27. HARIPRIYA B

Phishing-Alarm: Robust and Efficient Phishing Detection via Page Component Similarity

Abstract   Phishing is a form of social engineering in which an attacker mimics electronic communications to lure users to provide their confidential information. In a Web-based phishing attack, an attacker sets up scam Web pages (pretending to be an important Website such as a social network portal) to lure users to input their private information, such as passwords, social security numbers, credit card numbers, and so on. In fact, the appearance of Web pages is among the most important factors in deceiving users, and thus, the similarity among Web pages is a critical metric for detecting phishing Websites. In this paper, we propose a new technique, called Phishing Alarm, to efficiently and accurately detect phishing pages according to the similarity of page visual features. As the visual appearance of a web page is jointly determined by its elements and its CSS rules, our approach first evaluates the impact these elements and rules have on the final page appearance to users, and selects the effective features as the basis of our approach. Based on these effective features, we develop a new algorithm to effectively measure the similarity of page components between a suspicious page and its potential target. Phishing-Alarm is a new solution to detect phishing attacks using features that are hard to evade by attackers. In particular, we present an algorithm to quantify the suspiciousness ratings of Web pages based on the similarity of visual appearance between the Web pages. Since cascading style sheet (CSS) is the technique to specify page layout across browser implementations, our approach uses CSS as the basis to accurately quantify the visual similarity of each page element. As page elements do not have the same influence to pages, we base our rating method on weighted page-component similarity. The proof of concept implementation verifies the correctness and accuracy of this approach with a relatively low performance overhead.