キーボード入力の監視による不正利用者の判別方法

中國真教[†] 堂薗 浩[††] 野口義夫[††]

本研究では，UNIX ワークステーションをある利用者が利用する際，その利用者のキーボード入力の特徴をとらえることにより不正利用者の判別を行う．キーボード入力の癖は，利用者それぞれに特徴があり，たとえば，入力ミスの多いキーやキーボード入力の速度，実行する UNIX コマンドの種類やオプションの使い方など，キーボードの操作や UNIX の操作の熟練度の違いによりキーボード入力の特徴が異なる．このような判断基準を設け，利用者が実際にどのような操作を行ったのかをキーボードから直接情報を取得し，利用者の特徴や癖をコンピュータが学習することにより不正利用者の判別を試みる方法について実験し考察を行う．

A Method to Identify an Illegal User by Monitoring Keyboard Inputs

Masanori Nakakuni,[†] Hiroshi Douzono[††] and Yoshio Noguchi[††]

When a user uses a UNIX Workstation, the distinction of the illegal user will be possible by reading the characteristics of the user's keyboard input and learning. For example, the user will be identified by the habits: the UNIX command that it is carried out, the keys which are mistyped, and the speed of the keyboard inputs in the session. User's characteristics can be read from input of keyboard constancy and learned by computer, and the distinction of the automatic illegal user will be possible. We made some experiments of this method and the results are reported in this paper.

1. はじめに

近年，インターネットの急速な広がりと，数多くのネットワーククラッキングツールの流通のために，ネットワーククラッキングの発生が増加している．これはインターネットの利用において深刻な問題となっており，ネットワーククラッキング対策はコンピュータネットワークの管理者にとって重要な課題の 1 つとなっている．近頃のクラッキングツールは GUI ベースのものが多く，コンピュータの初心者でも操作が簡単で，クラッキングに関する多少の知識があれば，比較的容易にクラッキングを行うことが可能である．そのためコンピュータの初心者でさえクラッカーになることができる．クラッキングツールの種類は様々であるが，たとえば，ネットワーク上を流れるパケットの中身を盗み見るツールは，何者かがこのようなツールを利用することにより，ネットワーク利用者の個人データを盗み見られることがある．このように盗聴された個人情報には，インターネット上のネットワークサーバにアクセスするための利用者 ID とそのパスワードが含まれる場合があり，そのような情報が盗聴され，悪用される場合がある[5)~9)]．

一般に UNIX を搭載したコンピュータではユーザ名とパスワードを用いて正規利用者であることを認証する方式をとっているが，利用を許されていない人物が正規利用者のパスワードを上記のあるいはその他の方法を用いて盗み取り，コンピュータを不正に利用する場合がある．これらの不正利用を阻止するための現在の主なコンピュータセキュリティ対策はネットワークを経由したコンピュータへの不正アクセスを阻止する手法や様々な利用者認証方式を実装したものが多いが，不正アクセスが成功した場合の対策はあまり講じられていない．コンピュータへの不正アクセスが成功した後の不正利用者の発見方法はコンピュータ内部に蓄積する利用状況や利用時間などの履歴をそのコンピュータの管理者が手動で読み取り，不審な記録を発見するという方法が従来の主な手法である．しかし，不正利用者が不審な記録を残すことなく不正アクセスを行う場合もあり，このような記録を管理者が調べるだけで

† 佐賀大学大学院工学系研究科
Doctor Cource of Science and Engineering, Saga University
†† 佐賀大学理工学部
Faculty of Science and Engineering, Saga University

は不正利用者の発見ができない場合がある．

そこで，コンピュータのキーボード入力の様子から不正利用者を発見する手法として，粕川ら[1),2)]はコンピュータへのログイン時に行う利用者認証において，利用者IDとそれに対するパスワードの入力時に，キーが打たれてから次のキーを打つまでの時間を測定し，その特徴をとらえることにより個人認証を行う手法を提案した．しかし，この手法はネットワークを介した個人認証を行う場合，ネットワーク伝送速度に遅延が生じる可能性があり，キー入力を行う時間間隔を正確に測定することは困難であるため，ネットワークを介した個人認証を行う場合には不向きである．

次に，加藤ら[3)]はUNIX Shellなどで，コマンドを入力することなしに，リターンキーを打つ（空リターン）頻度や出現傾向のみを解析することにより，利用者それぞれの特徴をとらえて不正利用者の判別を行う手法を提案した．しかし，正規利用者と不正利用者の間で，空リターンに関する挙動が類似している場合は，不正利用者の発見は非常に困難であり，特に，正規利用者と不正利用者の両者が「空リターンは打たない」という特徴を持っていれば利用者のキーボード入力から空リターンの情報を検出するだけでは不十分であり，不正利用者の発見は不可能である．

また，高井ら[4)]は使用したコマンドやコマンド中で使用したオプション，参照したファイルなどに注目することにより，コンピュータの利用目的の違いから不正利用者を判別する手法と，キー入力の速さに注目することにより，キーボード操作の慣れや癖から不正利用者を判別する手法を提案した．使用したコマンドや参照したファイルなどの評価に関しては，不正利用者の判別に用いる基準として有効であると考えられるが，キー入力の速さの評価に関しては，粕川らの提案した手法のところで述べたとおり，利用者のキー入力の速さが正確に測定できない場合は，それに関する情報の信頼性は低いため，キー入力の速さは不正利用者の判別においてあまり有効ではなく，高井らの提案するこれらの手法では，実質的に，不正利用者の判別においてコンピュータの利用目的の違いだけに注目することになり，それ以外にコンピュータの利用目的とは無関係に現れる癖などの特徴に関しては注目していないことになる．

本研究では，UNIXマシンの利用者のキーボード入力から操作内容に関する様々な情報を直接取得し，コンピュータの利用目的の違いやキーボード操作の慣れや癖に関する利用者の特徴をコンピュータが学習することにより不正利用者の判別を行う実験[10)]を行い，その結果について考察を行う．

2. 従来の不正利用を阻止する手法とその問題点

本章では，これまでに提案されている不正利用を阻止する手法についてふれ，その問題点について述べる．コンピュータの不正利用を阻止する手法に関しては大別して，ネットワークを経由したコンピュータへのアクセス制御を行う手法と，利用者がコンピュータへのログイン時に行う利用者認証の手法があげられる．

まず，ネットワークを経由したコンピュータへのアクセス制限に関する手法とその問題点は以下のとおりである．

- ファイアウォールによるネットワーク内部の保護
  ファイアウォールによるアクセス制御では，ファイアウォールの外側からのアクセスによる侵入を防ぐことは可能であるが，ファイアウォールを介さない内部どうしのアクセスには無効である．
- 通信経路の暗号化によるネットワーク盗聴の防止
  ネットワーク上を流れるパスワードの盗聴を防ぐためにSSLやSSHを用い，ネットワークの通信経路を暗号化しても，通信経路上の盗聴以外の方法でパスワードが盗まれてしまえば，通信経路の暗号化は効力を発揮しない．

次に，利用者認証に関する手法とその問題点は以下のとおりである．

- ワンタイムパスワード
  ワンタイムパスワードとは，使い捨てパスワードとも呼ばれ，1度利用したパスワードは破棄し，それ以降の認証時にはそのパスワードが無効になるというシステムである．ワンタイムパスワードは1度しか利用できないため，ネットワークの通信経路上で盗聴されたとしても，そのパスワードは再利用できず，実質，パスワードの盗聴を防ぐことになる．しかし，ワンタイムパスワードを生成するためのパスフレーズが他人に盗まれてしまえば，他人がそのパスフレーズからワンタイムパスワードを生成し，それを利用することにより正規利用者への成りすましが可能になり，不正利用者の判別が不可能である．
- IDカード
  クレジットカードのような磁気テープが付属したIDカードなどをカードリーダで読み取ることにより認証を行う手法では，IDカードの所有者以外の人物がそのIDカードを利用することにより，簡単にIDカードの所有者に成りすますことが可

能であるので不正利用者の判別が不可能である．

- バイオメトリック個人認証技術
  指紋，虹彩，顔などの人間が生まれつき持つ生体特徴からカメラを搭載したハードウェアを用いて自動的に個人を特定するバイオメトリック個人認証技術が存在するが，認証の成功率はそのハードウェアが設置されている周辺の環境（温度や湿度，照明など）に非常に左右されやすい．そのうえ，認証に使用するハードウェアの多くは大変高価であるために入手が困難な場合があり，これらの技術の導入によりセキュリティの強度が高くなったとしても，それに投資する費用の面で問題が生じる．

これまでに述べたとおり，これらの手法では安全かつ手軽に利用者の認証を行うことができない．そこで，正規利用者以外の他人によって正規利用者の真似ができないような利用者自信が持つ特徴を，特別なハードウェアを用いることなく読み取り，正規利用者と不正利用者の判別を行う必要がある．

3. UNIX 利用者のキーボード入力の特徴

個々の利用者が行う UNIX マシンの操作内容はいくつかの要因によって利用者ごとに異なり，その操作内容の違いが利用者それぞれに現れる特徴であると考えられる．UNIX 利用者のキーボード入力において利用者それぞれに特徴が現れるその要因は，主に以下の 2 つであると考えられる．

- UNIX の利用目的の違い
- キーボード操作の慣れや癖の違い

次にそれぞれの要因について具体的に述べる．

(1) UNIX の利用目的の違い

UNIX の利用目的の違いによる UNIX 利用者の例は以下のとおりである．

- エディタや LaTeX を使い，メールを読む程度．使用する UNIX コマンドは `ls` や `mkdir` のような基本的なものを使用する．
- UNIX マシンを計算機サーバとして主にコンパイラを使用する．また，ソースファイルの編集のためにエディタを使用する．
- メールの読み書き，各種ファイルの作成をすべて UNIX 上で行う．また，UNIX コマンドや UNIX Shell を巧みに操る．
- UNIX の利用者でもあり，システム管理者でもあるため，一般利用者が使用するコマンドからシステム管理に関係するコマンドまで，様々なコマンドを実行する．
- UNIX に大変興味を持ち，好奇心旺盛で様々な UNIX コマンドの実行を試みる．

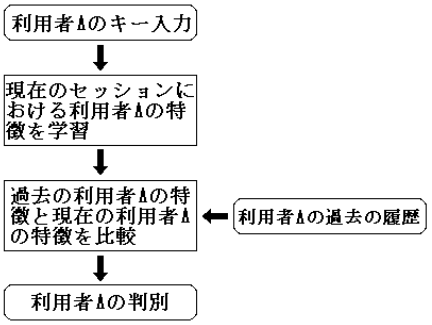

図 1 不正利用者の判別
Fig. 1 Identify an illegal user.

(2) キーボード操作の慣れや癖の違い

キーボード操作の慣れや癖の違いによる UNIX 利用者の例は以下のとおりである．

- キー入力が遅い．
- キー入力が速い．
- キー入力が遅いが，正確にキー入力を行う．
- キー入力が速いが，キー入力の間違いが多い．
- キー入力が速く，正確にキー入力を行う．
- 特定のキー入力のミスが多い．
- UNIX Shell に特有のキーバインドを多用する．

4. 利用者の特徴の学習

不正利用者を判別するためには利用者のキーボード入力をコンピュータが読み取り，その利用者の特徴や癖をコンピュータが学習し評価することによって不正利用者の判別を行う．利用者のキーボード入力から不正利用者の判別までのおおまかな流れを**図 1** に示す．

コンピュータが利用者から得た情報を学習するためには，それらの情報を何らかの方法で評価し，その評価の度合いを数値化する必要がある．そこで，不正利用者の判別に有効であると考えられる判定項目をいくつか設け，それらの判定項目に該当する情報を利用者のキーボード入力から取得する．不正利用者の判定を行うためには，それぞれの判定項目での評価を数値化する．ログインからログアウトまでに行う一連の作業を 1 つのセッションとし，利用者の特徴の学習においては，まず，初期の学習の段階では利用者のキーボード入力の特徴を学習し，正規利用者の特徴をとらえる．いくつかの判定項目における結果から総合的な評価を行い，正規利用者の過去の操作履歴と利用者の現在のセッションでの操作を比較することにより不正利用者の判別を行う．また，ある回数のコマンド実行までを区切りとし，1 つのセッションとして判別を行うこと

でログイン中の不正利用者の判別も可能になる．

5. 不正利用者を判別するための判定項目

不正利用者を判別するために判定項目をいくつか設ける．その項目は以下のとおりである．

(1) キー入力を行った回数
(2) 入力した文字の修正回数
(3) 入力した文字の種類（UNIX コマンドとともに使われる特殊記号）
(4) 実行したコマンドとその種類
(5) アクセスしたファイルとその種類
(6) 移動先のディレクトリ
(7) キー入力の速度

ここでは，これらの 7 つの判定項目を設けることにする．学習時にはこれらの判定項目をさらに細分化した判定項目を作成し学習を行う．たとえば，2 番目の判定項目「入力した文字の修正回数」を細分化する場合,「1 文字のみを削除し修正した文字とその回数」や「BS（バックスペース）キーの使用回数」などの項目に細分化し，それらの細分化されたものを，判定補助項目と呼ぶことにする．

そして，利用者のキーボード入力や UNIX の操作に関する熟練度の変化なども考慮しながら，コンピュータが利用者のキーボード入力を読み取り，それを学習する．

6. 判定項目の評価の数値化と利用者の学習

不正利用者の判定を行うためには，それぞれの判定項目での評価を数値化しなければならないが，その数値化の方法は，それぞれの判定項目で多少異なる方法を用いる．ここでは評価の度合いを数値化する例として,「1. キー入力を行った回数」,「2. 入力した文字の修正回数」,「3. 入力した文字の種類」,「4. 実行したコマンドとその種類」をあげて説明する．これらの説明のために，以下では次の記号を用いる．

N_i : セッションの回数
D_j : 判定項目の個数
D_{jk}: 判定項目 j の補助項目の個数
w_{jk}: 判定項目 jk における重み
X_{ijk}: セッション i における判定補助項目 jk の評価値
Y_{ij} : セッション i における判定項目 j の評価値
Z_i : セッション i におけるキー入力を行った回数
U_i : セッション i における実行したコマンドの個数
J_i : セッション i における評価値
J_{D_i} : セッション i における利用者の変化の評価値

ただし，i，j，k (>0) は以下のものとする．

i : セッションの番号
j : セッションにおける判定項目の番号
k : セッションにおける判定項目の補助項目の番号

そこで，セッション i における判定補助項目 jk の評価値 X_{ijk} は，利用者のキーボード入力から読み取った情報の中で，その判定項目に該当した回数とする．そして，それぞれの判定項目における評価式を以下のように定義する．

セッション i における判定項目「1. キー入力を行った回数」（判定項目番号 $j=1$）の評価値 Y_{i1} は式 (1) のように定義する．

$$Y_{i1} = Z_i \cdot w_{1k} \tag{1}$$

ただし，この判定項目では補助項目を 1 つとし，$k=1$ とする．また，重み w_{1k} は 0.0001 とする．これは，コマンドを実行する回数の評価を重視しないようにするために重みを小さな値に設定するためである．

セッション i における「2. 入力した文字の修正回数」（判定項目番号 $j=2$）の評価値 Y_{i2} は式 (2) のように定義する．

$$Y_{i2} = \sum_{k=1}^{D_{2k}} \frac{X_{i2k}}{Z_i} w_{2k} \tag{2}$$

ここでの重みは，過去から現在のセッションにおいて，判定補助項目 jk に初めて該当した場合は，その特徴を大きくとらえるために $w_{2k}=2.0$ とし，それ以外の場合は $w_{2k}=1.0$ とする．

セッション i における「3. 入力した文字の種類」（判定項目番号 $j=3$）の評価値 Y_{i3} は式 (3) のように定義する．

$$Y_{i3} = \sum_{k=1}^{D_{3k}} \frac{X_{i3k}}{Z_i} w_{3k} \tag{3}$$

ここでの重みは「2. 入力した文字の修正回数」の重みと同様に設定する．

セッション i における「4. 実行したコマンドの種類」（判定項目番号 $j=4$）の評価値 Y_{i4} は式 (4) のように定義する．

$$Y_{i4} = \sum_{k=1}^{D_{4k}} \frac{X_{i4k}}{U_i} w_{4k} \tag{4}$$

ここでの重みも,「2. 入力した文字の修正回数」の重みと同様に設定する．

このように，それぞれの判定項目 j の評価値を定義した．そして，これらの判定項目における評価値からセッション i における評価値は式 (5) のように定義

する．

$$J_i = \sum_{j=1}^{D_j} Y_{ij} \qquad (5)$$

利用者があるセッションで普段と違う振舞いをし，キーボード入力に変化があれば，そのセッションの前後で評価値が増減すると考えられる．

不正利用者の判別には，セッションの前後での評価値の差の大きさ（変化量）を利用者の変化の評価値とし，この値を判別に用いる．セッション i における利用者の変化の評価値 J_{D_i} は式 (6) のように定義する．

$$J_{D_i} = |J_i - J_{i-1}| \qquad (i > 1) \qquad (6)$$

本実験では，J_{D_i} の閾値を 0.1 に設定し，その値を超えた場合は不正利用者に判別する．

7. 不正利用者を判別する実験

今回，運用前の実験として，利用者が UNIX Shell のコマンドラインにおいてキーボード入力を行い，その様子を `script` コマンドを用いて記録し解析を行う．1 つのセッションをログインからログアウトまでとしたそのキーボード入力の記録を本実験のために作成した学習プログラムに入力し，その利用者のキーボード入力の特徴の学習を行う．

7.1 本実験における被験者の特徴

本実験では，まず，被験者を 3 人選択し，それぞれを利用者 A，B，C とする．いずれの被験者も普段から UNIX を利用しており，被験者それぞれの UNIX の使用経験は以下のとおりである．

利用者 A UNIX 使用歴約 6 年．UNIX 上ではメールの読み書き，文章作成，プログラムの作成など様々な作業を行う．システム管理者も兼ねる．

利用者 B UNIX 使用歴約 2 年．UNIX 上ではメールの読み書き，プログラムの作成を行う．

利用者 C UNIX 使用歴約 2 年．UNIX 上ではメールの読み書き，文章作成，プログラムの作成を行う．

7.2 実験時の条件

それぞれの利用者 A，B，C は以下の条件の下で実験を行う．

条件 1 被験者が UNIX マシンへログインし，ログアウトするまでに行われる UNIX マシンの操作におけるキーボード入力の様子のすべてを記録する．

条件 2 キーボード入力の方法はそれぞれの被験者が日頃行っている入力方法で実験を行う（片手もしくは両手でのキーボード入力などの方法は問わない）．

条件 3 被験者が実験に利用するキーボードは，被験者が日頃から利用し使い慣れているキー配列のキーボードを用いる．

条件 4 いずれの被験者も 1 つのセッションにおいてコマンドの実行回数には制限がないものとする．

7.3 実験の方法

実験では前述の条件の下に，次の 2 通りの実験を行う．

実験 1 正規利用者 A が 40 回のセッションにわたり 1 人でキーボード入力を行う．

実験 2 正規利用者 B に成りすました不正利用者 C が存在することを想定し，ある回数のセッションまで利用者 B がキーボード入力を行った後に，利用者 C が利用者 B と交代してキーボード入力を行う．利用者 B がキーボード入力を利用者 C に交代する時期は，コンピュータが利用者 B のキーボード入力の様子を学習し特徴をとらえた後とする．

これらの実験を行い，それぞれの利用者のキーボード入力の様子をコンピュータに学習させ，それぞれの実験において不正利用者の判別に関する評価を行う．

8. 不正利用者を判別する実験の結果

実験 1，2 のそれぞれの結果を**図 2**，**図 3** に示す．

まず，図 2 は実験 1 において利用者 A の 40 回にわたるセッションごとのキーボード入力の様子を評価し，それを数値化したグラフである．このグラフでは 1 回目のセッションで利用者の特徴を学習したときは，コンピュータが利用者 A のキーボード入力の特徴を知らないので評価値は高くなっている．1〜5 回目のセッションでは，セッション回数が増すにつれて評価値が減少し，利用者 A の特徴をコンピュータが学習している様子を示している．6〜11 回目のセッションでは評価値の変化量が以前に比べて減少しているが，これは利用者 A の大まかな特徴を学習したことを示している．そして，12 回目以降のセッションでは，それまでのセッションに比べて評価値の変化量がさらに減少し，評価値がほぼ一定になっているが，これはコンピュータが利用者 A の特徴をとらえたことを示している．実験 1 では，1 回のセッションあたりのコマンドの実行回数の平均は約 20 回であり，約 10 回のセッションで利用者のキーボード入力の特徴をとらえているので，利用者の特徴をとらえるためには，約 200 回のコマンドの実行を必要としたことが分かる．

また，19〜23 回目のセッションでの評価値が徐々に増加し，それ以降のセッションでは再び評価値の変化量が減少しているが，これは，19〜23 回目のセッションで利用者のキーボード入力に変化があったことを示し，24 回目以降でコンピュータが利用者の変化を学習

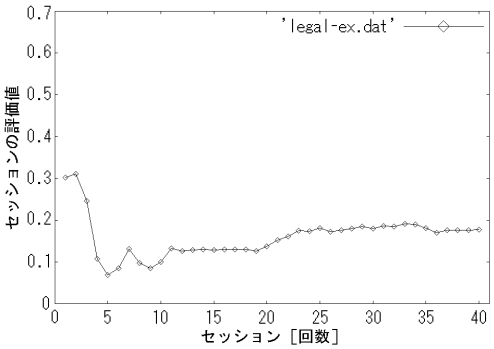

図 2 実験 1 におけるキーボード入力の評価
Fig. 2 Evaluation for legal user.

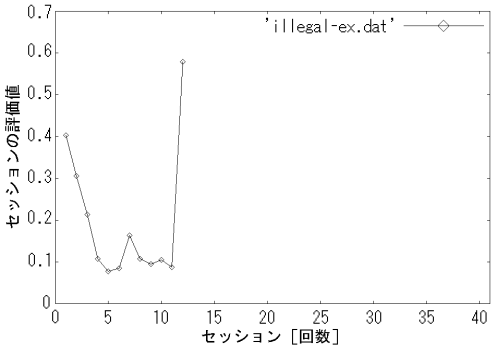

図 3 実験 2 におけるキーボード入力の評価
Fig. 3 Evaluation for illegal user.

したことを示している．この 19〜23 回目のセッションの評価値の増加は，利用者 A がこれまでのセッションで実行したことのないコマンドを実行したために評価値が増加した．

正規利用者が一貫してキーボード入力を行った様子の評価において，正規利用者の判別を正しく行うことができた場合のグラフの概形はいずれも，利用者の特徴の学習が完了した後にはセッションの前後において，評価値の変化量は閾値を超えることはなく，図 2 と同様に変化が小さなグラフとなる．

次に，図 3 は実験 2 において，ある回数のセッションまで正規利用者 B がキーボード入力を行った後に，不正利用者 C が正規利用者 B の代わりにキーボード入力を行った様子を評価し，それを数値化したグラフである．このグラフは，12 回目で利用者 B の代わりに利用者 C がキーボード入力を行った結果である．1〜11 回目までのグラフは実験 1 での 11 回目までのものと同様に増減しており，11 回目でコンピュータは利用者 B の特徴をとらえたことを示している．実験 2 では実験 1 と同様に，1 回のセッションあたりのコマンドの実行回数の平均は約 20 回であり，約 10 回のセッションで利用者のキーボード入力の特徴をとらえているので，利用者の特徴をとらえるためには，約 200 回のコマンドの実行を必要としたことが分かる．

11 回目のセッションでは評価値の変化量が減少していたが，12 回目のセッションでは評価値が急激に増加している．これは 1〜11 回目のセッションでキーボード入力を行った利用者と 12 回目のセッションでキーボード入力を行った利用者の特徴が異なり，コンピュータは 2 人が同一人物ではないと判断したことを示している．つまり，不正利用者 C が正規利用者 B に成りすましてキーボード入力を行ったことを示している．このセッションで評価値が急激に増加した要因は以下のものだった．

（1） 正規利用者 B と不正利用者 C では入力ミスの回数が多いキーが異なった．

（2） 正規利用者 B のセッションで使用回数が少なかったキーバインドを不正利用者 C は多用していた．

ここで，実験 2 における他の成功例を**図 4**，**図 5**，**図 6** に示す．

図 4 の実験では 1〜30 回目までのセッションにおいて，利用者 D がキーボード入力を行い，31 回目以降は利用者 E が利用者 D と交代してキーボード入力を行った場合の評価である．図 4 は，図 3 の結果と同様にキーボード入力を交代した 31 回目のセッションで評価値が急激に変化し，コンピュータが 2 人を同一人物でないと判断したことを示している．

図 5 の実験では 1〜25 回目までのセッションにおいて，利用者 F がキーボード入力を行い，26 回目以降は利用者 G が利用者 F と交代してキーボード入力を行った場合の評価である．図 5 は，図 3 の結果と同様にキーボード入力を交代した 26 回目のセッションで評価値が急激に変化し，コンピュータが 2 人を同一人物でないと判断したことを示している．

図 6 の実験では 1〜17 回目までのセッションにおいて，利用者 H がキーボード入力を行い，18 回目以降は利用者 I が利用者 H と交代してキーボード入力を行った場合の評価である．また，この実験では，利用者 H が過去のセッションにおいて実行していたコマンドを利用者 I が知っているものとし，利用者 I が実行したコマンドを利用者 H がその真似をすることにより，コンピュータがどのように評価を行うのかを実験した．図 6 は，図 3 の結果と同様にキーボード入力を交代した 18 回目のセッションで評価値が急激に変化し，コンピュータが 2 人を同一人物でないと判断したことを示している．このセッションで評価値が急

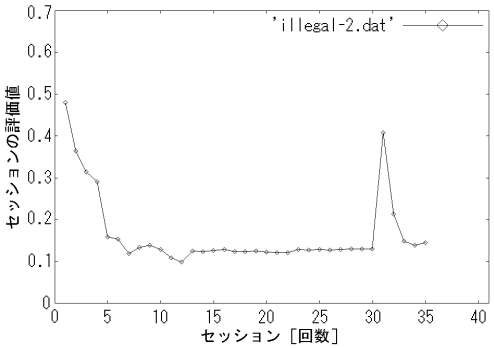

図 4 実験 2 における評価の成功例（その 1）
Fig. 4 Successfully identified example of experiment 1.

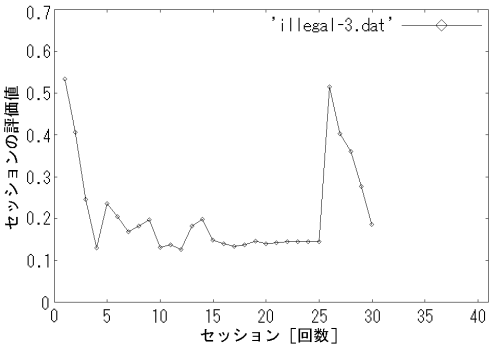

図 5 実験 2 における評価の成功例（その 2）
Fig. 5 Successfully identified example of experiment 2.

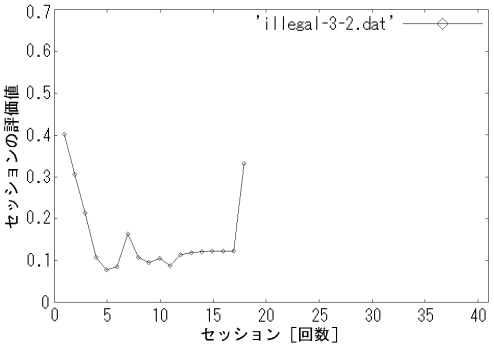

図 6 実験 2 における評価の成功例（その 3）
Fig. 6 Successfully identified example of experiment 3.

激に増加した要因は実験 2 の図 3 で示した結果と同様に，利用者 H と利用者 I では入力ミスの回数が多いキーが異なったことと，利用者 H のセッションで使用回数が少なかったキーバインドを利用者 I が多用していたことが要因である．つまり，利用者 H の実行するコマンドを利用者 I が真似たとしても，コンピュータが他の特徴を読み取ることにより，不正利用者を判別

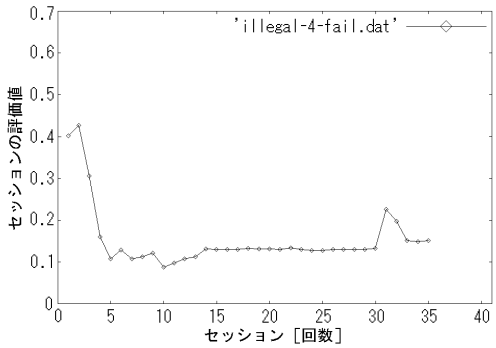

図 7 実験 2 における評価の失敗例（その 1）
Fig. 7 Unsuccessful identified example of experiment 1.

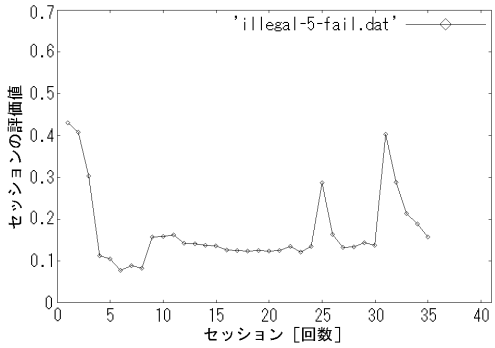

図 8 実験 2 における評価の失敗例（その 2）
Fig. 8 Unsuccessful identified example of experiment 2.

できたことを示している．

また，参考のために，**図 7**，**図 8** に失敗例を示す．

図 7 の実験では 1〜30 回目までのセッションにおいて，利用者 J がキーボード入力を行い，31 回目以降は利用者 K が利用者 J と交代してキーボード入力を行った場合の評価である．図 7 は，図 3 の結果と同様に利用者 K の学習を行っているが，キーボード入力を利用者 J から利用者 K に交代した 31 回目のセッションでは，これまでのセッションにおいて評価値の変化量が少し大きくなっているが，その変化量が本実験で設定した閾値を超えなかったために，不正利用者の判別に失敗している．誤って評価した原因は，利用者 K が 1 つのセッションにおいて，コマンドを実行した回数が少なく，利用者 K から得られる情報が少なかったために評価値の変化量が不正利用者の判別のための閾値を超えなかったからである．

図 8 の実験では 1〜30 回目までのセッションにおいて，利用者 L がキーボード入力を行い，31 回目以降は利用者 M が利用者 L と交代してキーボード入力を行った場合の評価である．図 8 は，図 3 の結果と同

様に利用者Lの学習を行っているが，1〜30回目までのセッションにおいて，利用者Lが1人でキーボード入力を行ったにもかかわらず，25回目のセッションで評価値が急激に増加し，その変化量が不正利用者の判別のための閾値を超えたために，誤って正規利用者のキーボード入力を不正利用者のキーボード入力に判別している．誤って評価した原因は，利用者Lが，日頃，実行しないコマンドを数多く実行したためである．

本実験では正規利用者が一貫してキーボード入力を行った実験1において，10人の被験者を用意し実験を行ったところ，8割（8人）の精度で正規利用者の判別を行うことができ，不正利用者のキーボード入力を想定した実験2において30組の被験者のペアを用意し，30回の実験を行ったところ，約7割（22回）の精度で不正利用者の判別を行うことができた．そして，利用者のキーボード入力において，「UNIXの利用目的の違い」と「キーボード操作の慣れや癖」に注目し利用者のキーボード入力の特徴の学習を行ったが，多くの利用者に後者の特徴が顕著に現れ，その特徴に注目することは不正利用者の判別に有効であることが分かった．

また，式(5)で定義したセッション i における評価式 J_i に関して，この定義式では判定項目の各項目 Y_{ij} の変化量が埋もれてしまい，正規利用者と不正利用者を判別する精度が低くなる可能性がある．そこで，式(5)の代わりに評価式 J_i を

$$J_i = \sum_{j=1}^{D_j} |Y_{ij} - Y_{i-1j}| \qquad (7)$$

ただし，$i = 1$ のとき $Y_{0j} = 0$

とすることで，利用者の判定の際に判定項目 Y_{ij} の変化量を反映させることができると考えられ，式(5)の変更以外はすべて同一の条件の下で，利用者を判別する実験1，2を行った．この場合，利用者のキーボード入力の変化に対して評価式が敏感になることが多く，不正利用者を判別する実験2において，不正利用者を正しく判別する精度は0.4割ほど向上したが，一方，正規利用者を判別する実験1においては，正規利用者を正しく判別する精度が2割低下した．不正利用者を正しく判別する精度が10割に限りなく近いことは大変重要ではあるが，不正利用者を正しく判別する精度の向上に比べ，正規利用者を正しく判別する精度が著しく低下したため，本研究では式(5)を用いて評価を行う方が，総合的に良い結果が得られると考え，式(7)を採用しなかった．

9. 考　　察

本章では，本手法の利点と問題点について考察を行う．まず，本手法の利点を以下にまとめる．

- 正規利用者への成りすましの発見が可能
  本手法では，正規利用者本人のUNIXの操作における特徴をとらえているため，万一，正規利用者の利用者IDとパスワードが漏洩し正規利用者以外の人物に利用された場合でも，その正規利用者に成りすました人物のUNIXの操作における特徴をとらえ，2人の特徴を比べることにより，2人が同一人物でない，つまり不正利用者であることを判断することが可能である．
- ファイアウォールを介さない不正利用の発見が可能
  コンピュータの不正アクセスはファイアウォールで守られたネットワークの外部からだけでなく，ファイアウォールを介さない，そのネットワークの内部における不正アクセス（ダイアルアップによる電話回線を介したネットワーク内部への侵入なども含む）が発生する可能性がある．本手法では，このようなファイアウォールを介さずに行う不正利用の発見にも有効である．
- 付加的なハードウェアは不要
  本手法では，利用者の認証を行うためのハードウェアを必要としない．したがって，本手法による利用者認証システムの導入は比較的容易である．

次に本手法の問題点を以下にまとめる．

- 判定項目「キー入力の速度」について
  利用者の「キー入力の速度」について評価を行った場合，利用者が回線速度の安定しない低品質のネットワークを経由してUNIXマシンにログインした場合，キー入力の速度を評価することは困難である．「キー入力の速度」について評価を行う場合は，利用者のログイン元の所在を明確にし，UNIXマシンとログイン元の間の回線速度とそのゆらぎが生じることを加味する必要がある．
- キーボード入力から得る不正利用者の情報の不足
  本手法では，不正利用者の侵入後，不正利用者のキーボード入力からその特徴をとらえ，即座に不正利用者の判別を行う必要があるが，不正利用者のキーボード入力から得られる情報量がきわめて少ない場合，不正利用者の判別が可能でない場合がある．

10. まとめと今後の課題

本研究では，UNIX マシンにおいて利用者のキーボード入力から利用者に関する情報を直接取得し，利用者のキーボード入力の特徴や癖をコンピュータが学習することにより正規利用者と不正利用者の判別を試みる方法について実験および考察を行った．

今後の課題としては，本手法のコンピュータへの実装の方法について検討を行い，利用者のキーボード入力の情報をカーネルを介して取得し，それらの情報を管理者権限を持った利用者以外には見ることができないよう考慮する必要がある．この際，実装時のコンピュータにかかる負荷は，不正利用者の判別を行う評価式は簡単なものであるため，コンピュータへの負荷は小さいと考えられ，リアルタイムで利用者のキーボード入力を監視することが可能となり，UNIX マシンの運用を妨げることなく本手法を実装することが可能であると考えられる．

謝辞 本研究の実験のために Sun Microsystems 社製 Ultra SPARC 10 を提供していただいた佐賀大学低平地防災研究センターのスタッフの皆様に深く感謝いたします．

参 考 文 献

1) 粕川正充，森 裕子，小松賢嗣，赤池英夫，角田博保：打鍵データに基づく個人認証システムの評価と改良，情報処理学会論文誌，Vol.33, No.5, pp.728–735 (1992).
2) 粕川正充，角田博保，森 裕子：アルペジオ打鍵列を利用した個人認証手法の提案，情報処理学会論文誌，Vol.34, No.5, pp.1198–1205 (1993).
3) 加藤，高田，小高，小倉：対話的計算機環境におけるキーボード入力系列のモデル化と認証への応用，電子情報通信学会論文誌（A），基礎・境界，Vol.J78-A, No.9, pp.1251–1254 (1995).
4) 高井，佐藤，宮崎：ユーザモデルを利用した継続的認証システム，電子情報通信学会技術研究報告，OFS，オフィスシステム，Vol.97, No.626, pp.37–42 (1998).
5) Garfinkel, S. and Spafford, G.: Practical UNIX Security, O'Reilly & Associates, Inc. (1991). 山口英（監訳）：UNIX セキュリティ，ASCII 出版局 (1993).
6) 山口 英，鈴木裕信：bit 別冊情報セキュリティ，共立出版 (2000).
7) 不正アクセスの動向，コンピュータ緊急対応センター (1999). http://www.jpcert.or.jp/nl/99-0004/99-0004-01.html
8) 情報処理振興事業協会セキュリティセンター. http://www.ipa.go.jp/SECURITY/index-j.html
9) Security of the Internet，CERT. http://www.cert.org/encyc_article/tocencyc.html
10) 中國真教，原 重臣，堂薗 浩，野口義夫：UNIX ワークステーションにおける不正利用者の判別方法，情報処理学会研究報告，2000-CSEC-7, pp.1–5 (2000).

(平成 12 年 5 月 1 日受付)
(平成 12 年 10 月 6 日採録)

中國 真教（学生会員）
1975 年生．1998 年 3 月佐賀大学理工学部物理学科卒業．2000 年 3 月同大学大学院工学系研究科博士前期課程修了．現在，同大学院工学系研究科博士後期課程在学中．情報セキュリティに関する研究に従事．

堂薗 浩（正会員）
1961 年生．1989 年 3 月京都大学大学院工学研究科博士後期課程修了．同年 4 月徳島大学工学部助手．1991 年熊本大学工学部助手．1992 年同大学工学部機械工学科講師．1994 年 3 月佐賀大学理工学部電子工学科助教授．現在に至る．ニューラルネット，遺伝的アルゴリズム，フローサイトメトリーソフトウエア工学に関する研究に従事．工学博士．計測自動制御学会，システム制御情報学会，International Neural Net Council 各会員．

野口 義夫
1941 年生．1965 年 3 月京都大学工学部数理工学科卒業．1965 年 4 月電子技術総合研究所勤務．1994 年 4 月佐賀大学理工学部電子工学科教授．現在に至る．工学博士．文字読取装置の開発，部分空間法，マルチスペクトル細胞画像に基づく自動細胞診断，染色体ソータの試作，ニューラルネットワーク等の研究に従事．