[研究ノート]

画像記憶のスキーマを利用した認証方式の改良：ストーリーの利用による記憶負荷の削減

An Improvement of User Authentication Using Schema of Visual Memory: Exploitation of "Schema of Story"

静岡大学創造科学技術大学院　日本学術振興会特別研究員（DC）　山本　匠
Graduate School of Science and Technology, Shizuoka University
Research Fellow of the Japan Society for the Promotion of Science (DC)　Takumi YAMAMOTO
三菱電機株式会社　原田篤史
Mitsubishi Electric Corporation　Atsushi HARADA
静岡大学情報学部　漁田武雄
Faculty of Informatics, Shizuoka University　Takeo ISARIDA
静岡大学創造科学技術大学院　科学技術振興機構，CREST　西垣正勝
Graduate School of Science and Technology, Shizuoka University
Japan Science Technology and Agency, CREST　Masakatsu NISHIGAKI

要　旨

近年，人間の画像認識能力の高さを利用して記憶負荷を軽減させる画像認証方式が注目されている．しかしながら，画像認証方式は毎回の認証時にパス画像がディスプレイ上に表示されるため，認証時の覗き見攻撃に対して脆弱であった．この問題に対し，覗き見をする攻撃者にとってパス画像の記憶が困難となるように，モザイク化等の不鮮明化処理を施した一見すると無意味に見える画像（不鮮明化画像）をパス画像として使用する「画像記憶のスキーマを利用したユーザ認証方式」が提案されている．しかしながら，不鮮明な画像であっても，毎回の認証で同じパス画像を用いている限り，覗き見攻撃者にそれを覚えられる可能性が残る．そこで本稿では，ユーザに m 枚のパス画像を記憶させた上で，その中の n 枚（m＞n）のパス画像を用いて認証を行うという改良を基本方式に加える．認証の都度，n 枚の画像を選び直すことで，覗き見攻撃は格段に困難になると期待される．本方式では，ユーザは，1 回の認証で利用するパス画像の枚数より多くパス画像を記憶する必要がある．そのため，いかにユーザに複数のパス画像を効率よく記憶してもらうかが重要になる．本稿では，パス画像を効率よく記憶する手段として，ユーザに一本の動画を覚えてもらい，動画中の各コマをパス画像として利用するという方法を採用する．これによりユーザは，ストーリーを持った複数のパス画像を容易に覚えることができるようになると考えられる．本稿では改良方式のプロトタイプシステムを実装し，比較実験により改良方式の有効性を確認する．

キーワード

画像認証　覗き見攻撃　スキーマ　不鮮明化画像　ストーリー　動画

1. はじめに

近年，人間の画像認識能力の高さを利用して記憶負荷を軽減させる画像認証方式[1,2]が注目されている．しかし，画像認証方式は毎回の認証時にパス画像が画面上に表示されるため，認証時の覗き見攻撃に対して脆弱であった．そのため，我々は既に，不鮮明な画像をパス画像に利用することで，覗き見攻撃に耐性を有する「画像記憶のスキーマを利用した認証方式」（以下，基本方式と呼ぶ）を提案している[3]．詳細は2章で述べるが，基本方式は従来の画像認証方式に比べて覗き見攻撃やパス画像の漏洩への耐性向上を果たしており，本人認証に関しても高い認証成功率を維持している．

しかし，不鮮明な画像であっても，毎回の認証で同じパス画像を用いている限り，覗き見攻撃者にそれを覚えられる可能性が残る．そこで本稿では，ユーザにm枚のパス画像を記憶させた上で，その中のn枚（m>n）のパス画像を用いて認証を行うという改良を基本方式に加える．認証の都度，n枚の画像を選び直すことで，覗き見攻撃は格段に困難になると期待される．

本方式では，ユーザは，1回の認証で利用するパス画像の枚数より多くパス画像を記憶する必要がある．そのため，いかにユーザに複数のパス画像を効率よく記憶してもらうかが重要になる．本稿では，パス画像を効率よく記憶する手段として，ユーザに一本の動画を覚えてもらい，動画中の各コマをパス画像として利用するという方法を採用する．これによりユーザは，ストーリーを持った複数のパス画像を容易に覚えることができるようになると考えられる．本稿では，改良方式の有効性を確かめるため，本改良方式と基本方式との比較実験を行う．

2. 基本方式のコンセプト

本章では，基本方式について簡単に概観する．詳細については文献[3]を参照されたい．

2.1 不鮮明化画像による認証

画像認証方式にとって覗き見攻撃が脅威となるのは，正規ユーザのみならず覗き見攻撃者にとっても画像の記憶は容易であるからである．そこで，基本方式では，覗き見をする攻撃者にとってパス画像の記憶が困難となるように，モザイク化等の不鮮明化処理を施した一見無意味な画像（以下，不鮮明化画像）をパス画像として使用する．人間は画像の記憶に優れているものの，それは有意味な画像を記憶する場合に限ってのことであり，無意味に見える（意味を言語化できない）画像を記憶することはやはり難しい[4]．ゆえに，他人のパス画像（不鮮明化画像）を覗き見て記憶することは，攻撃者にとって困難な作業となる．

ただし，無意味な画像を記憶することは正規ユーザにとっても困難であるため，正規ユーザにのみ，パス画像の登録時に不鮮明化画像のオリジナル画像（例：図1左）を見せ，当該画像に不鮮明化処理を施したパス画像（例：図1右）と合わせて記憶してもらう．不鮮明化画像にはオリジナル画像の特徴がある程度残されているため，オリジナル画像を見ることによって，正規ユーザは不鮮明化画像の中にオリジナル画像の持つ意味を見出せるようになる．この結果，正規ユーザは不鮮明化画像を有意味な画像として認識できるようになり，パス画像を容易に記憶することができる．これは，不鮮明なパス画像に対する「スキーマ」[5]を正規ユーザに学習させていることに相当する．ここでスキーマとは，人間が外界からの情報を知覚した際に無意識のうちに蓄積している「その情報をどのように認識・記憶したかという知識構造」を意味する認知心理学用語である．一度不鮮明化画像に対するスキーマを学習すれば，それ以降，当該不鮮明化画像を見た場合にも，スキーマにより簡単にその意味を再認識することが可能になる．

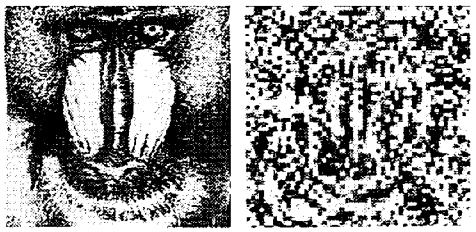

図 1 画像の不鮮明化処理

2.2 認証方式

認証の際には，システムは，パス画像である不鮮明化画像を囮画像とともに提示する．スキーマを有する正規ユーザは，複数の不鮮明画像の中からパス画像を選ぶことができる．

不鮮明化画像の生成手順および基本方式の認証手順の詳細に関しては文献[3]を参照されたい．図 2 に，9 択（囮画像 8 枚）システムの認証画面の例を示す．

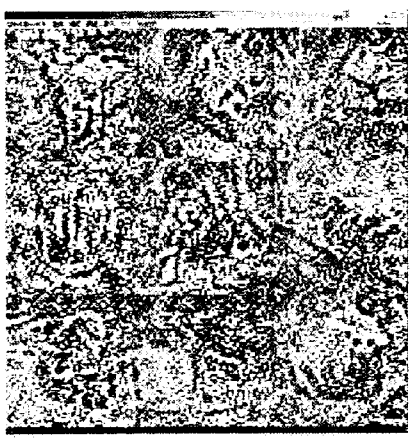

図 2 9 択認証システムにおける認証画面の例*

2.3 基本方式の有効性と課題

基本方式は，既存の画像認証方式（オリジナル画像をパス画像として利用する方式）と比べ，正規ユーザの認証成功率を高く維持したまま，攻撃耐性についても有望な結果を残している[3]．

文献[3]では覗き見攻撃者（実験実施者の認証行為を横で見ていた被験者）にとって非常に有利な条件である 2 択認証システム*を用いて覗き見実験が実施されているが，既存の画像認証方式のなりすまし成功率が 100%であったのに対し，基本方式ではなりすまし成功率を 90%に下げることができている．文献[3]では，パス画像の情報を他人に言葉で伝えることができるかどうかを測るパス画像漏洩の実験も実施された．攻撃者にパス画像の情報**を言葉で与えた上で，2 択認証システムによる認証試行を行わせたところ，既存の画像認証方式の認証成功率が 100%であったのに対し，基本方式の認証成功率は 74%であった．

しかしながら，基本方式では，毎回の認証におけるパス画像は常に同じものが使われる方式となっているため，攻撃者が覗き見した認証画面の中のパス画像がなりすましの際の認証画面にも必ず表示されることになる．不鮮明な画像であっても，同じパス画像を毎回の認証で用いている限り，攻撃者にそれを覚えられる可能性が残る．

そこで本論文では，ユーザに m 枚のパス画像を記憶させた上で，その中の n 枚（m＞n）のパス画像を用いて認証を行うという改良を加える．攻撃者が他者の認証を覗き見たとしても，次の認証で同じパス画像が現れるとは限らず，覗き見攻撃への耐性が向上すると期待される．ここで，正規ユーザが m 枚の不鮮明化画像を覚えるにあたっての記憶負荷を抑える方策が重要になる．

3. 動画を利用した改良方式

改良方式では，パス画像どうしに関連する意味を持たせる方法[6]に着目する．これは例えば，「女性」「椅子」「コーヒー」「ケーキ」の 4 枚のパス画像を，「花子さんが喫茶店の椅子に座ってコーヒーとケーキを注文した」というストーリー性のある文章と一緒に記憶することで記憶負荷を軽減する方法である．ただし，この方法を用いる上では，以下の 2 つの問題点に留意しなければならない．

---

* ユーザは 1 枚のパス画像を記憶する．認証画面に提示される画像は，パス画像と囮画像 1 枚の計 2 枚である．パス画像を選択することができたユーザを正規ユーザとして認証する．

** 文献[3]では動物の画像を不鮮明化したものをパス画像として用いたため，「動物の種類（例：犬）」，「正面か，横向きか」，「全身か，一部か」，「座っているか，立っているか」に関する情報を攻撃者に言葉で伝えた．

1)　パス画像の間のストーリー（関連）を推測されて攻撃される可能性がある．

2)　複数の画像に対する「記憶しやすいストーリー」を考えること自体が，正規ユーザにとって負荷となる．

1) は，「コーヒー」を覗き見られて，「ケーキ」がパス画像に含まれることが推測されてしまう，という問題である．しかし，本方式においては，攻撃者が不鮮明化画像（パス画像）を覗き見たとしても，その意味（オリジナル画像）を類推することは難しいため，そこから他のパス画像との関連を見抜くことは困難であり，この問題による脅威は小さいと考えられる．

2)に対しては，動画中の各コマをパス画像として利用することにより，この問題の解決を図る．認証システムは，ストーリーを有する任意の動画を一本準備する．そして，この動画の各コマを静止画に分解して，ある程度異なるシーンからｍ枚の静止画を抽出し，オリジナル画像セットとする．続いて，ｍ枚のオリジナル画像をそれぞれ不鮮明化処理することによりパス画像セットを生成する．正規ユーザには，登録時に，動画およびｍ枚のオリジナル画像とパス画像（不鮮明化画像）の組が提示される．

認証時には，ｍ枚のパス画像の内，ｎ枚（ｍ＞ｎ）が用いられる．毎回の認証ごとにｎ枚のパス画像は選び直され，その都度のパス画像ｎ枚を使って認証が行われる．改良方式のシステムの概観を図 3 に示す．

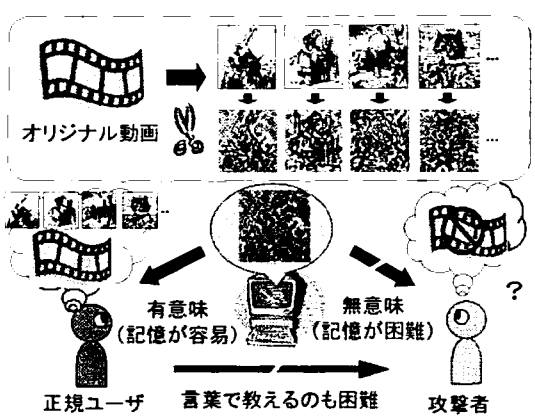

図 3　ストーリーのスキーマを利用する認証方式

4.　評価実験

本方式の有効性を，基本方式との比較実験を通じて評価する．被験者は，本学情報系学部学生 10 名である．

4.1 覗き見攻撃によるなりすまし実験

ストーリーの利用が覗き見攻撃への耐性に影響を与えるかどうか実験により確認する．

● 実験方法

本実験システムでは，正規ユーザが記憶すべきパス画像の枚数を 10 枚とし，2 択の認証フェーズ（認証画面中にパス画像 1 枚と囮画像 1 枚が提示される）を 1 ターン行って認証可否の判定を行う．すなわち，ｍ＝10，ｎ＝1 である．認証に使用されるパス画像は，認証の都度，10枚のパス画像セットの中からランダムに選択される．この設定は文献[3]の基本方式における覗き見攻撃の実験（ｍ＝1，ｎ＝1）の設定と同一であるので，両者をそのまま比較することができるというメリットがある．

実験実施者（正規ユーザ）が認証フェーズにおける 2 択の選択を 1 ターン行って認証を通過する認証画面を，各被験者（攻撃者）が間近から覗き見し，その直後に，正規ユーザへのなりすましを試みる．各被験者につき，同じパス画像セットについて 5 回ずつ認証試行を行ってもらった．文献[3]を参考に，2 択システムでは覗き見時間が 5 秒もあれば十分と考えられるため，各認証試行において，被験者の覗き見時間は 5 秒に設定した．なお，比較のために，動画から抽出した 10 枚のオリジナル画像をパス画像とした認証システム(ｍ＝10，ｎ＝1)も用意した．

動画には，時間の長さやストーリー性などの面で都合の良い，TVCM（コマーシャル）の動画を 10 種類用意した．動画ごとに，それぞれの中から適切な 10 コマを選んで 10 枚の静止画（オリジナル画像セット）を抽出し，これを不鮮明化処理して 10 枚の画像セットを生成した．被験者ごとに，実験実施者が無作為に 1 つの TVCM を選び，その TVCM から

生成された画像セットをパス画像セットとして用いて実験が行われる．

● 実験結果

実験の結果を表 1に示した．比較のために基本方式における覗き見実験（文献[3]の 3.3 節）の結果も一緒に示す．表中，「成功率」は 10 人の各被験者につき 5 回ずつ行った認証試行の全体の成功率（なりすまし成功率）を表し，「平均時間」は一回のパス画像選択に要した回答時間（認証画面が表示されてからマウスがクリックされるまでの時間）の平均値である．

表 1 覗き見攻撃実験の結果

	基本方式	改良方式 オリジナル画像	改良方式 不鮮明化画像
成功率	46/50 (92%)	43/50 (86%)	31/50 (62%)
平均回答時間（秒）	2.66	3.24	5.03

改良方式を用いることで，2 択の認証システムであっても，覗き見攻撃の成功率を 62%にまで抑えることができている．また，オリジナル画像を利用した場合の結果と比較しても，不鮮明化パス画像の使用が，覗き見して得た情報からストーリーを推測するという攻撃を困難にしていることが見てとれる．

4.2 言葉によるパス画像の漏洩実験

ストーリーの利用が言葉によるパス画像漏洩の耐性に影響を与えるかどうか実験により確認する．本実験システムは，4.1 節と同じものである（m＝10，n＝1）．

● 実験方法

文献[3]の基本方式における言葉によるパス画像漏洩の実験（m＝1，n＝1）と同様の設定で実験を行う．各被験者（攻撃者）は，実験実施者（正規ユーザ）からパス画像セットに対する動画（TVCM）のストーリーに関する情報を言葉*で教えられる．その直後に，与えられた情報をもとに正規ユーザへのなりすましを試みる．これを，同じパス画像セットに対して，各被験者につき 5 回ずつ行ってもらった．なお，比較のために，不鮮明化する前のオリジナル画像を用いた認証システム(m＝10，n＝1)も用意した．

● 実験結果

実験の結果を表 2に示した．比較のために基本方式における言葉によるパス画像の漏洩実験（文献[3]の 3.4 節）の結果も一緒に示す．表中の用語は表 1と同じである．

表 2 言葉よる漏洩実験の結果

	基本方式	改良方式 オリジナル画像	改良方式 不鮮明化画像
成功率	37/50 (74%)	49/50 (98%)	30/50 (60%)
平均回答時間（秒）	10.91	2.71	7.26

改良方式を用いることで，2 択の認証システムであっても，言葉によるパス画像漏洩の成功率を 60%にまで抑えることができた．また，オリジナルの動画を利用した場合の結果と比較しても，不鮮明化パス画像の使用が，ストーリーの内容を知った上でのパス画像を推測するという攻撃を困難にしていることが分かる．

4.3 本人認証の実験

ストーリーによって，本当に複数のパス画像を記憶しやすくなるかどうかを，比較実験を通じて検証する．4.1節で用いた実験システムを微修正し，9択

* 具体的には，「女性が自転車に乗って，砂浜までの長い坂を下っている」，「女性が壁の反対側にいる人に向けてボールを投げている」などの情報を攻撃者に与えた．

の認証フェーズを 4 ターン行って 1 回の認証とするシステムを構築した．正規ユーザが記憶すべきパス画像の枚数は 10 枚のままであり，認証に使用されるパス画像は，認証の都度，10枚のパス画像セットの中から 4 枚がランダムに選択される．すなわち，m＝10，n＝4 である．このような実験システムを使用した理由は，基本方式の本人認証実験（文献[3]の 4.1 節で行われた実験）の結果と比較するためである．基本方式では，被験者（正規ユーザ）が 4 枚のパス画像を覚え，9 択×4 ターンの認証(m＝4，n＝4) を行っている．本節の実験と基本方式の実験の本人認証率を比較することにより，「ストーリーのスキーマを利用して 10 枚のパス画像を覚える際の記憶負荷」と「関連のないパス画像を 4 枚覚える際の記憶負荷」を対比することができる．

● 実験方法

パス画像登録後，1 日後と 8 日後に，各被験者につき 5 回ずつ認証を行ってもらう．

● 実験結果

実験の結果を表 3に示した．比較のために基本方式における本人認証の実験（文献[3]の 4.1 節）の結果も一緒に示す．表中，「成功率」は 10 人の各被験者につき 5 回ずつ行った認証試行の全体の成功率を表し，「ターンごとの平均回答時間」とは，ターンごとのパス画像選択に要した時間の平均である．

表 3 本人認証実験の結果

	基本方式		改良方式	
認証実施日	1 日後	8 日後	1 日後	8 日後
成功率	50/50 (100%)	49/50 (98%)	46/50 (92%)	45/50 (90%)
ターンごとの平均回答時間（秒）	8.19	7.10	11.69	11.30

1 日後，8 日後とも，基本方式の本人認証率がほぼ 100%であるのに対し，改良方式は認証率の低下，および，回答に要する時間の増加が見られる．しかしながら，ストーリーのスキーマによって 10 枚もの不鮮明化画像を覚えさせた場合にも，依然として 9 割以上の本人認証率が維持できていることが確認できた．

5. まとめと今後の課題

本稿では，スキーマを利用した画像認証方式をベースに，ユーザに m 枚のパス画像を記憶させた上で，その中の n 枚 (m＞n) のパス画像を用いて認証を行うという改良を基本方式に加え，覗き見攻撃耐性をさらに高めた認証方式へと改良した．また，この改良にあたり問題となる記憶負荷に対して，一本の動画から複数のパス画像を生成し，パス画像セットにストーリーを持たせるという方式による対策を提案した．

本人認証実験の結果においては，改良方式は依然として記憶負荷の点で改良の余地を残している．そのため，今後，ストーリーの与え方を工夫し，正規ユーザの記憶負荷を軽減する必要がある．

例えば，TV コマーシャルではなく，正規ユーザが自分で撮影した動画や写真を利用することが考えられる．正規ユーザは自分のエピソードと関連付けてパス画像セットを記憶することができるので，さらに記憶負荷を軽減することができると考えられる．本改良に対して，今後早急に，その有効性を検証していきたい．

謝辞　本研究は科研費（No.20-6290）の研究助成を受けている．また，本研究は一部，（財）セコム科学技術振興財団の研究助成を受けている．

参考文献

[1] Rachna Dhamija, Adrian Perrig, "Deja Vu: A User Study Using Images for Authentication", 9th USENIX Security

Symposium, pp.45-58, 2002.

[2] Real User Corporation, "PassFace", http://www.realuser.com/ (2009 年 1 月 確認).

[3] 原田篤史, 漁田武雄, 水野忠則, 西垣正勝：画像記憶のスキーマを利用したユーザ認証システム, 情報処理学会論文誌, Vol.46, No.8, pp.1997-2013, 2005.

[4] 太田信夫, 多鹿秀継 編著, "記憶研究の最前線", 北大路書房, 2001.

[5] W. F. Brewer: Schemata, In R. A. Wilson & F. C.Keil (Eds.), MIT Encyclopedia of the Cognitive Sciences, pp.729-730, 1999.

[6] 株式会社ニーモニックセキュリティ, ニーモニックガード, http://www.mneme.co.jp/neme/neme.html (2009 年 1 月確認)

(受付日：2009 年 1 月 26 日)
(受理日：2009 年 10 月 5 日)

著者略歴

山本 匠（やまもと・たくみ） 2006 年静岡大学情報学部情報科学科卒業. 2007 年 9 月同大学大学院修士課程修了. 現在, 同創造科学技術大学院博士課程, 日本学術振興会特別研究員 (DC). 情報セキュリティに関する研究に従事.

原田篤史（はらだ・あつし） 2001 年静岡大学情報学部情報科学科卒業. 2003 年同大学大学院修士課程修了. 2006 年同博士課程修了. 同年 三菱電機株式会社 情報技術総合研究所入社, 情報セキュリティに関する研究に従事.

漁田武雄（いさりだ・たけお） 1950 年生. 1976 年広島大学大学院教育学研究科博士課程後期中退. 同年広島大学教育学部助手. 1988 年国立特殊教育総合研究所研究員. 1982 年静岡大学教養部講師. 現在, 静岡大学情報学部情報社会学科教授. 文学博士. 人間の記憶の文脈依存機構の解明に関する研究に従事. 著書等としては「目撃証言と文脈依存記憶」(現代のエスプリ 350, 目撃者の証言：法律と心理学の架け橋 至文堂) 等がある. 日本心理学会会員, 日本認知心理学会会員, 日本基礎心理学会会員, アメリカ心理学会国際会員.

西垣正勝（にしがき・まさかつ） 1990 年静岡大学工学部光電機械工学科卒業. 1992 年同大学院修士課程修了. 1995 年同博士課程修了. 日本学術振興会特別研究員 (PD) を経て, 1996 年静岡大学情報学部助手. 1999 年同講師, 2001 年同助教授. 2006 年より同創造科学技術大学院助教授. 2007 年より准教授. 博士 (工学). 情報セキュリティ, ニューラルネットワーク, 回路シミュレーション等に関する研究に従事.