Memoirs of the Faculty of Science
Kochi University (Information Science)
Vol.29 (2008), No.4

Web サーバログ解析によるセキュリティリスクの評価

小山 貴和子† 菊地 時夫††

要旨

近年，コンピュータとネットワークが発達してきており，ネットワーク社会の健全な発達を図っていく必要がある．セキュリティ対策を実施していても，不正アクセスによる被害を受ける場合がある．セキュリティ対策の一例として，定期的なログ監査の実施やアクセスログ解析の実施が挙げられる．本研究では，情報科学教室が運用している Web サーバのアクセスログを解析した．情報科学教室の Web サーバへの不正アクセスの攻撃パターン，アクセス元(国)，被害の有無を検証していった．また，日常的なログ監査方法についても提案した．本研究を通じて，情報科学教室のセキュリティ対策が十分機能していることが明らかとなった．

1 はじめに

1.1 背景

近年，コンピュータとネットワークが発達してきており，ネットワーク社会の健全な発展を図っていく必要がある．ソフトウェア開発支援や情報処理技術者試験，情報セキュリティ対策などを行っている経済産業省の独立行政法人である情報処理推進機構[1]は，2007 年度のコンピュータ不正アクセス届出の内容に，侵入・メール不正中継・DoS 攻撃・アドレス詐欺・ワーム感染・アクセス形跡(未遂)・ワーム形跡などがあると報告した[2]．

セキュリティ対策を実施していても，不正アクセスによる被害を受ける場合がある．不正アクセスによる被害は減少することがあっても，無くなることはないというのが現状である．

1.2 着眼点

不正アクセスをされても，被害を受けているかどうか気付かない場合がある．さらには，知らない間に被害が拡大し，加害者となってしまう場合もある．このことから，不正アクセスには早期発見，早期対処することが必要であると考えられる．そこで，定期的にログ監査を実施することで，早期発見が可能となる．また，システムの安全性を確認することもできる．以上のことか

† 高知大学理学部数理情報科学科
Department of Mathematics and Information Science, Faculty of Science, Kochi University
††高知大学理学部
Faculty of Science, Kochi University

ら，セキュリティ対策として，定期的にログ監査を実施し，アクセスログ解析を実施する必要があると考えられる．

1.3　目的

本研究の目的は，情報科学教室が運用しているWebサーバのアクセスログを解析し，不正アクセスによる被害の有無や，不正アクセスの攻撃パターン，アクセス元(国)などを検証し，不正アクセスの実態を明らかにすることである．また，本研究を通じて，情報科学教室のセキュリティ対策の安全性が明らかになると考えられる．

1.4　研究の方法

一般的にアクセスログ解析には，Analog[3]やwwwstat[4]やThe Webalizer[5]などの解析ツールが使用されている．しかし，本研究では解析ツールは使用せず，Python[6]で実装し解析を行った．Pythonの正規表現を使用することで，アクセスログの分類が容易に行える．

本研究では，2007年1月～12月のログデータを基にアクセスログ解析を行った．まず，どのようなアクセスが来ているか各月ごとに統計を取り，どのアクセスが不正アクセスであり，どの程度危険であるか検証した．さらに，アクセス元について各月ごとに統計を取り，どこの国からアクセスがあるか検証した．また，不正アクセスの攻撃パターン及びアクセス元(国)について2007年度年間の統計を取り，傾向を検証した．

2　不正アクセス

2.1　不正アクセスの定義

不正アクセスとは，あるコンピュータへの正規のアクセス権を持たない人が，他人のパスワードを用いたり，ソフトウェアの脆弱性などを悪用したりしてアクセス権を取得し，不正にコンピュータを利用する，あるいは試みることである．国内では1999年に不正アクセス行為の禁止等に関する法律[7](通称：不正アクセス禁止法)が成立し(施行は2000年より)，不正アクセス行為は犯罪行為として処罰される．

2.3　不正アクセスの対策

基本的なセキュリティ対策を実施していれば，不正アクセスによる被害はほとんど免れる．以下にその対策例を挙げる[2]．

- システム管理者
  - ✓IDやパスワードの厳重な管理及び設定
  - ✓セキュリティホールの解消(パッチ適用不可の場合は，運用による回避策も含む)
  - ✓ルータやファイアウォールなどの設定やアクセス制御設定
  - ✓定期的なログ監査
- 個人ユーザ
  - ✓OSやアプリケーションソフトのアップデート
  - ✓パスワードの設定と管理(複雑化，定期的に変更，安易に他人に教えないなど)
  - ✓無線LANやPC共有についてのセキュリティ設定確認
  - ✓ルータやパーソナルファイアウォールの活用

3　アクセスログ

3.1　Web サーバの構成

高知大学理学部情報科学教室の Web サーバの構成を表 1 に示す．サーバソフトとしては，フリー・オープンソース・ソフトウェアである Apache httpd [8]を用いている．Apache httpd は Netcraft[9]の調査でインターネット上の Web サーバの過半数を占めている．

表 1. Web サーバの構成

ハードウェア	富士通 PRIMEPOWER 250
オペレーティングシステム	Solaris 10
サーバソフトウェア	Apache httpd 2.0

情報科学教室の Web サーバでは，さらに VirtualHost を設定することで，複数の名前を持つ仮想 Web サーバとして動作させている．また，今回問題にするようなホスト名を指定しないアクセスに対しては，基本的に何もコンテンツを提供しないように，VirtualHost 設定の最初の項目がデフォルトであることを利用して，ここにダミーの設定を書いている．

```
NameVirtualHost *:80
<VirtualHost *:80>
   ServerAdmin webmaster@dummy-host.example.com
   DocumentRoot /www/docs/dummy-host.example.com
   ServerName dummy-host.example.com
   ErrorLog logs/dummy-error_log
   CustomLog logs/dummy-access_log combined
</VirtualHost>
```

図 1. VirtualHost の設定

3.2　アクセスログ

アクセスログとは，Web サーバの動作を記録したものである．図 2 に示すようにクライアントと Web サーバとのやり取りを記録している．ここでは，図 1. の設定に示されるように httpd.conf にデフォルトで定義されている combined の書式に従って，アクセス元の IP アドレス，日付と時刻，HTTP コマンド，サービス状態コード，送信バイト数，リンク元のページの URL，ユーザエージェント(訪問者の Web ブラウザ名や OS 名)が記録されている．1 回の動作につき，これらの項目を列挙した 1 行のログデータが生成される．

一般に，アクセスの多いサーバでは大量のアクセスログが生成されるため，定期的に整理する必要があるが，本サーバでは毎月 1 日に cron によって起動されたコマンドでログファイル名を書き換え，新しいログファイルに記録されるようにしている．

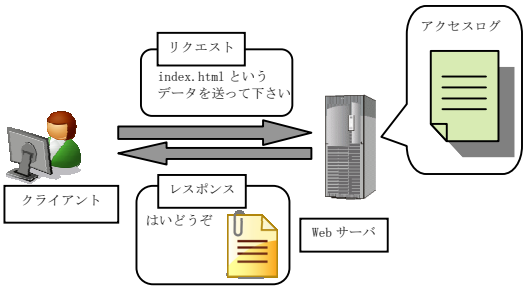

図 2.　HTTP のやり取り

3.3　HTTP メソッド

HTTP メソッドとは，リクエスト URI，つまりどのファイルを HTTP サーバから取得するかを示すコマンドである．クライアントからの HTTP リクエスト内に，HTTP メソッドが入る．HTTP メソッドは RFC1945 § 8[10]及び，RFC2616 § 9[11]で規定されている．以下に HTTP メソッドの種類を示す．

- GET

リクエスト URI で識別される情報を取得するためのメソッドで，最もよく使用され

る．1 行で書ける範囲で何でも書けるので，Web サーバの脆弱性を悪用する場合にも頻繁に使用される．

●HEAD

サーバがレスポンスに対してオブジェクトボディーを返してはならないことを除けば，GET メソッドと同一の意味のメソッドである．ファイルの最終更新日や属性を調べる場合に使用するが，GET によるコマンドが実行可能であるかどうかのチェックにも利用できる．

● POST

クライアントからの情報をサーバへ転送するために使用するメソッドである．GET メソッドでも同様のことができるが，POST メソッドではより多くのデータを転送することができる．

●PUT

FTP の PUT メソッドに似ていて，ローカルにあるファイルをサーバに転送するメソッドである．Apache httpd では，WebDAV で利用できるが，本研究に使用した Web サーバには実装していない．

● DELETE

PUT メソッドの逆で，URI で指定したリソースをサーバ上から削除するメソッドである．PUT 同様，本サーバには実装されていない．

● OPTIONS

リクエスト URI 先にあるリソースへの通信オプション(どのメソッドをサポートしているか)を通知したり調べたりする場合に使用するメソッドである．

● TRACE

特定のサーバに接続し，そこからループバックを起動するために使用するメソッドである．プロキシサーバの動作を確認するときなどに用いられる．

● CONNECT

プロキシや SSL 等のプロトコルで暗号化されたものにトンネリング接続を要求するメソッドである．プロキシサーバが有効になっているとき，このメソッドが悪用されることがある．

3.4 HTTP ステータスコード

コード	コメント
1xx(100 番台)	Informational
2xx(200 番台)	Success
3xx(300 番台)	Redirection
4xx(400 番台)	Client Error
5xx(500 番台)	Server Error

表 2.　HTTP ステータスコードの分類

HTTP ステータスコードは，Web サーバからのレスポンスの意味を表現する 3 桁の数字からなるコードである．大きく 5 種類に分類でき，先頭の 1 桁を見れば，リクエストの成功やエラーなどの状態が大体把握できる[表 2.]．HTTP ステータスコードは RFC1945 §9[10]及び，RFC2616 §10[11]で規定されている．以下に代表的なコードと本研究で頻繁に出てきたコードを示す．

- 200(OK)

  OK．リクエストは成功し，レスポンスとともに要求に応じた情報が返される．ブラウザでページが正しく表示された場合は，ほとんどがこのステータスコードを返している．

- 400(Bad Request)

  不正リクエスト．クライアントのリクエストの書式がおかしい(定義されていないメソッドを使用するなど)場合に返される．

- 403(Forbidden)

  禁止されている．リソースにアクセスすることを拒否された．アクセス権限がない場合やクライアントホストがアクセス拒否設定されている場合などに返される．

- 404(Not Found)

  見つからない．リソースが見つからなかった．

- 414(Request-URI Too Long)

  リクエスト URI が大きすぎる．URI が長過ぎるのでサーバが処理を拒否した場合に返す．

- 501(Not Implemented)

  未実装．サーバが実装されていないメソッドを使用した．

4　情報科学教室における 2007 年の不正アクセスの動向

4.1　メール不正中継

CONNECT メソッドを用いて，リモートホストの 25 番ポートに接続しようとするアクセスが多くあり，年間のアクセス数は合計 544 件あった．

これは，メール不正中継を試みようとしたアクセスであり，Apache HTTP Server では Proxy サーバを有効にして外部からのアクセスに制限をかけないと悪用されてしまう．本サーバには効果のない攻撃であるが，メール不正中継を許可してしまうと，攻撃者に踏み台サーバとして悪用され，攻撃者から送られてきた迷惑メール(SPAM)を第 3 者に送信し，加害者となってしまう恐れがある．また，大量の迷惑メールの送信で Web サーバが過負荷となり，通常のメールが送れなくなるといったサービス低下に陥る可能性もある．

これらのメール不正中継アクセスについて，クライアントホストの IP アドレスから whois コマンドを利用して，アクセス元の国名を調査した結果，中国(96%)とアメリカ(4%)からであった[図 3]．

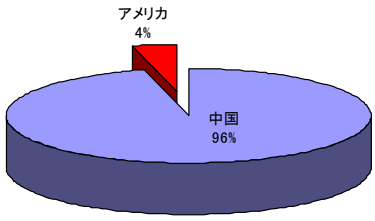

図 3.　メール不正中継のアクセス元

この不正アクセスは 2 月～3 月に集中しており，4 月以降はほとんど来ておらず，8 月以降には各月 0 件であった[図 4.]．Web サーバの脆弱性対策によりメール不正中継可能なサーバが新たに見つかることは無くなっているためかもしれない．

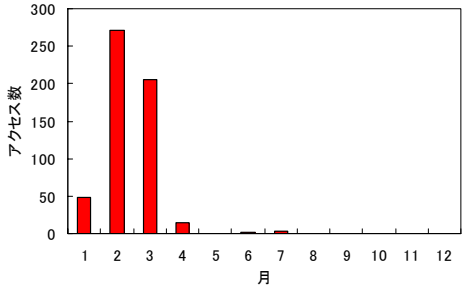

図 4.　メール不正中継の月別アクセス数(2007)

4.2　オープンプロキシチェック

オープンプロキシかどうかを確認しようとしたアクセスが毎月あった．オープンプロキシであると，メール不正中継の踏み台として悪用され易い．本研究では，年間で特に多かった事例の“GET http://www.anonymitytest.com/cgi-bin/azenv.pl”という攻撃について傾向を検証した．

この不正アクセスは，ほとんどが中国からであった[図 5]．また，毎月来ており，最近はアクセス数が減少傾向にある[図 6]．

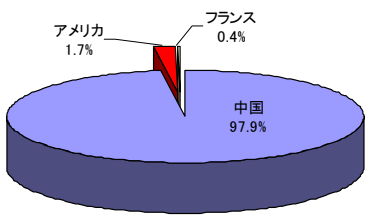

図 5.　オープンプロキシチェックのアクセス元

なお，このアクセスに対しては応答コード 403 が返されており，本サーバはこの攻撃に対して脆弱ではない．

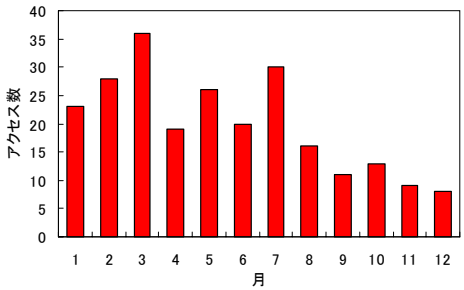

図 6. オープンプロキシチェックの月別アクセス数(2007)

4.3 CGI-BIN プログラムを狙った攻撃

CGI-BIN プログラムを狙った攻撃が毎月あった．本研究では，年間で特に多かった事例の“POST /cgi-bin/guestbook.cgi”という攻撃について傾向を検証した．このアクセスは 1995 年の「全国ニューメディア祭 in 高知」で使用されたゲストブック CGI に対するものではあるが，本サーバでは CGI の利用が許可されていないので，応答コード 403 が返される．

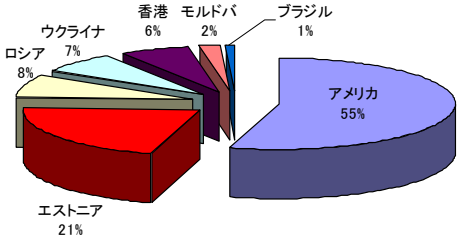

図 7.　CGI に対する攻撃（国別）

この不正アクセスは，約半数がアメリカからであった[図 7]．また，4 月以前は来ていたが，5 月以降は全く来ていない[図 8]．最近の傾向として，CGI-BIN プログラムに対する攻撃は下火になっていると思われる．

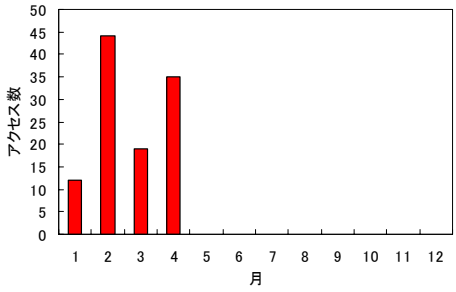

図 8.　CGI に対する攻撃(月別)

4.4　phpMyAdmin の脆弱性を狙った攻撃

MySQL 用の管理ツール phpMyAdmin[12]の脆弱性を狙った攻撃が最も多くあり，年間のアクセス数は合計 702 件あった．アクセス元を調査した結果，ドイツ(53%)，アメリカ(21%)，アイルランド(8%)，カナダ(8%)，香港(8%)，イタリア(2%)からであった[図 9]．

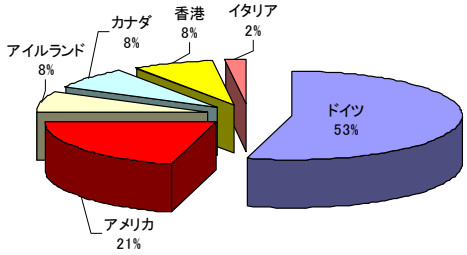

図 9.　phpMyAdmin に対する攻撃元

phpMyAdmin の脆弱性を狙った攻撃は，phpMyAdmin の中のプログラム main.php などにアクセスしようとしたものであった．アクセスログでは，“GET /phpmyadmin/main.php”などである．アクセス方法に特徴があり，phpMyAdmin の様々なプログラムに対して一度に連続的なアクセスを試みていた．これは，手当たり次第に脆弱性のあるプログラムにアクセスし，リモート実行できるかどうか調査していたと考えられる．傾向として，毎月来ている攻撃ではないが，一度に来るアクセス数が多い[図 10]．なお，本サーバにおいては phpMyAdmin はインストールされていないので，応答コード 404 が返される．

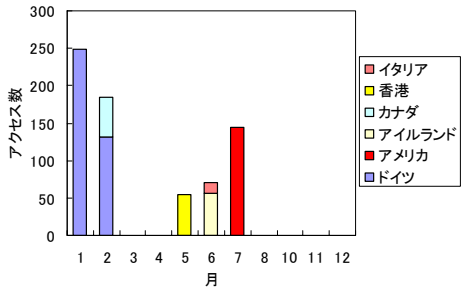

図 10　phpMyAdmin に対する攻撃

4.5　WebDAV の脆弱性及びバッファオーバーランを狙った攻撃

“SEARCH /¥x90¥x04H¥x04H･･･(省略)･･･¥x90¥x90”という不正アクセスがたまに来ていた．とても長いアクセスログであるため，少し厄介である．この不正アクセスは W32.HLLW.Gaobot.gen[13]というワームによるもので，Windows 2000，Windows NT，Windows XP の TCP ポート 80 番を使用している WebDAV の脆弱性を狙った攻撃であった．この攻撃に対して，本サーバは応答コード 414 を返す．

また，この不正アクセスと一緒に“POST /_vti_bin/_vti_aut/fp30reg.dll”というアクセスが来ていた．この不正アクセスは Microsoft FrontPage Server Extensions のデバッグコンポーネントでバッファオーバーランを狙った攻撃であった．本サーバでは応答コード 404 を返す．

アクセス数はごく少数で，年間のアクセス数は合計 12 件であった．また，アクセス元を調査した結果，2 月と 4 月は日本から，7 月はアルジェリアから，8 月はルー

マニアからであった[図 11]．2004 年に流行した不正アクセスで，ほとんど対策が行き届いてきているためか，最近はまれにしか見られないようである．

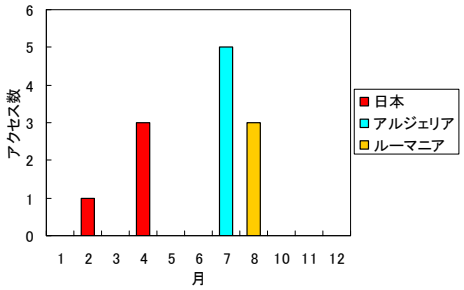

図 11　月別グラフ(WebDAV に対する攻撃)

4.6　アクセス元(国)

情報科学教室 Web サーバに不正アクセスをしてきた国は年間で 34 カ国あった．中国からの不正アクセスが最も多かったが，これは，メール不正中継やオープンプロキシチェックの攻撃回数が多かったためである．これに次ぐ，ドイツ・アメリカからの攻撃は phpMyAdmin の脆弱性を狙ったものが多かった．[図 12]．

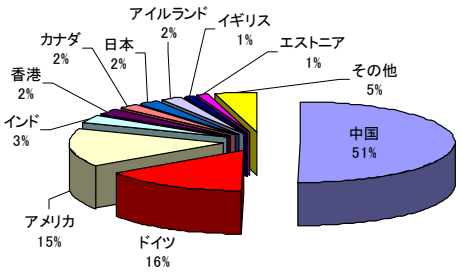

図 12　年間アクセス元国別グラフ

5　おわりに

本研究で調査した期間の結果を以下にまとめる．

●攻撃パターンの種類
　✓メール不正中継の試み
　✓オープンプロキシチェック
　✓CGI-BIN プログラムを狙った攻撃
　✓phpMyAdmin の脆弱性を狙った攻撃
　✓WebDAV の脆弱性を狙った攻撃
　✓バッファオーバーランを狙った攻撃
●どの攻撃も成功していなかった
●不正アクセスによる被害はなかった
●脆弱性を狙った攻撃が多かった
●アクセス元は，中国が最多であった

情報科学教室の Web サーバについて，セキュリティ対策はしっかりと実施されており十分に機能していると考えられる．しかし，セキュリティ対策が万全であっても，システムの安全性を保つために，今後も定期的なログ監査を行う必要はある．

本研究では過去のログデータについて調査を行ったので，日常的なログ監査方法について提案する．

●攻撃パターンに応じた正規表現を使って，攻撃の種類と回数をまとめる
●攻撃元（国）の回数をまとめる
●新規の攻撃パターンを報告する

これらを毎月のログローテートと同時に実行し，ログ監査を行う．また，毎月来ているアクセスで確実に安全であるものは別のログに記録し，問題がある可能性があるものだけをログファイルに残すようにすると，ログ監査が容易に行える．日常的なログ監査は大変であるが，セキュリティ対策やシステムの安全性を保つためにも必要である．

6 謝辞

本研究を進めるにあたりご指導，ご支援頂きました地球環境情報学研究室の皆様に感謝し，心より深く御礼申し上げます.

参考文献

[1] 独立行政法人情報処理推進機構, http://www.ipa.go.jp/, 2008.

[2] 独立行政法人情報処理推進機構, “2007 年コンピュータ不正アクセスの届出状況について”, http://www.ipa.go.jp/security/txt/2008/documents/2007all-cra.pdf, 2008.

[3] Analog, WWW logfile analysis, http://www.analog.cx/, 2005.

[4] wwwstat, HTTPd Logfile Analysis Software, http://ftp.ics.uci.edu/pub/websoft/wwwstat/, 1994-2001

[5] The Webalizer, What is your web server doing today?, http://www.mrunix.net.webalizer, 2008.

[6] Python Software Foundation, Python programming language official web site, http://www.python.org/, 1990-2008.

[7] 不正アクセス行為の禁止等に関する法律, http://law.e-gov.go.jp/htmldata/, H11/H11H0128.html, 1999.

[8] Apache Foundation, Apache HTTP Server Project, http://httpd.apache.org/, 1996-2008

[9]Netcraft, February 2008 Web server survey, http://news.netcraft.com/archives/2008/02/06/february_2008_web_server_survey.html, 2008.

[10] Berners-Lee, T., Fielding, R., and Frystyk, H., Hypertext Transfer Protocol HTTP/1.0 (RFC 1945), http://www.ietf.org/rfc/rfc1945.txt, 1996.

[11] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and Berners-Lee, T., Hypertext Transfer Protocol - HTTP/1.1 (RFC 2616), http://www.ietf.org/rfc/rfc2616.txt, 1999.

[12] phpMyAdmin, The phpMyAdmin Project -Effective MySQL Management-, http://www.phpmyadmin.net/, 2008

[13] Symantec co., Security Response W32.HLLW.Gaobot.gen, http://www.symantec.com/region/jp/sarcj/data/w/w32.hllw.gaobot.gen.html, 2003