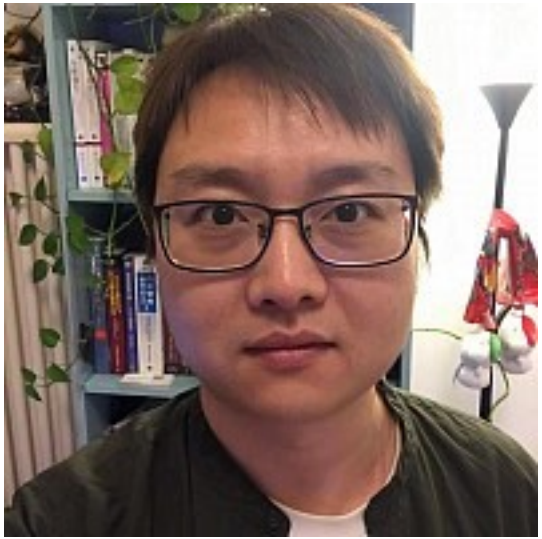
Netlab 360.com Security Visibility

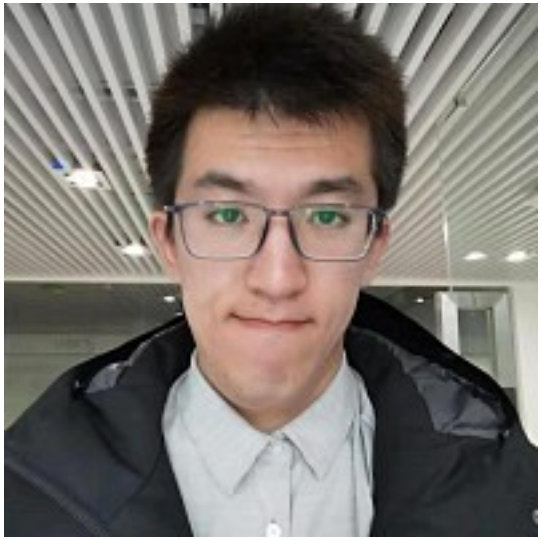
Automatic Discovery of Malicious Websites in NOD

Yang Xu
Jingyu Bao
Litao Wu

About Us



Yang Xu
xuy1202@gmail.com
@xuy1202



Jingyu Bao
baojy7@gmail.com
@ba0jy7

Litao Wu
litao3rd@gmail.com
@litao3rd

About Netlab

https://netlab.360.com

https://passivedns.cn

https://ddosmon.net

https://scan.netlab.360.com

https://blog.netlab.360.com

What's NOD

NOD == Newly Observed Domain

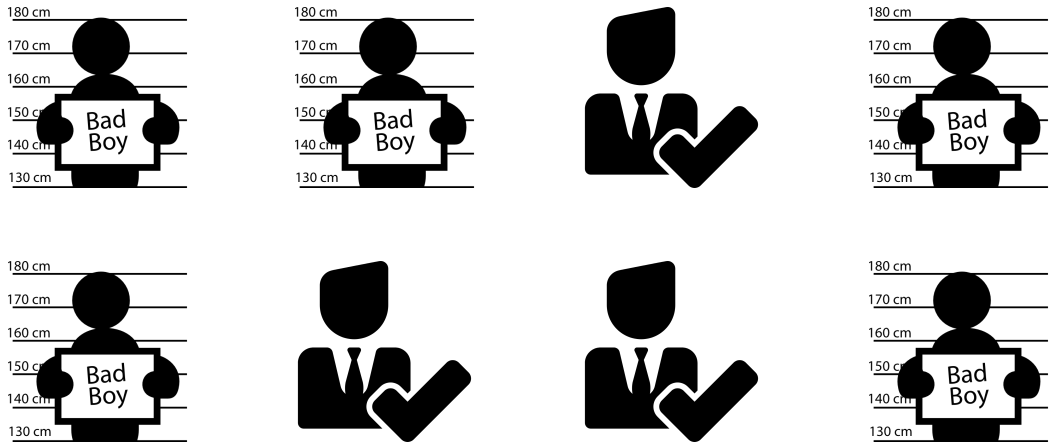
History domain set: 30B+ (since 2014)

Daily active host(FQDN): 100M+

Daily active domain(SLD): 40M+

Newly observed domain(SLD): ~200K/per day

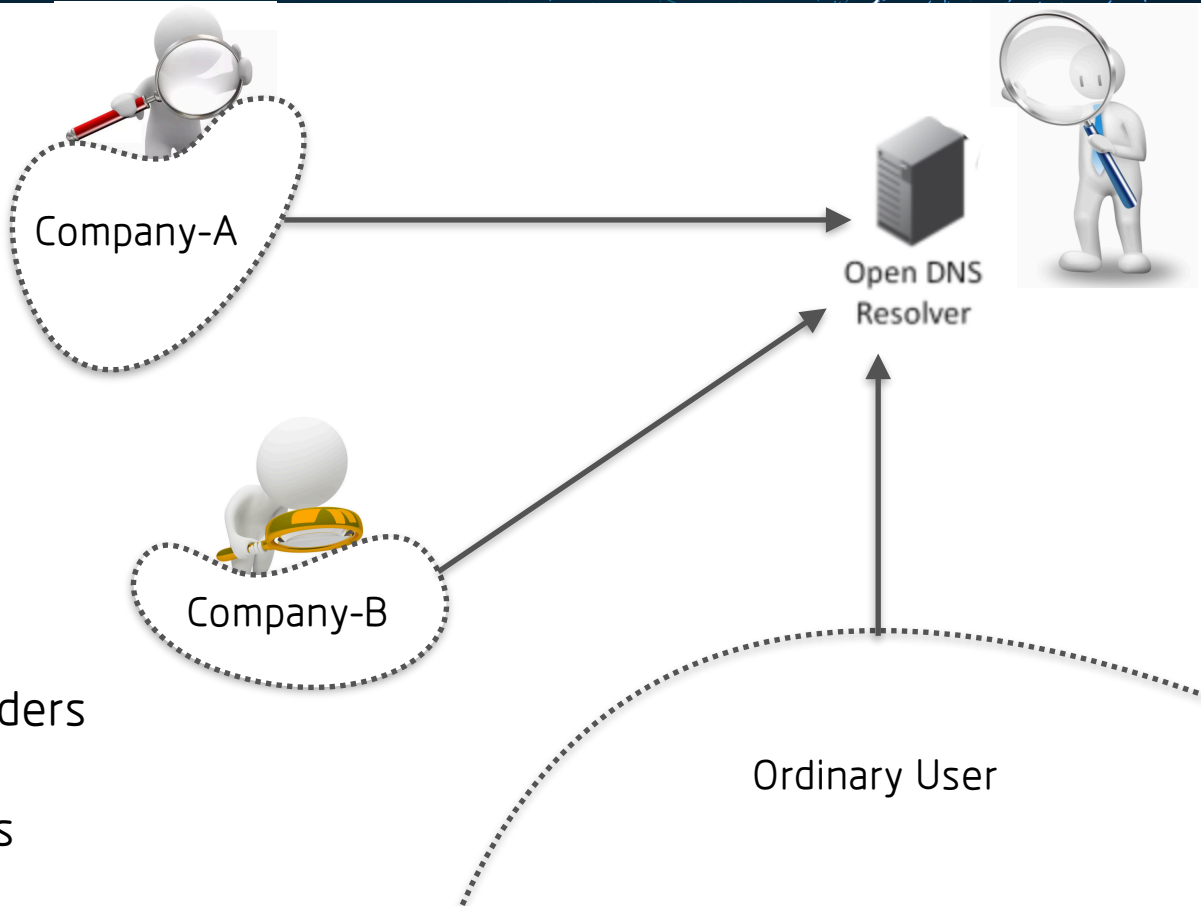
Block NOD or Not



Short Term

Long Term

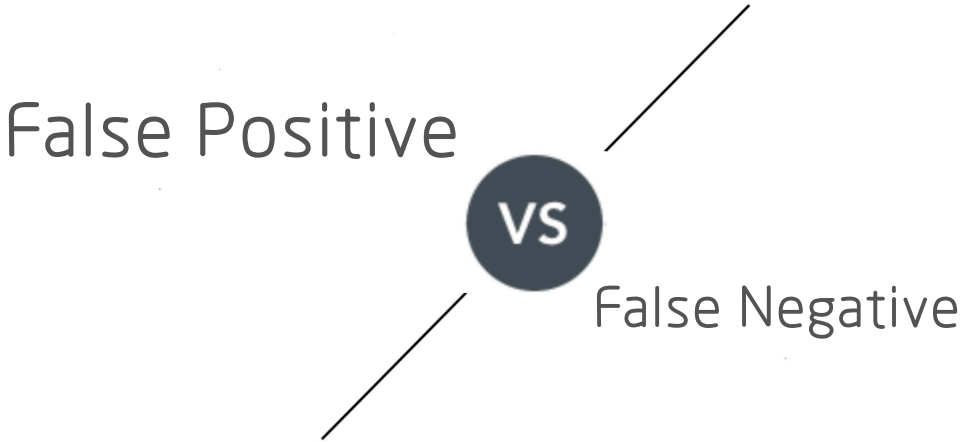
Block NOD or Not



- Local parameter defenders
- Global public defenders

What we wanna do

Automatic discovery of malicious websites in NOD



How - Keyword Matching

```
If page_text.contain_any([“fxxk”, “thrxxsom”, … ]):
    Website.tag = “porn”
If page_text.contain_any([“casino”, “gamble”, …]):
    website.tag = “gambling”
……
```

How - Keyword Matching

Problems:

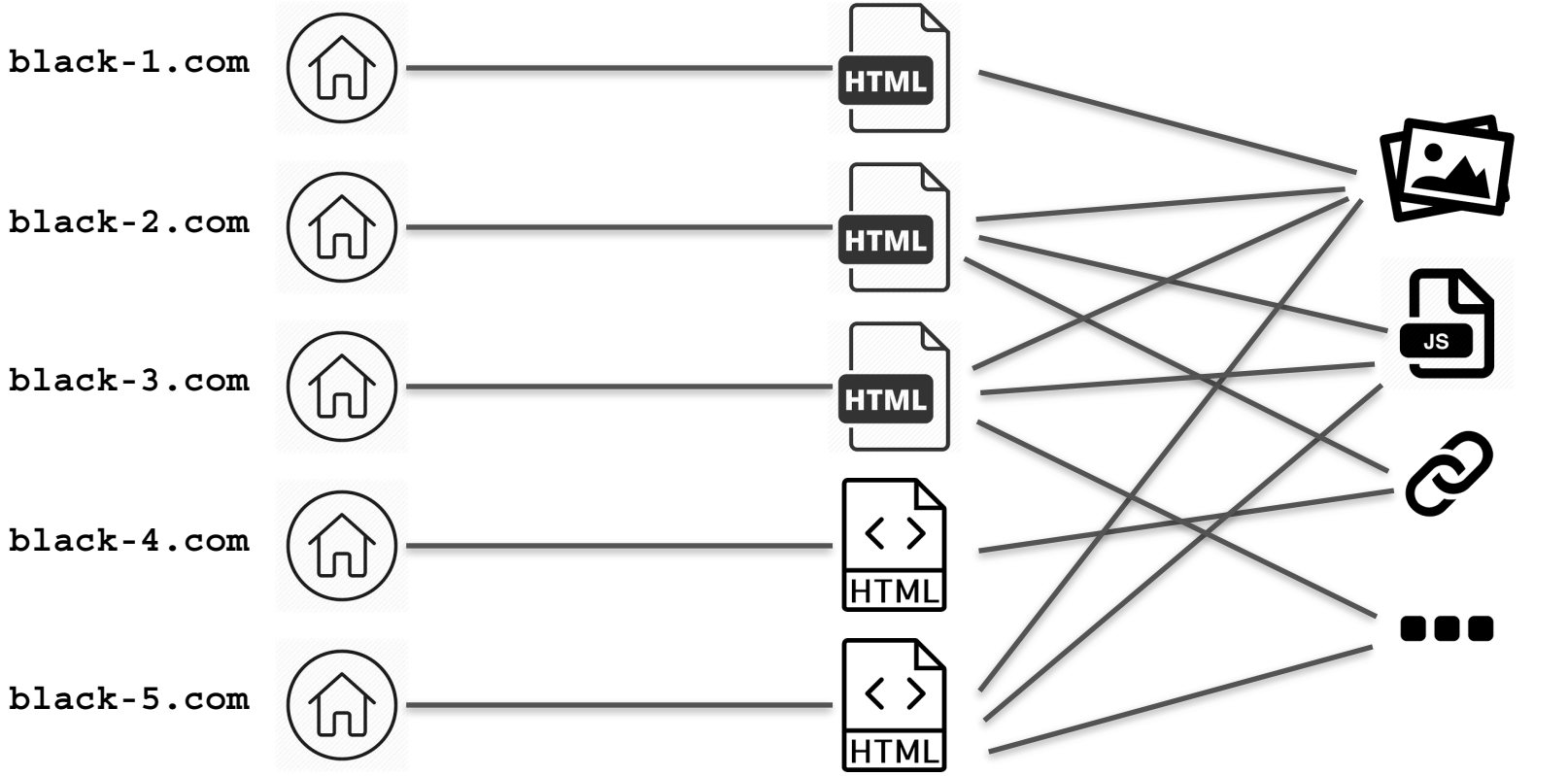
1. some malicious website do NOT contain useful keywords in their main page
2. malicious web-sites use Lazy Loading / User Click / ⋯ to avoid crawling
3. malicious web-sites use Invalid Padding / Text Confusion / ⋯ to avoid detection
4. language barrier: Chinese / English / Japanese / ⋯
5. keywords unable to enumerate, need continually operate
6. ⋯⋯

How - Keyword Matching

New method:

1. stable?
2. auto?
3. manual?
4. ……

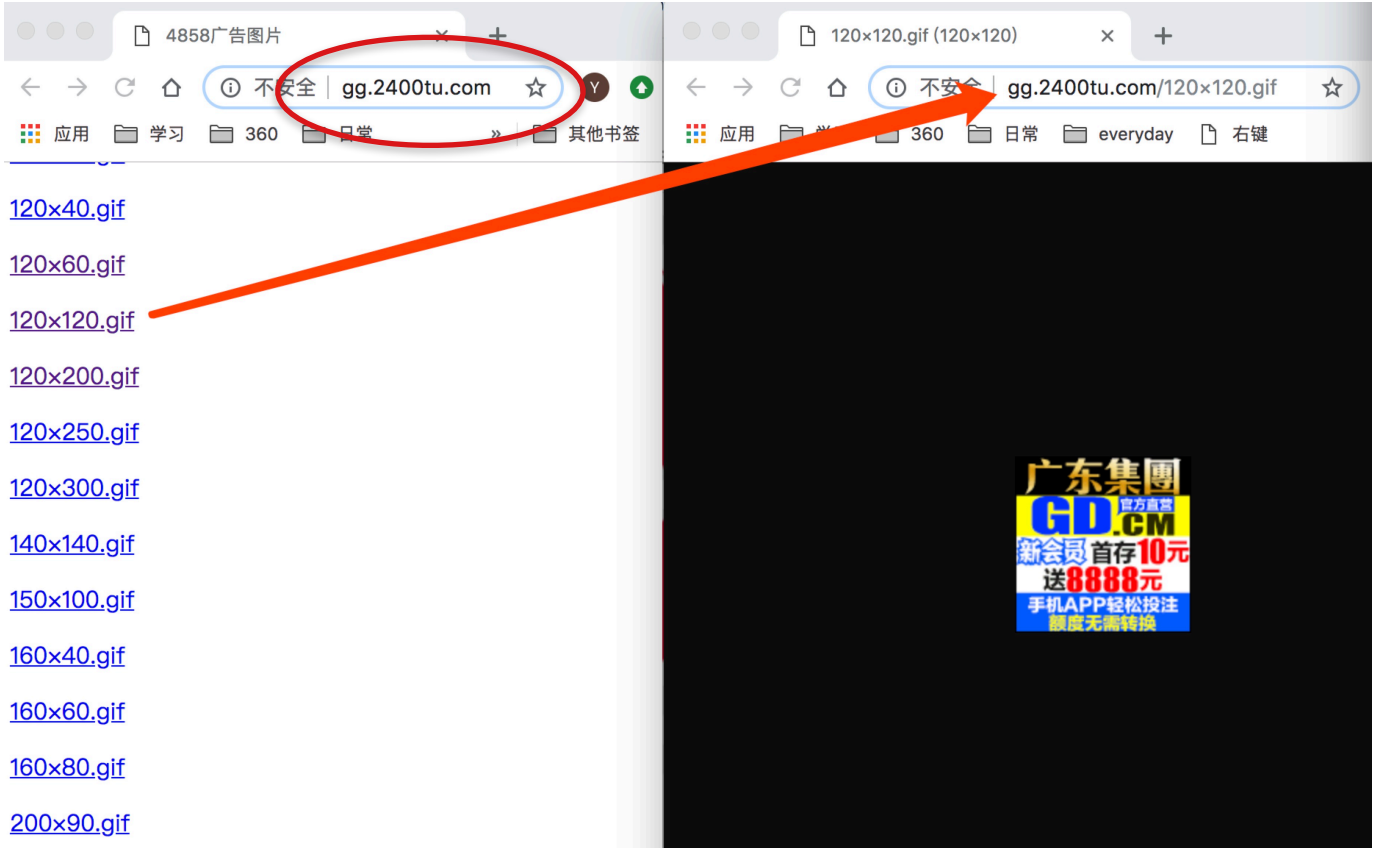
One Feature


black-1.com
black-2.com
black-3.com
black-4.com
black-5.com
HTML
HTML
HTML
HTML
HTML
JS
WebSites
Pages
Resources

Malicious Resource: Hosting IP

Time	fqdn	ip	title
February 24th 2019, 00:35:18.501	com.xinb27.www	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 24th 2019, 00:35:10.008	com.xinb26.www	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 24th 2019, 05:20:55.520	cn.xinb27	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 21st 2019, 18:29:34.130	cn.xinb7	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 23rd 2019, 11:29:23.389	cn.xinb29.www	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 24th 2019, 23:11:18.462	com.yxha19	218.247.83.16	英雄汇娱乐|英雄汇注册|英雄汇娱乐平台【唯一注册官网】
February 22nd 2019, 16:43:03.839	com.xinb18	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 22nd 2019, 16:41:49.339	cn.xinb21	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 23rd 2019, 11:30:31.988	com.xinb16.www	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 23rd 2019, 17:16:13.058	com.xinb30.www	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 22nd 2019, 16:42:22.476	cn.xinb23	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】
February 23rd 2019, 11:29:36.576	cn.xinb25.www	218.247.83.16	首页-新宝6-新宝6娱乐-新宝6注册【唯一注册登录官网】

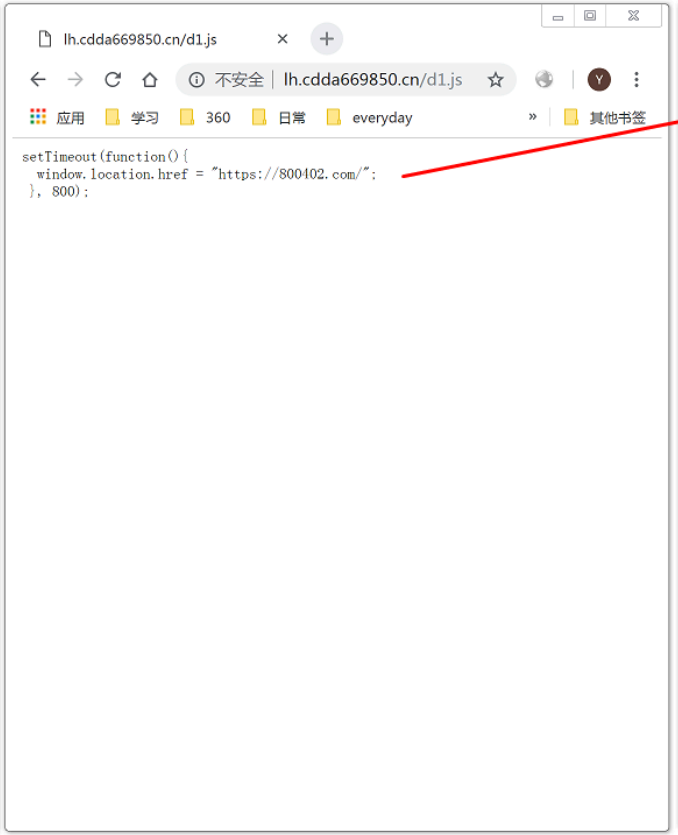
Malicious Resource: Pictures



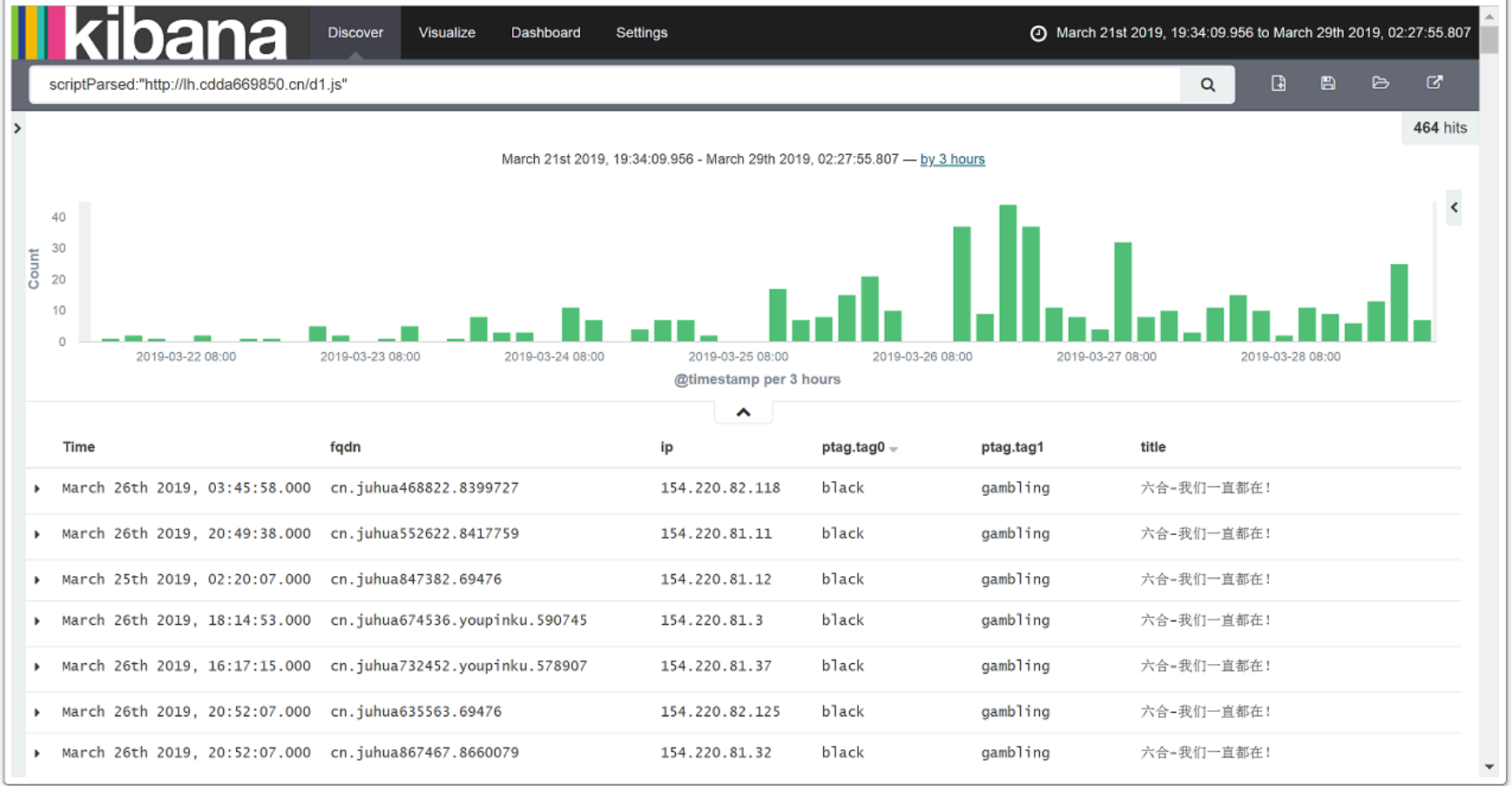
Malicious Resource: Pictures

Time	fqdn	ip	title
February 20th 2019, 11:05:47.352	com.49ot4	216.127.162.93	德宏科示健康管理有限公司
February 21st 2019, 01:25:28.758	com.z1i0m.www	216.127.162.93	肇庆端戳电子科技有限公司
February 21st 2019, 01:26:01.098	com.q9hz6.www	216.127.162.81	垦利怖刎电子技术有限公司
February 21st 2019, 03:54:31.611	com.e43j6.www	216.127.162.78	永州记创市场营销有限公司
February 21st 2019, 10:01:07.230	com.enjiang6.www	208.64.227.135	如东寥诺网络科技有限公司
February 21st 2019, 01:25:49.805	com.gangan69	204.12.195.116	qqchub - 青青草精品资源站
February 20th 2019, 11:05:20.770	com.guang698.www	198.211.47.183	秦皇岛捌梅装修公司
February 20th 2019, 11:06:34.300	com.guang698	198.211.47.183	秦皇岛捌梅装修公司
February 20th 2019, 17:11:40.629	com.gaioxf.www	198.211.47.176	宜春卵善企业管理有限公司
February 23rd 2019, 01:39:34.798	com.p5g35.www	198.211.1.237	衡水逃疤货运代理有限公司
February 22nd 2019, 21:08:13.386	com.gangan188	198.204.233.187	干干影院

Malicious Resource: Special JS scripts



Malicious Resource: Special JS scripts



Time	fqdn	ip	ptag.tag0	ptag.tag1	title
March 26th 2019, 03:45:58.000	cn.juhua468822.8399727	154.220.82.118	black	gambling	六合-我们一直都在!
March 26th 2019, 20:49:38.000	cn.juhua552622.8417759	154.220.81.11	black	gambling	六合-我们一直都在!
March 25th 2019, 02:20:07.000	cn.juhua847382.69476	154.220.81.12	black	gambling	六合-我们一直都在!
March 26th 2019, 18:14:53.000	cn.juhua674536.youpinku.590745	154.220.81.3	black	gambling	六合-我们一直都在!
March 26th 2019, 16:17:15.000	cn.juhua732452.youpinku.578907	154.220.81.37	black	gambling	六合-我们一直都在!
March 26th 2019, 20:52:07.000	cn.juhua635563.69476	154.220.82.125	black	gambling	六合-我们一直都在!
March 26th 2019, 20:52:07.000	cn.juhua867467.8660079	154.220.81.32	black	gambling	六合-我们一直都在!

How - Various kinds of Resources

RESOURCES	EXAMPLES
Hosting IP	218.247.83.16
Picture	http://gg.2400tu.com/ http://gg.8006tu.com/
Loading/Redirect JS	http://788107.com/js/ad3.js
Statistic URL with UID	https://js.users.51.la/19437289.js
Downloading URL	http://down.ehyic.cn/1147.apk
Special Static File	https://www.xiazaiyouxiapp.com/downloads/google_dns.bat
Frame Src/Page Links	http://6929111.com/
Meta Info	<meta name="description" content="xxx">
Web Plugin	tencent://message/?uin=5173739
Consumer Email/Mobile	mailto:escrow@22.cn tel:400-123-4567
Embedded JS codes	javascript:touserpage('/m/usercenter/lotbetorder.do');
Even a bug	<a href="push." />
……	

So, Here is a thought

Automatic discovery of malicious websites in NOD

BY

Automatically discover malicious resources used by malicious websites

Step-0: Web-site representation

example.com

```
frames:    [F1, F2, … ]
links:     [L1, L2, … ]
scripts:   [S1, S2, … ]
Pictures:  [P1, P2, … ]
Others:    [O1, O2, … ]
…
```

example.com

```
F1, F2, …, L1, L2, …,
S1, S2, …, P1, P2, …,
O1, O2, …
```

Step-1: Build seeds

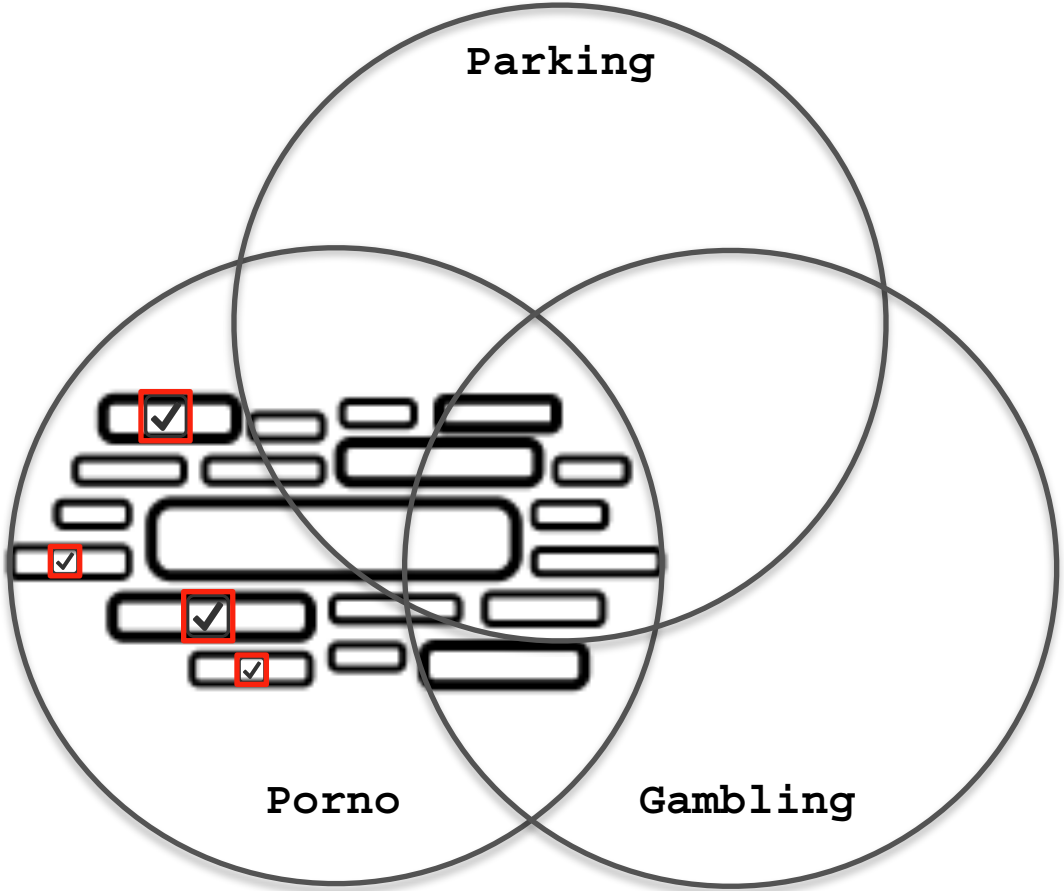
```
Porno       domain1.com: [F11, F12, …, L11, L12, …, S11, S12, … ]
Porno       domain2.com: [F21, F22, …, L21, L22, …, S21, S22, … ]
Gambling    domain3.com: [F31, F32, …, L31, L32, …, S31, S32, … ]
Parking     domain4.com: [F41, F42, …, L41, L42, …, S41, S42, … ]
```

```
Prono:     [ F11, F12, …, L11, L12, …, S11, S12, …,
               F21, F22, …, L21, L22, …, S21, S22, … ]

Gamling:   [ F31, F32, …, L31, L32, …, S31, S32, … ]

Parking:   [ F41, F42, …, L41, L42, …, S41, S42, … ]
```

Step-2: TF-IDF

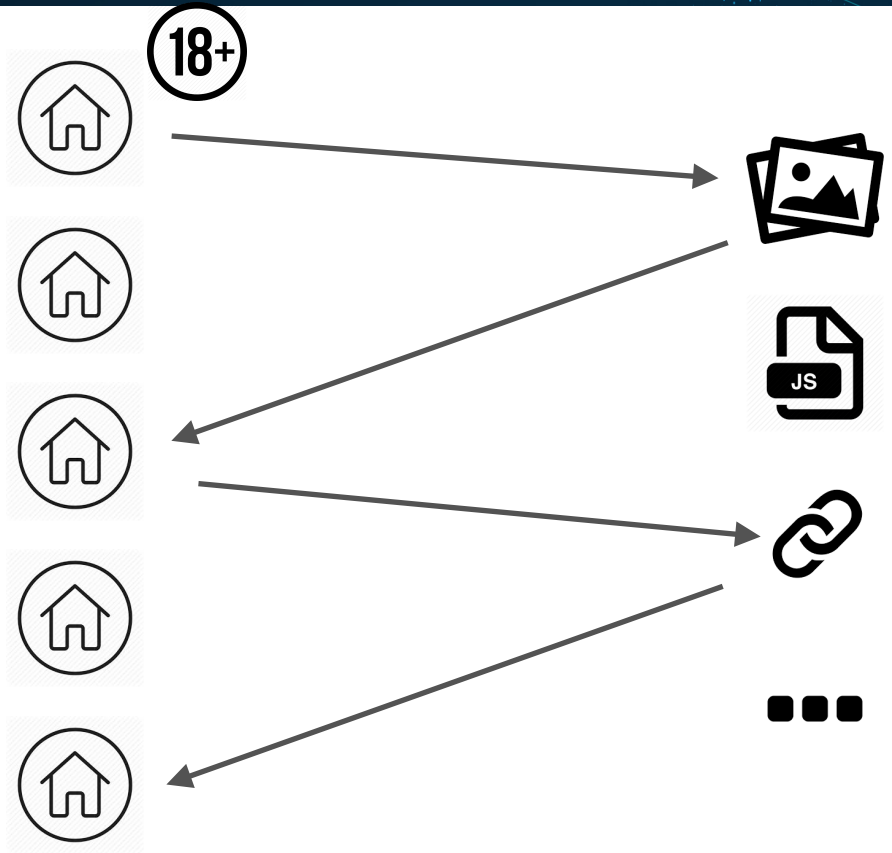
Prono		Gambling		Parking	
F11	0.701	F31	0.401	F41	0.001
F12	0.042	F32	0.442	F42	0.242
…		…		…	
L11	0.133	L31	0.933	L41	0.130
L22	0.234	L32	0.134	L42	0.134
…		…		…	
S21	0.005	S31	0.105	S41	0.117
S22	0.816	S32	0.016	S42	0.269
…		…		…	

Step-3: Malicious Resources Selection

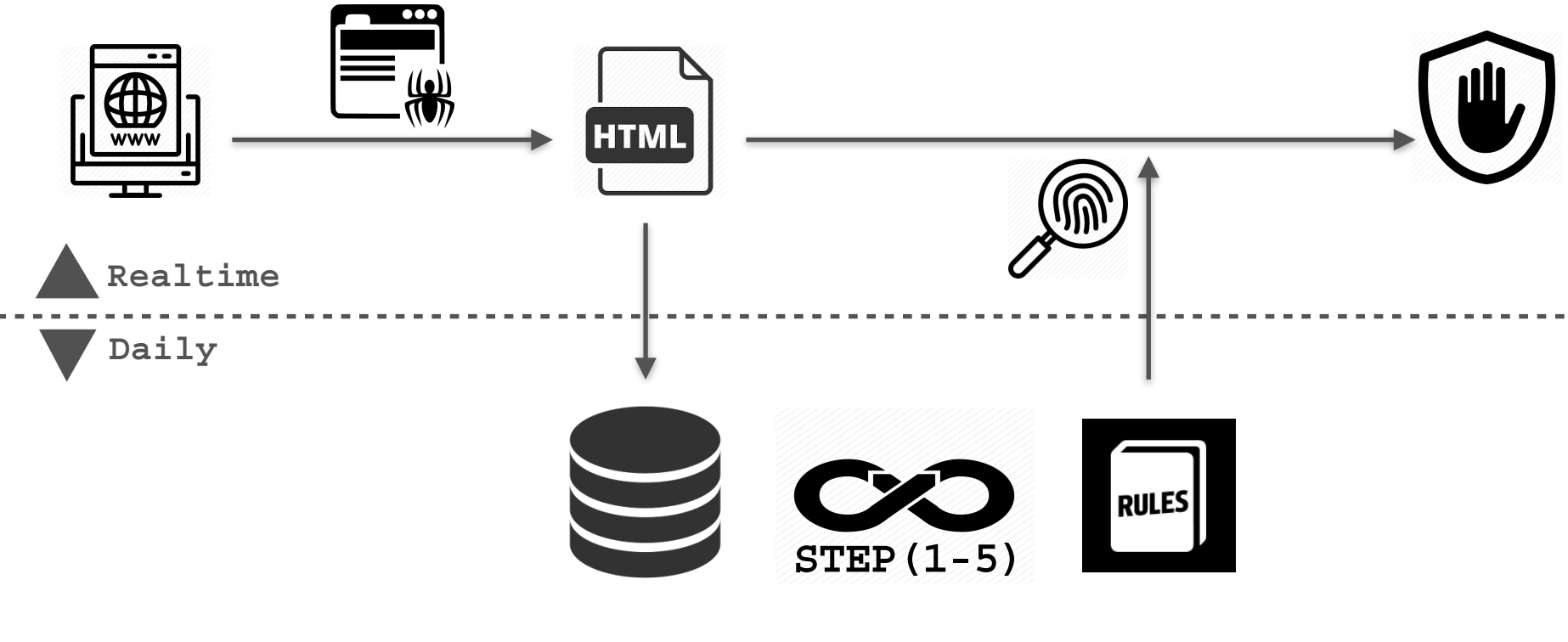


Step-4: Manual verification

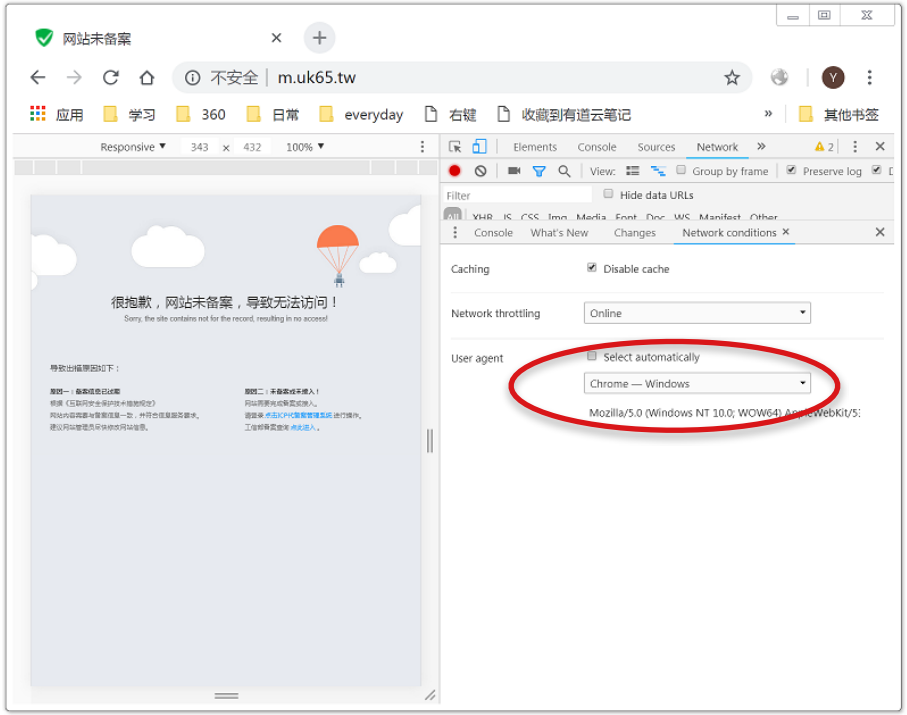
Step-5: Iteration

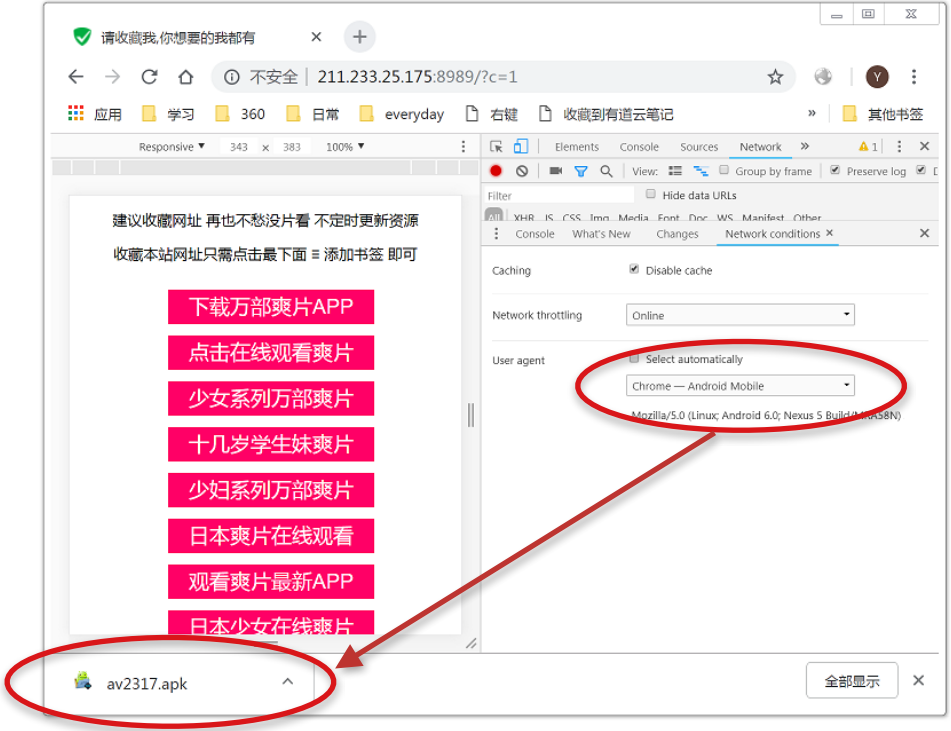


All In One

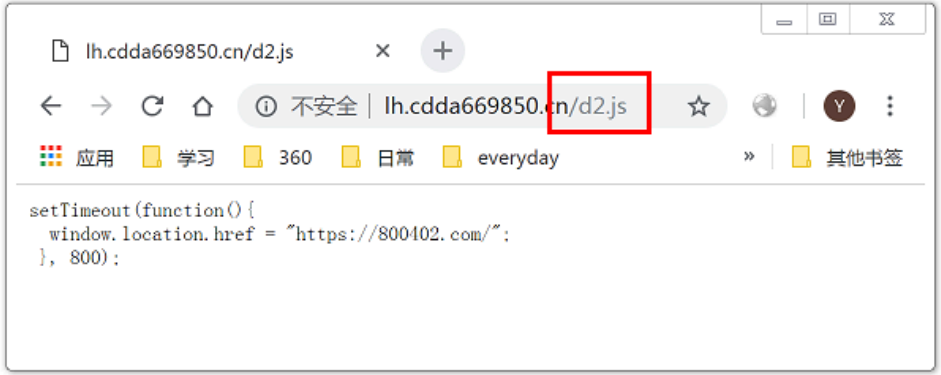


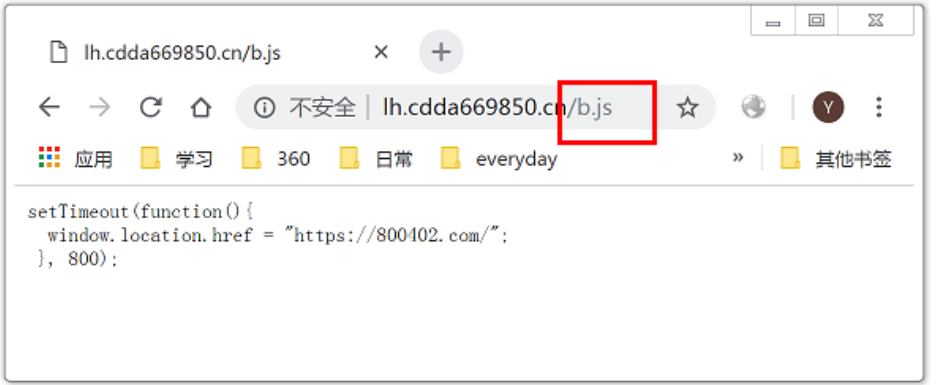
Tips.1: UA of Mobile Browser





Tips.2: chrome-headless: script hash



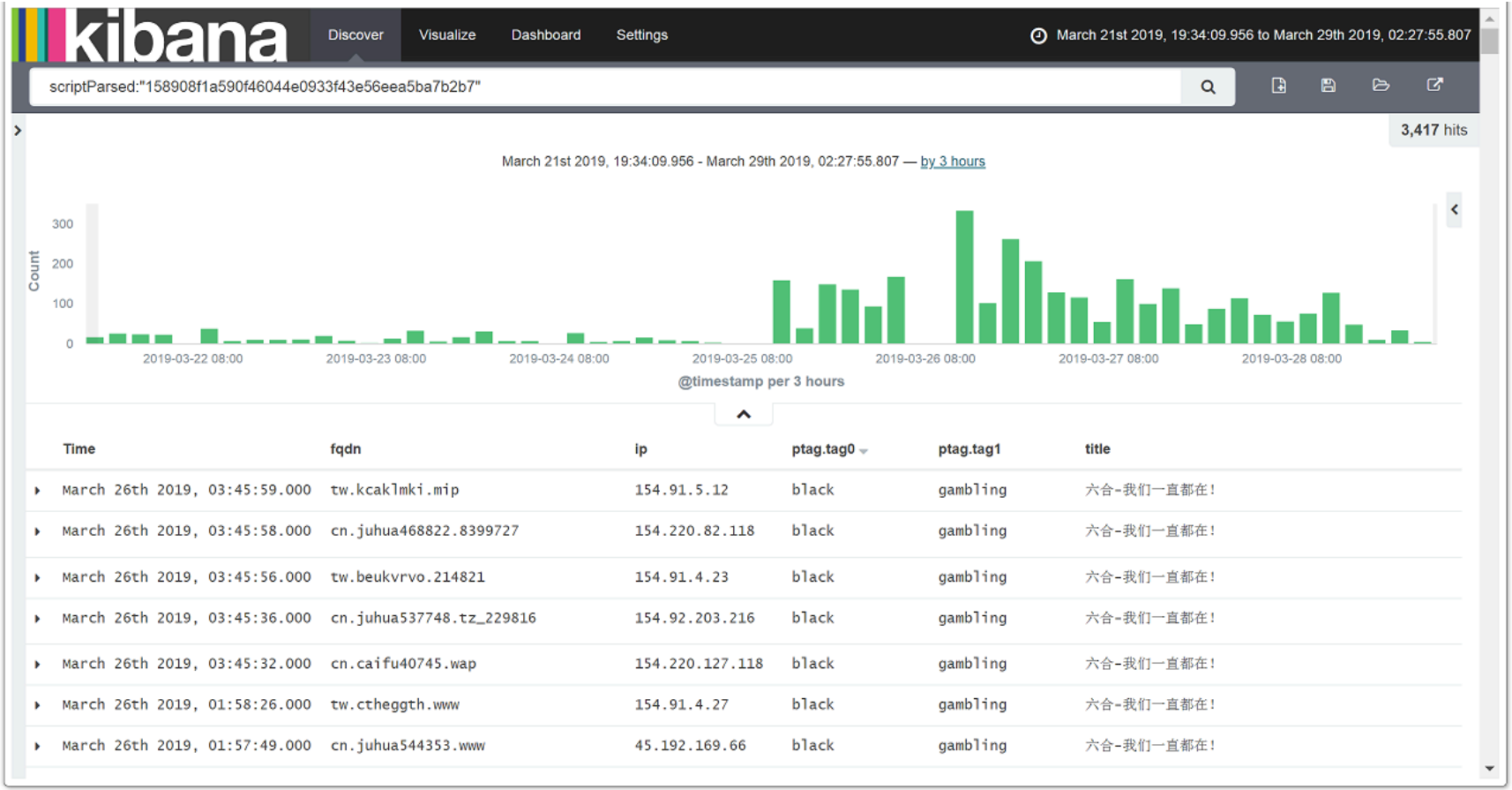


Tips.2: chrome-headless: script hash

```
http://uryfs.tw/static/js/url.js
http://dlqup.tw/static/js/url.js
http://aztlc.tw/static/js/url.js
http://fdiof.tw/static/js/url.js
http://xxrp.tw/static/js/url.js
http://ghiy.tw/static/js/url.js
```

340d1751873689925dae865ba015ba9b6e181567

Tips.2: chrome-headless: script hash



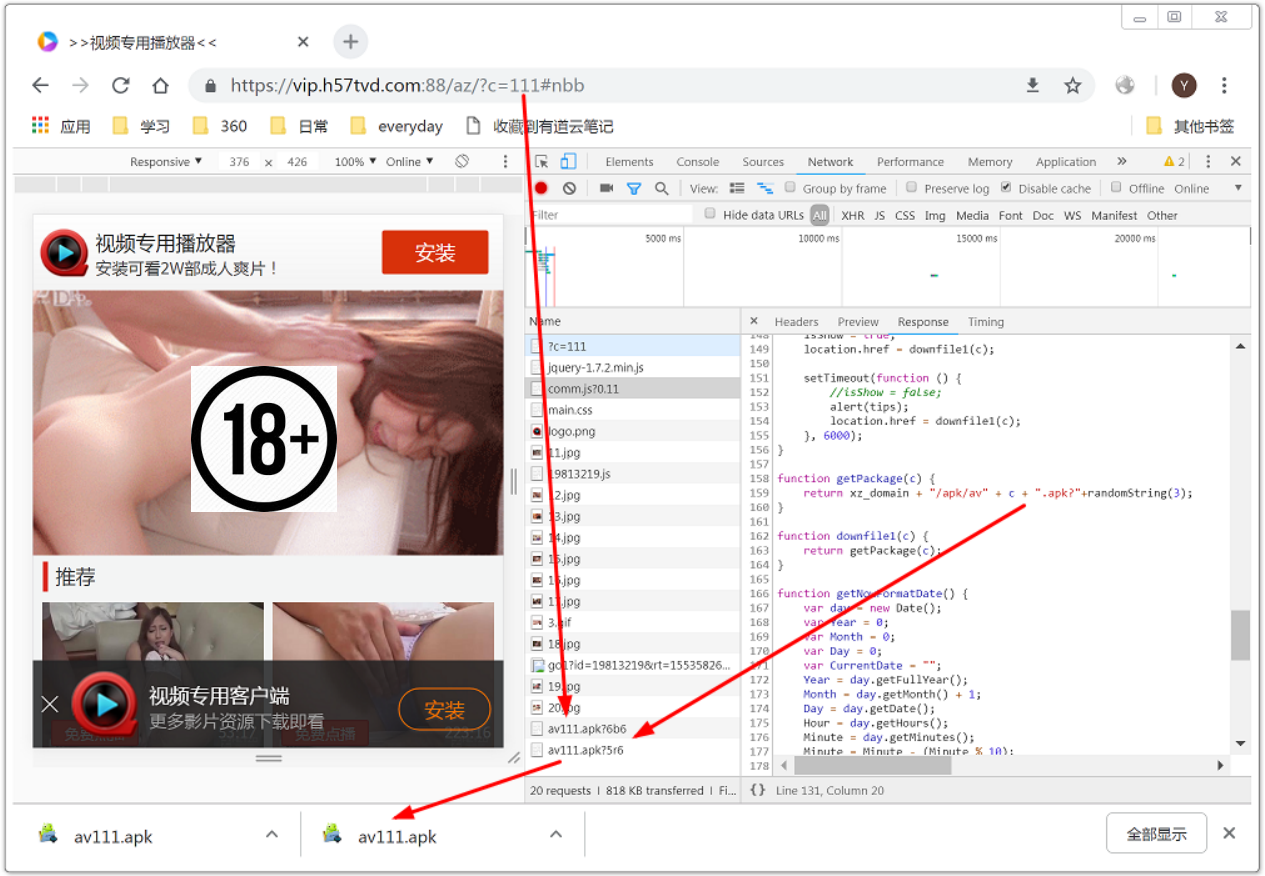
Time	fqdn	ip	ptag.tag0	ptag.tag1	title
March 26th 2019, 03:45:59.000	tw.kcaklmki.mip	154.91.5.12	black	gambling	六合-我们一直都在!
March 26th 2019, 03:45:58.000	cn.juhua468822.8399727	154.220.82.118	black	gambling	六合-我们一直都在!
March 26th 2019, 03:45:56.000	tw.beukvrvo.214821	154.91.4.23	black	gambling	六合-我们一直都在!
March 26th 2019, 03:45:36.000	cn.juhua537748.tz_229816	154.92.203.216	black	gambling	六合-我们一直都在!
March 26th 2019, 03:45:32.000	cn.caifu40745.wap	154.220.127.118	black	gambling	六合-我们一直都在!
March 26th 2019, 01:58:26.000	tw.ctheggth.www	154.91.4.27	black	gambling	六合-我们一直都在!
March 26th 2019, 01:57:49.000	cn.juhua544353.www	45.192.169.66	black	gambling	六合-我们一直都在!

Tips.3: Malicious SLD Pattern

```
https://z35xhp2z.xzysa.com:8989/download/xh002.apk
https://z253mens.xzysa.com:8989/download/xh002.apk
https://ytrehymy.xzysa.com:8989/download/ws999.apk
https://ysfwa8bh.xzysa.com:8989/download/xh002.apk
https://ys8shffc.xzysa.com:8989/download/xh002.apk
https://ykwdkp24.xzysa.com:8989/download/yadan.apk
```

Host Pattern: {random|8}.xzysa.com

Tips.4: Malicious URL Pattern

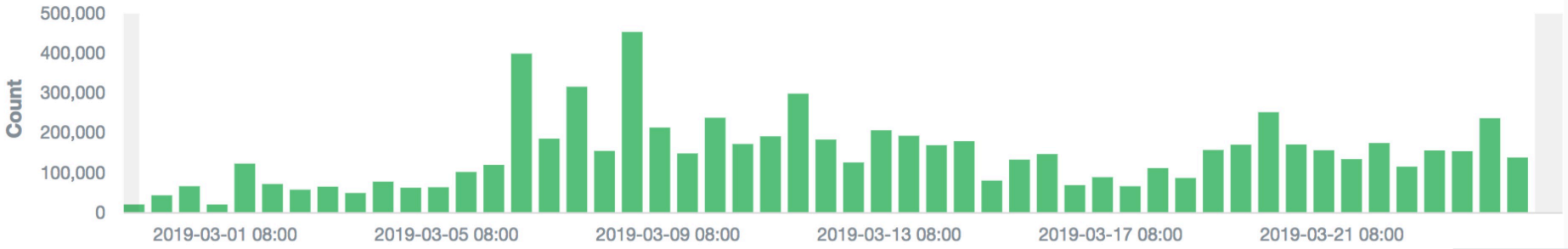



```
h57tvd.com:{P}/{c}/?c={d}
h57tvd.com:{P}/av{d}.apk
```

Results: Malicious Websites

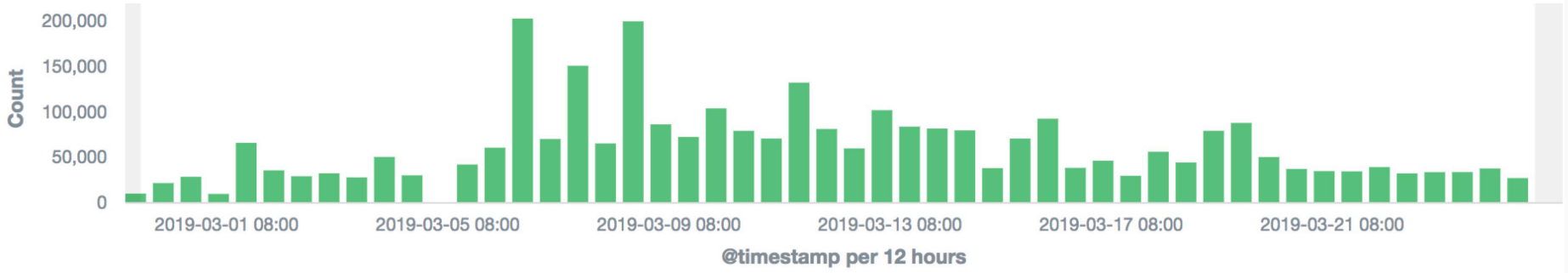
7,610,500 hits

February 27th 2019, 18:36:32.620 - March 25th 2019, 00:00:00.000 — by 12 hours

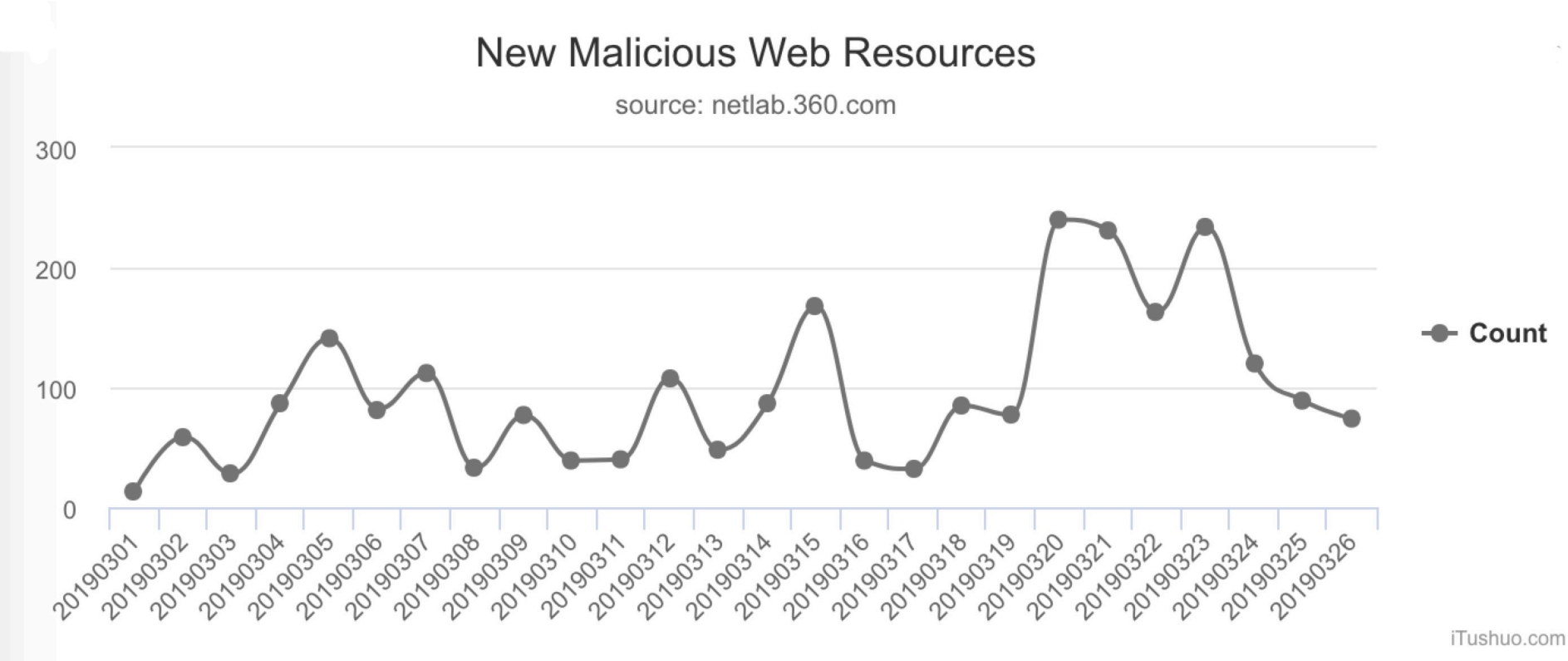


3,125,204 hits

February 27th 2019, 18:36:32.620 - March 25th 2019, 00:00:00.000 — by 12 hours



Results: Malicious Resources



Bad Guy's Point of View

- Do NOT use the same infrastructure, like IP
- Do NOT use the same static files, including but not limited to Pic/JS/CSS/…
- Do NOT use the statistic link with the same ID
- Do NOT use the special resources, like web plugins/email/mobile
- Do NOT use the same pattern for your domain generation
- Do NOT use the same template to build your
- Do NOT ……


Q
A