様式Ｃ－１９

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成　24 年　5 月　30 日現在

機関番号：１３８０１
研究種目：基盤研究(B)
研究期間：2009～2012
課題番号：21300021
研究課題名（和文）ネットワーク社会に有用な「覚え易く、忘れにくく、更に覗き見にも強い画像認証方式」
研究課題名（英文）An image-based user authentication that is easy to remember, difficult to forget, and robust against observing attack
研究代表者
西垣　正勝（NISHIGAKI MASAKATSU）
静岡大学・教授
研究者番号：20283335

研究成果の概要（和文）：
本研究は、人間の認知の特性である「スキーマ」の仕組みを利用することで、ユーザの記憶負荷軽減と覗き見攻撃の困難化という、矛盾する課題を両立させ得る画像認証方式を実現するものである。本方式の最大の特徴は、オリジナル画像の代わりに、オリジナル画像に対して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ことにある。人間は画像を記憶することに優れているが、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難しい。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。一方、正規ユーザにはパス画像登録時に不鮮明化画像とともにオリジナル画像を見ることが許される。これにより、正規ユーザの脳の中でオリジナル画像と不鮮明化画像の両者が関連付けられ、正規ユーザは不鮮明化されたパス画像の意味を簡単に認識することができる。本研究期間においては、特に、覗き見耐性の強化、囮画像の自動生成、利便性の向上の観点から、不鮮明化画像を用いた認証方式を深化させる。

研究成果の概要（英文）：
The image-based user authentication is effective since humans are significantly more efficient about recognition of previously seen images than precise recall of passwords. However, the use of images results in vulnerability against the observing attacks. To cope with these issues, we propose a user authentication using “unclear images” as pass-images. Even for humans it is hard to remember a meaningless image. That is why it is expected to be difficult for unauthorized users to memorize authorized user’s unclear pass-image. Only the authorized users are allowed to see the original images corresponding to their unclear pass-images in the enrollment phase. By seeing the original images, the authorized users can recognize the meaning in the unclear pass-images and can easily memorize and remember them by using the original images as a clue. This kind of clue is called as “schema” in cognitive psychology. In this research period, we tried to improve the unclear-images-based user authentication from the aspect of from the robustness against observing attack, automatic generation of the decoy images, and the enhancement of its usability.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
2009 年度	2,200,000	660,000	2,860,000
2010 年度	1,800,000	540,000	2,340,000
2011 年度	1,900,000	570,000	2,470,000
2012 年度	1,900,000	570,000	2,470,000
総　計	7,800,000	2,340,000	10,140,000

研究分野：総合領域
科研費の分科・細目： 情報学 ・ 計算機システム・ネットワーク
キーワード：ネットワークセキュリティ技術、ユーザ認証、画像認証、スキーマ、不鮮明化画像

１．研究開始当初の背景

組織内の機密データに対する不正アクセスや、それに伴うデータ破壊や情報流出などを防止するために、システム利用者の権限に応じたアクセス制御を徹底する必要があり、そのためのユーザ認証の確実な実施が重要となっている。

現在最も普及している認証方式は、汎用性の高さからパスワードが主流となっているが、人間にとって長くランダムな文字列を記憶することは容易ではない。そのため、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再認型の画像認証方式が注目されている。しかし、画像認証においては、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となる。

この問題に対し、申請者らは、有意味なオリジナル画像に対してモザイク化をはじめとする不鮮明化処理を施すことで一見して無意味な画像を作成し、それをパスワード画像として用いることで覗き見攻撃への耐性を有する画像認証方式を提案している。

正規ユーザは、パスワード画像登録時に不鮮明化処理を行う前のオリジナル画像を見ることができるため、不鮮明化されたパスワード画像をオリジナル画像の持つ意味と結びつけて記憶することができる。そのため、認証試行時には画面上に表示される自身の不鮮明なパスワード画像を容易に再認識することが可能となり、記憶の負荷は小さくなる。人間が有するこのような認知の特性は、認知心理学の分野で「スキーマ」と呼ばれる。

一方、オリジナル画像を見ることのできない他のユーザが当該ユーザの認証行為を覗き見たとしても、意味を認識できない不鮮明なパスワード画像を記憶に留めておくことは困難である。また、他者が正規ユーザからパスワード画像の意味を言葉で教えられたとしても、不鮮明なパスワード画像を教えられた意味どおりに認識することは困難であるため、本方式は正規ユーザからのパスワード漏洩に対してもあるレベルの耐性を有する。

しかし、たとえ不鮮明な画像であっても認識すること自体は不可能ではないため、カメラで盗撮しておいた不鮮明化画像と同じものを選ぶという攻撃には脆弱性が残る。最近はATMへの盗撮カメラ設置の事件も発生していることから、カメラ撮影を用いた覗き見にも一定レベルの耐性を有することが望まれる。

また、画像認証においては、パス画像を隠すために利用される囮画像（認証画面にパス画像と共に表示される複数の画像）を適切に更新していくことも重要な課題の１つである。正規ユーザにとって馴染みの無い新しい囮画像を適切に用意することができなければ、画像認証の安全性や正規ユーザの利便性の低下につながってしまう。

２．研究の目的

前述のとおり、画像認証には、「覗き見の問題」と「囮画像の問題」の２つの本質的な課題が残されており、現在までのところ正規ユーザの負荷を抑えたまま両課題を克服した研究は確認されていない。そこで本研究では、画像の記憶・再認の優位性は残したまま、両課題を同時に解決する新しい画像認証の実現を目指す。

その際、人間の計算能力の限界に鑑みるに、計算量的な安全性を礎とする暗号学的なアプローチによって画像認証を改良することは基本的に困難であると考えられる。このため本研究では、認知心理学的な観点からのアプローチ、すなわち、不鮮明化画像の特長に着目して研究を進めていく。

さらに、画像認証技術に対する認知心理学の知見の活用を加速することによって、ユーザ認証の利便性をリバレッジし、利便性と安全性を両立するような画像認証を模索する。

３．研究の方法

（１）覗き見耐性の向上：

以下の２つの方法によって、不鮮明化画像を用いたユーザ認証に対する覗き見耐性の強化を検討する。

①言語手がかり付き再認方式（RVC 方式）：

覗き見攻撃耐性の強化のためには、認証の度異なるパス画像セットを利用できることが望ましい。そこで本方式では、正規ユーザに m 枚のパス画像を記憶させた上で、1 回の本人認証にあたって n 枚（m＞n）のパス画像を用いて認証を行うという運用（m-n 対策）を導入する。しかし、正規ユーザに一回の認証に必要なパス画像の枚数よりも多くのパ

ス画像を記憶させることは、正規ユーザの負荷の増大につながる。そのため、m-n 対策の導入に対してはユーザ負荷増大の緩和対策が必須となる。

そこで本方式では、パス画像を思い出すにあたっての手がかりとなる言語情報を認証時に提示する（RVC 方式：Recognition with Verbal Cue）ことにより、m-n 対策を導入したときのユーザ負荷の増大を抑制する。RVC 方式では、スキーマを有する正規ユーザは、手がかり情報によって認証時の再認および想起の促進が期待される。ここで、画面に提示される手がかりは正規ユーザだけでなく、覗き見攻撃者にも与えられることになる。しかし、「不鮮明なパス画像に関する情報を言葉で与えただけでは，スキーマを持たない攻撃者はパス画像を特定するに足る情報量を得ることができない」という不鮮明化画像の特長から、認証時にパス画像に対する手がかり情報を言葉で与えたとしても、スキーマを持たない攻撃者にはその情報を有効に活用できないことが予想できる。

②暗示・応答型画像認証方式（Q&R 方式）：

再認型画像認証のチャレンジ＆レスポンス（以下、C&R）化を試みる。ただし、人間の計算能力には限界があるため、チャレンジからのレスポンス生成に暗号計算が必要となる通常の C&R 型のユーザ認証プロトコルを適用することは不可能である。そこで、チャレンジの意味を隠す（正規ユーザ本人にしかチャレンジを理解することができないようにする）というアプローチを採ることによって、認知心理学的にチャレンジを隠蔽（正規ユーザ本人にしか伝わらないようにする）することで、簡素なレスポンス生成処理を採用した場合であっても、あるレベルの安全性が担保される C&R 型画像認証を実現する。カメラを用いた覗き見であっても、攻撃者はチャレンジの意味が分からず、パス画像の特定が困難となる。

具体的には、「不鮮明なパス画像に関する情報を言葉で与えただけでは、スキーマを持たない攻撃者はパス画像を特定するに足る情報量を得ることができない」という不鮮明化画像の特長を活用し、認証の度に異なる質問（チャレンジ）を正規ユーザのみが理解することができる形で提示する。本方式においては、認証の度に異なる質問が正規ユーザのみが知覚できる形で提示される。

本研究では、このような暗示的なチャレンジを「キュー（Cue）」と呼び、本方式を暗示・応答型画像認証方式（以下、Q&R（Cue & Response）方式）と表現する。今回は、動物の画像をオリジナル画像として、これを不鮮明化した画像をパス画像として用いることとした。キューは、パス画像に対する部位情報を言語手がかりで示したもの（例えば、「左目」、「右耳」、「尻尾」、「左前足」等）であり、認証の度に変化する。ユーザは、囮画像に紛れているパス画像を見つけた上で、キューによって指示された部位に対応する場所をクリックすることによってレスポンスを返す。

スキーマを持たない攻撃者には、不鮮明化画像の意味を認識することは困難であるため、指示された部位に対応する場所を正しくクリックすることは容易なことではない。一方、不鮮明化画像の意味（スキーマ）を知っている正規ユーザであれば、指示された部位をクリックすることは容易である。

キューにより指定される部位は認証の度に変化する（ある認証フェーズで「左耳」をクリックしている瞬間を覗き見られたとしても、次回の認証においては例えば「左前足」という指示に変わる）ため、覗き見攻撃に対する耐性も増加する。また、この方法は、パス画像選択の総当たり数を増やすことを可能にするというメリットもある。

（２）囮画像の自動生成（ADG 方式）：

不鮮明化画像αに対応するオリジナル画像をユーザに見せてスキーマを獲得してもらった後に、その不鮮明化画像αを 90 度回転させた画像をユーザに見せても、ユーザはそれが不鮮明化画像αであるということに気付かない。言い換えると、「簡素な画像処理によって、オリジナル画像と不鮮明化画像の間のスキーマを切断することが可能である」という特長が不鮮明化画像には存在する。

この不鮮明化画像の特長を活用すれば、正規ユーザが記憶しているパス画像や正規ユーザにとって馴染みの深い画像からでも、正規ユーザがパス画像との混乱をきたすことの無い囮画像を生成することが可能だと考えられる。本研究では、このアイデアに基づいて、従来の写真や絵（オリジナル画像）を利用する画像認証方式では実現不可能な方法で、囮画像を生成する方式（ADG 方式：Automatic Decoy image Generation）を構築する。

ここでは、具体的に以下の 3 つの ADG 方式を検討する。

①囮画像生成法 1：

2 枚のオリジナル画像 A と B のそれぞれ上半分と下半分をつなげた上で、不鮮明化処理を施すことによって囮画像（不鮮明化画像）を生成する。

②囮画像生成法 2：

正立したオリジナル画像 B に回転したオリジナル画像 A を同じ割合で重ね合わせた上で、不鮮明化処理を施すことによって囮画像（不鮮明化画像）を生成する。

③囮画像生成法 3：

囮画像生成法 1 の要領で 2 種類のオリジナ

ル画像を作成し、その 2 種類を囮画像生成法 2 の要領で重ね合わせる。最後に不鮮明化処理を施すことによって囮画像（不鮮明化画像）を生成する。

（3）エンターテイメント性の付加：
不鮮明化画像の特徴を利用することによって、画像認証にエンターテイメント性を付加する方法を検討する。エンターテイメント性の付加によって、ユーザが楽しんで認証試行を行うことができるような画像認証方式を模索する。これによって、ユーザ認証における安全性と利便性の両立を目指す。

4．研究成果

（1）覗き見耐性の向上：

①言語手がかり付き再認方式（RVC 方式）：
RVC 方式の有効性を、本人認証および覗き見の実験により評価した。被験者は本学情報系学部学生 10 名である。

・本人認証実験
m-n 対策の導入（正規ユーザに m 枚のパス画像を記憶させた上で、1 回の本人認証にあたって n 枚（m＞n）のパス画像を用いて認証を行う）によって増大されるユーザの負荷が、手がかりの提示によってどの程度抑えることができるのかについて、本人認証率の比較を通じて検証した。使用した画像は 300×300 pixel である。比較のために、既存の不鮮明化画像方式、および、RVC 方式から手がかりの提示を除去したシステム（以下、比較方式 1 と呼ぶ）を構築して、RVC 方式と合わせて実験を行った。

表 1．本人認証実験結果

	不鮮明化画像方式（m＝4, n＝4, 手がかり＝無）		比較方式 1（m＝10, n＝4, 手がかり＝無）		RVC 方式（m＝10, n＝4, 手がかり＝有）	
	1 日後	8 日後	1 日後	8 日後	1 日後	8 日後
認証成功率	100 %	98 %	90 %	82 %	100 %	98 %

表 1 に実験結果を示す。1 日後、8 日後とも、RVC 方式の本人認証率は、不鮮明化画像方式と同様、100%を維持している。比較方式 1（手がかり無）の認証率が低下している事実より、手がかり情報を用いることで、認証画面に表示されているパス画像に対するスキーマの想起が促進され、認証成功率の低下が抑えられたのだと推測できる。

・覗き見実験
攻撃者が正規ユーザの認証作業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ないため、m-n 対策の導入が覗き見攻撃耐性を大きく向上させることは容易に想像ができる。しかし、RVC 方式では、m-n 対策の導入とともに、言語手がかりの提示を行っている。スキーマを有していない攻撃者は言語手がかりを十分活用することはできないと考えられるが、言語手がかりの提示が覗き見攻撃の脅威をどれほど増加させてしまうかについて実験により確認する必要がある。比較のために、RVC 方式と既存の不鮮明化画像方式に対して実験を行った。なお、覗き見実験では、攻撃者になるべく有利になるように、囮画像は 1 枚のみ（m=2、n=1）としている（鮮明な画像を用いる既存の画像認証の場合は、同様の実験環境での覗き見成功率はほぼ 100%となることに注意のこと）。
表 2 に実験結果を示す。RVC 方式のなりすまし成功率が不鮮明化画像方式のなりすまし成功率よりも低いことから、m-n 対策の覗き見に対する効果に認められた形となる。

表 2．覗き見実験結果

	不鮮明化画像方式	RVC 方式
成功率	92.0%	78.0%

②暗示・応答型画像認証方式（Q&R 方式）：
Q&R 方式の有効性を、本人認証および覗き見の実験により評価した。被験者は本学情報系学部学生 10 名である。

・本人認証実験
正規ユーザにとって、参照画像から与えられるキュー（パス部位）を正しく認識し、パス画像中のパス部位を的確に選択することが可能かどうかを、本人認証率の比較を通じて検証した。使用した画像は 300×300 pixel であり、被験者がパス部位の中心から半径 10 pixel の円内をクリックできた場合に認証成功とした。
表 3 に実験結果を示す。キューによって暗示されたパス部位を正しく認識し、パス画像上のパス部位を正確に選択できた割合は、1 日後、8 日後とも 90%以上であり、本方式の記憶負荷がそれほど高くないことが見て取れる。

表 3．本人認証実験結果

	1 日後	8 日後
認証成功率	93.0%	90.0%

・覗き見実験

攻撃者が過去に覗き見したキュー（参照画像上のパス部位）に対するレスポンス（正規ユーザがクリックした画像とその位置）の情報を用い、現在表示されているキュー（パス部位）に対するレスポンスを推測することが難しいかどうかを確認するために、認証試行を攻撃者に複数回覗き見されたことを想定した覗き見攻撃実験を行った。

表 4 に実験結果を示す。覗き見回数が増加するにつれて攻撃成功率も増加しており、覗き見情報無しでは約 3 割、覗き見情報 3 個では約 4 割程度の攻撃成功率であった。パス部位という秘密情報が提示される分、Q&R 方式の覗き見攻撃耐性は不鮮明化画像方式よりも当然高くなっている。なお、覗き見実験では、攻撃者になるべく有利になるように、囮画像は使用せずにパス画像のみを表示している（鮮明な画像を用いる既存の画像認証の場合は、同様の実験環境での覗き見成功率はほぼ 100%となることに注意のこと）。

表 4. 覗き見実験結果

	覗き見情報の数			
	0 個	1 個	2 個	3 個
成功率	31.0%	39.5%	44.0%	43.0%

（2）囮画像の自動生成（ADG 方式）：

ADG 方式により生成された囮画像の有効性を確かめるために、攻撃者にとって囮画像（本研究にて提案した囮画像生成法 1～3 によって生成された不鮮明化画像）とパス画像（オリジナル画像を不鮮明化することによって生成された不鮮明化画像）の見分けが困難であるか否かを測る識別実験を行った。使用した画像は 300×300 pixel である。被験者は本学情報系学部学生 10 名である。

実験の結果を表 5 に示した。いずれの囮画像生成法においても、被験者は 60%前後の割合で自然不鮮明化画像と加工不鮮明化画像との違いを認識していることが見てとれる。被験者が完全に当て推量で回答した場合の識別率が 50%であるため、囮画像生成法 1～3 によって作成された不鮮明化画像は概ね「オリジナル画像を不鮮明化することによって生成された不鮮明化画像」と同等の画像となっていると考えてよいと判断できる。

表 5. 識別実験結果

	囮画像生成法 1	囮画像生成法 2	囮画像生成法 3
識別成功率	58.5%	60.5%	57.5%

（3）エンターテイメント性の付加：

不鮮明化画像においては、「不鮮明度を上げるほど、不正者による覗き見に強くなるが、正規ユーザも認識が難しくなる」という不鮮明化の強度に対するトレードオフの問題が存在する。これを逆に利用し、ロールプレイングゲームに見られるゲームキャラクタの「レベルアップ」の概念を不鮮明化画像に導入することを提案した。

まず、不鮮明化画像の利用に関して初心者のユーザに対しては、不鮮明度の低い画像を与える。不鮮明度が低い（オリジナル画像に近い）画像であれば、初心者ユーザであっても、不鮮明化画像を用いたユーザ認証を容易に利用することができると期待される。そして、ある程度の期間が経過して、そのユーザが不鮮明化画像の認識に対して「慣れ」が生じてきた場合には、不鮮明度の高い不鮮明化画像の利用を認めるようにする。

このように、「不鮮明化画像に対する認知能力が高まったユーザほど、不鮮明度の高い画像の利用を許す」という運用を採り入れることによって、正規ユーザに不鮮明度の強い画像を利用することに対するインセンティブを与えることができると期待される。

また、不鮮明化画像だけでなく、通常のオリジナル画像を利用したエンターテイメント性の醸成の方法や、画像認証だけなく、CAPTCHA（通信相手が人間であるのかマルウエアであるのかを見分ける技術）に対してエンターテイメント性を導入する方法についても検討を行った。

５．主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 6 件）

① Junya Kani, Masakatsu Nishigaki: Gamified CAPTCHA, Proceedings of 2013 International Conference on Human-Computer Interaction, 査読有, 2013, 採録決定

② T. Yamamoto, Y. Kojima, M. Nishigaki : shoulder-surfing-resistant image-based authentication system with temporal indirect image selection, Proceedings of 2009 International Conference on Security & Management, 査読有, 2009, pp. 188-194,

③ T. Yamamoto, A. Harada, T. Isarida, M. Nishigaki : Advantages of user authentication using unclear images: automatic generation of decoy images, Proc. IEEE International Conference on Advanced Information Networking

and Applications 2009, 査読有, 2009, pp. 668-674

④ 山本匠, 漁田武雄, 西垣正勝：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, 日本セキュリティ・マネジメント学会誌, 査読有, 2009, vol.23, no.3, pp. 41-47

⑤ 山本匠, 漁田武雄, 西垣正勝：不鮮明化画像が実現するスキーマを利用した画像認証方式の改良, 日本セキュリティ・マネジメント学会誌, 査読有, 2009, vol.23, no.3, pp. 17-29,

⑥ 山本匠, 漁田武雄, 西垣正勝：不鮮明化画像を利用した暗示・応答型画像認証方式の提案, 情報処理学会論文誌, 査読有, 2009, vol.50, no.9, pp.2062-2076

〔学会発表〕(計８件)

① 西垣正勝：エンターテイメントセキュリティとメディア処理, 映像情報メディア学会メディア工学研究会(2012.12.12) 埼玉工業大学, (埼玉) (招待講演)

② Junya Kani, Harunobu Agematsu, Masakatsu Nishigaki：Gamified CAPTCHA, International Workshop on Security 2011 (Poster presentation) (2011.11.8), 九州大学西新プラザ (福岡)

③ 池谷勇樹, 可児潤也, 米山裕太, 西垣正勝:3DCG を用いたチャレンジ画像生成プラットフォームの提案～適用例１：画像 CAPTCHA～, コンピュータセキュリティシンポジウム 2012 (2012.11.1). くにびきメッセ (島根)

④ 西垣正勝：エンターテイメントセキュリティ, コンピュータセキュリティシンポジウム, (2011.10.21), 朱鷺メッセ (新潟)

⑤ 高田愛美, 鈴木徳一郎, 山本匠, 西垣正勝：興味対象への視線誘引を利用したアンコンシャス視線認証の精度改善に関する研究, 情報セキュリティ心理学とトラスト研究グループ研究会, (2011.2.18), キャンパスプラザ京都 (京都)

⑥ 高田愛美, 鈴木徳一郎, 山本匠, 西垣正勝：視線誘導型なりすまし検知方式の検討 (その２), マルチメディア, 分散, 協調とモバイルシンポジウム, (2010.7.7), 下呂温泉水明館 (岐阜)

⑦ 山本匠, 漁田武雄, 西垣正勝：不鮮明化画像と言語手がかりを用いた Challenge & Response 型画像認証の実現に対する試み, 映像メディア学会メディア工学研究会, (2009.12.4), 長野県工業技術総合センター (長野)

⑧ 山本匠, 西垣正勝：利便性と安全性を兼ね備えた画像認証方式の実現に向けて, 高度情報シンポジウム 2009, (2009.11.12), 静岡大学 (静岡)

〔その他〕
ホームページ等
http://minamigaki.cs.inf.shizuoka.ac.jp/index.html

６．研究組織
(1)研究代表者
西垣　正勝 (NISHIGAKI MASAKATSU)
静岡大学・創造科学技術大学院・教授
研究者番号：2028335
(2)研究分担者
なし
(3)連携研究者
なし