

**ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

Vũ Huy Thiện

**ĐÁNH GIÁ MỨC ĐỘ PHỔ BIẾN CỦA
QUẢNG CÁO VÀ QUẢNG CÁO ĐỘC HẠI
TRÊN WEBSITE**

Ngành: Công nghệ thông tin

Chuyên ngành: Truyền dữ liệu và mạng máy tính

Mã số:

**TÓM TẮT LUẬN VĂN THẠC SĨ CÔNG
NGHỆ THÔNG TIN**

Hà Nội - 2016

MỤC LỤC

CHƯƠNG 1. TỔNG QUAN VỀ QUẢNG CÁO TRỰC TUYẾN.....	4
1.1. Quảng cáo.....	4
1.2. Quảng cáo trực tuyến.....	4
1.3. Các tác nhân tham gia quảng cáo trực tuyến.....	4
1.4. Các hình thức quảng cáo trực tuyến.....	5
1.5. Phương pháp tính chi phí quảng cáo.....	9
1.6. Lợi ích của quảng cáo trực tuyến.....	9
1.7. Những vấn đề lo ngại về quảng cáo trực tuyến.....	9
1.8. Kỹ thuật đưa nội dung quảng cáo lên website.....	9
CHƯƠNG 2. QUẢNG CÁO ĐỘC HẠI.....	12
2.1. Khái niệm.....	12
2.2. Phương thức làm việc.....	12
2.3. Các loại quảng cáo độc hại.....	12
2.4. Kỹ thuật thực hiện mã độc.....	13
CHƯƠNG 3. ĐÁNH GIÁ MỨC ĐỘ PHỔ BIẾN CỦA QUẢNG CÁO VÀ QUẢNG CÁO ĐỘC HẠI.....	15
3.1. Mục đích.....	15
3.2. Quy trình.....	15
3.3. Chi tiết các bước thực hiện.....	15
KẾT LUẬN.....	19
TÀI LIỆU THAM KHẢO.....	20

MỞ ĐẦU

Ngày nay, với sự phát triển mạnh mẽ của Internet, các dịch vụ trên mạng Internet ngày càng phát triển không ngừng, trong đó dịch vụ quảng cáo trực tuyến trên website cũng trở lên phổ biến hơn trước. Các tổ chức, doanh nghiệp, cá nhân luôn muốn quảng cáo sản phẩm, dịch vụ của mình trên các trang Web nổi tiếng, có lượng người dùng truy nhập lớn nhằm thu hút sự quan tâm của người dùng, khách hàng. Việc phân phối quảng cáo trực tuyến trên website được thực hiện qua nhiều phương thức khác nhau: thông qua các thẻ HTML, mã kịch bản Javascript, Iframe,... Tuy nhiên, vấn đề đặt ra là trong số các quảng cáo trên website thì cái nào là an toàn cho người dùng, cái nào gây hại (quảng cáo độc hại) cho người dùng khi họ click chuột vào chúng. Quảng cáo độc hại có thể lây lan các phần mềm độc hại hay chuyển hướng người dùng đến các website độc hại nhằm thực hiện các hành vi độc hại của kẻ tấn công đối với người dùng và hệ thống.

CHƯƠNG 1. TỔNG QUAN VỀ QUẢNG CÁO TRỰC TUYẾN

1.1. Quảng cáo

Quảng cáo là hình thức tuyên truyền được trả tiền để thực hiện việc giới thiệu thông tin về sản phẩm, dịch vụ, công ty hay ý tưởng. Quảng cáo là hoạt động truyền thông không trực tiếp giữa người với người mà trong đó người muốn truyền thông phải trả tiền cho các phương tiện truyền thông đại chúng để đưa thông tin đến thuyết phục hay tác động đến người nhận thông tin.

1.2. Quảng cáo trực tuyến

Quảng cáo trực tuyến [1], còn được gọi là tiếp thị trực tuyến hoặc quảng cáo trên Internet, là một hình thức tiếp thị và quảng cáo trong đó sử dụng Internet để cung cấp các thông điệp tiếp thị quảng cáo cho người tiêu dùng.

1.3. Các tác nhân tham gia quảng cáo trực tuyến

1.3.1. *Publisher*

Publisher là các nhà xuất bản trang web, thực hiện việc bán không gian quảng cáo trên các trang web của mình, quản lý và thu lợi nhuận từ việc bán quảng cáo đó.

1.3.2. *Advertiser*

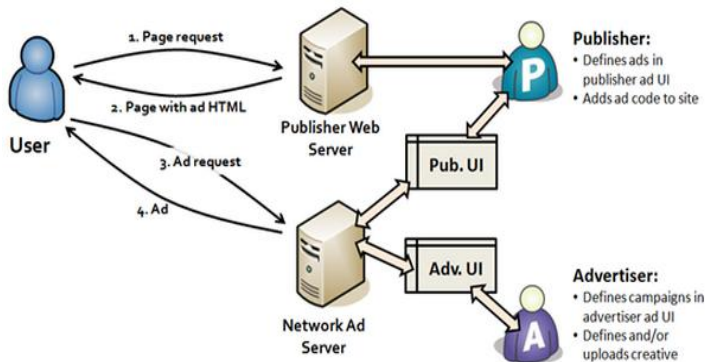
Advertiser (nhà quảng cáo) là một công ty hoặc cá nhân có Website bán hàng hóa hoặc dịch vụ trực tuyến,

Advertiser hợp tác với Publisher để quảng cáo hàng hóa, dịch vụ giúp họ.

1.3.3. Ad network - Advertising Network

Ad network chỉ một mạng quảng cáo liên kết nhiều website lại và giúp nhà quảng cáo (Advertiser) – có thể đăng quảng cáo cùng lúc trên nhiều website của nhiều Publisher khác nhau.

1.3.4. Mô hình hoạt động của quảng cáo trực tuyến



1.4. Các hình thức quảng cáo trực tuyến

1.4.1. Quảng cáo hiển thị

Quảng cáo hiển thị truyền tải thông điệp quảng cáo một cách trực quan bằng việc sử dụng văn bản, logo, hình ảnh động, video, hình ảnh, hoặc các phương tiện đồ họa khác.

1.4.2. Biểu ngữ quảng cáo

Biểu ngữ quảng cáo trên web điển hình là những quảng cáo đồ họa hiển thị trong một trang web. Nhiều biểu ngữ quảng cáo được phân phối bởi một máy chủ quảng cáo trung tâm.

1.4.3. Khung quảng cáo

Những khung quảng cáo là dạng đầu tiên của biểu ngữ quảng cáo trên web. Từ “biểu ngữ quảng cáo” thường đề cập đến khung quảng cáo truyền thống.

1.4.4. Quảng cáo Pop-ups/pop-unders

Kiểu quảng cáo pop-up được hiển thị trên một khung cửa sổ mới của trình duyệt và nằm trên cửa sổ trang web mà người dùng đang xem. Còn kiểu quảng cáo pop-under mở một cửa sổ trình duyệt nằm dưới trang web mà người dùng đang xem.

1.4.5. Quảng cáo trôi

Kiểu quảng cáo trôi (floating ad hoặc overlay ad) là một dạng của kiểu quảng cáo đa phương tiện xuất hiện chồng lên nội dung của trang web.

1.4.6. Quảng cáo mở rộng

Kiểu quảng cáo mở rộng (expanding ad) thay đổi kích thước theo một điều kiện đã định trước.

1.4.7. Biểu ngữ đánh lừa

Biểu ngữ đánh lừa (trick banners) là một dạng quảng cáo mô phỏng các thông báo thường gặp của các chương trình mà người xem hay sử dụng (ví dụ như thông báo có email mới) nhằm dụ dỗ người xem click vào biểu ngữ quảng cáo.

1.4.8. Quảng cáo chiếm khe

Kiểu quảng cáo chiếm khe (interstitial ad) hiển thị trước khi người dùng có thể truy cập vào nội dung cần xem, đôi khi họ còn phải đợi một khoảng thời gian để quảng cáo hiển thị.

1.4.9. Quảng cáo dạng văn bản

Quảng cáo dạng văn bản (text ads) hiển thị liên kết dựa trên văn bản, nó có thể hiển thị độc lập với trang web hoặc có thể chèn liên kết dựa trên nội dung mà người dùng đang xem.

1.4.10. Tiếp thị mô tơ tìm kiếm

Tiếp thị mô tơ tìm kiếm (Search Engine Marketing, viết tắt là SEM) được thiết kế để gia tăng khả năng hiển thị của một trang web trong các trang kết quả tìm kiếm.

1.4.11. Tiếp thị trên mạng xã hội

Tiếp thị trên mạng xã hội (Social media marketing – SMM) là dạng xúc tiến thương mại được thực hiện thông qua các dịch vụ mạng xã hội.

1.4.12. Quảng cáo trên thiết bị di động

Quảng cáo trên thiết bị di động (mobile advertising) bao gồm các dạng quảng cáo tĩnh hoặc sử dụng các đa phương tiện để hiển thị quảng cáo thông qua các phương tiện: tin nhắn SMS (Short Message Service) hay MMS (Multimedia Messaging Service), các trang web dành riêng cho di động, các ứng dụng hoặc game cho di động.

1.4.13. Quảng cáo bằng email

Quảng cáo bằng email là một dạng quảng cáo bằng cách gửi đi các thông tin giới thiệu sản phẩm & dịch vụ qua email đến nhiều người trong danh sách của nhà quảng cáo.

1.4.14. Quảng cáo trực tuyến đã được phân loại

Đây là dạng quảng cáo trên các trang sản phẩm/dịch vụ cụ thể và có tính chuyên biệt.

1.4.15. Phần mềm quảng cáo

Là một dạng phần mềm khi được cài đặt vào máy tính của người dùng, nó sẽ tự động hiển thị quảng cáo.

1.4.16. Tiếp thị liên kết

Tiếp thị liên kết xảy ra khi những nhà quảng cáo liên kết với các bên thứ 3 để tạo ra các khách hàng tiềm năng cho họ.

1.5. Phương pháp tính chi phí quảng cáo

1.5.1. CPM (Cost Per Mile)

1.5.2. CPC (cost per click)

1.5.3. CPA (Cost Per Action)

1.5.4. Chi phí cố định (Fixed cost)

1.6. Lợi ích của quảng cáo trực tuyến

1.6.1. Về vấn đề chi phí

1.6.2. Khả năng đo lường

1.6.3. Các định dạng (kiểu) quảng cáo

1.6.4. Khả năng hướng mục tiêu

1.6.5. Phạm vi phân phối

1.6.6. Tốc độ triển khai

1.7. Những vấn đề lo ngại về quảng cáo trực tuyến

1.7.1. Gian lận trong quảng cáo

1.7.2. Thay đổi về công nghệ

1.7.2.1. Máy khách không đồng nhất

1.7.2.2. Chương trình ngăn chặn quảng cáo

1.7.2.3. Kỹ thuật chống hướng mục tiêu

1.8. Kỹ thuật đưa nội dung quảng cáo lên website

Việc đưa nội dung quảng cáo lên website thực ra là việc hiển thị những đoạn văn bản, hình ảnh, âm thanh, video trên website để quảng bá hình ảnh, dịch vụ của tổ chức, cá nhân.

1.8.1. Sử dụng HTML

```
<a  
href="http://s.tuoitre.vn/DominoX/click.ashx?ID=6349;ff  
f4ce74-30df-4acb-bd09-c0e1da34211a;http://tuoitre.vn/"  
target="_blank">
```

```
</a>
```

1.8.2. Sử dụng Javascript

```
<a href="https://www.thegioididong.com/galaxy-  
note-7" onclick="jQuery.ajax({ url:  
'/www.thegioididong.com/bannertracking?bid=8626&amp;r  
='+ (new Date).getTime(), async: true, cache: false });">
```

```
</a>
```

1.8.3. Sử dụng XML

```
<?xml version="1.0" encoding="utf-8" ?>
```

```

    <Advertisements
xmlns="http://schemas.microsoft.com/AspNet/AdRotator-
Schedule-File">
    <Ad>
    <ImageUrl>~/images/Contoso_ad.gif</ImageUrl>
    <NavigateUrl>http://www.contoso-
ltd.com</NavigateUrl>
    <AlternateText>Ad for Contoso, Ltd. Web
site</AlternateText>
    <Impressions>100</Impressions>
    </Ad>
    </Advertisements>

```

1.8.4. Sử dụng ActionScript đối với Flash

```

<embed      src="banner.swf"    quality="high"
type="application/x-shockwave-flash"    width="950"
height="91"
pluginspage="http://www.macromedia.com/go/getflashpl
ayer" allowScriptAccess="always" />

```

CHƯƠNG 2. QUẢNG CÁO ĐỘC HẠI

2.1. Khái niệm

Quảng cáo độc hại [2] là việc sử dụng quảng cáo trực tuyến để lây lan phần mềm độc hại.

2.2. Phương thức làm việc

Các trang web hoặc các nhà xuất bản web vô tình kết hợp một quảng cáo bị hỏng hoặc độc hại vào trang của họ. Sau khi quảng cáo được đặt ra, và người dùng bắt đầu nhấp vào nó, máy tính của họ có thể bị nhiễm: “người dùng nhấp chuột vào quảng cáo để truy cập vào trang quảng cáo, và thay vào đó là trực tiếp bị nhiễm hoặc chuyển hướng đến một trang web độc hại. .

2.3. Các loại quảng cáo độc hại

2.3.1. Quảng cáo pop-up cho tải lừa đảo

Chẳng hạn như các chương trình chống virus giả mạo mà cài đặt phần mềm độc hại trên máy tính hay đơn giản chỉ là các cửa sổ bật lên (pop-up) giả mạo hiển thị máy tính của bạn đang bị nhiễm virus và yêu cầu người dùng phải cài đặt các phần mềm khác để diệt virus.

2.3.2. Trong văn bản hoặc trong nội dung quảng cáo

Các từ được gạch chân đôi trong quảng cáo trực tuyến được gọi là quảng cáo trong văn bản

2.3.3. Tải về tự động (Drive-by Download)

Tải về tự động làm việc thông qua việc khai thác lỗ hổng trình duyệt, các thành phần plug-in trên trình duyệt.

2.3.4. Thành phần giao diện

2.3.5. iframe ẩn

Iframe ẩn là một cách để những kẻ tấn công để ẩn các đối tượng được sử dụng để phát tán phần mềm độc hại.

2.3.6. Mạng phân phối nội dung

Một mạng phân phối nội dung (CDN) là một máy chủ quảng cáo của bên thứ ba cung cấp nội dung tới các khu vực khác nhau trên web.

2.3.7. Biểu ngữ độc hại

2.3.8. Quảng cáo của bên thứ ba

2.3.9. Các ứng dụng của bên thứ ba

2.3.10. Cướp liên kết

2.4. Kỹ thuật thực hiện mã độc

2.4.1. Che dấu mã

Để trốn tránh việc phân tích phát hiện các hành vi nguy hiểm, một số quảng cáo độc hại sử dụng các kỹ thuật che dấu để tránh quét dựa trên chữ ký phần mềm chống virus [4].

2.4.2. Chuyển hướng URL

Nhiều quảng cáo độc hại tự động chuyển hướng tới URL khác. Khi trình duyệt truy cập một URL, thời gian phản hồi của URL này sẽ tự động hướng dẫn các trình duyệt để truy cập vào một hoặc nhiều URL khác mà không ảnh hưởng đến nội dung của màn hình vào người sử dụng.

2.4.3. Khai thác các lỗ hổng

Hệ thống Web độc hại hoặc các lỗ hổng trình duyệt nguy hiểm là lợi thế của chương trình, khi người dùng truy cập vào các trang này, các thủ tục này có thể tải về trojan hoặc phần mềm độc hại khác trên máy chủ dẫn đến tình trạng không an toàn của máy chủ.

CHƯƠNG 3. ĐÁNH GIÁ MỨC ĐỘ PHỔ BIẾN CỦA QUẢNG CÁO VÀ QUẢNG CÁO ĐỘC HẠI

3.1. Mục đích

Mục đích của chương này là đánh giá mức độ phổ biến của quảng cáo và quảng cáo độc hại trên các websites quan tâm (trong phạm vi luận văn này sẽ khảo sát các website của Việt Nam). Sau khi có được danh sách các website sẽ tiến hành khảo sát, đánh giá nhằm đưa ra con số thống kê về quảng cáo, quảng cáo độc hại theo lĩnh vực, tên miền

3.2. Quy trình

- Đầu tiên, lập danh sách website (là đầu vào), sau đó dùng PhantomJS để lấy URL của từng website.

- Tiếp theo, sử dụng EasyList, AdblockPlus để lấy URL quảng cáo trong số các URL đã thu thập được ở bước đầu tiên.

- Sau đó, sử dụng VirusTotal để lấy quảng cáo độc hại trong số các URL quảng cáo ở bước thứ hai.

- Cuối cùng phân tích, đánh giá kết quả.

3.3. Chi tiết các bước thực hiện

3.3.1. Lập danh sách website được khảo sát

Đầu tiên, chúng tôi thu thập tập URLs các websites của Việt Nam theo một số lĩnh vực. Tiếp theo, chúng tôi phát triển một crawler (con bọ) để lấy các tài nguyên trên từng website. Crawler chúng tôi phát triển dựa vào Phantomjs. Phantomjs là một “headless browser” (trình duyệt web không sử dụng giao diện đồ họa người dùng).

Ở đây, chúng tôi thu thập tổng cộng 237 website thuộc các lĩnh vực: mua bán, rao vặt (45 website); giáo dục, đào tạo (53 website); du lịch (51 website); y tế (33 website); báo chí (55 website).

3.3.2 Thu thập quảng cáo

Sử dụng EasyList và Adblock Plus để thu thập URL, URL quảng cáo.

EasyList bao gồm các vùng và các mẫu URL cho các máy chủ liên quan đến quảng cáo, và được sử dụng bởi các plugin trình duyệt là Adblock Plus để chặn quảng cáo. Sau khi thu được hầu hết tất cả các URL của IFRAME trên webpage, chúng tôi sử dụng một API của Adblock Plus để xem trong những URL của iframe mà chúng tôi thu thập, thì URL nào là URL quảng cáo. Nếu kết quả trả về mà URL bị block (khóa), thì URL đó là URL quảng cáo, còn ngược lại không phải URL quảng cáo.

3.3.3 Nhận diện quảng cáo độc hại

Sau khi đã có được các URL quảng cáo, chúng tôi thực hiện lần lượt gửi từng URL quảng cáo đã thu thập được ở bước 3.3.2 lên VirusTotal [5] để biết quảng cáo là độc hại hay lành tính.

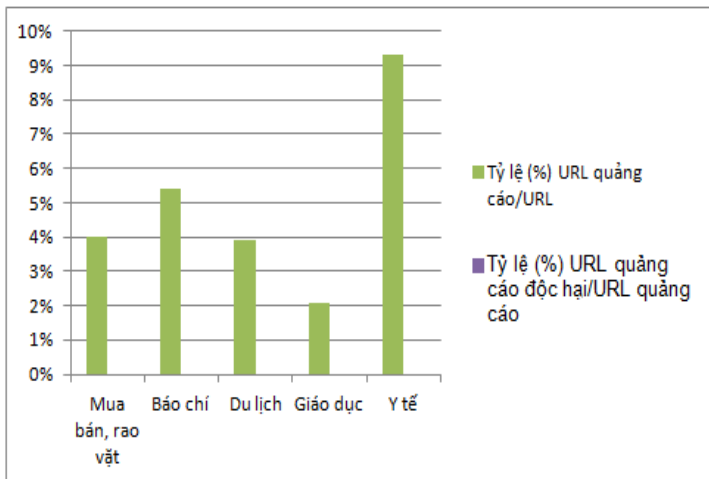
3.3.4. Tổng hợp và phân tích kết quả

Sau khi thu thập và chạy các kịch bản để nhận biết quảng cáo, quảng cáo độc hại; bảng dưới đây thống kê tập URL, URL quảng cáo, URL quảng cáo độc hại của các website theo lĩnh vực và tên miền.

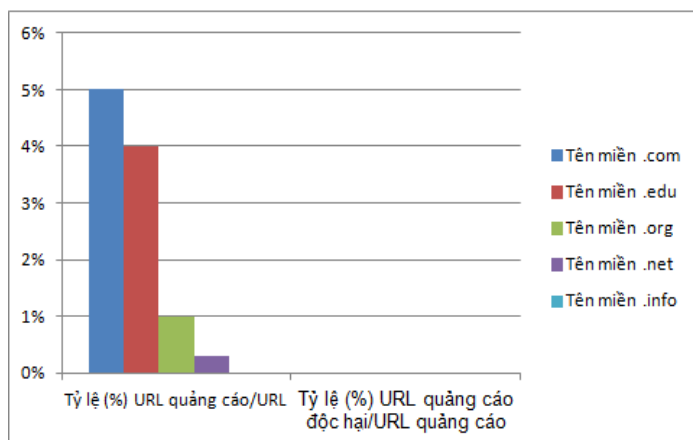
✓ **Thống kê dưới dạng bảng:**

✓ **Thống kê dưới dạng biểu đồ trực quan**

- Thống kê theo lĩnh vực



- Thống kê theo tên miền



KẾT LUẬN

Kết quả của luận văn sẽ giúp các cơ quan quản lý nhà nước về lĩnh vực Thông tin và Truyền thông (Bộ Thông tin và Truyền thông hoặc Sở Thông tin và Truyền thông các tỉnh/thành phố trực thuộc Trung ương) theo dõi, kiểm tra được danh sách các websites của các tổ chức, cá nhân thuộc phạm vi quản lý của cơ quan mình; làm cơ sở phục vụ cho công tác quản lý, chỉ đạo đối với các tổ chức, cá nhân là chủ sở hữu hoặc các tổ chức, cá nhân có liên quan đến thực hiện quảng cáo trên website, sao cho việc quản lý, vận hành quảng cáo trên website được tốt hơn, sạch hơn và đảm bảo an toàn cho người dùng duyệt web.

Mặt khác, kết quả của luận văn cũng tư vấn giúp các tổ chức, cá nhân trong việc lựa chọn đăng ký tên miền (.com, .net, .org,...) dựa trên những tên miền có tỷ lệ % quảng cáo/quảng cáo độc hại nhiều nhất, ít nhất.

THAM KHẢO

- [1] https://en.wikipedia.org/wiki/Online_advertising, theo Wikipedia website.
- [2] <https://en.wikipedia.org/wiki/Malvertising>, theo Wikipedia website.
- [3] Aditya K Sood, Richard J Enbody, Michigan State University, “Malvertising – exploiting web advertising”, Computer Fraud & Security, p11-p15, 2011.
- [4] Fuqiang Yu, “Malicious URL Detection Algorithm based on BM Pattern Matching”, International Journal of Security and Its Applications Vol.9, No.9 (2015).
- [5] The Dark Alleys of Madison Avenue, “Understanding Malicious Advertisements”, 2014.
- [6] Đỗ Gia Quân, *Phát hiện và ngăn chặn quảng cáo độc hại dựa vào URL*, Khoá luận tốt nghiệp 2016, Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội, Hà Nội, 2016.