

ホームページのアクセス履歴データを用いた交通情報獲得パターンの分析*

Analysis of Transport Information Acquisition Patterns Based on Access Log Data to Home Page

森井 紀裕** 藤原 章正*** 張 峻屹**** 岡村 敏之*****

by Norihiro MORII** Akimasa FUJIWARA*** Junyi ZHANG**** Toshiyuki OKAMURA*****

1. はじめに

交通機関を利用する人々にとって、通行規制情報や路面状況情報・バス運行情報などは交通機関を安全・快適に利用するために重要なものである。近年、情報通信技術の発展と、道路情報提供の規制緩和を背景に、カーナビゲーションシステム、インターネット、携帯電話などを活用した交通情報提供が活発化している。

中でもインターネットは、総務省「平成 14 年通信利用動向調査」¹⁾の結果によると平成 14 年末におけるインターネット利用者数は全国で 6,942 万人に上り、人口普及率 54.5% に達し、初めて 50% を突破した。また、世帯普及率は 81.4% に上り、社会の情報インフラとしての位置づけがより一層明確になりつつある。

そこで、本研究では、様々な情報提供のデバイスの中でインターネットに着目し、交通情報提供ホームページのアクセス履歴データを用いることによって、人々がネット上でどういった情報をいつ、どれだけの時間をかけて獲得しているのかを分析する。それによって、ネット上での情報獲得のパターンを見つけ出し、交通情報ホームページの改善に向けた基礎資料を提供することを目的とする。なお、ここでは対象ページへの立ち寄りをもって情報獲得と呼ぶこととする。

2. 対象データの概要

本研究では、『ひろこく - 国土交通省広島国道工事事務所 - 』²⁾(以下『ひろこく』)が管理しているホームページ(図 1)のうち、交通情報を提供している以下の ~ のページへのアクセス履歴データを対象として分析を行う。

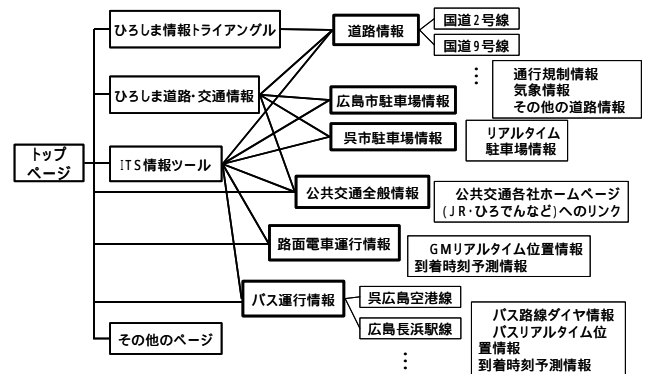


図 1. ホームページの構成図

- 規制情報・・・各路線の通行規制情報
- 気象情報・・・各路線に設けられた観測所の降雨・降雪・路面情報他
- 駐車場情報・・・リアルタイム駐車場情報
- 道路全般情報・・・ ~ 以外の道路情報
- バス路線ダイヤ情報・・・各路線の時刻表
- バス運行情報・・・バスリアルタイム位置情報・到着時刻予測情報
- 路面電車運行情報・・・グリーンムーバリアルタイム位置情報・到着時刻予測情報
- 公共交通全般情報・・・その他の公共交通情報(他社ページへのリンク他)

表 1. データ概要

期間	2002年4月1日(月)~4月30日(火)
延べアクセス件数	462,604件
対象ホームページ	広島交通情報提供ホームページ 『ひろこく - 国土交通省広島国道工事事務所 - 』
記録内容	アクセス日, アクセス時刻 アクセス者IPアドレス, アクセス先URL

* キーワーズ: 地域ITS, 交通情報, アクセス履歴データ
 ** 学生員, 修(工), 広島大学大学院国際協力研究科
 (連絡先: 〒739-8529 東広島市鏡山1-5-1 TEL&FAX:0824-24-6921)
 *** 正会員, 博(工), 広島大学大学院国際協力研究科
 **** 正会員, 博(工), 広島大学大学院国際協力研究科
 ***** 正会員, 博(工), 広島大学大学院国際協力研究科

データの概略を表 1 に示す。アクセス者の IP アドレスによって、個人を特定しアクセス履歴を分析することができる。さらに、曜日、天候などアクセス履歴データからでは判別できない外部データを統合し加工する。その加工データを用いて集計分析を行うことで利用者の情報獲得パターンを把握する。

3. アクセス件数変動パターンの分析

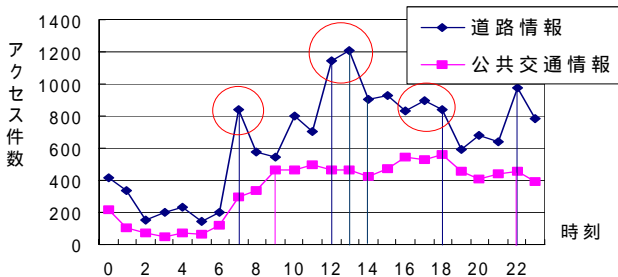


図 2. 情報ごとに見た時刻ごとのアクセス件数の変動

(1) 時間帯とアクセス件数との関係

図 2 は、～ の情報をまとめて道路情報 (15,590 件)、～ の情報をまとめて公共交通情報 (8,379 件) としてアクセス件数を図 2 に示す。図 2 より、道路情報は公共交通情報に比べて、アクセス件数の変動が大きい。特にアクセスが多い時間帯は 12～13 時の昼休み中であり、次に朝 7 時台と 14～18 時に集中している。このことは業務交通や通勤・帰宅交通の発生直前に情報アクセスが集中するためと考えられる。一方、公共交通情報は 10～22 時の間に安定したアクセス件数が見られる。これは、バスや路面電車の位置情報・到着時刻予測情報などのリアルタイム情報を提供しており、利用する直前にアクセスが多いためと考え、朝ピーク時よりはむしろオフピークから終電までの移動において情報のニーズが高いことが分かる。

(2) 土日と平日の違いとアクセス件数との関係

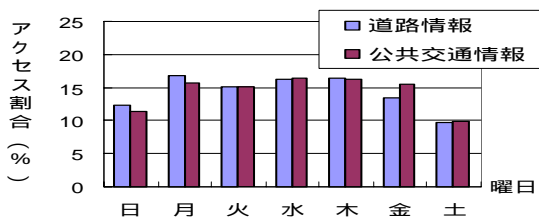


図 3. 道路情報と公共交通情報の曜日ごとのアクセス割合

図 3 に曜日ごとのアクセス割合を示す。図 3 より、平日と土日のアクセス割合は、道路情報と公共交通情報のアクセス割合に大きな違いは見られずどちらの情報も平日のアクセス割合が高い。

(3) ページ構造とアクセス件数との関係

図 4 に階層数ごとのアクセス件数を示す。ここで、階層とはトップページからそのページに辿り着くまでのクリックの数である。図 4 より、6 階層より深い階層においては明らかにアクセス件数が減少している。これより、6 階層より深いところに情報提供ページを置くことは望ましくないと言える。

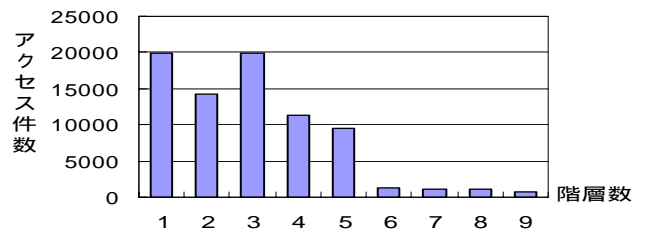


図 4. 階層数ごとのアクセス件数

4. 複数情報獲得パターンの分析

表 3. 複数情報獲得者数

獲得回数	1	2	3	4	5	6	7	8	計
獲得者数	4898	2200	1250	1030	611	132	65	21	10207
割合 (%)	47.9	21.5	12.2	10	5.9	1.2	0.6	0.2	100

表 3 に、1 回のアクセスにおいて複数の情報を獲得している利用者数を示す。ここで、同じ IP アドレスの人でアクセスから次のアクセスの間隔が 10 分以上の場合、1 回のアクセスは終了したとしている。1 回のアクセスで 2 つ以上の情報を獲得している利用者は全体の 5 割以上を占めており、これらの利用者に対する改善策が重要である。

表 4 は 1 回のアクセスにおいて 2 種類の情報を獲得している利用者の集計結果を示す。この中で、利用者数が比較的多い組合せは、() 規制情報と気象情報 () バス運行情報と路面電車情報 () バス運行情報と公共交通全般情報 () 駐車場情報と路面電車情報である。

ここで、規制情報と気象情報の組合せ獲得についてみると、この組合せ獲得が多いことより、規制情報と気象情報のページ間の移動をしやすくするか、検

表 4. 2 種類の情報獲得者数割合

	規制	気象	駐車場	道路全般	バス路線ダイヤ	バス運行	路面電車	公共交通全般
規制								
気象	32(329)							
駐車場	0.1(18)	0(4)						
道路全般	0.5(61)	0.2(21)	0(5)					
バス路線ダイヤ	0(3)	0(5)	0.1(11)	0(0)				
バス運行	0.1(18)	0.1(15)	1.1(119)	0(5)	1(112)			
路面電車	0(9)	0(10)	1.4(152)	0(4)	0.1(18)	1.5(155)		
公共交通全般	0.2(27)	0.4(47)	0.6(70)	0(2)	0.1(17)	1.4(150)	1(109)	

2種類の総情報獲得者数・・・1,496人(全体の14.7%)

表内の数値は $\frac{\text{組合せ獲得者数}}{\text{総獲得者数}} \times 100 (\%)$ 括弧内の数値は人数

索した地域の道路情報ページに規制・気象の両方の情報を提供する改善策が有効であると考えられる。また、図 5 は規制情報と気象情報組合せ獲得者の時間変動を示している。図 5 より、この組合せ情報は出勤・帰宅時間に獲得されている。

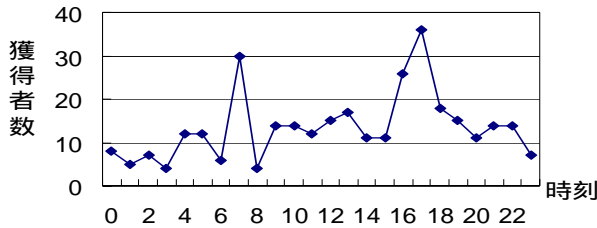


図 5. 規制情報と気象情報組合せ獲得者の時間変動

公共交通情報の組合せ獲得についてみると、上記の()、()に加えて、バス路線ダイヤ情報とバス運行情報、路面電車情報と公共交通全般情報との組合せ獲得者も比較的多く、公共交通情報は組合せて獲得されることが多い。これより、特定の公共交通情報ページ間にリンクを設けページ間の移動をしやすくする改善策が利用者の利便性向上に有効であると考えられる。

5. 共分散構造分析による情報獲得パターンの把握

利用者の獲得した情報の種類による「利用特性」が、対象ホームページにアクセスする頻度にどのように影響を及ぼすかを時間帯ごとと曜日ごとに把握する。また、時間帯ごとと曜日ごとに分かれたアクセス頻度に時間帯ごとと曜日ごとの一回のアクセスにおける平均滞在時間が影響するかどうかを検討する。これらのアクセス頻度に対する影響要因を表す評価

構造モデルを共分散構造分析により構築した。図 6 にモデルのパス図を示す。

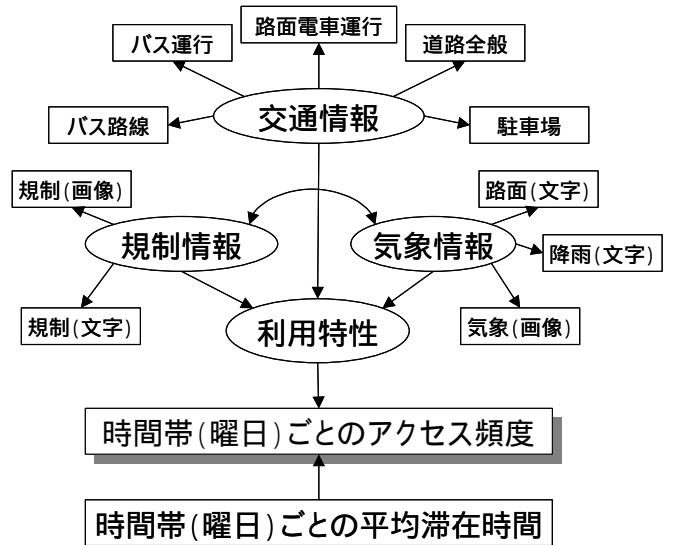


図 6. モデルのパス図

共分散構造分析において用いたデータは、アクセス者 IP アドレスにより、個人ごとのアクセス頻度と平均滞在時間に注目している。ただし、IP アドレスからでは携帯電話からのアクセスにおいて個人の特定が不可能であるためパソコンからアクセスしているデータのみを取り扱うこととする。サンプル数は、4,275 人となっている。ここでの、「規制(画像)」、「駐車場」などは、利用者が一ヶ月の間にそのページを訪れたかどうかのダミー変数である。

ここにおいて、アクセス頻度とはホームページに訪問した頻度であり、ホームページにアクセスして閲覧を終了するまでを一回のアクセスとする。また、ここでの平均滞在時間とは、ある利用者がホームページにアクセスしていた時間をアクセス回数で除したものである。

(1) 時間帯ごとにみた共分散構造分析

時間帯ごとのアクセス頻度について構成したモデルより得られた結果を表 5 に示す。GFI 値、AGFI 値はそれぞれ、0.880、0.855 となっておりこのモデルは適合度が高いといえる。

表 5 のパラメータ推定結果より、情報の「利用特性」はアクセス頻度に大きな影響を及ぼしているといえる。しかし、時間帯ごとのパラメータ推定値に大きな違いはなくパソコンからの情報獲得行動では時間帯ごとに特徴はないといえる。このことから、

表 5 . 分析結果 (時間帯ごと)

		標準化 パラメータ	t値
規制情報	利用特性	0.06	0.36
	規制情報ページ(文字情報)	0.72	14.45 **
	規制情報ページ(画像入り)	0.23	
気象情報	利用特性	0.44	2.70 **
	気象情報ページ(画像入り)	0.15	
	降雨情報ページ(文字情報)	0.90	9.30 **
	路面情報ページ(文字情報)	0.75	9.26 **
交通情報	利用特性	0.18	
	バス路線情報ページ	-0.01	-0.28
	バス運行情報(画像入り)	-0.14	-4.50 **
	路面電車情報(画像入り)	0.05	1.78
	道路全般情報	0.16	4.85 **
	駐車場情報	0.55	
利用特性	アクセス頻度(0~6時)	0.65	
	アクセス頻度(6~9時)	0.78	44.82 **
	アクセス頻度(9~12時)	0.88	49.31 **
	アクセス頻度(12~15時)	0.84	47.87 **
	アクセス頻度(15~18時)	0.85	48.30 **
	アクセス頻度(18~21時)	0.84	47.80 **
	アクセス頻度(21~0時)	0.71	42.09 **
	平均滞在時間(0~6時)	0.10	8.21 **
平均滞在時間(6~9時)	0.11	10.70 **	
平均滞在時間(9~12時)	0.15	18.45 **	
平均滞在時間(12~15時)	0.14	16.53 **	
平均滞在時間(15~18時)	0.13	15.72 **	
平均滞在時間(18~21時)	0.17	20.11 **	
平均滞在時間(21~0時)	0.19	17.39 **	

**=1%有意 *=5%有意

提供される情報の種類や内容が決まれば情報獲得に対するニーズは時間帯によらずほぼ一定であると考えられる。

また、平均滞在時間に関しては、パラメータ推定値の t 値が大きく全ての時間帯においてアクセス 1 回当たりの滞在時間が長いほど、同じ時間帯でのアクセス頻度が高いことがわかる。

(2) 曜日ごとにみた共分散構造分析

表 6 . 分析結果 (曜日ごと)

		標準化 パラメータ	t値
規制情報	利用特性	0.18	0.98
	規制情報ページ(文字情報)	0.72	14.43 **
	規制情報ページ(画像入り)	0.23	
気象情報	利用特性	0.36	1.86
	気象情報ページ(画像入り)	0.15	
	降雨情報ページ(文字情報)	0.90	9.40 **
	路面情報ページ(文字情報)	0.75	9.37 **
交通情報	利用特性	0.18	
	バス路線情報ページ	-0.02	-0.71
	バス運行情報(画像入り)	-0.15	-4.66 **
	路面電車情報(画像入り)	0.05	1.71
	道路全般情報	0.15	4.65 **
	駐車場情報	0.55	
利用特性	アクセス頻度(日曜)	0.68	45.20 **
	アクセス頻度(月曜)	0.78	
	アクセス頻度(火曜)	0.76	51.70 **
	アクセス頻度(水曜)	0.64	42.41 **
	アクセス頻度(木曜)	0.69	45.83 **
	アクセス頻度(金曜)	0.70	47.33 **
	アクセス頻度(土曜)	0.62	41.22 **
	平均滞在時間(日曜)	0.16	13.45 **
	平均滞在時間(月曜)	0.14	13.12 **
	平均滞在時間(火曜)	0.10	9.02 **
平均滞在時間(水曜)	0.12	10.69 **	
平均滞在時間(木曜)	0.19	16.61 **	
平均滞在時間(金曜)	0.19	16.04 **	
平均滞在時間(土曜)	0.11	9.44 **	

**=1%有意 *=5%有意

曜日ごとのアクセス頻度について構成したモデルより得られた結果を表 6 に示す。GFI 値, AGFI 値はそれぞれ, 0.834, 0.799 となっておりこのモデルも適合度が高いといえる。

表 6 のパラメータ推定結果より、曜日ごとのアクセス頻度についても時間帯ごとと同様に大きな特徴はないといえる。また、曜日ごとにおいても平均滞在時間はアクセス頻度に影響を与えているといえる。

時間帯ごと曜日ごとの両パラメータ推定値ともいえる事は、規制情報、気象情報ともに文字情報に強く影響を受けていることである。これより、パソコンからのアクセスであっても文字だけの簡単な情報が好んで利用されると考えられる。また、バス運行情報のパラメータ推定値が負であることよりヶ月に一度でもバス運行情報にアクセスした利用者はアクセス頻度が低くなるのが分かる。

6. 結論

交通情報提供ホームページにおける情報獲得パターンとして、情報が提供されているページまでの階層数が深いとその情報を獲得する人は少ないこと、ある特定の情報においては多種類の情報を組合せて獲得している人が多いことが明らかとなった。そこで、階層数を浅くすることや情報ページ間の移動を容易にするなどの改善策が考えられる。

また、情報提供の方法としては文字情報のような簡潔な情報提供が好まれることがわかった。

謝辞

今回の研究にあたり、広島国道工事事務所清水課長にはデータを提供していただいたり質問に答えていただいたりしました。ここに記して謝意を表します。

参考文献

- 1) 総務省：平成 14 年通信利用動向調査，2003 年
- 2) ひろこく 国土交通省広島工事事務所：
<http://www.hirokoku-mlit.go.jp/>