

3500 と 1 日前からのフォーラム

【あらまし】1999年5月に開設し、現在も継続しているフォーラム（BBS）について、投稿件数および文字数の推移から投稿活動の変遷を定量的に明らかにする。また、2006年5月以降のアクセスログを用いて、投稿活動とアクセスの中長期的な状況を明らかにする。1999年5月から数えて2015年12月が「200期」にあたること、2006年6月1日から2015年12月31日の期間の日数が「3500と1日」であることがわかったので分析する。ただし本稿はドラフトである。

【キーワード】mb_strlen()関数、ページビュー数、ユニーク訪問者数、累積確率、ロジスティック曲線

A Study for Long-term Analysis for Volunteer-based Knowledge Sharing; 3,501 Days of Our Forum

Abstract This is a draft. There is a forum (BBS on the Internet) in <http://atos.neorail.jp/atos3/forum/> (in Japanese). This forum opened in May 1999 and still continues. Many varieties of topics about computer-based systems and services in railways around Tokyo have been on track, it has attracted many students of engineering or informatics. In reading this report, we describe characteristics of articles posted within just 200 months, and we contrast between bibliometric indices of 3,132 posts and actual access logs by 146,422 unique visitors in latest 3,501 days.

Keywords mb_strlen(), pageviews, unique visitors, cumulative probability, logistic curves

1. はじめに

インターネット上のサービスやサイトの状況を測定する指標として、一般にページビュー数（PV）およびユニーク訪問者数（Unique Visitors：またはUU）が用いられている。これらの指標を広告料金の設定などに用いる商業的なサイトに限らず、個人のサイトにおいても、サイトやページの認知度や反響を推定する指標として、ページビュー数およびユニーク訪問者数は有用であるとみなされている。しかし、条件の異なるページ間やサイト間で指標の実数を比較したり、ページごとに計測した指標をサイト全体で合算したりするなど不適切に取り扱われる例や、もっぱら短期的な動向にばかり注目される例などが少なくない。2,000ページを超える数のページについて個別に分析した例や、5年（60箇月）を超える中長期的なアクセスの状況について分析した例が公表されることがほとんどないことが遠因になっていると考えられる。

本分析では、1999年5月に開設し、現在も運営を継続しているフォーラム（<http://atos.neorail.jp/atos3/forum/>）について、投稿件数および文字数の推移から投稿活動の変遷を定量的に明らかにする。また、2006年5月以降のアクセスログを用いて、投稿活動とアクセスの中長期的な状況を明らかにする。1999年5月を「1期」として月単位で数えて2015年12月が「200期」にあたること、および、2006年6月1日から2015年12月31日の期間の日数が「3500と1日」であることがわかったので分析する。以下、2節で背景について説明し、3節で分析方法とその結果について詳述し、4節で考察し、5節でまとめる。

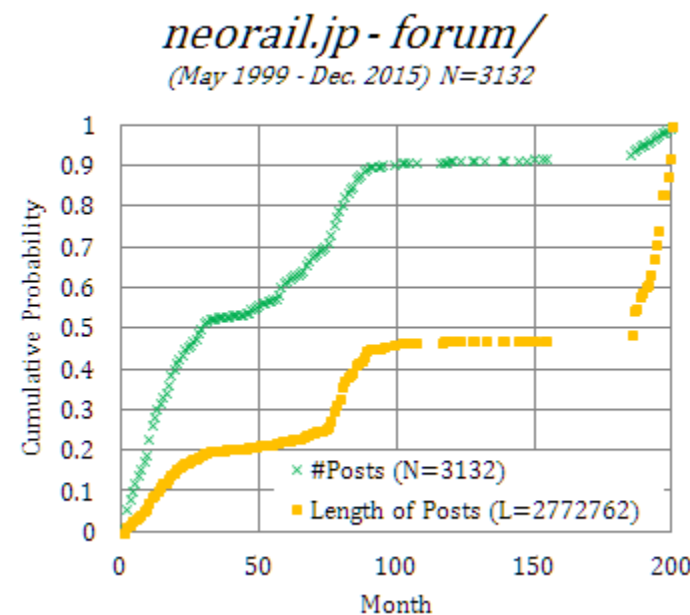


図1 月ごとの投稿件数と合計文字数の推移

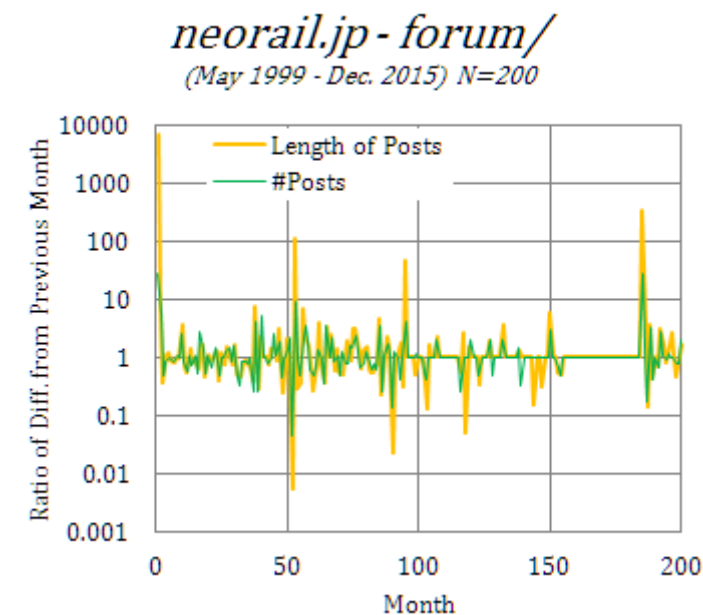


図2 投稿件数と合計文字数の前月に対する増減率

2. 背景

1999年5月から2005年9月まで（1-77期）の投稿について分析した結果を、2005年10月に公表している^{*1}。しかし、この時点では、投稿1件ごとのアクセスログを蓄積していなかったため、投稿活動とアクセスの状況を詳細に分析することはできなかった。これを受け、アクセスを記録するプログラムを新たに作成するとともに、フォーラムの過去の投稿を探しやすくする機能の開発を段階的に進めてきた。

フォーラムのアクセスログは、Cookieを使用してユニーク訪問者を識別する方式としている。アクセスを記録するプログラムは、2006年3月22日までにサイト全体で使用を開始していたが、フォーラムの投稿1件ごとに対する記録を開始したのは2006年5月25日からである。本分析では2006年6月1日以降のアクセスログを用いる。フォーラムの投稿番号とアクセスログの利用可能な期間の対応を表2に示す。

フォーラムにおける運営上の主な変更点を表3に示す。2006年9月までに現在のフォーラムのデザインや機能が完成している。これらの段階的な開発は、新規の訪問者の支援が目的である。

2003年2月（46期）には、累計投稿件数は1,703、累計文字数は582,218に達しており、新規の訪問者がプログラムによる支援なしでフォーラムを閲覧することは非常に困難になっていた。フォーラムの過去の投稿を探しやすくするため、2003年3月（47期）から全文検索機能を利用可能としている。また、相互に参照関係にある投稿の間を遷移できるようにするため、2005年9月に引用索引機能を利用可能としている。このほか、2006年2月までに掲載基準の確立などを行ない、利用規約を整備している。なお、投稿は管理者による閲読を経て掲載する体制としている。

フォーラムへのアクセスに関係することがらについて、以下にまとめて述べる。フォーラムを含むサイト全体について、1999年10月に「Yahoo! Japan」のディレクトリに掲載されたほか、2006年5月からは、新しい投稿を掲載したときに主要なブログ検索エンジンに対して「Ping」を送信している。また、投稿1件ごとの個別表示ページについて、検索エンジンによる収集を許可している。この結果、少なくともGoogleでは、遅くとも3日以内にフォーラムの投稿が検索可能となっており、実際に、大多数の利用者が検索エンジンを経由して訪れる状況になっている。近年、ソーシャルメディアからの遷移もみられることから、2014年12月には投稿1件ごとの個別表示ページに「ソーシャルボタン」を設けている。

*1 前回の分析：http://atos.neorail.jp/atos3/forum/forum_statistics.html

Date	May 1999 -	Jul. 2003 -	Sep. 2007 -	- Dec. 2015
Month	1 - 50	51 - 100	101 - 150	151 - 200
IDs of Post	1 - 1773	1774 - 2866	2867 - 2903	2904 - 3155

Tab.1 Periods for every 50month

Date	May 2006 -	Sep. 2009 -	- Dec. 2015
Month	1 - 84	85 - 124	125 - 184
IDs of Post	1 - 2748	2749 - 2886	2887 - 2903
Access Logs	N/A	Cookie-based Logs are Available	

Tab.2 Extracting Periods from Access Logs

Month	Date	Improvements
6	Oct. 1999	This website listed in the Yahoo! Japan directory.
47	Mar. 2003	Full-text search for all of older posts had been available.
75	Jul. 2005	Instead of aggregated views per every 10 posts, all posts are shown in individual pages; measuring pageviews of each posts became possible.
77	Sep. 2005	Citation Index for all posts had been available; non-linear access to older or later posts became easier.
86	May 2006	Site renewed to the design now in use. RSS had been available, and sending a Ping to major search engines are started.
89	Sep. 2006	Associative Content Navigation is embedded; all hyperlinks are built automatically.
188	Dec. 2014	Social sharing buttons are added to individual pages.
190	Feb. 2015	Selected news articles are shown in root index page.

Tab.3 History of This Forum

3. 分析

フォーラムの運営期間は、1999年5月から2016年12月まで200月にわたる。本分析では、1箇月を1期とする。50期ごとの投稿件数を表1に示す。また、アクセスログは、85期の途中にあたる2006年5月25日から蓄積を開始していたが、本分析では86-200期におけるアクセスログを分析対象とした。フォーラムの投稿番号とアクセスログの利用可能な期間の対応を表2に示す。

アクセスログでは、Cookieによりユニーク訪問者を識別している。同一の利用者であっても、PCとモバイル端末などを個別に計上するほか、PC上で異なるブラウザを使い分けた場合も個別に計上する。また、Cookieの有効期限は100日に設定していた。同一の利用者であっても、100日を超えてアクセスしなかった期間を挟む場合は、新たに別のユニーク訪問者として計上される。最大で、1名の利用者が年間に3ユニーク訪問者として計上されうる。計上されたユニーク訪問者数を約57.6(年間3回×9.6年×端末2種類)で除した値が実際の利用者(定期的に再訪する利用者)の数と推定され、2,544人となる。

3.1 投稿件数と文字数

月ごとの投稿件数と合計文字数の推移を図1に示す。1-50期、51-100期、101-150期と、約50期のスパンで投稿活動が変動していることが示唆される。150期以降は投稿件数が0の月が続き、185期から新しい投稿活動が立ち上がっている。投稿件数の累積確率は、29期に50.1%、50期に55.9%、85期に87.0%、99期に90.8%、154期に92.0%で、185-200期は8.0%にあたる。合計文字数の累積確率は、50期に21.5%、99期に46.3%、154期に47.5%で、185-200期は52.5%にあたる。なお、文字数は、Shift_JISで蓄積している投稿の本文(見出しを含まない)をUTF8に変換し、PHPのmb_strlen()関数により得られた値を用いる。

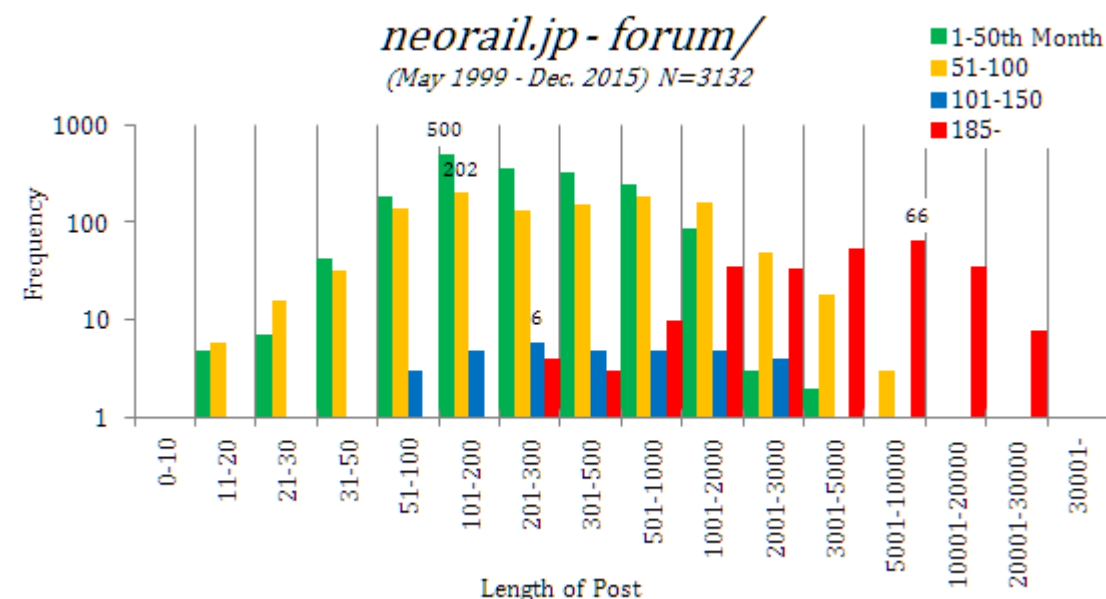


図3 投稿あたりの文字数の分布

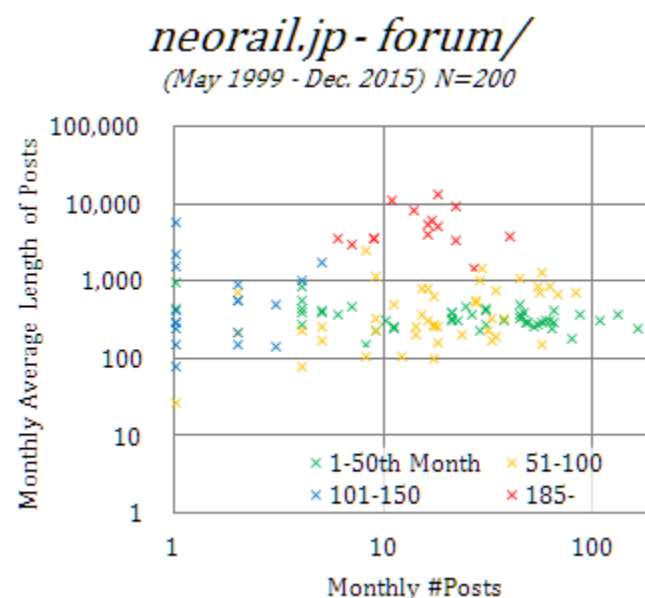


図4 月ごとの投稿件数と平均文字数

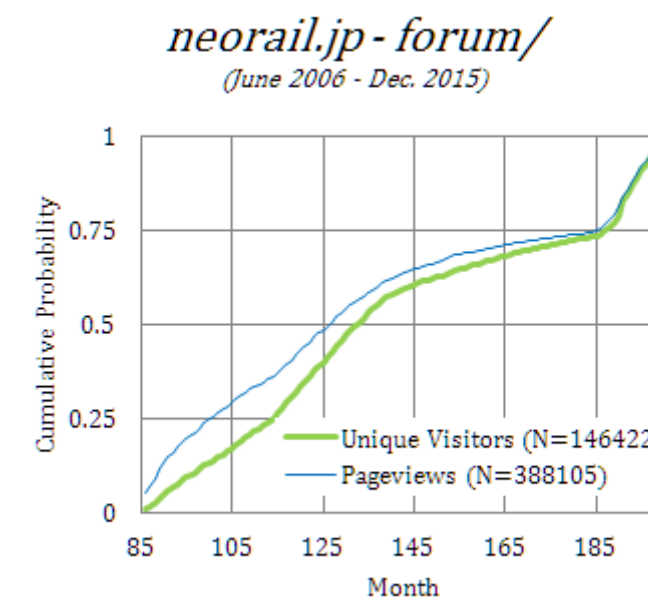


図5 ページビュー数、ユニーク訪問者数の推移

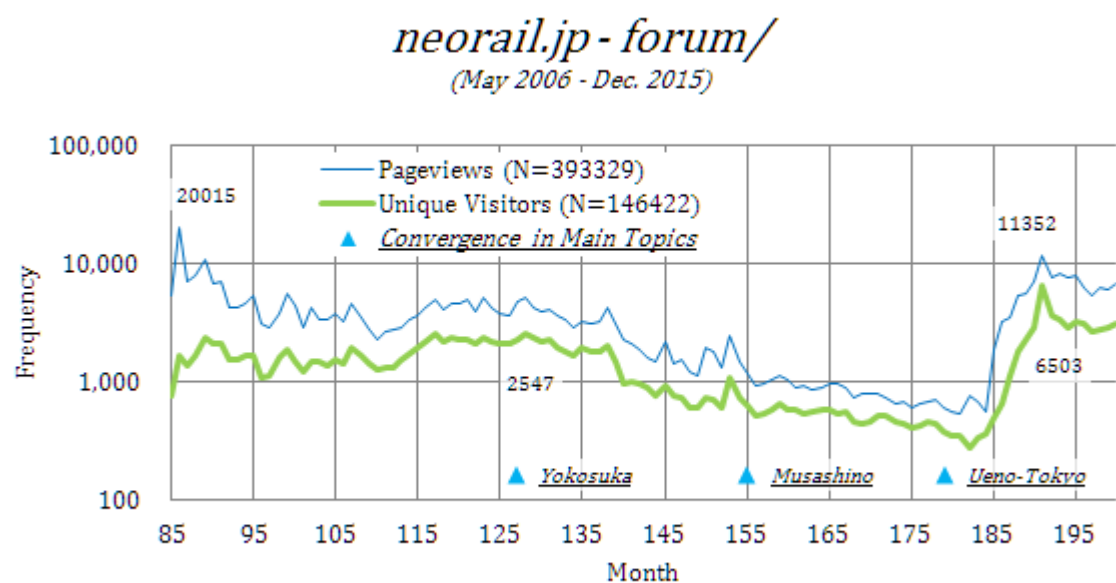


図6 月ごとのページビュー数およびユニーク訪問者数

件である。151-184期は除外した。185-200期では、201-30,000字の範囲に分布しており、5,001-10,000字の投稿が最多で66件である。

月ごとの投稿件数と平均文字数の対応を図4に示す。表1に示した期間ごとに色分けして示した。全体では、投稿件数と平均文字数の間に相関は認められない。表1に示した期間ごとでは、図1で示唆された投稿活動の変動が、平面上の領域として観測できる。

3.2 ページビュー数とユニーク訪問者数

86-200期におけるページビュー数、ユニーク訪問者数の推移を図5に示す。ページビュー数は月ごとの値を累計、ユニーク訪問者数は、86期から当期までの値を用いた。85期のアクセスログはデータ日数が6日であることから、分析に用いない。月ごとのページビュー数およびユニーク訪問者数を図6に示す。ページビュー数は、86期に20,015で最多、次いで191期に11,352で2位である。ユニーク訪問者数は、191期に6,503で最多、次いで128期に2,547で2位である。ユニーク訪問者数は、182期に275で最小となっている。図1に示したように、新しい投稿活動は185期から始まっているが、月ごとのユニーク訪問者数は183期から上昇に転じている。

フォーラムの投稿活動およびアクセス状況には、訪問者の関心が影響していると考えられる。フォーラムの訪問者が関心を持つと考えられる主なトピックを表4に示す。トピックについては、収束した時点基準として扱う。このため、現在までに収束したとみなせないトピックは除外する。155-184期の投稿件数は0であるが、ページビュー数の平均は793.0、月ごとのユニーク訪問者数の平均は474.9となっている。

投稿およびアクセスを表2に示した期間ごとに分け、図7に期間ごとのユニーク訪問者数、図8に投稿あたりの平均ページビュー数を示す。図7に示すユニーク訪問者数は、一例として、185-200期における1-2748の投稿へのアクセスにおいて、Cookieにより識別されたユニーク訪問者を重複なく集計した値が10,645であることを示している。185-200期の期間中のユニーク訪問者数は、図5に示したように38,843であり、このうち、75.5%にあたる29,309のユニーク訪問者が、当期中の投稿である2904-3155の投稿へアクセスしている。同様に、27.4%にあたる10,645のユニーク訪問者が、1-84期に投稿された1-2748の投稿へアクセスし、3.8%にあたる1,467のユニーク訪問者が、85-124期に投稿された2749-2886の投稿へアクセスし、0.4%にあたる149のユニーク訪問者が、当期の直前である

Month	Date	Main Topics
4	Aug. 1999	"JR East's 'ATOS' has gone away!" made headlines; in fact, communication was disrupted by burnout of a LSI chip in Fujino station. We didn't know anything about this at that time. What is 'ATOS' was unknown to most people.
32	Dec. 2001	The 'Shonan-Shinjyuku Line' started. Prior to this time, JR East's automated route controlling system 'ATOS' has been deployed in Sobu, Yokosuka, and Tokaido lines.
75	Jul. 2005	JR East replaced an old route controlling system by 'ATOS' in Saikyo Line; the old but computer-based system had been used for 30 years and this is the first full-scale replacement of such system in Tokyo area.
127	Nov. 2009	JR East's intensive route controlling system 'ATOS' has been deployed in Yokosuka branch line (Ofuna to Kurihama). Local control center in Zushi closed.
143	Mar. 2011	The East Japan Great Earthquake shocked railways. The Tohoku Shinkansen suspended around 49 days.
155	Mar. 2012	JR East's 'ATOS' has been deployed in Musashino line. The operation of freight trains has been efficient.
179	Mar. 2014	The new rail tracks 'Tohoku Jukan Sen' (extension of intercity line from north Kanto region; Tochigi, Gunma, and Ibaraki) between Ueno and Tokyo (3.6 km; 2.2 miles) started with one-year delay.

Tab.4 Main Topics on This Forum and the Surroundings

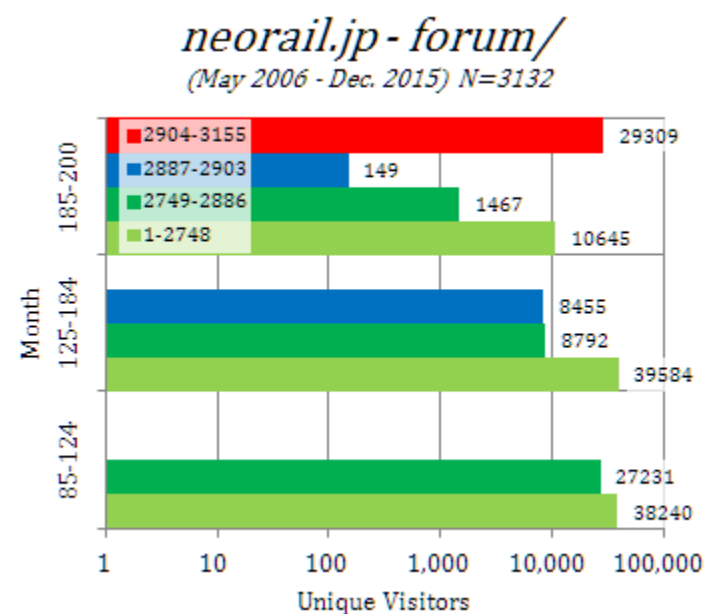


図7 期間ごとのユニーク訪問者数

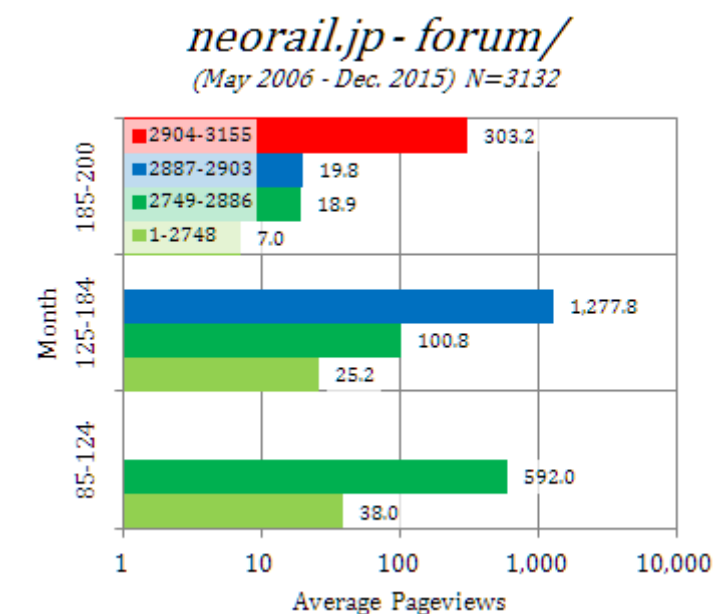


図8 投稿あたりの平均ページビュー数

125-184期に投稿された2887-2903の投稿へアクセスしている。図8では、期間ごとのページビュー数を、対象の投稿件数で除した平均値を示している。

図7から、過去の投稿に多くのユニーク訪問者がアクセスしていることがわかったが、投稿件数に対しては少ないことを図8が示している。図8は、より新しい投稿がより多くアクセスされる傾向があることを示唆している。

3.3 投稿ごとのユニーク訪問者数および累計ページビュー数

投稿を表2に示した期間ごとに分け、投稿ごとのユニーク訪問者数とページビュー数の対応を図9-12に示す。1訪問者が同一の投稿に複数回アクセスすることがほとんどないことを示している。投稿ごとのユニーク訪問者数の最小値は、1-2748の投稿では0、2749-2886の投稿では78、2887-2903の投稿では102、2904-3155の投稿では2である。

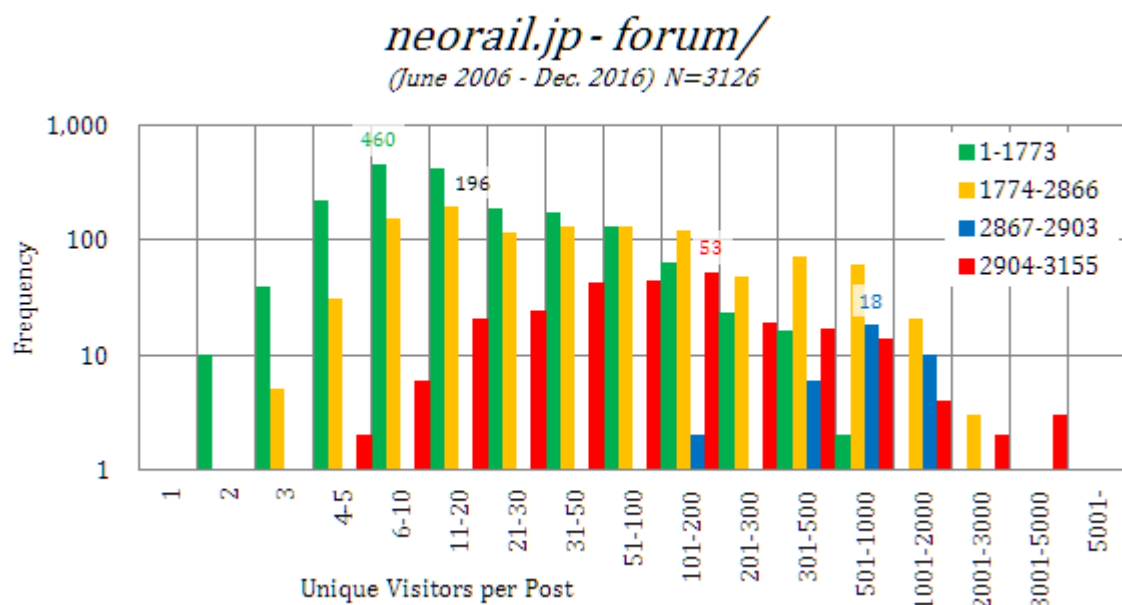


図9 投稿ごとのユニーク訪問者数の分布

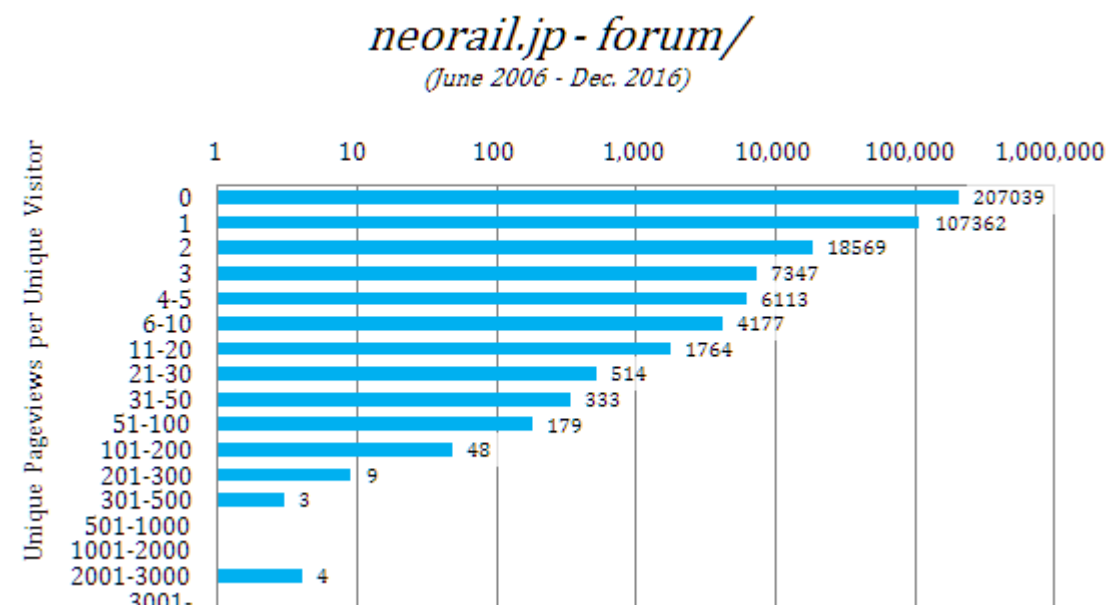


図11 ユニーク訪問者あたりのユニークページビュー数の分布

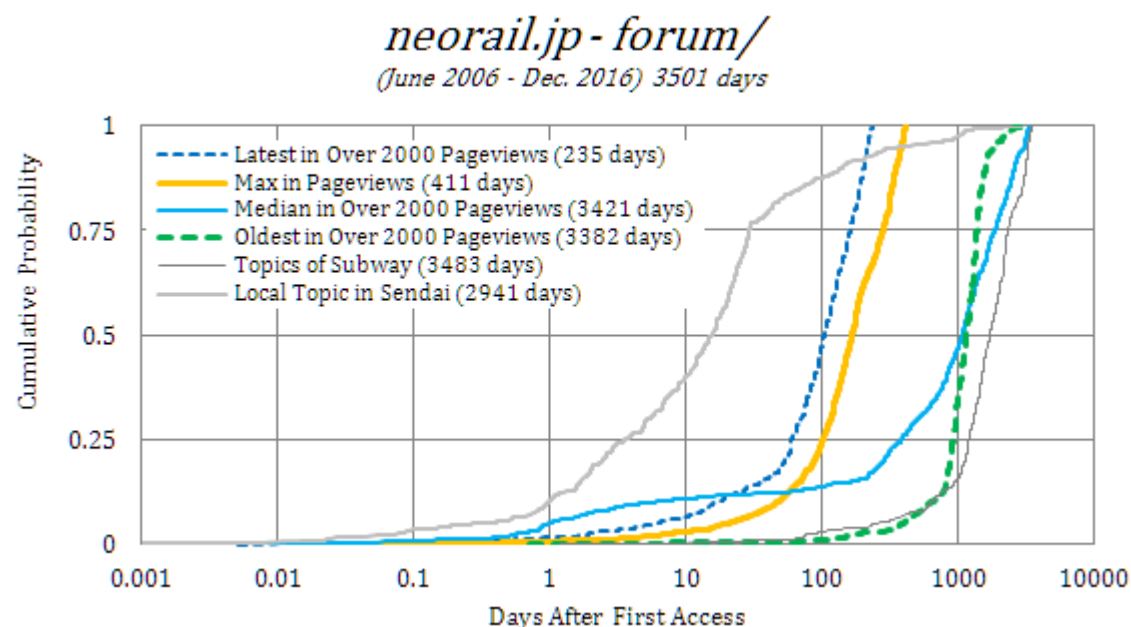


図10 初回のアクセスからの日数と累計ページビュー数

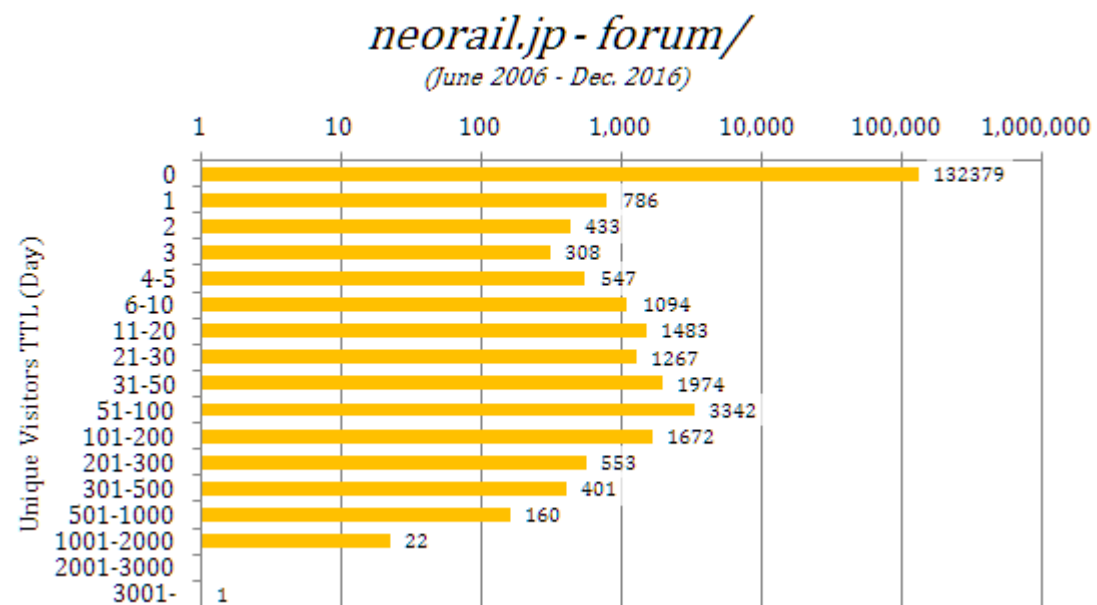


図12 ユニーク訪問者のアクセスログ上での存在期間 (TTL) の分布

投稿ごとのユニーク訪問者数の分布を図9に示す。投稿は表1に示した期間ごとに区分している。1-50期では6-10が最多で460件、51-100期では11-20が最多で196件、101-150期では501-1000が最多で18件、151-200期では101-200が最多で53件である。

投稿ごとのページビュー数について、アクセスログ上における初回のアクセスからの日数と累計ページビュー数の対応を図10に示す。ページビュー数が2,000を超え、ユニーク訪問者数が1,300を超える投稿から、投稿日時の最大値と最小値、ページビュー数の最大値と中央値となった投稿を、今後のアクセス状況に影響をおよぼすことを防ぐため投稿番号を伏せて示す。また、比較のため、ページビュー数が1,000-2,000の投稿のうち、曲線の変曲点が横軸上で最大値および最小値となった2件の投稿を、同様に投稿番号を伏せて示す。

ページビュー数が2,000を超える投稿のうち最も古い投稿(緑色の破線)では、ほぼロジスティック曲線に従った

推移を示している。比較のために示した仙台圏に関する投稿(薄いグレーの実線)でも、ページビュー数の増加が29日後で止まり、1,000日後には累積確率がほぼ1になっている。これに対し、ほかの投稿では累積確率の0.25から1まで直線になっている。ページビュー数が中央値となった投稿(水色の実線)では、約30日後までに一度、ページビュー数の増加が止まった後、再び増加に転じている。

アクセスログの期間は3,501日であるが、図10では最長で3,483日後までアクセスのあった投稿(濃いグレーの細かい実線)における累計ページビュー数の推移を示している。2006年6月1日より前の投稿については、3,501日より前のアクセスの状況が不明である。

ユニーク訪問者あたりのユニークページビュー数の分布を図11に示す。アクセスログは、フォーラムとそのほかのページで共通としている。サイトにアクセスがありフォーラムへのアクセスはなかった場合を0とする。ただし、フ

オーラムの目次やナビゲーションだけにアクセスし、投稿の本文にはアクセスしなかった場合も 0 とする。0 が最多で 207,039、301-500 まで単調に減少している。2001-3000 については、このうち 2-3 訪問者は管理者で、1-2 訪問者は利用者によるサイトの一括ダウンロードである。なお、よく知られた検索エンジンのクローラーによるアクセスは、すべて事前に除外している。

ユニーク訪問者のアクセスログ上での存在期間 (TTL) の分布を図 12 に示す。存在期間は、ユニーク訪問者ごとにアクセスログの日時の最大値から最小値を減じて得た値で、日数で示す。0 が最多で 132,379、次いで 51-100 が 3,342 で、1001-2000 まで分布している。3001-については管理者である。なお、Cookie の有効期限は 100 日としており、期間中に Cookie の有効期限は変更していない。このため、利用者のアクセス頻度が 100 日を越える場合、新しいユニーク訪問者として計上される。155-184 期には、ユニーク訪問者数が過大に計上されている可能性があることに留意する必要がある。

4. 考察

本節では、投稿件数と文字数、ページビュー数とユニーク訪問者数、投稿ごとのユニーク訪問者数および累計ページビュー数、ユニーク訪問者ごとのユニークページビュー数および TTL について考察する。

4.1 投稿件数と文字数

月ごとの投稿件数と合計文字数の推移を図 1 に示したところ、1-50 期、51-100 期、101-150 期と、約 50 期のスパンで投稿活動が変動していることが示唆された。このことから、表 1 に示す期間の分割を採用した。185-200 期には、管理者による長文の投稿を軸とする「新しい投稿活動」が立ち上がっている。185-200 期に投稿された文字数は、全期間の 52.5% を占めている。

投稿の文字数は、投稿に際する参加者の負荷を意味するとみなすことができる。図 4 は、投稿件数と平均文字数は、投稿活動や参加者の構成に大きな変動がない限り、一定の範囲に収まることを示唆している。図 2 に示した、前の月に対する増減率では、1 期におけるフォーラムの立ち上げ、53 期における前月の減少の埋める増加、185 期における投稿活動の再開などの動静をわかりやすく示したものである。図 2 は電車のモーターにおける電流値の挙動*2 と似ており、それぞれ、静止状態からの始動時の突入電流 (1 期) 空転・滑走 (52-53 期) 上りこう配での再加速 (185 期) などとなぞらえれば直感的である。

なお、185-200 期における管理者の投稿では、投稿の時点では正確な文字数はカウントしておらず、Shift_JIS におけるバイト数のみが表示されていた。また、バイト数の表示を見て文字数を調節するようなことはいっさい行なわなかった。このため、185-200 期に観測される文字数の分布は、自然な状態で観測された分布であるとみなして問題ないと考えられる。図 4 では、185-200 期の投稿において、投稿件数、平均文字数のいずれかが突出した月があったのではなく、投稿件数が 16 前後、平均文字数が 5,000 前後となる点を中心とした一定の領域に分布していることがわかる。図 3 からは、文字数が 20,001-30,000 の投稿が 8 件、10,001-20,000 の投稿が 36 件あり、185-200 期の投稿 252 件のうち、17.5% において文字数が 10,000 を超えていたことがわかる。ただし、195 期 (2015 年 7 月) には、投稿者が本文中の任意の箇所に小見出しを設け、投稿の個別表示ページ内において小見出しの目次から当該の段落へジャンプできるようにする機能を新たに設けている。管理者による投稿では、小見出しは概ね 5,000 字ごとに 1 つ設けており、実質的には 5,000 字程度で 1 つのトピックを述べている。オンラインメディアの連載記事では 4,000 から 5,000 字程度を目安としていることが知られており*3、この慣行に反しないものである。

*2 電車のモーターの電流値 : <http://www.rtri.or.jp/rd/division/rd42/rd4210/images/rd42100109f02.gif>

*3 ネットから長文が消えたいいくつかの理由 : <http://www.itmedia.co.jp/lifestyle/articles/0702/05/news015.html>

4.2 ページビュー数とユニーク訪問者数

アクセスログは、85 期の途中にあたる 2006 年 5 月 25 日から蓄積を開始していたが、本分析では 86-200 期におけるアクセスログを分析対象とした。アクセスログでは、Cookie によりユニーク訪問者を識別している。

ユニーク訪問者数とページビュー数の累積確率の推移を示した図 5 で、ページビュー数において、86-125 期の約 30 箇月で 0.5、125-185 期の約 60 箇月で 0.25、185-200 期の約 15 箇月で 0.25 となっている。このことから、アクセスログにおいては、85-200 期を、85-124 期、125-184 期、185-200 期の 3 つの期間に分けることとした。この期間と投稿番号の対応を表 2 に示した。図および表で共通の色分けにより期間の区別を示している。

表 3 に示したように、投稿活動やアクセスの状況に大きな影響を与える運営上の変更は、85-125 期の初頭および 185-200 期の初頭に集中して行なっている。このため、ほとんどの期間におけるアクセスの状況は、主に、フォーラムの参加者や訪問者が関心 (主に居住地域や年齢などによる) を持ったトピックに依存しているともみなすことができる。フォーラムの参加者や訪問者がリアルタイムに関心を持ったと考えられる主なトピックを表 4 に示した (ただし、現在の視点でまとめたものである) 85-125 期には、期間中に収束する短期的で具体的なトピックはなく、「優先席」や「バリアフリー」など、長期的でやや抽象的なトピックが投稿された。

投稿およびアクセスを表 2 に示した期間ごとに分け、図 7 に期間ごとのユニーク訪問者数、図 8 に投稿あたりの平均ページビュー数を示した。図 7 から、過去の投稿に多くのユニーク訪問者がアクセスしていることがわかるが、投稿件数に対しては相対的に少ないことを図 8 が示している。図 8 からは、より新しい投稿がより多くアクセスされる傾向が示唆される。125-184 期においては、トピックは定期的に発生していたものの、例えば「横須賀線 (大船 - 久里浜間)」については、他の線区でのトピックとして過去に投稿された内容との相違が明らかではなかった (利用者が関心を持つ時期と、技術的詳細が公表される時期のタイムラグが大きい) ことなどから、訪問者が漸減したと考えられる。図 6 に示したように、投稿件数が 0 であった 155-184 期において、ページビュー数の平均は 793.0 (最小は 181 期の 520) 月ごとのユニーク訪問者数の平均は 474.9 (最小は 182 期の 275) となっている。この期間のページビュー数およびユニーク訪問者数は、すべて当期 (アクセスされた月) より前の期 (月) の投稿が閲覧されたことによるものである。(本分析では、目次やナビゲーションページへのアクセスを除外しているため、新着の投稿の有無を確かめるためのアクセスは指標に含まれない。)

4.3 投稿ごとのユニーク訪問者数および累計ページビュー数

ある投稿が、投稿された時点から時間の経過とともにページビュー数を得ていく過程を図 10 に示した。図 10 では、ページビュー数が 2,000 を超え、ユニーク訪問者が 1,300 を超える投稿から、投稿日時の最大値と最小値、ページビュー数の最大値と中央値となった投稿を、投稿番号を伏せて示した。投稿番号を伏せるのは、今後のアクセス状況に影響をおよぼすことを避けるためである。また、比較のため、ページビュー数が 1,000-2,000 の投稿のうち、曲線の変曲点が横軸上で最大値および最小値となった 2 件の投稿を、同様に投稿番号を伏せて示した。ページビュー数が 2,000 を超えユニーク訪問者が 1,300 を超える投稿のうち最も古い投稿 (緑色の破線 : ページビュー数は 2,169、ユニーク訪問者数は 1,453) では、片対数グラフ上でほぼロジスティック曲線に従った推移を示している。比較のために示した仙台圏に関する投稿 (薄いグレーの実線 : ページビュー数は 1,036、ユニーク訪問者数は 896) でも、ページビュー数の増加が 29 日後で止まり、1,000 日後には累積確率がほぼ 1 になっており、ほぼロジスティック曲線に従った推移を示している。ページビュー数が中央値となった投稿 (水色の実線) では、約 30 日後までに一度、ページビュー数の増加が止まった後、再び増加に転じている。ほかの投稿では累積確率の 0.25 から 1 までほぼ直線になっている。

ロジスティック曲線*4 は、生物の成長や個体数の増加、製品の普及などを説明したり、成長などの限界や市場規模を推定したりするときに使われることが知られている*5。フォーラムの投稿についても、長期的には累計ページビュー数の推移がロジスティック曲線となることが示唆される。

4.4 ユニーク訪問者ごとのユニークページビュー数および TTL

ユニーク訪問者あたりのユニークページビュー数の分布を図 11 に示した。アクセスログは、フォーラムとそのほかのページで共通としている。サイトにアクセスがありフォーラムへのアクセスはなかった場合を 0 とする。ただし、フォーラムの目次やナビゲーションだけにアクセスし、投稿の本文にはアクセスしなかった場合も 0 とする。

ユニークページビュー数が 0 のユニーク訪問者を含めると、ユニーク訪問者の総数は 353,461 である。このうち、ユニークページビュー数が 0 または 1 であるユニーク訪問者は 88.9% である。ユニークページビュー数が 2 であるユニーク訪問者を含めると 94.2% に達し、大多数の利用者がすぐに立ち去っていることがわかる。図 12 に示したユニーク訪問者のアクセスログ上での存在期間 (TTL) から、ユニークページビュー数が 1-3 であるユニーク訪問者の 99.3% にあたる 132,379 のユニーク訪問者において、TTL が 0 日であることがわかる。なお、TTL においては 51-100 日に分布のピークがみられるが、これは Cookie の有効期限を 100 日としていたことで、100 日を越える期間を挟んで再びアクセスした利用者を追跡できていないことにより生じた偽のピークであると考えられる。

ユニーク訪問者ごとのユニークページビュー数は、利用者が能動的に複数の投稿を閲覧しようとしたり、定期的に新着の投稿を閲覧したりしない限り、大きな値にはなりえないと考えられる。このようなアクセスを以下、フォーラムの体系的な閲覧と称する。ただし、非常に大きな値は外れ値として除外することが必要となる。実際に、ユニークページビュー数が 2,001-3,000 であるユニーク訪問者の数は 4 であるが、このうち 2-3 訪問者は管理者で、1-2 訪問者は利用者がプログラムを用いてサイトを一括ダウンロードしていたものであることがわかっている。

体系的な閲覧とみなす閾値を決定することは困難であるが、仮に 185-200 期の投稿の 2 割以上を閲覧したとみなすことができる、ユニークページビュー数が 51-500 であるユニーク訪問者の数は 239 である。前節で推定した利用者の実数 2,544 人に対し、9.4% のユニーク訪問者がフォーラムの投稿を体系的に閲覧しているとみなすことができる。

5. おわりに

本稿では、これまで管理者により直感的に把握されていたフォーラムの投稿活動やアクセスの状況を、データにより明示的に示す方法について検討した。

本分析の対象としたフォーラムは、1999 年 5 月 25 日に開設された。2006 年 5 月 25 日からは、フォーラムの投稿 1 件ごとに対するアクセスログを記録している。また、2006 年 9 月までに、新規の訪問者の支援を目的とするフォーラムのデザインや機能が完成している。本分析では、1999 年 5 月 25 日から 2015 年 12 月 31 日の期間の投稿 3,132 件と、2006 年 6 月 1 日から 2015 年 12 月 31 日の期間のアクセスログを用いた。アクセスログの期間全体でのユニーク訪問者数は 146,422、ページビュー数 (目次やナビゲーションのページを除く) は 388,105 である。

本分析にあたり、投稿の文字数を初めて正確に集計した。投稿件数と合計文字数の前月に対する増減率を示した図 2 は管理者の直感に反しないものである。1 期 (1999 年 5 月) には、投稿件数の増減率は 28.0 倍、合計文字数の増減率は 6672.0 倍で、いずれも最大値である。185 期 (2014 年 9 月) は、投稿件数の増減率は 27.0 倍、合計文字数の増減率は 348.5 倍で、いずれも 1 期に次いで 2 位である。1 期での立ち上げ、185 期での再立ち上げに要した労力を、データにより明示的に示すことができていると考えられる。

次に、ユニーク訪問者数とページビュー数については、ページビュー数の累積確率における推移から、85-124 期、125-184 期、185-200 期の 3 つの期間に分けて検討した。85-125 期には、期間中に収束する短期的で具体的なトピックはなかった。125-184 期には、過去に投稿された内容との相違が明らかではなかったことなどから、訪問者が漸減したと考えられる。

さらに、ある投稿が、投稿された時点から時間の経過とともにページビュー数を得ていく過程を示した図 10 で、ペ

*4 ゴンペルツ曲線とロジスティック曲線 : <http://www.kogures.com/hitoshi/webtext/stat-seicho-kyokusen/>

*5 Wikipedia 「Diffusion of innovations」: https://en.wikipedia.org/wiki/Diffusion_of_innovations

Date	Important Topics in the Future
May 2005, Mar. 2016, - Mar. 2019	In the city of Kawasaki, Musashi-Kosugi and Oda-Sakae stations open for local transit; the Keihin industrial area and the area around Haneda airport will convert to R&D, commercial and residential areas.
- Mar. 2016	The council for transport will report a plan to build railway networks in Tokyo area in the period through 2030.
2018 - 2036	Following the decision of the Radio Regulatory Council, JR railway companies will probably deploy Radio Block (CBTC, ATACS, or 'Wireless ATC').
2020 -	The 'new' modal transits (includes urban ropeway or gondolas, hovercraft, dual mode vehicles, and LRT) will catch a great deal of attention. The congestion of low-altitude local airway in Tokyo will become a problem because private helicopters and 'Drone's will increase rapidly.
2027	JR Central (JR Tokai) will serve the Linear Chuo Shinkansen between Tokyo and Nagoya. In Tokyo area, Shinagawa and Hashimoto stations will open.
2045	JR Central will extend the Linear Chuo Shinkansen to Osaka.
2030 - 2045	Sagami line's track doubling is postulated by local government and communities expectantly. It is estimated that passengers from Odawara and Takasaki to Hashimoto will increase rapidly.
- 2050	The rebuilding of bayside freight rail tracks in Tokyo and Yokohama is proposed by local government and city developers with great expectations. The fast and reliable mass transits to airports are required. The experts work on research for mathematical optimization of railway networks. The reducing and smoothing of congestion of commuter trains in Tokyo area is required strongly. The experts of labor policy offer 'flexible' working and 'work-life balance'.
2050 -	Because of labor shortage in Japan, it is expected that skill transfer in domestic railway industries will be more difficult. When many railways are served by robots, general public could confuse by the full-automated services. The experts of transport have an idea for Demand-Responsive Transit, but R&Ds for ITS and AGT are still continued separately.

Tab.5 Important Topics in the Future

付録 1 今後の主なトピック

ページビュー数が 2,000 以上の投稿の一部が、ほぼロジスティック曲線に従った推移を示した。投稿ごとの累計ページビュー数が、投稿の「普及率」や「賞味期限」を推定する指標として応用できる可能性が示唆された。

4 節では、3 節で分析した各指標に加え、ユニーク訪問者について考察した。長期間かつトピックが複雑なフォーラムでは、新規の訪問者が過去の投稿を体系的に閲覧することが困難になりやすいという問題がある。185-200 期 (2014 年 9 月から 2015 年 12 月) の期間では、この期間の投稿の 2 割以上を閲覧したとみなすことができる、ユニークページビュー数が 51-500 であるユニーク訪問者の数は 239 で、一定の方法で推定された実訪問者の概数に対して、9.4% に留まることがわかった。

本分析では、投稿の内容 (テキスト) は分析対象としていない。今後、テキストを用いた分析を行ない、長期間にわたるフォーラムでのナビゲーションや情報推薦における課題を明らかにする予定である。

謝辞 本分析の機会を提供して下さった 146,422 のユニーク訪問者に謝意を示す。本分析に際し、いかなる第三者に対しても本件アクセスログを開示していないこと、いかなる金銭の授受も行っていないことをここに記してオンラインコミュニティへの敬意を示す。(2016 年 1 月 3 日セルフ受理、2016 年 1 月 7 日セルフ条件付採録、2016 年 4 月 2 日セルフ採録)