city for one month(Aug. 2014) and fitting of the model(red) to the empirical data(black), where the red curve is separated into 3 parts by 2 boundary point, $t^{mq}_0 = 130$ and $t^{mq}_1 = 300$. These 3 parts correspond to the 3 fitting functions shown in (1). The upper inset and lower inset display the power law and exponential part of CCDF. Straight lines represent the analytical prediction.



(a) (b)

Figure 8: (a) and (b) are the CCDFs of ten-minute quantized traffic in BSs near city centre and suburb, respectively, for one month(Aug. 2014).

BSs is neither too low nor too high, which is reasonable. Besides, when changing timescales of traffic measurement, we observe that both one-week and one-month traffic follow the same distribution, as shown in Fig. 6. The goodness of fit is measured quantitatively by the R-square statistics, which are summarized in Table 1(b). The average adjusted R-square statistics are all over 99% for the three parts. This confirms the accuracy of our trimodal distribution model. In conclusion, spatial distribution of traffic in different BSs follows the trimodal distribution, with most BSs belonging to the compound and power law part. Since the compound part is similar to the exponential part, these results indicate the spatial differentiation of traffic in power law part is more obvious than that of others.

In terms of temporal distribution, we count ten-minute traffic value for the whole month. After quantization, their aggregated empirical CCDF, i.e., $P(T^{mq}_{K,l} > t^{mq}_{K,l})$, with K as the BS set, are plotted using log-log scale in Fig. 7. The exponents of curve fitting are also shown in Table 1(a). It can be observed from the figure that 90% of $t_{K,l}^{mq}$ are less than 90. Comparing with Fig. 5, it seems that the compound part occupies a more dominated position and the hump between power law part and exponential tail is less distinct, but we still find a trimodal distribution temporally. Further, we select a radius of 10 km of the region in centre and suburb of *Shanghai*, respectively, and plot the CCDFs of ten-minute quantized traffic in BSs located in these 2 regions in Fig. 8. Similar to spatial distribution, the temporal distribution of ten-minute traffic also follows the same invariant pattern. The goodness of fit is also shown in Table 1(b), with the adjusted R-square statistics all over 99%.

In conclusion, on basis of above quantitative analysis, we find a universal model of spatial-temporal traffic distribution, which shows robustness in different scales of time and space. Furthermore, our study shows uniformity of spatial and temporal pattern, with both of them following the t-rimodal distribution. With this uniformed spatio-temporal traffic distribution model, we can characterize the operational status of BS both spatially and temporally. For example, after setting a series of thresholds, we can divide the operational status of BS into several types such as no-load, low-load, full-load, off-load, etc. Then, we are able to calculate the proportion of each status from the proposed model. In terms of the proportion of full-load and off-load status, we can utilize this information to aid the design of traffic congestion management scheme. Further, it provides insightful guidelines when deploying future cellular networks, *e.g.* deciding geographical placement and traffic dimensioning. In addition, when considering spatio-temporal inhomogeneity of traffic distribution, our model provides important insights on the dynamic energy-efficient operation of BS, where the operators shut some BSs and transfer the corresponding load to neighboring cells during periods of low utilization. Moreover, as the extension of the time-dependent pricing, the policies of the space-dependent pricing can be developed, in which the service providers made discounts in areas where the capacity of BSs is not sufficiently utilized to encourage users to generate larger data demand in these area. Thus, the utilization of the mobile network as well as the profit of the service providers can be improved.

4. RELATED WORK

Works related to our work can be divided by two topics: measurement analysis about the spatial traffic distribution, and measurement analysis about the temporal traffic distribution.

The spatial distribution of the cellular traffic has been studied in the literature [2–4]. Gotzner *et al.* [3] found that voice traffic in different cells of GSM networks can be described by a log-normal distribution. Michalopoulou *et al.* [4] presented a large scale data-driven traffic analysis based on data set from GSM networks in Germany. They found that data traffic loads in different cells of GPRS/EDGE networks can be approximated by log-normal mixtures. Different with them, we use the Complementary Cumulative Distribution Function (CCDF) rather than Probability Distribution Function (PDF) to represent the distribution, and describe it by the trimodal distribution to further character its long tail. Thus, our model is with more accuracy.

There have also been some works about the temporal distribution of mobile traffic, such as [6, 8]. Nan *et al.* [8] analyzed and statistically modeled the downlink throughput per cell distributions over time and over different cells based on real network throughput data. However, their data only records downlink throughput every 15 minutes for a week. While in our data, the timestamps are accurate to seconds and the duration is over a month, which is more large-scale and fine-grained. In addition, we analyze the characters of mobile traffic by combining both the spatial and temporal distributions in our work.

5. CONCLUSION AND FUTURE WORK

In this paper, we present our work on quantitatively characterizing inhomogeneity of the spatio-temporal distribution of data traffic in a large scale mobile cellular data network. Based on the large-scale, long-time and fine-grained data set, we provide a systematic measurement study on the spatio-temporal of mobile traffic. We find that the mobile traffic loads follow a trimodal distribution, which is the combination of compound-exponential, power-law and exponential distributions, in terms of both spatial and temporal dimension.

Our future work will address the fundamental questions of what are the key factors that influence the degree of inhomogeneity of the spatio-temporal traffic distribution. In addition, we will investigate the relationship between the spatial and temporal distribution of mobile traffic, and further develop a traffic model combining them to help us better understand the characters of mobile traffic. Meanwhile, we plan to study how to utilize our modelled behaviors of traffic to solve various practical issues in mobile networks, such as reducing congestion and improving energy efficiency.

6. REFERENCES

[1] Cisco Visual Networking Index, White Paper, Feb. 2015. available online at http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf

[2] D. Lee, S. Zhou, X. Zhong, Z. Niu, X. Zhou, and H. Zhang, "Spatial modeling of the traffic density in cellular networks," *Wireless Communications, IEEE*, vol. 21, no. 1, pp. 80-88, 2014.

[3] U. Gotzner, and R. Rathgeber, "Spatial Traffic Distribution in Cellular Networks," in *Proc. IEEE VTC*, 1998, pp. 1994–1998.

[4] M. Michalopoulou, J. Riihijarvi, and P. Mahonen, "Towards characterizing primary usage in cellular networks: A traffic-based study," In *Proc. DySPAN*, 2011, pp. 652–655

[5] S. Sen, C. Joe-Wong, S. Ha, and M. Chiang, "Incentivizing timeshifting of data: a survey of time-dependent pricing for internet access," *Communications Magazine, IEEE*, vol. 50, no. 11, pp. 91–99, 2012.

[6] M. Laner, P. Svoboda, S. Schwarz, "Users in cells: a data traffic analysis," *Proc. IEEE WCNC*, 2012, pp. 3063–3068.

[7] C. Peng, S. B. Lee, S. Lu, "Traffic-driven power saving in operational 3G cellular networks," in *Proc. ACM MOBICOM*, 2011, pp. 121–132.

[8] E. Nan, X. Chu, W. Guo, "User data traffic analysis for 3G cellular networks," in *Proc. IEEE CHINACOM*, 2013, pp. 468–472.