Broadband Traffic Report
Traffic Growth Slows to a Degree

1.1 Overview

In this report, we analyze traffic over the broadband access services operated by IIJ each year and present the results[*1][*2][*3][*4][*5][*6][*7][*8]. Here, we will once again report on changes in traffic trends over the past year, based on daily user traffic and usage by port.

Figure 1 shows the average monthly traffic trends for IIJ's overall broadband services and mobile services. The IN/OUT traffic indicates directions from the ISP perspective. IN represents uploads from users, and OUT represents user downloads. Because we cannot disclose specific traffic numbers, we have normalized the latest values for each set of OUT data to 1.

For broadband, over the past year, IN traffic has increased by 10%, while OUT traffic has increased by 25%. A year ago, each had increased by 18% and 47%, respectively, so it could be said that growth has slowed to a degree.

For mobile, we only have data for the past three years, but over the past year, IN traffic increased by 103% and OUT traffic increased by 70%. Although the increase has slowed compared to the 125% and 137% increases from a year ago, this still represents significant growth. However, the total volume of mobile traffic is still an order of magnitude lower than broadband.

1.2 About the Data

As with our previous reports, for broadband traffic, the survey data utilized here was collected using Sampled NetFlow from the routers that accommodate the fiber-optic and DSL broadband customers of our personal and enterprise broadband access services. For mobile traffic, access gateway billing information was applied to determine usage volumes for personal and enterprise mobile services, while Sampled NetFlow data from the routers used to accommodate these services was employed to determine the ports used.

Because traffic trends differ between weekdays and weekends, we analyze traffic in one-week chunks. In this case, we used data for the week spanning May 29 to June 4, 2017. For comparison, we used the data for the week spanning May 30 to June 5, 2016, which we analyzed in the previous report.

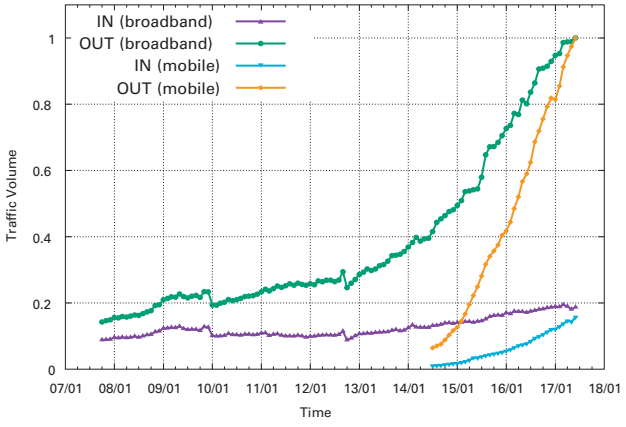

Figure 1: Trends in Monthly Traffic for Broadband and Mobile

Results are aggregated by subscription for broadband traffic, and by phone number for mobile traffic, as some subscriptions cover multiple phone numbers. The usage volume for each broadband user was obtained by matching the IP address assigned to users with the IP addresses observed. We gathered statistical information by sampling packets using NetFlow. The sampling rate was set as 1/8192 to 1/16382, taking into account router performance and load. We estimated overall usage volumes by multiplying observed volumes with the reciprocal of the sampling rate.

*1 Kenjiro Cho. Broadband Traffic Report: Traffic Growth is Accelerating. Internet Infrastructure Review. Vol.32. pp28-33. August 2016.
*2 Kenjiro Cho. Broadband Traffic Report: Comparing Broadband and Mobile Traffic. Internet Infrastructure Review. Vol.28. pp28-33. August 2015.
*3 Kenjiro Cho. Broadband Traffic Report: Traffic Volumes Rise Steadily Over the Past Year, and HTTPS Use Expands. Internet Infrastructure Review. Vol.24. pp28-33. August 2014.
*4 Kenjiro Cho. Broadband Traffic Report: The Impact of Criminalization of Illegal Downloads was Limited. Internet Infrastructure Review. Vol.20. pp32-37. August 2013.
*5 Kenjiro Cho. Broadband Traffic Report: Traffic Trends over the Past Year. Internet Infrastructure Review. Vol.16. pp33-37. August 2012.
*6 Kenjiro Cho. Broadband Traffic Report: Examining the Impact of the Earthquake on Traffic on a Macro Level. Internet Infrastructure Review. Vol.12. pp25-30. August 2011.
*7 Kenjiro Cho. Broadband Traffic Report: Traffic Shifting away from P2P File Sharing to Web Services. Vol.8. pp25-30. August 2010.
*8 Kenjiro Cho. Broadband Traffic Report: Increasing Traffic for General Users. Internet Infrastructure Review. Vol.4. pp18-23. August 2009.

IIJ provides both fiber-optic and DSL access as broadband services, but fiber-optic access now accounts for the vast majority of use. 97% of users observed in 2017 used fiber-optic connections, which makes up 99% of the overall broadband traffic volume.

1.3 Daily Usage Levels for Users

First, we will examine the daily usage volumes for broadband and mobile users from several perspectives. Daily usage indicates the average daily usage calculated from a week's worth of data for each user.

Figure 2 and Figure 3 show the average daily usage distribution (probability density function) per broadband and mobile user. Each compares data for 2016 and 2017 split into IN (upload) and OUT (download), with user traffic volume on the X axis, and user frequency on the Y axis. The X axis shows volumes between 10 KB (10^4) and 100 GB (10^{11}) using a logarithmic scale. Most users fall within the 100 GB (10^{11}) range, with a few exceptions.

The IN and OUT distribution for broadband traffic (Figure 2) shows an almost log-normal distribution, which looks like a normal distribution in a semi-log graph. A linear graph would show a long-tailed distribution, with the peak close to the left end and a slow gradual decrease towards the right. The OUT distribution is further to the right than the IN distribution, indicating that the download volume is more than an order of magnitude larger than the upload volume. When comparing 2016 and 2017, the peak distribution for both IN and OUT traffic has moved slightly to the right, which indicates that overall user traffic volumes are increasing.

Looking at the OUT distribution on the right, the peak has been steadily moving to the right over the past few years. However, the usage levels of heavy users on the right end have not increased much, and the distribution is becoming less symmetric. Meanwhile, the IN distribution on the left is symmetric, and closer to log-normal distribution.

The data for mobile traffic (Figure 3) indicates that usage volumes are significantly lower than broadband. Additionally, because there are limits on data usage, the ratio of heavy users to the right of the distribution is lower, and it is asymmetric. There are also no extremely heavy users. Due to those who use mobile only when going out, as well as limits on data usage, there is a greater variance in daily usage volumes for each user when compared to broadband. For this reason, when you look at the daily average for a week worth of data, there is less variance between users than when examining individual days. Plotting the distributions for individual days in the same way results in slightly lower peaks, and raises the tails on both sides, but the basic spread and modal values of the distribution remain largely unchanged.

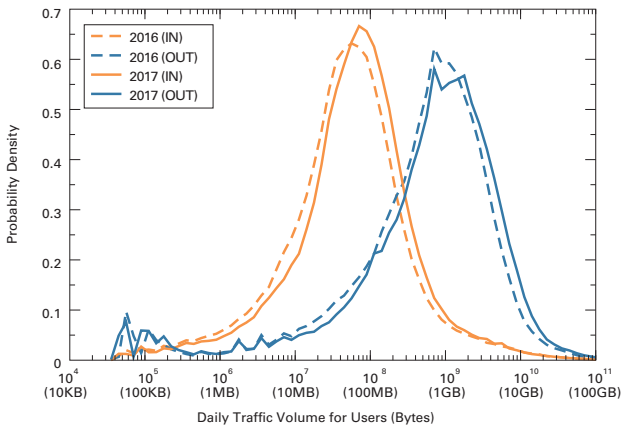

Figure 2: Daily Broadband User Traffic Volume Distribution Comparison of 2016 and 2017

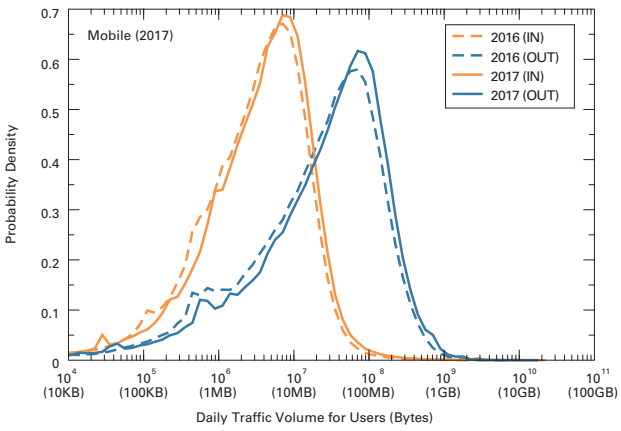

Figure 3: Daily Mobile User Traffic Volume Distribution Comparison of 2016 and 2017

Table 1 shows trends in the mean and median daily traffic values for broadband users as well as the modal value (the most frequent value that represents the peak of the distribution). The peak position had shifted slightly from the center of the distribution, so the modal value was corrected towards the center of the distribution. Comparing the modal values in 2016 and 2017, IN rose from 56 MB to 79 MB, and OUT rose from 1,000 MB to 1,260 MB. This means that IN and OUT traffic grew by a factor of 1.4 and 1.3, respectively. Meanwhile, because mean values are influenced by the heavy users to the right of the graph, they were significantly higher than the modal values, with the mean IN value being 520 MB and the mean OUT value being 2,624 MB in 2017. In 2016, these were 475 MB and 2,081 MB, respectively. For mobile traffic (Table 2), the mean and modal values are close due to the lack of heavy users. In 2017, the modal values were 7 MB for IN and 79 MB for OUT, while the mean values were 12 MB for IN and 77.4 MB for OUT. The modal values were identical to the previous year for IN traffic, and grew by a factor of 1.3 for OUT traffic.

Figure 4 and Figure 5 plot per user IN/OUT usage volumes for a random sample of 5,000 users. The X axis shows OUT (download volume) and the Y axis shows IN (upload volume), with both using a logarithmic scale. Users with identical IN/OUT values are plotted on the diagonal line.

The cluster below spread out parallel to the diagonal line represents general users with download volumes an order of magnitude higher than upload volumes. For broadband traffic, there was previously a clearly-recognizable cluster of heavy users spread out thinly on the upper right of the diagonal line, but this is now no longer discernible. There are also differences in the usage levels and the IN/OUT ratio for each user, pointing to the existence of diverse forms of usage. Here, almost no difference can be discerned compared to 2016.

The trend for OUT traffic being an order of magnitude larger also applies to mobile, but usage volumes are lower than broadband, and there is less variance between IN/OUT. The

	IN (MB/day)			OUT (MB/day)		
Year	Mean	Median	Mode	Mean	Median	Mode
2005	430	3	3.5	447	30	32
2007	433	5	4	712	58	66
2008	483	6	5	797	73	94
2009	556	7	6	971	88	114
2010	469	8	7	910	108	145
2011	432	9	8.5	1,001	142	223
2012	410	12	14	1,026	173	282
2013	397	14	18	1,038	203	355
2014	437	22	28	1,287	301	447
2015	467	33	40	1,621	430	708
2016	475	48	56	2,081	697	1,000
2017	520	63	79	2,624	835	1,260

Table 1: Trends in Mean and Modal Values for the Daily Traffic Volume of Broadband Users

	IN (MB/day)			OUT (MB/day)		
Year	Mean	Median	Mode	Mean	Median	Mode
2015	6.0	2.7	5.5	46.6	19	40
2016	7.8	3.6	7	63.0	27	63
2017	12.0	4.3	7	77.4	35	79

Table 2: Trends in Mean and Modal Values for the Daily Traffic Volume of Mobile Users

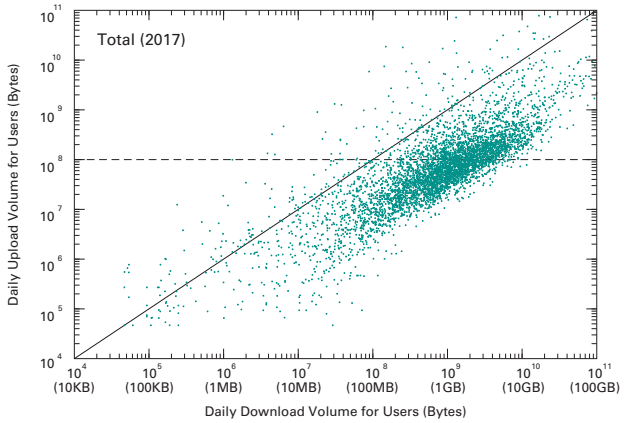

Figure 4: IN/OUT Usage for Each Broadband User

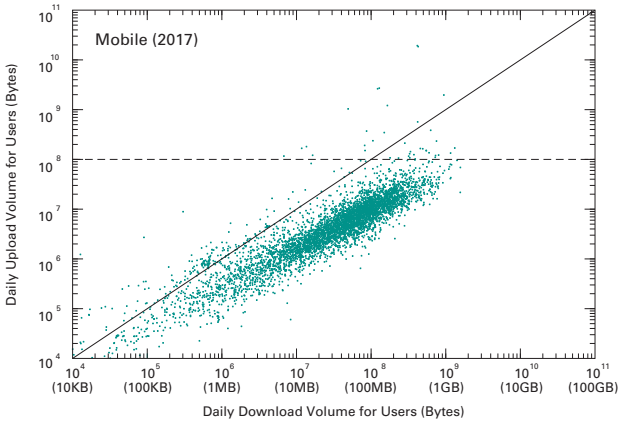

Figure 5: IN/OUT Usage for Each Mobile User

slope of the cluster is also now lower than the diagonal line, indicating that download ratios are relatively higher for users with higher usage levels. Compared to last year, there are more users with a higher upload volume, which can be seen scattered across the top away from the cluster. This is probably due to a small group of users beginning to stream video from mobile devices.

Figure 6 and Figure 7 show the complementary cumulative distribution of the daily traffic volume for users. This indicates the percentage of users with daily usage levels greater than the X axis value on the Y axis in a log-log scale, which is an effective way of examining the distribution of heavy users. The right side of the graph falls linearly, showing a long-tailed distribution close to a power-law distribution. It can be said that heavy users are distributed statistically, and are by no means a special class of user. On mobile, heavy users exhibit power-law distribution for OUT traffic, but the linear slope for IN traffic is in more disarray than last year, and a larger proportion of users are uploading large volumes of data.

There is a great deal of deviation in traffic usage levels between users, and as a result, traffic volume from a small portion of users accounts for the majority of overall traffic. For example, the top 10% of broadband users account for 60% of the total OUT traffic and 87% of the total IN traffic, while the top 1% of users account for 25% of the OUT traffic and 59% of the IN traffic. However, along with the decrease in the ratio of heavy users over the past few years, the distribution bias has been in a slightly declining trend. Meanwhile, on mobile the top 10% of users account for 48% of OUT traffic and 62% of IN traffic, while the top 1% account for 13% of OUT traffic and 39% of IN traffic. Although the ratio of heavy IN traffic users has increased over the past few years, this demonstrates that the overall proportion of heavy users is smaller on mobile than broadband.

1.4 Usage by Port

Next, we will look at a breakdown of traffic and examine usage levels by port. Recently, it has been difficult to identify applications by port number. Many P2P applications use dynamic ports on both ends, and a large number of client/server applications utilize port 80, assigned to HTTP in order to avoid firewalls. For this reason, generally speaking, when both parties use a dynamic port that is port 1024 or higher, there is a high possibility of it being a P2P application, and when one of the parties uses a well-known port lower than port 1024, it is likely to be a client/server application. In light of this, here we will look at usage levels for TCP and UDP connections by port number, taking the lower port number of the source and destination ports.

As heavy users account for the majority of overall traffic, in order to identify the trends for general users, we have taken a rough approach by extracting data for users with a daily upload volume of less than 100 MB, and treating them as light users. This constitutes users below the horizontal line IN=100 MB point in Figure 4, which roughly corresponds to mobile user usage levels.

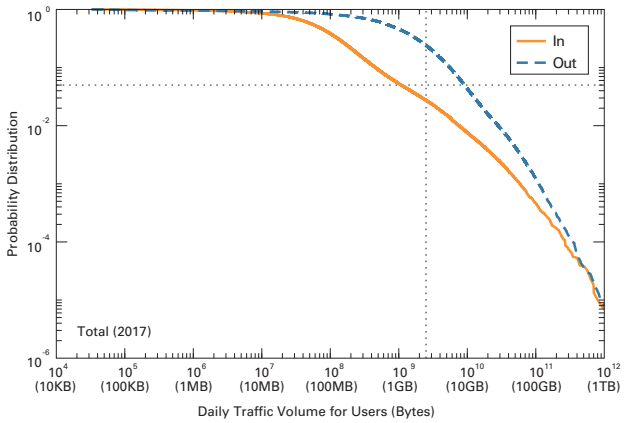

Figure 6: Complementary Cumulative Distribution of the Daily Traffic Volume for Broadband Users

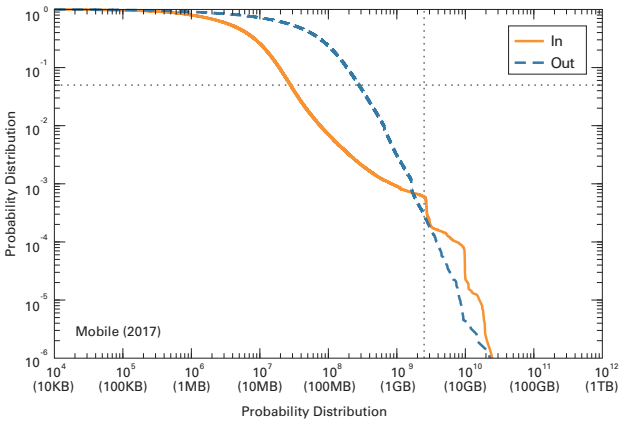

Figure 7: Complementary Cumulative Distribution of the Daily Traffic Volume for Mobile Users

Table 3 compares port usage ratios for broadband users in 2016 and 2017, showing all users and light users. 84% of the overall traffic in 2017 is TCP based. The ratio of port 443 (HTTPS) traffic has increased from 31% to 43%, and has finally overtaken port 80 (HTTP) traffic. The ratio of port 80 (HTTP) traffic decreased from 37% in 2016 to 28%, demonstrating that the gradual migration from HTTP to HTTPS has progressed further. TCP dynamic port traffic, which has been on the decline, dropped from 14% in 2016 to 11% in 2017. The ratio of individual dynamic port numbers is tiny, with port 1935 used by Flash Player being the highest at 1% of the total, and the next highest is under 0.3%. Aside from TCP, there is port 443 (HTTPS) UDP traffic, which is thought to be Google's QUIC protocol. The rest is mostly VPN related.

Looking exclusively at light users, port 443 HTTPS traffic, which accounted for 40% of the total in 2016, increased 13 percentage points to 53% in 2017, and port 80 traffic that accounted for 49% of the total in 2016, decreased 14 percentage points to 35% in 2017. The differences between port usage ratios for light users and overall traffic are now disappearing.

Table 4 shows port usage ratios for mobile users, which as a whole were also close to the values for broadband users, suggesting that mobile users are also using applications in a similar manner to broadband users.

The growth in the use of HTTPS is due to more and more services mainly based in the United States making regular use of encrypted HTTPS communications, since the existence of a controversial U.S. National Security Agency (NSA) program for intercepting communications came to light in June 2013. Looking at HTTPS traffic volumes broken down by provider in the data for 2017, about half were related to Google. Google accounted for around 70% of the volume last year, so we can see that other companies have also made progress in the shift to HTTPS.

Figure 8 compares trends in TCP port usage over a week for overall broadband traffic in 2016 and 2017. Trends in TCP port usage are shown for four categories: port 80, port 443, other well-known ports, and dynamic ports. The peak overall traffic volume is shown as 1. Compared with 2016, we can see that the overall ratio of port 443 usage has increased further, and the use of dynamic ports is decreasing. The overall peak is between 19:00 and 23:00, with the peak for port 443 occurring slightly earlier than that for port 80. Traffic increases in the daytime on Saturday and Sunday, reflecting times when the Internet is used at home.

protocol port	2016		2017	
	total (%)	light users	total (%)	light users
TCP	**82.8**	**93.3**	**83.9**	**92.3**
(< 1024)	63.3	89.9	72.9	88.6
443 (https)	30.5	39.6	43.3	52.5
80 (http)	37.1	49.2	28.4	35.2
182	0.3	0.2	0.3	0.3
81	0.4	0.7	0.2	0.2
993 (imaps)	0.1	0.1	0.2	0.1
22 (ssh)	0.2	0.0	0.1	0.0
110 (pop3)	0.1	0.1	0.1	0.1
(>= 1024)	13.7	3.2	11.0	3.7
1935 (rtmp)	1.5	1.7	1.1	1.2
8080	0.2	0.1	0.3	0.1
UDP	**11.1**	**4.0**	**10.5**	**4.9**
443 (https)	2.4	2.8	3.8	3.7
4500 (nat-t)	0.2	0.1	0.2	0.1
ESP	**5.8**	**2.6**	**5.1**	**2.7**
IP-ENCAP	**0.2**	**0.0**	**0.3**	**0.0**
GRE	**0.1**	**0.0**	**0.1**	**0.0**
ICMP	**0.0**	**0.0**	**0.0**	**0.0**

Table 3: Usage by Port for Broadband Users

protocol port	2016	2017
	total (%)	total (%)
TCP	**94.4**	**84.4**
443 (https)	43.7	53.0
80 (http)	46.8	27.0
31000	0.2	1.8
993 (imaps)	0.5	0.4
1935 (rtmp)	0.3	0.2
81	0.5	0.1
UDP	**5.0**	**11.4**
443 (https)	1.5	7.5
12222	0.1	0.1
4500 (nat-t)	0.2	0.2
53 (dns)	0.2	0.1
ESP	**0.4**	**0.4**
GRE	**0.1**	**0.1**
ICMP	**0.0**	**0.0**

Table 4: Usage by Port for Mobile Users protocol port

Figure 9, mobile data shows trends for port 80 and port 443, which account for the majority of overall traffic. On mobile, the ratio of port 443 traffic has increased further. Compared to broadband, high traffic levels continue from morning to night. On weekdays, there are three peaks representing the morning commute hours, lunch breaks, and the evening from 17:00 to 22:00, demonstrating that usage times are different from broadband.

1.5 Conclusion

One of the broadband traffic trends for this year we can identify is that the increase in traffic that had been accelerating in recent years has started to slow down. Over the past year download volumes climbed 25% and upload volumes went up by 10%, showing continued growth, but growth rates are down compared to the 47% and 18% increases each saw last year. One reason for this is there were a small number of big updates or popular new streaming services appearing this year. Also, on the whole, the shift towards larger and more frequent software updates has settled down, and the adoption of flat-rate music and video streaming services has come full circle.

Despite the growth rate of mobile traffic falling slightly, it has still flourished over the past three years. It is different from broadband traffic in several ways, such as there being fewer heavy users, and higher levels of usage during weekday commute hours and lunch breaks.

The use of HTTPS has expanded greatly compared to three years ago, and finally exceeds HTTP, with 43% of broadband and 53% of mobile traffic sent using HTTPS. But a large amount of commercial content still hasn't been converted to HTTPS, so we expect that more progress will be made in this area in the coming years.

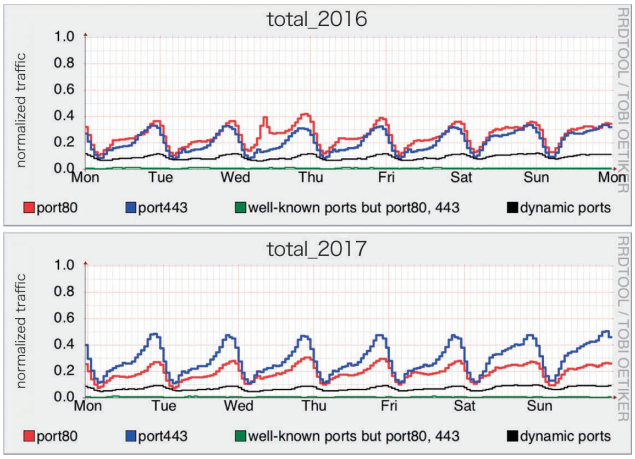

Figure 8: Weekly TCP Port Usage Trends for Broadband Users 2016 (top) and 2017 (bottom)

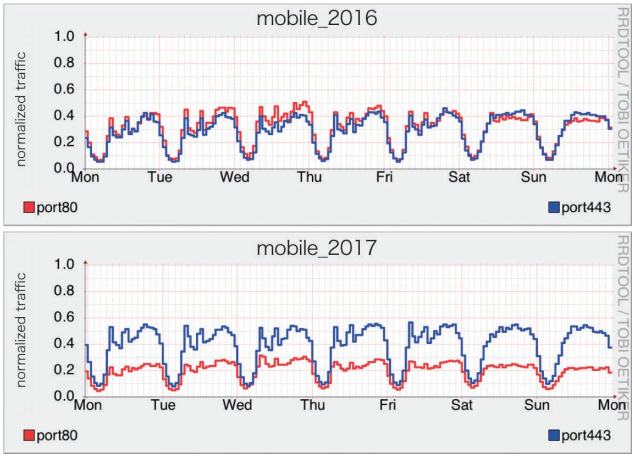

Figure 9: Weekly TCP Port Usage Trends for Mobile Users 2016 (top) and 2017 (bottom)

Author:
Kenjiro Cho
Research Director, Research Laboratory, IIJ Innovation Institute Inc.