Vol.2016-EIP-72 No.1
2016/6/2

# Web アクセスログを活用した組織風土変化の把握に関する考察

上河内 栄治[†1]　森 滋男[†1]　後藤 厚宏[†1]

**概要**：情報セキュリティ対策として企業活動で発生する各種のログを監視する企業が増加している．ログ監視の主な目的はインシデント発生を検知する，またはインシデントの原因や影響範囲を把握することである．これらのログにはインシデント以外にも多くの企業活動情報が含まれている可能性が高く，そこには組織のマネジメント状態を知る指標が含まれているのではないかと考えた．仮説を検証するため実際の Web アクセスログを分析した結果，内部不正発生リスクに関係する日常的なルール違反の発生状況が把握できる可能性がわかった．

# A Study on Web access log utilization for understanding of organizational climate changes

Eiji KAMIGAUCHI[†1]　Shigeo MORI[†1]　Atsuhiro GOTO[†1]

## 1. はじめに

現在，企業における情報資産の価値はますます高まりつつあり，「大切な情報をどのように守るか」が重要な企業活動として認識されてきている．その取り組みのひとつにセキュリティの監視を行う，セキュリティ・オペレーション・センター(以下，SOC)を自社内または，サービスプロバイダ内に設置する動きが活発化している．

NPO 日本ネットワークセキュリティ協会が発表した「2015 年度 国内情報セキュリティ市場調査速報」によれば，セキュリティ運用・管理サービス全体は前年度から 14.5%増加している．なかでもセキュリティ総合監視・運用支援サービスは前年比で 28.4%増加しており，今後も堅調な伸びが期待される[1]としている．

SOC では情報セキュリティインシデントの発生に備え，日々の企業活動で収集される様々なログ情報を監視し，異常の早期発見や予防に努めているが，これらのログには企業活動から記録された多くの情報が含まれており，そのなかに組織のマネジメント状態を知る指標が含まれているはずである．この仮説が正しい場合，収集したログから組織風土[a]やマネジメントに関する指標が発見できる可能性がある．

一方，2015 年度に情報セキュリティ大学院大学の原田研究室が実施したアンケート調査[2]により，現状では内部の不正が情報セキュリティリスクにおける最も大きな脅威として認識されていることがわかっている．

甘利は企業における内部不正対策のひとつに「弁解余地の排除」をあげ，その中で「ルールが形骸化し，守られない状況が常態化すると，ルール違反をすることに口実を与えてしまう」[3]と述べており，島は情報処理推進機構(IPA)の「組織における内部不正防止ガイドライン[4]」策定の際に行ったアンケート調査に基づく分析・考察のうち組織ルールに関する項目について「従業員等が順守していることが不正対策として効果的である」[5]と述べている．

この「ルール違反」の発生が SOC の監視業務で扱うログから検出できるのではないかと考えた．

## 2. Web アクセスログと社内ルール違反

### 2.1 SOC で収集できるログ情報

まず，セキュリティ監視業務に従事している技術者へのヒアリング調査から，SOC で監視しているログ種別の洗い出しを行った．結果は下記のとおりである．

① IPS/IDS[b]ログ
② ファイアウォール(F/W)ログ
③ Proxy ログ(Web アクセスログ)
④ セキュリティ関連機器の死活監視
⑤ ネットワーク帯域
⑥ アンチウイルスログ
⑦ メール送受信ログ
⑧ ファイルサーバ監査ログ
⑨ サーバ，ネットワーク機器の動作ログ
⑩ パケットキャプチャ情報

なお，通常はログに記録された情報全てを詳細に監視しているわけではなく IPS/IDS などが異常検知時に発報するアラームを監視し，その状況に応じて上記のログを詳細調査している．

†1 情報セキュリティ大学院大学

a) 本稿での組織風土は企業全体の風土ではなく，企業内で部や課といった単位で区別された組織毎の風土である

b) IPS:Intrusion Detection System(不正侵入検知システム)，IPS:Intrusion Prevention System(不正侵入防御システム)

### 2.2 企業における「ルール違反」発生状況の把握

つづいて，企業における「ルール違反」の種類を洗い出し，ログからルール違反の発生が判定できるか否かを検討した．

なお，ここでターゲットにしたルール違反は，「2012 年度版　職業上の不正と濫用に関する国民への報告書」[6]による分類で「資産の不正流用」にあたるものとし，「汚職」や「財務諸表不正」については除外した．

洗い出した項目について，発生頻度，情報の入手可否，データからルール違反が判定できるか，SOC 監視業務で扱うログなのかという視点で整理したものが表 1 である．

表 1　企業で発生するルール違反と特徴の整理結果

判例:○多い、容易、実施　△少ない、可能　×非常に少ない、困難、未実施

| 項目 | 発生頻度 | 情報入手可否 | データからの判定 | SOC監視業務 |
|---|---|---|---|---|
| ①社内電話・社給携帯電話の私的利用 | ○ | △ | △ | × |
| ②社内ネットワークからのインターネット私的利用 | ○ | ○ | ○ | ○ |
| ③社給備品の私的利用 | △ | × | × | × |
| ④勤怠情報の不正修正 | △ | △ | ○ | × |
| ⑤遅刻・無断欠勤 | △ | △ | △ | × |
| ⑥各種手当ての不正請求 | △ | △ | △ | × |
| ⑦事務所内ルールを守らない | ○ | △ | × | × |

この結果，日常的に発生していることが想定できるルール違反の種類は多いが，SOC による監視業務で発生状況を把握できる可能性があるものは，社内ネットワークからのインターネット私的利用のみに絞り込むことができた．

Proxy サーバを設置している場合，Web アクセスに関する送信元や接続先 URL などの状況が記録される．この「Web アクセスログ」からインターネットの私的利用を把握できる可能性がある．

しかし，社有電話や携帯電話の私的利用の判定が困難であるのと同様，Proxy サーバの Web アクセスログに残る膨大な通信記録を全て確認することは非現実的である．また，単純に接続先の URL 情報だけで私的利用に該当する通信か否かを判定することも難しい．

### 2.3 Web アクセスログを使ったインターネットの私的利用把握手法

そこで今回の研究では Proxy サーバの Web アクセスログを詳細に調査するのではなく，通信ブロックの発生件数に着目して調査を進めた．なお，Proxy サーバに設定した URL フィルター[c]の条件により通信がブロックされる．

実際に Proxy サーバに残された Web アクセスログを確認したところ，Web アクセスを許可したか，または通信をブロックしたかといった結果とともに，ブロック理由がログに残されていることが確認できた．

この「Web 通信のブロック」がインターネットを私的に利用した場合に多く発生する傾向がわかれば，ブロック件数やブロック発生率によりルール違反（私的利用）の発生状況を把握できることになる．さらに組織毎に発生状況が異なる場合は組織風土の違いがブロックの発生に関連することを説明できる可能性がある．

すなわち Web アクセスに対するブロック発生状況をモニタリングすることで，企業で日常的に発生している一部のルール違反発生や内部不正防止策を実行したことによる組織風土変化の把握につなげることができる．

次の章からは，実際の Web アクセスログを分析した結果から，「ルール違反」の発生が定量的に把握できる可能性を考察する．

## 3. Web アクセスログの分析結果

### 3.1 企業プロフィールとログの概要

Web アクセスログの提供企業といただいたデータのプロフィールは表 2 のとおりである．

表 2　ログの提供企業プロフィールとデータの概要

| 企業プロフィール | |
|---|---|
| 業態 | ICTの運用，保守，関連するコールセンタ業務 |
| 従業員数 | 約2,000人 |
| **ログデータの概要** | |
| ログの種類 | Webアクセスログ(Proxyログ) |
| 収集期間 | 2015年9月～11月(3ヶ月間) |
| ログ件数 | 約90,000,000件/月 |
| 記録情報 | 日時，送信元アドレス，アクセス結果，アクション，通信量，<br>接続先情報，ポート番号，カテゴリー，ポート番号，カテゴリー，<br>フィルタリング結果，フィルタリング理由，等29項目 |

まず，提供された 3 か月間のログを分析し，Web アクセスログからインターネット閲覧の私的利用発生が把握できるか否かを確認した．

なお，この企業ではファイアウォール統合型の Proxy サーバを利用していた．

### 3.2 2015 年 10 月の Web アクセス状況

2015 年 10 月の Web アクセス状況を日付別と時間帯別でまとめたものが図 1，図 2 である．

以降は，Proxy サーバによりブロックされた件数を「ブロック数」とし，時間単位のブロック件数をアクセス総数で除算したものを「ブロック率」と表現する．

なお，ブロック件数のうち，「接続先が見つからない」理由でブロックされたものは件数から除外して集計した．

c) Web サイトの閲覧を制限する機能で，企業では業務に関係がない Web サイトやコンピュータウイルスなどのマルウェアが埋め込まれた危険な Web サイトへのアクセスをブロックするよう設定するケースが多い

Web アクセス件数と発生したブロック率の状況から，休日･夜間はブロック率が高くなる傾向を知ることができた．この傾向は 9 月，11 月についても同様であった．

なお，月により若干変動はあるものの，平均ブロック率は 0.5%～0.6%であった．

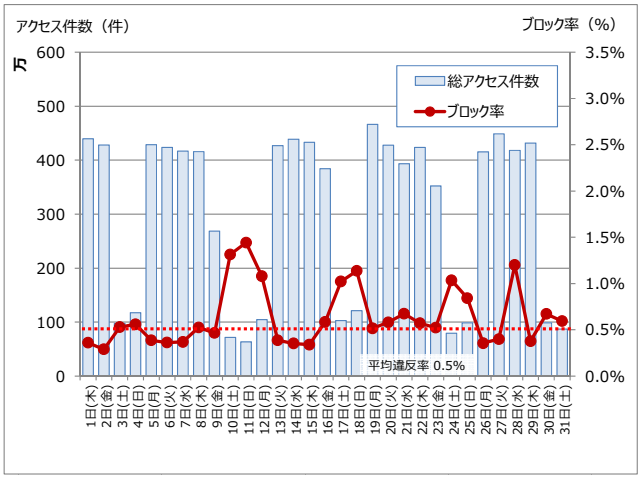


図 1　Web アクセス件数とブロック率(2015/10，日別)

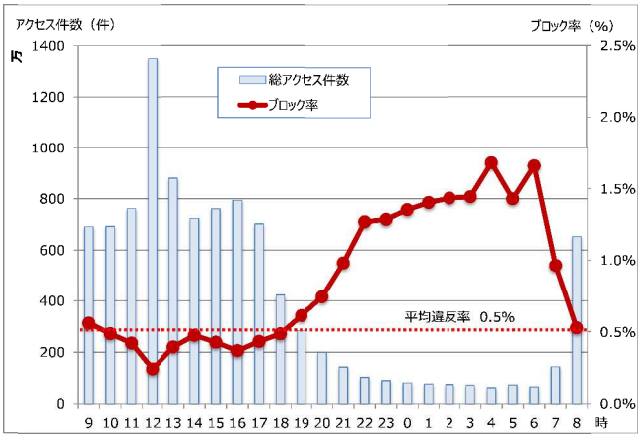


図 2　Web アクセス件数とブロック率(2015/10，時間別)

### 3.3　Web アクセスログの統計的分析結果

前節では，Web アクセス件数とブロック率について日付別，時間帯別にグラフ化した結果を紹介した．

グラフから休日や深夜帯にブロック率が高くなる傾向が見えたが，このことからブロックは意図的なアクセスで発生している可能性が高く，偶発的な発生ではないことが予測できた．この仮説を検証するためさらに統計的な観点で分析を進めた．

分析には 2015 年 8 月 30 日～2015 年 11 月 29 日の Web アクセスログを使い，統計ソフトウェアに Minitab®を利用，有意水準は 5%を採用した．

#### 3.3.1 Web アクセス件数と勤務人数の相関

まず，勤務人数の変化で Web アクセス件数がどのように変化するのかを確認した．

この企業は ICT の運用保守を行っており日常業務で PC を使用するため，勤務人数が多いほど Web アクセス件数も増加すると想定した．

なお，勤務人数を正確に把握する手段がなかったことから，1 時間単位の集計で通信が検出できたユニークな IP アドレス数を勤務人数として代用した．(以降も勤務人数に関しては同様に算出したデータを使用する)

分析には 1 時間単位で集計した Web アクセス件数と，同じく 1 時間単位で集計した IP アドレス数を用い，従属変数に Web アクセス件数，独立変数に IP アドレス数を設定して回帰分析をおこなった．結果は図 3 および，表 3 のとおりである．

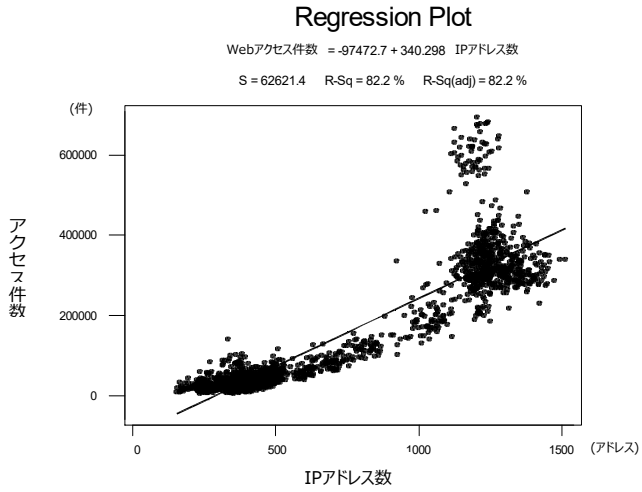


図 3　Web アクセス件数と IP アドレス数の散布図

表 3　Web アクセス件数と IP アドレス数の回帰分析結果

**Regression Analysis: Webアクセス数 versus IPアドレス数**

The regression equation is
アクセス総数 = - 97473 + 340 IPアドレス数

| Predictor | Coef | SE Coef | T | P |
|---|---|---|---|---|
| Constant | -97473 | 2591 | -37.62 | 0.000 |
| IPアドレス数 | 340.298 | 3.393 | 100.29 | 0.000 |

| S = 62621 | R-Sq = 82.2% | R-Sq(adj) = 82.2% |
|---|---|---|

Analysis of Variance

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 3.94387E+13 | 3.94387E+13 | 10057.21 | 0.000 |
| Residual Error | 2182 | 8.55658E+12 | 3921440586 | | |
| Total | 2183 | 4.79953E+13 | | | |

結果は調整済みの決定係数 $R^2$(adj)が 82.2%と，IP アドレス数が多いほど Web アクセスが増加する「強い正の相関」が確認できた．(P 値=0.000)

このことから勤務人数が多い時間帯には Web アクセス数も増加することがわかった．

#### 3.3.2 Web アクセス件数とブロック率の相関

つづいて，Web アクセス件数とブロック率の相関を回帰

分析で確認した.

目的は，ブロックが無意識の Web アクセスで発生する「エラー」なのか，それとも Web 閲覧者が会社のルール違反であることを承知で行った結果発生することが多いのかをブロック発生状況から推察することである.

確認方法として，1 時間単位で集計した Web アクセスのブロック発生率と，同じく 1 時間単位で集計した Web アクセス件数の相関を利用する．違反率が一定であればブロックは故意のルール違反ではなく，ヒューマンエラーの結果発生していることになる.

従属変数にブロック率，独立変数に Web アクセス件数を設定して回帰分析を行った．(図 4，表 4)

結果は調整済みの決定件数 $R^2$(adj)は 19.5%で，ブロック率と Web アクセス件数にはやや負の相関があることが確認できた．(P 値=0.000)

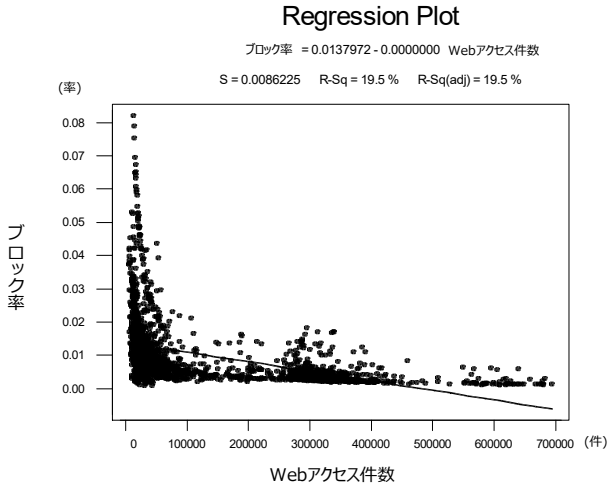


図 4　ブロック率と Web アクセス件数の散布図

表 4　ブロック率と Web アクセス件数の回帰分析結果

**Regression Analysis: ブロック率 versus Webアクセス件数**

The regression equation is
**ブロック率 = 0.0138 -0.000000 Webアクセス件数**

| Predictor | Coef | SE Coef | T | P |
|---|---|---|---|---|
| Constant | 0.0137972 | 0.0002413 | 57.18 | 0.000 |
| Webアクセス件数 | -0.00000003 | 0.00000000 | -23.02 | 0.000 |

| S = 0.008623 | R-Sq = 19.5% | R-Sq(adj) = 19.5% |
|---|---|---|

Analysis of Variance

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 0.039394 | 0.039394 | 529.86 | 0.000 |
| Residual Error | 2182 | 0.162227 | 0.000074 | | |
| Total | 2183 | 0.201621 | | | |

前述したとおり Web アクセスにおけるブロックが Web 閲覧者のヒューマンエラーで発生している場合，ブロック率はアクセス件数に関係なく一定になるはずである．しかし，実データからは Web アクセス件数が多くなるとブロック率が低下する負の相関が示された.

このことから，Web アクセスにおけるブロックの多くは，ルール違反であることを承知でアクセスした結果発生している可能性が高いことがわかった．そうであればブロックは Web 閲覧者のミスなどでたまたま発生したエラーではなく，故意のアクセスにより発生していると考えられる.

### 3.3.3 勤務人数とブロック率の相関

Web アクセス数とブロック率には負の相関があることが確認できたが，同様に IP アドレス数とブロック率の関係についても確認を行った.

仮説としては，Web アクセスと IP アドレス数には強い正の相関が確認できていることから Web アクセスとブロック率と同様に負の相関を示すと予想し，1 時間単位で集計した IP アドレス数と 1 時間単位で集計したブロック率を分析に用いた．従属変数にブロック率，独立変数に IP アドレス数を設定して回帰分析を行った結果が図 5 および，表 5 である.

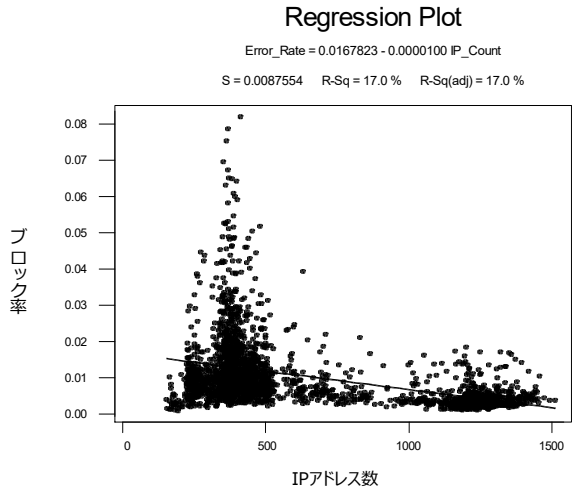


図 5　ブロック率と IP アドレス数の散布図

表 5　ブロック率と IP アドレス数の回帰分析結果

**Regression Analysis: ブロック率 versus IPアドレス数**

The regression equation is
**ブロック率 = 0.0168 -0.000010 IPアドレス数**

| Predictor | Coef | SE Coef | T | P |
|---|---|---|---|---|
| Constant | 0.0167823 | 0.0003623 | 46.32 | 0.000 |
| IPアドレス数 | -0.00001004 | 0.00000047 | -21.17 | 0.000 |

| S = 0.008755 | R-Sq = 17.0% | R-Sq(adj) = 17.0% |
|---|---|---|

Analysis of Variance

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 0.034357 | 0.034357 | 448.20 | 0.000 |
| Residual Error | 2182 | 0.167264 | 0.000077 | | |
| Total | 2183 | 0.201621 | | | |

調整済みの決定件数 R2(adj)は 17.0%で，ブロック率の高さと IP アドレス数には弱い負の相関があることが確認できた．(P 値=0.000)

この分析から，勤務人数の少ない時間に，ブロックが発

生する Web アクセスが行わる傾向が確認できた．この事実から周囲に勤務者が多い場合,ブロックが発生するような,なるべく人にみられたくない Web サイトの利用が抑制されていると考えられる．(インターネットの私的利用)

今回の分析結果からも 2.3 で述べた「Web アクセスログのブロック件数からインターネットの私的利用判定」ができ，一部のルール違反発生状況の把握にブロック率が活用できる可能性が高いことがわかった．

ここまでの分析で，ブロックは私的利用というルール違反の結果発生しているケースが多いと判断し，以降はブロック率を違反率，ブロック件数を違反件数と表現する．

### 3.4 組織風土と違反率の相関

ここまでの分析から，Proxy サーバによる Web アクセスのブロックは Web 閲覧者がルール違反であることを承知しながらアクセスした結果発生しており，ブロック数のカウントは日常の「ルール違反」発生状況把握に利用できる可能性が高いことがわかった．

ただし，現時点ではあくまで Web 閲覧者個人のルール違反であり，組織全体の風土を測る指標として利用する場合は組織毎のブロック発生傾向に差がなくてはならない．

そこで違反率を従属変数に設定し，組織間で統計的に有意差が確認できる独立変数を見つけ出す作業を行った．

なお,以降の分析は IP アドレスにより所属組織が判断できた 33 の組織[d]に絞って調査を行った．

#### 3.4.1 違反率に差を生む要因(独立変数)の検討

分析を進めるにあたり，違反率に影響を及ぼすと考えられる要因を洗い出した．仮説として組織風土が違反率に影響すると考えており，要因は組織風土との関係が予想できた変数を中心に検討した．結果は下記のとおりである．

①測定月

測定月についてはデータのプロフィールを確認した際，月毎の平均ブロック発生率に若干の違いがあったことから，分析対象に含めた．

②時間

時間についてもデータのプロフィールを確認した際，通常の勤務時間帯と深夜・早朝で大きな違いがあることがわかっており，統計的にも有意な差があることを確認する目的で分析対象に含めた．

③組織

仮説としてルールが守られないことが常態化しているとルール違反が増加すると考えている．

インターネットの私的利用もルール違反に該当するため，違反率は組織風土により異なると予測した．

④職種

データ提供を受けた企業では組織毎に従業員の主な職種が異なっている．また，職種により 24 時間交代勤務を行っているケースもあり違反率に影響があると考えた．

⑤組織の所属人数

組織長がマネジメントできる人数は 10 から 15 名程度とされることが多い．このため組織に所属する人数が多い場合マネジメントがおろそかになり，違反を許す風土につながっている可能性がある．なお，分析は単純に所属人数を量的変数として設定したケースと，15 名を基準にそれより多いか否かを質的変数として設定したケースで調査した．

⑥退職者数

違反には直接関係しないと考えたが，組織風土を測る指標として退職率が使われることがあることから，本要因についても量的変数として退職者数，質的変数として退職者の有無を設定して調査した．

⑦長欠者数

長欠者数についても退職者数と同様である．

本要因についても量的変数として長欠者数，質的変数として長欠者の有無を設定して調査した．

⑧懲戒者数

懲戒については組織風土の影響で発生し組織全体が処分されるケースと，個人の行動により，該当者のみが処分されるケースがある．この懲戒についても違反を許す組織風土が関係している可能性があるため影響を確認した．

#### 3.4.2 月別違反率の分散

表 6　測定月に対する違反率の分散分析結果

**One-way ANOVA: 違反率 versus 測定月**

Analysis of Variance for Deny(Int

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| 測定月 | 2 | 0.0000419 | 0.0000209 | 1.91 | 0.154 |
| Error | 96 | 0.0010535 | 0.0000110 | | |
| Total | 98 | 0.0010954 | | | |

| Level | N | Mean | StDev | Individual 95% CIs For Mean Based on Pooled StDev |
|---|---|---|---|---|
| 11月 | 33 | 0.004252 | 0.003492 | (--------*--------) |
| 10月 | 33 | 0.003871 | 0.003388 | (--------*--------) |
| 9月 | 33 | 0.002722 | 0.003042 | (---------*--------) |

Pooled StDev = 0.003313　　0.0024　0.0036　0.0048

d) 複数の組織でDHCPスコープを共用している場合が調査対象外となった．これは提供いただいたデータだけではIPアドレスと組織を紐づけることができないためである．

### 3.4.3 時間に着目した違反率の分散

つづいて時間と時間帯の違いで違反率に差が見られるか否かを確認した.

まずは 1 時間毎の違反率の分散を確認した．測定時間を独立変数に設定して分散分析した結果が表 7 である．結果, P 値は 0.05 以下であり，時間により違反率に有意差があることが確認できた.

なお，時間は勤務が始まる 9 時から翌日 8 時までを 1 日と考え，9 時から 32 時までの 24 時間で表現している.

表 7　時間と違反率の分散分析結果

**One-way ANOVA: 違反率 versus 時間**

Analysis of Variance for Error ra

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| 違反率 | 23 | 0.0012962 | 0.0000564 | 6.73 | 0.000 |
| Error | 48 | 0.0004019 | 0.0000084 | | |
| Total | 71 | 0.0016981 | | | |

Individual 95% CIs For Mean Based on Pooled StDev

| Level | N | Mean | StDev |
|---|---|---|---|
| 09H | 3 | 0.005608 | 0.000750 |
| 10H | 3 | 0.004896 | 0.000772 |
| 11H | 3 | 0.004601 | 0.001264 |
| 12H | 3 | 0.002378 | 0.000564 |
| 13H | 3 | 0.003792 | 0.000817 |
| 14H | 3 | 0.004571 | 0.001275 |
| 15H | 3 | 0.004159 | 0.001159 |
| 16H | 3 | 0.003916 | 0.001120 |
| 17H | 3 | 0.004432 | 0.001047 |
| 18H | 3 | 0.005304 | 0.000862 |
| 19H | 3 | 0.006318 | 0.000541 |
| 20H | 3 | 0.007578 | 0.000818 |
| 21H | 3 | 0.009521 | 0.002026 |
| 22H | 3 | 0.011346 | 0.001757 |
| 23H | 3 | 0.012447 | 0.002303 |
| 24H | 3 | 0.012492 | 0.003648 |
| 25H | 3 | 0.013203 | 0.003797 |
| 26H | 3 | 0.013869 | 0.004607 |
| 27H | 3 | 0.014128 | 0.006150 |
| 28H | 3 | 0.015742 | 0.005111 |
| 29H | 3 | 0.013580 | 0.005657 |
| 30H | 3 | 0.014764 | 0.005475 |
| 31H | 3 | 0.009608 | 0.001224 |
| 32H | 3 | 0.006011 | 0.001156 |

Pooled StDev = 0.002894　　0.0000　0.0060　0.0120　0.0180

つづいて時間帯別にも有意差があるか否かを確認した．時間帯を始業前(8:00 から 8:59)，午前中(9:00 から 11:59)，昼休み(12:00 から 12:59)，午後(13:00 から 17:59)，定時間後(18:00 から 21:59)，深夜(22:00 から翌日 7:59)にわけて違反率との相関を調査した結果が表 8 で，違反率は時間帯で有意差があることがわかった．(P 値=0.000)

表 8　時間帯と違反率の分散分析結果

**One-way ANOVA: 違反率 versus 時間帯**

Analysis of Variance for 違反率

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| 時間帯 | 5 | 0.0011913 | 0.0002383 | 31.03 | 0.000 |
| Error | 66 | 0.0005068 | 0.0000077 | | |
| Total | 71 | 0.0016981 | | | |

Individual 95% CIs For Mean Based on Pooled StDev

| Level | N | Mean | StDev |
|---|---|---|---|
| 始業前 | 3 | 0.006011 | 0.001156 |
| 午前中 | 9 | 0.005035 | 0.000943 |
| 昼休み | 3 | 0.002378 | 0.000564 |
| 午後 | 18 | 0.004362 | 0.001028 |
| 定時間後 | 9 | 0.007806 | 0.001794 |
| 深夜 | 30 | 0.013118 | 0.003951 |

Pooled StDev = 0.002771　　0.0000　0.0050　0.0100　0.0150

特に深夜帯で違反率が高くなっているように見えるが，定常的に深夜勤務がない組織があることを考慮し，以降の分析では定時間内(9:00 から 17:59)の違反率に限定して分析を進めた.

### 3.4.4 組織別違反率の分散

次に組織別の違反率に差があるか否かを確認した．違反率が組織毎の風土を反映しているとすれば，組織毎に有意差があることが統計的に確認できるはずである．組織を独立変数に設定し分散分析を行った結果が表 9 である.

表 9　組織と違反率の分散分析結果

**One-way ANOVA: 違反率(定時間) versus 組織**

Analysis of Variance for 違反率(定時間)

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| 組織 | 32 | 0.0009738 | 0.0000304 | 13.22 | 0.000 |
| Error | 66 | 0.0001519 | 0.0000023 | | |
| Total | 98 | 0.0011257 | | | |

Individual 95% CIs For Mean Based on Pooled StDev

| Level | N | Mean | StDev |
|---|---|---|---|
| EDU | 3 | 0.001497 | 0.000975 |
| OKA | 3 | 0.002371 | 0.001271 |
| OKB | 3 | 0.002425 | 0.000394 |
| OMA | 3 | 0.005176 | 0.002861 |
| REA | 3 | 0.001645 | 0.000551 |
| REB | 3 | 0.005075 | 0.001261 |
| REC | 3 | 0.001845 | 0.000314 |
| RED | 3 | 0.002829 | 0.001134 |
| REE | 3 | 0.002394 | 0.000689 |
| REF | 3 | 0.002839 | 0.001694 |
| RH | 3 | 0.003829 | 0.002697 |
| RI | 3 | 0.003909 | 0.001066 |
| RJ | 3 | 0.005299 | 0.001247 |
| RKA | 3 | 0.002572 | 0.000211 |
| RKB | 3 | 0.002667 | 0.001470 |
| RKC | 3 | 0.000909 | 0.000235 |
| RKD | 3 | 0.001965 | 0.000765 |
| RLA | 3 | 0.000602 | 0.000117 |
| RLB | 3 | 0.000605 | 0.000324 |
| RLC | 3 | 0.001307 | 0.000407 |
| RWA | 3 | 0.002659 | 0.000905 |
| RWB | 3 | 0.002975 | 0.001133 |
| RWC | 3 | 0.003628 | 0.001579 |
| RWD | 3 | 0.004440 | 0.004296 |
| RWE | 3 | 0.002078 | 0.000457 |
| RWF | 3 | 0.002682 | 0.000165 |
| SA | 3 | 0.001615 | 0.000408 |
| SB | 3 | 0.002659 | 0.000691 |
| SC | 3 | 0.004405 | 0.003034 |
| SD | 3 | 0.001976 | 0.000956 |
| SE | 3 | 0.004367 | 0.001109 |
| SF | 3 | 0.004020 | 0.001874 |
| SG | 3 | 0.019538 | 0.002554 |

Pooled StDev = 0.001517　　0.0000　0.0070　0.0140　0.0210

分析の結果，P 値は 0.05 以下であり組織の違いで違反率に有意差があることがわかった.

仮に違反率がその組織のルール違反が発生しやすい組織風土を反映しているとすれば，URL フィルターでブロックされた違反件数や違反率を定期的に観測することで組織風土の変化を知ることができる可能性を示した結果であると考えられる.

### 3.4.5 職種別違反率の分散

次に職種の違いが違反率に影響しているか否かを確認した.分析に関しては 33 の組織を職種毎に 7 グループに分類して独立変数とした．分散分析の結果が表 10 である.

表 10　職種と違反率の分散分析結果

**One-way ANOVA: 違反率(定時間) versus 職種**

Analysis of Variance for 違反率(定時間)

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| category | 6 | 0.0002014 | 0.0000336 | 3.34 | 0.005 |
| Error | 92 | 0.0009243 | 0.0000100 | | |
| Total | 98 | 0.0011257 | | | |

| Level | N | Mean | StDev | Individual 95% CIs For Mean Based on Pooled StDev |
|---|---|---|---|---|
| 職種A | 9 | 0.003324 | 0.002102 | (-------*--------) |
| 職種B | 12 | 0.002028 | 0.001026 | (------*------) |
| 職種C | 9 | 0.000838 | 0.000442 | (-------*--------) |
| 職種D | 9 | 0.004346 | 0.001733 | (-------*--------) |
| 職種E | 36 | 0.002924 | 0.001638 | (----*---) |
| 職種F | 3 | 0.001497 | 0.000975 | (-------------*-------------) |
| 職種G | 21 | 0.005511 | 0.006148 | (----*-----) |

Pooled StDev = 0.003170　　0.0000　0.0025　0.0050

結果は P 値が 0.05 以下であり，職種の違いでも違反率に有意差があることが確認できた．

### 3.4.6 組織の所属人数

組織と職種で違反率に違いがあることが確認できたが，組織のマネジメント状況が違反率に影響しているのであれば，組織に所属する従業員数の影響が考えられる．

まずは従業員数を独立変数，違反率を従属変数に設定し回帰分析で相関を確認した．結果が図 6 および，表 11 である．

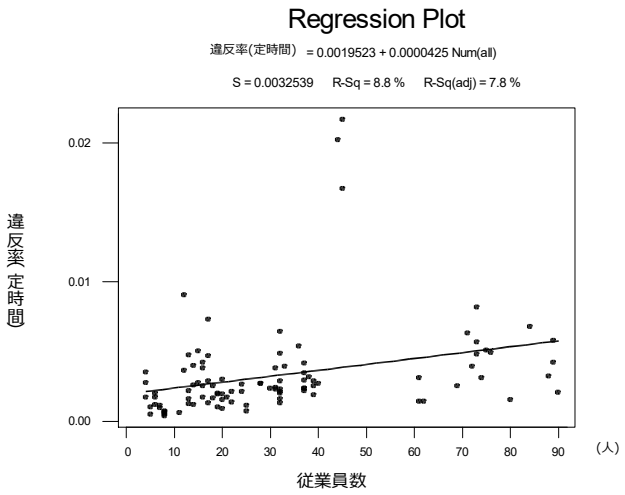


図 6　違反率と従業員数の散布図

表 11　違反率と従業員数の回帰分析結果

**Regression Analysis: 違反率(定時間) versus 従業員数**

The regression equation is
違反率(定時間) = 0.00195 +0.000042 従業員数

| Predictor | Coef | SE Coef | T | P |
|---|---|---|---|---|
| Constant | 0.0019523 | 0.0005487 | 3.56 | 0.001 |
| 違反率(定時間) | 0.00004245 | 0.00001391 | 3.05 | 0.003 |

S = 0.003254　R-Sq = 8.8%　R-Sq(adj) = 7.8%

Analysis of Variance

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| Regression | 1 | 0.00009865 | 0.00009865 | 9.32 | 0.003 |
| Residual Error | 97 | 0.00102705 | 0.00001059 | | |
| Total | 98 | 0.00112570 | | | |

回帰分析の結果，P 値は 0.05 以下で有意であり，従業員数が多くなると違反率が高くなる弱い正の相関があることが確認できた．

組織に所属する従業員数が多くなると一人一人の行動までマネジメントが行き届かず，ルール違反の発生が管理しきれないため違反が多く発生している可能性がある．

回帰分析で従業員数と違反率に相関があることが確認できたが，次に組織の人数で大小を定め，違反率に影響があるのか否かを確認した．

ここでは一般的に 1 名のマネージャがマネジメントできる従業員数はせいぜい 15 名までといわれることが多いことから，15 名以下の組織を小規模，16 名以上の組織を大規模として区別し，分散分析を行った．結果が表 12 である．

表 12　組織の大小と違反率の分散分析結果

**One-way ANOVA: 違反率(定時間) versus 組織の大小**

Analysis of Variance for 違反率(定時間)

| Source | DF | SS | MS | F | P |
|---|---|---|---|---|---|
| 組織の大小 | 1 | 0.0000496 | 0.0000496 | 4.47 | 0.037 |
| Error | 97 | 0.0010761 | 0.0000111 | | |
| Total | 98 | 0.0011257 | | | |

| Level | N | Mean | StDev | Individual 95% CIs For Mean Based on Pooled StDev |
|---|---|---|---|---|
| 小規模 | 27 | 0.002141 | 0.001928 | (----------*----------) |
| 大規模 | 72 | 0.003730 | 0.003714 | (------*------) |

Pooled StDev = 0.003331　　0.0010　0.0020　0.0030　0.0040

分散分析の結果，P 値は 0.05 以下であり組織の大小で違反率に差があることが統計的に確認できた．組織に所属する人数が多い場合，マネジメントが行き届かずルール違反を見逃すことが多くなり，違反率が高くなっている可能性がある．

この結果からも Web アクセス時に発生するブロック件数を組織風土変化の指標として利用できる可能性が高いことが示されている．

### 3.4.7 その他の項目

退職者，長欠者，懲戒の有無を独立変数に設定し違反率に統計的な有意差があるか否かを確認したが，全て P 値が 0.05 を超え，これらの要因と違反率に相関があるとはいえないことがわかった．

ただし，退職，長欠，懲戒については期間中の発生がわずかであり，このことが分析結果に影響した可能性も考えられる．

## 4. 考察

現在までの分析で Web アクセス時のブロックは Web 閲覧者の故意の私的利用により発生している可能性が高いことがわかった．

これまで，Web アクセスログからインターネットアクセスの私的利用を調査しようとした場合，サンプル調査または，時間を絞って調査するログの量を少なくする必要があり，全数確認は実質不可能であった．

しかし，今回の調査で用いた Web アクセスのブロック件数をカウントする手法を用いることで，企業におけるインターネット私的利用の発生が効率的に把握できる可能性が高いことがわかった．

さらには組織毎の違反率に有意差があることも判明したことで，組織風土の違いがブロック率に影響している可能性も示すことができたと考えている．

ただし違反率には職種や時間帯の違いにより有意差が生じることもわかり，仮に違反率が「違反の発生しやすい組織風土」の把握に活用できるとしても，職種の異なる組織同士や勤務形態の異なる組織間で単純に違反率比較し，組織風土やマネジメントの良し悪しを判断することはできないと考えている．

それよりも定期的な Web アクセスログの分析から，違反率がどのように変化したかを把握し「組織風土の状況変化を測る指標」として活用すれば，組織マネジメントをサポートするツールとしてより有効に活用できるのではないだろうか．

## 5. まとめ

ここまでの研究で Web アクセスログに残る通信のブロック情報から，インターネットの私的利用という「ルール違反」の発生が把握できる可能性が高く，「決められたルールを順守する組織風土」の状態を把握することにも活用できる可能性を示すことができた．

この手法は内部不正発生防止策の効果測定に利用することができると考えている．そうであれば SOC による監視業務の付加価値になる．

ただし，現段階では違反の発生傾向から「故意のルール違反によりブロックが発生している可能性が高い」としか主張できないため，今後，ブロック理由やブロック時に接続しようとした URL のサンプル調査まで範囲を広げ，仮説を検証する．

## 謝辞

本研究に賛同していただきログをご提供くださった企業の皆様に感謝いたします．さらに研究を進めるにあたり，活発な議論，助言をいただいた情報セキュリティ大学院大学後藤研究室の皆様に感謝いたします．

## 参考文献

[1] “2015 年度 国内情報セキュリティ市場調査 速報(公開用資料)”.http://www.jnsa.org/result/2016/surv_mrk/data/2015_mrk-report_sokuhso.pdf,(参照 2016-03-12)
[2] “2015 年情報セキュリティ アンケート調査結果”. http://lab.iisec.ac.jp/~harada_lab/survey/2015/2015_questionnaire_result.pdf,(参照 2016-03-12)
[3] 甘利康文.内部不正抑制に応用できる犯罪理論.内部不正対策 14 の論点,2015,193p
[4] “組織における内部不正防止ガイドライン（日本語版）第 3 版”. http://www.ipa.go.jp/files/000044615.pdf,(参照 2016-03-22)
[5] 島成佳.職場環境の整備で防ぐ内部不正.内部不正対策 14 の論点,2015,102p
[6] “職業上の不正と濫用に関する国民への報告書(日本語訳)”. http://www.acfe.jp/upload/RTTN2012_J.pdf.,(参照 2016-03-13)