"OPENNESS OF SEARCH ENGINE": A CRITICAL FLAW IN SEARCH SYSTEMS; A CASE STUDY ON GOOGLE, YAHOO AND BING

Katuru SM Kalyana Chakravarthy

ISTQB Certified QA Engineer
Hyderabad, A.P, India
Kalyan.Katuru@gmail.com

ABSTRACT

There is no doubt that Search Engines are playing a great role in Internet usage. But all the top search engines Google, Yahoo and Bing are having a critical flaw called "Openness of a Search Engine". An Internet user should be allowed to get the search results only when requested through Search engine's web page but the user must not be allowed to get the search results when requested through any web page that does not belong to the Search Engine. Only results of a search engine should be available to the Internet user but not the Search Engine. This paper explains the critical flaw called "Openness of Search Engine" with a case study on top 3 search engines 'Google', 'Yahoo' and 'Bing'. This paper conducts an attack based test using J2EE framework and proves that 'Google' passed the test and it strongly protects its Critical Search System, where 'Yahoo' and 'Bing' are failed to protect their Search Engines. But previously 'Google' also had other high severity issues with the Openness of search engine; this paper reveals those issues also. Finally this paper appeals strongly to the all top Search Engines to fix their critical flaws of "Openness of Search Engine".

1. INTRODUCTION

1.1 What is Openness of a Search Engine?

This paper defines Openness of a Search Engine as "Ability of a search Engine to provide the Search results to any HTTP page's request".

1.2 Critical loss with openness of a Search Engine

A flaw with openness of a Search Engine is that user can retrieve the search results by requesting from any HTML page that does not belong to search engine. When user is able to retrieve the search results by requesting from any HTML page, he can modify the branding simply and can own the search results with his branding. Critical loss with Openness of Search Engine is that hacker's engine can display the search results of an original search engine with fake branding.

Natarajan Meghanathan, et al. (Eds): ITCS, SIP, JSE-2012, CS & IT 04, pp. 303–314, 2012.
© CS & IT-CSCP 2012
DOI : 10.5121/csit.2012.2128

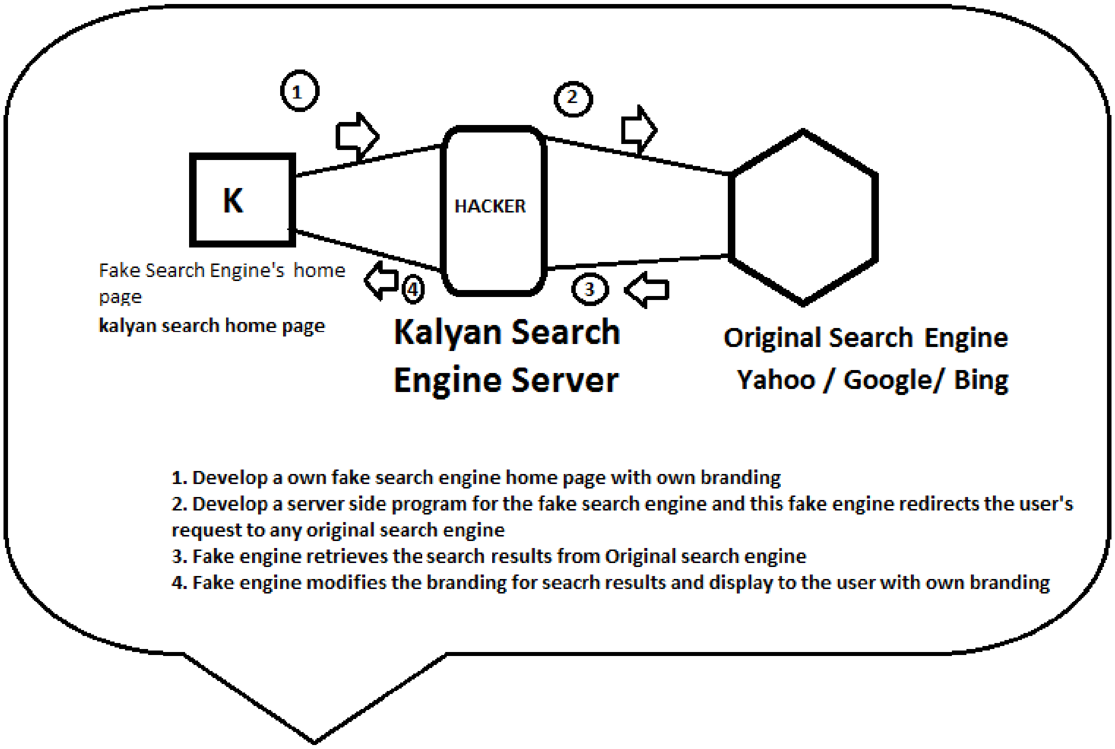
This causes hacker to own a complete search engine by just developing own branded HTML page and simple java code.

1.3 What is wrong?

"Search Engine should not be open to a HTML page that does not belong to the Search Engine".

1.4 Technique to hack the Search Results

In this paper a technique called “Own a Search Engine” is proposed to prove the Openness of top 3 Search Engines "Google, Yahoo and Bing". Following figure gives the overview of the proposed technique to hack the search results from a search engine and to display with the fake branding.


1
2
K
HACKER
Fake Search Engine's home page
kalyan search home page
4
3
Kalyan Search Engine Server
Original Search Engine Yahoo / Google/ Bing

1. Develop a own fake search engine home page with own branding
2. Develop a server side program for the fake search engine and this fake engine redirects the user's request to any original search engine
3. Fake engine retrieves the search results from Original search engine
4. Fake engine modifies the branding for seacrh results and display to the user with own branding

Fig 1.1 Overview of proposed technique to hack search results

1.5 How to solve this critical issue

Google has the solution but others do not have. But Google also had some loop holes. Following sections discuss all these issues in a detailed manner. Finally this paper strongly appeals to all top search engines to fix this critical issue.

2. OWN A SEARCH ENGINE

2.1 A proposed technique "Own A Search Engine" to hack the search results of a search engine

This paper proposes a technique called “Own A Search Engine” to hack the search results of a Search Engine and to display with own branding.

2.1.1 Steps involved in the technique "Own a Search Engine"

Setp-1:
Develop a search engine client HTML Page called 'ownsearchengine.html'.

Following is the designed page:

```
<HTML>
<HEAD>
</HEAD>
<BODY bgcolor=pink>
<CENTER>
<form action="./ownsearchengine.jsp">
<H1>KALYAN SEARCH ENGINE</H1>
<Input type=text name=searchstring>
<Input type=submit>
</CENTER>
</BODY>
</HTML>
```

Setp-2: Develop a server side program (a JSP) named as ‘ownsearchengine.jsp’ and write the pieces of code in the JSP such that it should do following actions.

Step-2a: Read the input string from the page 'ownsearchengine.html'.

Following is the piece of code that can be used to read the input string from the page 'ownsearchengine.html'

```
        String ss=request.getParameter("searchstring");
```

Here “searchstring” is the Input parameter name that is passed to the JSP from the HTML page.

Step-2b: Redirect the input string to an original search engine.

Following is the piece of code that can be used to redirect the input string to an original search engine.

```
        String url="http://search.yahoo.com/search?q="+ss;
        URL urlobj= new URL(url);
        URLConnection uc =urlobj.openConnection();
```

This step is heart of this technique. Generally all search engines' request uses 'q' as input name for search string. So if the search engine is called by providing the “q” as parameter in the url, search results can be retrieved.

Step-2c: Catch the search results from the original search engine.
Following is the piece of code that can be used to catch the search results from the original search engine

```
        BufferedReader in = new BufferedReader(

        new InputStreamReader(

        uc.getInputStream()));

        String inputLine="", output= “”;

        while ((inputLine = in.readLine()) != null)

        output=output+inputLine;

        in.close();
```

Step-2d: Modify the branding of the search results to display own branding. Following is the piece of code that can be used to modify the branding of the search results to display own branding

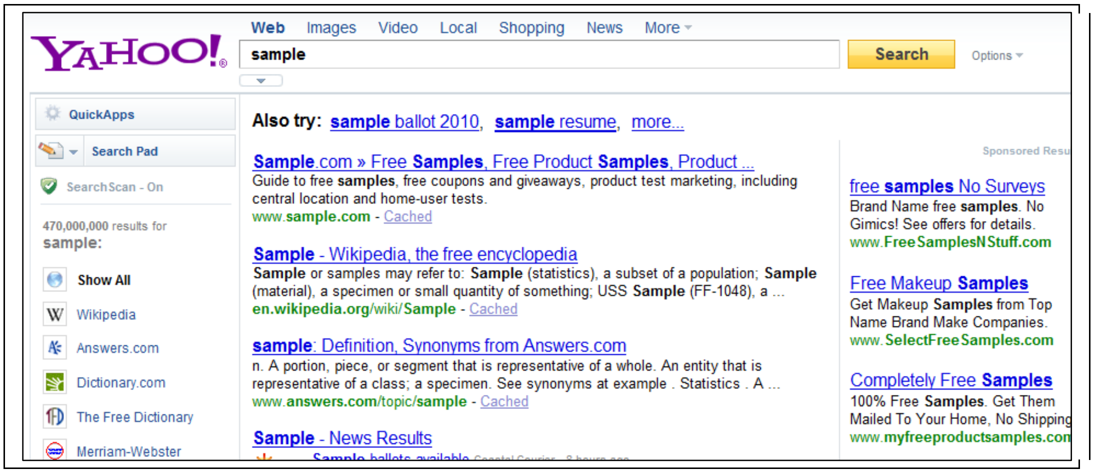
```
        output=output.replaceAll("Yahoo!", “Kalyan!")
        /* This is example code to replace branding of yahoo search results*/
```

Step-2e: Display the hacked search results to the end user with own branding. Following is the piece of code used to display the hacked search results to the end user with own branding

```
<%=output %>
```

3. CASE STUDY-I: OWN THE ENGINE "SEARCH.YAHOO.COM"

In this section proposed technique “Own a Search Engine” is applied to hack the search results of Seach.Yahoo.com. First observe the result of Yahoo search engine when searched for some test string “sample”.



sample

470,000,000 results for **sample**:

Also try: **sample** ballot 2010, **sample** resume, more...

Sample.com » Free **Samples**, Free Product **Samples**, Product ...
Guide to free **samples**, free coupons and giveaways, product test marketing, including central location and home-user tests.
www.**sample**.com - Cached

Sample - Wikipedia, the free encyclopedia
Sample or samples may refer to: **Sample** (statistics), a subset of a population; **Sample** (material), a specimen or small quantity of something; USS **Sample** (FF-1048), a ...
en.wikipedia.org/wiki/**Sample** - Cached

sample: Definition, Synonyms from Answers.com
n. A portion, piece, or segment that is representative of a whole. An entity that is representative of a class; a specimen. See synonyms at example . Statistics . A ...
www.answers.com/topic/**sample** - Cached

Sample - News Results

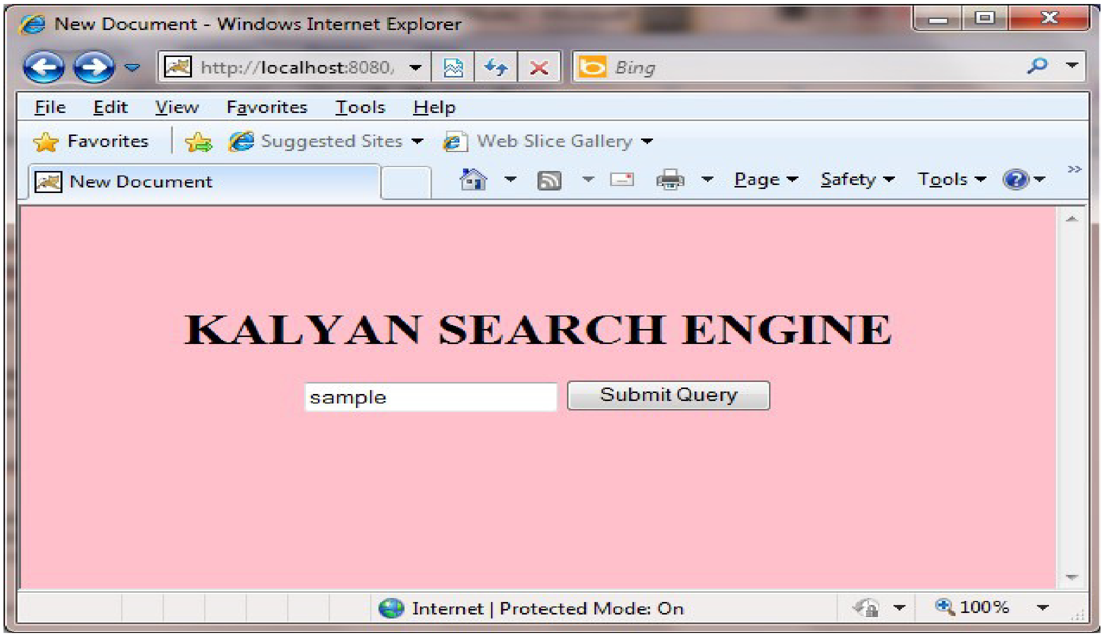
free **samples** No Surveys
Brand Name free **samples**. No Gimics! See offers for details.
www.FreeSamplesNStuff.com

Free Makeup **Samples**
Get Makeup **Samples** from Top Name Brand Make Companies.
www.SelectFreeSamples.com

Completely Free **Samples**
100% Free **Samples**. Get Them Mailed To Your Home, No Shipping
www.myfreeproductsamples.con

Fig 3.1 Original results of Yahoo search engine

Now apply the “Own a Search Engine” technique, when user access a fake search engine and tried to search for a string “sample”, the fake engine internally redirects the search request to the original Yahoo engine and retrieve the results and modify the branding of results to display own branding. Finally user will see the results of Yahoo search engine with the fake branding.
Following figures explains the hacking of Yahoo search results.

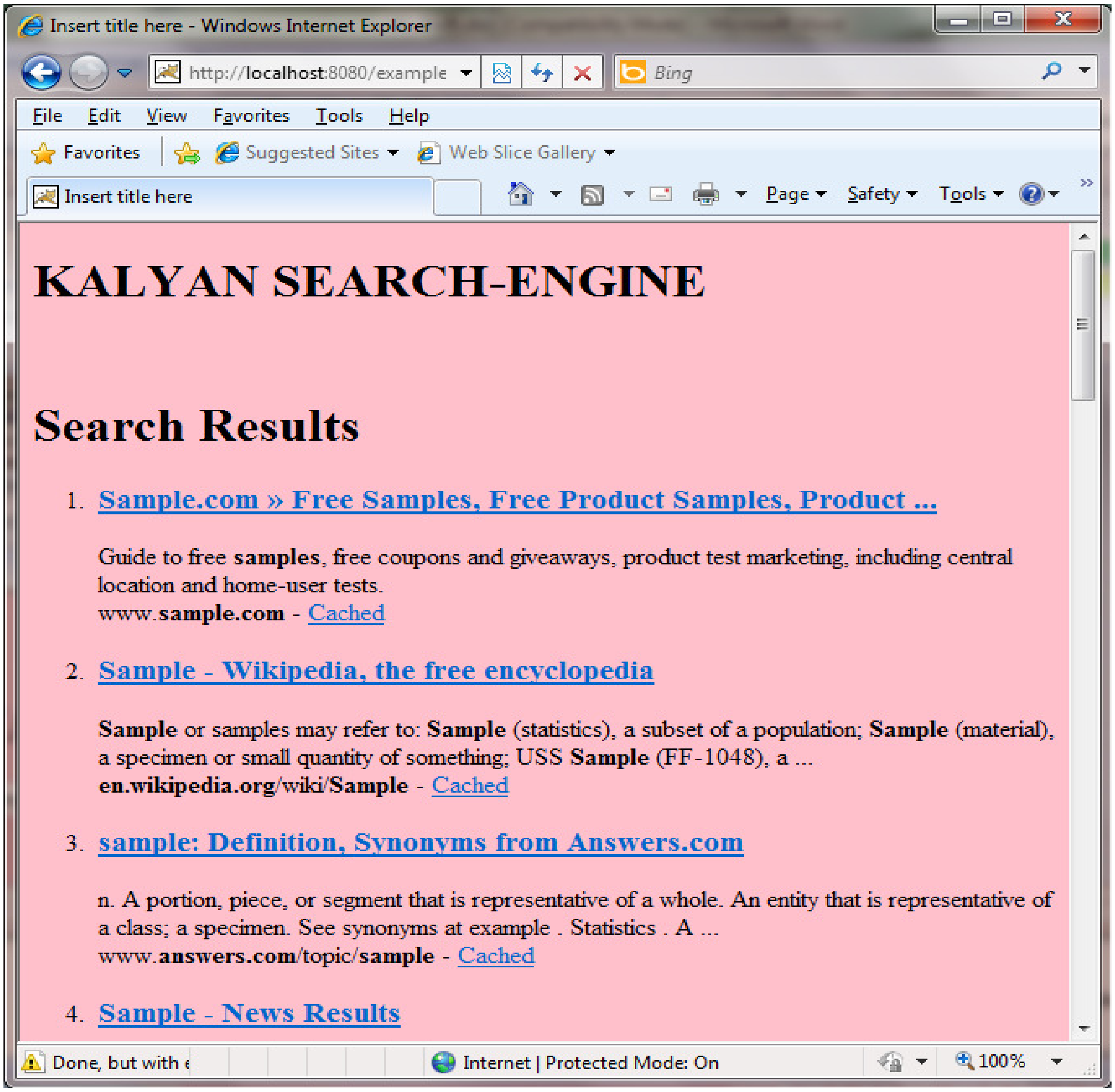


KALYAN SEARCH ENGINE

sample

Submit Query

Fig 3.2 User tries to search from a fake engine



KALYAN SEARCH-ENGINE

Search Results

1. [Sample.com » Free Samples, Free Product Samples, Product ...]

   Guide to free **samples**, free coupons and giveaways, product test marketing, including central location and home-user tests.
   www.**sample.com** - Cached

2. Sample - Wikipedia, the free encyclopedia

   Sample or samples may refer to: **Sample** (statistics), a subset of a population; **Sample** (material), a specimen or small quantity of something; USS **Sample** (FF-1048), a ...
   en.wikipedia.org/wiki/**Sample** - Cached

3. sample: Definition, Synonyms from Answers.com

   n. A portion, piece, or segment that is representative of a whole. An entity that is representative of a class; a specimen. See synonyms at example . Statistics . A ...
   www.**answers.com**/topic/**sample** - Cached

4. Sample - News Results

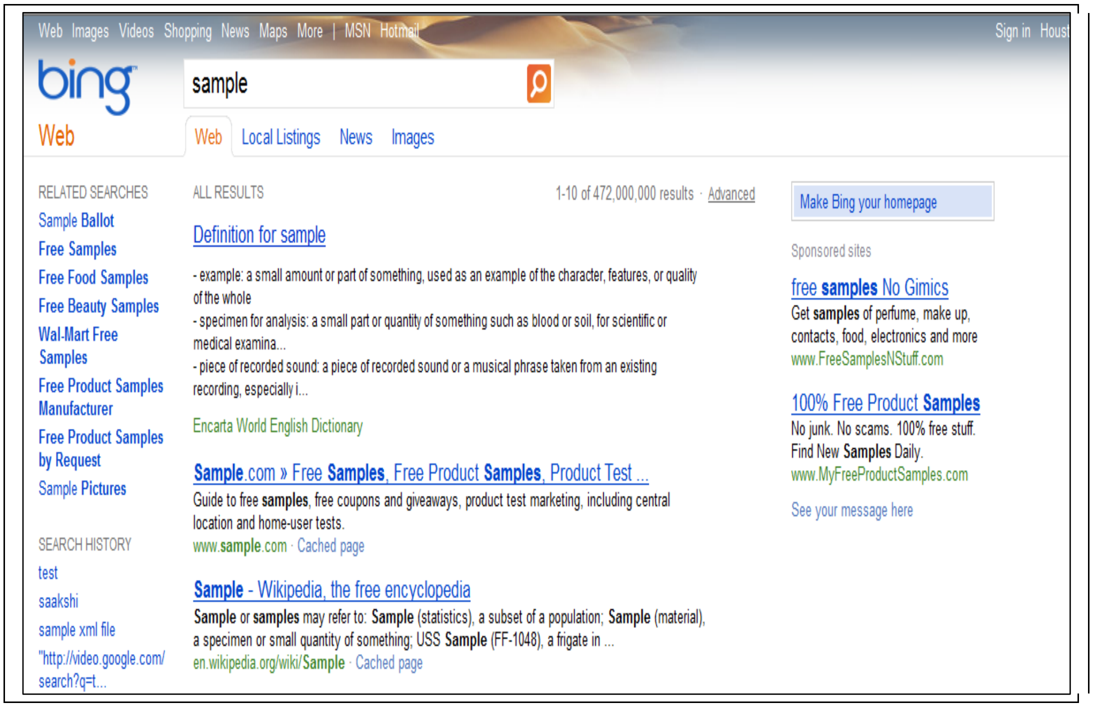
Fig 3.3 Fake engine internally uses Yahoo search engine and display results with fake branding

Compare the search results displayed in fig 3.1and 3.3. It is observed that same results are shown in the both cases because fake engine internally uses Yahoo search engine. From the above figures it is proved that Yahoo search engine is open to any HTML page and any user can hack the results of Yahoo and can show the Yahoo results with fake branding.
Hence this paper is concluding that:

“Yahoo failed to protect its Search Engine from Openness”.
Yahoo search engine suffers from the Openness of the Search Engine. This paper strongly appeals to Yahoo to fix this issue, else any user can hack Yahoo’s results. I.e. indirectly hacker owns a search engine which is Yahoo in real.

4. CASE STUDY-II: OWN THE ENGINE "BING.COM"

In this section proposed technique “Own a Search Engine” is applied to hack the search results of Bing.com. First observe the result of Bing search engine when searched for some test string “sample”.



Definition for sample

- example: a small amount or part of something, used as an example of the character, features, or quality of the whole
- specimen for analysis: a small part or quantity of something such as blood or soil, for scientific or medical examina...
- piece of recorded sound: a piece of recorded sound or a musical phrase taken from an existing recording, especially i...

Encarta World English Dictionary

Sample.com » Free **Samples**, Free Product **Samples**, Product Test ...

Guide to free **samples**, free coupons and giveaways, product test marketing, including central location and home-user tests.

www.**sample**.com · Cached page

Sample - Wikipedia, the free encyclopedia

Sample or **samples** may refer to: **Sample** (statistics), a subset of a population; **Sample** (material), a specimen or small quantity of something; USS **Sample** (FF-1048), a frigate in ...

en.wikipedia.org/wiki/**Sample** · Cached page

free **samples** No Gimics

Get **samples** of perfume, make up, contacts, food, electronics and more

www.FreeSamplesNStuff.com

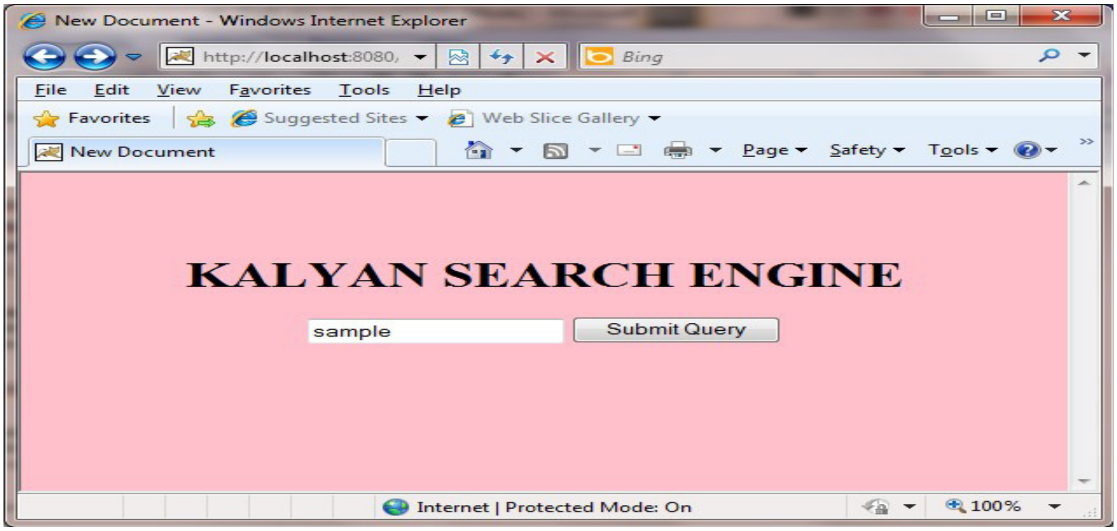
100% Free Product **Samples**

No junk. No scams. 100% free stuff. Find New **Samples** Daily.

www.MyFreeProductSamples.com

Fig 4.1 Original results of Bing search engine

Now apply the “Own a Search Engine” technique, when user access a fake search engine and tried to search for a string “sample”, the fake engine internally redirects the search request to the original Bing engine and retrieve the results and modify the branding of results to display own branding. Finally user will see the results of Bing search engine with the fake branding.

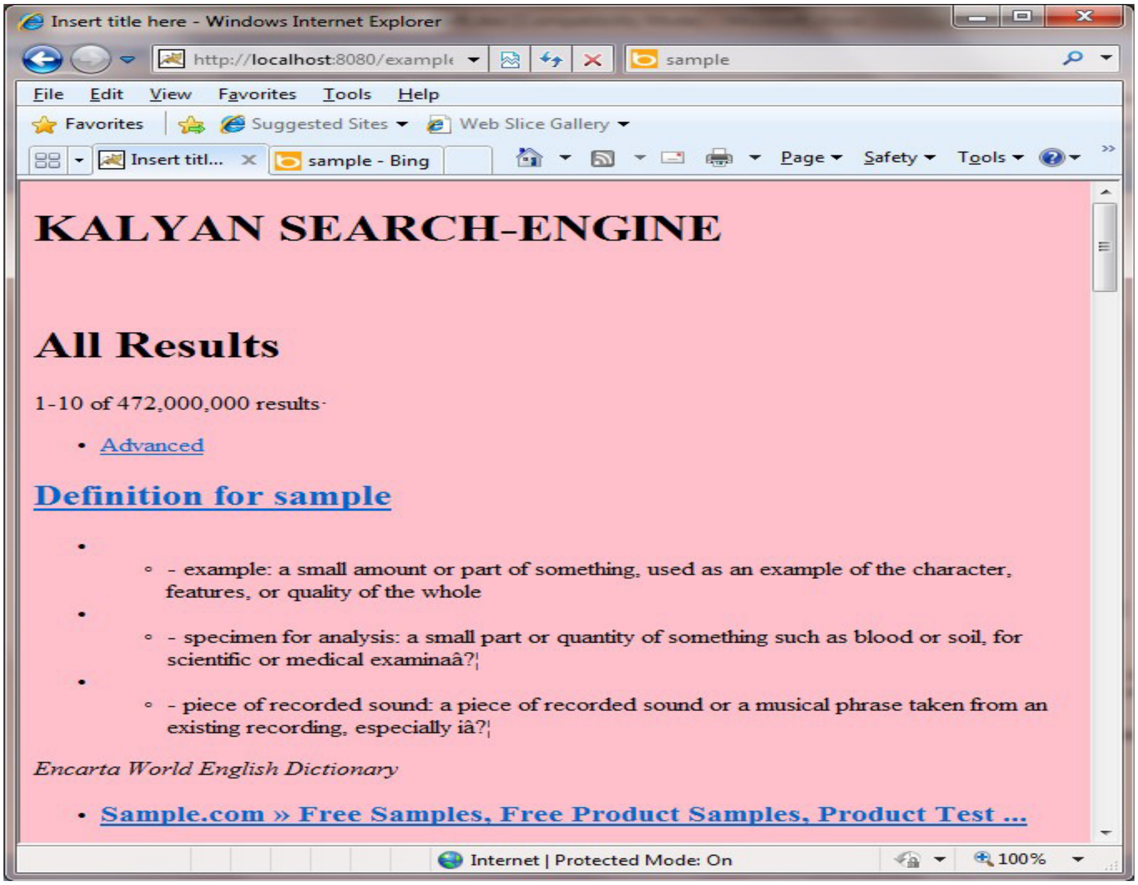

New Document - Windows Internet Explorer
http://localhost:8080/
Bing
File Edit View Favorites Tools Help
Favorites
Suggested Sites
Web Slice Gallery
New Document
Page Safety Tools
Internet | Protected Mode: On
100%

KALYAN SEARCH ENGINE

sample

Submit Query

Fig 4.2 User tries to search from a fake engine


Insert title here - Windows Internet Explorer
http://localhost:8080/example
sample
File Edit View Favorites Tools Help
Favorites
Suggested Sites
Web Slice Gallery
Insert titl...
sample - Bing
Page Safety Tools
Internet | Protected Mode: On
100%

KALYAN SEARCH-ENGINE

All Results

1-10 of 472,000,000 results·

- Advanced

Definition for sample

-
    - - example: a small amount or part of something, used as an example of the character, features, or quality of the whole
-
    - - specimen for analysis: a small part or quantity of something such as blood or soil, for scientific or medical examinaâ?¦
-
    - - piece of recorded sound: a piece of recorded sound or a musical phrase taken from an existing recording, especially iâ?¦

Encarta World English Dictionary

- Sample.com » Free Samples, Free Product Samples, Product Test ...

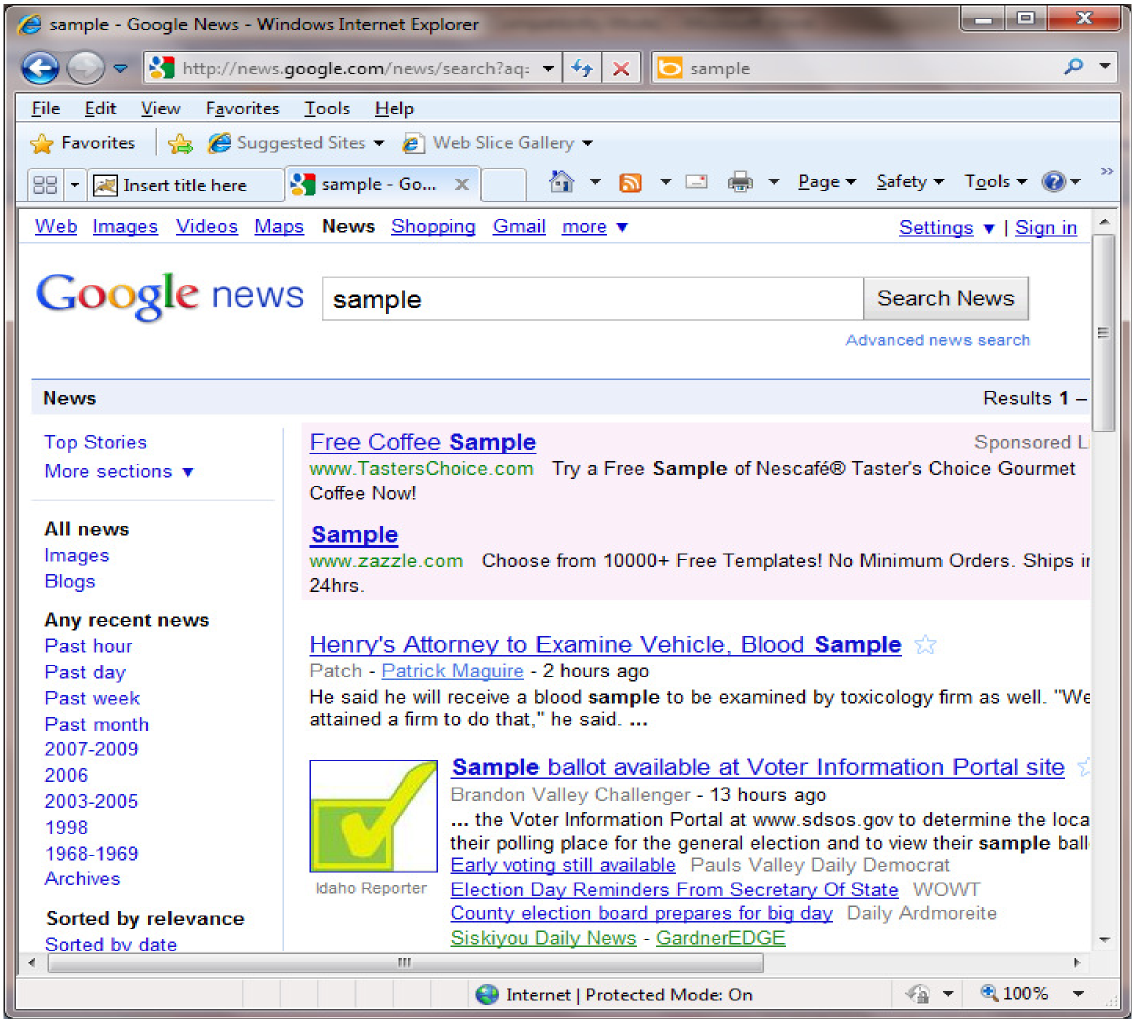
Fig 4.3 Fake engine internally uses Bing search engine and displays the results with fake branding.

Compare the search results displayed in fig 4.1and 4.3. It is observed that same results are shown in the both cases because fake engine internally uses Bing search engine.

From the above figures it is proved that Bing search engine is open to any HTML page and any user can hack the results of Bing and can show the Bing results with fake branding.
Hence this paper is concluding that:
“Bing failed to protect its Search Engine”.

5. CASE STUDY-III: OWN THE ENGINE "NEWS.GOOGLE.COM"

Though Google is strongly protecting its search engine, it had some loop holes in the past. Google was failed to protect its other search applications like “News.Google.com”.
First observe the result of news.google.com when searched for some test string “sample”.



Free Coffee Sample Sponsored L
www.TastersChoice.com Try a Free Sample of Nescafé® Taster's Choice Gourmet Coffee Now!

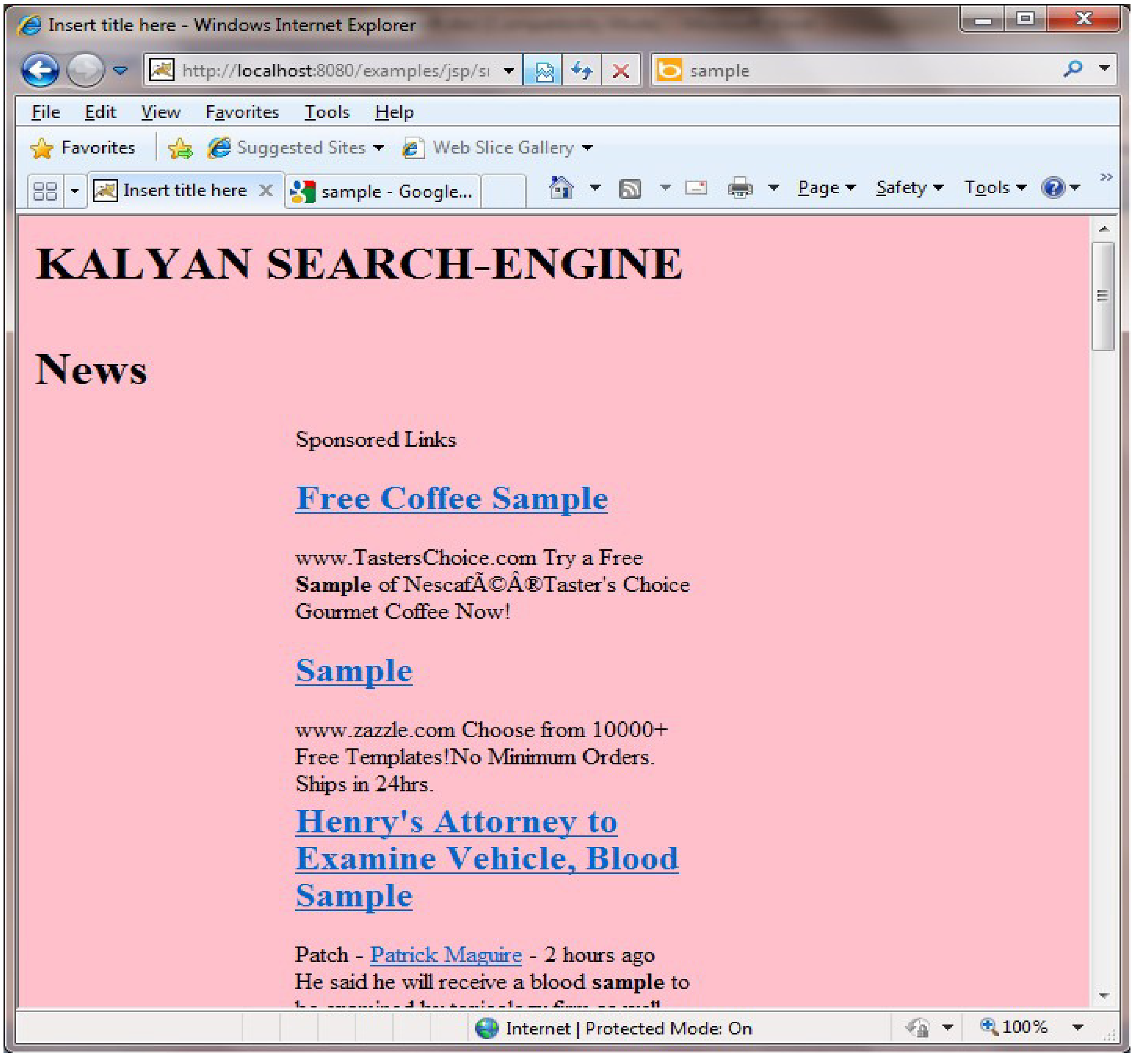
Sample
www.zazzle.com Choose from 10000+ Free Templates! No Minimum Orders. Ships ir 24hrs.

Henry's Attorney to Examine Vehicle, Blood Sample
Patch - Patrick Maguire - 2 hours ago
He said he will receive a blood sample to be examined by toxicology firm as well. "We attained a firm to do that," he said. ...

Sample ballot available at Voter Information Portal site
Brandon Valley Challenger - 13 hours ago
... the Voter Information Portal at www.sdsos.gov to determine the loca their polling place for the general election and to view their sample ball
Early voting still available Pauls Valley Daily Democrat
Election Day Reminders From Secretary Of State WOWT
County election board prepares for big day Daily Ardmoreite
Siskiyou Daily News - GardnerEDGE

Fig 5.1 Original results of Google news search engine

Apply the “Own a Search Engine” technique, when user access a fake search engine and tried to search for a string “sample”, the fake engine internally redirects the search request to the original news.google.com engine and retrieve the results and then modifies the branding of results to display own branding. Finally user will see the results of nesws.google.com search engine with the fake branding. Following figure explains the hacking of news.google.com search results.


Insert title here - Windows Internet Explorer
http://localhost:8080/examples/jsp/s
sample
File Edit View Favorites Tools Help
Favorites
Suggested Sites
Web Slice Gallery
Insert title here
sample - Google...
Page
Safety
Tools
Internet | Protected Mode: On
100%

KALYAN SEARCH-ENGINE

News

Sponsored Links

Free Coffee Sample

www.TastersChoice.com Try a Free **Sample** of NescafÃ©Â®Taster's Choice Gourmet Coffee Now!

Sample

www.zazzle.com Choose from 10000+ Free Templates!No Minimum Orders. Ships in 24hrs.

Henry's Attorney to Examine Vehicle, Blood Sample

Patch - Patrick Maguire - 2 hours ago
He said he will receive a blood **sample** to

Fig 5.2 Fake engine internally uses Google-News search engine and displays the results with fake branding

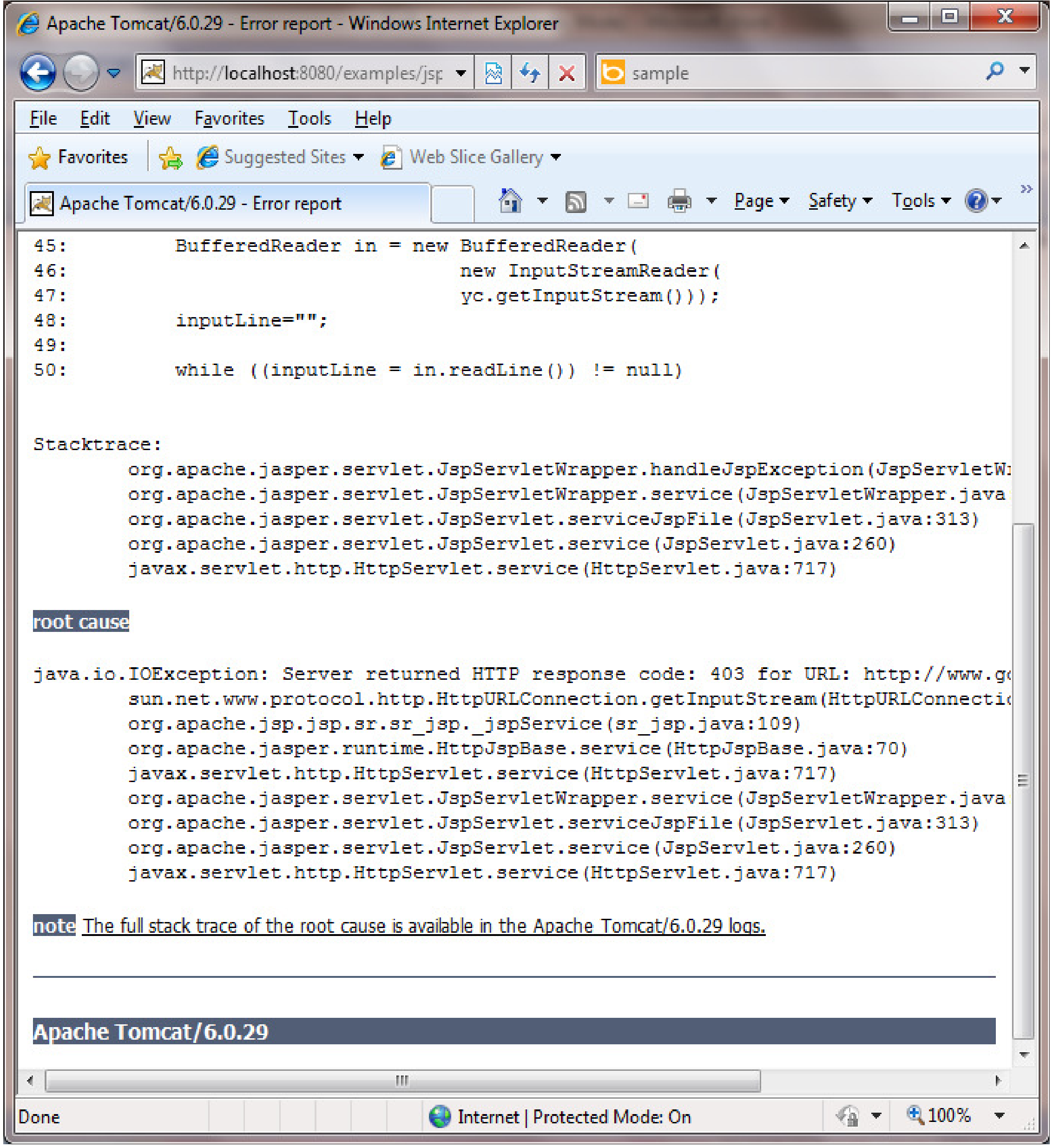
Compare the search results displayed in fig 5.1and 5.2. It is observed that same results are shown in the both cases because fake engine internally uses Google-News search engine. From the above figures it is proved that Google-News search engine was open to any HTML page and any user can hack the results of Google-News and can show the news.google.com results with fake branding. Hence this paper is concluding that:

"Google-News was failed to protect its Search Engine",
But currently Google has strong protection for its search engine.

6. CASE STUDY-IV: OWN THE ENGINE "SEARCH.GOOGLE.COM"

In this section proposed technique "Own a Search Engine" is applied to hack the search results of search.google.com but Google has strong protection to save their search engine from this kind of attack. When we tried to apply the "Own a Search Engine" technique on seacrh.Google.com, an exception is thrown to the fake search engine with the HTTP 403 error code.

Following figure shows how the Google is protecting its search engine.



Apache Tomcat/6.0.29 - Error report - Windows Internet Explorer
http://localhost:8080/examples/jsp
sample
File Edit View Favorites Tools Help
Favorites
Suggested Sites
Web Slice Gallery
Apache Tomcat/6.0.29 - Error report
Page
Safety
Tools

```
45:         BufferedReader in = new BufferedReader(
46:                                 new InputStreamReader(
47:                                 yc.getInputStream()));
48:         inputLine="";
49:
50:         while ((inputLine = in.readLine()) != null)


Stacktrace:
        org.apache.jasper.servlet.JspServletWrapper.handleJspException(JspServletW
        org.apache.jasper.servlet.JspServletWrapper.service(JspServletWrapper.java
        org.apache.jasper.servlet.JspServlet.serviceJspFile(JspServlet.java:313)
        org.apache.jasper.servlet.JspServlet.service(JspServlet.java:260)
        javax.servlet.http.HttpServlet.service(HttpServlet.java:717)
```

root cause

```
java.io.IOException: Server returned HTTP response code: 403 for URL: http://www.go
        sun.net.www.protocol.http.HttpURLConnection.getInputStream(HttpURLConnecti
        org.apache.jsp.jsp.sr.sr_jsp._jspService(sr_jsp.java:109)
        org.apache.jasper.runtime.HttpJspBase.service(HttpJspBase.java:70)
        javax.servlet.http.HttpServlet.service(HttpServlet.java:717)
        org.apache.jasper.servlet.JspServletWrapper.service(JspServletWrapper.java
        org.apache.jasper.servlet.JspServlet.serviceJspFile(JspServlet.java:313)
        org.apache.jasper.servlet.JspServlet.service(JspServlet.java:260)
        javax.servlet.http.HttpServlet.service(HttpServlet.java:717)
```

note The full stack trace of the root cause is available in the Apache Tomcat/6.0.29 logs.

Apache Tomcat/6.0.29

Fig 6.1 shows the exception thrown from the Google Search Engine to the fake engine

7. CONCLUSION

This paper conducts a test to the search engines Google, Yahoo and Bing with a proposed technique and proves the openness of the top 3 search engines. The proposed test technique ‘Own a search engine’ proves that Google strongly protects its search engine where Bing and Yahoo failed to protect their search engines from “Openness” flaw.

But Google also had loop holes in the past; applications like ‘news.google.com’ were suffered from Openness problem. As a conclusion, this paper proves a critical loop hole of top search engines called “Openness of a Search Engine” with a proposed technique. And this paper strongly appeals to Bing and Yahoo to fix the critical flaws of their Search Engines; else anyone can own the search results of their search engines.

8. REFERENCES

[1] “Ad-Kick attack on Google home page”, poster presented in the conference ICISS 2009, in Kolkata, India, by “Katuru SM Kalyana Chakravarthy”.

[2] “Internet & World Wide Web How to Program” book by Deitel & Deitel.

[3] “New Attack Ad-Kick Attack on GOOGLE & Failure of MICROSOFT Security Essentials”, paper by “Katuru SM Kalyana Chakravarthy”.

[4] “MICROSOFT Security Essentials Fails to protect the Hosts File”, paper by “Katuru SM Kalyan Chakravarthy”.

[5] “New attack: ‘Ad-Kick attack’ on a website using hosts file hijacking”, paper by “Katuru SM Kalyan Chakravarthy”.

Authors

Name: Katuru SM Kalyana Chakravarthy

Short Biography:

“Katuru SM Kalyana Chakravarthy” is a Software Test Engineer, having 5+ years of experience in Software Testing and Quality Assurance Services. Currently he is working as Automation Engineer in Oracle India Private Ltd, Hyderabad, India and he has worked in AppLabs, India for five Years including one year of on-site experience. He is much interested in Research & Development, submitted white papers to the international conferences and presented a Poster in the conference ICISS 2009, Kolkata, India.