**Detecting BGP Anomalies Using Machine Learning Techniques**

**Student:** Zhida Li. **Instructor:** Steve Whitmore.

**Simon Fraser University, Vancouver, British Columbia, Canada**

**BORDER GATEWAY PROTOCOL**
- Border Gateway Protocol (BGP) is an interdomain routing protocol used in networks consisting of a large number of Autonomous Systems (ASs).
- Propagation of the BGP routing information is susceptible to misconfigurations, power outages, malicious attacks, and worms.
- The main function of BGP is to select the best routes between ASes based on routing algorithms and network policies.
- Determining the anomalies and their causes is useful for assessing loss of data and connectivity.
- BGP anomaly detection system design relies on machine learning techniques.
- We use well-known classifiers and exploit their ability to reliably detect network anomalies in datasets of known BGP network anomalies.

**BGP DATASETS**
- Analyzed Internet routing data are acquired from two projects that provide valuable information to networking research:
  - Routing Information Service (RIS) project initiated in 2001 by the Réseaux IP Européens (RIPE) Network Coordination Centre (NCC)
  - These projects collect and store routing data that provide a unique view of the Internet topology.
- Anomalous events considered in this project:
  - Slammer
  - Nimda
  - Code Red I

**APPROACHES**
- In this project we attempt to detect various BGP anomalies by applying:
  - Support Vector Machine (SVM) models, and
  - A Long Short-Term Memory (LSTM) recurrent neural networks.
  - We use feature scoring algorithms for SVM to select the most relevant features:
    - minimum Redundancy Maximum Relevance (mRMR)
    - Mutual Information Difference (MID)
    - Mutual Information Quotient (MIQ)
    - Mutual Information Base (MIBASE).

<table>
<thead>
<tr>
<th>Class</th>
<th>Date</th>
<th>Duration (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slammer</td>
<td>Anomaly</td>
<td>January 25, 2003</td>
</tr>
<tr>
<td>Nimda</td>
<td>Anomaly</td>
<td>September 18, 2001</td>
</tr>
<tr>
<td>Code Red I</td>
<td>Anomaly</td>
<td>July 19, 2001</td>
</tr>
</tbody>
</table>

**EXPERIMENTAL PROCEDURE**
- **Step 1:** Train and test the three SVM and LSTM models using 37 features.
- **Step 2:** Select the 10 most relevant features using the three feature selection algorithms: MID, MIQ, and MIBASE. Train and test the three SVM models using datasets with and without anomalies. Skip this Step for generating the LSTM models.
- **Step 3:** Evaluate the SVM and LSTM models using the accuracy and F-score measures.
- **Step 4:** Tune the SVM and LSTM model parameters to achieve the best performance.

**PERFORMANCE MEASURES**
- F-score = 2 × (precision × sensitivity) / (precision + sensitivity)
- Sensitivity: ratio of identified anomalies (TP) and all labeled anomalies (true).
- Precision: ratio of identified anomalies (TP) and all data points identified as anomalous.
- TP: number of anomalous training data points classified as anomaly
- FP: number of regular training data points classified as anomaly
- FN: number of anomalous training data points classified as regular
- TN: number of regular training data points classified as regular

**PERFORMANCE EVALUATION**

<table>
<thead>
<tr>
<th>Unbalanced Datasets</th>
<th>SVM 1</th>
<th>SVM 2</th>
<th>LSTM 1</th>
<th>LSTM 2</th>
<th>LSTM 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>75.79%</td>
<td>77.26%</td>
<td>77.98%</td>
<td>63.98%</td>
<td>63.98%</td>
</tr>
<tr>
<td>Precision</td>
<td>75.79%</td>
<td>77.26%</td>
<td>77.98%</td>
<td>63.98%</td>
<td>63.98%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>75.79%</td>
<td>77.26%</td>
<td>77.98%</td>
<td>63.98%</td>
<td>63.98%</td>
</tr>
<tr>
<td>F-score</td>
<td>75.79%</td>
<td>77.26%</td>
<td>77.98%</td>
<td>63.98%</td>
<td>63.98%</td>
</tr>
</tbody>
</table>

**REFERENCES**

**CONCLUSION**
- Feature selection and classification algorithms were used to detect BGP anomalies.
- The SVM 2 models based on the combination of the Slammer and Code Red I training datasets achieve better accuracy and F-score than results reported.
- The SVM classifier achieved the highest F-score using balanced datasets. In case of the unbalanced datasets, the accuracy is higher due to the large number of the regular testing data.
- Using the SVM classifier may be a feasible approach for detecting BGP anomalies in communication networks.

**WRITING FOR PUBLICATION**

**ENSC 803**

**August 2016, Burnaby, Canada**