# ỨNG DỤNG MÁY HỌC ĐỂ PHÁT HIỆN URL GIẢ MẠO

**Nguyễn Phúc An Tim**
**Trường Cao đẳng Cộng đồng Hậu Giang**

***Tóm tắt:*** *Phát hiện URL giả mạo là một vấn đề nghiêm trọng trong an ninh mạng. Các phương pháp truyền thống dựa vào danh sách đen thường không đủ nhanh chóng và linh hoạt để phát hiện các URL mới giả mạo. Trong bài báo này, tác giả trình bày cơ sở lý luận về URL giả mạo, những tác dụng và tiềm năng hiệu quả của máy học đối với bài toán phát hiện URL giả mạo.*

***Từ khóa:*** *phishing, URL giả mạo, máy học, Random Forest, an ninh mạng*

# APPLICATION OF MACHINE LEARNING TO DETECTING FAKE URLS

**Nguyen Phuc An Tim**
**Hau Giang Community College**

***Abstract:*** *Detecting fake URLs is a serious problem in network security. Traditional methods based on blacklists are often not fast and flexible enough to detect new fake URLs. In this paper, the author presents the theoretical basis of fake URLs, the effects and potential effectiveness of machine learning for the problem of detecting fake URLs.*

***Keywords:*** *phishing, fake URLs, machine learning, Random Forest, network security*

*Nhận bài: 11/05/2025* *Phản biện: 04/06/2025* *Duyệt đăng: 09/06/2025*

## I. ĐẶT VẤN ĐỀ

Trong kỷ nguyên số hiện nay, sự bùng nổ của các cuộc tấn công giả mạo trực tuyến đang trở thành mối đe dọa nghiêm trọng đối với an ninh mạng toàn cầu. Đặc biệt, các cuộc tấn công giả mạo URL (phishing URL) đang phát triển với tốc độ chóng mặt và mức độ tinh vi ngày càng cao, nhắm vào người dùng Internet, tổ chức và doanh nghiệp để đánh cắp thông tin nhạy cảm, tài sản kỹ thuật số và dữ liệu cá nhân.

Các cuộc tấn công qua URL giả mạo đã gây ra những hậu quả nghiêm trọng trên phạm vi toàn cầu. Kẻ tấn công liên tục cải tiến phương thức, tạo ra các URL có vẻ hợp pháp nhưng thực chất dẫn người dùng đến các trang web độc hại, từ đó chiếm đoạt thông tin đăng nhập, dữ liệu tài chính, và thậm chí cài đặt phần mềm độc hại vào hệ thống của nạn nhân. Điều đáng báo động là các phương pháp tấn công ngày càng khó phát hiện, khai thác tinh vi các lỗ hổng trong nhận thức của người dùng.

Trước thách thức này, các phương pháp phòng thủ truyền thống như danh sách đen (blacklist) và phân tích dựa trên quy tắc (rule-based analysis) đã bộc lộ nhiều hạn chế nghiêm trọng. Chúng thường phản ứng chậm, không có khả năng thích ứng với các mẫu tấn công mới, và đặc biệt bất lực trước các biến thể URL giả mạo chưa từng xuất hiện trong cơ sở dữ liệu. Hạn chế này tạo ra khoảng trống lớn trong hệ thống bảo mật, đặt ra yêu cầu cấp thiết về một giải pháp mới hiệu quả hơn.

Máy học (Machine Learning) là kỹ thuật thiết kế và phát triển các thuật toán cho phép máy tính đánh giá hành vi dựa trên dữ liệu thực nghiệm, chẳng hạn như dữ liệu cảm biến hoặc cơ sở dữ liệu. Một chương trình học có thể tận dụng các mẫu (dữ liệu) để nắm bắt các đặc điểm quan tâm, dữ liệu có thể được xem như là ví dụ minh họa mối quan hệ giữa các biến quan sát được. Trọng tâm chính của nghiên cứu học máy là tự động học cách nhận ra các mẫu phức tạp và đưa ra quyết định thông minh dựa trên dữ liệu. Học máy có thể được chia thành các nhánh như sau: học có giám sát, học nửa giám sát và học không giám sát. Còn học sâu (Deep Learning) nổi lên như một giải pháp đột phá, mang đến cách tiếp cận chủ động và thông minh hơn trong việc nhận diện URL giả mạo. Với khả năng phân tích các đặc trưng phức tạp của URL, học hỏi từ dữ liệu lịch sử, và tự thích nghi trước các mẫu tấn công mới, các mô hình máy học hứa hẹn vượt qua những hạn chế của phương pháp truyền thống.

Trong bối cảnh đó, nghiên cứu việc ứng dụng các mô hình máy học để nâng cao hiệu quả nhận diện URL giả mạo là cần thiết. Thông qua việc phát triển, đánh giá và so sánh các mô hình khác nhau, cũng như phân tích các đặc trưng quyết định trong quá trình phân loại, nghiên cứu hướng tới xây dựng một hệ thống phòng thủ trước làn sóng tấn công giả mạo ngày càng phức tạp trong thời đại số. Kết quả của nghiên cứu không chỉ góp phần nâng cao năng lực bảo mật mạng mà còn đặt nền móng cho các giải pháp phòng thủ thông minh trong tương lai.

## II. NỘI DUNG NGHIÊN CỨU

### 2.1. Một số vấn đề về URL

URL (Uniform Resource Locator), hay còn gọi là Định vị tài nguyên thống nhất, là địa chỉ duy nhất của một tài nguyên trên Web. Một URL hợp lệ sẽ dẫn đến một tài nguyên cụ thể, có thể là trang HTML, tệp CSS, hình ảnh, video, tài liệu PDF... Tuy nhiên, trong một số trường hợp, URL có thể trỏ đến các tài nguyên đã bị xóa hoặc di chuyển sang địa chỉ khác. Mỗi URL được tạo ra đều có mục đích giúp người dùng truy cập nhanh chóng vào nội dung họ cần trên Internet. Nếu không có URL, việc định vị và truy xuất thông tin trên web sẽ trở nên khó khăn và kém hiệu quả hơn rất nhiều.

URL bao gồm nhiều thành phần khác nhau, trong đó hostname (tên máy) giúp ánh xạ đến địa chỉ IP của tài nguyên trên Internet. Ngoài ra, nó còn chứa các thông tin bổ sung để trình duyệt và máy chủ có thể xử lý dữ liệu đúng cách. Có thể hình dung địa chỉ IP giống như số điện thoại, còn hostname tương đương với tên người sở hữu số đó. Hệ thống tên miền (DNS - Domain Name System) đóng vai trò như một danh bạ điện thoại, giúp chuyển đổi hostname thành địa chỉ IP để mạng có thể định tuyến lưu lượng truy cập chính xác.

Về cấu trúc, URL lần đầu tiên được xác định vào năm 1994 bởi Sir Tim Berners-Lee, cha đẻ của World Wide Web. Về nguyên tắc, URL kết hợp giữa tên miền và đường dẫn tệp để xác định chính xác vị trí của tài nguyên trên máy chủ. Cách hoạt động của URL tương tự như đường dẫn tệp trong hệ điều hành Windows, chẳng hạn như C:\Documents\Personal\myfile.txt, nhưng có thêm các yếu tố bổ sung giúp xác định đúng máy chủ trên Internet và sử dụng giao thức phù hợp để truy cập dữ liệu.

URL giả mạo là một đường dẫn được thiết kế để đánh lừa người dùng, trông giống như một trang web hợp pháp nhưng thực chất được dùng để thực hiện hành vi giả mạo, phát tán mã độc hoặc thu thập dữ liệu cá nhân.

Các loại URL giả mạo phổ biến

URL giả mạo bằng lỗi chính tả (Typosquatting): Lợi dụng sự bất cẩn của người dùng khi nhập sai tên miền.

Ví dụ:

gogle.com thay vì google.com

amaz0n.com thay vì amazon.com

URL giả mạo bằng ký tự tương tự (Homograph Attack): Sử dụng các ký tự Unicode để tạo tên miền giống hệt tên miền thật.

Ví dụ: apple.com (chữ "a" là ký tự Cyrillic, không phải chữ "a" Latin).

URL giả mạo bằng subdomain: Kẻ tấn công thêm subdomain để làm cho URL trông đáng tin cậy hơn.

Ví dụ:

paypal.com.secure-login.com → Tên miền thực là secure-login.com, không phải paypal.com.

URL rút gọn (Shortened URL): Sử dụng dịch vụ rút gọn URL (bit.ly, tinyurl, v.v.) để che giấu URL thực.

Ví dụ: bit.ly/abc123 có thể trỏ đến một trang giả mạo mà người dùng không biết trước.

URL chứa mã độc (Malicious URL): URL dẫn đến trang web chứa mã độc (malware, spyware).

Ví dụ: http://freedownloads.com/install.exe (tải xuống tệp tin chứa mã độc).

Về danh sách đen ("blacklist") là danh mục URL đã được xác định là nguy hiểm và bị hệ thống chặn truy cập. Tuy nhiên, hàng ngày có hàng nghìn URL giả mạo mới được tạo ra, khiến danh sách này không thể bảo đảm bao quát. Ngoài ra, nhiều URL nguy hiểm có đọi tên linh hoạt, làm danh sách đen nhanh chóng trở nên lỗi thời. Máy học (machine learning) cho phép mô hình tự động nhận diện mẫu trong dữ liệu và dự đoán URL giả mạo dựa trên các đặc trưng như: độ dài URL, các từ khóa nghi ngờ, cấu trúc tên miền, vị trí ký tự đặc biệt. Khác với danh sách đen, mô hình máy học có khả năng nhận diện URL giả mạo chưa từng xuất hiện. Các thuật toán như Decision Tree, Random Forest, SVM, và XGBoost được áp dụng rộng rãi do khả năng xử lý tốt các đặc trưng phi tuyến tính, giám sát cân đối dữ liệu. Việc huấn luyện và đánh giá hiệu suất dựa trên các tiêu chí: độ chính xác, độ nhạy, và tỷ lệ dương tính giả. Các đặc trưng của URL được trích xuất bao gồm: độ dài chuỗi, số lần xuất hiện ký tự lạ, tỷ lệ ký tự số, số lượng dấu điểm, có hay không dùng IP, tên miền phụ, các từ nhạy cảm như "login", "bank", "update", "verify", v.v. Việc chuẩn hoá và chọn lọc đặc trưng phù hợp là yếu tố quan trọng đối với hiệu quả mô hình.

### 2.2. URL giả mạo, Phishing và các hình thức tấn công phishing phổ biến

- URL giả mạo là địa chỉ trang web được thiết kế nhằm giả mạo các trang web hợp pháp để lừa người dùng cung cấp thông tin nhạy cảm. Hình thức tấn công này thường diễn ra thông qua email, tin nhắn, hoặc các trang web được chia sẻ trên

mạng xã hội. Các URL giả mạo có thể khác nhau về độ dài, cách viết tên miền, và thậm chí che giấu địa chỉ thật.

- Phishing là một hình thức tấn công giả mạo mà kẻ gian mạo danh một tổ chức hoặc cá nhân đáng tin cậy để đánh cắp thông tin nhạy cảm như tên đăng nhập, mật khẩu, thông tin thẻ tín dụng.

Hình thức phổ biến nhất là gửi email có chứa URL giả mạo dẫn đến trang web giả.

Các hình thức tấn công phishing phổ biến:

- Email Phishing: Tin tặc gửi email giả mạo từ các tổ chức uy tín như ngân hàng, PayPal, Google. Email chứa một URL giả mạo yêu cầu người dùng nhập thông tin đăng nhập.

Ví dụ: Một email từ "support@secure-paypal.com" yêu cầu cập nhật tài khoản PayPal.

- Spear Phishing: Phishing có mục tiêu cụ thể, thường nhắm vào cá nhân hoặc tổ chức. Tin tặc nghiên cứu trước thông tin về nạn nhân để tạo email giả thuyết phục hơn.

Ví dụ: Một nhân viên công ty nhận email từ "IT-Support@company.com" yêu cầu đổi mật khẩu nội bộ.

- Whaling (Nhắm vào cá nhân cấp cao): Phishing nhắm vào lãnh đạo cấp cao trong doanh nghiệp để đánh cắp thông tin quan trọng.

Ví dụ: CEO của một công ty nhận email từ "CFO" yêu cầu chuyển tiền vào tài khoản ngân hàng giả mạo.

- Smishing (SMS Phishing): Gửi tin nhắn SMS chứa liên kết giả mạo.

Ví dụ: "Ngân hàng A: Bạn có giao dịch đáng ngờ, hãy xác nhận tại bank-security.com".

- Vishing (Voice Phishing): Giả mạo qua cuộc gọi điện thoại, giả danh nhân viên ngân hàng hoặc tổ chức chính phủ.

Ví dụ: Một cuộc gọi từ "bộ phận an ninh ngân hàng" yêu cầu cung cấp mã OTP.

**2.3. Tình hình giả mạo URL ở nước ta**

Trong thời gian gần đây, tình trạng giả mạo URL và các hình thức lừa đảo trực tuyến đang có xu hướng gia tăng mạnh mẽ tại Việt Nam, gây ảnh hưởng nghiêm trọng đến an toàn thông tin và tài sản của người dùng. Các đối tượng xấu không ngừng tìm cách tạo ra những trang web giả mạo với giao diện và nội dung gần giống các trang chính thống của các cơ quan nhà nước, tổ chức tài chính, doanh nghiệp lớn, sàn thương mại điện tử… nhằm lừa đảo, đánh cắp thông tin cá nhân, tài khoản ngân hàng hoặc thực hiện các giao dịch bất hợp pháp.

Theo thống kê từ Cục An toàn thông tin (Bộ Thông tin và Truyền thông), chỉ trong tháng 3/2024, đã có 100 website giả mạo thương hiệu bị phát hiện. Tính đến hết quý 1/2024, tổng số trang web giả mạo được ghi nhận đã lên tới 124.579, cho thấy mức độ nguy hiểm ngày càng gia tăng. Đáng lo ngại, nhiều trang web trong số này nhắm vào các cơ quan chính phủ, ngân hàng và tập đoàn lớn để tạo lòng tin với người dùng, từ đó thực hiện các hành vi lừa đảo tinh vi.

Hình thức lừa đảo tinh vi, kết hợp với các chiến thuật lừa đảo khác nhau như gửi tin nhắn SMS, email giả mạo, hoặc quảng cáo trên mạng xã hội. Trước thực trạng này, người dùng cần hết sức cảnh giác khi truy cập các trang web, đặc biệt là các trang liên quan đến tài chính, giao dịch trực tuyến hoặc thông tin cá nhân.

Nghiên cứu ứng dụng máy học để phát hiện URL giả mạo này tập trung vào việc ứng dụng máy học để nhận diện URL giả mạo nhằm phát hiện các trang giả mạo, phát tán mã độc hoặc tấn công thay đổi giao diện. Thay vì dựa vào danh sách đen (blacklist) hoặc phân tích nội dung trang web, nghiên cứu sử dụng đặc trưng của chính URL để phân loại. Hai tập dữ liệu được sử dụng là balanced_urls.csv (cân bằng giữa URL hợp lệ và độc hại) và malicious_phish.csv (gồm nhiều loại URL giả mạo khác nhau), giúp đánh giá hiệu suất mô hình trên các nguồn dữ liệu khác nhau.

Các đặc trưng quan trọng của URL được phân tích bao gồm: độ dài, số lượng dấu đặc biệt, loại domain, giao thức HTTP/HTTPS. Đặc biệt trích xuất thêm 2 đặc trưng mới là lỗi chính tả so với URL hợp lệ, từ khóa đáng ngờ (login, secure, verify). Mô hình sẽ được huấn luyện với nhiều thuật toán máy học (Logistic Regression, Random Forest, XGBoost, ANN) và đánh giá bằng các chỉ số như Accuracy, Precision, Recall, F1-score. Đặc biệt, SHAP được sử dụng để xác định các thuộc tính quan trọng, giúp tối ưu hóa mô hình.

**2.4. Tác dụng của ứng dụng máy học để phát hiện URL giả mạo**

Tác động của ứng dụng máy học để phát hiện URL giả mạo không chỉ góp phần nâng cao năng lực bảo mật mạng mà còn đặt nền móng cho các giải pháp phòng thủ thông minh trong tương lai.

Đối với người dùng cá nhân: Người dùng có thể tránh bị giả mạo khi truy cập vào các trang web nguy hiểm, đặc biệt trong các giao dịch ngân hàng, mua sắm trực tuyến hoặc sử dụng email. Hệ thống giúp giảm thiểu nguy cơ mất cắp thông tin

cá nhân, tài khoản và dữ liệu tài chính. Ngoài ra, nó còn giúp bảo vệ thiết bị khỏi các phần mềm độc hại được phát tán thông qua các trang web giả mạo.

Đối với doanh nghiệp và tổ chức: Doanh nghiệp có thể sử dụng hệ thống để ngăn chặn các cuộc tấn công giả mạo nhằm vào nhân viên hoặc khách hàng. Điều này đặc biệt quan trọng đối với các ngân hàng, công ty thương mại điện tử và các tổ chức xử lý dữ liệu nhạy cảm. Việc triển khai hệ thống giúp bộ phận IT giám sát an ninh mạng hiệu quả hơn, giảm thiểu rủi ro rò rỉ dữ liệu từ các cuộc tấn công phishing. Ngoài ra, nó cũng góp phần bảo vệ uy tín doanh nghiệp, ngăn chặn việc kẻ xấu lợi dụng thương hiệu để giả mạo khách hàng.

Đối với hệ thống an ninh mạng: Ứng dụng này có thể được tích hợp vào các hệ thống an ninh như tường lửa, phần mềm bảo mật hoặc hệ thống phát hiện xâm nhập (IDS) để chặn truy cập vào các trang web độc hại. Nó cũng có thể hỗ trợ cơ quan thực thi pháp luật trong việc theo dõi và phân tích các chiến dịch giả mạo trực tuyến, giúp giảm thiểu tội phạm mạng. Bên cạnh đó, các nhà cung cấp dịch vụ Internet (ISP) có thể tận dụng công nghệ này để bảo vệ khách hàng khỏi các mối đe dọa từ web độc hại.

## III. KẾT LUẬN

Sự gia tăng nhanh chóng của các URL giả mạo đang đặt ra thách thức lớn trong bảo mật thông tin trên Internet. Các URL này được thiết kế tinh vi để đánh lừa người dùng, đánh cắp dữ liệu cá nhân. Trong khi các phương pháp truyền thống như danh sách đen không đủ hiệu quả trước các URL mới, nghiên cứu đề xuất ứng dụng học máy để nhận diện URL giả mạo dựa trên các đặc trưng nội tại.

Nghiên cứu sử dụng các mô hình như Logistic Regression, Random Forest, XGBoost và ANN để huấn luyện trên hai tập dữ liệu khác nhau, kiểm tra hiệu suất và tính ổn định. Đặc biệt, công cụ SHAP được tích hợp để xác định các đặc trưng quan trọng, từ đó tối ưu hóa mô hình, giảm thời gian xử lý và tăng khả năng diễn giải kết quả.

Mô hình học máy cho phép phát hiện URL giả mạo chưa từng xuất hiện, nâng cao hiệu quả phòng chống tấn công. Hướng phát triển tương lai bao gồm tích hợp deep learning (LSTM, BERT), phân tích hành vi người dùng, cập nhật đặc trưng tự động và nâng cao nhận thức cộng đồng để bảo vệ an toàn thông tin mạng.

## TÀI LIỆU THAM KHẢO

1. Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). "Machine learning based phishing detection from URLs." Expert Systems with Applications.

2. Dang Thi Mai. (2024). APPLYING THE AUTOENCODER MODEL FOR URL PHISHING DETECTION. *Vinh University Journal of Science,* 53(4A), 124–132. https://doi.org/10.56824/vujs.2024a143a

3. Hamidi, H., & Sayah, A. (2024). Combining Machine Learning Algorithms to Detect Phishing URLs: A Stacking Approach. *1nInternational Journal of Engineering,* 38(8), 1939–1952. https://doi.org/10.5829/ije.2025.38.08b.18

4. Haq, Q. E. U., Faheem, M. H., & Ahmad, I. (2024). Detecting Phishing URLs Based on a Deep Learning Approach to Prevent Cyber-Attacks. *Applied Sciences, 1*4(22), 10086. https://doi.org/10.3390/app142210086