Internet Censorship Detection and Data Recovery

mgr inż. Paweł Koszut
Internet Censorship Detection and Data Recovery

- Introduction and example of censored website
- Practical problems with censorship detection
- Visualisation of website changes
- Other applications
Example of censored website
The page you've requested is an excerpt from a book by Brent Scowcroft and George H. W. Bush titled *A World Transformed*, which appeared in the March 2, 1998, issue of *TIME* magazine under the title "Why We Didn't Remove Saddam*. It has been removed from our site because the publisher did not grant us rights to sell the piece online through the TIME archive.
Why We Didn't Remove Saddam

By GEORGE BUSH AND BREN'T SCOWCROFT

The end of effective Iraqi resistance came with a rapidity which surprised us all, and we were perhaps psychologically unprepared for the sudden transition from fighting to peacemaking. True to the guidelines we had established, when we had achieved our strategic objectives (ejecting Iraqi forces from Kuwait and eroding Saddam's threat to the region) we stopped the fighting. But the necessary limitations placed on our objectives, the fog of war, and the lack of "battleship Missouri" surrender unfortunately left unresolved problems, and new ones arose.

We were disappointed that Saddam's defeat did not break his hold on power, as many of our Arab allies had predicted and we had come to expect. President Bush repeatedly declared that the fate of Saddam Hussein was up to the Iraqi people. Occasionally, he indicated that removal of Saddam would be welcome, but for very practical reasons there was never a promise to aid an uprising. While we hoped that popular revolt or coup would topple Saddam, neither the U.S. nor the countries of the region wished to see the breakup of the Iraqi state. We were concerned about the long-term balance of power at the head of the Gulf. Trying to eliminate Saddam, extending the ground war into an
Practical problems
Internet Censorship Detection and Data Recovery

Internet Explorer Multiple Code Execution Vulnerabilities

- **Security Professionals**
- **Security Vendors**
- **Free Solutions For**
  - Open Communities
  - Journalists & Media
- **Software Inspectors**
  - Scan Online
  - Personal (IPS)
  - Network (NSI)
- **Secunia Advisories**
  - Search
  - Historic Advisories
  - Listed By Product
  - Listed By Vendor
  - Statistics / Graphs
  - Secunia Research
  - Report Vulnerability
  - About Advisories
- **Virus Information**
  - Chronological List
  - Last 10 Virus Alerts
  - About Virus Information

**Data Loss Prevention**
Learn more about Proofpoint’s data loss prevention & email security
www.proofpoint.com

**Description**
Some vulnerabilities have been reported in Internet Explorer, which can be exploited by malicious people to compromise a user’s system.

1) A use-after-free error in mshtml.dll when handling "setExpression()" method calls can be...
ETag
## FileETag Directive

### Description:
File attributes used to create the ETag HTTP response header

### Syntax:
`FileETag component ...`

### Default:
`FileETag INode MTime Size`

### Context:
server config, virtual host, directory, .htaccess

### Override:
FileInfo

### Status:
Core

### Module:
core

The `FileETag` directive configures the file attributes that are used to create the ETag (entity tag) response header field when the document is based on a file. (The ETag value is used in cache management to save network bandwidth.) In Apache 1.3.22 and earlier, the ETag value was *always* formed from the file's inode, size, and last-modified time (mtime). The `FileETag` directive allows you to choose which of these -- if any -- should be used. The recognized keywords are:

- **INode**
  The file's i-node number will be included in the calculation

- **MTime**
  The date and time the file was last modified will be included

- **Size**
  The number of bytes in the file will be included

- **All**
  All available fields will be used. This is equivalent to:

  ```
  FileETag INode MTime Size
  ```

- **None**
  If a document is file-based, no ETag field will be included in the response
Internet Censorship Detection and Data Recovery
Internet Censorship Detection and Data Recovery
**Internet Censorship Detection and Data Recovery**

---

**Document Information**

**General**

- **Title:** China's President Hu concludes inspection tour in Anhui_English_Xinhua
- **URL:** [http://news.xinhuanet.com/english/2008-01/14/content_7422102.htm](http://news.xinhuanet.com/english/2008-01/14/content_7422102.htm)
- **Last modified:** Mon, 14 Jan 2008 19:33:43 GMT
- **Document encoding:** gb2312

**HTTP Headers**

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accept-Ranges</td>
<td>bytes</td>
</tr>
<tr>
<td>Connection</td>
<td>keep-alive</td>
</tr>
<tr>
<td>Content-Encoding</td>
<td>gzip</td>
</tr>
<tr>
<td>Content-Length</td>
<td>4428</td>
</tr>
<tr>
<td>Content-Type</td>
<td>text/html; charset=GB2312</td>
</tr>
<tr>
<td>Date</td>
<td>Mon, 14 Jan 2008 20:12:49 GMT</td>
</tr>
<tr>
<td>ETag</td>
<td>“4748a4e-4a3b-c0ff33c0”</td>
</tr>
<tr>
<td>Last-Modified</td>
<td>Mon, 14 Jan 2008 19:33:43 GMT</td>
</tr>
<tr>
<td>Server</td>
<td>Apache</td>
</tr>
<tr>
<td>Vary</td>
<td>Accept-Encoding</td>
</tr>
</tbody>
</table>
ETagFix V1.00

ETagFix alters the IIS etag that is generated for each page so that it does not change each time your server is restarted.

Etags are a method of informing web browsers and crawlers when a web page has changed. Using etags, your web browser or crawler will determine when a cached page may be used and when the page must be downloaded to present the current information.

Sounds simple and useful, right? Well, yes and no. True, it’s important for browsers to download a page when it really has changed. However, it’s not necessary to take the time and bandwidth required to download a page when the etag has been updated due to restarting the web server, but the page information hasn’t changed.

*ETagFix stabilizes the etag that Internet Information Services (IIS) generates so that it doesn’t change each time the web server is restarted.*
Possible solution
Internet Censorship Detection and Data Recovery
Internet Censorship Detection and Data Recovery
Internet Censorship Detection and Data Recovery

![Graph showing lines over time](image)
Internet Censorship Detection and Data Recovery

![Graph showing lines over time](image)
Internet Censorship Detection and Data Recovery
Iraqi oil production is now running at 2.3 million barrels per day, compared with 1.9 million barrels at the start of this year.

It puts the rise down to the improving security situation in Iraq, especially in the north of the country.

But the IEA warned that attacks on Iraqi oil facilities remain a threat.

In southern Iraq, more than 85% of the residents of Basra believe British troops have had a negative effect on the Iraqi province since 2003, according to a BBC poll.

The survey for BBC Newsnight of nearly 1,000 people also suggests that 56% believe their presence has increased the overall level of militia violence.

**Sabotage attacks**

In its latest monthly Oil Market Report, the IEA puts the Iraqi oil sector's investment in maintaining its infrastructure at $15 billion. It says Da'wa, a group that has secured permission to give significant aid to the Iraqi oil sector, and has a $100 million credit line from the World Bank.

Iraq's oil infrastructure appears to be getting back on track.
Internet Censorship Detection and Data Recovery

BBC NEWS | Business | International version

Last Updated: Friday, 14 December 2007, 17:31 GMT

News services
Your news when you want it

THE STRUGGLE FOR IRAQ
KEY STORIES
Zawahiri says UK 'fleeing Basra'
UK troops return Basra to Iraqis
US marine guilty of Iraq killing
US Senate passes Iraq funds bill
Basra residents blame UK troops
Triple car bombs hit south Iraq

BACKGROUND AND ANALYSIS
Basra uncertainty
The future for the city as British troops hand over control

Basra's new era
Heading back home to Iraq
Sunni 'neighbourhood watch'
Political delay on death row
Is Iraq getting better?
Iraq violence in figures

ANDREW NORTH DIARY
Surface security

Iraq's oil infrastructure appears to be getting back on track.

The [XXX] xxx xxxxxxxx
xxxxxxx xxxxxxxx xxxxxxxx
xxx

BBC poll: UK troops

Last Updated: Friday, 14 December 2007, 17:31 GMT

E-mail this to a friend
Printable version

Africa
Americas
Asia-Pacific
Europe
Middle East
South Asia
UK
Business
Market Data
Economy
Companies
Health
Science/Nature
Technology
Entertainment
Also in the news
---------
Video and Audio
---------
Have Your Say
In Pictures
Country Profiles
Special Reports
RELATEB BBC SITES
Internet Censorship Detection and Data Recovery
http://news.bbc.co.uk/2/hi/technology/7149588.stm

step: 10 minutes
Algorithm

The operation of diff is based on solving the Longest Common Subsequence (LCS) problem. In the LCS problem, you have two sequences of items:

```plaintext
  a b c d f g h j q z
```

```plaintext
  a b c d e f g i j k r x y z
```

and you want to find the longest sequence of items that is present in both original sequences in the same order. That is, you want to find a new sequence which can be obtained from the first sequence by deleting some items, and from the second sequence by deleting other items. You also want this sequence to be as long as possible. In this case it is

```plaintext
  a b c d f g j z
```

From the longest common subsequence it's only a small step to get diff-like output:

```plaintext
  e h i k q r x y
+ - + + - +++
```
Internet Censorship Detection and Data Recovery

http://news.bbc.co.uk/2/hi/technology/7149588.stm

step: 10 minutes
Internet Censorship Detection and Data Recovery
Internet Censorship Detection and Data Recovery
Internet Censorship Detection and Data Recovery
More applications
French Embassy site for Libya said to be serving malware

The French Embassy website for Libya has been compromised and is serving up malware to visitors, according to McAfee.

By Ellen Messmer Framingham | Monday, 17 December, 2007

McAfee researcher Francois Paget discovered this on Thursday and the company says it has reported its findings to the French government. The site has been attacked using an iFrame exploit that inserts an invisible frame in the page in order to re-direct some web browser connections to another location, which serves up a "downloader," code that attempts to reside on the victim machine. If the downloader is successful, the attacker can then remotely attempt to download other malware, "typically a bot or a password-stealing Trojan," says Dave Marcus, McAfee security researcher and communications manager.

Marcus says Paget, a researcher with tools to scan scripts and investigate code behaviour, happened by chance to be looking at the French Embassy website for Libya and discovered the attack code on it. The incident is similar to discoveries made by security researchers of other compromised websites spewing attack code, including that of the Bank of India and the MySpace page of Alicia Keys.
Internet Censorship Detection and Data Recovery
Internet Censorship Detection and Data Recovery
Internet Censorship Detection and Data Recovery

Thank you