



US 20230231879A1

(19) **United States**

(12) **Patent Application Publication**
Li et al.

(10) **Pub. No.: US 2023/0231879 A1**

(43) **Pub. Date: Jul. 20, 2023**

(54) **MACHINE LEARNING FOR VISUAL
SIMILARITY-BASED PHISHING
DETECTION**

(71) Applicant: **Fortinet, Inc.**, Sunnyvale, CA (US)

(72) Inventors: **Haitao Li**, Coquitlam (CA); **Lisheng
Ryan Sun**, Burnaby (CA)

(73) Assignee: **Fortinet, Inc.**, Sunnyvale, CA (US)

(21) Appl. No.: **18/125,916**

(22) Filed: **Mar. 24, 2023**

Related U.S. Application Data

(63) Continuation-in-part of application No. 16/583,707,
filed on Sep. 26, 2019.

Publication Classification

(51) **Int. Cl.**
H04L 9/40 (2006.01)

(52) **U.S. Cl.**
CPC **H04L 63/1483** (2013.01)

(57) **ABSTRACT**

In one embodiment, a similarity index is calculated from characteristics of a suspected phishing web page to a database of known phishing web pages. The characteristics derive from both HTML tags of the suspected phishing web page and a screenshot of the suspected phishing web page. With machine learning using the similarity index as an input, a probability is estimated that the suspected web page comprises a known phishing web page from the database of known phishing web pages. A known phishing web page is selected from one or more candidates known phishing web pages, based on having a highest probability.

100

