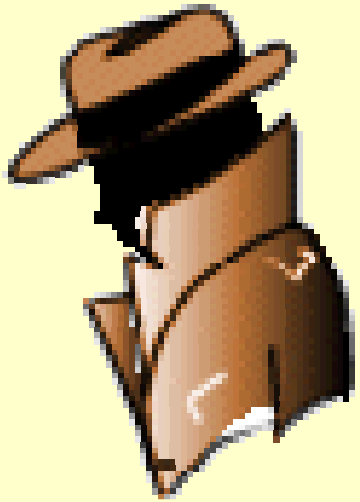
Google Search Obfuscator

Mauro Rappa



Personal Profile

Ingegneria TLC

Sistemista *NIX

Docente Sophos Puremessage

Appassionato di sicurezza informatica

Talk presso: LinuxDay – Smau – Infosecurity – WebbIT – Italian Debian Conference

Vi siete mai chiesti…

Le tracce delle nostre ricerche..

Ricerca semplice sul Sito

http://www.google.it/search?hl=it&q=**firenze**

Ricerca dal plugin di FF

http://www.google.it/search?q=**firenze**+&ie=utf-8&aq=t&rls=org.mozilla:it:official&client=firefox

N.B. Vi sono diversi parametri ‘accessori’

Idea

Tentare di ‘nascondere’ i termini di ricerca rintracciabili nei log di un proxy, rendendo vana una semplice lettura delle URL registrate.

Idea

Tentare di ‘nascondere’ i termini di ricerca rintracciabili nei log di un proxy, rendendo vana una semplice lettura delle URL registrate.

Risultato

Un programma in Javascript 1.3, che offusca le query richieste a Google ed utilizza tecniche di evasione per rallentare la ricerca nei log

N.B. l'URL richiesta è, purtroppo, riproducibile (scoprendo i termini di ricerca usati)

Analisi di Google

Ho studiato, dal punto di vista applicativo,

come funziona la ricerca su Google.

Sono emerse diverse peculiarità che ho sfruttato nel mio programma:

1. I termini di ricerca sono passati tramite la variabile **q**, che può essere ripetuta
2. Vi sono moltissimi parametri accessori, non strettamente necessari

Analisi di Google 2

3. La variabile **hl=it** specifica la lingua dell'utente
4. Le variabili possono essere codificate in 'URL Encode', e sono perfettamente interpretate da Google
5. La chiave di ricerca **non** è case-sensitive

Prima versione di G.S.O.

Semplice encoding della query:

http://www.google.it/search?q=&rls=org.mozilla:it:official&q=%66%69%72%65%6E%7A%65&hl=it&oe=utf-8

Evoluzione di G.S.O.

Per evitare la ricerca ‘semplice’ attraverso le RegExp, i parametri vengono mescolati in maniera casuale.

http://www.google.it/search?q=&safe=images&s_filetype=&s_filetype=&hl=it&s_filetype=&q=%46%69%52%45%6E%7A%45&rls=org.mozilla:it:official

Gestione chiavi di ricerca multiple

Generalmente si utilizzano più chiavi di ricerca assieme per affinare la ricerca.

La tecnica attuata è stata quella di spezzettare la stringa di ricerca in singole parole e di riproporle a Google con diverse variabili **q**.

Problema: non devo cambiare l'ordine dei termini di ricerca!

Esempio

Ricerca per “Firenze Eprivacy”

http://www.google.it/search?q=&hl=it&as_occt=any&q=%66%69%52%65%6E%7A%65&as_epq=&q=%45%70%52%49%76%41%43%79&

Interroghiamo Google Fiji!

Possiamo interrogare qualunque server Google, basta solo specificare il nostro linguaggio di preferenze tramite il parametro **hl=it**.

In questo modo le ricerche dentro il log saranno ulteriormente complicate poiché dirette ad indirizzi diversi.

Interroghiamo Google Fiji! (2)

I server di Google sono numerosi ed ognuno di loro fornirà i risultati nella lingua che impostiamo, quindi possiamo randomizzare anche i server specificando la lingua desiderata

http://72.14.221.104/search?q=&rls=org.mozilla:it:official&q=%66%49%72%45%4E%7A%65+%45%50%52%69%56%61%63%59&hl=it&oe=utf-8

N.B. la ricerca per server contattato diviene vana.

Termini di ricerca Fake

Per ‘complicare’ l’URL, possiamo utilizzare la variabile di Google che ricerca esclude un termine

http://www.google.it/search?q=&q=%46%49%72%45%6E%5A%65&as_rights=&hl=it&&q=%45%50%72%49%76%61%43%59&as_oq=&as_eq=-LKNLHG

Ulteriori evoluzioni

Usare altri servizi di Google per ricerche di altro tipo:

http://images.google.it/search?hl=it&q=firenze

http://news.google.it/search?hl=it&q=firenze

Ulteriori evoluzioni (2)

Fare le richieste usando Google come traduttore ed reinterrogandolo:

http://translate.google.com/translate?u=http%3A%2F%2Fwww.google.it%2Fsearch%3Fhl%3Dit%26q%3Dfirenze%26meta%3D&langpair=en%7Cit&hl=it&newwindow=1&ie=UTF-8&oe=UTF-8&prev=%2Flanguage_tools

Riepilogo

- Chiavi codificate
- Distribuzione randomica delle variabili
- Server di richiesta variabili
- Interrogazioni ‘nascoste’ come traduzioni, ricerca immagini o news

Da:

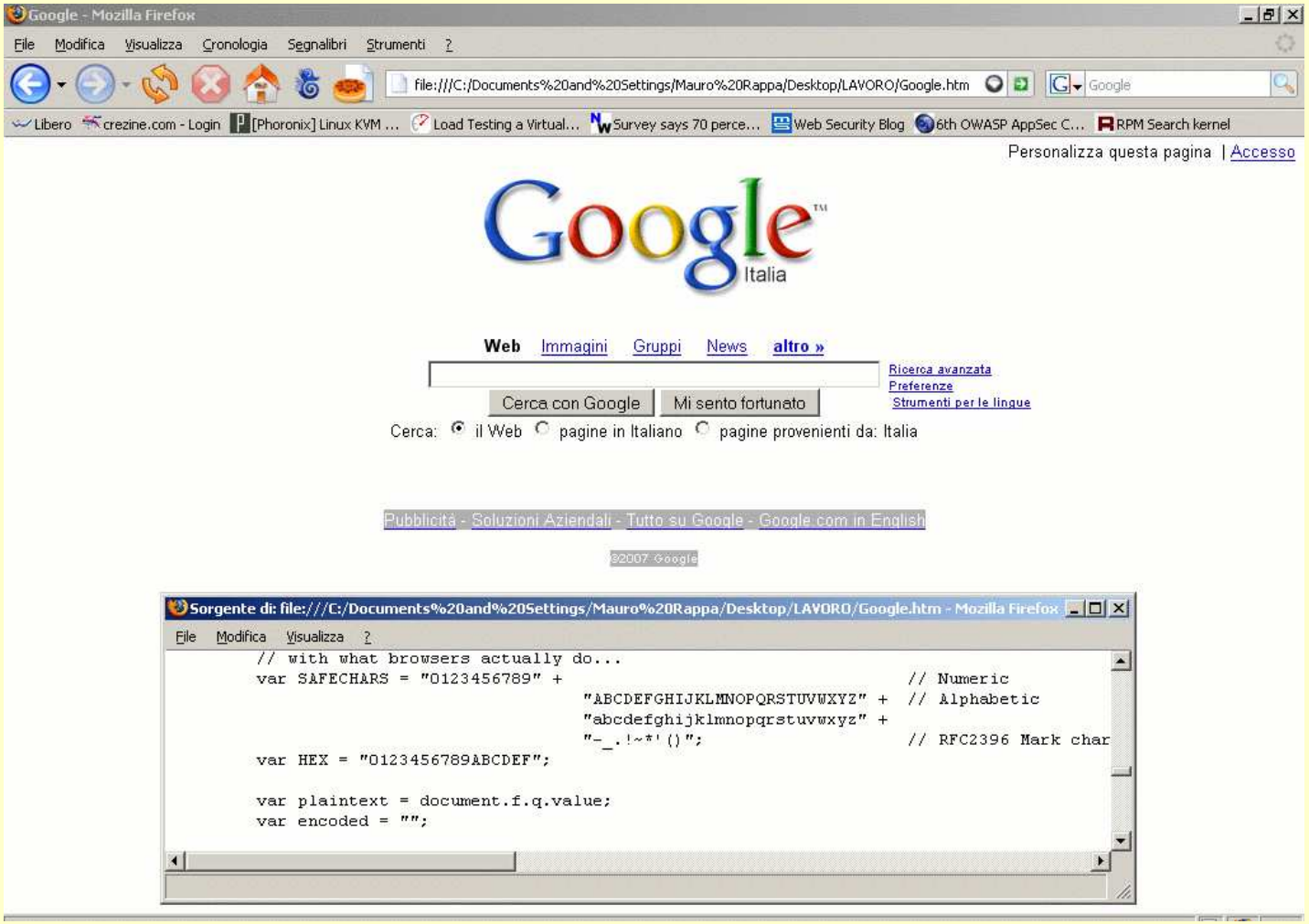
http://www.google.it/search?hl=it&q=firenze

A:

http://72.14.221.104/search?q=& as_occt=any &

q=%66%49%72%45%4E%7A%65&ie=utf-8

Volete usare G.S.O?



Creiamo un Firefox Plugin?

Il Javascript non richiede nessun adeguamento per ‘girare’ come estensione.

Il lavoro per l’integrazione in FF, non è complicato, è solo laborioso (Xml di configurazione, costruzione package, test di retrocompatibilità…).

Ogni aiuto è ben accetto!!

Google….

“Google è il dominio tecnologico derivato dalla ricerca scientifica che si fa strumento di gestione della conoscenza, espressione diretta della tecnocrazia.”

Tratto da:

The Dark Side of Google a.k.a. luci e ombre di Google – Ippolita

www.ippolita.net

Grazie

Per l’attenzione e la pazienza.

Scrivetemi a mauro@crezine.com

Slide e programma disponibili su:
www.firefoxsolutions.com/gso/