

What is claimed is:

1. A system for detecting page impersonation in phishing attacks, comprising:
 - a computer having a processor and a network communication; and
 - a program product comprising machine-readable program code for causing, when executed, the computer to perform the following process steps:
 - automatically analyzing the body of an e-mail message to detect an embedded universal resource locator (URL);
 - automatically extracting the embedded URL;
 - automatically capturing a screenshot of a website referenced by the embedded URL;
 - automatically comparing the captured screenshot with a record screenshot without any preprocessing of the captured screenshot, wherein the record screenshot corresponds to a trusted site; and
 - when the captured screenshot does not match the record screenshot, marking the embedded URL as safe.
2. The system of claim 1, further comprising:
 - when the captured screenshot matches the record screenshot, determining if a domain of the embedded URL corresponds to a trusted domain.
3. The system of claim 2, further comprising:
 - when the domain of the embedded URL corresponds to the trusted domain, marking the embedded URL as safe.
4. The system of claim 3, further comprising:
 - when the domain of the embedded URL does not correspond to the trusted domain, marking the e-mail message as a page impersonation attempt.

5. The system of claim 1, further comprising:

a page impersonation database storing data associated with the trusted site, wherein the trusted site data includes: a trusted URL, a trusted domain corresponding to the trusted URL, and the record screenshot.

6. The system of claim 5, further comprising:

receiving a URL designating a contributed site from a user; and
storing the contributed site in the page impersonation database.

7. The system of claim 6, further comprising:

automatically capturing a screenshot of the contributed site; and
storing the screenshot for the contributed site in the page impersonation database.

8. A method for detecting a page impersonation phishing attempt presented by an e-mail message, comprising:

automatically analyzing the body of an e-mail message to extract an embedded universal resource locator (URL);

automatically capturing a screenshot of a website referenced by the embedded URL;

automatically comparing the captured screenshot with a record screenshot without any preprocessing of the captured screenshot, wherein the record screenshot corresponds with a trusted site; and

when the captured screenshot does not match the record screenshot, marking the embedded URL as safe.

9. The method of claim 8, further comprising:

when the captured screenshot matches the record screenshot, determining if a domain of the embedded URL corresponds to a trusted domain associated with the trusted site.

10. The method of claim 9, further comprising:

when the domain of the embedded URL corresponds to the trusted domain, marking the embedded URL as safe.

11. The method of claim 10, further comprising:

when the domain of the embedded URL does not correspond to the trusted domain, marking the e-mail message as a page impersonation attempt.

12. The method of claim 9, further comprising:

storing the trusted site in a page impersonation database, wherein the trusted site includes a trusted URL, a trusted domain corresponding to the trusted URL, and the record screenshot.

13. The method of claim 12, further comprising:

receiving a URL designating a contributed site from a user; and
storing the contributed site in the page impersonation database.

14. The method of claim 13, further comprising:

automatically capturing a screenshot of the contributed site; and
storing the screenshot for the contributed site in the page impersonation database.

15. A non-transitory computer-readable memory adapted to detect page impersonation phishing attacks, the non-transitory computer readable memory used to direct a computer to perform process steps, comprising:

 automatically analyzing the body of an e-mail message to extract an embedded universal resource locator (URL);

 automatically capturing a screenshot of a website referenced by the embedded URL;

 automatically comparing the captured screenshot with a record screenshot without any preprocessing of the captured screenshot, wherein the record screenshot corresponds with a trusted site; and

 when the captured screenshot does not match the record screenshot, marking the embedded URL as safe.

16. The non-transitory computer-readable memory of claim 15, wherein the process steps further comprise:

 when the captured screenshot matches the record screenshot, determining if a domain of the embedded URL corresponds to a trusted domain associated with the trusted site.

17. The non-transitory computer-readable memory of claim 16, wherein the process steps further comprise:

 when the domain of the embedded URL corresponds to the trusted domain, marking the embedded URL as safe.

18. The non-transitory computer-readable memory of claim 17, wherein the process steps further comprise:

 when the domain of the embedded URL does not correspond to the trusted domain, marking the e-mail message as a page impersonation attempt.

19. The non-transitory computer-readable memory of claim 18, wherein the process steps further comprise:

storing the trusted site in a page impersonation database, wherein the trusted site includes a trusted URL, a trusted domain corresponding to the trusted URL, and the record screenshot.

20. The non-transitory computer-readable memory of claim 19, wherein the process steps further comprise:

receiving a URL designating a contributed site from a user;

automatically capturing a screenshot of the contributed site; and

storing the contributed site and the screenshot of the contributed site in the page impersonation database.