

Quero

An Effective Defense against Intrusive Web Advertising

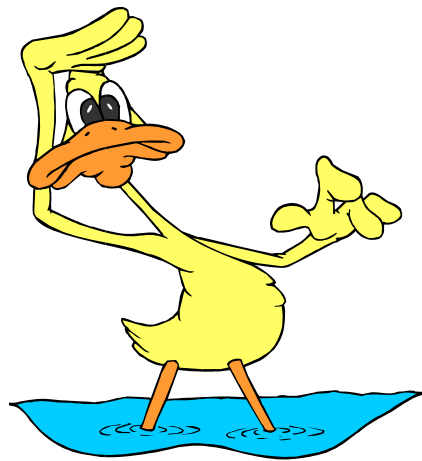
Viktor Krammer ^{1,2}
1 Secure Business Austria
2 Vienna University of Technology

<http://www.quero.at/>

Quero

Advertising is the art of convincing people to spend money they don't have for something they don't need.

Will Rogers



Quero **Agenda**

- **What is Quero**
- **Web Attacks**
- **Ad Categorization**
- **Content Filter**
- **Web Study**
- **Ad Blocking Discussion**
- **Conclusion**

Quero **What is Quero**

Quero is an add-on for Windows Internet Explorer 5.5+

Quero is a combined navigation/search/find toolbar

Quero is an ad blocker

Quero is a set of Browser Helper Objects written in VC++

CQueroBand : IObjectWithSite, IDeskBand, IInputObject

CQueroFilter : IInternetProtocol, IInternetProtocolSink



Quero **Agenda**


- What is Quero
- **Web Attacks**
- Ad Categorization
- Content Filter
- Web Study
- Ad Blocking Discussion
- Conclusion

Quero **Web Attacks**

Categories

- Malicious code execution / injection
- Information disclosure / identity theft / spyware
- **Intrusive advertising** / adware

Methods

- Vulnerability-based (browser / Web app / input validation / SQL injection / XSS / session hijacking, etc.)
 - Configuration mistakes (directory browsing, source code disclosure, etc.)
 - Social Engineering (Phishing)
 - Obfuscation (polymorphism)
- 

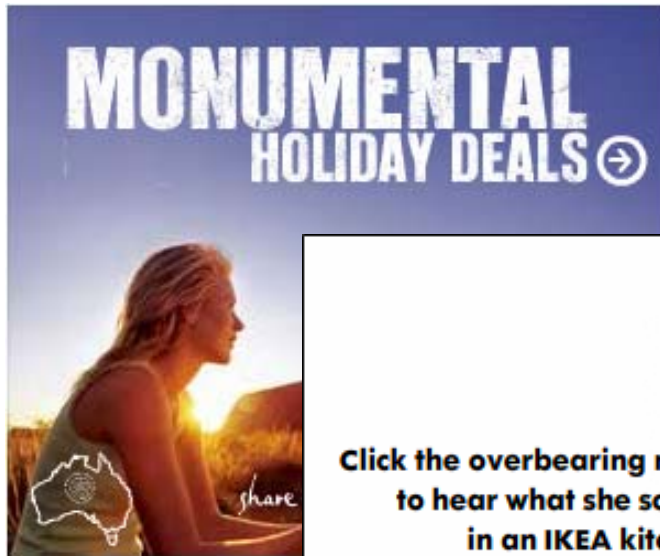
Quero **Agenda**

- What is Quero
- Web Attacks
- **Ad Categorization**
- Content Filter
- Web Study
- Ad Blocking Discussion
- Conclusion

Quero **Dimensions**

- Medium
 - text, image, video, audio
- Size
- Integration with Web page
- Interactivity
- Intrusiveness
- Privacy Impact

Quero Banners



Real Estate



CLOSE

Click the overbearing mother-in-law to hear what she sounds like in an IKEA kitchen.

SOUND ON/OFF



IKEA
Love your home

LOST 57.2LBS*
WEIGHT WATCHERS
NE.
LEARN MORE ▶
Results not typical



Canada and the USA on sale.

CANADA
USA

BOOK NOW!

From: CALGARY

MONTREAL
\$13
ONE WAY

FREDERICTON
HALIFAX
ONE WAY

50% Bonus Aeroplan® Miles.

aircanada.com



Quero **Video ads**



Quero Text ads

Cadillac For

ContentLink™ Advertisement

Post Reply

10-05-06, 01:38 AM

 **fpmesiIII**
Cadillac Owners Fanatic
Cadillac(s): 1996 [Cadillac Seville](#) STS (164k)

get rid of contentlink?

does anyone know how to get rid of annoying content link randomly put in messages. i knew how to get rid of the old link junk, intellitxt i believe. but this new one has me fondled.

Ads by Google

[Traffic Secrets 2.0](#)
Get yours today for only \$397 USD with a 100% MoneyBack Guarantee
www.LatestTrafficSecrets.com

[100% Free AntiSpyware](#)
Free Spyware Scan. Save your PC from Harmful Spyware. 100% Free
spywareremover.com

[Want Website Visitors?](#)
Special: 10,000 Visitors Only \$7.99 Traffic Includes Free 24/7 Stats
www.hitscheap.com



[Want to Quit your Job?](#)
Make 3k-5k per Week From Home. See How
www.dmeadath.com

Debt Consolidation

Need help with debt consolidation?
Receive free information & advice.
www.debtreliefusa.org

International Wholesale

Thousands of Prequalified Suppliers
Trade Leads, Products & Companies
www.Alibaba.com

Bad Credit Loans

Don't Let a Bad Score Stop You -
Refinance or Apply for a Loan Today
www.superpages.com

Sponsored Links

Quero Pop-ups



Quero

Sticky ads

本本实惠

BenQ 250G 移动硬盘 799元
天敏炫影2 “包”你满意
鼎好5周年真情回馈
选本如何做到不惧奸商
三星MP3 DV超值拍卖周

产品报价 产品论坛

软件下载
视频
冠军
硬盘799
关闭
IBM X3010/ X3100 促销

台电16G杀毒加密U盘199元
个人信息安全隐患调查
Stix闪控升级PC
商用PC破天荒夏日超低价
金士顿大眼king仔抢先拿
七喜3599商用

Canon 佳能
&c. 电众

DELL
YOURS IS HERE
Studio 15 (R510909CN)
销售热线
800-858-2035
400-884-9428

打破
• 8系C
• 15"

你的力量

关闭

Quero Ad Games



Wie heißt die
Hauptstadt von China?

A

Neu-Delhi

B

Peking

C

Tokio

D

Seoul

Quero Interstitials

IGN.COM

Continue to [IGN.com: Video Games, Cheats, Movies and More](#)

[Disable this ad](#)



<ADVERTORIAL>

MuseumsQuartier Wien



Das MuseumsQuartier Wien ist ein einzigartiges und atemberaubendes Areal zeitgenössischer Kunst und Kultur

<ADVERTORIAL> Bezahlte Werbung

FONDS-SPEZIAL

Immobilienfonds mit Vertrauens-Malus

[9]



Die US-Immo- und -Finanzkrise lässt die Anleger auch vor Immobilienfonds zurückschrecken - zu Unrecht, wie nicht nur die Branche meint

■ THEMENSCHWERPUNKT

In Pictures

IN
ASSOC
WITH

Nikon



DAILY SNAPSHOT

Through the lens

CNN highlights the best pictures each day from around the world

Destinations



Vienna

[All 242 hotels in Vienna »](#)

★★★★★	Radisson SAS Palais Hotel	From € 119
★★★★☆	Embassy	From € 85
★★★★☆	NH Vienna Airport	From € 99
★★★★☆	Hotel Astoria	From € 108
★★★★☆	Hotel Park Inn Vienna	From € 89

Quero

Example Web site

猫扑通行证 用户名 密码 自动登录 登录 游客 | 注册 | 忘记密码 北京·上海·广州

大杂烩 我的空间 猫游记 猫扑 www.mop.com 贴贴 大话战国 宽频

新闻 | 体育 | 奥运 | NBA | 汽车 | 车型 | 科技 | 财经 | 房产 | 社会 | 军事 | 人物 | 图片 | 论坛 | 访谈 | 家电 | 彩铃 | 汽车大杂烩
娱乐 | 明星 | 影视 | 音乐 | 爱听 | 潮流 | 爱鞋 | 时尚 | 女性 | 母婴 | 型男 | 游戏 | 魔兽 | 动漫 | 视觉 | 城市 | Flash | 在线看漫画

雪碧 透心凉 心飞扬

“雪碧”及“Sprite”是可口可乐公司的注册商标

北京2008年奥运会合作伙伴

国家地震局称北京今晚将发生...
部分贷款...
4月CPI同比...
德媒造假...
阜阳白...
南京女子...
沈阳性文...
何洁雷...
明星,请穿好衣服!

海马3运动版正式上市

猫扑网友第一时间发布全国地震见闻
追忆红色苏联暴力美学 华语电影获奖...
带着想法去旅行!
余嘉静:骗亲戚两百万 游圣地亚哥的...
我老婆不可告人的隐私 吃兰州拉面的...
秀秀新认识的夜店公主 外国人违规了照样罚...
P当猪肉... 上海“电梯...
在线... 校长领导...
阵容秀... 一对90后...
论坛

蜀山OL

ideapad U110

送

当上奥运火

汽车大杂烩

想就仙踪

Unknown Zone

Intrusiveness (subjective)

Type	Intrusiveness
Sticky / Layer	+++++
Pop-ups	+++++
Interstitials	++++
Banner, dynamic	+++
Video	+++
Ad Games	+++
Banner, static	++
Text	+
Content Sponsoring	+

Pop-ups and overlapping layer ads are most intrusive for me.

Quero **Agenda**

- What is Quero
- Web Attacks
- Ad Categorization
- **Content Filter**
- Web Study
- Ad Blocking Discussion
- Conclusion

Quero **Content Filter [1]**

- Allows the user to opt-out of online ads
 - Flash ads
 - Banner ads
 - Text ads
 - Layer ads
 - Frame-based ads
 - Pop-ups



Ad blocking software is a response to abusive activity by advertisers.

Quero **Content Filter [2]**

- Web browser add-on (implemented for IE 5.5+)
- Works by a simple but effective set of rules
- No daily filter rule updates necessary
- Static and behavioral analysis (code could be obfuscated)
- Blocks the content from being downloaded
- User interface plays an important part of the solution since the user has to deal with false positives
 - Toggle filter on or off
 - Whole sites can be whitelisted
 - Temporarily unblock a Web site (Version 4.5)

Quero **Feature selection**

- **Media Type** (HTML element type, Object classid)
- **Size** (width x height)
- **Dynamic Creation** (document.createElement etc.)
- **Different Domain** (content domain != current domain)
- **Different Host** (content host != current host)
- **URL tokens** (delimiters [_=;:/.-*?])
- **Target URL** (Link analysis)
- **Presence of HTTP redirection**
- **DOM tree and page position**
- **Element attributes**
- **Image analysis**
- **Surrounding text, ...**

Quero **Rule-based classifier**

- Block all Flash-based content by default
- Block unwanted pop-ups
- Block ad banners based on their size
- Block content that comes from well-known ad providers
- Block images based on ad-related keywords in their URL
- Block absolute-positioned DIV or IFRAME elements that are dynamically created
- Do not block content on sites that are whitelisted

Currently, about 30 fine-grained rules are sufficient to block over 90% of online ads.

Quero **Implementation**

- Asynchronous pluggable **MIME filter** for „text/html“
- **Interpositioning script calls**
window.open
document.write
document.createElement
etc.
not supported by Internet Explorer platform
exploited COM architecture
hacked vtable of several objects ;-)
- **URL pattern matching**
 $F = \{ \text{Patterns} \}, |F|=m, U, |U|=n, \text{IsAddURL}(U,F)?$
 $O(n)$

Quero **Agenda**

- What is Quero
- Web Attacks
- Ad Categorization
- Content Filter
- **Web Study**
- Ad Blocking Discussion
- Conclusion

Quero Web Study

- Crawled front page of Alexa Global Top 500 Web sites
- Semi automatic classification into ad & non-ad content
- Focused on image classification

Type	Count	Sites	in %	Avg	Median
Pop-ups	29	26	5.2%	1.1	1
Flash	690	215	42.8%	3.2	2
non-ads	110	66	13.1%	1.7	1
ads	580	182	36.3%	3.2	2
Images	15981	456	90.8%	35.0	27
non-ads	14350	452	90.0%	31.7	24
ads	1631	249	49.6%	6.6	3
Text ads	72	44	8.8%	1.6	1
Google	68	42	8.4%	1.6	1
IntelliTXT	3	3	0.6%	1	1
DIV layers	534	75	14.9%	7.1	2
Web bugs	730	230	45.8%	3.2	2

84% of Flash animations were ads.

10% of images were ads.

Table 1. Object Types

Quero **Testing Hypotheses**

Are script generated images likely to be ads?

Are images hosted on another server are likely to be ads?

Are images with a query string in their URL likely to be ads?
etc.

Feature	Precision	Recall
Script generated	28.5%	32.1%
Different 2nd level domain	15.0%	30.6%
Different 3rd level domain	13.8%	83.6%
Query string in URL	34.4%	13.3%
Image dimensions in URL	34.9%	28.6%
Ad pattern in URL	95.5%	63.3%

Table 2. Characteristics of Image Ads

$$\text{precision}(A \Rightarrow B) = P(B|A) = \frac{|\{A \wedge B\}|}{|\{A\}|}$$

$$\text{recall}(A \Rightarrow B) = P(A|B) = \frac{|\{A \wedge B\}|}{|\{B\}|}$$

„Ad pattern in URL“ is still by far the best indicator.

Keyword	Precision	Recall
ads	98.0%	31.5%
banner	87.0%	16.4%
adv	79.5%	6.6%
click	76.5%	4.7%
upload	15.2%	4.2%
adimages	100.0%	3.3%
banners	94.4%	3.2%
doubleclick	100.0%	2.9%
adimg	79.5%	2.2%
adserver	100.0%	2.1%

Table 3. Keyword Analysis

Banner Dimensions

Width	Height	Precision	Recall
106	50	98.7%	6.0%
300	250	100.0%	5.7%
120	60	80.72%	5.4%
728	90	100.0%	5.0%
468	60	100.0%	4.4%
88	31	56.52%	4.2%
114	23	80.85%	3.0%
120	90	14.91%	1.4%
186	47	80.0%	1.3%
120	600	100.0%	1.3%

Table 4. Banner Dimensions

Country →	us	cn	jp	de	tw	uk	hk	br	cz	vn
Sites	188	84	23	16	15	11	9	9	9	8
in %	37.5%	16.7%	4.6%	3.2%	3.0%	2.2%	1.8%	1.8%	1.8%	1.6%
Sites with ads	103	69	16	10	11	7	5	5	8	7
in %	54.8%	82.1%	69.6%	62.5%	73.3%	63.6%	55.6%	55.6%	88.9%	87.5%
# Ads	557	1071	74	77	148	52	26	52	38	194
Avg per site	5.4	15.5	4.6	7.7	13.5	7.4	5.2	10.4	4.8	27.7
Ad Pixels										
Avg per site	179,442	211,459	99,480	195,174	193,093	188,719	264,134	77,966	147,705	605,949
Avg per object	55,888	31,588	35,421	42,444	19,544	36,419	70,381	25,994	86,316	21,208
Image Filter										
Precision	96.5%	98.7%	98.0%	96.9%	95.7%	88.9%	94.1%	100.0%	100.0%	94.9%
Recall	97.1%	73.9%	94.1%	91.2%	82.6%	94.1%	100.0%	100.0%	50.0%	83.6%

Table 5. Ads per Country

Filter	Type	Count	Ads	Blocked	FP	Precision	Recall
Quero Version 3.4	Pop-ups	29	29	29	0	100.0%	100.0%
	Flash	690	580	690	110	84.1%	100.0%
	Images	15981	1631	1454	52	96.4%	86.0%
	Text	71	71	71	0	100.0%	100.0%
	Overall	16771	2311	2244	162	92.8%	90.1%
Adblock Plus Filter: EasyList 495 rules (2007-06-07)	Pop-ups	29	29	29	0	100.0%	100.0%
	Flash	690	580	316	4	98.7%	53.8%
	Images	15981	1631	1103	151	86.3%	58.4%
	Text	71	71	71	0	100.0%	100.0%
	Overall	16771	2311	1519	155	89.8%	59.0%
Adblock Plus Filter: Dr. Evil 525 rules (2007-06-05)	Pop-ups	29	29	29	0	100.0%	100.0%
	Flash	690	580	204	1	99.5%	35.0%
	Images	15981	1631	792	150	81.1%	39.4%
	Text	71	71	71	0	100.0%	100.0%
	Overall	16771	2311	1096	151	86.2%	40.9%

Table 6. Filter Results and Comparison

Quero

Ad Blocker: turned off

猫扑通行证 用户名 密码 自动登录 登录 游客 | 注册 | 忘记密码 北京·上海·广州

大杂烩 我的空间 猫游记 猫扑 www.mop.com 贴贴 大话战国 宽频

新闻 | 体育 | 奥运 | NBA | 汽车 | 车型 | 科技 | 财经 | 房产 | 社会 | 军事 | 人物 | 图片 | 论坛 | 访谈 | 家电 | 彩铃 | 汽车大杂烩
娱乐 | 明星 | 影视 | 音乐 | 爱听 | 潮流 | 爱鞋 | 时尚 | 女性 | 母婴 | 型男 | 游戏 | 魔兽 | 动漫 | 视觉 | 城市 | Flash | 在线看漫画

雪碧 透心凉 心飞扬

“雪碧”及“Sprite”是可口可乐公司的注册商标

关闭 2008 梦想

猫扑网友第一时间发布全国地震见闻

追忆红色苏联暴力美学 华语电影获奖力作 喝雀巢咖啡

带着想法去旅行!

余嘉静：骗亲戚两百万 游圣地亚哥的... 送

我老婆不可告人的隐私 吃兰州拉面的... 当上奥运火

秀秀新认识的夜店公主 外国人违规了照样罚... 汽车大杂烩

当猪肉... 上海“电梯... 论坛

蜀山OL

ideaPad U110

明星,请穿好衣服!

部分贷款... 4月CPI同... 德媒造假... 阜阳白富... 南京女子... 沈阳性文... 何洁雷人...

海马3运动版正式上市

Unknown Zone

猫扑通行证
用户名
密码
☐ 自动登录
登录
[游客](#) | [注册](#) | [忘记密码](#)
北京·上海·广州

大杂烩

我的空间

猫游记

猫扑
www.mop.com

贴贴

大话战国

宽频

[新闻](#) | [体育](#) | [奥运](#) | [NBA](#) | [汽车](#) | [车型](#) | [科技](#) | [财经](#) | [房产](#) | [社会](#) | [军事](#) | [人物](#) | [图片](#) | [论坛](#) | [访谈](#) | [家电](#) | [彩铃](#) | [汽车大杂](#)
[娱乐](#) | [明星](#) | [影视](#) | [音乐](#) | [爱听](#) | [潮流](#) | [爱鞋](#) | [时尚](#) | [女性](#) | [母婴](#) | [型男](#) | [游戏](#) | [魔兽](#) | [动漫](#) | [视觉](#) | [城市](#) | [Flash](#) | [在线看漫画](#)

news

国家地震局称北京今晚将发生余震传言不属实

明星,请穿好衣服!

海马3运动版正式上市 售价10.78万元

- 部分贷款炒房者月供二三十万被套牢
- 4月CPI同比涨8.5% 食品价格涨
- 德媒造假图片诬称中国警察监视喇嘛
- 阜阳白富美举报人死亡鉴定被指有缺陷
- 南京女子称被城管队长骑在身上暴打
- 沈阳性文化节女模情趣内衣秀(组图)
- 何洁雷人照(图) 姚明再惨遭PS(图)

原汁原味街拍大盘点

猫扑网友第一时间发布全国地震见闻

- 追忆红色苏联暴力美学
- 华语电影获奖大片赏析
- 媒体诬蔑女友死于游戏
- 金莎冒充粉丝当托儿?
- 余嘉静:骗亲戚两百万
- 游圣地亚哥的海洋公园
- 我老婆不可告人的隐私
- 吃兰州拉面的风险太大
- 秀秀新认识的夜店公主
- 外国人违规了照样罚款

猫扑搜索
网页
论坛
博客
网友
圈子

选秀 猫王 李嘉欣 课本
蜀山OL

2008 梦想

喝雀巢咖啡

当上奥运火

汽车大杂烩

Quero **Agenda**

- What is Quero
- Web Attacks
- Ad Categorization
- Content Filter
- Web Study
- **Ad Blocking Discussion**
- Conclusion

Quero **Ad Blocking Discussion**

Ad blocking has become an integral part of

- **Web browsers**
 - Pop-up Blocker
 - IE8 InPrivate Blocking
 - Opera Content Blocking
- **Add-ons**
 - Adblock Plus (Firefox)
 - SafariBlock (Safari)
 - Quero (IE)
- **Internet Security suites**
- **Proxies**

Quero **IE8 InPrivate Blocking**

- New feature in IE8 aimed for blocking tracking scripts, 1x1 tracking pics, etc.
- Addition to InPrivate browsing (prevents recording history entries, new cookies, form data, passwords, etc.)
- Self-learning algorithm or subscription based
- Blocking rule: third-party content „seen“ on more than 10 different sites is regarded as a potential privacy threat
- Can also be used to block common ad servers



Why Web advertising works, almost

- Natural business model of information based media
- Keeps the Web essentially free
- Cheaper than conventional advertising
- Interactive possibilities (at least *Link* to advertised content)
- Success measurable to some degree (conversion rate)
- Personalizable (may affect privacy)
- Different pay models based on „real“ impressions or clicks

Quero **What's wrong though**

- The interactive nature of the Web has been abused for intrusive advertising and user tracking/profiling
- Web advertising business model is driven by greed (increasing impressions, clicks) rather than balancing the interests of readers, publishers and advertisers
- Rich-media ads slow down the actual Web page (waste of bandwidth?)
- Online ads deliver most hacks [Finjan]
- The average European click-through rate (CTR) has dropped to 0.18% according to [ADTECH 2007]
- Excessive Web advertising has lead to *Banner Blindness* [Benway]

Quero **Agenda**

- What is Quero
- Web Attacks
- Ad Categorization
- Content Filter
- Web Study
- Ad Blocking Discussion
- **Conclusion**

Quero **Conclusion**

- Intrusive Web advertising is regarded as one of the major annoyances of today's Web
- Ad blocking software has a long tradition: Google made pop-up blocking popular; Adblock Plus most popular extension for Firefox; IE8 has InPrivate Blocking
- The URL is still by far the best indicator for content filtering
- Only a small number of rules is sufficient to block most ads
- Some sites are enforcing regulations on ads served to them
- Ad blocking users are usually geeks that would not click on ads anyway

Quero

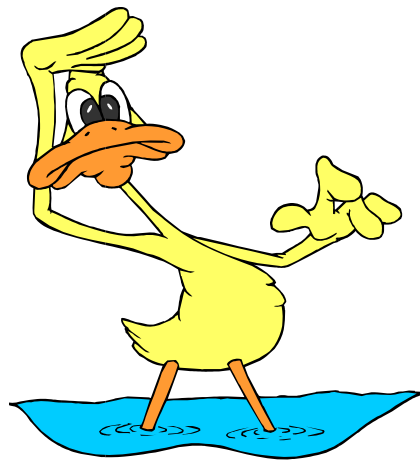
IE Team Chat

Q: IE7 crashes on these websites: {...}

A: Do you have phishing filter on?

Q: yes

A: Turn it off.



Quero



Viktor Krammer

support@quero.at

<http://www.quero.at/>