Verificarea Rețelelor Neuronale Folosind Alpha-Beta CROWN și NeuralSAT pentru benchmark-ul ACAS XU

Bordea Liviu-Valentin, Ciurdea Roberta-Carla, Burcea Adina Mădălina
Coordonator: Conf. Dr. Erașcu Mădălina

Universitatea de Vest
din Timişoara

Cuprins

- Motivația
- Descrierea Dataset-ului
- Instalarea și Rularea Tool-urilor
- Rezultate
- Concluzii

Motivația

- Într-un sistem critic pentru securitate precum ACAS Xu, un atacator poate exploata cu ușurință o situație limită gestionată incorect pentru a cauza daune semnificative, punând în pericol mii de vieți.
- Metodele existente pentru testarea rețelelor neurale în fața situațiilor limită se concentrează pe identificarea exemplelor adversare.

Descrierea Problemei:

- Reproducerea și înțelegerea rezultatului din Competiția de Verificare a Rețelelor Neurale (VNN-COMP) din 2023 pentru benchmark-ul ACAS Xu s-au realizat folosind Alpha-Beta-CROWN și NeuralSAT

Descrierea Dataset-ului

ACAS XU(Automated Collision Avoidance System)

- este un sistem de evitare a coliziunilor proiectat pentru aeronave fără pilot
- folosește rețele neuronale pentru a anticipa cele mai bune acțiuni optime în funcție de locația și viteza avioanelor atacatoare din apropiere

Structura:

- **fisiere onnx** - 45 rețele neuronale;
- **fisiere vnnlib** - 10 proprietăți;
- **instances.csv** - conține lista completă a instanțelor de benchmark, câte una pe linie: onnx_file, vnn_lib_file, timeout_secs.
- **generate.py** - creeaza fișierele .vnnlib și instances.csv;

Instalarea și rularea Tool-urilor: Alpha-Beta CROWN

Instalare:

- Am folosit miniconda3.
- Am clonat de pe GitHub folderele "alpha-beta-CROWN" și "auto_LIPRA".
- Am creat environmentul și am activat alpha-beta-crown.
- Am instalat pachetele necesare pentru a putea rula benchmark-ul.

Rulare:

- primele tentative de rulare au rezultat 2 erori diferite:
  - "killed": din cauza memoriei insuficiente.
  - "Recovery": blue recovery screen.
- Rularea a durat 5 ore.

Instalarea și rularea Tool-urilor: NeuralSAT

Instalare:

- Am folosit miniconda3.
- Am clonat de pe github folderul "neuralsat".
- Am creat environmentul și am activat "neuralsat".

Rulare:

- Am folosit un script python ca să automatizăm rularea individuală.
- Rularea a durat 1 ora și 17 minute.

Rezultate

Score = 10*Verified + 10*Falsified - 150*Penalty

#	Tool	Verified	Falsified	Fastest	Penalty	Score	Percent
1	α-β-CROWN	112	47	0	27	-2460	0%
2	neuralSAT	120	46	0	19	-1190	0%

Table: Benchmark 2024-acasxu

- Timp mediu α-β-CROWN: 88.362 sec
- Timp mediu neuralSAT: 10.10 sec
- Număr de rezultate "Timeout": 27 α-β-CROWN; 20 neuralSAT

Rezultate: Comparații

Competition	2023		2024	
Tools	alpha-beta-CROWN	neuralSAT	alpha-beta-CROWN	neuralSAT
Verified	139	138	112	120
Falsified	47	46	47	46
Fastest	0	0	0	0
Penalty	0	0	27	19
Score	1860	1840	-2460	-1190
Percent	100%	98.9%	0%	0%

Table: Comparație Rezultate

Concluzii

- Instalarea cu succes a ambelor tool-uri
- Verificarea benchmark-ului folosind cele două tool-uri
- Rezultate asemanatoare competiției trecute
- Penalizări datorate incapacității dispozitivelor

Mulțumim!