Characterize and Quantify Cyber Attack Pattern by Granger Causality

Van Trieu
Richard Garcia, Shouhuai Xu, Yusheng Feng

University of Texas at San Antonio
Motivation:
Cybersecurity Statistics[1]

- 3 billion Yahoo accounts were hacked (2016)
- Damage related to cybercrime is estimated to hit $6 trillion (2021)

Enhance the performance of alert correlation and minimize damage from attacks is necessary.

Goal:
Study the phenomenon of the time series data in IPv4 address space utilization based on Granger Causality.

Apply the learned phenomenon into Long Short Term Memory model to improve prediction in dynamic time series data.
Framework

Data: /8 subnet
/16 subnet

X Granger Cause Y
Linear Regression
Sum Square Error
- Dependent variable and lagged values
- Independent variables and lagged values

Hypothesis F-test

- Network of subnets
- Communities of subnets (Clusters)

Characterize and identify the future state

7/10/2019
Granger connectivity between subnets

Network of /8 subnet (96 days)

Granger Causal results:
• 199/207 nodes
• 2592 edges

Seven different communities:
**Insight 1:** The higher the number of degree range in nodes, the similar in pattern of their time series data.
Insight 2: The nodes with highest degree range have the most popular pattern in the data set.
Insight 3: VAR of nodes (VAR between nodes), which have same source and similar p value, can give a better prediction compare to the original VAR (VAR between source and node)
Statistical Comparison

### MAPE and RMSE

<table>
<thead>
<tr>
<th></th>
<th>S66</th>
<th>S208</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MAPE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC(S123)</td>
<td>69.88831</td>
<td>52.74795</td>
</tr>
<tr>
<td>GC</td>
<td>38.54982</td>
<td>31.48487</td>
</tr>
<tr>
<td><strong>RMSE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC(S123)</td>
<td>43.44967</td>
<td>40.06773</td>
</tr>
<tr>
<td>GC</td>
<td>26.960052</td>
<td>19.413949</td>
</tr>
</tbody>
</table>
LSTM (Encoder-Decoder)

Encoder’s output vector:

\[ o_{at} = \beta_{ao} + \beta_{a1}x_{at-1} + \alpha_{a1}x_{bt-1} + \beta_{a2}x_{at-2} + \alpha_{a2}x_{bt-2} \]

Encoder-Decoder LSTM summary model