Integrated Detection Mechanism

Leandros Maglaras & Jianmin Jiang
University of Surrey - Surrey Centre for Cyber Security (SCCS)
SCADA Cybersecurity Workshop 16th December 2014
Scada systems - Cyber attacks

Cyber-attacks can come from any part of the infrastructure:
1. **FIELD Network** as SCADA systems
2. **OPERATION Network** as Telco system or monitoring management system
3. **IT Network** as enterprise devices and services

Kinds of **cyber attacks**:
1. Denial of Service (**DoS**)
2. Accidental or malicious infections by **worms**
3. **Spoofing** attacks/**Man-In-The-Middle** attacks
4. Authentication violation
Network monitoring
Detection and classification

- Feature extraction
- Per packet – per flow analysis
- Parameter calibration
- Performance evaluation metrics (TP, TN, FP,…)
- Machine learning algorithms
  a. Naïve Bayes
  b. Clustering
  c. Markov chains
  d. Support Vector Machines

Events are analysed and patterns are detected
If patterns are known, the relationships between the data elements are identified
If the relationships are known, context of data elements are identified
If the context is known, then the meaning of the data is understood (i.e. whether the data corresponds to normal or abnormal behaviour of the system)

Threat identification by machine learning
OCSVM for SCADA systems

- **OCSVM** does not require any **signatures** of data to build the detection model.
- **OCSVM** is capable of detection both known and unknown (**novel**) attacks.
- In practice training data, taken from **SCADA** environment, could include **noise** samples - **OCSVM** detection approach is robust to noise samples.
- Algorithm **configuration** can be controlled by the user to regulate the percentage of anomalies expected.
- **OCSVM** detectors can operate **fast** enough for online detection.
- **OCSVM** is capable of handling **multiple** attributed data (many features).
IT- OCSVM: Integrated detection mechanism

- **Pre-processing** of raw input data, feed the OCSVM module
- Selection of the most appropriate **features** for training of the OCSVM
- Creation of **cluster of OCSVM** models trained on discrete datasets
- **Testing** of the traffic dataset that contain malicious attacks
- **Ensemble of Classifiers**
- **Social analysis** based on network traces
- **Fusion** of the information gathered OCSVMs
- Creation of **IDMEF** files that describe the nature of the alert, in terms of importance, the position in the system, time.
Architecture of the detection mechanism

Social analysis

Central OCSVM

Ensemble of outcomes

Additional Weight on outcomes

Aggregation of outcomes

K-means clustering

Final alerts IDMEF

Cluster of split OCSVMs
## Features

### Central OCSVM

<table>
<thead>
<tr>
<th>A/A</th>
<th>Network Data feature</th>
<th>Type of feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Packet size</td>
<td>Content based</td>
</tr>
<tr>
<td>2</td>
<td>Rate</td>
<td>Time based</td>
</tr>
<tr>
<td>3</td>
<td>Num_packets_dst</td>
<td>Time based</td>
</tr>
<tr>
<td>4</td>
<td>Num_packets_src_dst</td>
<td>Time based</td>
</tr>
<tr>
<td>5</td>
<td>Num_ARP_packets</td>
<td>Time based</td>
</tr>
</tbody>
</table>

\[
Packet_{scaled} = \frac{\text{packet size}}{\text{Max packet size}}
\]

\[
Rate_{scaled} = \frac{\text{Time difference}}{\text{Max time difference}}
\]

\[
\text{Num}\_packets\_dst = \sum_{k=1}^{10} a \times 0.1, \text{where } \begin{cases} 
  a = 1 \text{ if } \text{destination}_\text{packet}(i-k) = \text{destination}_\text{packet}(i) \\
  a = 0 \text{ if } \text{destination}_\text{packet}(i-k) \neq \text{destination}_\text{packet}(i)
\end{cases}
\]

\[
\text{Num}\_packets\_src\_dst = \sum_{k=1}^{10} a \times 0.1, \text{where } \begin{cases} 
  a = 1 \text{ if } \text{destination}_\text{packet}(i-k) = \text{destination}_\text{packet}(i) \text{ and } \text{source}_\text{packet}(i-k) = \text{source}_\text{packet}(i) \\
  a = 0 \text{ if } \text{destination}_\text{packet}(i-k) \neq \text{destination}_\text{packet}(i) \text{ or } \text{source}_\text{packet}(i-k) \neq \text{source}_\text{packet}(i)
\end{cases}
\]

\[
\text{Num}\_ARP\_packets = \sum_{k=1}^{10} a \times 0.1, \text{where } \begin{cases} 
  a = 1 \text{ if } \text{packet}_\text{protocol}(i-k) = \text{ARP} \\
  a = 0 \text{ if } \text{packet}_\text{protocol}(i-k) \neq \text{ARP}
\end{cases}
\]
Cluster of OCSVMs
Ensemble system

SCADA SYSTEM

Ensemble System

Cc – Central OCSVM
C1-n – Split OCSVM
IT-OCSVM
IT-OCSVM in operation

$$q_e(i,j) = \sum_{n=1}^{N} w_id_t(i,j)$$
Social metrics

Spearman’s correlation coefficient – based on used protocol

\[ p = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \]

The final output is a number that indicates whether there is a differentiation in the way that each source behaves during the training and the testing period.

\[ q_s(i, j) = \frac{q_e(i, j)}{p_j}, \forall q_e(i, j) \text{ with source node } j \]

Most used protocols used by a node during normal (left) and abnormal (right) operation:
- modtcp
- udp:data
- icmp:data
- arp
- tcp:FIN
- tcp:SYN
Fusion of alarms

K-means OCSVM

1. Read data
2. Extract features
3. OCSVM with default parameters
4. K-means clustering
5. $K_{\text{thresh}}$ or 1 cluster
   - Yes: Send IDMEF
   - No: Multiple OCSVMs per source

Split OCSVM

1. Read data
2. Important source?
   - No: Add source to association file
     /search source in file
   - Yes: Create training / testing set
     Create model / test accuracy
     All sources?
       - No
         - Send IDMEF
       - Yes
Fusion of alarms

1\textsuperscript{st} Stage : Aggregation :

\[ q_{aj} = \sum_{i} q_{s}(i, j), \quad q_{bj} = \sum_{i} 1, \forall q_{s}(i, j) \text{ with source node } j \]

2\textsuperscript{nd} Stage : Clustering - Categorization

\[ SSE = \sum_{k=1}^{K} \sum_{j=1}^{N_k} ||q_{aj} - \mu_k||^2 \]

- Aggregated alerts
- K-means clustering (instances)
- K-means clustering (Values)

<table>
<thead>
<tr>
<th>K-means (instances)</th>
<th>K-means (Value)</th>
<th>Final classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severe</td>
<td>Severe</td>
<td>Severe</td>
</tr>
<tr>
<td>Severe</td>
<td>Possible</td>
<td>Medium</td>
</tr>
<tr>
<td>Possible</td>
<td>Severe</td>
<td>Medium</td>
</tr>
<tr>
<td>Possible</td>
<td>Possible</td>
<td>Possible</td>
</tr>
</tbody>
</table>

Meeting – 30/10/2013
Communication - Integration

IDMEF messages produced by IT-OCSVM

IDMEF files inform about source, destination, time and classification of the event

Sources file created by IT-OCSVM
Nature of the trial

A. Network scan attack
B. ARP spoofing - MITM attack
C. DoS attack

<table>
<thead>
<tr>
<th>Frame number</th>
<th>Frame time_epoch</th>
<th>Source</th>
<th>Destination</th>
<th>Frame protocols</th>
<th>Frame.len</th>
<th>Col.info</th>
</tr>
</thead>
</table>

Operations Level

Field Level

Cockpit CI
Transformed datasets

Central OCSVM

| 1 1: 0.1866122987018105 2: 0.0517955801049724 3: 0.0 4: 0.0 5: 0.1 |
| 1 1: 4.6638459674666064-4 2: 0.05041436464088398 3: 0.1 4: 0.1 5: 0.1 |
| 1 1: 5.011205741489403E-4 2: 0.04765193370165746 3: 0.2 4: 0.2 5: 0.1 |
| 1 1: 0.07182604724705569 2: 0.041436466408839779 3: 0.0 4: 0.0 5: 0.1 |
| 1 1: 0.0067294657508171 2: 0.04143646408839779 3: 0.0 4: 0.0 5: 0.1 |
| 1 1: 0.04598406510677064 2: 0.04143646408839779 3: 0.2 4: 0.2 5: 0.1 |
| 1 1: 0.03813842506418002 2: 0.04143646408839779 3: 0.0 4: 0.0 5: 0.2 |
| 1 1: 0.14369331420862086 2: 0.04558011049723757 3: 0.2 4: 0.2 5: 0.2 |
| 1 1: 0.02196640269067456 2: 0.04419889502762431 3: 0.1 4: 0.1 5: 0.2 |
| 1 1: 0.03862565135253925 2: 0.04558011049723757 3: 0.2 4: 0.2 5: 0.2 |
| 1 1: 0.014140214648189736 2: 0.04419889502762431 3: 0.0 4: 0.0 5: 0.2 |

Split OCSVM

| 1 1: -4.618338530824944 2: 0.6122448979591837 3: 0.0 |
| 1 1: -4.747598641215887 2: 0.6122448979591837 3: 0.0 |
| 1 1: -4.22820822378137 2: 0.6122448979591837 3: 0.0 |
| 1 1: -4.47037156022453 2: 0.6122448979591837 3: 0.0 |
| 1 1: -4.282215789767468 2: 0.6122448979591837 3: 0.0 |
| 1 1: -4.585233406666963 2: 0.6122448979591837 3: 0.0 |
| 1 1: -4.539874534291136 2: 0.6122448979591837 3: 0.0 |
| 1 1: -4.565755053271114 2: 0.6122448979591837 3: 0.0 |
| 1 1: -4.703581812231512 2: 0.6122448979591837 3: 0.0 |
| 1 1: -4.22820822378137 2: 0.6122448979591837 3: 0.0 |

**Testing data** consists of normal data and attack data and the composition of the data sets are as follows:

- **Testing set-A':** 1- 5000 Normal data records
- **Testing set-B':** 5000- 10000- Normal data records + Arp spoofing attack + Network scan attack
- **Testing set-C':** 10000 – 25000 Normal data records + Dos attack + Network scan attack
- **Testing set-D':** 25000- 41000 Normal data records + MITM attack
Rate of packets

DOS attack
ARP spoofing (overall – split datasets)
Precision - accuracy

Detection accuracy = \( \frac{\text{True positives} + \text{True negatives}}{\text{Sample size}} \times 100\% \)

False alarm rate = \( \frac{\text{False positives}}{\text{True negatives} + \text{False positives}} \times 100\% \)

<table>
<thead>
<tr>
<th>Testing Data set</th>
<th>DA</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>98.81%</td>
<td>1.18%</td>
</tr>
<tr>
<td>B</td>
<td>94.6%</td>
<td>3.25%</td>
</tr>
<tr>
<td>C</td>
<td>95.20%</td>
<td>1.51%</td>
</tr>
<tr>
<td>D</td>
<td>96.37%</td>
<td>2.3%</td>
</tr>
<tr>
<td>FULL</td>
<td>96.3%</td>
<td>2.5%</td>
</tr>
</tbody>
</table>
Impact of the fusion mechanism

Aggregated alarms produced by IT-OCSVM are significantly decreased compared to the initial alarms.

IT-OCSVM categorizes aggregated alarms.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Initial alarms</th>
<th>Aggregated alarms</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>129</td>
<td>16</td>
</tr>
<tr>
<td>B</td>
<td>658</td>
<td>21</td>
</tr>
<tr>
<td>C</td>
<td>9273</td>
<td>18</td>
</tr>
<tr>
<td>D</td>
<td>203</td>
<td>16</td>
</tr>
<tr>
<td>All</td>
<td>10507</td>
<td>22</td>
</tr>
</tbody>
</table>
The number of created OCSVMs affects the performance of the detection mechanism.

The system, under the default Configuration can be adopted in soft real-time applications.
• Integrated detection mechanism
• Based on OCSVM, Social network analysis
• Automatic Creation of a cluster of split OCSVMs
• Ensemble, aggregation, k-means clustering

Conclusions – Discussion


**Publications**


Any question ?
Thank you for your attention
Cybersecurity on SCADA: risk prediction, analysis and reaction tools for Critical Infrastructures

Any question?
Thank you for your attention