Vol.2015-IS-131 No.1
2015/3/6

情報漏えいリスクを低減するアカウント管理手法

大谷　和也[1,a)]　柿崎　淑郎[1,b)]　佐々木　良一[1,c)]

概要：現在，パスワードリスト攻撃が流行している．一方でユーザーは多くのパスワードを覚えきれず，パスワードの使い回しが頻発している．また，ユーザーが複数の Web サイトで同じパスワードを使い回していることから，ユーザーが記憶できるパスワードの数には限界がある．そこでパスワードの使い回しを前提とした情報漏えいリスクのコントロールを目的としたアカウント管理手法について述べる．本提案手法を使うことによって，パスワードリスト攻撃を受けて，情報漏えいしても，限定された情報しか漏えいしない，さらに，無配慮にパスワードを使い回すよりは，ユーザーが記憶すべきパスワードの数を減らしつつ，情報漏えいリスクを低減できる．

1. はじめに

現在，パスワードリスト攻撃が流行している．パスワードリスト攻撃とは，図 1 に示すように，悪意ある攻撃者が大量の ID とパスワードを組み合わせたリストを使い，Web サイトへ不正にログインを試みる不正アクセスの一種である．また，IPA で公開されているオンライン本人認証方式の実態調査によると，パスワードリスト攻撃の認知件数が平成 24 と平成 25 年を比べると 7 倍にも増加しており，パスワードリスト攻撃は激化している [5].

パスワードリスト攻撃の対策としては，Web サイトごとに異なる ID とパスワードの設定，パスワードの定期変更，二要素認証の導入などがある．野村総合研究所の調査でユーザーが Web サービスを使う際に確実に覚えられる ID とパスワードの組み合わせの数は，平均 3．1 個という結果が出ている [4]．そのため，複数の ID とパスワードの組み合わせを覚えきれなかったり，それらの記憶にかかるコストが高いなどユーザーの負担が大きいという問題がある．

エムオーテックの調査により 3 人に 1 人は同じパスワードを設定し，複数の Web サイトで利用しているという結果が出ている [3]．そのため，ユーザーは同じパスワードを複数の Web サイトで使い回すことが多い．図 1 に示すように，ユーザーが複数のサイトに同じパスワードを使い回している状況で，パスワードリスト攻撃を受けて，攻撃者が不正ログインに成功してしまうと，他の Web サイトでも同様に不正ログインが成功してしまう恐れがある．そのため

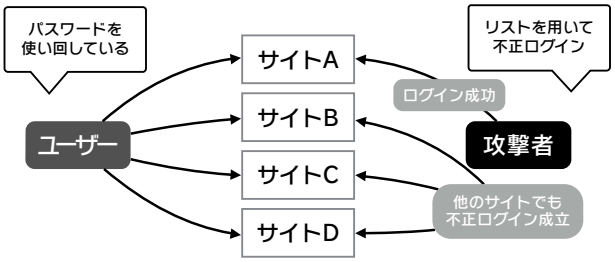

図 1　パスワードリスト攻撃

本研究では，同じパスワードを設定しているアカウントについては，1 つのアカウントが不正ログインされた場合は，他のサイトも同様に不正ログインされることを仮定する．

本稿では，パスワードの使い回しを前提にした情報漏えいリスクを低減するアカウント管理手法について述べる．本提案手法では，パスワードリスト攻撃による情報漏えいリスクをコントロールし，被害を最小限に抑えることを目的とする．ここでの情報漏えいリスクとはパスワードリスト攻撃を受け，不正ログインをされたときに個人情報が漏えいしてしまうリスクのこととする．本提案手法はアカウントに登録されている情報に基づいて，類似の情報を保持しているアカウント同士でグループ分けを行う．このグループは類似の情報しか保持していないため，不正ログインによって情報漏えいが発生しても，限定された情報しか漏えいしない．よって，このグループには同じパスワードを設定しても情報漏えいリスクは増加しないため，ユーザーが記憶すべきパスワードの数を減らしつつ，情報漏えいリスクを低減できる．

2. 関連研究

Florêncio らは，パスワードについての研究文献の多くが

[1] 東京電機大学
Tokyo Denki Uniersity
[a)] 11fi024@ms.dendai.ac.jp
[b)] kakizaki@im.dendai.ac.jp
[c)] sasaki@im.dendai.ac.jp

Web サービスの安全確保やパスワードの基準の設定において実際には役に立たないことから，パスワードの耐久性，強力なパスワードの耐久性の基準を現実的に調査した [1]．その結果，強力なパスワードの作成は記憶するコストに対して，得られるメリットが少ないということがわかった．文献 [1] ではユーザーに対して重要な情報を保有していないアカウントは手間をかけず，重要な情報を保有しているアカウントに焦点を絞ると良いと述べられている．

また，安全なパスワードについて，以下の 2 つが通説とされている [2]．

(1) パスワードはランダムで強固なものであるべきだ．
(2) パスワードはアカウント全体で再利用するべきではない．

しかし，実際にはユーザーはこれを満たすことはできておらず，複数の強固なパスワードを記憶する決定的な方法はない．そこで文献 [2] では，パスワード管理におけるユーザーの労力と情報漏えいしたときに予想される被害を最小化するために，アカウントのグループ分けを行い，それらにパスワードの使い回しを許容としたときに，どのような影響があるかを調査をした．その結果，パスワードの使い回しを許容するとパスワードは強固になり，パスワードの使い回しを許容しないとパスワードは強固ではなくなり，パスワードの使い回しの許容とパスワードの強度とにトレードオフの関係であることが分かった．また，グループ分けをしてパスワードの使い回しを許容した場合，情報漏えいの確率は上がるが予想される被害は低くなることが示された．Florêncio らは，ランダムで強固なパスワードを大量に管理することは実際的には困難であり，脆弱なパスワードまたはパスワードの使い回しを許容しないパスワード管理手法は最適ではないと主張している．

3. アカウント情報に関する調査

3.1 概要

まず，アカウントにログインした場合どのような情報が確認できるか，つまり，攻撃者が不正にログインした場合に漏えいする情報を調査する．この調査により，どのアカウントがどのような情報を保有しており，また，どのような情報が漏えいする恐れがあるかを明らかにする．

3.2 方法

確認できる情報を調査するため，実際のサイトでアカウント登録を行う．調査したサイトは Alexa Top 500（2015 年 1 月 5 日時点）より，日本からのアクセスが多い上位 50 サイトを対象とした．調査の一例として，Amazon のアカウント登録ページを図 2 に，会員情報ページを図 3 に示す．またここで，本稿で用いる用語を以下のように定義する．

登録されている情報

登録されている情報とは，アカウント作成時に必須とされている項目とする．例えば，図 2 における赤枠で囲われている名前，フリガナ，E メールアドレス，パスワードが該当する．なお，アカウント作成時に要求されるが必須ではない項目については，**登録されていない情報**とした．

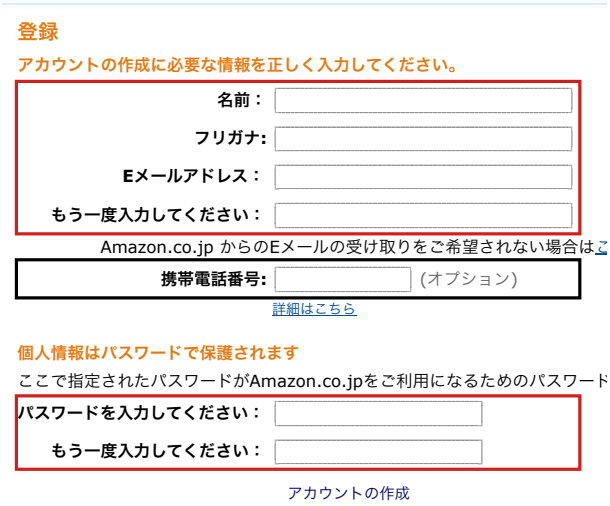

図 2 Amazon の新規登録ページ

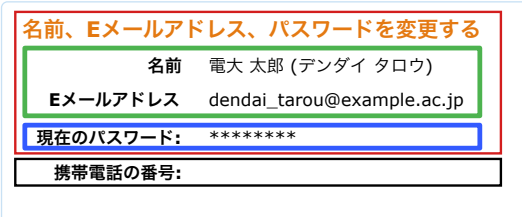

図 3 Amazon の名前，E メールアドレス，パスワード変更ページ

登録されていない情報

登録されていない情報とは，アカウント作成時に要求されない情報および，必須ではない項目とする．例えば，図 2 における黒枠で囲われている携帯電話番号が該当する．また，**登録されていない情報**は**確認できない情報**である．

確認できる情報

確認できる情報とは，マイページ等にログインした際に確認することができる情報であり，「登録されている情報」である．アカウントには様々な情報が記録されているが，その全てを確認できるわけではない．例えば，マイページなどにおいて，ユーザー名のように表示される情報の他にも，変更は可能だが表示されない情報などがある．このうち，ログイン後には確認することができる情報を**確認できる情報**とする．例えば，図 3 における緑枠で囲われている名前，E メールアドレスの情報が該当する．

確認できない情報

確認できない情報とは，登録されていない情報および登録されている情報のうちマイページ等で確認することができない情報のことである．例えば，図 3 におけるパスワード，携帯電話の番号が該当する．

表 1 調査したサイト（一部抜粋）

サイト名	ユーザー名	名前	メールアドレス	携帯番号	国	性別
Google	○	○	○	○	○	△
楽天	○	○	○	×	×	×
Amazon	○	○	○	×	×	×
mixi	○	○	○	×	×	○

3.3 結果

調査した 50 サイトは，検索エンジンサイト，ショッピングサイト，ブログサイト，SNS，グルメサイトなど多種多様のサービスを提供しているサイトからなる．調査結果として各サイトにおける**確認できる情報**の一部抜粋を表 1 に示す．表 1 中の○が**確認できる情報**，△が**確認できない情報**，×が**登録されていない情報**である．

4. 提案手法

4.1 概要

類似の情報を元に同じパスワード設定した場合を図 4 に，無配慮に同じパスワード設定した場合を図 5 に示す．図 4 のように，サイト A，C に同じパスワードを設定し，パスワードリスト攻撃をされた場合，名前，性別，生年月日が攻撃者に漏えいする．一方で図 5 のように，サイト B，C に同じパスワードを設定し，パスワードリスト攻撃をされた場合，名前，郵便番号，住所，性別，生年月日が攻撃者に漏れる．前者と後者を比較すると，前者の方が漏えいする情報が少ないことがわかる．

本提案手法では，このように**確認できる情報**に基づいて，類似したアカウント同士をグループ化することで，グループ化されたアカウントに同じパスワードを設定したとしても，漏えいする情報を最小限にすることを目指す．

ここで，本研究が目指す達成要件を挙げる．

要件 1 パスワードリスト攻撃を受けても漏えいする情報を最小限とする
無配慮に複数のサイトで同じパスワードを使い回すよりも，漏えいする情報を最小限にすることで，情報漏えいリスクを低減する．

要件 2 記憶すべきパスワードの個数を極力減らす
ユーザーの記憶できるパスワードの個数は限られているため [4]，全てのアカウントに異なるパスワードを設定することは困難である．よって，情報漏えいリスクを高めることなく，記憶すべきパスワードの個数が少なくなることが望ましい．

4.2 アカウントのクラスタリング

類似の情報を持つ Web サイトの組み合わせを得るために確認できる情報を用いて，それぞれのアカウントの類似度を算出し，クラスタリングを行う．

クラスタリングとは，ある集合を似たような特徴を持ついくつかの集合に分けるものであり，分けられた集合のことをクラスタと呼ぶ．

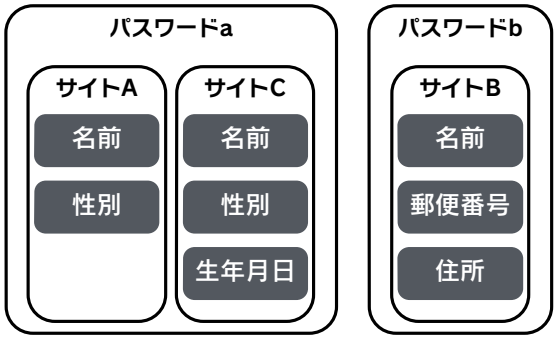

図 4 類似の情報を元に同じパスワードを設定した場合

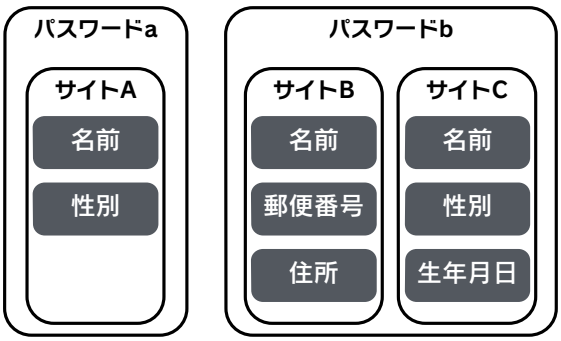

図 5 無配慮に同じパスワードを設定した場合

表 2 サイト x，y の確認できる情報，確認できない情報

サイト名	ユーザー名	名前	メールアドレス	電話番号	性別
サイト x	1	0	1	0	1
サイト y	1	0	1	1	1

アカウントのクラスタリングをすることにより，類似の情報を持つ Web サイトの組み合わせを得ることができる．

4.2.1 類似度

類似の情報を持つ Web サイト同士に分けるために，アカウント同士の類似度を計算する．類似度の算出にはコサイン類似度を使う．コサイン類似度の計算式は以下の通りである．

$$\cos\theta = \frac{x \cdot y}{|x||y|} \tag{1}$$

式 1 の x と y は**確認できる情報**と**確認できない情報**を要素としたベクトルである．また，コサイン尺度なので，1 に近ければ類似度が高いといえる．

計算例として，サイト x，y の**確認できる情報**，**確認できない情報**を表 2 に示す．表 2 中の 1 が**確認できる情報**，0 が**確認できない情報**である．

まず，表 2 からサイト x，y の行毎にベクトルとする．

$$x = (1, 0, 1, 0, 1)$$
$$y = (1, 0, 1, 1, 1)$$

次に，ベクトル x，y を式 1 に代入し，コサイン類似度を

算出する．

$$\cos\theta = \frac{(1\times1+0\times0+1\times1+0\times1+1\times1)}{\sqrt{1^2+0^2+1^2+0^2+1^2}\sqrt{1^2+0^2+1^2+1^2+1^2}} = \frac{1+0+1+0+1}{\sqrt{3}\sqrt{4}} = \frac{3}{\sqrt{12}} = 0.866$$

4.2.2 クラスタリング

類似の情報を持ったアカウント同士でグループ分けを行うため，k-means 法でクラスタリングを行う．k-means 法の手順は以下の通りである．

(1) ある集合に対して，無作為に中心点を k 個作成する．この k がクラスタ数となる．
(2) 各データを中心点に一番近いもの同士でグループ分けする．
(3) 各中心点をグループの中心に移動する．
(4) 中心点が動かなくなるまで，手順 2，3 を繰り返す．

5. 実験

5.1 目的

本提案手法によって，**要件 1**「パスワードリスト攻撃を受けても漏えいする情報を最小限とする」を満たすかどうかを検証するために，ユーザーが実際に利用しているアカウントを利用して実験を行う．

5.2 方法

ユーザーによって，所有しているアカウントが異なるため，ユーザーに合わせてアカウントを選定しなければならない．本実験では調査した 50 サイトの中から第一著者が所有しているアカウント 17 個を用いる．この 17 サイトに対して，本提案手法を適用し，クラスタリングを行う．ここで，クラスタ数 k は野村総合研究所の調査 [4] より，$k = 3$ とした．

5.3 結果

図 6 に $k = 3$ のクラスタリング結果を示す．図 6 中の 1，2，3 と番号が振られた 3 つの楕円をそれぞれ，クラスタ 1，クラスタ 2，クラスタ 3 とする．

表 3 にクラスタ毎のアカウント内訳を示す．また，表 4，5，6 に，各クラスタに属するアカウントがどのような情報を保持しているかを示す．表 4，5，6 中における○は**確認できる情報**，×は**確認できない情報である**．

6. 考察

各クラスタが保持している情報から，クラスタリング結果が適切であるか，つまり，**要件 1**「パスワードリスト攻撃を受けても漏えいする情報を最小限とする」が達成されているかを考察する．ここでの評価基準は，各クラスタに対して同じパスワードに設定することを許容できるかどうか．

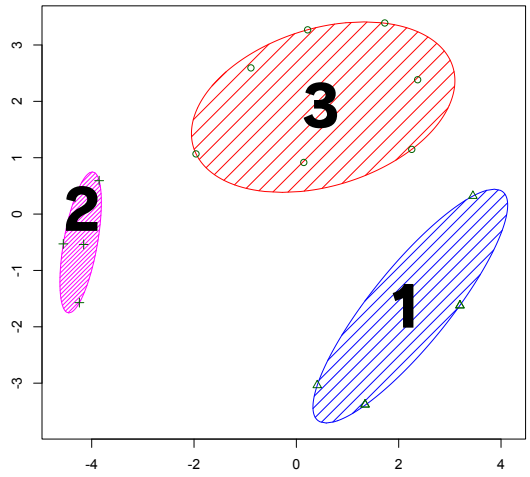

図 6 $k = 3$ のクラスタリング結果

表 3 $k = 3$ におけるクラスタ毎のアカウント内訳

クラスタ 1	Google	楽天	Amazon
	DMM.com	Twitter	Instagram
クラスタ 2	ヨドバシ.com	Apple ID	Biglobe
	クロネコメンバーズ		
クラスタ 3	yahoo	mixi	Ameba
	Microsoft	ニコニコ動画	pixiv
	FC2		

考察の前提として，個人を特定するに至る重要な情報として**機微情報**を設定する．ここでは，「電話番号」「携帯番号」および都道府県，市区町村，番地までを含む「住所」を機微情報とする．ただし，クラスタに属する全てのアカウントが同じ機微情報を保持していた場合は，同じ情報漏えいリスクであると見做して，例外とする．

6.1 クラスタ 1 の考察

表 4 のクラスタ 1 が保持している情報より，クラスタ 1 のそれぞれのアカウントに同じパスワードを設定し，パスワードリスト攻撃をされた場合，ニックネーム，名前，ユーザー名，携帯番号，メールアドレス，国情報が漏えいすることが分かる．各アカウントが保持している情報に着目すると，全てのサイトがユーザー名とメールアドレスを持っていて，名前を持っているのが Google と楽天と Amazon，ニックネームを持っているのが Twitter，携帯電話と国情報を持っているのが Google ということが分かる．

ここで，Google は機微情報である「携帯番号」を保持しているため，クラスタ内の他のアカウントと同じパスワードを設定してしまうと，機微情報が漏えいしてしまう．したがって，Google をクラスタ 1 から除外し，独立したパスワードを設定した方が良いと考えられる．また，名前とニックネームについては情報漏えいしたとしても，大きなリスクにはならない．

よって，クラスタ 1 から Google を外すことで，クラスタ 1 に属するアカウントから情報漏えいしても，限定され

表 4 クラスタ 1 が保持している情報

サイト名	ニックネーム	名前	ユーザー名	携帯番号	メールアドレス	国
Google	×	○	○	○	○	○
楽天	×	○	○	×	○	×
Amazon	×	○	○	×	○	×
DMM.com	×	×	○	×	○	×
Twitter	○	×	○	×	○	×
Instagram	×	×	○	×	○	×

表 5 クラスタ 2 が保持している情報

サイト名	名前	ユーザー名	生年月日	性別	電話番号	郵便番号
ヨドバシ.com	○	○	○	×	×	○
Apple ID	○	○	○	×	×	×
Biglobe	○	○	○	○	×	○
クロネコメンバーズ	○	○	×	○	○	○

サイト名	メールアドレス	携帯メールアドレス	秘密の質問	使用言語
ヨドバシ.com	○	×	×	×
Apple ID	○	×	○	○
Biglobe	○	×	×	×
クロネコメンバーズ	○	○	×	×

サイト名	都道府県	市区町村	番地	アパート・マンション・ビル
ヨドバシ.com	○	○	○	○
Apple ID	○	○	○	○
Biglobe	○	○	○	×
クロネコメンバーズ	○	○	○	○

た情報しか漏えいしないため，情報漏えいのリスクを低減することができる．

6.2 クラスタ 2

表 5 のクラスタ 2 が保持している情報より，クラスタ 2 のそれぞれのアカウントに同じパスワードを設定し，パスワードリスト攻撃をされた場合，名前，ユーザー名，生年月日，性別，電話番号，郵便番号，メールアドレス，携帯メールアドレス，秘密の質問，使用言語，都道府県，市区町村，番地，アパート・マンション・ビルが漏えいすることが分かる．また，全てのサイトが名前，ユーザー名，メールアドレス，都道府県，市区町村，番地を保持している．ここで，全てのサイトが機微情報である「住所」を持っているため，これらのどのアカウントから情報漏えいが発生しても同じリスクであるため，情報漏えいリスクを高める要因にはならない．

クロネコメンバーズは機微情報である「電話番号」を保持しているため，クラスタ内の他のアカウントと同じパスワードを設定してしまうと，機微情報が漏えいしてしまう．したがって，クロネコメンバーズをクラスタ 2 から除外し，独立したパスワードを設定した方が良いと考えられる．また，性別，使用言語，秘密の質問，郵便番号が情報漏えいしたとしても，大きなリスクにならない．さらに，生年月日はクロネコメンバーズ以外は保持しているため，情報漏えいリスクを高める要因にはならない．

よって，クラスタ 2 からクロネコメンバーズを外すことで，クラスタ 2 に属するアカウントから情報漏えいしても，限定された情報しか漏えいしないため，情報漏えいのリスクを低減することができる．

6.3 クラスタ 3

表 6 のクラスタ 3 が保持している情報より，クラスタ 3 のそれぞれのアカウントに同じパスワードを設定し，パスワードリスト攻撃をされた場合，ニックネーム，名前，ユーザー名，電話番号，メールアドレス，国情報，生年月日，性別，郵便番号，都道府県，使用言語が漏えいすることが分かる．また，全てのサイトがユーザー名，メールアドレス，性別を保持している．

Microsoft は機微情報である「電話番号」を保持しているため，クラスタ内の他のアカウントと同じパスワードを設定してしまうと，機微情報が漏えいしてしまう．したがって，Microsoft をクラスタ 3 から除外し，独立したパスワードを設定した方が良いと考えられる．また，ニックネーム，名前，国，生年月日，郵便番号，都道府県，使用言語は情報漏えいしたとしても，大きなリスクにならない．

よって，クラスタ 3 から Microsoft を外すことで，クラスタ 3 に属するアカウントから情報漏えいしても，限定された情報した漏えいしないため，情報漏えいのリスクを低減することができる．

表 6　クラスタ 3 が保持している情報

サイト名	ニックネーム	名前	ユーザー名	電話番号	メールアドレス	国
yahoo	×	×	○	×	○	×
mixi	○	○	○	×	○	×
Ameaba	○	×	○	×	○	×
Microsoft	×	○	○	○	○	○
ニコニコ動画	○	×	○	×	○	○
pixiv	○	×	○	×	○	×
FC2	×	×	○	×	○	×

サイト名	生年	生年月日	性別	郵便番号	都道府県	使用言語
yahoo	×	○	○	○	×	×
mixi	○	×	○	×	○	×
Ameaba	×	○	○	×	×	×
Microsoft	×	○	○	○	×	×
ニコニコ動画	×	○	○	×	○	×
pixiv	×	×	○	×	×	○
FC2	×	×	○	×	×	×

6.4 まとめ

以上より，クラスタ 1 から Google，クラスタ 2 からクロネコメンバーズ，クラスタ 3 から Microsoft をそれぞれ除外することで，各クラスタに同じパスワードを設定することが，情報漏えいリスクの観点から許容できた．また，各クラスタから除外を行った結果，記憶すべきパスワードは 6 個となった．

ここで，無配慮にクラスタ数 $k = 6$ で分類し，各クラスタに同じパスワードを設定した場合，いくつかのクラスタでは本提案手法よりも漏えいする情報を減らすことができるかもしれないが，全クラスタを総合的に見た場合，情報漏えいリスクは増大する．また，$k < 6$ の場合においても，各クラスタから漏えいする情報が増加するため，情報漏えいリスクは増加する．よって，本提案手法を用いることにより，**要件 1** は満たされた．

一方で，**要件 2** の達成は十分ではない．今回の実験では $k = 6$ が最適であったが，目標であった $k = 3$ には届かなかった．これは利用しているアカウントや機微情報の設定により変化すると考えられるため，必ずしも $k = 3$ を達成できるとは限らないと考えられる．クラスタ数 k が増加するということは，パスワードの使い回しを許容しないことと見做すことができる．つまり，文献 [2] で示されているように，クラスタ数 k が増加するとトレードオフでパスワード強度が低下することが懸念される．今後は，この観点からさらに研究を進めることが必要であると考える．

7. おわりに

本稿ではパスワードリスト攻撃による情報漏えいリスクを低減するアカウント管理手法について述べた．本提案手法で類似の情報を持つアカウント同士をクラスタリングをすることで，クラスタに属する全てのアカウントに同じパスワードを設定したとしても，限定された情報しか漏えいせず，さらに，ユーザーが記憶すべきパスワードを減らすことができた．本提案手法により，ユーザーは自身で情報漏えいリスクをコントロールしながら，重要なアカウントと重要でないアカウントを分けて，適切なパスワード管理ができる．

今後の展望として，調査する情報を増やし，より現実的な利用状況でクラスタリングできるようにする必要がある．今回調査したデータはアカウント登録したときの必須入力項目に限ったので，アカウント登録後に追加できる情報についての調査を行いたい．また，パスワードを使い回してしまう人が本提案手法を利用できるように，所有しているアカウントを選択することにより，自動的にクラスタリングを行うツールが必要である．

参考文献

[1] Dinei Florêncio, Cormac Herley, and Paul C. van Oorschot. An administrator's guide to internet password research. In *28th Large Installation System Administration Conference (LISA14)*, pp. 44–61, Seattle, WA, November 2014. USENIX Association.

[2] Dinei Florêncio, Cormac Herley, and Paul C. van Oorschot. Password portfolios and the finite-effort user: Sustainably managing large numbers of accounts. In *23rd USENIX Security Symposium (USENIX Security 14)*, pp. 575–590, San Diego, CA, August 2014. USENIX Association.

[3] エムオーテック株式会社. パスワード情報漏えい“パス漏れ”に要注意!!3 人に 1 人が同じパスワードで複数の Web サービスを利用！ http://www.motex.co.jp/nomore/report/1105/, (2015).

[4] 株式会社野村総合研究所. 利用者登録する商品・サービスを選別する傾向が強まった生活者と顧客情報の鮮度維持を望む事業者〜生活者と事業者を対象とした ID に関する実態調査〜. http://www.nri.com/jp/news/2012/120208.html, 2012.

[5] IPA 独立行政法人情報処理推進機構. オンライン本人認証方式の実態調査. http://www.ipa.go.jp/security/fy26/reports/ninsho/, August 2014.