
Visual-Similarity-Based Phishing Detection

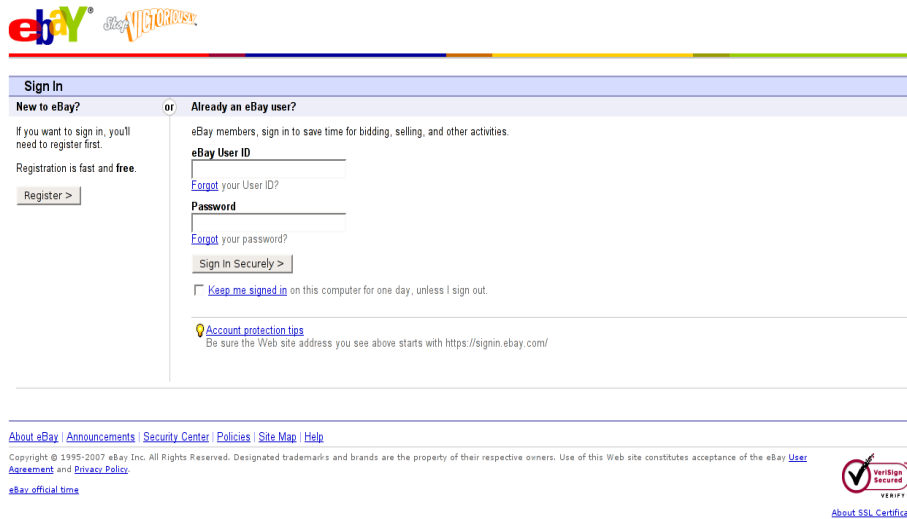
Eric Medvet, University of Trieste (IT)

Engin Kirda, Christopher Kruegel, Technical University of
Vienna (AT)

What is phishing?

- A form of **online fraud** aimed at stealing user's sensitive information
- The attacker crafts a web page that mimicks the original page and spreads its URL using emails
- An important and current problem:
 - More than 25000 phishing attacks in Dec. 2007
 - "Do yourself" phishing kits
 - Can be combined with other online attacks

Phishing: an example



The screenshot shows the genuine eBay sign-in page. At the top left is the eBay logo with the text "eBay MARKETPLACE". Below the logo is a navigation bar with the text "Sign In". The main content area is divided into two sections: "New to eBay?" and "Already an eBay user?". The "New to eBay?" section includes a "Register >" button and text stating "Registration is fast and free." The "Already an eBay user?" section includes fields for "eBay User ID" and "Password", with "Forgot" links for both. There is also a "Sign In Securely >" button and a checkbox for "Keep me signed in". At the bottom, there is a footer with links for "About eBay", "Announcements", "Security Center", "Policies", "Site Map", and "Help". A copyright notice and a "VeriSign Secured" logo are also present.

The genuine page



The screenshot shows a phishing page that mimics the genuine eBay sign-in page. At the top left is the eBay logo with the text "eBay.co.uk". Below the logo is a navigation bar with the text "Please Sign In:". The main content area is divided into two sections: "New to eBay?" and "Already an eBay user?". The "New to eBay?" section includes a "Register >" button and text stating "Registration is fast and free." The "Already an eBay user?" section includes fields for "eBay User ID" and "Password", with "Forgot" links for both. There is also a "Sign In Securely >" button and a checkbox for "Keep me signed in". At the bottom, there is a footer with links for "About eBay", "Announcements", "Safety Centre", "VeriSign Protecting IP", "Policies", "Feedback Forum", "Site Map", and "Help". A copyright notice and a "VeriSign Secured" logo are also present.

The phishing page

Against phishing: state of the art

- At the email level
 - Blocking suspected emails (similarly to antispam techniques)
 - Sender authentication (*not widely used*)
- Black lists (*uneffective vs. zero day attack*)
- Browser-integrated solutions
 - Domain-specific passwords
 - Detect suspected URLs
 - AntiPhish, prevents sending credentials to untrusted unknown websites...

Antiphish

1. Store user's credentials and link to corresponding sites:
 - <username1, passwd1, url1>
 - <username2, passwd2, url2>
 - . . .
2. Whenever credentials are to be sent to a website, check if it is trusted
 - Problem: what if the user uses the **same credentials** on many websites? False positives
 - Solution: check if the website is **suspiciously similar** to the trusted one

Contribution of this work

- A novel visual comparison technique for phishing detection
- **Compare** the suspected web page with the corresponding genuine one: if they are too similar, raise an alert
 - Thought to be part of AntiPhish
 - ... but can also be integrated in other solutions: e.g., in a mail server to block suspected emails

Visual comparison: overview

- Why "visual"?
- Users get convinced to be on a legitimate page mainly by perceived page content
- We compare page features that **affect user perception**:
 - Text blocks
 - Images
 - Overall image (the page **as rendered by the browser**)
- We obtain a visual similarity index

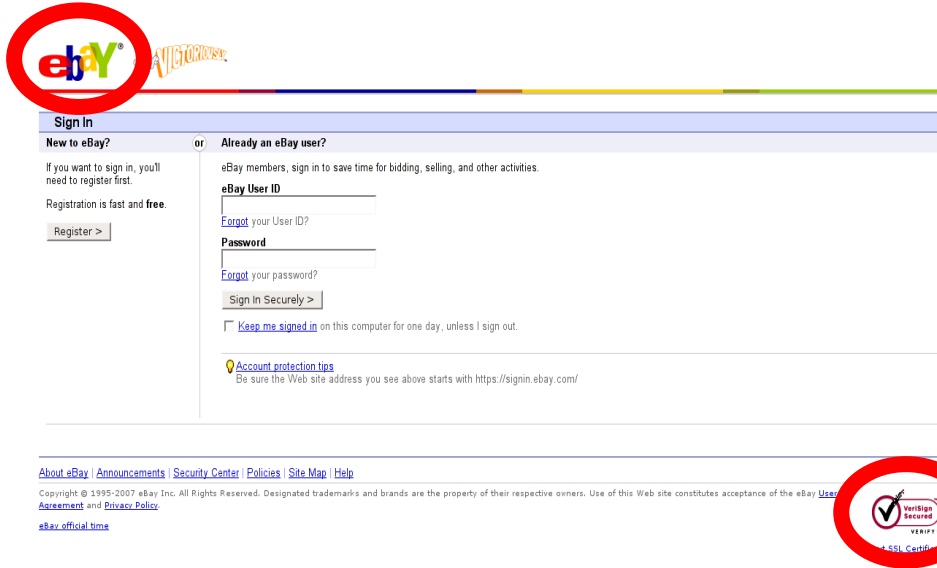
What we consider

- Text blocks:
 - textual content, foreground color, background color, font size, font family name, position
- Images:
 - `src` attribute, area ($w \times h$), color histograms, 2D Haar transform, position
- Overall image:
 - color histograms, 2D Haar transform

Comparison

1. We compute distances between elements of each text block pairs, image pairs and the overall image pair:
 - a similarity index for each pair

Comparison



Sign In

New to eBay? or **Already an eBay user?**

If you want to sign in, you'll need to register first.

Registration is fast and free.

[Register >](#)

eBay members, sign in to save time for bidding, selling, and other activities.

eBay User ID

[Forgot](#) your User ID?

Password

[Forgot](#) your password?

[Sign In Securely >](#)

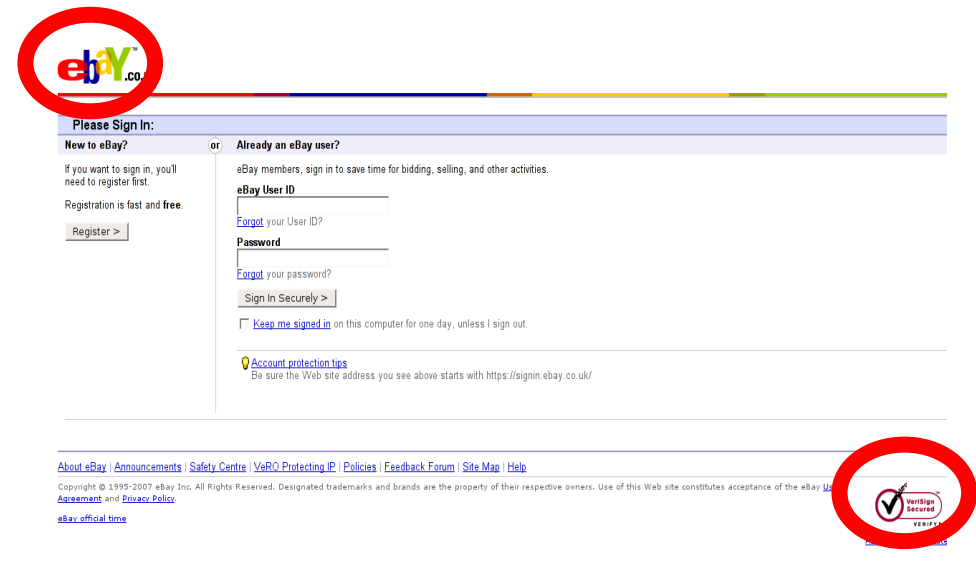

[Keep me signed in](#) on this computer for one day, unless I sign out.

[Account protection tips](#)
Be sure the Web site address you see above starts with <https://signin.ebay.com/>

[About eBay](#) | [Announcements](#) | [Security Center](#) | [Policies](#) | [Site Map](#) | [Help](#)

Copyright © 1995-2007 eBay Inc. All Rights Reserved. Designated trademarks and brands are the property of their respective owners. Use of this Web site constitutes acceptance of the eBay [User Agreement](#) and [Privacy Policy](#).

[eBay official time](#)



Please Sign In:

New to eBay? or **Already an eBay user?**

If you want to sign in, you'll need to register first.

Registration is fast and free.

[Register >](#)

eBay members, sign in to save time for bidding, selling, and other activities.

eBay User ID

[Forgot](#) your User ID?

Password

[Forgot](#) your password?

[Sign In Securely >](#)


[Keep me signed in](#) on this computer for one day, unless I sign out.

[Account protection tips](#)
Be sure the Web site address you see above starts with <https://signin.ebay.co.uk/>

[About eBay](#) | [Announcements](#) | [Safety Centre](#) | [VeriSign Protecting IP](#) | [Policies](#) | [Feedback Forum](#) | [Site Map](#) | [Help](#)

Copyright © 1995-2007 eBay Inc. All Rights Reserved. Designated trademarks and brands are the property of their respective owners. Use of this Web site constitutes acceptance of the eBay [User Agreement](#) and [Privacy Policy](#).

[eBay official time](#)



1. ebay logo vs ebay logo
2. ebay logo vs verisign logo
3. verisign logo vs ebay logo
4. verisign logo vs verisign logo
5. ...

Comparison

1. We compute distances between elements of each text block pairs, image pairs and the overall image pair:
 - a similarity index for each pair
2. We consider only the first 10 **best-matching** text block pairs, the 5 best-matching image pairs, the 2 overall images:
 - three similarity indexes for text, images, overall image
3. We obtain the final similarity index averaging indexes of text blocks, images and overall image
 - Some weights (coefficients) are set by domain knowledge, the others by optimization on a small training set
 - Some algorithm optimization in order to reduce computational effort

Experimental evaluation: dataset

Dataset

- **41 positive pairs** <genuine page, phishing page>, partitioned according to visual similarity
 - 20 level 0: almost perfect visual match
 - 14 level 1: small differences
 - 7 level 2: noticeable differences
- **161 negative pairs** <genuine page, different page>
 - 115 with a login form
 - 46 without a login form

Positive pairs: example

Sign in

New to eBay? If you want to sign in, you'll need to register first. Registration is fast and free. [Register >](#)

Already an eBay user? eBay members, sign in to save time for bidding, selling, and other activities. [Log in](#) your User ID? [Log in](#) your password? [Sign in Securely >](#)

[Keep me signed in](#) on this computer for one day, unless I sign out.

[Account protection tips](#)
Be sure the Web site address you see above starts with <https://signin.ebay.com/>

[About eBay](#) | [Announcements](#) | [Security Center](#) | [Feedback](#) | [Site Map](#) | [Help](#)
Copyright © 1995-2007 eBay Inc. All Rights Reserved. Designated trademarks and brands are the property of their respective owners. Use of this Web site constitutes acceptance of the [eBay User Agreement](#) and [Privacy Policy](#).
[eBay official time](#)

[About SSL Certificate](#)

The genuine page

Please Sign In:

New to eBay? If you want to sign in, you'll need to register first. Registration is fast and free. [Register >](#)

Already an eBay user? eBay members, sign in to save time for bidding, selling, and other activities. [Log in](#) your User ID? [Log in](#) your password? [Sign in Securely >](#)

[Keep me signed in](#) on this computer for one day, unless I sign out.

[Account protection tips](#)
Be sure the Web site address you see above starts with <https://signin.ebay.co.uk/>

[About eBay](#) | [Announcements](#) | [Security Center](#) | [VeriSign Protecting ID](#) | [TurboKey](#) | [Feedback Forum](#) | [Site Map](#) | [Help](#)
Copyright © 1995-2007 eBay Inc. All Rights Reserved. Designated trademarks and brands are the property of their respective owners. Use of this Web site constitutes acceptance of the [eBay User Agreement](#) and [Privacy Policy](#).
[eBay official time](#)

[About SSL Certificate](#)

Place Your Credit Card on File

Use this form to place your credit card on file for automatic monthly payment of your eBay seller fees. These are the fees eBay charges sellers for listing and selling items. The information is protected using the industry standard [SSL](#).

Contact name

Credit or debit card number
Visa, MasterCard, American Express, or Discover

Card Identification number
This is the 3-digit number on the back of your credit or debit card. For American Express cards, use the 4-digit number on the front of the card. [Learn more](#)

Expiration date
--Month-- --Year--

Card PIN Number
4-Digit code used in ATM's

Cardholder name

Billing address (must match monthly statement address)

E-mail address

City

State **Zip code** **County** (United States)

Pin **Mother's Maiden Name** **Social Security Number** **Date of Birth**
--Month-- --Year--
--Day--
(mm/dd/yyyy)

[Continue >](#)
Please click the Continue button only once.

[About eBay](#) | [Announcements](#) | [Security Center](#) | [Feedback](#) | [Site Map](#) | [Help](#)
Copyright © 1995-2007 eBay Inc. All Rights Reserved. Designated trademarks and brands are the property of their respective owners. Use of this Web site constitutes acceptance of the [eBay User Agreement](#) and [Privacy Policy](#).
[eBay official time](#)

Testing methodology

- Portion of the dataset used for training
 - 14 on 41 positive pairs
 - 21 on 161 negative pairs
- FPR and FNR computed on the remaining part

Results

- Satisfying results:

Levels	t	f_p	n_n	FPR	f_n	n_p	FNR
All (0, 1 and 2)	0.956	0	140	0%	2	27	7.4%
Only 0 and 1	0.956	0	140	0%	0	20	0.0%
Only 0	0.956	0	140	0%	0	11	0.0%

- **No false positives**
- Only 2 false negatives (level 2, rather difficult to detect)

False negative: a sample

Bad Login :: Empty Fields...



Member Number:
Enter your member number in the field above.

Password:
(First-Time Users, use the last four digits of your Social Security Number as your password.)
Enter the Security Code Shown Below:



Attention First-Time Users of 1st United's New Internet Banking Site:

1. To enter the new site, enter your member number in the Member Number field.
2. Enter the last four digits of the Primary Member's Social Security Number in the Password field.
3. Enter the Security Code with shown in the Security Code field.
4. Click submit and follow the instructions to continue.

Unauthorized use of these systems is strictly prohibited and is a subject of prosecution under the Computer Fraud and Abuse Act of 1986 and Title 18, U.S. Code, Sec. 1030 and 1030a. First Systems Corporation may monitor and audit usage of this system. All passwords are hereby notified to the use of this system is subject to change without notice.



Welcome to 1st United Services Credit Union Login
Please reauthorize your account immediately. Complete the form below and click "Submit" to finish. * Denotes required field.

Card Number:

Expiration Date:

Electronic Signature: (max 50)

Email Address:

Copyright © 1st United Services Credit Union. All Rights Reserved.

Thanks

Questions?
