Benchmarking bioinspired machine learning algorithms with CSE-CIC-IDS2018 network intrusions dataset

CIIC, School of Technology and Management - Polytechnic of Leiria; CRACS – INESC-TEC - Portugal

Paulo Ferreira, Mário Antunes

2180047@my.ipleiria.pt, mario.antunes@ipleiria.pt

Motivation

- Network IDS is a **widely studied** and **very popular** research field.
- Biology mimicking is a good spot for network IDS
- Datasets used are mostly (too) artificial
- CSE-CIC-IDS2018 dataset is promising to evaluate ML methods
- Benchmark based on popularity and availability of WEKA

Bioinspired Algorithms

- Mimic biological systems applied to computer science problems
- **Artificial Neural Networks**
  - Multi-Layer Perceptron (MLP)
  - Learning Vector Quantization(LVQ)
- **Artificial Immune Systems**
  - **CLON**al Selection **ALG**orithm (CLONALG)

CSE-CIC-IDS2018 Dataset

- Labeled dataset
- Attack types – DoS/DDoS, Brute-force, SQL Injection, Infiltration, Botnet
- Victims: 420 PC+30 servers; 50 attackers
- Period: 14/02/2018 to 02/03/2018
- 80 attributes (network flows)

Day	Time		Type of attack	Tool
	Start	End		
16/02/2018	10:12	11:08	DoS	SlowHTTPTest
	13:45	14:19	DoS	Hulk
21/02/2018	10:09	10:43	DDoS	LOIC-UDP
	14:05	15:05	DDoS	HOIC

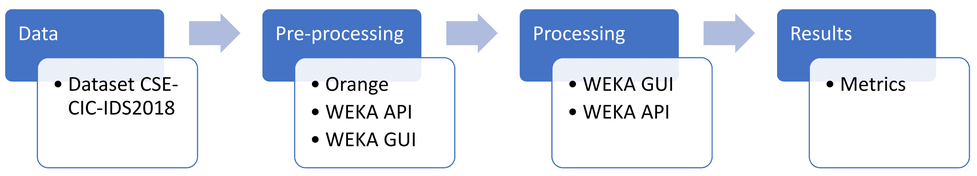
Goals and contributions

- To evaluate CSE-CIC-IDS2018 dataset
- To benchmark a set of supervised bioinspired algorithms
- To compare individual results with an ensemble algorithm

Java CLI application to benchmark IDS, using WEKA and CSE-CIC IDS2018

`https://github.com/paulo-ferreira-mcif/benchmarkids`

Methodology



- **Pre-processing**
  - To eliminate unnecessary attributes
  - To deal with missing values
  - To normalize values
  - To apply class reduction
- **Processing**
  - Tests run 10 times for each dataset

Test Scenarios

Scenario	Training		Test	
	Day	Traffic	Day	Traffic
1	16/02/2018	Normal + Attack1	16/02/2018	Normal + Attack2
2	16/02/2018	Normal + Attack1	21/02/2018	Normal + Attack1
3	16/02/2018	Normal + Attack2	21/02/2018	Normal + Attack2
4	16/02/2018	Normal + Attacks	21/02/2018	Normal + Attacks

Results

- **Scenarios 1 and 2** – not promising
- **Scenario 3**

Algorithm	TPR	TNR	FPR	FNR	Precision	Recall	Accuracy	F1
MLP	1,0000	0,9998	0,0002	0,0000	0,9999	1,0000	0,9999	0,9999
Ensemble	1,0000	0,0026	0,9974	0,0000	0,6559	1,0000	0,6562	0,7922
CLONALG	1,0000	0,0026	0,9974	0,0000	0,6559	1,0000	0,6562	0,7922
LVQ	1,0000	0,0004	0,9996	0,0000	0,6554	1,0000	0,6554	0,7918

- **Scenario 4**

Algorithm	TPR	TNR	FPR	FNR	Precision	Recall	Accuracy	F1
MLP	0,8977	0,9996	0,0004	0,1023	0,9284	0,8977	0,9327	0,8987
Ensemble	0,9992	0,0032	0,9968	0,0008	0,6564	0,9992	0,6565	0,7923
LVQ	1,0000	0,0008	0,9992	0,0000	0,6560	1,0000	0,6561	0,7923
CLONALG	0,9992	0,0030	0,9970	0,0008	0,6564	0,9992	0,6564	0,7923

Conclusions

- Rich dataset with few entries in the bibliography
- Scenarios 1 and 2: Results not promising
- Scenarios 3 and 4 : ANN (MLP) shows better performance
- Ensemble: Not significative influence in final results with these models.

RECPAD'20

UNIVERSIDADE DE ÉVORA