

# Internet Gambling Investigations

Architecture of the World Wide Web



---

---

---

---

---

---

---

---

## Objectives

During this session we will discuss:

- The term 'world wide web'
- User interaction on the world wide web
- The purpose of gateways
- The purpose of a proxy server
- The structure and syntax of a URI / URN / URL
- The purpose and types of obfuscation
- Common URI/URL obfuscation techniques

---

---

---

---

---

---

---

---

## Structure of the World Wide Web

*What is the World Wide Web (WWW)?*

"An information space in which the items of interest, referred to as resources, are identified by global identifiers called Uniform Resource Identifiers (URI)."

The Architecture of the World Wide Web, Vol. 1  
W3C Recommendation 15 December 2004

---

---

---

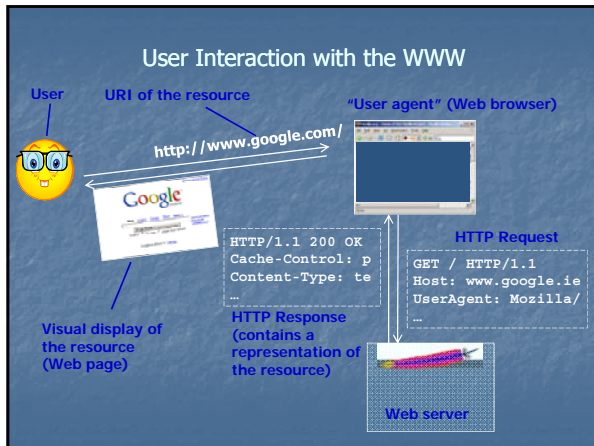
---

---

---

---

---



---

---

---

---

---

---

---

---

### Structure of the World Wide Web

*Where do servers get the page content?*

- Read from file(s) on disk
- Generated on-the-fly using another program
- Forwarded via HTTP request from another server

---

---

---

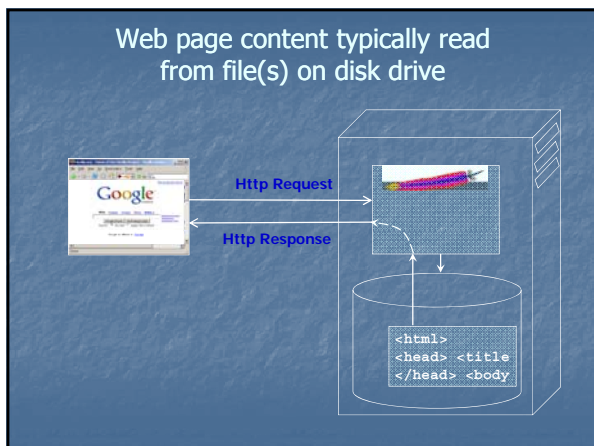
---

---

---

---

---



---

---

---

---

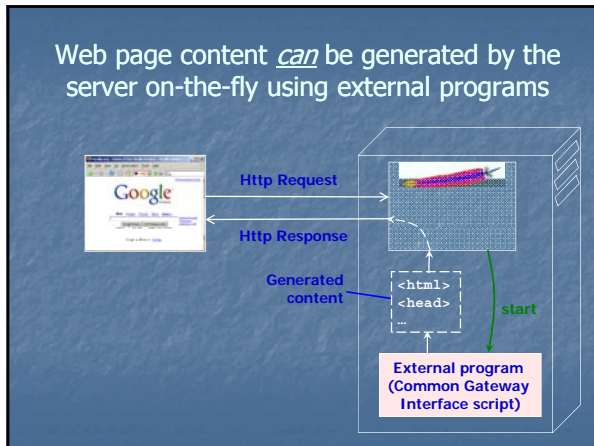
---

---

---

---

Web page content *can* be generated by the server on-the-fly using external programs



---

---

---

---

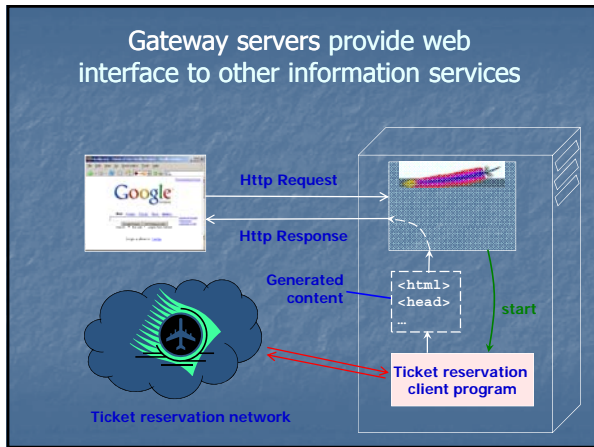
---

---

---

---

Gateway servers provide web interface to other information services



---

---

---

---

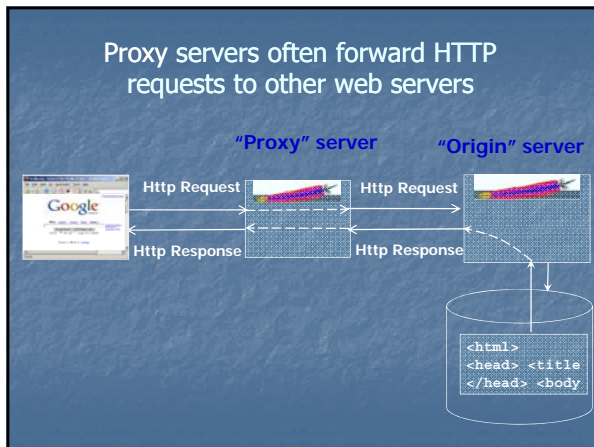
---

---

---

---

Proxy servers often forward HTTP requests to other web servers



---

---

---

---

---

---

---

---

Structure of the World Wide Web

*Why use proxies?*

- Most proxy servers exist to reduce Internet traffic and user response time through "caching"
- Caching proxies retain copies of frequently requested web resources
- Before forwarding the HTTP request, the caching proxy checks if a valid copy of the requested resource is available locally and - if it is available - returns that copy to the user.

---

---

---

---

---

---

---

---

Structure of the World Wide Web

Other uses of proxy servers:

- Proxy servers can be used to control access between an Intranet and the Internet.
- Proxy servers can be used to conceal the identity of web users from the origin web server.
  - "Anonymizing" proxies clean HTTP requests of information that may identify the end user.

---

---

---

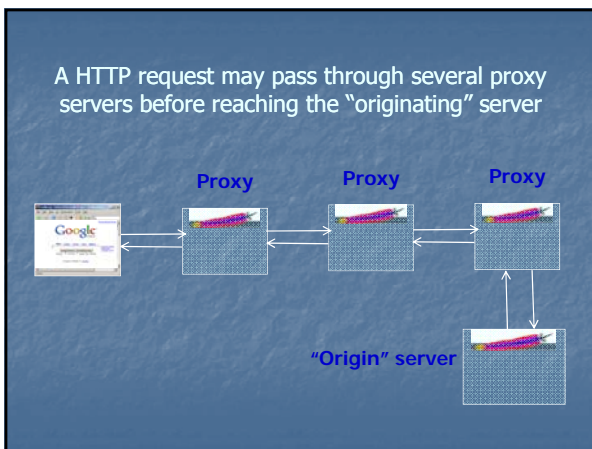
---

---

---

---

---



---

---

---

---

---

---

---

---

Structure of the World Wide Web

Logging

- In most instances the web server creates a log file of requests it receives, along with how it responded to those requests.
- The log files can provide invaluable investigative data.
- Proxies can also maintain similar logs.
- The format of the log files usually depends on the type of web server.

---

---

---

---

---

---

---

---

Structure of the World Wide Web

Encryption/SSL

- Communication between the client and server can be encrypted using SSL (Secure Socket Layer)
- This allows sensitive information, such as banking and credit card details to be transferred securely across the internet.

---

---

---

---

---

---

---

---

Structure of the World Wide Web

Hyper-Links

- Web pages often contain "references" to other pages
- These references are known as "hyperlinks" or "links"
- The web browser renders links in an identifiable way
- Users utilize the links to travel between pages

---

---

---

---

---

---

---

---

Structure of the World Wide Web

Web 2.0

- A phrase used to refer to a perceived "second generation" of internet based services
- Specifically pertains to collaborative/sharing websites, such as
  - Social networking sites
  - Wikis

---

---

---

---

---

---

---

---

Structure of the World Wide Web

Client Side Content

- When a web server delivers a page to a client, the page can contain code which is executed by the web browser,
  - javascript
- This allows the web page to provide dynamic content to the user without needing to be in regular communication with the web server.
- This is a different thing to a script or external program which is run on the web server itself.

---

---

---

---

---

---

---

---

Structure of the World Wide Web

Uniform Resource Identifier (URI)

---

---

---

---

---

---

---

---

Uniform Resource Identifier

### Uniform Resource Identifier

*Definition:*

- A "URI" is a string of characters used to identify or name a resource on the Internet.
- This identification enables interaction with representations of the resource over a network (typically the World Wide Web) using specific protocols.

---

---

---

---

---

---

---

---

Uniform Resource Identifier

### Uniform Resource Identifier

- A general purpose method for referring to many types of TCP/IP resources
- Generally they are divided into two primary categories, based on how they describe a resource:
  - Uniform Resource Locators (URL)
  - Uniform Resource Names (URN)

---

---

---

---

---

---


---

---

Uniform Resource Identifier

Uniform Resource Locator

- The URL refers to a resource through the combination of a protocol and a specific resource location.
- A URL begins with the name of the protocol being used for accessing the resource and then contains sufficient information to how it can be obtained.



The diagram shows a light blue rounded rectangle labeled 'URI' at the top. Below it, a horizontal line divides the space. Underneath, there are two rounded rectangles side-by-side: a yellow one labeled 'URL' on the left and a purple one labeled 'URN' on the right. A vertical dashed line separates the two boxes.

---

---

---

---

---

---

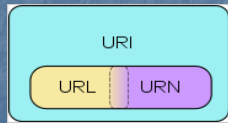
---

---

### Uniform Resource Identifier

#### Uniform Resource Name

- The URN provides a way of uniquely naming a resource without specifying an access protocol or specific location.



---

---

---

---

---

---

---

---

### Uniform Resource Identifier

#### *In other words....*

The URN defines an item's identity, while the URL provides a method for finding it.



---

---

---

---

---

---

---

---



### URL Obfuscation



---

---

---

---

---

---

---

---



### URL Obfuscation

**ob-fus-cate**

Pronunciation [ob-fuh-skeyt, ob-fuhs-keyt]

–verb (used with object), -cat-ed, -cat-ing.

1. to confuse, bewilder, or stupefy.
2. to make obscure or unclear: to obfuscate a problem with extraneous information.
3. to darken.

---

---

---

---

---

---

---

---

### Purpose of Obfuscation

- Phishing scams rely on victim’s belief they are accessing a genuine website.
- The URL of the phishing website is usually disguised to look similar to the real website.
- Let’s examine some tricks used to achieve that:
  - Username based obfuscations
  - %-encoding based obfuscations
  - Misspelled URLs
  - Homographic URLs

---

---

---

---

---

---

---

---

### Username-Obfuscated URL

Consider URL:

[http://cnn.example.com&story=breaking\\_new@10.0.0.1/top\\_story.htm](http://cnn.example.com&story=breaking_new@10.0.0.1/top_story.htm)

An uninformed user might assume that the host is ‘cnn.example.com’, while it is actually part of the username.

The actual host address (after @) is [10.0.0.1](http://10.0.0.1)

---

---

---

---

---

---

---

---

### %-encoding

Characters in URL can be specified using %-notation %xx, where xx is the hexadecimal ASCII code of the character

`http://www%2Fgoogle%2Ecom/`

`http://www.google.com/`

Note: special characters like "@" "/" "." and "?" lose their special meaning when encoded.

---

---

---

---

---

---

---

---

### Hexadecimal Codes of ASCII Characters

Second digit (6)

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2	SP	!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

First digit (7)

Example: %76      %76 = character 'v'

---

---

---

---

---

---

---

---

### %-Obfuscated URL

`http://www.paypal.com%2E%75%73%65%72%73%65%74%2E%6E%65%74:%34%39%30%33/%63/%69%6E%64%65%78%2E%68%74%6D`



`http://www.paypal.com.userset.net:4903/c/index.htm`

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2	SP	!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

---

---

---

---

---

---

---

---

### De-Obfuscating %-Encoded URL

Free Tools:

- <http://www.gooby.ca/decrypt/>
- <http://www.dnsstuff.com/tools/tools/>
- <http://wepawet.iseclab.org/>
- <http://www.id4com.com/toolset/URLObfuscate.aspx>

---

---

---

---

---

---

---

---

---

---

*What is the host name in the following URL?*

<http://www.ebay.com@%61%2E%63%6F%6D>

0	NUL	SOH	STX	ETX	EOT	ENO	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2	SP	!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

a.com

---

---

---

---

---

---

---

---

---

---

### URL Obfuscation by Misspelling

usbank.com	ussbank.net
firstusa.com	firstusaonline.biz
washingtonmutual.com	washingttonmutual.com

---

---

---

---

---

---

---

---

---

---

### Homographic\* URL Obfuscation

WWW.GOOGLE.COM      WWW.G00GLE.COM

Letter 'O'      Digit '0' (Zero)

\*homograph is one of two or more words spelled alike but different in meaning (Merriam-Webster Online Dictionary)

---

---

---

---

---

---

---

---

### Obfuscation via redirection

Uses ability of some well known web-sites to redirect web browser to a different website when given appropriate URI.

Example:

*<http://r.aol.com/cgi/udir?http://www.iacis.com/>*

---

---

---

---

---

---

---

---

### Obfuscation by Sub-Domain

*What's the actual address?*

<http://www.ebay.com.forensics.com>

Actual Address

---

---

---

---

---

---

---

---

### Obfuscation by Decimal (Integer) Conversion

**Formula 1:** Start by breaking IP address into four octets.

For example, IP 134.39.248.56

First Octet: 134  
 Second Octet: 39  
 Third Octet: 248  
 Fourth Octet: 56

To calculate the decimal address from a dotted string, perform the following calculation:

(first octet) \* 256 + (second octet) = \* 256 + (third octet) = \* 256 + (fourth octet) = Decimal Integer

$134 * 256 + 39 = * 256 + 248 = * 256 + 56 = 2250766392$

4/15/2010 International Association of Computer Investigative Specialists 37

---

---

---

---

---

---

---

---

---

---

---

---

### Obfuscation by Decimal (Integer) Conversion

**Formula 2:** Start by breaking IP address into four octets.

For example, IP 134.39.248.56

First Octet: 134  
 Second Octet: 39  
 Third Octet: 248  
 Fourth Octet: 56

To calculate the decimal address from a dotted string, perform the following calculation:

(first octet \* 256<sup>3</sup>) + (second octet \* 256<sup>2</sup>) + (third octet \* 256) + (fourth octet) = Decimal Integer

$134 * 16777216 + 39 * 65536 + 248 * 256 + 56 = 2250766392$

4/15/2010 International Association of Computer Investigative Specialists 38

---

---

---

---

---

---

---

---

---

---

---

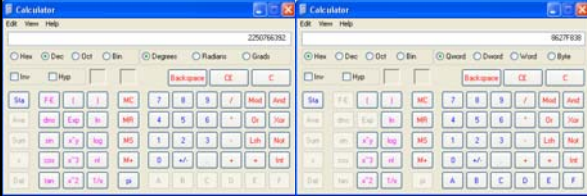
---

### Obfuscation by Decimal Conversion (Integer - Hexadecimal)

To convert an IP integer address to Hexadecimal:

For example, IP 134.39.248.56 = 2250766392 = 0x8627F838

Use Scientific calculator:




---

---

---

---

---

---

---

---

---

---

---

---

Structure of the World Wide Web

QUESTIONS?

---

---

---

---

---

---

---

---

Structure of the World Wide Web

Time for a Practical Demonstration

---

---

---

---

---

---

---

---

888casino.com

- Domain Name: 888casino.com
- IP Address: 213.52.252.59
- Decimal Integer: 3577019451
- Hexadecimal: D534FC3B

---

---

---

---

---

---

---

---

For example, IP **213.52.252.59**

First Octet: **213**  
Second Octet: **52**  
Third Octet: **252**  
Fourth Octet: **59**

To calculate the decimal address from a dotted string, perform the following calculation:

(first octet) \* 256 + (second octet) = \* 256 + (third octet) = \* 256 + (fourth octet)  
= Decimal Integer

**213** \* 256 + **52** = \* 256 + **252** = \* 256 + **59** = **3577019451**

**Decimal – 3577019451**    <http://3577019451>  
**Hexadecimal – D534FC3B**    <http://0xD534FC3B>

---

---

---

---

---

---

---

---