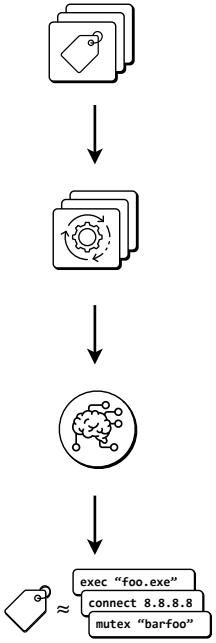
Technische Universität Braunschweig

TagVet: Vetting Malware Tags Using Explainable Machine Learning

Lukas Pirch, Alexander Warnecke, Christian Wressnegger and Konrad Rieck

System Overview



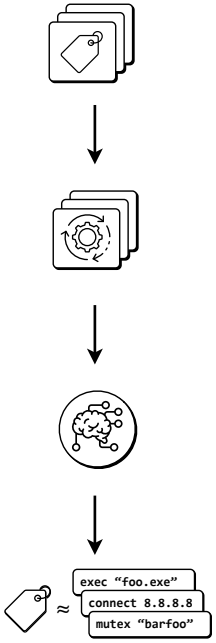
Tagged Malware Samples

Behavior Monitoring

Tag Prediction

Generating Explanations

System Overview



Tagged Malware Samples

Behavior Monitoring

Tag Prediction

Generating Explanations

Behavior Reports - Consolidation

```
<gfn_file_create id="gfn_557" timestamp="29514" symbol_name="CreateFileW"
    post_symbol_name="CreateFile" symbol_offset="0"
    function_call_id="2497" new_process_id="87"
    monitored_process_id="1" thread_id="3" file_obj="e_id_538"
    desired_access="GENERIC_WRITE" create_disposition="CREATE_ALWAYS"
    file_attributes="FILE_ATTRIBUTE_NORMAL" success="True">
  <gob_file id="e_id_538" handle="0xa4"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe"/>
</gfn_file_create>
<gfn_file_write id="gfn_558" timestamp="29515" symbol_name="WriteFile"
    post_symbol_name="WriteFile" symbol_offset="0" function_call_id="2499"
    new_process_id="87" monitored_process_id="1" thread_id="3" file_obj="e_id_538"
    data="BINARY(offset=0x9c426,skipped=False,size=0x11600)" size="71168" success="True"/>
```

Behavior Reports - Consolidation

```
<gfn_file_create id="gfn_557" timestamp="29514" symbol_name="CreateFileW"
    post_symbol_name="CreateFile" symbol_offset="0"
    function_call_id="2497" new_process_id="87"
    monitored_process_id="1" thread_id="3" file_obj="e_id_538"
    desired_access="GENERIC_WRITE" create_disposition="CREATE_ALWAYS"
    file_attributes="FILE_ATTRIBUTE_NORMAL" success="True">
  <gob_file id="e_id_538" handle="0xa4"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe"/>
</gfn_file_create>
<gfn_file_write id="gfn_558" timestamp="29515" symbol_name="WriteFile"
    post_symbol_name="WriteFile" symbol_offset="0" function_call_id="2499"
    new_process_id="87" monitored_process_id="1" thread_id="3" file_obj="e_id_538"
    data="BINARY(offset=0x9c426,skipped=False,size=0x11600)" size="71168" success="True"/>
```

Behavior Reports - Consolidation

```
<gfn_file_create id="gfn_557" timestamp="29514" symbol_name="CreateFileW"
    post_symbol_name="CreateFile" symbol_offset="0"
    function_call_id="2497" new_process_id="87"
    monitored_process_id="1" thread_id="3" file_obj="e_id_538"
    desired_access="GENERIC_WRITE" create_disposition="CREATE_ALWAYS"
    file_attributes="FILE_ATTRIBUTE_NORMAL" success="True"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe">
</gfn_file_create>
<gfn_file_write id="gfn_558" timestamp="29515" symbol_name="WriteFile"
    post_symbol_name="WriteFile" symbol_offset="0" function_call_id="2499"
    new_process_id="87" monitored_process_id="1" thread_id="3" file_obj="e_id_538"
    data="BINARY(offset=0x9c426,skipped=False,size=0x11600)" size="71168" success="True"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe"/>

```

Behavior Reports - Consolidation

```
<gfn_file_create id="gfn_557" timestamp="29514" symbol_name="CreateFileW"
    post_symbol_name="CreateFile" symbol_offset="0"
    function_call_id="2497" new_process_id="87"
    monitored_process_id="1" thread_id="3" file_obj="e_id_538"
    desired_access="GENERIC_WRITE" create_disposition="CREATE_ALWAYS"
    file_attributes="FILE_ATTRIBUTE_NORMAL" success="True"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe">
</gfn_file_create>
<gfn_file_write id="gfn_558" timestamp="29515" symbol_name="WriteFile"
    post_symbol_name="WriteFile" symbol_offset="0" function_call_id="2499"
    new_process_id="87" monitored_process_id="1" thread_id="3" file_obj="e_id_538"
    data="BINARY(offset=0x9c426,skipped=False,size=0x11600)" size="71168" success="True"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe"/>

```

Behavior Reports - Consolidation

```
<gfn_file_create symbol_name="CreateFileW"
    post_symbol_name="CreateFile" symbol_offset="0"
    desired_access="GENERIC_WRITE" create_disposition="CREATE_ALWAYS"
    file_attributes="FILE_ATTRIBUTE_NORMAL" success="True"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe">
</gfn_file_create>
<gfn_file_write symbol_name="WriteFile"
    post_symbol_name="WriteFile" symbol_offset="0" success="True"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe"/>
```

Behavior Reports - Consolidation

```
<gfn_file_create symbol_name="CreateFileW"
    post_symbol_name="CreateFile" symbol_offset="0"
    desired_access="GENERIC_WRITE" create_disposition="CREATE_ALWAYS"
    file_attributes="FILE_ATTRIBUTE_NORMAL" success="True"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe">
</gfn_file_create>
<gfn_file_write symbol_name="WriteFile"
    post_symbol_name="WriteFile" symbol_offset="0" success="True"
    file_name="c:\users\5p5nrgjn0js halpmcxz\appdata\roaming\microsoft\xjvfbd.exe"
    file_name_orig="C:\Users\5p5NrGJn0jS HALPmcxz\AppData\Roaming\Microsoft\xjvfbd.exe"/>
```

Behavior Reports - Consolidation

```
<gfn_file_create symbol_name="CreateFileW"
    post_symbol_name="CreateFile" symbol_offset="0"
    desired_access="GENERIC_WRITE" create_disposition="CREATE_ALWAYS"
    file_attributes="FILE_ATTRIBUTE_NORMAL" success="True"
    file_name="c:\users\*\appdata\roaming\microsoft\*.exe"
    file_name_orig="C:\Users\*\AppData\Roaming\Microsoft\*.exe">
</gfn_file_create>
<gfn_file_write symbol_name="WriteFile"
    post_symbol_name="WriteFile" symbol_offset="0" success="True"
    file_name="c:\users\*\appdata\roaming\microsoft\*.exe"
    file_name_orig="C:\Users\*\AppData\Roaming\Microsoft\*.exe"/>
```

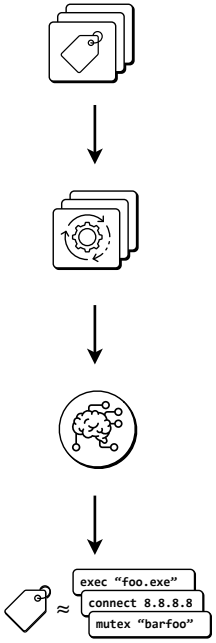
Behavior Reports - Consolidation

```
<gfn_file_create symbol_name="CreateFileW"
    post_symbol_name="CreateFile" symbol_offset="0"
    desired_access="GENERIC_WRITE" create_disposition="CREATE_ALWAYS"
    file_attributes="FILE_ATTRIBUTE_NORMAL" success="True"
    file_name="c:\users\*\appdata\roaming\microsoft\*.exe"
    file_name_orig="C:\Users\*\AppData\Roaming\Microsoft\*.exe">
</gfn_file_create>
<gfn_file_write symbol_name="WriteFile"
    post_symbol_name="WriteFile" symbol_offset="0" success="True"
    file_name="c:\users\*\appdata\roaming\microsoft\*.exe"
    file_name_orig="C:\Users\*\AppData\Roaming\Microsoft\*.exe"/>
```

Behavior Reports - Consolidation

```
<gfn_file_create post_symbol_name="CreateFile"
    desired_access="GENERIC_WRITE" create_disposition="CREATE_ALWAYS"
    file_attributes="FILE_ATTRIBUTE_NORMAL" success="True"
    file_name="c:\users\*\appdata\roaming\microsoft\*.exe"/>
<gfn_file_write post_symbol_name="WriteFile" success="True"
    file_name="c:\users\*\appdata\roaming\microsoft\*.exe"/>
```

System Overview



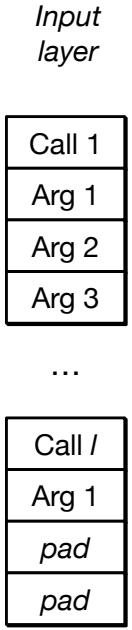
Tagged Malware Samples

Behavior Monitoring

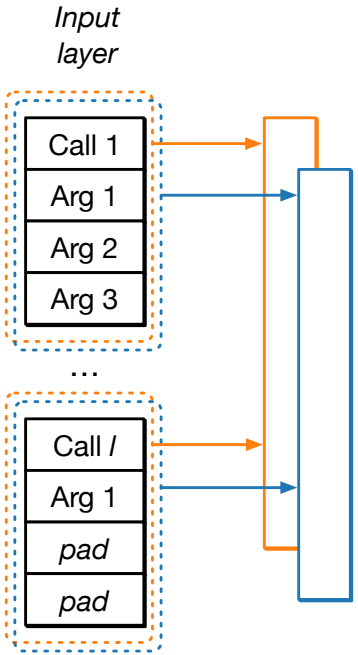
Tag Prediction

Generating Explanations

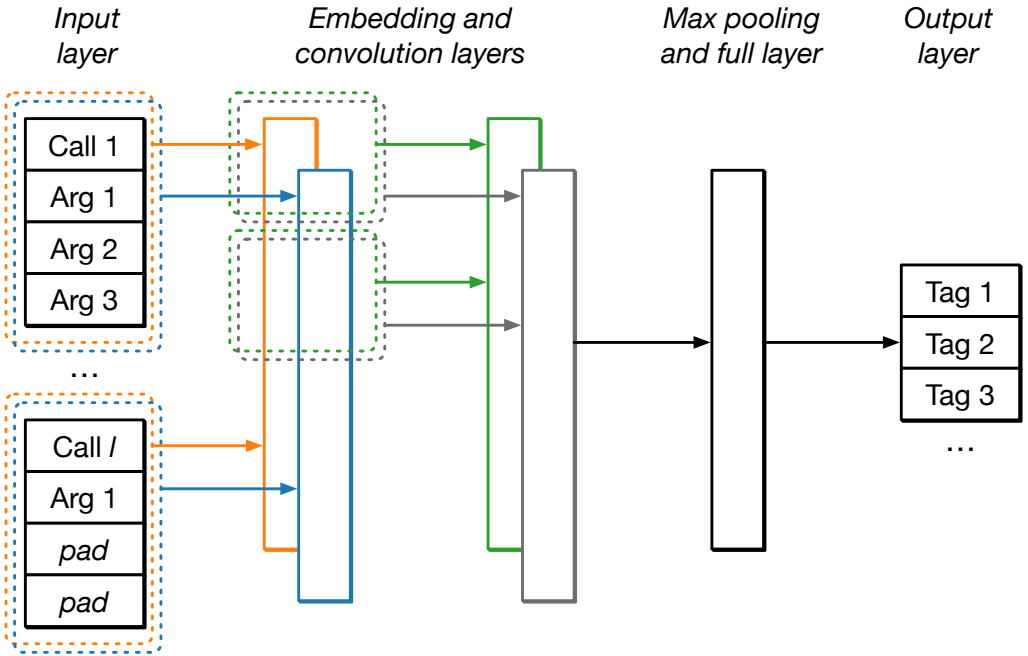
Tag Prediction - CNN



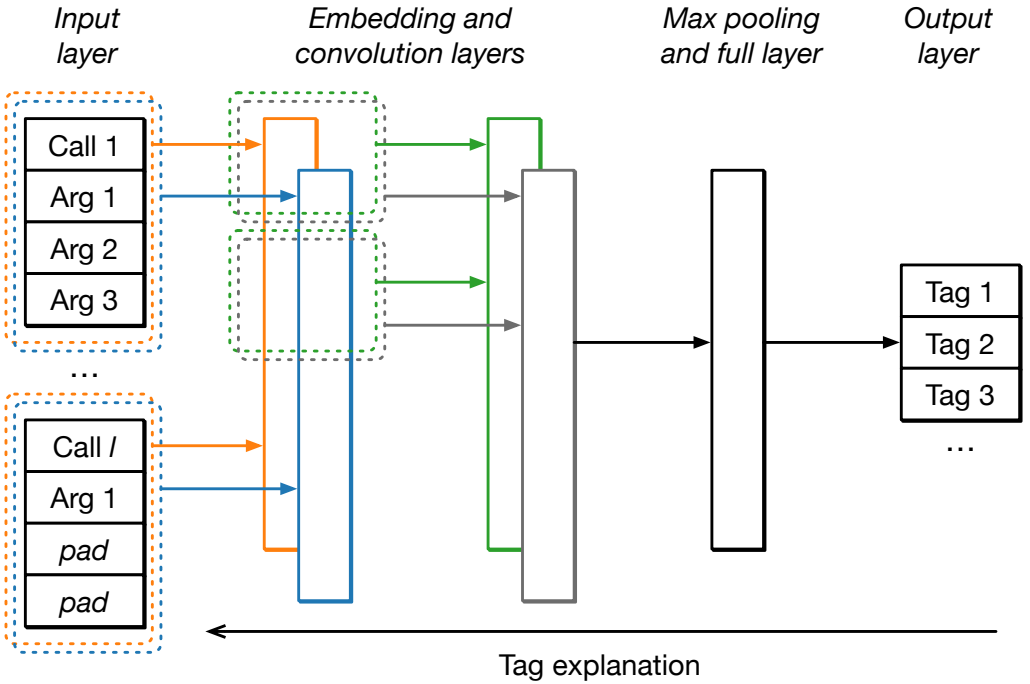
Tag Prediction - CNN



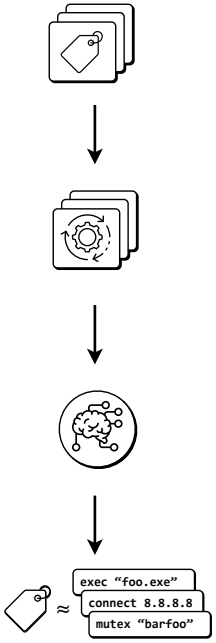
Tag Prediction - CNN



Tag Prediction - CNN



System Overview



Tagged Malware Samples

Behavior Monitoring

Tag Prediction

Generating Explanations

Generating Explanations

- **Layer-wise Relevance Propagation**
  - Gradient-based relevance assignment with conservation property

$$\sum_i R_i^1 = \sum_i R_i^2 = \cdots = \sum_i R_i^L$$

Generating Explanations

- **Layer-wise Relevance Propagation**
  - Gradient-based relevance assignment with conservation property

$$\sum_i R_i^1 = \sum_i R_i^2 = \cdots = \sum_i R_i^L$$

Id	Token	Id	Token
0	proc_create		
1	symbol_name	2	createprocess
3	creation_flags	4	create_suspended
5	show_window	6	sw_hide
7	Success	8	true

Generating Explanations

- **Explanation Aggregation**
  - Aggregate relevances per tag
  - Select top k most relevant features across samples
  - Reconstruct context from known XML structure

Id	Token	Id	Token
0	proc_create		
1	symbol_name	2	createprocess
3	creation_flags	4	create_suspended
5	show_window	6	sw_hide
7	Success	8	true

Generating Explanations

- **Explanation Aggregation**
  - Aggregate relevances per tag
  - Select top k most relevant features across samples
  - Reconstruct context from known XML structure

Id	Token	Id	Token
0	proc_create		
1	symbol_name	2	createprocess
3	creation_flags	4	create_suspended
5	show_window	6	sw_hide
7	Success	8	true

Examples

- **Sandbox Tag “Attempts to connect to unavailable TCP servers”**

Id	Tokens in context
1	sck_connect(*)
2	sck_connect(in:post_symbol_name=“connect”)
3	sck_connect(out:success=“false”)
4	sck_connect(in:remote_port=“*”)

Examples

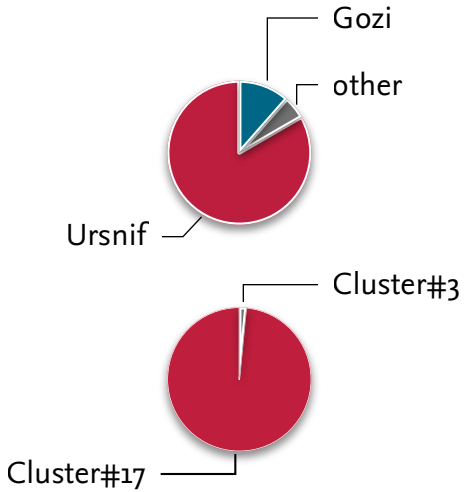
- **Explanation for cluster #17**

Id	Tokens in context
1	mod_map(in:commit_size="0")
2	mod_map(in:zero_bits="0")
3	mod_map(in:protection="page_execute_readwrite")
4	mod_map(out:section_offset="0")
5	mod_get_proc_address(in:function="memset")

Examples

- **Explanation for cluster #17**

Id	Tokens in context
1	mod_map(in:commit_size="0")
2	mod_map(in:zero_bits="0")
3	mod_map(in:protection="page_execute_readwrite")
4	mod_map(out:section_offset="0")
5	mod_get_proc_address(in:function="memset")



Key Takeaways

- **TagVet explanations help at**
  - building trust in external information sources
  - unveiling inconsistencies between different approaches
  - provide concise and meaningful explanations

Key Takeaways

- **TagVet explanations help at**
  - building trust in external information sources
  - unveiling inconsistencies between different approaches
  - provide concise and meaningful explanations
- **Inherent limitation:**
  - Discrepancy between behavior-based explanations and the features used by the original tagging

Thanks for your attention!