ネットワークグループ活動報告 2002

工学部 Web サーバの運用とホームページアクセス解析

中村勝（工学部　技術部　第二系第四班）

1　はじめに

技術部ネットワークグループは，平成 12 年 4 月より工学部 Web サーバの維持・管理業務を引き継ぎ，これまで運用を行っている．また，平成 12 年度技術報告集の「ネットワークグループの現状報告及びホームページのアクセス解析」では，アクセス解析について報告しているが，運用から約半年後ということもあり，一部の解析にとどまったので一年を通してのアクセス解析を行い，これまでの活動についても含め報告する．

2　工学部 Web サーバの運用

工学部 Web サーバの運用では大別して二つに分けることが出来る．これを以下に示す．

2.1　工学部ホームページ管理

工学部ホームページ(図 2.1)では入試案内，学科及び大学院博士前期課程の紹介，大学院博士後期課程の紹介，研究活動など様々なコンテンツがあり，更新・追加依頼があれば，作業をネットワークグループメンバーで行っている．今年度は工学部ホームページ見直しということもあり，学科及び大学院博士前期課程の紹介も更新した．

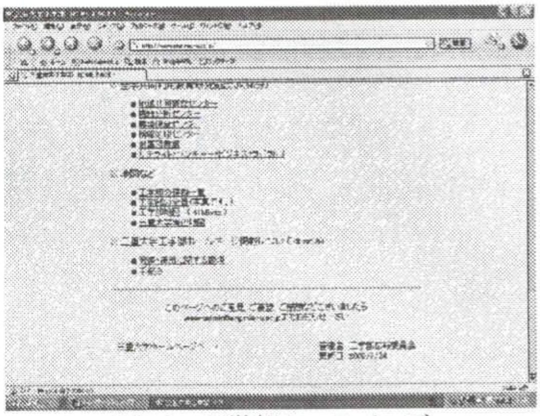

図 2.1　工学部ホームページ

しかしながら，工学部ホームページに掲載する手続きが明文化されていなかったため，更新時の手続きの流れやファイル・データの授受がまちまちであったので，申請の方法など工学部ホームページに掲載する手続きを明文化し，工学部ホームページに掲載することとなり，「手続き」項目が定められた．

2.2　サーバの維持・管理業務

サーバを停止することなく稼働させるため，以下のような作業を毎日あるいは随時行っている．

- システムログ等の日常監視
  サーバの状態，不正アクセス等の監視行う
- データ・ログのバックアップ
  ホームページの内容や設定ファイルの保存を行う
- セキュリティホールに対するパッチ当て
  ソフトウェアに潜むバグ・セキュリティホールに対しパッチ及びアップグレードを行う

3　アクセスログの解析

3.1　アクセスログ解析の目的

アクセスログ解析の目的とは閲覧者の知りたい情報を解析し，ホームページで提供する情報の充実を図ることである．

3.2　アクセスログとは

ホームページにアクセスするとサーバには図 3.1 に示すログが記録される．ログには

- リモート・ホスト名
- アクセス日時
- HTTP リクエストヘッダー
- 転送バイト数
- ファイル名

などが記録される．

```
127.0.0.1 - - [09/Sep/2002:17:34:34 +0900] "GET / HTTP/1.1" 200 3325 "-" "Mozilla/5.0 (X11; U; Linux i686; ja-JP; rv:0.9.8+) Gecko/20020328 - Vine Linux(0vl8)"
127.0.0.1 - - [09/Sep/2002:17:34:34 +0900] "GET /PoweredByVine.png HTTP/1.1" 200 2833 "http://localhost/" "Mozilla/5.0 (X11; U; Linux i686; ja-JP; rv:0.9.8+) Gecko/20020328 - Vine Linux(0vl8)"
127.0.0.1 - - [09/Sep/2002:17:34:34 +0900] "GET /icons/apache_pb.gif HTTP/1.1" 200 2326 "http://localhost/" "Mozilla/5.0 (X11; U; Linux i686; ja-JP; rv:0.9.8+) Gecko/20020328 - Vine Linux(0vl8)"
```

図 3.1　アクセスログの例

3.3　アクセスログの解析の手順

アクセスログではリモートホスト名が IP アドレスで記述されている. このログを使用すれば, どこの国からのアクセスかを知ることは困難である. そこで IP アドレスをドメイン名に変換する必要がある. 「DNSTran」はアクセスログの IP アドレスをドメイン名に変換するソフトウェアである. これを用い, ドメイン名に変換した.

また, アクセスログの解析には analog を用いた. analog では

- 時間別レポート
- 時間別概要
- 成功したリクエスト件数
- 上記の日別平均
- ホームページへの成功したリクエスト件数

などが図3.2の画面のようにグラフィカルに表示される.

アクセスログは 2001/9/1～2002/8/31 のログ(2002/7/1～7/9 が欠落)を使用して解析を行った.

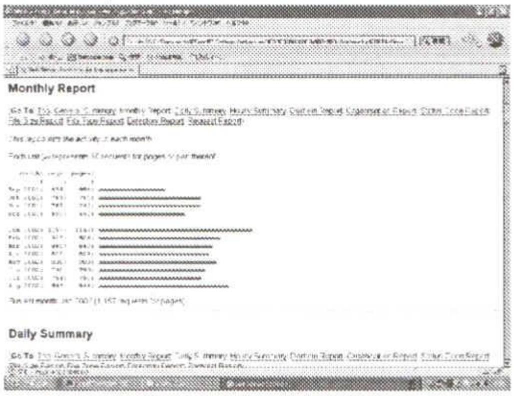

図 3.2　analog の画面

3.4　アクセスログ解析の結果

工学部の一年間のアクセス数は 214989 件で一日あたり 589 件のアクセスがあった. うち国内のアクセス数が 127619 件で全体の約 6 割, 学内からのアクセスは20984件で全体の1割ほどであった. (表 3.1), (図 3.3)

表 3. 1　一年間のアクセスデータ

	アクセス件数(回) 一日平均アクセス数(回)	総アクセス数に占める割合
総アクセス数	214989	−
日本国内からのアクセス数	127619	59.4%
三重大学内からのアクセス数	20984	9.8%

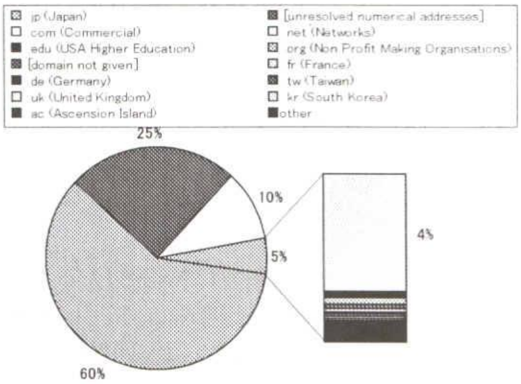

図 3.3　トップレベルドメイン別のアクセス割合

時刻毎アクセス数では日本のアクセスが多いため, 日本のアクセス事情による特徴がでていると思われる. (図 3.4)

曜日別アクセス数では, 土曜, 日曜のアクセスがほかの曜日と比較すると少ない(6 割程度)ことがわかる. 月曜から金曜まではほぼ同じアクセス数である. (図 3.5)

以下に示すものは前回(平成 12 年度)行われたアクセスログ解析の結果と同様である.

- 総アクセス数に占める日本国内からのアクセス割合(約 6 割)
- 総アクセス数に占める学内のアクセス割合(約 1 割)
- 土曜・日曜と他の曜日とのアクセス数の割合(約 6 割)

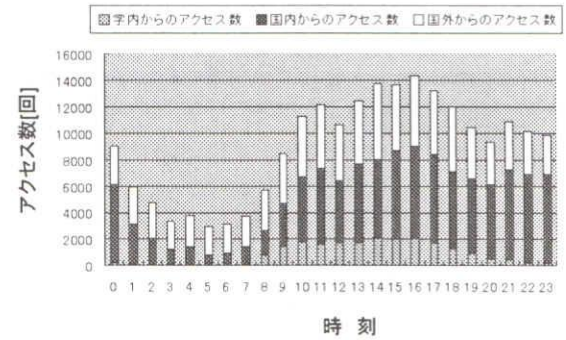

図 3.4　時刻毎アクセス数

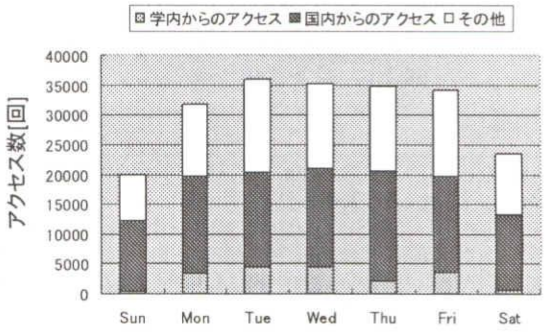

図 3.5　曜日別アクセス数

次にアクセス先の割合では，トップページ(index1.html)へのアクセス数が全体の 2 割になることがわかる．以下，各学科のページ(kikai.html,denki.html.buturi.html,bunshi.html) へのアクセスが続き，学部入試案内のページ(gakubu_j.html)の順になっている．(図 3.6)

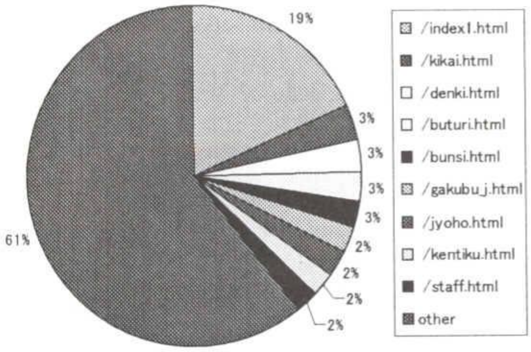

図 3.6　アクセス先の割合

では，今年度掲載したページのアクセス数の推移を

- 三重大学工学研究報告第 26 巻(2001)へのアクセス数
- リカレント教育講座へのアクセス数
- 公開講座へのアクセス数

についてみることとする．

三重大学工学部研究報告では掲載日が 2002 年8月7日ということもありアクセス数がそれほど多くはないが，日数が経過する毎に学外からのアクセス数が増加していることがわかる．

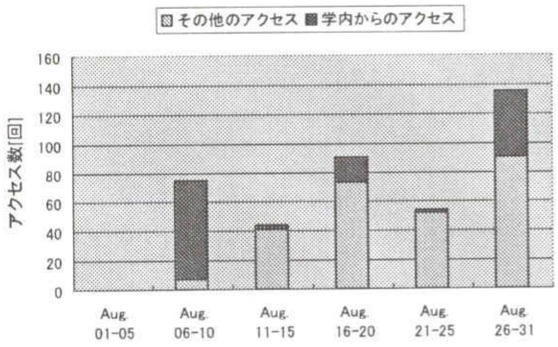

図 3.7　三重大学工学部研究報告
第 26 巻(2001)へのアクセス数

リカレント教育講座へのアクセス数ではピークが 7 月 26～31 日である．これは「物理工学科社会人向けリカレント教育講座募集」の締切りが 8 月 2 日であるからと思われる．(図 3.8)

公開講座へのアクセス数では締切り 7 月 29 日であったため掲載日(2002 年 7 月 15 日)から増加し，締切り日以降減少していることがわかる．(図 3.9)

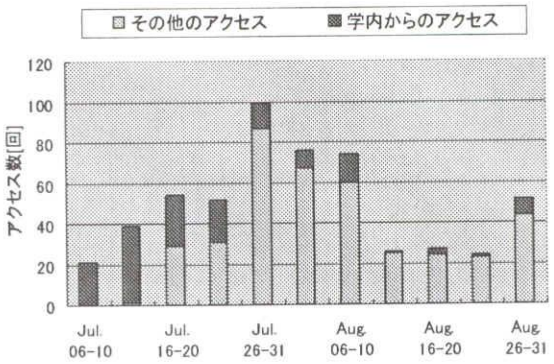

図 3.8　リカレント教育講座へのアクセス数

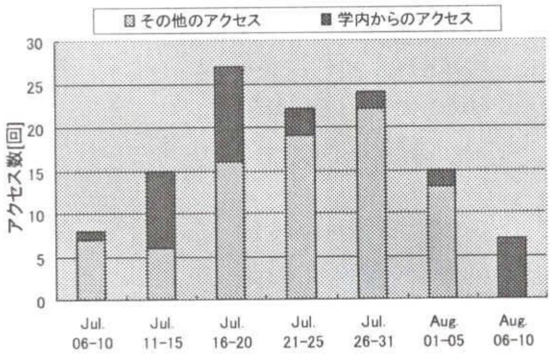

図 3.9　公開講座へのアクセス数

各月のアクセス数を 1 日～5 日,6 日～10 日,11 日～15 日,16 日～20 日,21 日～25 日,26 日以降の 6 つに分けてアクセス数の合計を図 3.10 に，その割合を図 3.11 に示す．1 月 21～25 日と 1 月 26 日以降にアクセスが多いことがわかる．

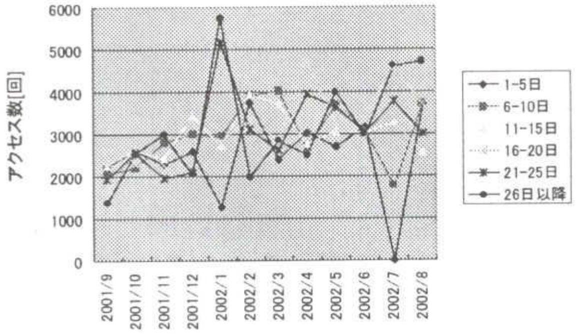

図 3.10　5 日間毎のアクセス数

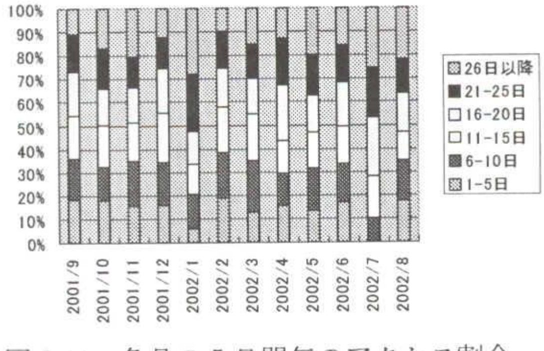

図 3.11　各月の 5 日間毎のアクセス割合

これはセンター試験後から，願書提出までの期間に一致すると思われる．この期間のアクセス先について調べることとする．

この時期にアクセスが多いと思われるページは，入試関係だと思われる．1 月 15 日～2 月 10 日間のアクセス先の割合を図 3.12 に，1 月 15 日～2 月 10 日間の全アクセス数，入試関係のアクセス数と学科のアクセス数の推移を図 3.13 に示す．図 3.6 と比較すると入試関係(gakubu_j.html)のアクセスが増加しているが，学科のアクセス数も増加していることがわかる．図 3.13 より大学入試センター試験後(1/19)以降に入試関係，学科のページのアクセス数，共に増加していることがわかる．これは閲覧者が入試関係のページのみでなく各学科のページまで閲覧しているということになる．

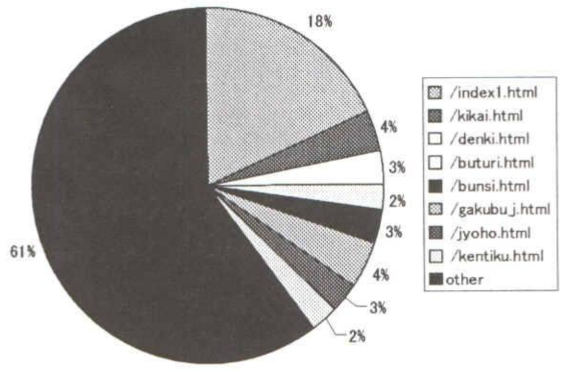

図 3.12 1/15～2/10 のアクセス先の割合

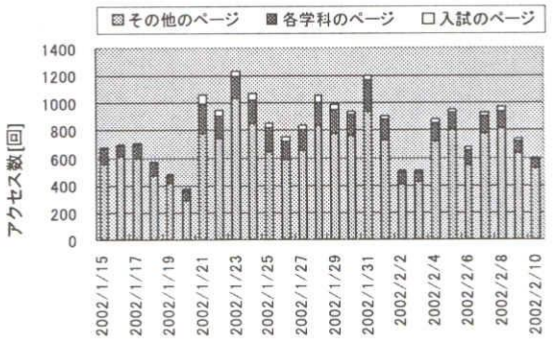

図 3.13 1/15～2/10 のアクセス数の推移

4 おわりに

今回は一年間分のログにて解析を行った．トップレベルドメイン別アクセスログの解析については，前回行ったログ解析の結果と同様になっている．また大学入試センター試験後にはアクセス数が通常の約 1.5 倍になることがわかり，入試関係のページ，各学科のページのアクセス数が増加することがわかった．

今後は複数年分のログにて解析を行い，特徴を見つけることが必要である．

また，サーバの OS のバージョンが古くなっているため，早急にリプレイスを行う必要がある．

[参考文献]

[1]日本 analog ユーザの会ホームページ:
http://www.tokyo-kasei.ac.jp/~matsuki/analog/
[2]アクセスログの有効活用法:
http://www.rfs.jp/server/apache/05.html
[3] 日本 Analog ユーザ会:
http://www.tokyo-kasei.ac.jp/~matsuki/analog/
[4]山本好弘　平山かほる:工学部技術部ネットワークグループ活動報告　ネットワークグループの現状報告とホームページのアクセス解析,技術官等による技術報告集第 9 号,pp22-28