Inserção pelos operadores de rede de conteúdo falso em websites selecionados

ADAPTADO DE WEBSITE-TARGETED FALSE CONTENT INJECTION BY NETWORK OPERATORS

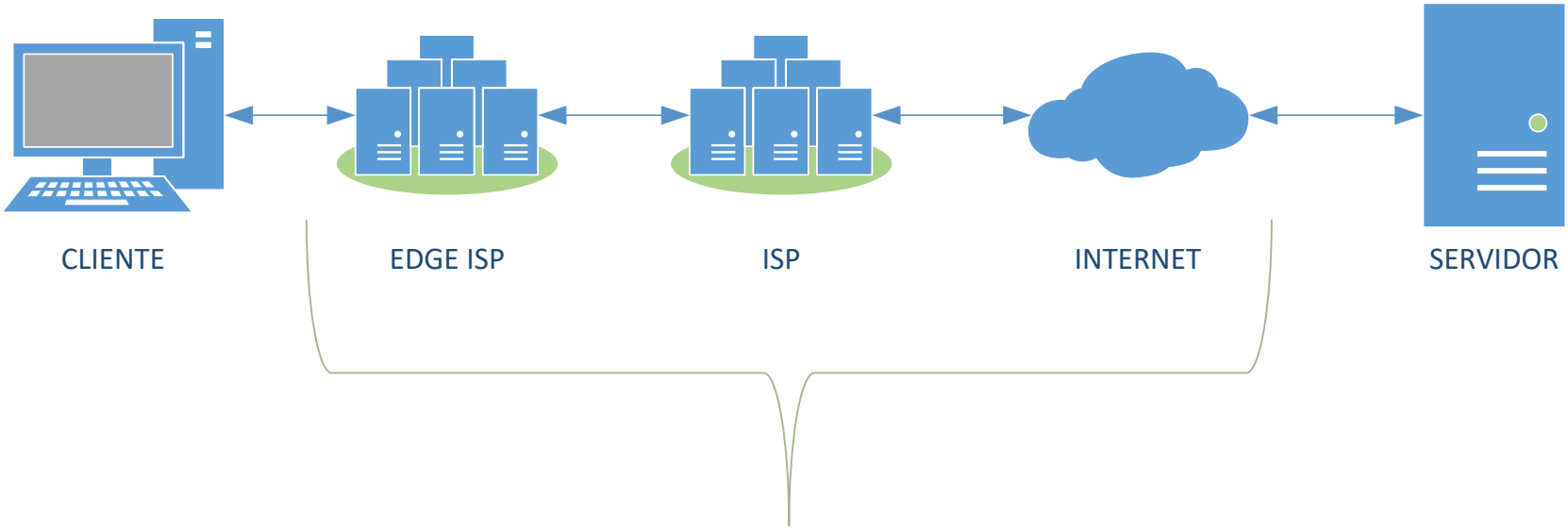
GABI NAKIBLY , JAIME SCHCOLNIK E YOSSI RUBIN - ISRAEL

LUIZ FERNANDO PUTTOW SOUTHIER

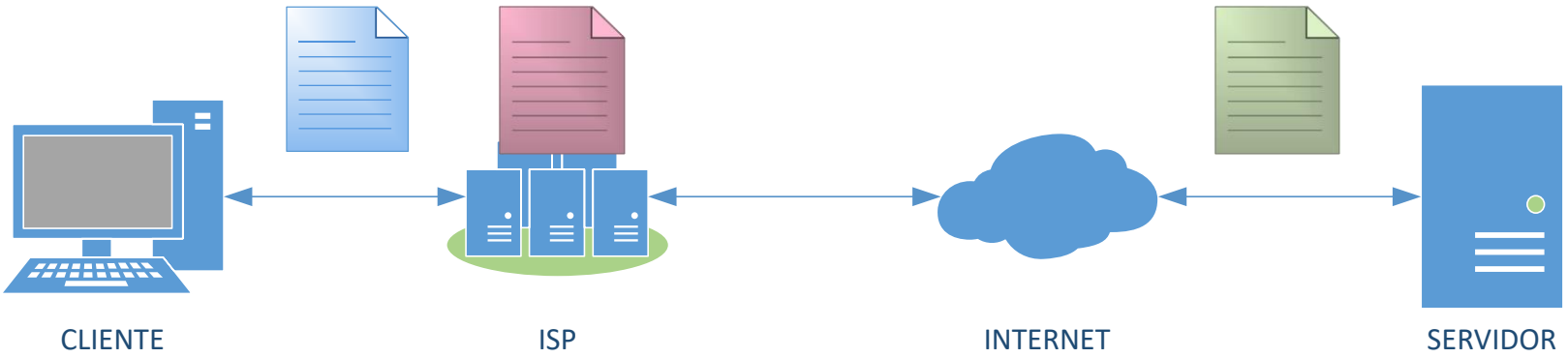
ISP

- **Provedor de serviço internet** (*Internet Service Provider*, **ISP**) é uma empresa ou organização que oferece principalmente serviço de acesso à Internet.

Edge ISP

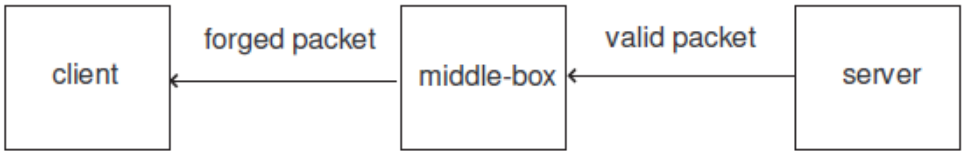


Alteração de conteúdo

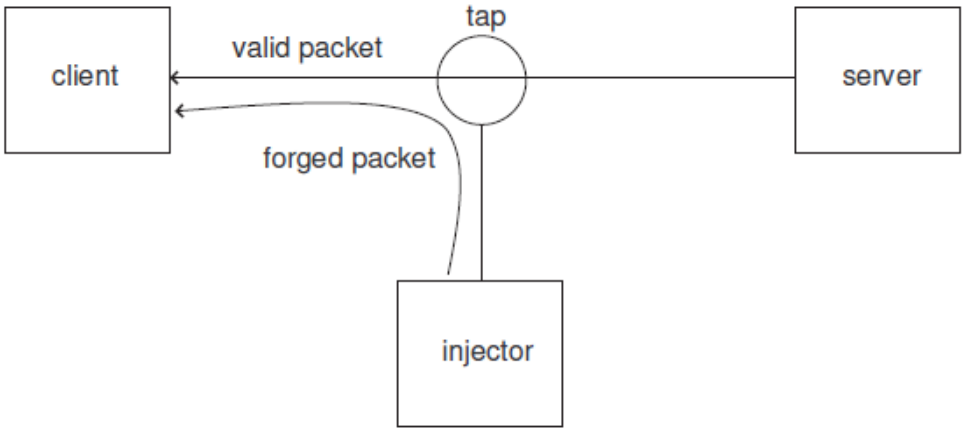


- Facilitação de cache
- Inserção de propaganda
  - erros de DNS
  - erros de HTTP
- Compressão de conteúdo
- CENSURA

In-band vs Out-of-band



(a) In-band alteration of packet by a middle-box. Only a single packet arrives at the client.



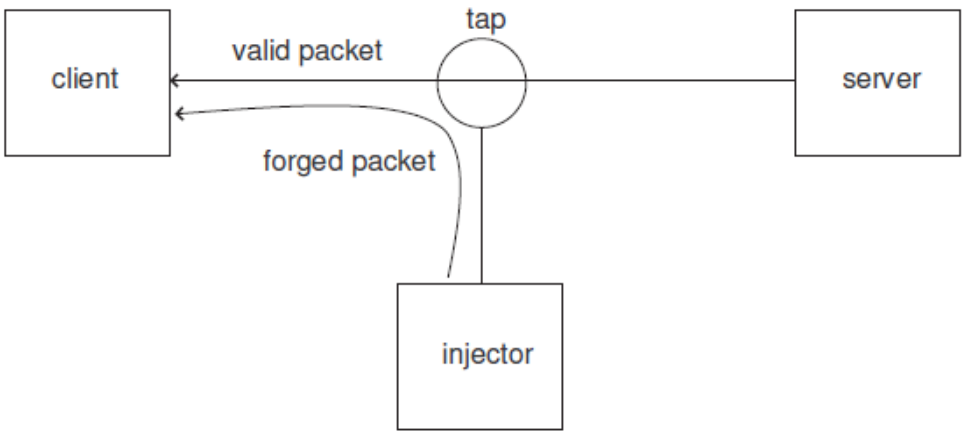
(b) Out-of-band injection of a forged packet. Two packets arrive at the client.

not carried by TLS

not authenticated using TCP authentication

Out-of-band

De acordo com a especificação de TCP, o primeiro byte de dados recebidos para um determinado número de sequência é aceito. Um byte de dados subsequente que tem o mesmo número de sequência é sempre rejeitadas independentemente do seu valor.

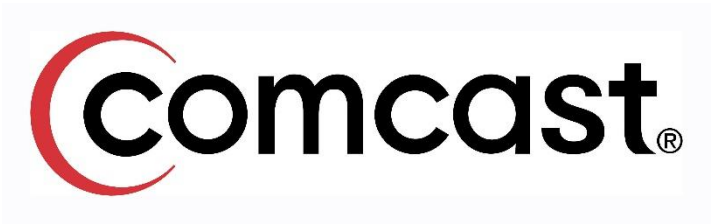


Um cliente HTTP receba apenas uma resposta de HTTP. Quando a resposta forjada é mais curta e chegou antes da resposta válida, então o cliente recebe o fluxo de bytes que inclui a resposta forjada, seguida por bytes da resposta válida.

Conteúdo injetado normalmente faz com que o restante seja ignorado

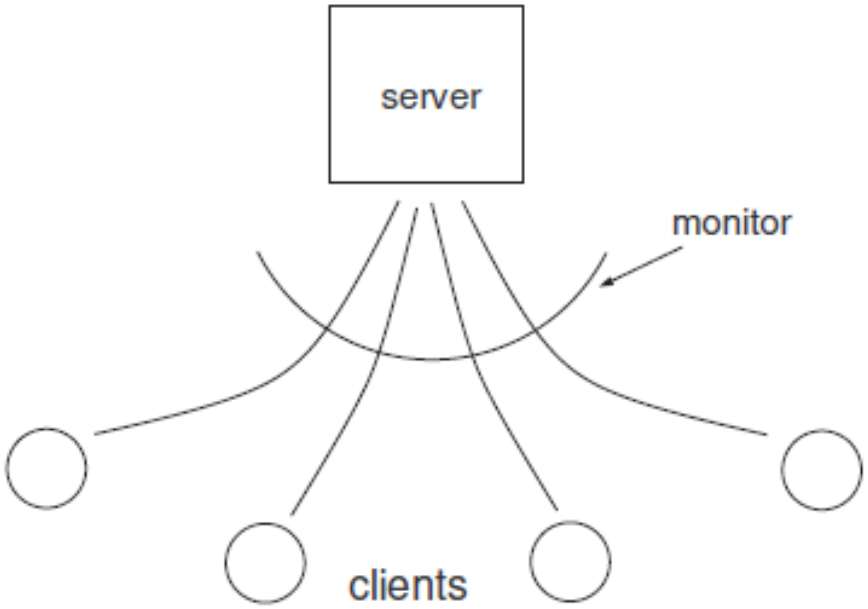
ISP – estudos inserções conhecidas

- CMA Communications in 2013
- Comcast in 2012
- Mediacom in 2011
- WOW! in 2008
- Rogers in 2007



Server-centric approach

- Um ou mais servidores possuem um código específico
- Vários clientes requisitam o conteúdo
- JavaScript analisa o código entregue aos clientes comparando-o com o dos servidores
- Verifica-se as incoerências



Server-centric approach

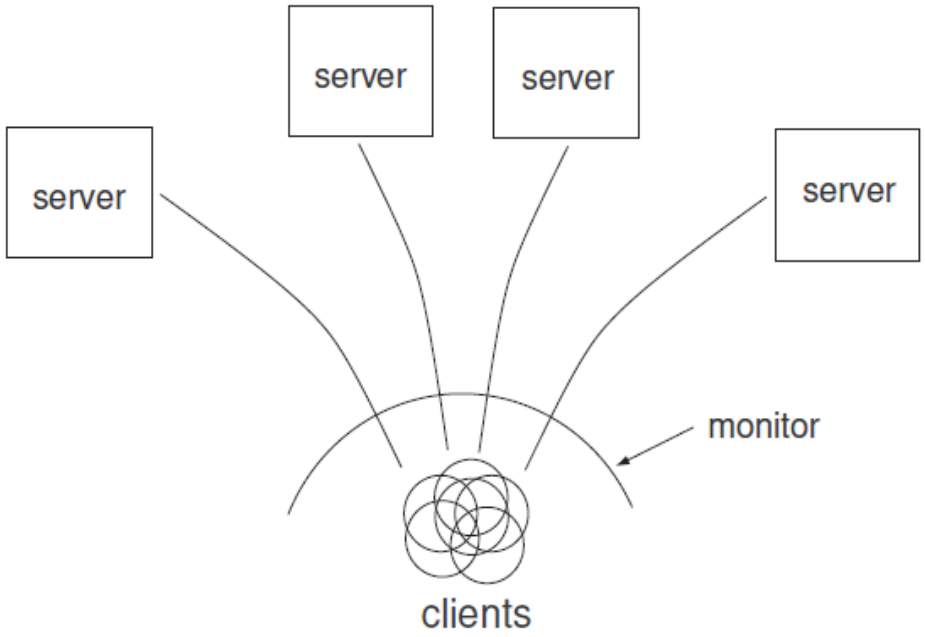
- Vantagem: Diversos clientes de diversas regiões geográficas que são atendidos por diferentes edge ISPs.
- Desvantagem: conteúdo buscado pelos clientes é muito específico (arquivos de teste em servidores de teste). Todos os clientes obtém o mesmo conteúdo dos mesmos servidores web de testes e as ISPs costumam detectar isso*. Não representa a realidade.
- **In some cases these network entities modify all internet traffic originating from very popular websites such as google.com, apple.com, and bing.com or all Internet traffic originating from .com.*

Resumo do trabalho

- Server-centric approach não é eficiente, pois perde uma grande porção de alterações
- Monitoramento durante várias semanas
- HTTP porta 80
- Petabytes
- Mais de 1,5 milhão de endereços IP
- Revelaram operadores de rede que alteram o conteúdo com base no site acessado
- Alterações incluem inserção de propagandas e conteúdo malicioso
- *Toda conexão de internet está sujeita a alteração de conteúdo pelos operadores de rede*

Client-centric approach

- Out-of-band tenta entregar os dois pacotes, o válido e o alterado
- É possível detectar o conteúdo alterado



Exemplo de out-of-band

- HTTP GET - cnzz.com – (site chinês de estatísticas) - javascript

```
GET /core.php?show=pic&t=z HTTP/1.1
User-Agent: Mozilla/5.0 (Windows NT 6.1; WOW64)
Host: c.cnzz.com
Accept-Encoding: gzip
Referer: http://tfkp.com/
```

Exemplo de out-of-band

- Respostas: legítima e alterada (TCPs com mesmo nº de sequência)

```
HTTP/1.1 200 OK
Server: Tengine
Content-Type: application/javascript
Content-Length: 762
Connection: keep-alive
Date: Tue, 07 Jul 2015 04:54:08 GMT
Last-Modified: Tue, 07 Jul 2015 04:54:08
GMT
Expires: Tue, 07 Jul 2015 05:09:08 GMT
!function(){var
p,q,r,a=encodeURIComponent,c=...
```

```
HTTP/1.1 302 Found
Connection: close
Content-Length: 0
Location:
http://adcpc.899j.com/google/google.js
```

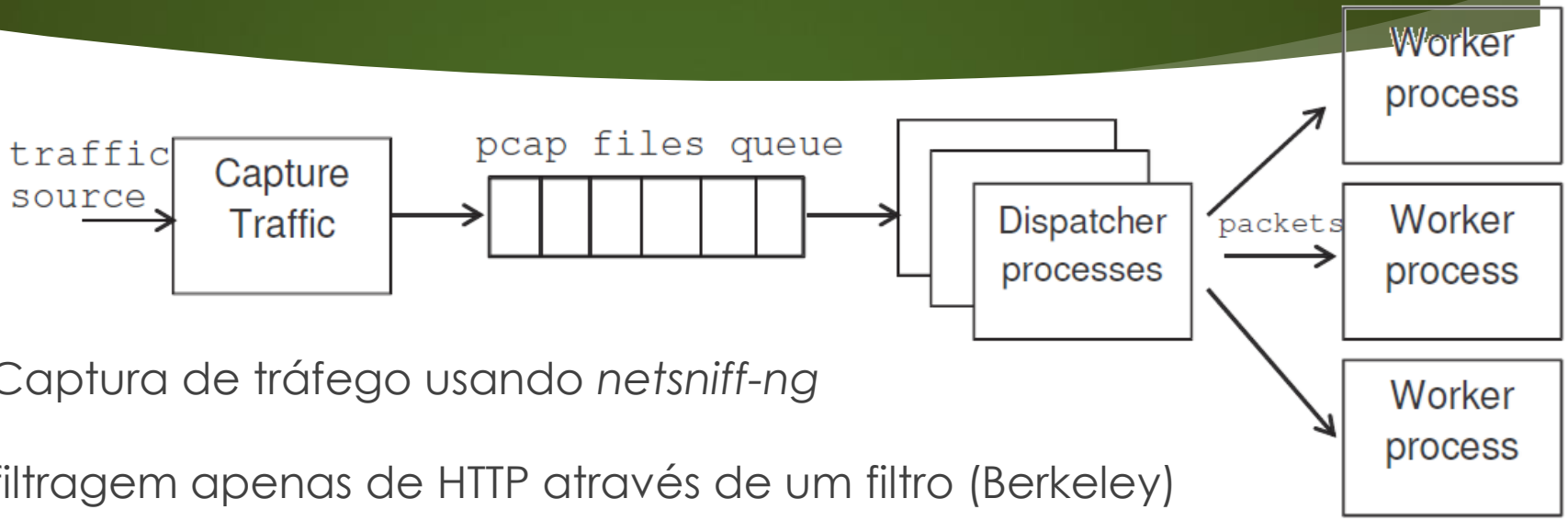
JavaScript redireciona o usuário por uma serie de redes te propaganda e termina da Google's ad network que gera uma propaganda.

Pacote alterado chegou antes que o pacote legítimo, mostrando ao usuário uma propaganda no lugar o conteúdo original

Contribuições

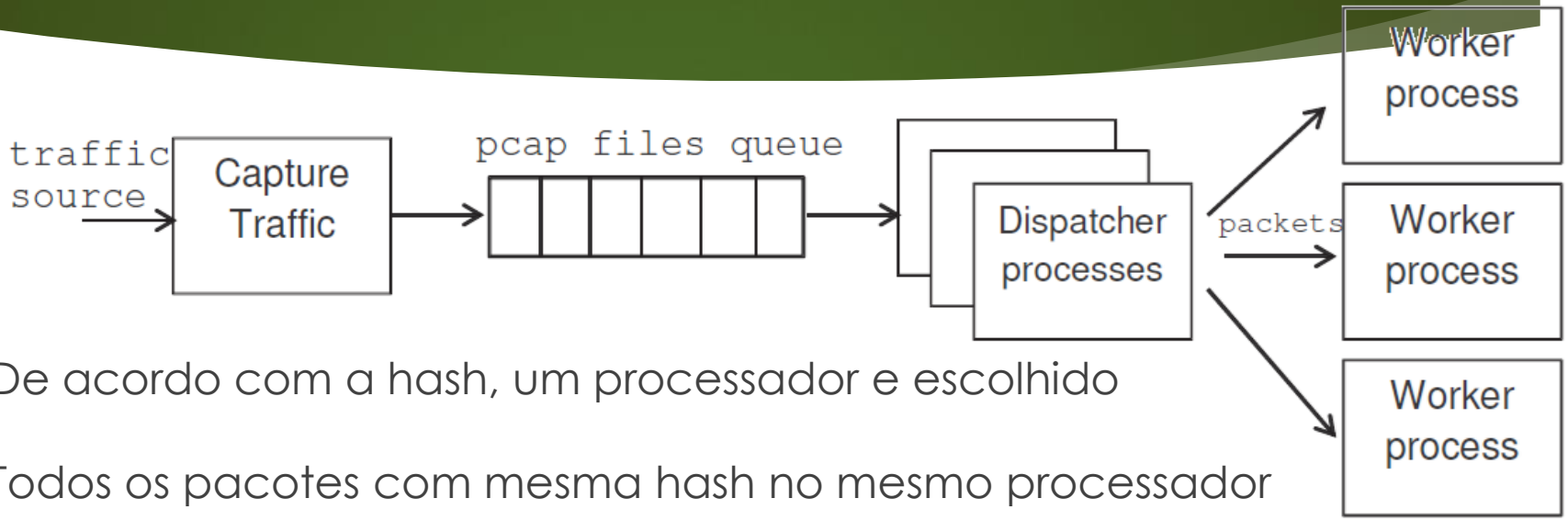
- Provam de que os operadores de rede injetam conteúdo falso out-of-band
- Investigam as identidades dos operadores de rede que fazem isso.
- Analisam a característica das inserções e o propósito delas

Metodologia



- Captura de tráfego usando *netsniff-ng*
- filtragem apenas de HTTP através de um filtro (Berkeley)
- Produção de arquivos com 200 mil pacotes cada
- Arquivos colocados em fila para processamento
- Arquivos processados por um dispatcher: cada pacote é lido nos arquivos e uma hash é calculada com IP e porta TCP – pacotes da mesma sessão possuem mesma hash

Metodologia



- De acordo com a hash, um processador e escolhido
- Todos os pacotes com mesma hash no mesmo processador
- Pacotes com mesma hash são armazenados em uma estrutura de dados que representa a sessão

Metodologia

- Detectar dois pacotes dentro da mesma sessão com o mesmo nº de sequência mas com diferentes *payloads*
- Checados apenas pares de pacotes que chegam em um período menor que 200ms
- Máximo 30 pacotes por sessão

Institution	User base	Monitoring period [week]	Traffic volume [Tb]	Number of sessions [Million]
University A	20,000	2	80	8
University B & University C	50,000	16	1400	120
Enterprise D	5,000	3	24	0.8

Análise das Inserções

- 400 inserções que alteram conteúdo de sites selecionados
- Centenas de inserções com os objetivos de cache e de censura de conteúdo
- Inserções agrupadas em 14 categorias de acordo com o conteúdo inserido
  - Nome do grupo (identificador)
  - Site de detino, tipo do site e onde está hospedado
  - Recurso injetado e propósito

Análise das Inserções

Group name	Destination site(s)	Site type	Location	Injected resource	Purpose
szzhengan	wa.kuwo.cn	Ad network	China	A JavaScript that appends content to the original site	Malware
taobao	is.alicdn.com	Ad network	China	A JavaScript that generates a pop-up frame	Advertise-ment
netsweeper	skyscnr.com	Travel search engine	India	A 302 (Moved) HTTP response	Content filtering
uyan	uyan.cc	Social network	China	A redirection using 'meta-refresh' tag	Advertise-ment
icourses	icourses.cn	Online courses portal	China	A redirection using 'meta-refresh' tag	Advertise-ment
uvclick	cnzz.com	Web users' statistics	Malaysia/China	A JavaScript that identifies the client's device	Advertise-ment
adcpc	cnzz.com	Web users' statistics	Malaysia/China	A 302 redirection to a JavaScript that opens a new window	Advertise-ment
jiathis	jiathis.com	Social network	China	A redirection using 'meta-refresh' tag	Advertise-ment

Análise das Inserções

server erased	changsha.cn	Travel	China	Same as legitimate response but the value of HTTP header 'Server' is changed	Content filtering
gpwa	gpwa.org	Gambling	United States	A JavaScript that redirects to a resource at qpwa.org	Malware
tupian	www.feiniu.com www.j1.com	e-commerce	China	A JavaScript the directs to a resource at www.tupian6688.com	Malware
mi-img	mi-img.com	Unknown	China	A 302 redirection to a different IP	Malware
duba	unknown	Unknown	China	A JavaScript that prompts the user to download an executable	Malware
hao	02995.com	Adware-related	China	A 302 (Moved) HTTP response	Advertise-ment

Análise das Inserções

uyan	uyan.cc	Social network	China	A redirection using 'meta-refresh' tag	Advertise-ment
icourses	icourses.cn	Online courses portal	China	A redirection using 'meta-refresh' tag	Advertise-ment
jiathis	jiathis.com	Social network	China	A redirection using 'meta-refresh' tag	Advertise-ment

www.baidu.com/?tn=95112007_hao_pg

hao123.com

hao	02995.com	Adware-related	China	A 302 (Moved) HTTP response	Advertise-ment

Análise das Inserções

gpwa	gpwa.org	Gambling	United States	A JavaScript that redirects to a resource at qpwa.org	Malware

qpwa.org - Romênia

duba	unknown	Unknown	China	A JavaScript that prompts the user to download an executable	Malware

Botão colorido "Download" que baixa um executável identificado como vírus por vários antivírus

mi-img	mi-img.com	Unknown	China	A 302 redirection to a different IP	Malware

Usuários Android. Redireciona para um apk malicioso conhecido como Android/Gepew.A!tr

Timing

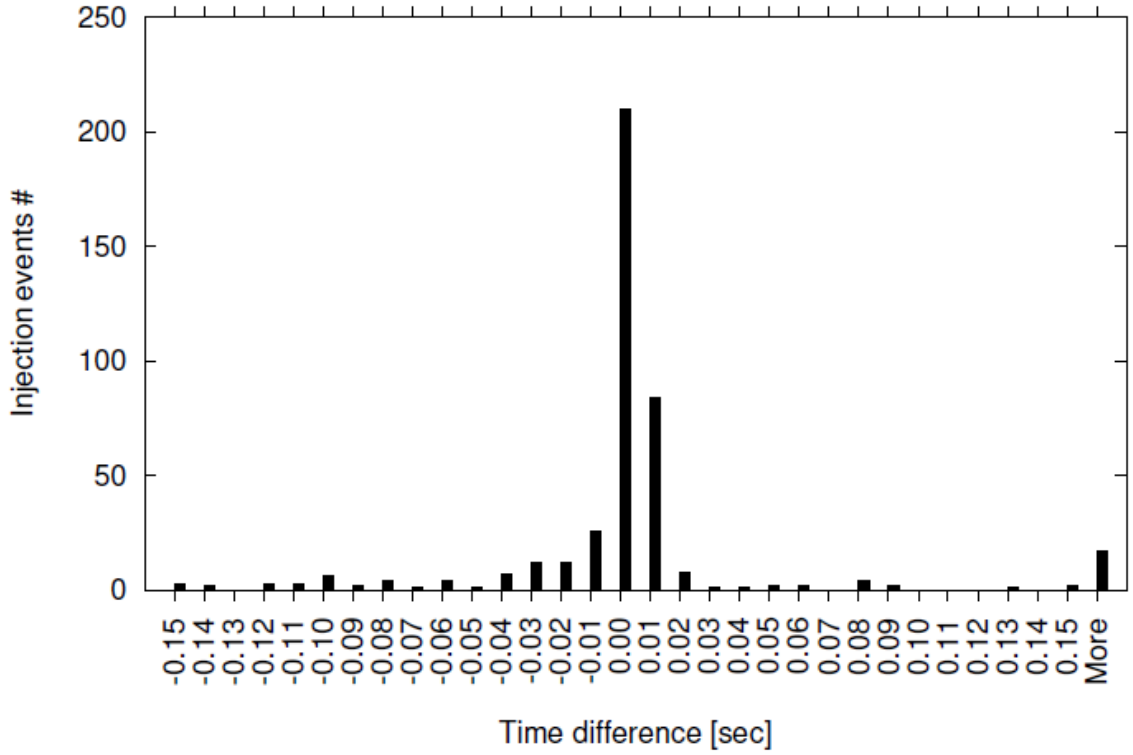
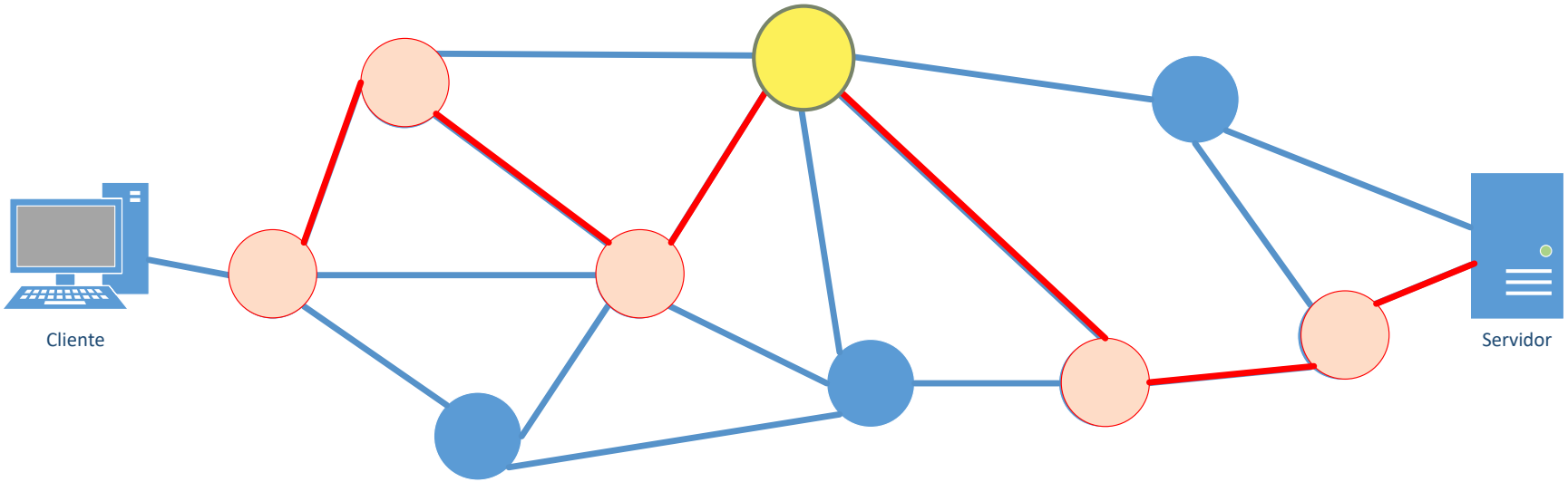


Figure 4: Arrival time difference between the forged and legitimate packets

Identificando os operadores da rede que injetaram conteúdo

- Estimar número de saltos do pacote forjado (comparando com o original) **Ex 3**
- Estimar o caminho do cliente até o servidor usando *traceroute*
- Combinando as duas informações inferir o enlace naquele caminho
- Identificar à quem pertence aquele IP (*BGP advertisements*)



Conclusão

AS number	Operator
17816, 4837	China Unicom
4134, 4812	China Telecom
38182	Extreme Broadband (Malaysia)
6943	Information Technology Systems (US)

Table 4: The operators for each suspected injecting autonomous system

- Como evitar? HTTPS

Referências

Website-Targeted False Content Injection by Network Operators
Gabi Nakibly , Jaime Schcolnik e Yossi Rubin - Israel