Broadband Traffic Report: Download Growth Slows for a Second Year Running

1.1 Overview

In this report, we analyze traffic over the broadband access services operated by IIJ and present the results each year[*1][*2][*3][*4][*5][*6][*7][*8][*9]. Here, we again report on changes in traffic trends over the past year, based on daily user traffic and usage by port.

Figure 1 shows the overall average monthly traffic trends for IIJ's broadband services and mobile services. IN/OUT indicates the direction 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 data, setting the latest OUT observation in each dataset to 1. Starting with this edition of the report, the broadband data include IPv6 IPoE traffic. The thin line labeled "broadband-IPoE" excludes IPv6 IPoE traffic. IPv6 traffic on IIJ's broadband services comprises both IPoE and PPPoE traffic[*10], but IPoE traffic does not pass directly through IIJ's network as we use Internet Multifeed Co.'s transix service for IPoE, and IPoE is therefore excluded from the analysis that follows here. As of June 2018, IPoE accounted for 12% of IN and 8% of OUT broadband traffic overall.

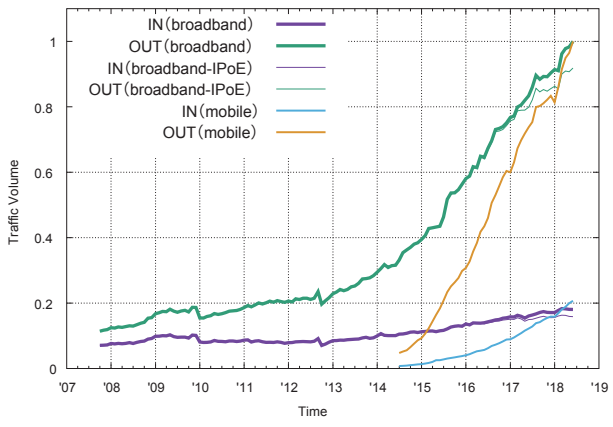

Figure 1: Monthly Broadband and Mobile Traffic over Time

Growth in both broadband and mobile traffic slowed temporarily in the latter half of last year, but that growth has picked up again this year and returned to its original trajectory. Over the past year, broadband IN traffic increased 12% and broadband OUT traffic increased 20%. The respective figures a year earlier were 10% and 25%, and two years earlier 18% and 47%, so growth in download volume has been slowing for two years running. For mobile, we only have data for the past four years. Mobile IN traffic increased 69% and OUT traffic increased 36% over the past year. Although these figures represent a slowing of growth compared with 103% and 70% a year ago, the level of growth remains high. That said, the total volume of mobile traffic remains an order of magnitude lower than broadband traffic.

1.2 About the Data

As with previous reports, for broadband traffic, our analysis uses data sampled 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, we use access gateway billing information to determine usage volumes for personal and enterprise mobile services, and we use Sampled NetFlow data from the routers used to accommodate these services to determine the ports used.

Because traffic trends differ between weekdays and weekends, we analyze traffic in one-week chunks. In this report, we look at data for the week of May 28 through June 3, 2018, and compare those data with data for the week of May 29 through June 4, 2017, which we analyzed in the previous edition of this report.

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

*1 Kenjiro Cho. Broadband Traffic Report: Traffic Growth Slows to a Degree. Internet Infrastructure Review. Vol.36. pp4-9. August 2017.
*2 Kenjiro Cho. Broadband Traffic Report: Traffic Growth is Accelerating. Internet Infrastructure Review. Vol.32. pp28-33. August 2016.
*3 Kenjiro Cho. Broadband Traffic Report: Comparing Broadband and Mobile Traffic. Internet Infrastructure Review. Vol.28. pp28-33. August 2015.
*4 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.
*5 Kenjiro Cho. Broadband Traffic Report: The Impact of Criminalization of Illegal Downloads was Limited. Internet Infrastructure Review. Vol.20. pp32-37. August 2013.
*6 Kenjiro Cho. Broadband Traffic Report: Traffic Trends over the Past Year. Internet Infrastructure Review. Vol.16. pp33-37. August 2012.
*7 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.
*8 Kenjiro Cho. Broadband Traffic Report: Traffic Shifting away from P2P File Sharing to Web Services. Vol.8. pp25-30. August 2010.
*9 Kenjiro Cho. Broadband Traffic Report: Increasing Traffic for General Users. Internet Infrastructure Review. Vol.4. pp18-23. August 2009.
*10 Akimichi Ogawa. Appendix A.3 "IPv6 PPPoE and IPv6 IPoE" in Professional IPv6 (in Japanese). Lambda Note. July 2018.

address assigned to users with the IP addresses observed. We gathered statistical information by sampling packets using NetFlow. Sampling rates were set between 1/8,192 and 1/16,382, taking into account router performance and load. We estimated overall usage volumes by multiplying observed volumes with the reciprocal of the sampling rate.

IIJ provides both fiber-optic and DSL broadband services, but fiber-optic access now accounts for the vast majority of use. Of users observed in 2018, 97% were using fiber-optic connections and accounted for 99% of overall broadband traffic volume.

1.3 Users' Daily Usage

First, we examine daily usage volumes for broadband and mobile users from several angles. 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 distributions (probability density functions) for broadband and mobile users. Each compares data for 2017 and 2018 split into IN (upload) and OUT (download), with user traffic volume plotted along the X-axis and user frequency along the Y-axis. The X-axis shows volumes between 10KB (10^4) and 100GB (10^{11}) using a logarithmic scale. Most users fall within the 100GB (10^{11}) range, with a few exceptions.

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

The peak of the OUT distribution, which appears toward the right in the plot, has steadily been moving rightwards over the past few years, but heavy-user usage levels have not increased much, and as a result, the distribution is becoming less symmetric. The IN distribution on the left, meanwhile, is generally symmetric and closer to a log-normal distribution.

The data for mobile traffic (Figure 3) indicate that usage volumes are significantly lower than for broadband. And limits on mobile data usage mean that heavy users, which fall on the right-hand side of the distribution, account for only a small proportion of the total, so the distribution is asymmetric. There are also no extremely heavy users. The variability in each user's daily usage volume is higher for mobile than for broadband owing to those users who only using mobile data when out of the home/office as well as the limits on mobile data. Hence, the daily average for a week's worth of data shows less variability between users than the data for individual days. Plotting the distributions for individual days in the same way results in slightly lower peaks and correspondingly higher tails on both sides, but the basic shape and modal values of the distribution remain largely unchanged.

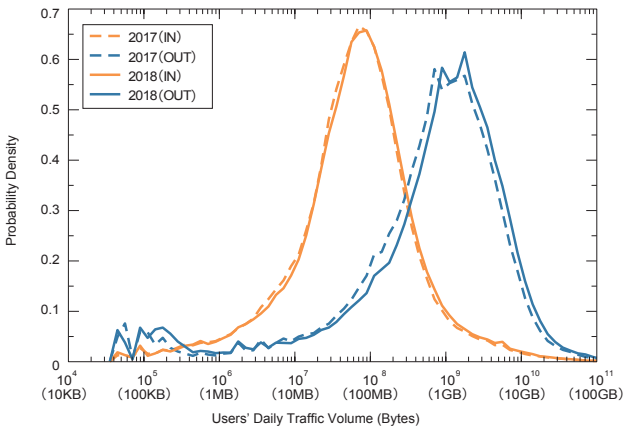

Figure 2: Daily Broadband User Traffic Volume Distribution Comparison of 2017 and 2018

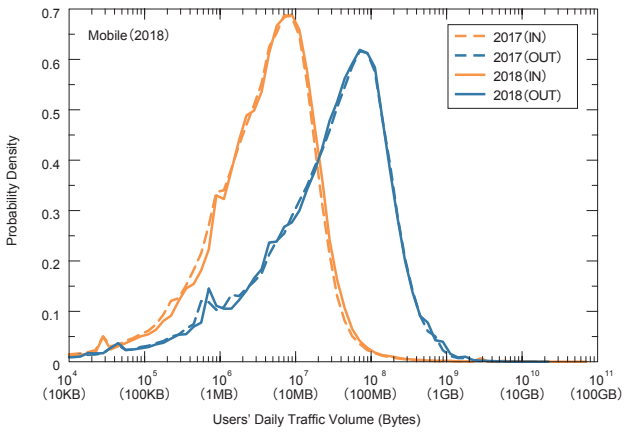

Figure 3: Daily Mobile User Traffic Volume Distribution Comparison of 2017 and 2018

Table 1 shows trends in the mean and median daily traffic values for broadband users as well as the mode (the most frequent value, which represents the peak of the distribution). The peak was slightly off from the center of the distribution, so the distribution was adjusted to bring the mode toward the center.

Comparing the values for 2017 and 2018, the IN mode was unchanged at 79MB, while the OUT mode rose from 1,260MB to 1,413MB, translating into growth factors of 1 and 1.1, respectively. Meanwhile, because the means are influenced by heavy users (on the right-hand side of the distribution), they were significantly higher than the corresponding modes, with the IN mean being 582MB and the OUT mean being 3,139MB in 2018. The 2017 means were 520MB and 2,624MB, respectively. For mobile traffic (Table 2), the mean and modal values are close owing to the lack of heavy users. In 2018, the IN mode was 7MB and the OUT mode was 79MB, while the means were IN 17.0MB and OUT 81.9MB. The modes for both IN and OUT traffic were identical to those for the previous year. The means increased despite there being very little change in the medians and modes, which indicates an increase in heavy users, particularly for IN traffic.

Figure 4 and Figure 5 plot per-user IN/OUT usage volumes for random samples of 5,000 users. The X-axis shows OUT (download volume) and the Y-axis shows IN (upload

Table 1: Trends in Mean and Mode of Broadband Users' Daily Traffic Volume

	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
2018	582	67	79	3,139	1021	1,413

Table 2: Trends in Mean and Mode of Mobile Users' Daily Traffic Volume

	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
2018	17.0	4.7	7	81.9	36	79

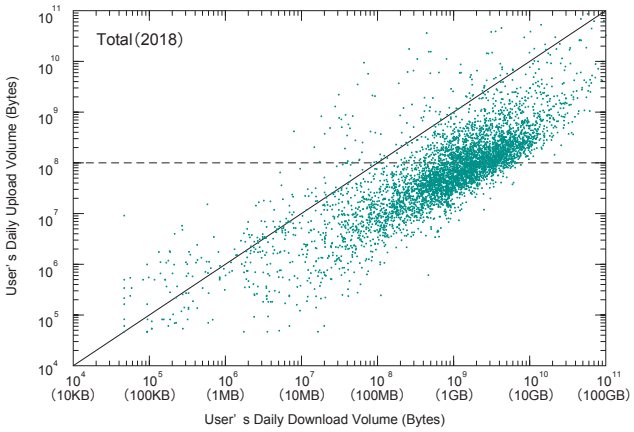

Figure 4: IN/OUT Usage for Each Broadband User

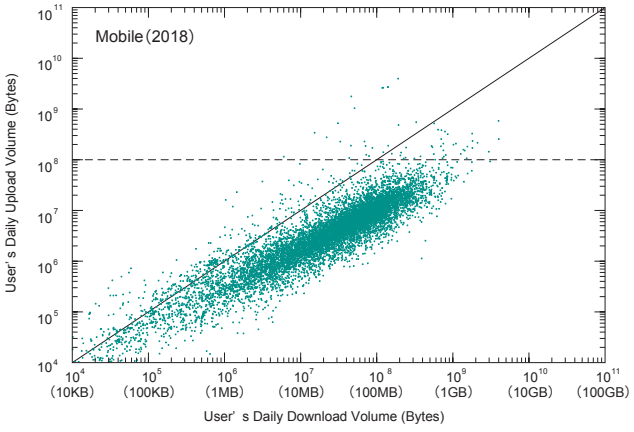

Figure 5: IN/OUT Usage for Each Mobile User

volume), with both using a logarithmic scale. Users with identical IN/OUT values fall on the diagonal.

The cluster spread out below and parallel to the diagonal in each of these plots represents typical 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 about the upper right of the diagonal, but this is now no longer discernible. Variability between users in terms of usage levels and IN/OUT ratios is wide, indicating that there is a diverse range of usage styles. Almost no difference can be discerned when these plots are compared with those for 2017.

For mobile traffic, the pattern of OUT being an order of magnitude larger also applies, but usage volumes are lower than for broadband, and there is less variability between IN and OUT. The slope of the mobile cluster is also less steep than the diagonal, indicating that download ratios tend to be higher at higher usage levels.

Figure 6 and Figure 7 show the complementary cumulative distribution of users' daily traffic volume. On these log-log plots, the Y-axis values represent the proportion of users with daily usage levels greater than the corresponding X-axis values. These plots are an effective way of examining the distribution of heavy users. The linear-like decline toward the right-hand side of the plots indicates that the distributions are long-tailed and close to a power-law distribution. Heavy users appear to be distributed statistically and do not appear to constitute a separate, special class of user. On mobile, heavy users appear to be distributed according to a power-law for OUT traffic, but the linear-like slope for IN traffic is more out of shape than it was last year, with a larger proportion of users uploading large volumes of data.

Traffic is heavily skewed across users, such that a small proportion of users accounts for the majority of overall traffic volume. For example, the top 10% of broadband users account for 60% of total OUT and 86% of total IN traffic, while the top 1% of users account for 25% of OUT and 59% of IN traffic. As the proportion of heavy users has declined over the past few years, the skew has also decreased, albeit only slightly. As for mobile, the top 10% of users account for 50% of OUT and 70% of IN traffic, while the top 1% account for 15% of OUT and 52% of IN traffic. The proportion of heavy users has steadily been increasing over the past few years.

1.4 Usage by Port

Next, we look at a breakdown of traffic and examine usage levels by port. Recently, it has become difficult to identify applications by port number. Many P2P applications use dynamic ports on both ends, and a large number of client/server applications use port 80, which is assigned to HTTP, to avoid firewalls. Hence, generally speaking, when both parties are using a dynamic port numbered 1024 or higher, the traffic is likely to be from a P2P application, and when one of the parties is using a well-known port lower than 1024, the traffic is likely to be from a client/server

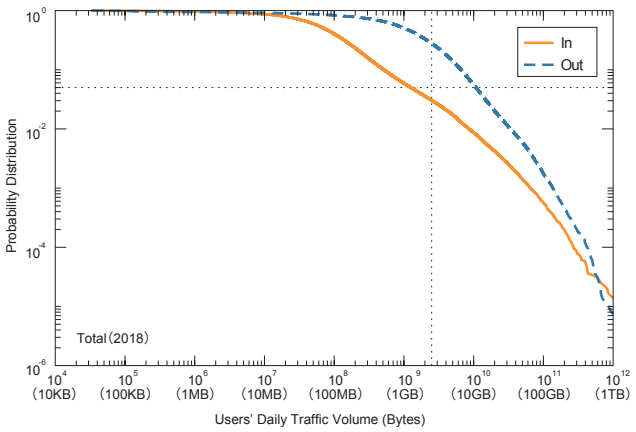

Figure 6: Complementary Cumulative Distribution of Broadband Users' Daily Traffic Volume

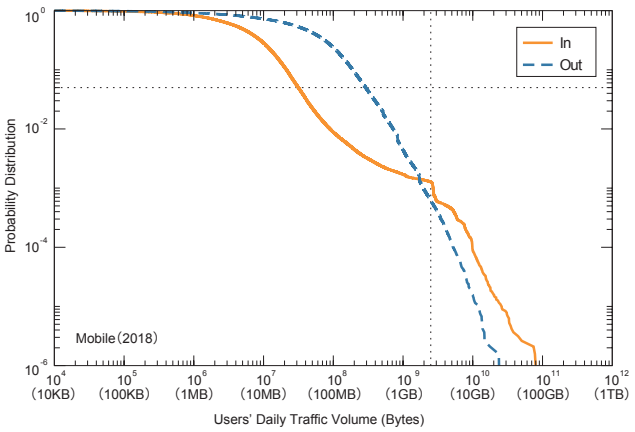

Figure 7: Complementary Cumulative Distribution of Mobile Users' Daily Traffic Volume

application. In light of this, we take the lower of the source and destination port numbers when breaking down TCP and UDP usage volumes by port.

Table 3 shows the percentage breakdown of broadband users' usage by port over the past four years. In 2018, 79% of all traffic was over TCP connections. The proportion of traffic over port 443 (HTTPS) had continued to increase up until the previous edition of this report, but it fell from 2017's 43% to 41% here. The proportion of traffic over port 80 (HTTP) also fell from 2017's 28% to 27% here, while the figure for UDP port 443, which is used by Google's QUIC protocol, rose from 11% to 16%. These figures demonstrate that the ongoing transition from HTTP to HTTPS is now turning toward QUIC. TCP dynamic port traffic, which has been on the decline, fell from 11% in 2017 to 10% in 2018. The proportion accounted for by individual dynamic port numbers is tiny, with port 1935, used by Flash Player, accounting for the largest share at around 0.7%, and the remaining port numbers accounting for less than 0.3%. As for non-TCP traffic, almost all of the traffic over ports other than UDP port 443 is VPN related.

Table 4 shows the percentage breakdown by port for mobile users. HTTPS accounts for a greater proportion of traffic here than with broadband, but the figures are close to those for broadband on the whole, suggesting that mobile users use applications in a manner similar to broadband users.

Figure 8 compares overall broadband traffic for key port categories across the course of the week from which observations were drawn in 2017 and 2018. We break the data into four port buckets: TCP ports 80 and 443, dynamic ports (1024 and up), and UDP port 443. In this edition, we take out the "well-known ports" bucket, since usage has dwindled, and add in UDP port 443 instead. The data are normalized so that peak overall traffic volume on the plot is 1. The overall peak is between 19:00 and 23:00 hours, with the peak for

Table 3: Broadband Users' Usage by Port

year	2015	2016	2017	2018
protocol port	(%)	(%)	(%)	(%)
TCP	**80.8**	**82.8**	**83.9**	**78.5**
(< 1024)	63.3	69.1	72.9	68.5
443(https)	23.3	30.5	43.3	40.7
80(http)	37.9	37.1	28.4	26.5
182	0.4	0.3	0.3	0.3
993(imaps)	0.1	0.1	0.2	0.2
22(ssh)	0.2	0.2	0.1	0.1
(>= 1024)	17.5	13.7	11.0	10.0
1935(rtmp)	1.8	1.5	1.1	0.7
8080	0.3	0.2	0.3	0.3
UDP	**11.4**	**11.1**	**10.5**	**16.4**
443(https)	0.9	2.4	3.8	10.0
4500(nat-t)	0.2	0.2	0.2	0.2
ESP	**7.4**	**5.8**	**5.1**	**4.8**
IP-ENCAP	**0.2**	**0.2**	**0.3**	**0.2**
GRE	**0.2**	**0.1**	**0.1**	**0.1**
ICMP	**0.0**	**0.0**	**0.0**	**0.0**

Table 4: Mobile Users' Usage by Port

year	2015	2016	2017	2018
protocol port	(%)	(%)	(%)	(%)
TCP	**93.8**	**94.4**	**84.4**	**76.6**
443(https)	37.4	43.7	53.0	52.8
80(http)	52.5	46.8	27.0	16.7
31000	0.0	0.2	1.8	2.9
993(imaps)	0.5	0.5	0.4	0.3
1935(rtmp)	0.5	0.3	0.2	0.1
UDP	**5.2**	**5.0**	**11.4**	**19.4**
443(https)	1.0	1.5	7.5	10.6
4500(nat-t)	0.3	0.2	0.2	4.5
12222	0.0	0.1	0.1	2.3
53(dns)	0.1	0.2	0.1	0.1
ESP	**0.7**	**0.4**	**0.4**	**3.9**
GRE	**0.3**	**0.1**	**0.1**	**0.1**
ICMP	**0.0**	**0.0**	**0.0**	**0.0**

port 443 coming just slightly earlier than the peak for port 80. Traffic increases during the daytime on Saturday and Sunday, reflecting household Internet usage times.

Figure 9 shows the trend for TCP ports 80 and 443 and UDP port 443, which account for the bulk of mobile traffic. When compared with broadband, we note that mobile traffic levels remain high throughout the day, from morning through night. The plot shows that usage times differ from those for broadband, with three separate mobile traffic peaks occurring on weekdays: morning commute, lunch break, and evening from 17:00 to 22:00 hours.

1.5 Conclusion

One identifiable trend in broadband traffic over the past year is that growth slowed somewhat in the latter half of last year but picked up again and returned to its upward trajectory in 2018. Over the past year, download volumes climbed 20% and upload volumes rose 12%, but growth in downloads has been slowing for two years in a row now.

Although the mobile traffic growth rate has fallen slightly, mobile traffic has still grown substantially over the past four years. Differences in comparison with broadband include the paucity of heavy mobile users and notably higher levels of mobile usage on weekdays during commute and lunch-break hours.

The use of HTTPS has expanded greatly since about four years ago, with TCP and UDP port 443 traffic combined accounting for 51% of broadband and 63% of mobile traffic. Given the increasing pressure to transition to HTTPS recently, with web browsers displaying messages saying the HTTP is unsafe and HTTP-only sites being pushed down the search engine rankings, we expect the decline in HTTP traffic to continue ahead.

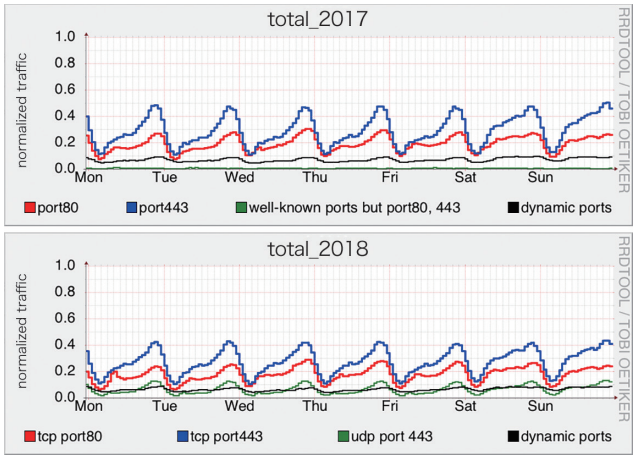

Figure 8: Broadband Users' TCP Port Usage Over a Week
2017 (top) and 2018 (bottom)

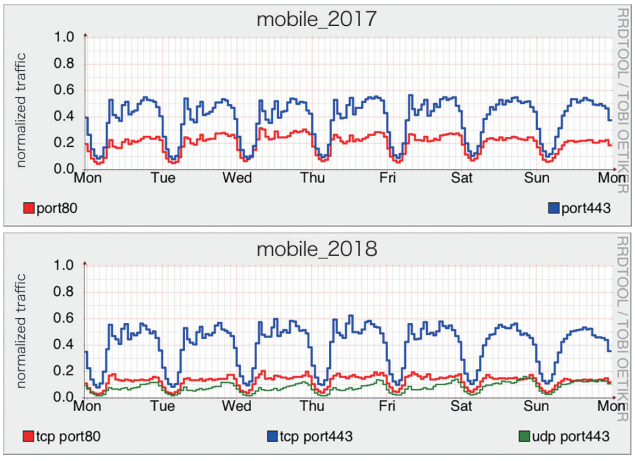

Figure 9: Mobile Users' TCP Port Usage Over a Week
2017 (top) and 2018 (bottom)

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