**SCIS 2017** 2017 Symposium on
Cryptography and Information Security
Naha, Japan, Jan. 24 - 27, 2017
The Institute of Electronics,
Information and Communication Engineers

# 共有アカウントは内部不正を誘発するか？(2)
# Does sharing credentials cause insider threat ? (2)

新原 功一 * 山田 道洋 † 菊池 浩明 †
Niihara Koichi Yamada Michihiro Kikuchi Hiroaki

**あらまし** 昨今，大規模な情報漏えい事案等が発生したことが契機となり，内部不正のリスクを低減することは急務の課題となっている．内部不正はいくつかの誘発要因の影響を受けて発生することがあるが，特にシステムにログインするためのアカウントを共有した場合，利用者を識別することが出来ないため内部不正を誘発することが多いといわれている．そこで本研究は，アカウントを共有することで内部不正を誘発する大きさがどれくらいになるかを検証する．ランサーズ社のクラウドソーシングサービスにより集めた 198 名の被験者は筆者らが構築した検索エンジン評価用の疑似環境にアクセスし，利用した感想を報告する．被験者にはランダムに共通アカウント，個別アカウントのいずれかを付与する．作業完了条件は個人毎に与えられた 70 語の検索ワードのうち，50 語以上の検索を完了することとし，当該作業を途中で放棄（検索したワードが 50 語未満）した場合や利用規則を違反した行為を不正事象とする．不正事象は利用者毎に測定した．測定結果を統計解析の手法を用いて分析し，共有アカウントと不正事象の相関関係を明らかにする．

**キーワード** 内部不正，共有アカウント，情報漏えい

## 1 はじめに

大手教育会社の情報漏えい事案等を契機に内部不正に対する脅威への対策は急務となっている [1]．内部不正を防ぐための管理策は労働環境や待遇の改善，従業員に対する教育など様々なものが存在する [2]．中でも，システムにログインするためのアカウントを利用者へ個別に払い出すと内部不正が抑制できる生じにくくなるといわれている [3]．Hausawi らがセキュリティ専門家に対して行ったインタビューによると，エンドユーザが行うセキュリティに対する最も否定的な振舞いは認証情報の共有 (Sharing credential) であった [4]．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ひとつの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．

本研究は共有アカウントが内部不正を誘発する度合いはどれくらいなのかを検証する．組織は，内部不正の発生に抑制効果が高い手法を識別できれば，それらの対策に費用を集中することができ，効果的にリスクを低減することができる．

これらの実験は実在する企業等の従業員を対象に実施することが望ましいが，実験結果を第三者に提供することは，当該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．さらに内部不正の発生頻度は低く，その過程を詳細に観察することは困難である．

この問題に対して，筆者らは疑似環境において，被験者が行った作業で生じる不正事象の発生数を観測する実験（以下，前回実験とする）を行った [5]．被験者には，4 つのグループをランダムに割り当てた．被験者に払い出すアカウントは共通 ID と個別 ID のいずれかとし，さらに常にアカウント名が画面に表示されるグループと表示されないグループに分類した．共有 ID を利用するユーザがより多くの不正事象を発生させることを期待したが，いくつかの不正事象は個別 ID を利用するユーザの方が多くの不正を犯した．個別 ID は，疑似環境が独自に払い出したものであり被験者はあまり警戒せずに作業を行ったことが，原因の一つと考えている．そこで，前回実験における反省点を踏まえ，筆者らは新たな実験（以下，本実験とする）を行った．本稿は，本実験の結果を報告する．前回実験と本実験の差異を表 1 に示す．

本実験は，前回実験と同様に不正事象の発生数を測定する．大きな違いは個別アカウントを被験者の LancersID

* 明治大学大学院先端数理科学研究科, 〒 164-8525 東京都中野区中野 4-21-1, Meiji University Graduate School of Advanced Mathematical Sciences, 4-21-1 Nakano, Nakano-ku, Tokyo, Japan 164-8525
† 明治大学総合数理学部, 〒 164-8525 東京都中野区中野 4-21-1, Meiji University Undergraduate School of Interdisciplinary Mathematical Sciences, 4-21-1 Nakano, Nakano-ku, Tokyo, Japan 164-8525

表 1: 前回実験 [5] と本実験の差異

| 項目 | 前回実験 [5] | 本実験 |
|---|---|---|
| 実験環境 | 疑似環境 (1 サイト) | 疑似環境 (2 サイト) |
| 作業内容 | アンケート，データ入力 | WEB サイトの評価 |
| クラウドソーシングサービス | Crowdworks | Lancers |
| 個別アカウントのユーザ名 | 独自 ID(使い捨て) | 被験者の LancersID |
| 共有 ID が内部犯行を誘発する効果 | 有意な差はなかった | 30 代に有意な差があった |

にした点である．前回実験は個別 ID が疑似環境で払い出された使い捨て ID であったため，被験者は監視がされていることとの因果関係を強く感じることがなく，内部不正を抑制する効果が限定的であったと推察する．

以下，2 章では，関連研究の調査について述べる．3 章では提案方式，4 章で実験結果，評価を記す．5 章で考察を与え，最後に 6 章でまとめる．

## 2 関連研究

### 2.1 内部不正の誘発要因

いくつかの研究では，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，傾向をモデル化している．Cappli らは，MERIT[1] を提案している [6]．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 [7]．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 [8]．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．

### 2.2 共通アカウントの利用と内部不正

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 [4]．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 (17%) 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ため心理的に重要情報を持出しやすい環境となると指摘している [3](p.31)．

[1] Management and Education of the Risk of Insider Threat

## 3 提案方式

### 3.1 仮説

本研究は，内部不正を誘発する要因として，次の 2 つの仮説を立てる．

仮説 $H_1$(共有 ID)：共有アカウント (例:guest アカウント) を利用していると内部不正を犯す．

仮説 $H_2$(ID 未表示)：作業中に常時アカウント名が表示[2] されていないと内部不正を犯す．

### 3.2 困難性

実環境では，内部不正の発生確率は低く，たとえ発生した場合でも観測が難しい．仮に観測ができた場合でも，それらの開示は企業内のセキュリティポリシに抵触する可能性がある．また，疑似環境において共有 ID と内部不正の関係を識別するためには以下の困難性が存在する．

1. 不正事象を誘発する要因の制御

   被験者が報酬を受け取ってタスクを遂行する際，数多くの内部不正を観測することは期待できない．一方，報酬を支払わない場合，被験者を収集することは容易ではない．そのため，優良な被験者に対して不正事象を誘発するための仕掛けが必要となる．

2. 共有アカウントを利用したユーザの識別

   被験者が共有アカウントを利用する場合，アカウント毎の操作履歴で被験者を識別することは出来ない．アカウント以外の方法で，被験者を一意に識別することは自明でない．

3. 個別アカウントと共有アカウントの差別化

   被験者にとっては，疑似環境で独自に払い出した個別アカウントは，一度しか使わないもの，言い換えれば “使い捨て ID” である．“使い捨て ID” はマイナンバーや SNS のアカウントのように長期間利用するものと比べて，被験者にとっての価値は低いと推察する．そのため，“使い捨て ID” を利用

[2] WEB サイトの各ページの上端に常時ユーザ名が表示されている状態

する被験者は，監視がされていると強く感じることもなく，今後の社会活動にも支障をきたす可能性が少ないため，内部不正を抑制する効果は薄いと考える．

### 3.3 本研究の新規性

本研究の新規性は，3.2 節の困難性を次の様に改善したことである．

1. 不正事象を誘発する要因の制御
   (a) 被験者にストレスを与えることで，モチベーションを低下させる (3.8.1)(3.8.2)
   (b) 被験者が真面目にやらなくても記録が残らないようにみせる (3.8.3)
   (c) 被験者は途中で作業を終わらせてもよいようにみせる (3.8.4)
2. 共有アカウントを利用したユーザの識別
   (a) 被験者が疑義作業で使うデータは被験者毎に一意に与える．本実験で被験者が入力したデータはすべて疑似環境に記録し，誰に払い出したデータであるかを確認することで被験者を識別する (3.9)
3. 個別アカウントと共有アカウントの差別化
   (a) 被験者が日常的に利用するアカウントを本実験の個別アカウントとして活用する (3.10)

### 3.4 目的

本研究は，共有アカウントの利用やアカウントの未表示が不正事象の発生に与える影響を明らかにすることを目的とする．

### 3.5 実験概要

筆者らは，組織の業務環境を再現した擬似作業用 WEB サイト（以下，本サイト）を構築した．本サイトは，検索ワード案内サイト（以下，案内サイト）と検索エンジン評価サイト（評価サイト）で構成する．仮説を検証するため，被験者は 4 つのグループに分類した．グループ毎に異なる刺激を与えることで不正事象の発生数にどれだけの差があるのかを観測した．

表 2 に被験者グループと仮説との関係を示す．

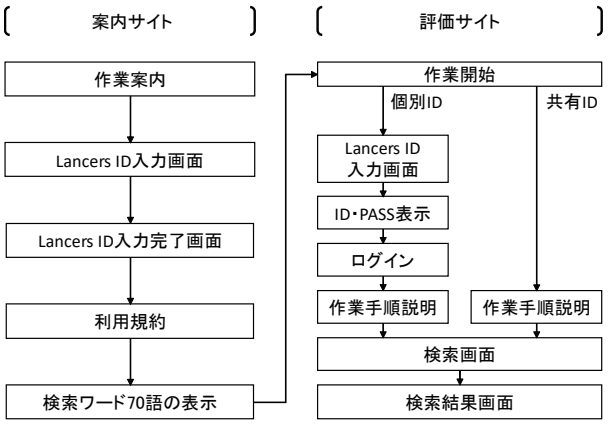


図 1: 画面遷移図

#### 3.5.1 被験者

本実験の母集団は，国内におけるすべての雇用者とする．標本はランサーズ社のクラウドソーシングサービスに登録したユーザのうち作業を完了した 198 名である．ユーザは本実験（タスク）を完了した順番で抽出した．被験者の質を確保するため，募集要件はランサーズ社で本人確認書類の提出が確認されていることとした．無作為抽出は実施していないが，クラウドソーシングサービスには様々なユーザが登録されており，多様な属性を持ったユーザの代表を抽出できると期待した．

#### 3.5.2 作業の流れ

まず，被験者はランサーズ社から本実験の作業を受託する．次に案内サイトにアクセスして，ランサーズ社のクラウドソーシングサービスにおける被験者の個別アカウントである LancersID を入力する．本サイトは被験者が作業を開始する前に利用規約を表示した．被験者は同意ボタンを押下しないと作業が開始できないようにした．利用規約を確認後，70 語の検索ワードを確認する．

その後，評価サイトにアクセスする．評価サイトは，個別 ID，共有 ID のいずれかを被験者に付与する．個別 ID が付与された被験者は，再び LancersID を入力して評価サイトの ID，PASS を確認し，評価サイトにログインする．その後，作業手順説明を確認する．一方，共有 ID が払い出された被験者は，ログイン等の作業は要求されず，作業手順説明を確認する．検索画面では案内サイトで表示された検索ワードを入力し，検索結果を確認する．図 1 は本サイトの画面遷移図である．

作業完了後，被験者はランサーズ社に完了報告を行う．筆者らは完了報告に基づいて利用状況を本サイトのアクセスログから確認し，問題がなければ作業を承認する．作業承認後，ランサーズ社は筆者らが事前に支払をしていた費用の一部を被験者に支払う．

表 2: 仮説とグループの関係

| グループ | $H_1$(共有 ID) | $H_2$(ID 未表示) | $N$ |
|---|---|---|---|
| $A$ | 共有 | 未表示 | 52 |
| $B$ | 個別 | | 52 |
| $C$ | 共有 | 表示 | 46 |
| $D$ | 個別 | | 48 |

表 3: 検索ワードの例

| | | | | | |
|---|---|---|---|---|---|
| ユーザ $A$ | 地所 | 坂田 | 基 | 墓地 | 多摩川 … |
| ユーザ $B$ | 少し | 岡 | 常善 | 延享 | 宸翰 … |
| ユーザ $C$ | 時信 | 晴雪 | 更迭 | 書 | 最後 … |

#### 3.5.3 仮説とグループの関係

評価サイトでは，以下の仕組みを構築することで仮説を検証した．まず，仮説 $H_1$(共有 ID) を検証するため，評価サイトではグループ $A$ と $C$ の被験者には共有アカウントを払い出した．アカウント名は “Guest” である．なお，本サイトは “Guest” アカウントを払い出した場合でも被験者毎の行動を識別ができるようにした．グループ $B$ と $D$ の被験者には個別アカウントを払い出した．詳細は 3.9 節を参照．仮説 $H_2$(ID 未表示) を検証するため，グループ $A$ と $B$ の被験者はアカウント名を一切表示させなかった．一方，グループ $C$ と $D$ の被験者には画面の上端に利用中のアカウント名を常時表示させた．

### 3.6 タスクの定義

(1) 検索エンジンの評価

被験者は案内サイトでユーザ毎に一意の 70 語の検索ワードを与えられる．評価サイトで 50 語以上を検索せよ (検索ワードの例は表 3 を参照)．被験者が入力した検索ワードの検索結果は，Google 社の検索 API を利用して表示する．検索エンジンの検索に関する作業は，クラウドソーシングサービスでは一般的な作業であり作業自体も簡単なことから，様々な属性の被験者が集まることを期待した．

(2) アンケートの回答

被験者は Lancers の作業完了報告画面で，評価サイトを利用した感想や被験者自身の属性などのアンケートに回答せよ．感想は，被験者の作業を完了させるために求める．被験者の属性 (年代，性別，職業) は，不正事象の発生が属性毎に差があるかを識別するために確認する．

表 4: 不正事象と検知方法の関係

| 不正事象 | 検知方法 |
|---|---|
| 途中放棄 | 回答内容の分析 |
| でたらめ | 回答内容の分析 |
| 越権行為 | php によるログ取得 |

### 3.7 不正事象

#### 3.7.1 不正事象の定義

本実験では 3 種類の不正事象を定義する．

(1) 途中放棄

評価サイトで検索したキーワードが 50 語未満である．

(2) でたらめ

評価サイトで検索したキーワードが，案内サイトで提示した文字列と異なる場合やキーワード自体が未入力である．

(3) 越権行為

評価サイトにおける管理者画面のリンクを押下した．これは，利用規約にて定めた禁止事項「管理者画面にアクセスすること」に該当する．

#### 3.7.2 検知方法

不正事象は以下のようにして検知する．不正事象と検知方法の関係を表 4 に示す．

1. 被験者の回答内容を分析して判定

   検索した全ての文字列と案内サイトが与えた検索ワードを比較する．

2. php によるアクセスログの取得

   管理者画面へのアクセスは，php を利用してログをデータベースに出力する．

### 3.8 不正事象を誘発する要因の制御

被験者に不正事象を誘発させるため，本サイトに実装した仕組みを以下に記す．

#### 3.8.1 応答時間の遅延

評価サイトの検索処理は Google 社の検索 API を利用しているため，本来の処理速度は 1 秒未満であるが，javascript によって被験者の検索回数をカウントし，回数に応じて人工的に遅延を生じさせる．遅延時間により，被験者のモチベーションを低下させ，被験者がより多くの不正事象を発生させることを期待した．被験者の検索回数を $s$ とすると，表 5 の遅延時間を加える．

表 5: 被験者の検索回数 $s$ と遅延時間，貼付制限の関係

| 検索回数 $s$ | 遅延時間 (秒) | 貼付制限 |
|---|---|---|
| 1〜5 | 0 | |
| 5〜13 | 1 | |
| 13〜19 | 2 | |
| 19〜23 | 3 | |
| 23〜31 | 4 | |
| 31〜33 | 20 | |
| 33〜37 | 2 | |
| 37〜41 | 9 | |
| 41〜43 | 20 | ○ |
| 43〜45 | 5 | ○ |
| 45〜47 | 9 | ○ |
| 47〜 | 5 | ○ |

#### 3.8.2 貼付制限

案内サイトが表示した検索ワードの中には読み方が比較的難しい単語[3]を含めている．読み方が分からない場合，多くの被験者は検索ワードを一旦コピーして，評価サイトに貼り付ける．そこで評価サイトでは，javascript によって被験者の検索回数をカウントし，41 回目以降の検索では，ブラウザ上での貼付行為を javascript で無効化した．作業に対する難易度を上げることで，被験者がより多くの不正事象を発生させる．

被験者の検索回数 $s$ と貼付制限の関係を表 5 に示す．

#### 3.8.3 検索回数と作業完了の関係

評価サイトの検索画面や検索結果画面では，被験者が検索した回数等を表示しない．また，被験者はたとえ 50 回以上を検索しなくともランサーズ社に作業完了を報告できる．作業完了を被験者の自己申告とすることで，多くの被験者が不正事象を発生させる．

#### 3.8.4 ログイン認証の未実施

被験者は案内サイトで LancersID を入力したが，評価サイトでは共有 ID を利用する被験者に対してログイン等の認証を不要とする．案内サイトと評価サイトは別のサイトであるような印象を与えるため，ドメイン名やサイトの背景色，フォントなどを異なるものとすることで，共有 ID を利用する被験者に対して自らの操作ログが記録されないと印象付ける．

### 3.9 共有アカウントを利用したユーザの識別

本実験では，被験者にユーザ毎に一意の検索ワードを与える．表 3 は検索ワードの例である．検索ワードの掲載数は 70 語である．検索された検索ワードを全て記録して，共有アカウントを利用したユーザであっても，誰がアクセスしたのかを識別することが出来る．

### 3.10 個別アカウントと共有アカウントの差別化

評価サイトが払い出す個別 ID は，被験者が入力した LancersID とする．パスワードは新たに乱数で生成し，被験者のプライバシー情報を取得しないようにする．

LancerID は，被験者がクラウドソーシングサービスで報酬を得るためにタスクの作業を行う際に必要なアカウントである．被験者は，Lancers での作業が拒否された場合，作業承認率が低下してしまい，受託できる作業が制限されてしまう．LancerID を利用して不正な作業を行うとクラウドソーシングサービスにおける自らの立場が悪化するため，使い捨て ID を使う場合と比べて，真面目に作業を行うことを期待する．

## 4 実験結果

### 4.1 ユーザ数

被験者は 198 名である．被験者は，ランサーズの作業完了報告時に自らの属性を回答した．表 6 は属性別のユーザ数である．グループ $C$ は，グループ $A$，$B$ と比べて 6 名少ない．グループは，評価サイトが被験者のアクセス順に割当てており，乱数による差ではない．ユーザが作業自体を途中で止め，作業完了報告も行わなかったと想定する．

### 4.2 検索回数と所要時間

1 番目の検索から $i$ 番目の検索までの所要時間を $T_i$[s] とする．被験者の検索回数 $s$ と所要時間 $T_i$ の関係は，概ね次の 4 つのパターンに分類された．

1. 途中で作業を止めた（赤：不正事象）
2. 応答時間の遅延や貼付制限が出現したタイミングで止めた（緑：不正事象）
3. 50 語を超過した時点で検索を止めた（青）
4. 70 語近くまで検索を続けた（水色）

これらの関係を図 2 に示す．被験者の検索回数 $s$ が $s < 50$ の場合，不正事象 (1) 途中放棄に該当するため，図 2 のうち赤線と緑線が不正事象である．

### 4.3 検索回数 (グループ毎)

図 3 はグループ毎における被験者の検索回数 $s$ の累積相対頻度 $Cu(s)$ である．例えば，グループ A の $s = 50$ は，$Cu(s < 50)$ が 0.21 であり，グループ A の被験者 52 名中の 21%が 50 回未満で入力を完了したことを表す．

[3] 例：宸翰 (表 3 のユーザ B の 5 番目)

表 6: ユーザ数 ($A$:共有/ID 未表示, $B$:個別/ID 未表示, $C$:共有/ID 表示, $D$:個別/ID 表示)

| グループ | $A$ | $B$ | $C$ | $D$ | Total |
|---|---|---|---|---|---|
| 男性 | 28 | 30 | 23 | 28 | 109 |
| 女性 | 24 | 22 | 23 | 20 | 89 |
| 19 歳以下 | 0 | 0 | 1 | 0 | 1 |
| 20 歳〜29 歳 | 8 | 2 | 7 | 6 | 23 |
| 30 歳〜39 歳 | 18 | 19 | 17 | 22 | 76 |
| 40 歳〜49 歳 | 16 | 24 | 14 | 14 | 68 |
| 50 歳〜59 歳 | 6 | 5 | 6 | 5 | 22 |
| 60 歳〜 | 4 | 2 | 1 | 1 | 8 |
| 会社員 | 16 | 17 | 6 | 9 | 48 |
| 公務員 | 1 | 0 | 0 | 0 | 1 |
| 自営業 | 13 | 13 | 15 | 16 | 57 |
| パート，アルバイト | 7 | 5 | 2 | 5 | 19 |
| 専業主婦，専業主夫 | 6 | 10 | 13 | 8 | 37 |
| 学生 | 0 | 0 | 1 | 1 | 2 |
| 無職 | 5 | 6 | 4 | 6 | 21 |
| その他 | 4 | 1 | 5 | 3 | 13 |
| $N$ | 52 | 52 | 46 | 48 | 198 |

図の中心にある縦点線は，検索回数 $s = 50$ である．被験者への依頼事項は検索を 50 回以上することであり，点線より右側で被験者数が急激に増えている．

### 4.4 不正事象

#### 4.4.1 不正事象別発生ユーザ

表 7 は，不正事象別の発生ユーザ数である．不正事象 (1) 途中放棄が多く発生した．(1) 途中放棄の発生ユーザ数は，共有 ID を利用したユーザ (グループ A+C) は 20 ユーザ，個別 ID を利用したユーザ (グループ B+D) は 15 ユーザであり，個別 ID より共有 ID を利用したユーザの方が多くの不正事象を発生させた．(2) でたらめの発生ユーザ数は 11 ユーザであり比較的少なく，(3) 越権行為は 2 ユーザであり，ほとんど発生しなかった．

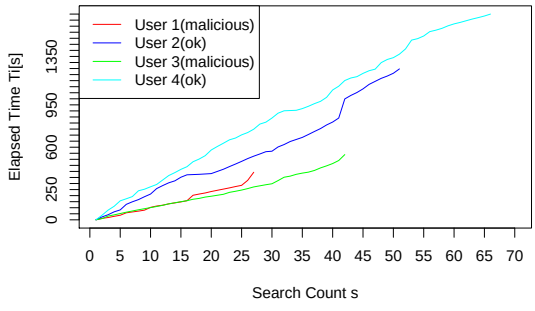


図 2: 所要時間と検索回数 (ユーザ毎)

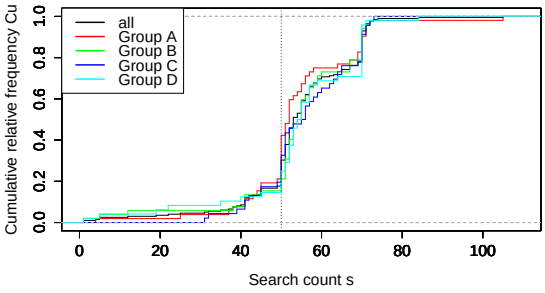


図 3: グループごとの検索回数

表 7: 不正事象別ユーザ数

| 分類 | A | B | C | D | Total |
|---|---|---|---|---|---|
| (1) 途中放棄 | 11 | 8 | 9 | 7 | 35 |
| (2) でたらめ | 5 | 3 | 1 | 2 | 11 |
| (3) 越権行為 | 1 | 1 | 0 | 0 | 2 |

(1) 途中放棄の発生ユーザ数が多いことから，属性別の発生ユーザ数を表 8 に示す．特に年代は，グループ毎の不正事象発生ユーザ数にばらつきがあった．30 歳〜39 歳 (以下，30 代) はグループ A の被験者が不正事象を多く発生させている．そこで，30 代の被験者におけるグループ毎の検索回数 $s$ の累積相対頻度 $Cu(s)$ を図 4 に示す．グループ A，B，C，D の検索回数 $s = 50$ は $Cu(s < 50)$ がそれぞれ，0.33，0.05，0.18，0.14 であり，グループ A は検索回数 $s$ が他のグループと比べて 50 回未満で作業を完了させるユーザが多い．

#### 4.4.2 独立性の検定

各不正事象の発生ユーザ数は，有意な差があるのか検定する．3.1 節で定めた仮説について，グループごとに

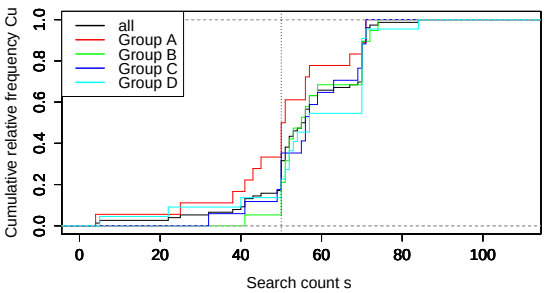


図 4: グループごとの検索回数（30 代のみ）

表 8: 不正事象 (1) 途中放棄の発生ユーザ数

| グループ | $A$ | $B$ | $C$ | $D$ | Total |
|---|---|---|---|---|---|
| 男性 | 7 | 5 | 6 | 6 | 24 |
| 女性 | 4 | 3 | 3 | 1 | 11 |
| ～19 歳 | 0 | 0 | 1 | 0 | 1 |
| 20 歳～29 歳 | 1 | 0 | 1 | 2 | 4 |
| 30 歳～39 歳 | 6 | 1 | 3 | 3 | 13 |
| 40 歳～49 歳 | 0 | 3 | 2 | 1 | 6 |
| 50 歳～59 歳 | 1 | 2 | 1 | 0 | 4 |
| 60 歳～ | 3 | 2 | 1 | 1 | 7 |
| 会社員 | 3 | 3 | 2 | 2 | 10 |
| 公務員 | 1 | 0 | 0 | 0 | 1 |
| 自営業 | 4 | 0 | 3 | 3 | 10 |
| パート，アルバイト | 1 | 0 | 0 | 0 | 1 |
| 専業主婦，専業主夫 | 1 | 2 | 1 | 0 | 4 |
| 学生 | 0 | 0 | 1 | 1 | 2 |
| 無職 | 1 | 2 | 0 | 1 | 4 |
| その他 | 0 | 1 | 2 | 0 | 3 |
| Total | 11 | 8 | 9 | 7 | 35 |

有意な差があるかを統計的に検定するため，次の $H_0$ と $H_1$ についてフィッシャーの直接確率検定を行う．

- 仮説 $H_1$(共有 ID)
  - 帰無仮説 ($H_0$): 共有 ID(A,C) と個別 ID(B,D) の不正発生は独立である．
  - 対立仮説 ($H_1$): 共有 ID と個別 ID の不正発生は独立ではない．
- 仮説 $H_2$(ID 未表示)
  - 帰無仮説 ($H_0$): ID 表示 (C,D) と ID 未表示 (A,B) の不正発生は独立である．
  - 対立仮説 ($H_1$): ID 表示と未表示の不正発生は独立ではない．

検定対象は，表 7 の各不正事象の発生ユーザ数およびグループ毎に発生のばらつきが大きい表 8 の 30 代の不正事象 (1) 途中放棄の発生ユーザ数とする．検定対象を表 2 の分類に集計したものが表 9 である．検定結果を表 10 に示す．30 代の不正事象 (1) 途中放棄は，$H_1$(共有 ID) の値が 0.1 以下であり，10%の有意水準で帰無仮説 $H_0$ は棄却され，共有 ID と個別 ID に有意な差がある．

#### 4.4.3 ロジスティック回帰分析

30 代の被験者において，どの要因が大きく誘発しているかを識別するため，個別 ID を基準として，共有 ID，

表 9: 不正事象の発生ユーザ数 (仮説毎)

| 分類 | 不正 | $H_1$(共有 ID) 共有 A+C | $H_1$(共有 ID) 個別 B+D | $H_2$(ID 未表示) 未表示 B+D | $H_2$(ID 未表示) 表示 C+D |
|---|---|---|---|---|---|
| 途中放棄 | あり | 20 | 15 | 19 | 16 |
| | なし | 78 | 85 | 85 | 78 |
| でたらめ | あり | 6 | 5 | 8 | 3 |
| | なし | 92 | 95 | 96 | 91 |
| 越権行為 | あり | 1 | 1 | 2 | 0 |
| | なし | 97 | 99 | 102 | 94 |
| 途中放棄 | あり | 9 | 4 | 7 | 6 |
| (30 代のみ) | なし | 26 | 37 | 30 | 33 |

表 10: フィッシャーの直接確率検定の分析結果

| 分類 | 仮説 | P value |
|---|---|---|
| 途中放棄 | $H_1$(共有 ID) | 0.3551 |
| | $H_2$(ID 未表示) | 0.8539 |
| でたらめ | $H_1$(共有 ID) | 0.7662 |
| | $H_2$(ID 未表示) | 0.2201 |
| 越権行為 | $H_1$(共有 ID) | 1.0000 |
| | $H_2$(ID 未表示) | 0.4987 |
| 途中放棄 | $H_1$(共有 ID) | 0.0763 . |
| (30 代のみ) | $H_2$(ID 未表示) | 0.7659 |

性別，職業の説明変数に対してロジスティック回帰分析を行った．目的変数を不正事象「途中放棄」の発生ユーザ数，説明変数を共有 ID，性別，職業としたロジスティック回帰分析の分析結果を表 11 に示す．共有 ID の p 値が 0.1 以下であり，90%の有意水準を下回っており，不正事象の発生に影響を与えていることが分かる．不正事象の発生確率を $p$，共有 ID，性別，職業ごとの推定値（偏回帰係数）を $x_1$，$x_2$，$\cdots$，$x_7$ とした場合ロジスティック関数の逆関数であるロジット関数は，

$$\log \frac{p}{1-p} = -16.75 + 1.189x_1 + 1.189x_2 + \cdots 12.90x_2$$

であり，P 値が 90%の有意水準を下回った推定値 (偏回帰係数) が共有 ID であるため，

$$\log \frac{p}{1-p} \doteq -16.75 + 1.189x_1$$

と近似できる．$\frac{p}{1-p}$ はオッズ比 (odds ratio) であり，共有 ID のオッズ比は 3.285 倍，すなわち ID を共用すると，しないときに対して 3.28 倍不正が生じやすくなる．

表 11: ロジスティック回帰分析の分析結果

| | 変数 | 推定値 (Estimate) | 標準誤差 (Std.Error) | z Value | Pr(> \|z\|) | |
|---|---|---|---|---|---|---|
| | (Intercept) | -16.75 | 1455.39 | -0.012 | 0.991 | |
| 1 | 共有 ID | 1.189 | 0.675 | 1.760 | 0.0784 | . |
| 2 | 男性 | 1.165 | 0.902 | 1.292 | 0.196 | |
| 3 | パート，アルバイト | 14.40 | 1455.40 | 0.010 | 0.992 | |
| 4 | 会社員 | 13.94 | 1455.40 | 0.010 | 0.992 | |
| 5 | 自営業 | 13.80 | 1455.40 | 0.009 | 0.992 | |
| 6 | 専業主婦，専業主夫 | 14.17 | 1455.40 | 0.010 | 0.992 | |
| 7 | 無職 | 12.90 | 1455.40 | 0.009 | 0.993 | |

## 5 考察

### 5.1 年代毎の傾向と対策

ロジスティック回帰分析の分析結果によると，30 代は個別 ID を利用した場合と比べて，共有 ID を利用した際は約 3.285 倍の確率で不正事象が誘発されることが分かった．この世代は，セキュリティ研修等で内部不正があった場合に罰せられる事例等を知っているためではないかと想定する．60 代は，8 名中 7 名のユーザが不正を犯している．世代毎に内部不正の発生における傾向が異なる場合，内部不正への対策も世代毎に変えていく必要があるかもしれない．

## 6 おわりに

本研究は，共通アカウントと内部不正の関係を明らかにするために疑似環境による実験を行った．被験者は 4 つのグループに分けて，グループ毎に利用 ID や ID 表示などの条件を変えて不正事象の発生数を観測した．フィッシャーの直接確率検定による独立性の検定により，30 代の被験者では共有 ID と個別 ID の利用が内部不正に関係性があることを確認した．また，ロジスティック回帰分析の分析結果より 30 代の被験者が共有 ID を利用すると内部不正が約 3.285 倍，誘発されることが分かった．

不正事象毎に内部不正の発生数が異なった原因やより実環境に近い形での実験の実施についてを今後の課題とする．

## 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, (2016.08.19 参照).

[2] Dawn Cappelli and Andrew Moore and Randall Trzeciak : The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes (Theft, Sabotage, Fraud), Addison-Wesley Professional (2012).

[3] 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, p. 31 (2015).

[4] Hausawi, Yasser M. : Current Trend of End-Users' Behaviors Towards Security Mechanisms, Human Aspects of Information Security, Privacy, and Trust: 4th International Conference, pp. 140-151 (2016).

[5] 新原功一, 山田道洋, 菊池浩明: 共有アカウントは内部不正を誘発するか？, 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).

[6] 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).

[7] 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).

[8] 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).