Fingerprinting mobile browser

Link: <https://helda.helsinki.fi/bitstream/handle/10138/273478/fingerprinting-mobile-browsers.pdf>

1. **Introduction**

Trong điện toán nối mạng, lấy dấu vân tay có thể được coi là một tập hợp các kỹ thuật thu thập thông tin về các thiết bị từ xa nhằm mục đích nhận dạng. Lần lượt, thông tin được thu thập này được sử dụng để tạo dấu vân tay của thiết bị. Nhìn chung, phương pháp lấy dấu vân tay cho rằng dấu vân tay rất đa dạng và ổn định. Bản chất đa dạng của dấu vân tay đảm bảo rằng không có hai máy nào có cùng dấu vân tay và độ ổn định đòi hỏi dấu vân tay vẫn giữ nguyên theo thời gian. Trong thực tế, sự đa dạng và ổn định không hoàn toàn đạt được trong thực tế.

Các phương pháp lấy dấu vân tay có thể được phân loại thành hai nhóm dựa trên các chế độ thu thập dữ liệu của chúng; cụ thể là thụ động và chủ động. Trong nhóm thụ động, phương pháp lấy dấu vân tay không truy vấn trực tiếp các thiết bị để thu thập thông tin được sử dụng để lấy dấu vân tay. Chúng thường dựa trên sự khác biệt tinh tế trong các tham số truyền thông đáp ứng yêu cầu của máy khách-máy chủ. Ví dụ, vân tay trình điều khiển thiết bị không dây [1] và dấu vân tay của thiết bị vật lý từ xa với độ lệch của đồng hồ phần cứng [2]. Trong nhóm hoạt động, phương pháp lấy dấu vân tay chủ động truy vấn các thiết bị được nhắm mục tiêu để thu thập thông tin được sử dụng để lấy dấu vân tay. Chẳng hạn, các phương thức này cài đặt hoặc chạy các tập lệnh trên thiết bị để thu thập thông tin từ thiết bị và gửi thông tin đó đến các máy chủ tạo dấu vân tay. Một ví dụ về hoạt động lấy dấu vân tay là kỹ thuật lấy dấu vân tay trình duyệt của Eckersley et al. [3].

Cả hai kỹ thuật lấy dấu vân tay chủ động và thụ động đều có thể được sử dụng trong một số ứng dụng, như theo dõi người dùng, xác thực lớp thứ hai và tùy chỉnh trải nghiệm người dùng. Các công ty quảng cáo Internet sử dụng dấu vân tay của thiết bị để theo dõi các hoạt động trực tuyến của người dùng cả trong một trang web và trên các trang web [4]. Ngoài ra, dấu vân tay của thiết bị có thể được sử dụng trong xác thực Web. Theo Alaca et al. [5], họ đã xác định được 29 cơ chế lấy dấu vân tay thiết bị để tăng cường xác thực web dựa trên mật khẩu.

Vân tay trình duyệt, một kỹ thuật lấy dấu vân tay phổ biến, có thể được sử dụng ở chế độ thụ động và chủ động. Phương pháp lấy dấu vân tay trình duyệt thụ động dựa trên các yêu cầu Web để thu thập thông tin nhận dạng về thiết bị khách. Ví dụ: họ có thể đặt số nhận dạng duy nhất trong cookie HTTP và trích xuất địa chỉ IP và tác nhân người dùng từ các yêu cầu HTTP. Trong chế độ hoạt động, các phương thức lấy dấu vân tay triển khai một trang web chạy mã JavaScript trên máy khách. Các tập lệnh này trích xuất thông tin liên quan đến cài đặt, cấu hình và các đặc điểm khác của trình duyệt. Ví dụ: lấy dấu vân tay dựa trên sự khác biệt tinh tế trong các đặc điểm hiển thị văn bản của phần tử canvas HTML5 [7]. Những phương pháp lấy dấu vân tay này có thể được áp dụng cả trong môi trường di động và máy tính để bàn.

Ngày nay, hàng tỷ người truy cập Internet trên điện thoại di động và một phần lưu lượng đáng kể đến từ các trình duyệt di động. Do đó, trình duyệt di động có thể được sử dụng làm cổng vào thiết bị di động để truy cập các tài nguyên cơ bản cho mục đích lấy dấu vân tay và sau đó dấu vân tay có thể được sử dụng để khử tên người dùng trên Internet. Tuy nhiên, các hệ thống vân tay trình duyệt được sử dụng rộng rãi được thiết kế cho môi trường máy tính để bàn, ví dụ: Panopticlick [3]. Do đó, thông tin nhận dạng được thu thập bằng các hệ thống này không bao gồm đầy đủ các tính năng độc đáo của điện thoại di động, ví dụ: cảm biến thiết bị. Mục tiêu chính của luận án này là khám phá các tính năng có thể lấy dấu vân tay bổ sung trong ngữ cảnh di động và đánh giá tác động của chúng trên trình duyệt di động.

Công việc này được thúc đẩy bởi ba tác phẩm sau đây. Đầu tiên, chúng tôi được thúc đẩy bởi công việc của Eckersley et al. [3] từ năm 2010. Công việc này cung cấp thuật toán vân tay rất hiệu quả, nhưng nó không bao gồm các API trình duyệt hiện đại. Hơn nữa, công việc này nó không nhằm vào các trình duyệt di động. Thứ hai, công trình nghiên cứu gần đây của Al-Fannah et al. [8] so sánh độ vân tay của các trình duyệt phổ biến nhất cho cả thiết bị di động và máy tính để bàn. Công việc này bao gồm các API hiện đại và nó đã bao phủ điện thoại di động. Tuy nhiên, công việc này không định lượng được tác động của các tính năng bổ sung đối với dấu vân tay được tính toán. Thứ ba, Bojinov và cộng sự. [9] sử dụng bài đọc gia tốc của thiết bị cho các thiết bị lấy dấu vân tay. Cụ thể, họ quan sát rằng các API cảm biến trình duyệt có thể được tận dụng bằng thuật toán vân tay. Công việc này mở rộng các tác phẩm này bằng cách định lượng tác động của các tính năng bổ sung lên dấu vân tay được tính toán

Phần còn lại của công việc luận án được tổ chức như sau. Phần 2 thảo luận về các công việc liên quan, Phần 3 mô tả số lượng các kỹ thuật trích xuất tính năng, Phần 4 thảo luận ngắn gọn về công cụ lấy dấu vân tay, Phần 5 trình bày về sự tiến hóa thời gian của dấu vân tay của trình duyệt di động, Phần 6 thảo luận về dấu vân tay trong tự nhiên và cuối cùng, Phần 7 trình bày Kết luận.

1. **Related Work**
2. **Mobile Browser Feature Extraction Techniques**
3. **HTTP Header**

Giao thức truyền siêu văn bản (HTTP) là một giao thức cấp ứng dụng cho các hệ thống thông tin hypermedia phân tán, hợp tác [14]. Nó là nền tảng để truyền thông dữ liệu cho World Wide Web và nó xác định phương thức yêu cầu, mã lỗi và tiêu đề của nó. Các tiêu đề HTTP cho phép máy khách và máy chủ truyền thông tin bổ sung với yêu cầu hoặc phản hồi bằng các cặp giá trị tên. Tên của tiêu đề được phân tách khỏi giá trị bằng dấu hai chấm đơn. Ví dụ: một thông báo yêu cầu có thể chứa các thông tin sau: accept\_encoding: "gzip, deflate, br". Tên tiêu đề tên là accept\_encoding và giá trị của nó là "gzip, deflate, br". Các trình duyệt web gửi các tiêu đề HTTP đến mọi trang web mà chúng kết nối và chúng không kiểm soát được người dùng. Những tiêu đề này tiết lộ một số thông tin xác định về trình duyệt. Do đó, tiêu đề HTTP có thể được sử dụng để tạo dấu vân tay trình duyệt. Trong phần này, chúng tôi đã thảo luận về các tiêu đề HTTP có liên quan đến dấu vân tay của trình duyệt web.