

本学ホームページの稼動集計から見た本学への期待
ーセキュリティ上の必要性を考慮した統計分析ー

橋本明浩¹⁾, 吉山直樹²⁾, 中川 泉³⁾, 中島紀恵子³⁾

1) 新潟県立看護大学 (情報科学), 2) 〃 (病態学), 3) 〃 (基礎看護学)

A Study on Statistical Analysis of WEB

Akihiro Hashimoto¹⁾, Naoki Yoshiyama²⁾, Izumi Nakagawa³⁾ and Kieko Nakazima³⁾

1) Niigata College of Nursing (Information Science), 2) 〃 (Clinical Pathophysiology),

3) 〃 (Fundamentals of Nursing)

キーワード: WWW, 統計分析(statistical analysis)

要旨

本学の WWW 稼動集計をもとに本学に期待されるキーワード, 参照ページ等を分析し, 本学に期待される分野等の検討をおこなった.

はじめに

インターネットで公開される様々な情報は, 入試, 大学案内だけでなく, 学術研究活動の成果も含み社会への大きな貢献の 1 つである. 本学ホームページには一日平均約 2000-3000 件のアクセスがある. これらの稼動集計, 傾向を綿密に分析することにより, 本学が社会から望まれる情報等を考察するだけでなく, インターネットからの不正な攻撃を回避するための方法も示す.

稼動集計の方法

本学のホームページは Linux (Redhat 社製) 上のプログラム Apache (Version 2.0.40) で実現されている. Apache プログラムの動作は設定記述ファイル (apache configuration file) でログの形式等の詳細を指定することができる. 出力されたログ情報を基礎にして分析を行うことにした. 分析ソフトウェアとして GPL (Gnu Public license) に従う無償のソフトウェア Webalizer (図 1) および Analog (図 2) を使用した. 解析対象期間は: 2004 年 1 月 16 日 (金) 17 時 28 分 ~ 2004 年 5 月 09 日 (日) 01 時 25 分 (113.33 日) である.

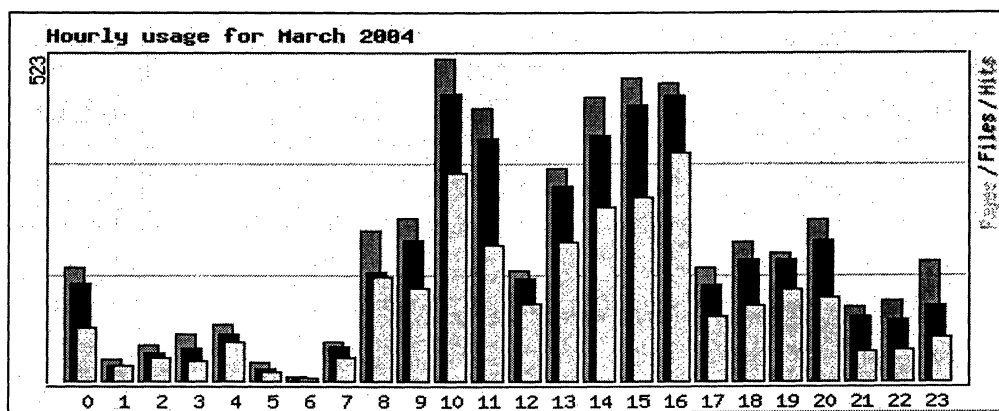


図 1 Webalizer の出力例

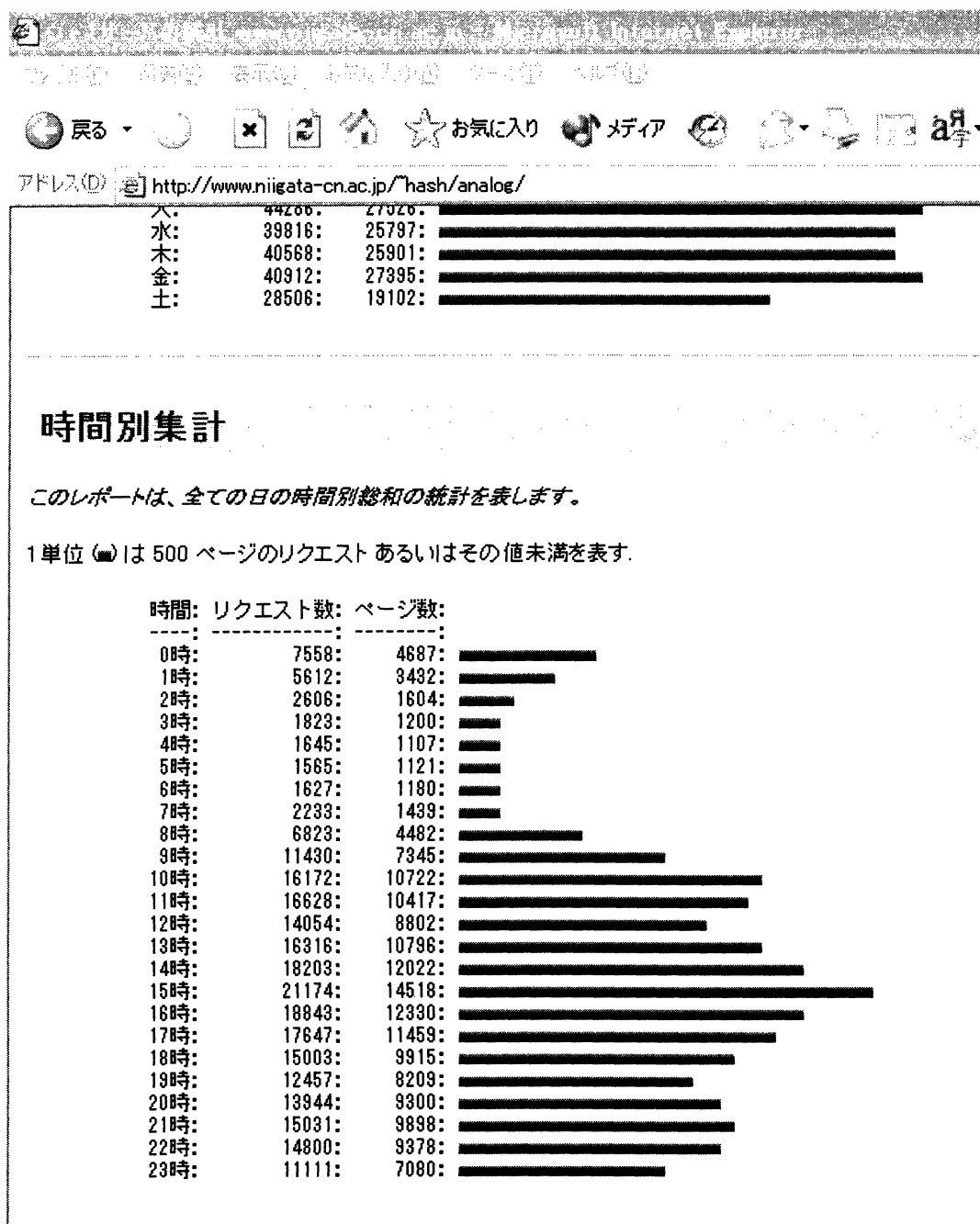


図 2 Analog 出力例

集計結果および基本統計

集計結果および基本統計を以下に示す。

1 日における時間変化

1 日におけるホームページのアクセスと利用ページ数の時間変化を表 1 時間別リクエストおよびページを図 4 に与える。図表からわかるように深夜 2 時から早朝 7 時まででも利用がある。このアクセスは、Yahoo, Google, Goo などの検索エンジンの Robot が本学ホームページをアクセスしたものである。

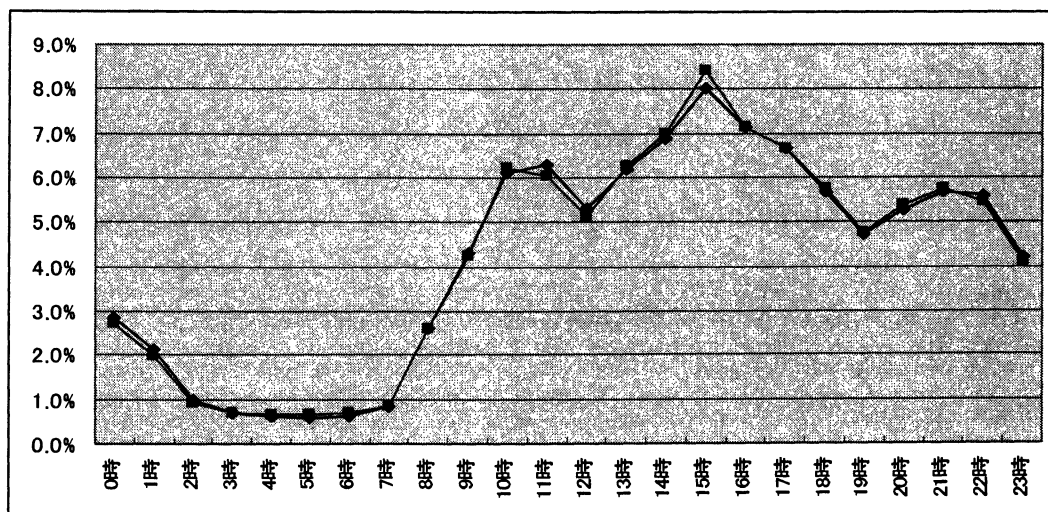


図 3 時間別リクエストおよびページのパーセント (■はページ, ◇はリクエスト)

表 1 時間別リクエストおよびページ

時間	リクエスト数	パーセント	ページ数	パーセント
0 時	7,558	2.9%	4,687	2.7%
1 時	5,612	2.1%	3,432	2.0%
2 時	2,606	1.0%	1,604	0.9%
3 時	1,823	0.7%	1,200	0.7%
4 時	1,645	0.6%	1,107	0.6%
5 時	1,565	0.6%	1,121	0.7%
6 時	1,627	0.6%	1,180	0.7%
7 時	2,233	0.8%	1,439	0.8%
8 時	6,823	2.6%	4,482	2.6%
9 時	11,430	4.3%	7,345	4.3%
10 時	16,172	6.1%	10,722	6.2%
11 時	16,628	6.3%	10,417	6.0%
12 時	14,054	5.3%	8,802	5.1%
13 時	16,316	6.2%	10,796	6.3%
14 時	18,203	6.9%	12,022	7.0%
15 時	21,174	8.0%	14,518	8.4%
16 時	18,843	7.1%	12,330	7.2%
17 時	17,647	6.7%	11,459	6.6%
18 時	15,003	5.7%	9,915	5.7%
19 時	12,457	4.7%	8,209	4.8%
20 時	13,944	5.3%	9,300	5.4%
21 時	15,031	5.7%	9,898	5.7%
22 時	14,800	5.6%	9,378	5.4%
23 時	11,111	4.2%	7,080	4.1%
合計	264,305	100.0%	172,443	100.0%

1 週間における曜日によるアクセス変化

1 週間における曜日によるアクセス変化を図 4 曜日別アクセス状況、表 2 曜日別アクセス状況に与える。月一金曜日に於いて約 15－16%の範囲に入り、休日の土日は 10%程度であることがわかる。

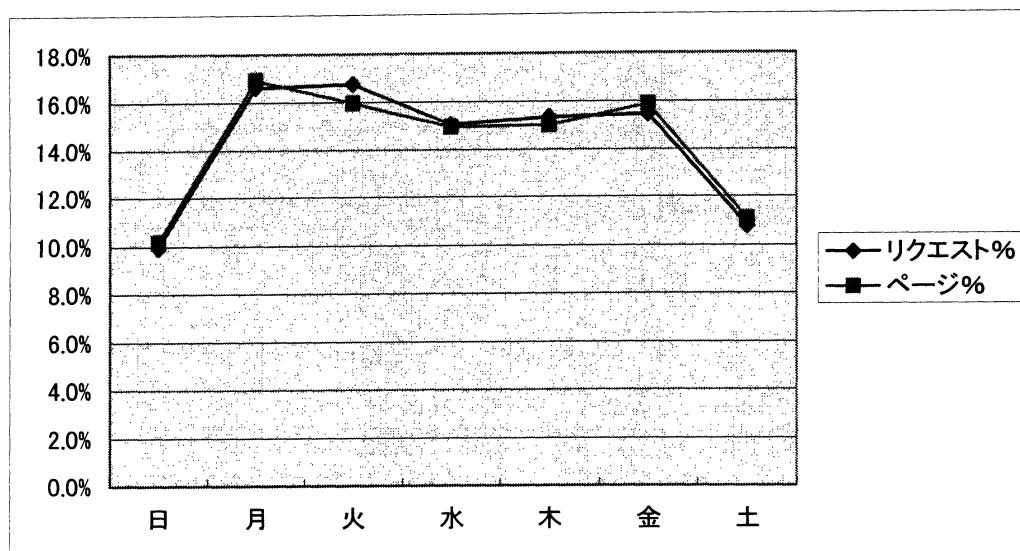


図 4 曜日別アクセス状況

表 2 曜日別アクセス状況

曜日	リクエスト数	パーセント	ページ数	パーセント
日曜日	26,299	10.0%	17,540	10.2%
月曜日	43,918	16.6%	29,182	16.9%
火曜日	44,286	16.8%	27,526	16.0%
水曜日	39,816	15.1%	25,797	15.0%
木曜日	40,568	15.3%	25,901	15.0%
金曜日	40,912	15.5%	27,395	15.9%
土曜日	28,506	10.8%	19,102	11.1%
計	264,305	100.0%	172,443	100.0%

月別変化

アクセスの月別変化を表 3 に示す。5 月にアクセスが少ないのは、前節に述べたように休日連休では 30%程度減少する効果によるものである。

表 3 月別アクセス状況

	月	日数	リクエスト数	リクエスト平均	ページ数	ページ平均
2004 年	1 月	16	44,163	2760.2	30,329	1895.6
2004 年	2 月	29	68,057	2346.8	44,879	1547.6
2004 年	3 月	31	70,120	2261.9	44,882	1447.8
2004 年	4 月	30	67,598	2253.3	43,291	1443.0
2004 年	5 月	9	14,367	1596.3	9,062	1006.9
	計	115	264,305	2298.3	172,443	1499.5

アクセスの分析

国別アクセス元分析

アクセス元の国別ドメインを表 4 アクセス元ドメイン状況に与える。本学のホームページが主に日本語で記述されているので、アクセス元の 62% は日本ドメイン (JP) となっているが、少数ではあるが驚くほど様々な国よりアクセスされていることがわかる。

表 4 アクセス元ドメイン状況

リクエスト数	リクエスト%	ドメイン
163738	61.950%	.jp (日本)
60431	22.864%	[未解決 IP アドレス]
24955	9.442%	.net (ネットワークサービス)
14400	5.448%	.com (商業団体)
118	0.045%	.edu (米国高等教育機関)
112	0.042%	.nl (オランダ)
62	0.023%	.tv (ツバル)
51	0.019%	.fr (フランス)
45	0.017%	[未知ドメイン名]
35	0.013%	.tw (台湾)
33	0.012%	.nz (ニュージーランド)
33	0.012%	.org (非営利団体)
28	0.011%	.sg (シンガポール)
27	0.010%	.de (ドイツ連邦)
27	0.010%	.bz (ベリーズ)
24	0.009%	.ca (カナダ)
22	0.008%	.arpa (旧アーパネット)
16	0.006%	.be (ベルギー)
13	0.005%	.au (オーストラリア)
13	0.005%	.it (イタリア)
12	0.005%	.pl (ポーランド)
11	0.004%	.hk (香港)
11	0.004%	.il (イスラエル)
10	0.004%	.uk (英国)
10	0.004%	.fi (フィンランド)
68	0.026%	その他
264305	1.00	合計

組織別アクセス元分析

組織別アクセス元分析を図 5 アクセス元組織状況に与える。図 5 凡例の”OC.NE.JP”は NTT 系列のネットワークプロバイダ, ”BBTEC.NET”はソフトバンクの系列のネットワークプロバイダである。

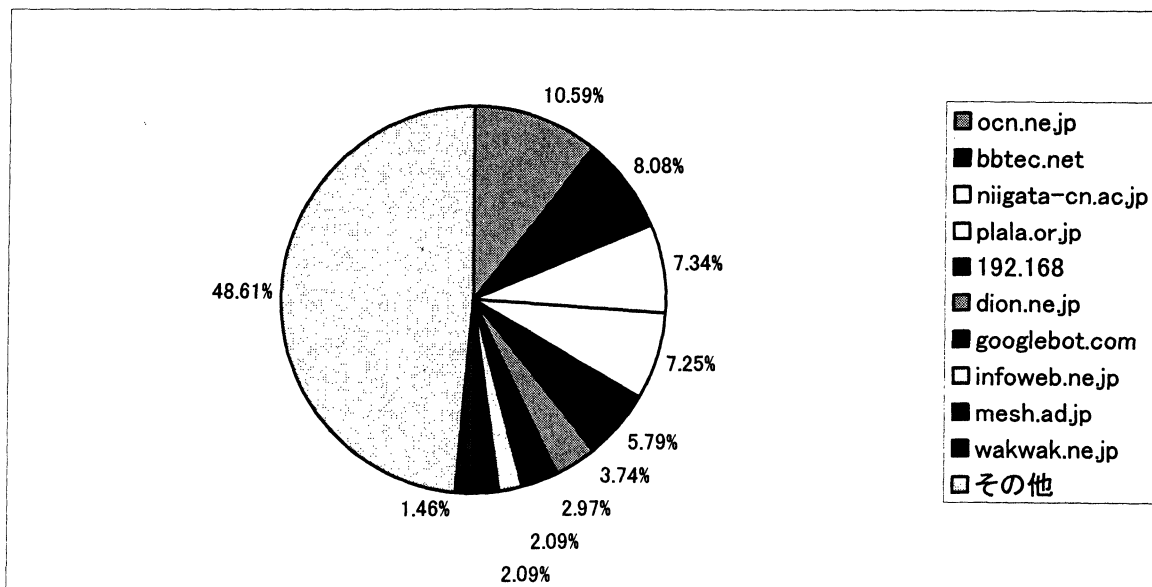


図 5 アクセス元組織状況

アクセス分野別の状況

本学のホームページのどのような項目にアクセスをしたかを図 6 に示す。大学案内, 入試および短期大学の情報へのアクセスが全体の 35%程度を占めていることがわかる。

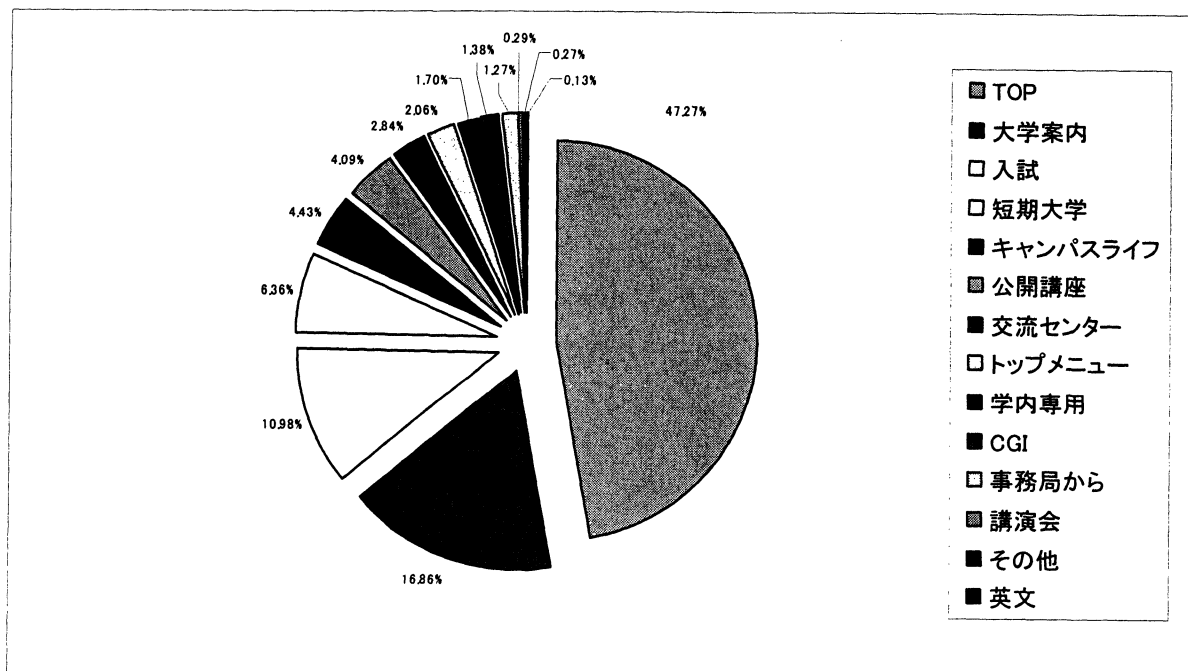


図 6 アクセス分野別状況

本学ホームページへのキーワード分析

次に本学のホームページにどのような経緯でアクセスに至ったかは大きな関心事である。インターネットユーザの多くは検索エンジンを利用する。検索エンジンからどのようなキーワードで本学ホームページに来たかを表 5 に示す。本学の名称“新潟県立看護大学”で検索し来たものが 1 位であるが、次に多いのはキーワード“看護研究”である。

キーワード	アクセス回数
新潟県立看護大学	1,717
看護研究	1,194
新潟県立看護短期大学	188
看護	179
看護大学	136
新潟県立看護大	134
基礎看護技術	96
新潟県立看護	88
cn.ac.jp	84
その他	10,871
合計	14,687

表 5 参照検索キーワード

むすび

紙面の都合で詳細まで言及できなかったが、日々の稼働分析により大学に求められているものが、研究、教育の 2 面であることについて再度認識される結果となったのは興味深い。稼働集計から本学ホームページに来た Refers 先を検討し、本学についての誹謗中傷などがインターネット掲示版に書かれた場合も直ちに発見できるなど重要な働きもあること、インターネットからのウィルス攻撃をも未然に防ぐことができる上で重要であることをも付記したい。