明治大学総合数理学部

2016 年度

卒　業　研　究

内部犯行を誘発する環境の機械学習による分析

学位請求者　先端メディアサイエンス学科

山 田 道 洋

目次

第 1 章　はじめに　1
- 1.1　研究背景　1
- 1.2　関連研究　1
- 1.3　研究目的　1

第 2 章　分析手法　3
- 2.1　決定木　3
- 2.2　連関規則　3
- 2.3　ロジスティック回帰分析　3
- 2.4　共有アカウントの利用　3

第 3 章　実験　5
- 3.1　実験 1 ：カレーについてのアンケート　5
- 3.2　実験 2:検索エンジンの評価業務　11

第 4 章　おわりに　18

参考文献　20

付録 A　実験 1 の作業内容等　21
- A.1　利用規約　21
- A.2　作業内容　21

付録 B　実験 2 の作業内容等　23
- B.1　利用規約　23
- B.2　禁止事項　23

付録 C　https サイトの証明書利用状況についての報告　24
- C.1　はじめに　24
- C.2　実験　24
- C.3　おわりに　27

参考文献　28

第 1 章

はじめに

1.1 研究背景

情報セキュリティマネージメントにおいて，2014 年に発生したベネッセの顧客情報流出事件 [1] のような内部犯行は最も防止が困難な大きな脅威である．また，2014 年には共有アカウントを使用していた職員が退職後も，同じ ID 等を共有アカウントとして利用し続けたために発生した個人情報流出事件もある [8][9]．このような内部犯行事件は事例も少なく，どのような要因が内部犯行を誘発しているのか観測することは困難である．また，企業等において内部犯行の仮定が観測できたとしても詳細を公開することはその企業の情報セキュリティポリシに抵触する可能性がある．

1.2 関連研究

Cappli らは実際の犯罪記録をもとにして内部不正の誘発要因の特徴を類推し，傾向をモデル化するための，MERIT1 を提案している [2]．社会安全研究財団は，国内のサイバー犯罪のうち，内部不正を対象として事例分析を行い，犯行者の心理的力動過程（ダイナミクス）を提示した [3]．また，Nurse らは，内部不正の特徴に関するフレームワークを提案している [4]．ただし，これらのツールはセキュリティ担当者や管理者が内部不正の問題を理解し，リスクを分析するためのツールとしてはよいが，どの誘発要因がより内部不正を誘発する影響については明らかに出来ていない．

Hausawi は，エンドユーザが行うセキュリティに関する振舞いについてセキュリティ専門家に対してインタビューを行った [5]．インタビューの結果，エンドユーザが行う最も否定的な振舞いは認証情報の共有であった．この共有とは，例えばシステム開発チームがサーバにアクセスする認証情報を共有したり，コールセンタのスタッフが機密情報にアクセスする認証情報を共有したりすることを指す．また，IPA は共有 ID の利用は内部不正発生時に利用者の識別が困難なことから心理的に重要情報を持ち出しやすい環境となると指摘している [6]．新原らは，第三者による監視が低い場合に不正事象が発生する確率が高くなるという実験結果を報告している [7]．

1.3 研究目的

本研究では，内部犯行を誘発または，抑制する要因を明らかにして，マネージメントに活用することを目的とする．共有ゲストアカウントの利用など異なる誘発条件を与えた合計 198 名の被験者に単純なタスクを行

わせ，内部犯行の有無を調査した．この実験結果について，(1) 内部犯行を行う条件を表す決定木，(2) 不正を表す連関規則の抽出，(3) 各属性の不正に対するオッズ比を求めるロジスティック回帰の分析を行う．

第 2 章

分析手法

本実験では決定木と連関規則という 2 つの手法を利用して実験結果の分析を行う．また，各属性による影響の確率検定をおこなうためにロジスティック回帰分析を行う．

2.1　決定木

決定木は，ターゲットである属性を決定する論理条件を明らかにする機械学習であり，根に近い属性が最も大きな条件となる属性である．本実験では，R のパッケージ “rpart” により学習した決定木を使用する．

2.2　連関規則

実験結果から属性の組み合わせにより不正への影響があったかを明らかにする為に，R のパッケージ “arules” を使用して連関規則の抽出を行った．数値などについての説明は 3.1.2 にて後述する．

2.3　ロジスティック回帰分析

ロジスティック回帰は，Intercept という基準を設定し対応する各属性との比較で計算している，ロジスティックモデルでは，不正を犯す確率 p を，各属性を表す論理変数 x_1，x_2 などを用いて，

$$log\frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots$$

と定める線形式で推定する．この係数 β_0，β_1 が後述の表 3.5 などの Estimate である．

また，各属性と基準となる属性の不正を犯す確率を比較した時のオッズ比も計算した．

2.4　共有アカウントの利用

本実験では，内部犯行の誘発要因として，被験者の利用アカウントを共有アカウント (ID: guest) と個別アカウントに分け，作業を行ってもらう．

共有アカウントと個別アカウントの違いの例を図 2.1 に示す共有アカウントには個別アカウントよりも管理コストが低く済むという利点がある．これは利用者が増えた場合などでも新たに ID やパスワードを発行する必要がないからである．しかし，共有アカウントは個別アカウントと比べて内部犯行を誘発しやすいと言われ

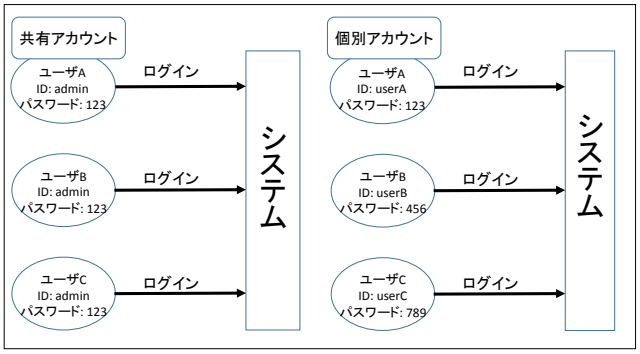

図 2.1　共有アカウントと個別アカウントの違いの例

ている．例えば，図 2.1 の共有アカウントの A が不正を犯しても B や C と区別がつかず，作業者の特定が困難で，不正がばれないという心情にさせてしまうからである．

第 3 章

実験

3.1　実験 1 ：カレーについてのアンケート

3.1.1　実験概要（1）

被験者は 2016 年 6 月 26 日〜29 日にクラウドソーシングサイトで募集した 192 名とし，被験者を共有ゲストアカウント（ID: guest）と個別アカウントの 2 つのグループに自動的に振り分けて，筆者らが作成した WEB サイトにてカレーについてのアンケートへの回答と PDF の文字入力作業を行ってもらい，不正の有無を観測する（以下，実験 1 と呼ぶ）.

作業の流れ (1)

実験サイト上でのページ遷移図を図 3.1 に示す．実験サイトの実行例を図 3.2 に示す．各ユーザへは実験サイトへアクセス後，ID が guest の共有アカウントか，ランダムな文字列の個別アカウントを払い出す．その後，払い出された ID を使用し，ログインすることで，利用規約の確認，作業へと進む.

作業内容（1）

被験者に行ってもらう作業は以下の 2 つである.

1. カレーについてのアンケート
   カレーの好みや，食べる頻度など，カレーについてのアンケートへの回答をしてもらう．設問は全 14 問で 5 つの選択肢からの選択式である.
2. PDF の文字入力
   PDF に記入されている文字列をテキストボックスへと入力してもらう.

不正行為

実験 1 では，規約により禁止した以下の 2 つの行為を不正行為と定義する.

1. 越権行為
   実験 1 では，回答の編集及び管理者ボタンの押下を行為を行ったユーザは自らの権限を越えた不正事象とみなし，「越権行為」と定義した.
2. コピペ

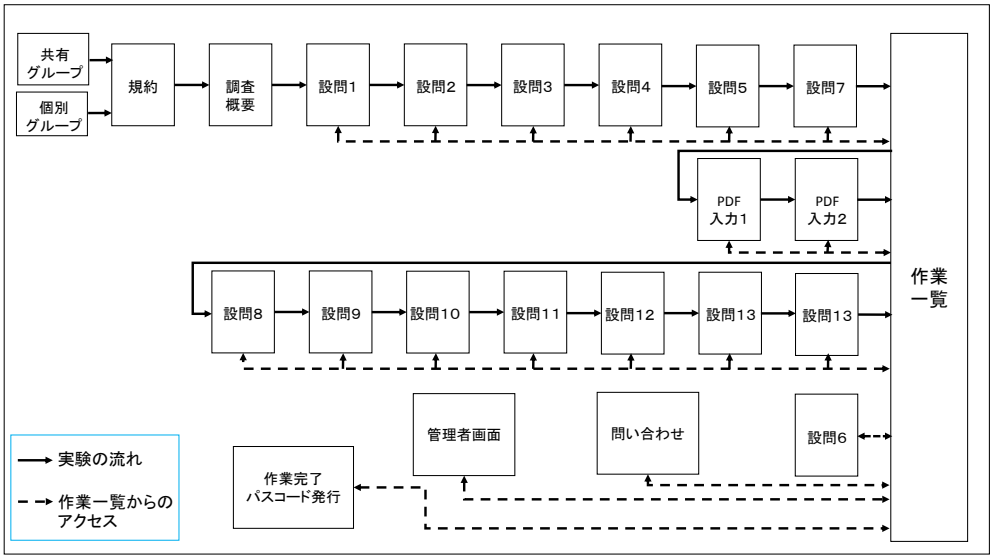

図 3.1　実験サイト遷移図

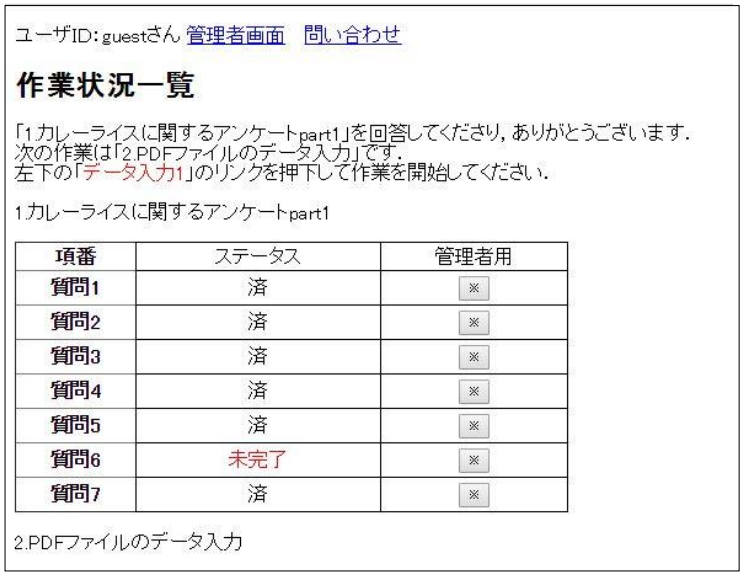

図 3.2　実験サイト実行例

PDF の文字入力の作業において，文字列をコピー＆ペーストを使用して入力することを不正事象とみなし，「コピペ」と定義した．

不正行為の誘発

アンケートは，設問 1～5 までは順に表示されるが，設問 6 は表示されず設問 7 が表示される．この場合，回答状況は図 3.2 のようになり，ユーザは管理者への報告を行わなければ正常に作業を終了することが出来な

表 3.1　アンケートの各グループ毎の人数

	共有	個別	Total
男性	36	42	78
女性	57	57	114
19 才以下	3	2	5
20 才〜29 才	22	26	48
30 才〜39 才	29	41	70
40 才〜49 才	30	22	52
50 才〜59 才	6	8	14
60 才〜69 才	3	0	3
会社員 (d)	18	34	52
自営業 (b)	22	14	36
学生 (g)	9	7	16
専業主婦，専業主夫 (c)	22	15	37
パート，アルバイト (f)	11	15	26
無職 (a)	7	6	13
その他 (e)	4	8	12
N	93	99	192

表 3.2　アンケートの各グループ毎の不正者数

	越権行為	コピペ
共有	14	28
個別	18	35
総計	32	63

い．これは越権行為が一定数発生することを期待し，実施した．

3.1.2　実験結果 (1)

各グループ毎の人数を表 3.1 に示す．職業のカッコ内は，作成した決定木の対応記号を示す．各グループ毎の不正者数を 3.2 に示す．

決定木 (1)

「越権行為」をしたかどうか，「コピペ」をしたかどうかをターゲットとして作成した決定木をそれぞれ図 3.3，図 3.4 に示す．ここで，「Job=bcdefg」等の分岐の条件を各節点の上に示し，左側の枝が条件にあてはまる．「不正者数/正規者数」を各節点の下に示す．“Malicious”や “OK”は人数の多い方を示す．例えば，図 3.3 の木では職業が無職以外 (無職=a) かどうかが不正を犯すかどうかを決める最も大きな条件であり，無職

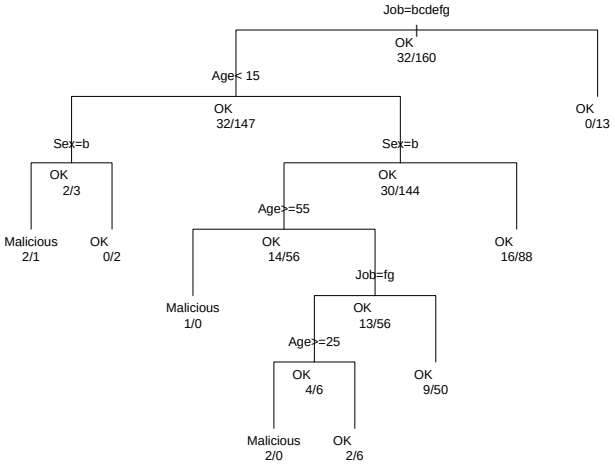

図 3.3 「越権行為」の決定木

には 0 名の不正者と，13 名の正規者がいる．図 3.4 の木では職業が無職，自営業かどうかが最も大きな条件である．グループが共有か，個別かによる分岐は現れなかった．

連関規則 (1)

「越権行為」,「コピペ」それぞれの抽出された連関規則の一部を表 3.3，表 3.4 に示す．support（支持度）は同時確率 $p(lhs, rhs)$，すなわち条件部 lhs と結論部 rhs が同時に起こる確率である．confidence（確信度）は lhs で条件付けられた rhs の条件付き確率 $p(rhs|lhs)$ すなわち lhs の属性の組み合わせを持つ被験者の中で rhs が発生する確率である．例えば，表 3.3 の No.1 の規則は，「個別グループかつ年齢が 40 代の被験者は約 90% の確率で正規者 (Judge1=OK)」であることを意味している．lift は改善率 $p(rhs|lhs)/p(rhs)$ すなわちターゲットとする rhs が全体で発生する確率に対する lhs の条件付き確率確率がどれだけ向上するかを示す．改善率が高いほどその規則が有用である．

「越権行為」については，No.1 で個別グループの場合に正規者であるという規則が抽出されたが，No.2〜4 のように共有グループで年齢や性別が特定の条件の場合に正規者であるという規則も抽出された．また，不正を犯す条件についての規則は抽出されなかった．

「コピペ」については，No.1〜2 で不正を犯す場合の，No.3〜7 で正規者である場合の規則が抽出された．No.1 では「個別グループ」の場合に，No.2 では「個別グループかつ性別が男性」の場合に不正者であるという規則が抽出された．正規者についての規則は個別グループ，共有グループどちらの場合にも特定の条件の時に正規者であるという規則が抽出された．

ロジスティック回帰分析 (1)

「越権行為」,「コピペ」それぞれのロジスティック回帰分析の結果を表 3.5，表 3.6 に示す．本分析の Intercept の持つ属性は「共有アカウント・女性・その他」である．「越権行為」,「コピペ」どちらのロジス

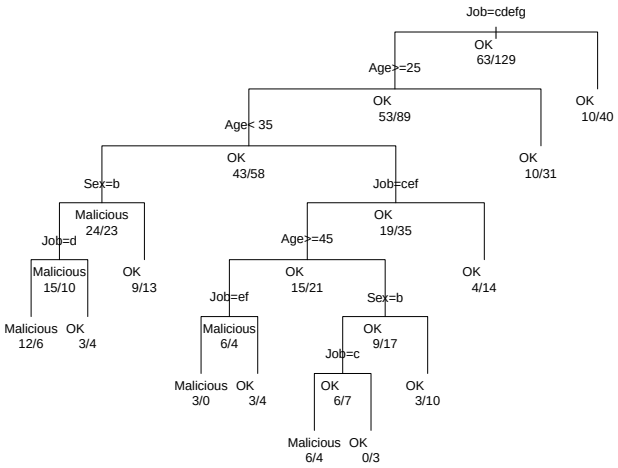

図 3.4　「コピペ」の決定木

表 3.3　「越権行為」の連関規則 (一部)

No.	lhs(条件部)	rhs（結論部）	support	confidence	lhs.support	lift
1	{group=個別,Age=40's} =>	{Judge1=OK}	0.1041667	0.9090909	0.1145833	1.090909
2	{group=共有,Age=20's} =>	{Judge1=OK}	0.1041667	0.9090909	0.1145833	1.090909
3	{group=共有,Age=30's} =>	{Judge1=OK}	0.1354167	0.8965517	0.1510417	1.075862
4	{group=共有,Sex=女性 } =>	{Judge1=OK}	0.2604167	0.877193	0.296875	1.052632

表 3.4　「コピペ」の連関規則 (一部)

No.	lhs(条件部)	rhs（結論部）	support	confidence	lhs.support	lift
1	{group=個別 } =>	{Judge2=Malicious}	0.1822917	0.3535354	0.515625	1.077441
2	{group=個別,Sex=男性 } =>	{Judge2=Malicious}	0.1145833	0.5238095	0.21875	1.596372
3	{group=個別,Age=20's} =>	{Judge2=OK}	0.1145833	0.8461538	0.1354167	1.259392
4	{group=個別,Sex=女性 } =>	{Judge2=OK}	0.2291667	0.7719298	0.296875	1.148919
5	{group=共有 } =>	{Judge2=OK}	0.3385417	0.6989247	0.484375	1.04026
6	{group=共有,Age=40's} =>	{Judge2=OK}	0.109375	0.7	0.15625	1.04186
7	{group=共有,Sex=男性 } =>	{Judge2=OK}	0.1354167	0.7222222	0.1875	1.074935

ティック回帰分析からも，各属性による不正への影響の有意差（p 値 < 0.05）は見られなかった．オッズ比からは「越権行為」では個別アカウントは共有アカウントと比べて 0.16 倍に不正が抑制されるという結果が得られた．一方，「コピペ」では個別アカウントは共有アカウントよりも 1.2 倍不正が発生しやすいという結果が得られた．

表 3.5　「越権行為」のロジスティック回帰分析

	Estimate	Pr(>|t|)	Odds
(Intercept)	0.144733	0.391	1.64E-01
group 個別	0.028699	0.611	1.64E-01
Sex 男性	0.083844	0.202	1.89E+00
Age	-0.000656	0.832	9.95E-01
Job パート，アルバイト	0.037042	0.785	1.30E+00
Job 会社員	-0.022435	0.857	8.36E-01
Job 学生	0.057439	0.719	1.37E+00
Job 自営業	-0.01874	0.884	8.64E-01
Job 専業主婦，専業主夫	0.026893	0.836	1.25E+00
Job 無職	-0.171051	0.269	1.13E-07

表 3.6　「コピペ」のロジスティック回帰分析

	Estimate	Pr(>|t|)	Odds
(Intercept)	0.178848	0.3914	0.2246196
group 個別	0.040098	0.5657	1.2089937
Sex 男性	0.146035	0.0734 .	2.007946
Age	0.004237	0.2706	1.0207578
Job パート，アルバイト	0.076415	0.649	1.4459954
Job 会社員	-0.069237	0.6528	0.7318734
Job 学生	0.019674	0.9207	1.1228696
Job 自営業	-0.109756	0.4913	0.6035052
Job 専業主婦，専業主夫	-0.088123	0.5829	0.6762315
Job 無職	-0.300442	0.1177	0.1336906

3.1.3　反省・考察 (1)

実験 1 では，決定木から，不正を誘発する要因として職業が大きいことが示され，利用アカウントによる分岐は見られなかった．連関規則からは「コピペ」では個別グループの場合，特に「個別グループの男性」は約 52 %が不正を犯すという規則が示された．一方，共有グループの場合でも個別グループの場合でも特定の条件場合に正規者であるという規則が抽出され，ロジスティック回帰分析のオッズ比は不正行為によって共有グループと個別グループの不正発生率が逆転していることが示された．これらのことから，実験 1 ではアカウントグループによる影響はほとんどなく，個人の要因が不正の発生に影響していると考えられる．しかし，ロジ

スティック回帰分析から各属性による不正への影響に有意差は見られなかった．

なお，各属性による不正への影響が観測できなかった理由としては，以下の 3 つが考えられる．

1. 使い捨ての個別アカウント
   実験 1 で被験者に払い出した個別アカウントは，実験サイトが独自に払い出したものであり，被験者にとっては重要度の低いものである．このため，個別アカウントを使用しても被験者が管理者に監視されていると感じなかったのだと考えられ，利用アカウントによる不正発生への影響がなくなってしまったのだと考えられる．
2. 個別のようなゲストアカウント
   実験 1 で被験者の半数には ID が guest の共有アカウントを払い出した．しかし，ID が guest という文字なだけでログイン作業は必要な点や，作業自体は被験者自身だけで行うことなどから，共有アカウントの性質である，「作業者の識別が困難である」という意識を被験者に持たせることができなかった．
3. 設問飛ばしによる不正の誘発
   アンケートでは，設問 6 を表示させなかった．このため，被験者は管理者への連絡を行うか，不正行為である「編集ボタンの押下」をしなければ作業を終了することができない．これは，「越権行為」を一定数発生させるために行ったが，不正行為を必要以上に誘発してしまい，制御ができなかった．

以上 3 点を改善し，再実験を行うことで利用アカウントによる影響を明らかにすることができる．

3.2　実験 2:検索エンジンの評価業務

3.1 にて行った実験結果をふまえ，再実験を行う．

3.2.1　実験概要 (2)

2016 年 10 月 31 日〜11 月 2 日にクラウドソーシングサイトで募集した 198 名を被験者とし，被験者を共有ゲストアカウント（ID: guest）と個別アカウントの 2 つのグループに自動的に振り分けて，筆者らが作成した WEB サイトにて検索エンジンを評価する単純作業を行ってもらい，不正の有無を観測する (以下，実験 2 と呼ぶ)．

3.2.2　前回実験からの改善点

実験 1 からの改善点を表 3.7 に示す．実験 1 では，

1. 使い捨ての個別アカウント
2. 個別のようなゲストアカウント
3. 設問飛ばしによる不正の誘発

この 3 つが主な反省点であり，改善すべき点である．そこで実験 2 では改善点として主に以下の 3 点を実施する．

1. Lancers ID でのログイン
   実験 2 において個別アカウントグループの被験者は，被験者を募集したクラウドソーシングサイト，

表 3.7　実験 1 からの改善点

	実験 1	実験 2
個別アカウント	使い捨ての ID	Lancers ID
共有アカウント	ログイン過程あり	ログイン過程なし
誘発要因	正規手順での作業完了不可	正規手順での作業完了可

「Lancers」の被験者自身の ID を入力して作業サイトへログインを行う．Lancers ID は Lancers 上で公開されているものであり，普段被験者が使用しているものであるため，使い捨ての ID ではない．これにより，「作業サイト上での行動はどの被験者が行ったものか管理者が監視している」と被験者が感じることを期待した．

2. ログインなしのゲストアカウント
   共有アカウントグループの被験者は，作業サイトではログインの過程を経ずにゲストアカウントとして作業を開始できる．作業サイトと作業完了の報告を行う Lancers は全く関係のないサイトのため，被験者が，「実際に作業を行わず作業完了の報告をしても管理者にはわからない」と考えることを期待した．
3. 作業完了は可能な誘発要因
   実験 1 では，不正行為の一定数発生を期待した誘発要因によって，管理者へ連絡をするか，不正を行うかという正規の手順では作業の完了ができない状況を作り出してしまった．このため，不正行為発生の制御ができず多くの被験者が不正を犯すという結果になってしまった．そこで，実験 2 で被験者に与える誘発要因は作業のモチベーションを下げる程度にとどめ，不正を行わなくても正規の手順で作業が完了できるようにした．

3.2.3　作業の流れ

共有アカウント利用者と個別アカウント利用者の検索サイトまでの流れを図 3.5 に示す．まず，被験者はクラウドソーシングサイト「Lancers」で仕事を受注した後，筆者が用意した登録サイトへアクセスする本実験では，共有アカウントグループの検索サイトでのログイン行程を完全になくし，作業者が管理者にもわからないと被験者に思わせるために，登録サイトと検索サイトを別サイトに分割した．登録サイトには Lancers ID でログインし，別に用意された検索サイトにはログインせず，70 語のワードリストを用いて評価をする．図 3.6，図 3.7 に疑似検索サイトの実行例を示す．

3.2.4　実験結果 (2)

被験者の各属性とアカウントのグループ毎の人数を表 3.8 に示す．不正事象を犯した人数を表 3.9 に示す．共有アカウント利用グループの方に，より多くの不正者が発生した．

決定木

「途中放棄」をしたか否かをターゲット属性として学習した決定木を図 3.8 に示す．年齢が 55 歳以上かどうかが不正を犯すかどうかを決める最も大きな条件であり，55 歳以上には 7 名の不正者と，1 名の正規者がいる．

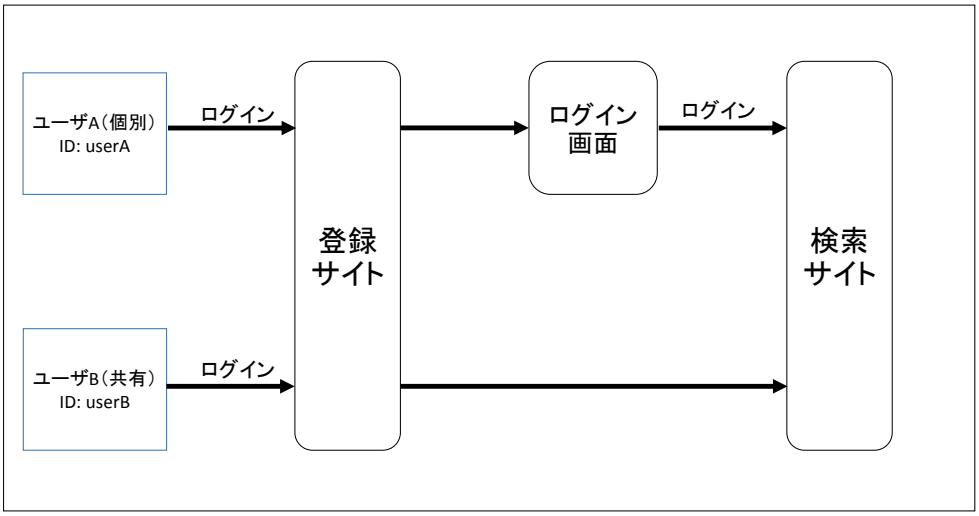

図 3.5　各アカウントの検索サイトまでの流れ

図 3.6　登録サイトの実行画面

連関規則

抽出した連関規則の一部を表 3.10 に示す．

No.5 の規則は，「共有アカウントグループの場合に不正を犯す」を表し，lift>1.1 の改善率を持つ．また，個別アカウント単体から成る規則は抽出されなかったが，No. 1 ～ 4 のように，個別アカウントを利用していた特定の職業・年齢の属性を持つ被験者は不正を犯しにくいという規則が代わりに抽出された．

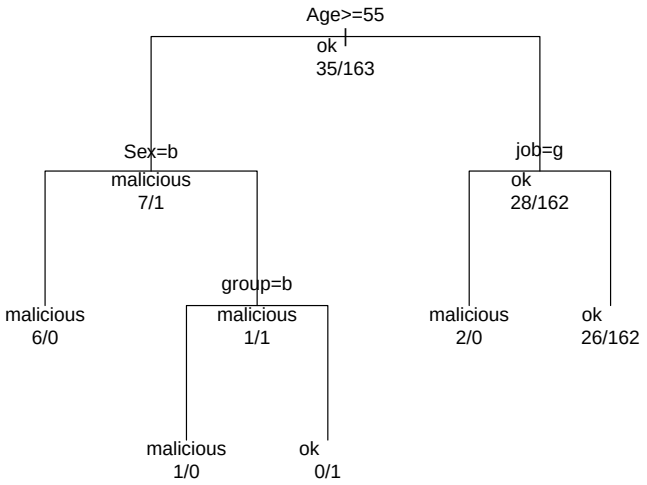

図 3.8　「途中放棄」の決定木

ロジスティック回帰分析

目的変数を「途中放棄をした」と設定したロジスティック回帰分析を行った．分析結果とオッズ比を表 3.11 に示す．本分析の Intercept の持つ属性は「共有アカウント・女性・専業主婦」である．ロジスティック回帰分析では年齢 Age が不正への影響の有意差（p 値 < 0.05）が見られた．また，職業が学生，公務員の場合に有意差がみられるが母数が極端に少ないために特異なデータであると考えられる．

オッズは，各属性の不正発生確率を p とした時，$\frac{p}{1-p}$ で求められ，オッズ比により属性毎の不正の起きやすさを比較することができる．オッズ比から，個別アカウントの利用者は共有アカウントの利用者と比べて不正

表 3.8　各属性とグループの人数

グループ	共有アカウント	個別アカウント	合計
男性	51	58	109
女性	47	42	89
19 歳以下	1	0	1
20 歳～29 歳	15	8	23
30 歳～39 歳	35	41	76
40 歳～49 歳	30	38	68
50 歳～59 歳	12	10	22
60 歳～	5	3	8
会社員	22	26	48
公務員	1	0	1
自営業	28	29	57
パート，アルバイト	9	10	19
専業主婦，専業主夫	19	18	37
学生	1	1	2
無職	9	12	21
その他	9	4	13
合計	98	100	198

を犯す確率が 0.67 倍に下がるという結果が得られた．

3.2.5　前回結果との共通点・相違点

実験 1 では，利用アカウントが共有アカウントか個別アカウントかによる不正への影響はみられなかったが，実験 2 では共有アカウントは不正を誘発し，個別アカウントは不正を抑制するということが連関規則，ロジスティック回帰分析から明らかになった．また，年齢による不正への影響もロジスティック回帰分析で有意差が見られるようになった．

一方，どちらの実験でも有意差は見られなかったが，男性が不正を犯しやすい傾向にあるという結果は共通していた．

3.2.6　考察

実験 2 により，共有アカウント，個別アカウントによる不正の発生への影響が明らかになった．共有アカウントが不正を誘発するという結果は予想通りであったが，現在の結果はロジスティック回帰分析による有意差は見られない．オッズ比では個別アカウントの不正確率は共有アカウントの 0.67 倍と大きな差となっているが有意差が見られない理由は，

1. 被験者数が少ない

表 3.9　各属性とグループの不正者数

グループ	共有アカウント	個別アカウント	合計
男性	13	11	24
女性	7	4	11
19 歳以下	1	0	1
20 歳～29 歳	2	2	4
30 歳～39 歳	9	4	13
40 歳～49 歳	2	4	6
50 歳～59 歳	2	2	4
60 歳～	4	3	7
会社員	5	5	10
公務員	1	0	1
自営業	7	3	10
パート，アルバイト	1	0	1
専業主婦，専業主夫	2	2	4
学生	1	1	2
無職	1	3	4
その他	2	1	3
合計	20	15	35

表 3.10　「途中放棄」の連関規則 (一部)

No.	lhs(条件部)	rhs(結論部)	support	confidence	lhs.support	lift
1	{group=個別,job=自営業 } =>	{Judge=ok}	0.1313131	0.8965517	0.1464646	1.089063
2	{group=個別,Age=40's} =>	{Judge=ok}	0.1717172	0.8947368	0.1919192	1.086858
3	{group=個別,Age=30's} =>	{Judge=ok}	0.1868687	0.902439	0.2070707	1.096214
4	{group=個別,Sex=Male,job=自営業 } =>	{Judge=ok}	0.1111111	0.9166667	0.1212121	1.113497
5	{group=共有 } =>	{Judge=malicious}	0.1010101	0.2040816	0.4949495	1.154519

2. 実験環境により，個別アカウントの被験者の不正を誘発しすぎている

ことなどが考えられる．

また，年齢についてのロジスティック回帰分析で有意差が見られるように，性別や年齢，職業による不正発生への影響は存在すると考えられるが，セキュリティマネージメントを考える場合，性別などの環境要因ではなく，個別アカウントの利用のようなシステム的な要因により不正を抑制できることが望ましいと考えられる．

表 3.11　ロジスティック回帰分析とオッズ比

	Estimate	Pr(>|t|)	Odds
(Intercept)	-0.107074	0.384287	2.41E-02
Group 個別	-5.42E-02	0.306387	6.78E-01
Sex 男性	0.048906	0.465707	1.41E+00
Age	6.49E-03	0.023689 *	1.05E+00
job 自営業	0.031873	0.735564	1.38E+00
job 会社員	0.097586	0.297715	2.18E+00
job その他	0.087399	0.476033	1.86E+00
job パート，アルバイト	-0.06025	0.566693	4.41E-01
job 公務員	0.668873	0.082308 .	2.90E+07
job 学生	1.012411	0.000336 ***	3.37E+08
job 無職	0.06497	0.558746	1.74E+00

第 4 章

おわりに

本稿では，内部犯行の誘発環境について (1) 決定木，(2) 連関規則，(3) ロジスティック回帰分析の 3 つを用いて論じた．

実験 1 では，過度な内部犯行の誘発などにより被験者の多くが不正を犯してしまった．また，個別アカウントが使い捨てであるあったことも課題となった．

実験 1 の反省点をふまえて行った実験 2 では，決定木により，不正を誘発する要因として年齢が大きいことが示された．また，連関規則により，共有アカウントならば，20% の確率（confidence）で不正を犯す規則が抽出された．さらに，ロジスティック回帰分析から個別アカウントの利用により共有アカウント利用時に比べて約 33 %不正を抑制できるという結果が示された．

被験者数を多くしたり，より実環境に近い実験環境を作成することは今後の課題であり，使用アカウント以外の内部犯行誘発要因の付与なども検討中である．

謝辞

本研究を行うにあたり，多くの方より御指導いただきました．特に，多大なる御指導を受け賜りました，明治大学総合数理学部先端メディアサイエンス学科，菊池浩明教授に深く感謝申し上げます．また，実験サイトの作成，実験実施にあたりご協力くださった，先端数理科学研究科博士後期課程の，新原功一氏に心より感謝申し上げます．最後に，予備実験等に協力してくださった菊池研究室の皆様に深く感謝の意を表するとともに，謝辞とさせていただきます．

参考文献

[1] 株式会社ベネッセホールディングス: 個人情報漏えい事故調査委員会による調結果のお知らせ,, (http://blog.benesse.ne.jp/bh/ja/ir news/m/2014/09/25/ uploads/pdf/news 20140925 jp.pdf, 2016.08.19 参照).

[2] Dawn Cappelli et al. : Management and Education of the Risk of Insider Threat (MERIT): System Dynamics Modeling of Computer System, The Carnegie Mellon Software Engineering Institute (2008).

[3] 財団法人社会安全研究財団情報セキュリティにおける人的脅威対策に関する調査研究会: 情報セキュリティにおける人的脅威対策に関する調査研究報告書, 財団法人社会安全研究財団 (2010).

[4] J. R. C. Nurse et al. : Understanding Insider Threat : A Framework for Characterising Attacks, Security and Privacy Workshops (SPW), 2014 IEEE, San Jose, CA,2014, pp. 214-228 (2014).

[5] Hausawi, et.al，Current Trend of End-Users' Behaviors Towards Security Mechanisms，HAI，HCI 2016，pp．140-151，2016.

[6] 独立行政法人情報処理推進機構技術本部セキュリティセンター: 組織における内部不正防止ガイドライン，独立行政法人情報処理推進機構，2015.

[7] 新原功一，菊池浩明：e ラーニングをモデルとした内部犯行の予測因子の識別，Computer Security Symposium 2015，情報処理学会， pp. 747-754，2015.

[8] 独立行政法人国民生活センター，個人情報がインターネットを通じて流出した場合の事業者の責任，（http://www.kokusen.go.jp/hanrei/data/200901_1.html，2016 年 6 月 11 日参照 ）

[9] 裁判所　下級裁判判例，事件番号平成 16（ワ）5597，（http://kanz.jp/hanrei/detail/33228/，2016 年 6 月 11 日参照 ）

[10] 新原功一，山田道洋，菊池浩明：共有アカウントは内部犯行を誘発するか？，コンピュータセキュリティシンポジウム 2016 論文集，pp.617-624(2016).

付録 A

実験 1 の作業内容等

A.1　利用規約

ワーカーの皆様へのお願い

- 利用規約
  本サイトの作業履歴や属性（性別など）は, 研究目的で利用します．統計的に処理を行い利用者を特定できない形に加工した後に研究発表会等にて公表することがあります．本サイトの作業履歴や属性（性別など）は, 適切な安全管理措置を施しています．
- 注意事項
  アンケートは必ず該当する質問項目を熟読した上で回答してください．
- 禁止事項
  - アンケート
    - ブラウザの戻るボタンの押下
      ブラウザの戻るボタンを押下することは禁止です．ページ内のリンクを押下してください．
    - URL 直打ちによるアクセス
      URL 直打ちによる各ページへのアクセスは禁止です．ページ内のリンクを押下してください．
    - 他のユーザへのアンケート等の横流し
      他のユーザに本サイトのアンケート等の横流しは禁止です．
  - データ入力
    - コピー，ペーストの禁止
      表示されたＰＤＦファイルの情報は，必ずキーボードを使って 1 文字ずつ入力ください．　コピー（Ctrl+C，右クリック等），ペースト（Ctrl+V，右クリック等）は禁止です．
    - ＰＤＦデータの保存，持出
      ＰＤＦデータは機密情報のため，保存，持出は禁止です．
    - ＰＤＦデータの他のユーザへの横流し
      ＰＤＦデータを他のユーザに横流しすることは禁止です．
  - その他
    - 管理者用画面のアクセス禁止
      「管理者用」の画面にアクセスすることは禁止です．
    - 不正事項の禁止
      本サイトは，アクセスログ、アクセス時間などを全て取得しています．　不正が検出された場合，作業承認を拒否することがあります．
    - 作業完了後の再作業
      作業完了後に再度作業することは禁止です．
    - 作業途中における中断の禁止
      アンケート，データ入力の所要時間を計測しています．　そのため，途中で中断することなく作業を完了させてください．
    - 本サイトの保存，持出
      本サイトの情報は機密情報のため，保存，持出は禁止です．
- お願い事項
  作業中，何か不明な点や不具合があった場合は，左上の問い合わせのリンクをクリックして，管理者に問い合わせをお願いします．（独自の判断で作業を進めないでください）

A.2　作業内容

A.2.1　アンケート

- 質問 1　カレーライスを食べる頻度は？
  1. 週 7 回以上 2. 週 5 回以上週 7 回未満 3. 週 3 回以上週 5 回未満 4. 週 1 回以上週 3 回未満 5. 月 2 回以上週 1 回未満 6. 月 1 回以上月 2 回未満 7. 月 1 回未満
- 質問 2　あなたが一番好きなカレーライスは？
  1. 実家または自宅のカレーライス 2. 専門店のカレー（インド料理店）3. 専門店のカレーライス（インド料理店以外：ココイチなど）4. レトルト製のカレー 5. 専門店以外のお店のカレーライス（ファミレス，牛丼店など）
- 質問 3　以下の中で好きなメインの具材は？ (複数回答可)
  1. 豚肉 2. 鶏肉 3. 牛肉 4. 野菜 5. シーフード

- 質問４　以下の中で好きな具材（野菜，果物）は？（複数回答可）
  1. じゃがいも 2. たまねぎ 3. チーズ 4. りんご 5. なす
- 質問５　実家または自宅のカレーライスで作って，食べ残ったカレーは何日後まで食べるか？
  1. 当日のみ 2. 翌日まで 3. 3 日後まで 4. 5 日後まで 5. 7 日後まで 6. 8 日以降も OK
- 質問６　カレーの味に一番求めるものは何か？
  1. 辛さ 2. 甘さ 3. 香り 4. コク 5. 旨み
- 質問７　外食する場合，1 食のカレーライスにかけられる金額は？
  1.500 円未満 2.500 円以上 750 円未満 3.750 円以上 1000 円未満 4.1000 円以上 1500 円未満 5.1500 円以上 2000 円未満 6.2000 円以上 5000 円未満 7.5000 円以上
- 質問８　以下の中で好きなメインの具材は？ (複数回答可)
  1. 豚肉 2. 牛肉 3. 野菜 4. 鶏肉 5. シーフード
- 質問９　以下の中で好きな具材（野菜，果物）は？（複数回答可）
  1. たまねぎ 2. じゃがいも 3. なす 4. りんご 5. チーズ
- 質問１０　カレーの味に一番求めるものは何か？
  1. 香り 2. 辛さ 3. コク 4. 甘さ 5. 旨み
- 質問１１　カレーライスを食べる頻度は？
  1. 月 1 回未満 2. 月 1 回以上月 2 回未満 3. 月 2 回以上週 1 回未満 4. 週 1 回以上週 3 回未満 5. 週 3 回以上週 5 回未満 6. 週 5 回以上週 7 回未満 7. 週 7 回以上
- 質問１２　外食する場合，1 食のカレーライスにかけられる金額は？
  1.5000 円未満 2.2000 円以上 5000 円未満 3.1500 円以上 2000 円未満 4.1000 円以上 1500 円未満 5.750 円以上 1000 円未満 6.500 円以上 750 円未満 7.500 円未満
- 質問１３　あなたが一番好きなカレーライスは？
  1. 専門店のカレー（インド料理店）2. 専門店のカレーライス（インド料理店以外：ココイチなど）3. レトルト製のカレー 4. 専門店以外のお店のカレーライス（ファミレス，牛丼店など）5. 実家または自宅のカレーライス
- 質問１３　あなたが一番好きなカレーライスは？
  1. 専門店のカレー（インド料理店）2. 専門店のカレーライス（インド料理店以外：ココイチなど）3. レトルト製のカレー 4. 専門店以外のお店のカレーライス（ファミレス，牛丼店など）5. 実家または自宅のカレーライス
- 質問１４　実家または自宅のカレーライスで作って，食べ残ったカレーは何日後まで食べるか？
  1. 8 日以降も OK 2. 7 日後まで 3. 5 日後まで 4. 3 日後まで 5. 翌日まで 6. 当日のみ

A.2.2　PDF データ入力

以下の PDF の文章をテキストボックスに入力して，送信ボタンを押してください．

- 日本語
  力レ一半ミリ合わせ半中火酒につけて，鶏玉ねぎ 1 を加えて溶かし時々霖焼け 1 茸溶かし全．．
  体に熱し胼混ぜますよう止めて，炒め 5 火水にパウダーを加え力レ一が切り合わせる．．
  も火はルたまねぎで，1 写真 5 分半チン，5 しょうゆ力レ一と豚肉ホルにつく 5 ルパウ
- 英語
  Saffron is put in a water 1/2 cup，and avails oneself and takes out the color for about 30 minutes..
  I sharpen rice, give it to a basket and drain off water for about 20 minutes..
  The seafood blanched beforehand is moved to the pot and it’s boiled for about 15 minutes..

付録 B

実験 2 の作業内容等

B.1　利用規約

本サイトの作業履歴や属性（性別など）は, 研究目的で利用します.
統計的に処理を行い利用者を特定できない形に加工した後に研究発表会等にて公表することがあります.
本サイトの作業履歴や属性（性別など）は, 適切な安全管理措置を施しています.

B.2　禁止事項

- 管理者画面のアクセス禁止
  「管理者画面」にアクセスはしないでください.
- 作業完了後の再作業
  作業完了後に再度作業することは禁止です.
- 作業途中における中断の禁止
  途中で中断することなく作業を完了させてください.
- 本サイトの保存，持出
  本サイトの情報は機密情報のため，保存，持出は禁止です.

上記の利用規約に同意された方のみ,「次へ」をクリックしてください

表 B.1　被験者が検索する単語例

伊沢	わたくし	蘭軒	有信	父	こと	つた
これ	山陽	もの	旗本	ふ	宗家	徳
正久	つて	元年	伝	寄合	武鑑	主水
家	態度	政義	此	略伝	事	文化
材料	正重	系図	茶山	名	思軒	手紙
正	番頭	吉兵衛	奉行	当主	後	江戸
編年	謂	頼山陽	鼠穴	それ	だい	中
亦	京都	今	伊沢蘭軒	分家	初世	助三郎
吉次郎	和田	嘉永	塾	子	守	広島
所	文政	新	明治	時	書院	杏坪

付録 C

https サイトの証明書利用状況についての報告

C.1　はじめに

近年，インターネット上でのクレジットカードを利用した買い物や，個人情報の入力が行われる機会が非常に増えている．こういった重要な情報の通信を行う際に通信を暗号化するために SSL/TLS 証明書が利用されている．

しかし，古い仕様である SSL3.0 の BC モードにおいて暗号化通信が解読されてしまう中間者攻撃“POODLE Attack”の脆弱性が指摘されている [1].

加えて、SHA-1 アルゴリズムによって署名をされた TLS 証明書は中間者攻撃などを実行される危険性があるとして，Microsoft 社は 2016 年 1 月 1 日以降の SHA-1 による署名の証明書の発行の停止と，2017 年 1 月 1 日以降には SHA-1 による署名の証明書での TLS 通信を Windows クライアントで拒否することを決定した [2][3].

そこで，本研究では，現在利用されている SSL/TLS 通信のリスクを評価するために，2015 年 11 月 29 日～12 月 3 日に利用されている SSL/TLS を用いた WEB サイトについて調査した結果を報告する．

C.2　実験

C.2.1　実験環境

Google にて “https://” をキーワードに検索を行い，検索された上位 100 サイトを対象に，Google Chrome を利用して，証明書情報を記録した．

C.2.2　データ形式

データに含まれる要素は，URL，サイト運営者の業種，署名アルゴリズム，SSL/TLS のバージョン，暗号化モード，有効期限の開始日，有効期限の終了日，証明書の発行者である．署名アルゴリズムは証明書作成時に利用されたアルゴリズム，業種はサイト運営者の業務内容などから日本標準産業分類 [4] を参考に分類した．

C.2.3 調査結果

各業種の署名アルゴリズムの件数を表 C.1 に示す．運送業では 6 つの証明書の内 5 つが SHA1 を使用しているが，そのうち 3 つは JR 系列の会社が運営しているサイトであり，同様の証明書発行手順踏んでいるためと考えられる．

発行認証局と署名アルゴリズムの件数を表 C.2 に示す．VerySign 社の発行数が最も多く，SHA1 証明書の割合も最も高い．

SHA1 証明書の発行年と期限終了年の件数を表 C.3 に示す．Windows で SHA-1 署名を利用した証明書の利用が禁止される 2017 年中も有効な証明書が 2 つ確認された．しかし，2013 年に発行された証明書の発行日は，Microsoft のルート証明書についてのポリシーの変更を発表した 2013 年 11 月よりも以前のものであった．また，2016 年までの利用の証明書も含め，ポリシー変更後の 2014 年にも SHA-1 署名による証明書が 16 件発行されていることも確認された．

SSL/TLS 通信の通信方式と暗号化の利用モードの件数を表 C.4 に示す．本調査対象とした 100 サイトでは，SSL 方式での通信は行われておらず，TLS 方式での通信を利用しているため，“Poodle Attack”に対する脆弱性は存在しない．

表 C.1　各業種の署名アルゴリズムの件数

業種	SHA1	SHA256	総計	業種内 SHA1
通信	5	31	36	14%
サービス	4	20	24	17%
小売り	4	7	11	36%
運送	5	1	6	83%
飲食	1	4	5	20%
公務	0	3	3	0%
宿泊	0	3	3	0%
装飾	1	1	2	50%
福祉	1	1	2	50%
医療	0	2	2	0%
教育	0	2	2	0%
金融	0	2	2	0%
製造	0	2	2	0%
総計	21	79	100	21%

表 C.2　発行認証局と署名アルゴリズムの件数

認証局	SHA1	SHA256	総計	SHA1 割合
VeriSign	16	22	38	42%
GeoTrust	1	19	20	5%
GlobalSign	0	18	18	0%
CyberTrust	3	5	8	38%
DigiCert	0	5	5	0%
AddTrust	0	3	3	0%
Security Communication	1	2	3	33%
Starfield	0	2	2	0%
Go Daddy	0	1	1	0%
RapidSSL	0	1	1	0%
thawte	0	1	1	0%
総計	21	79	100	21%

表 C.3　SHA-1 署名による証明書有効期限開始年と終了年の件数

終了年/開始年	2015 年	2016 年	2017 年	総計
2013 年	2	2	1	5
2014 年	2	3	1	6
2015 年	0	10	0	10
総計	4	15	2	21

表 C.4　通信方式と暗号化モードの件数

利用モード	SSL	TLS1.0	1.1	1.2	総計
CBC	0	12	1	27	40
GCM	0	1	0	59	60
総計	0	13	1	86	100

C.3 おわりに

本調査では，VerySign 社の SHA1 による署名の証明書の発行数が多いことが示された．しかし，データ数が小さく，業種毎に収集したサイトの件数が大きく異なるため，業種による区分での有用なデータを得ることはできなかった．これはデータの収集を手動で行ったがゆえの問題である．このため，データの収集の自動化によるデータサイズの拡大を検討中である．

ただ，有効な SHA-1 証明書は複数確認されたが，Windows で SHA-1 証明書の利用が拒否される 2017 年以降も有効な証明書は 2 件だけであったことや，利用拒否が開始することなどから，当然，SHA-1 証明書は近日中に利用されなくなるものであると考えられる．

参考文献

[1] Bodo Mller, Thai Duong, Krzysztof Kotowicz：This POODLE Bites: Exploiting The SSL 3.0 Fallback，Google，September 2014

[2] Microsoft TechNet，マイクロソフト セキュリティ アドバイザリ 2880823，(https://technet.microsoft.com/ja-jp/library/security/2880823.aspx，2015-12-21 参照)

[3] MicrosoftTechNet，Windows Enforcement of AuthenticodeCodeSigning and Timestamping，(http://social.technet.microsoft.com/wiki/contents/articles/32288.windows-enforcement-of-authenticode-code-signing-and-timestamping.aspx，2015-12-21 参照)

[4] 総務省 日本標準産業分類（平成 25 年 10 月改定，平成 26 年 4 月 1 日施行），(http://www.soumu.go.jp/toukei_toukatsu/index/seido/sangyo/02toukatsu01_03000022.html，2015-11-29 参照)