テクニカルノート

コマンド利用の周期性に基づく個人認証手法の提案

小高知宏[†] 白井治彦[†]
西野順二[†,☆] 小倉久和[†]

コンピュータの利用者が正当な利用者であることを調べる技術である認証について，利用者の挙動に基づく新しい方法を提案する．本方法では，利用者のコンピュータに与える操作コマンド系列の周期性を解析することで利用者のモデルを作成する．一定期間について入力された特定のコマンドについて周期を測定し，その結果を利用者のモデルとする．このモデルをセッションごとに作成し，あらかじめ正当な利用者について過去に作成したモデルと比較することで認証を行う．

An Authentication Method Based on the Cycle of the Command Repetition

Tomohiro Odaka,[†] Haruhiko Shirai,[†] Junji Nishino[†,☆] and Hisakazu Ogura[†]

In this paper, we propose a new authentication method that is based on the analysis of computer user behavior. The authentication system analyzes a series of user input commands for shell command interpreter program, and constructs user models that includes the cycle of command repetition. The authentication system constructs user models for all sessions and compare the user models at each other. If the user model differs from former models, the system warns of intrusion. This authentication method is compatible with password authentication system. The results of our experimentation shows that the intrusion detection rate is 84%.

1. はじめに

コンピュータネットワークの発展にともない，コンピュータセキュリティ技術の重要性が拡大しつつある．本稿では，コンピュータの利用者が正当な利用者であることを調べる技術である認証技術について，新しい方法を提案する．本稿で提案する方法では，利用者のコンピュータに与える操作コマンド系列の周期性を解析することで利用者のモデルを作成する．ここでコマンドの周期とは，あるコマンドが入力されてから次に同じコマンドが入力されるまでの間のコマンド投入数のことである．一定期間について入力された特定のコマンドについて周期を測定し，その結果を利用者のモデルとする．このモデルをセッションごとに作成し，あらかじめ正当な利用者について過去に作成したモデルと比較することで認証を行う．

利用者の認証には，利用者番号とパスワードを用いる認証方式（以下，パスワード方式とする）を用いるのが一般的である[1)]．しかし，パスワードが何らかの方法で他人に知られ，他人が不正にセッションを開始した場合には，パスワード方式は認証の能力を完全に失う．これに対して本方式は，セッションの途中でもつねに利用者の入力を監視し，その入力系列から利用者を認証することができる．また，本方式はパスワード方式の認証結果にかかわらず利用することが可能である．したがって，本方式をパスワード方式と組み合わせて用いることで，パスワード方式の欠点を補完することが可能である．

2. 利用者モデルの構築方法

本認証方式では，利用者がシステムに与えるコマンドの系列を，コマンドの投入周期に着目してモデル化する．

† 福井大学
Fukui University
☆ 現在，電気通信大学
Presently with The University of Electro-Communications

利用者が計算機にログインしてからログオフするまでを1セッションと定義する．セッションの開始からi番目に投入したコマンドを C_i とする．C_i 以降に初めて投入された $C_i = C_j$ となるコマンド C_j が与えられたとき，C_i におけるコマンドの周期 T_i を

$$T_i = j - i$$

と定義する．ただし，$C_i = C_j$ となるコマンドが存在しない場合には T_i は定義しない．N 個の入力コマンドからなるセッションについて，あらかじめ決めた1ないし2種類程度のコマンドについてそれぞれ周期を求める．次に，コマンドごとに求めた各周期の頻度を周期ごとにそれぞれ合計し，これらを要素とする特徴ベクトル v を作成する．

$$v = (v_1, v_2, v_3, ...v_k, ...v_n) / \sum_{i=1}^{n} v_i$$

ここで，v_k は周期 k となるコマンドの周期の，セッション全体での頻度である．また n は特徴ベクトル v の要素数である．本手法では特徴ベクトル v を利用者のモデルとする．

特徴ベクトルの計算例を示す．**図1** のようなセッションにおいて，lsコマンドとcdコマンドに着目して特徴ベクトル v を計算する場合を考える．図1の場合，入力コマンドの総数 N は8である．また，特徴ベクトル v の要素数 n は7としている．図1で，周期の計算に関係するlsコマンドまたはcdコマンドが出現するのは1行目，2行目，3行目，6行目，7行目，および8行目である．このうち，7行目と8行目のコマンドにはこれ以降に対応するコマンドが存在しないので周期計算の対象とならない．この結果，T_1，T_2，T_3，T_6 を扱うことになる．

図1で，C_1 はlsである．次にlsコマンドが出現するのは3行目であるから（すなわち C_3 =ls），

$$T_1 = 3 - 1 = 2$$

となる．同様に

$$T_2 = 5,\ T_3 = 3,\ T_6 = 2$$

となる．以上から，図1には周期2となる場合が2回（T_1 および T_6），周期3となる場合が1回（T_3），および周期5となる場合が1回（T_2）含まれることが分かる．これより，特徴ベクトル v は，

$$v = (0, 2, 1, 0, 1, 0, 0)/4$$
$$= (0, 0.5, 0.25, 0, 0.25, 0, 0)$$

となる．

認証においては，あらかじめ蓄積した特徴ベクトルの組と，認証を行いたいセッションから作成した特徴ベクトルを比較する．この際，両者の一致性を評価する必要がある．本稿では3層ニューラルネットワークをあらかじめ蓄積した特徴ベクトルを用いてBP法を用いてトレーニングすることで，未知の特徴ベクトルに対して一致性を評価することにした．

```
ls
cd work
ls -l
pwd
cp a b
ls
cd
ls
```

図1 セッションの例
Fig. 1 An example of a session.

3. 実　　験

本稿で提案する特徴ベクトルの有効性を示すために，8名のUNIX利用者がコマンドインタプリタtcshに対して入力したデータを用いて認証実験を行った．データはtcshを改造することで取得した．全利用者とも，コンピュータソフトウェア研究とプログラム開発を行う大学院生および学部4年生である．データ採取は2カ月間行った．極端に短いセッションでは周期性の検出が困難なため，採集したデータから50コマンド以上の入力を行ったセッションを選別して実験に用いることにした．各セッションにおける解析の対象コマンド数 N はセッションの先頭から50コマンドとし，特徴ベクトル v の長さ n を50とした．また，特徴ベクトル計算の対象とするコマンドは，あらかじめ各利用者のセッションを目視により観察することで，セッション中に頻繁に現れるコマンドを指定した．

3.1 特定の利用者グループ内での認証結果

採取した8名分のデータからセッション数の多いもの4名を選び，各セッションについて特徴ベクトルを作成した．さらに，これらの特徴ベクトルをデータ採取実験を行った初めの月と次の月に分けてそれぞれ10ずつ用意し，4人分のそれぞれ本人の特徴ベクトル（正例）10例，他人の特徴ベクトル（負例）30例からなるデータセットを作成した．はじめの1カ月間のデータを学習セットとし，次の1カ月間のデータを検査セットとした．特徴ベクトルを求める対象のコマンドとして，**表1** 右欄に示すようなコマンドを選んだ．

表 1 学習結果と検査セットに対する認証結果
Table 1 Results of authentication.

	学習結果		検査結果		対象としたコマンド
	認証成功率 (%)	排除率 (%)	認証成功率 (%)	排除率 (%)	
被験者 A	90	100	70	73	vi
被験者 B	100	100	60	83	ls
被験者 C	100	100	50	90	ls, jlatex
被験者 D	100	93	90	83	ls
全体の平均	98	98	68	83	

表 2 未知の利用者に対する排除結果
Table 2 Results of authentication (for unknown subjects).

	10 回の認証に対して，本人であるとした場合の回数				E〜H 合計の排除率 (%)
	被験者 E	被験者 F	被験者 G	被験者 H	
被験者 A	1	5	1	1	80
被験者 B	3	3	1	2	78
被験者 C	4	0	2	0	85
被験者 D	0	2	1	0	93
E〜H の平均					84

表 3 性質の異なる 94 名の 940 セッションに対する排除結果
Table 3 Results of authentication (for 94 unknown subjects).

	被験者 A	被験者 B	被験者 C	被験者 D	平均
排除したセッション数	638	743	652	783	704
排除率	67.9	79.0	69.4	83.3	74.9

学習セットデータを用いて 3 層ニューラルネットワークを BP 法を用いてトレーニングした．ネットワークは入力層，中間層および出力層のニューロン数がそれぞれ 50，50，1 とした．4 名それぞれについて別々にネットワークを用意し，それぞれのネットワークが本人の特徴ベクトルに対して 1，他人の特徴ベクトルに対して 0 を出力するようにトレーニングした．トレーニングは，学習セットについて学習が飽和するまで行った．トレーニング完了後，検査セットをネットワークに与え，出力のしきい値 Z_0 を $Z_0 = 0.5$ として判別を行った．表 1 に学習結果と検査結果を示す．表で，本人を正しく本人と認識する割合を認証成功率と呼び，本人以外を他人として正しく排除する割合を排除率としている．表 1 にあるように，学習セットについては，認証成功率，排除率ともほぼ 100%となった．また検査セットについては，認証成功率は平均で約 70%であるが，最低の 50%から最高の 90%まで，被験者によるばらつきが大きかった．排除率については，被験者により 73%から 90%となり，全体の平均で約 83%となった．

3.2 未知の利用者を含む場合の認証結果

次に，学習を終えたニューラルネットワークを用いて，学習セットに出現しない利用者の特徴ベクトルを排除できるかどうかを調べた．8 名の被験者のうち学習セットに含まれない被験者 E から被験者 H までについて同様にして特徴ベクトルを作成し，ニューラルネットワークに与えた結果を**表 2** に示す．

3.3 性質の異なる集団に対する認証結果

本手法による認証がある程度の規模を持った性質の異なる集団に対しても有効であるかどうかを調べるために，学部 2 年生と 3 年生からなる 94 名の UNIX 利用者の各 10 セッションについて同様にして特徴ベクトルを作成し，学習済みのニューラルネットワークに与えた結果を**表 3** に示す．表 3 に示すように，排除率の全体の平均は 74.9%である．

4. 考察とまとめ

本稿では，計算機利用者が計算機に与えるコマンド列の周期性に利用者の特徴が含まれていることを指摘し，この特徴が認証に利用できる可能性を示した．本方式は従来研究されているエキスパートシステムなどを応用した侵入監視システム[2]と異なり，利用者の挙動を直接監視するものである．また，キーボード打鍵間隔や文字の連鎖を監視する方法[5)〜7)]と異なり，キーボードやその他のハードウェアに依存しない方法である．

本稿の実験結果から，4 名の被験者のうちの 1 名である被験者 D において，本人を正しく本人と認識する認証成功率が 90%，学習セットに含まれる他人の未知データに対する排除率が 83%，まったく未知の利用者

に対する排除率が 93%，性質の異なる 94 名の集団に対する排除率が 83%となることを示した．この程度の認証能力があれば，パスワード方式を補完する認証手法として用いることが十分可能である．ただし，まったく未知の利用者に対する排除率が 93%と高率となったのは，たまたま被験者 D のコマンド入力傾向が被験者 E〜H と大きく異なっていたためであると考えられる．通常は学習セットに含まれる被験者に対する排除率を超えることは期待できない．

今回提案した手法では，排除率と比較して認証成功率が低い．すなわち，表 1 の検査結果に示すように，被験者 A，B，C，D に対してそれぞれ 70%，60%，50%，90%となった．したがって，被験者 D 以外に対して本手法を単独で認証に用いることは実用上難しい．このことからも，本手法を認証に用いる場合には単独で用いるのではなく，パスワード方式や本人を確認するための質問などと合わせて認証に用いるのが適当である．そうすれば，認証成功率の低さを補いつつ，本手法を適用することで侵入者排除のきっかけを作ることが可能である．

本稿における手法以外にも，計算機利用者の挙動から利用者の認証を行う手法がいくつか存在する[9)〜11)]．特に，文献 11) で言及したコマンド出現頻度と本稿で提案した特徴ベクトルとの間には関係が深く，今後，両者の関連を追及する必要がある．利用者の挙動から認証を行う手法では，複数の手法を融合して用いることが効果的である[12)]．今後，これらの手法と本手法を融合し，より認証能力の高い手法を検討する予定である．

謝辞 本研究の一部は，矢崎科学技術振興記念財団の研究助成金により支援を受けました．ここに感謝の意を表します．

参 考 文 献

1) Steiner, J.G., Neuman C. and Schiller J.I.: Kerberos: An Authentication Service for Open Network Systems, *USENIX Winter Conference*, pp.191–202 (1988).
2) Lunt, T.F.: A survey of intrusion detection techniques, *Computers & Security*, 12, pp.405–418 (1993).
3) Denault, M., Gritzalis, D., Karagiannis, D. and Spirakis, P.: Intrusion detection: Approach and performance issues of the SECURENET system, *Computers & Security*, 13, pp.495–508 (1994).
4) Tsai, W.T., Keefe, T.F., Thomsen, D.J. and Thuraisingham, M.B.: AI Applications in Multilevel Database Security, *Computers Security Journal*, 6, pp.63–80 (1990).
5) Marcus, B. and Samuel, J.R.: A Practical Approach to user Authentication, *10th Annual Computer Security Applications Conference*, pp.108–116 (1994).
6) 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).
7) 岡本忠士，白石善明，大家隆弘，森井昌克：なりすましに対する不正侵入検知システム（IDS-M），情処学知能と複雑系研報，99-ICS-117-6 (1999).
8) Calvin, K., George, F. and Karl, L.: Automated Detection of Vulnerabilities in Privileged Programs by Execution Monitoring, *10th Annual Computer Security Applications Conference*, pp.134–144 (1994).
9) 白井治彦，小高知宏，西野順二，小倉久和：対話的計算機環境におけるコマンド入力連鎖を用いた認証手法の提案，信学論（A），Vol.J82-A, No.10, pp.1602–1611 (1999).
10) 加藤友彦，高田光男，小高知宏，小倉久和：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，信学論（A），Vol.J78-A, No.9, pp.1251–1254 (1995).
11) 小高知宏，加藤友彦，高田光男，西野順二，小倉久和：計算機利用者のシステム操作入力文字列に基づく認証手法の検討，信学論（A），Vol.J79-A, No.4, pp.1001–1003 (1996).
12) Murthy, V.K.: Probabilistic security protocol for real-time authentication in distributed databases, *Proc. 6th IASTED/ISMM International Conference, Parallel and Distributed Computing and Systems*, pp.253–256 (1994).

(平成 13 年 4 月 18 日受付)
(平成 13 年 9 月 12 日採録)