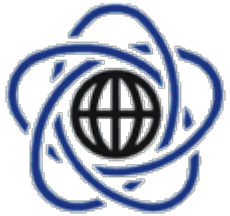


FACULTY OF MATHEMATICS AND INFORMATION SCIENCE
WARSAW UNIVERSITY OF TECHNOLOGY

LEARNING ATTACKER'S BOUNDED RATIONALITY MODEL IN SECURITY GAMES

Adam Żychowski and Jacek Mańdziuk

{a.zychowski, j.mandziuk}@mini.pw.edu.pl

December 2021

CONSIDERED PROBLEM

Stackelberg Security Games playing against not perfectly rational opponent

---

CONTRIBUTION

- the first successful neural network application to the defender's strategy estimation in Security Games
- end-to-end neuroevolutionary system (NESG) for finding high quality leader's strategies
- generic system which does not use any assumption about the opponent’s bounded rationality model or knowledge about his/her payoff distribution

---

RESULTS

experimental results in the cybersecurity domain outperform state-of-the-art methods in terms of computation time and quality of results

PROBLEM DEFINITION

- **Security Games** – a game model with **Stackelberg equilibrium** widely applicable in many real-world scenarios (e.g. surveillance, homeland security, poaching prevention, smuggling detection, cybersecurity)
- played by two non-symetrical players: Defender, Attacker
- the Defender (D) commits to a certain (mixed) strategy first, then the Attacker (A) chooses their strategy

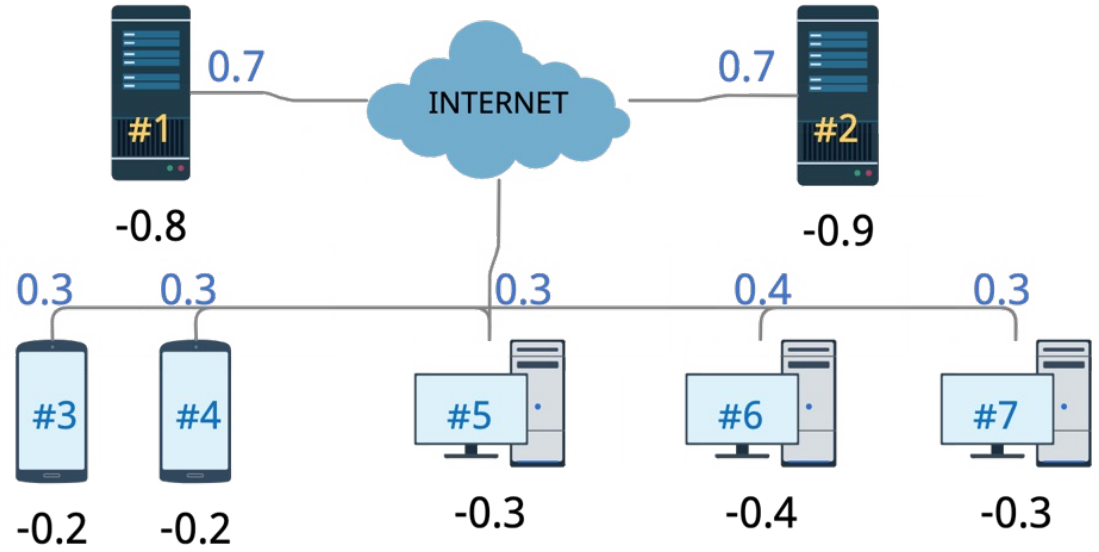
$$BR(\pi^D) = \arg\max_{\pi^A \in \Pi^A} U^A(\pi^D, \pi^A)$$

- goal: **maximize the Defender's payoff**

$$\arg\max_{\pi^D \in \Pi^D} U^A(\pi^D, BR(\pi^D))$$

- finding optimal Defender's mixed strategy is an NP-hard problem
- **perfect rationality** of both players is assumed

CYBERSECURITY SCENARIO



deep packet inspections

- *n* targets, *m* steps
- the detection system (the Defender) chooses a subset of hosts (targets) and inspects packets sent to them in order to detect a potential attack (malicious packets)
- the Defender has no knowledge about potential invaders, their goals, preferences or capabilities

BOUNDED RATIONALITY

Bounded rationality – taking non-optimal actions due to limitations of decision-makers (e.g. cognitive bias, partial knowledge, limited resources)

- introduced in 1957 by Herbert Simon, gained lots of intrest in 1990'
- bounded rationality ≠ irrationality

- there is no widely-agreeable BR formulation ⟶ several popular BR models are proposed

BOUNDED RATIONALITY MODELS

- **Anchoring theory** – humans have a tendency to flatten probabilities. Options with low probabilities are overestimated while those with high probabilities are underestimated

$$q'(i) = (1 - \alpha)q(i) + \alpha/M$$

- **Quantal Response** - humans choose a decision stochastically, the higher the payoff, the higher the chance for a decision to being chosen

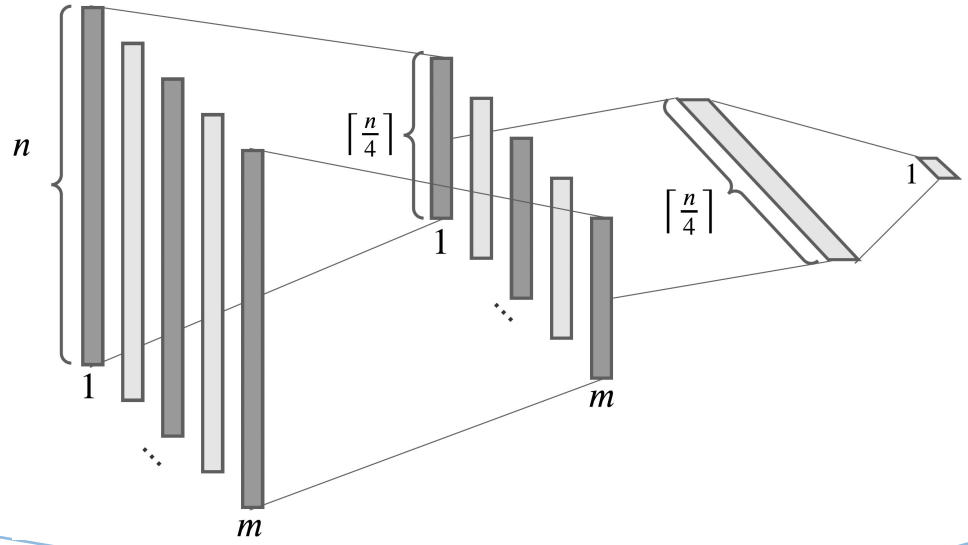
- **Prospect theory** - loss aversion and risk aversion are not symmetric. Instead of maximizing the expected payoff humans tend to maximize the *prospect* which describes people's perception of the probability and the outcome

MOTIVATION

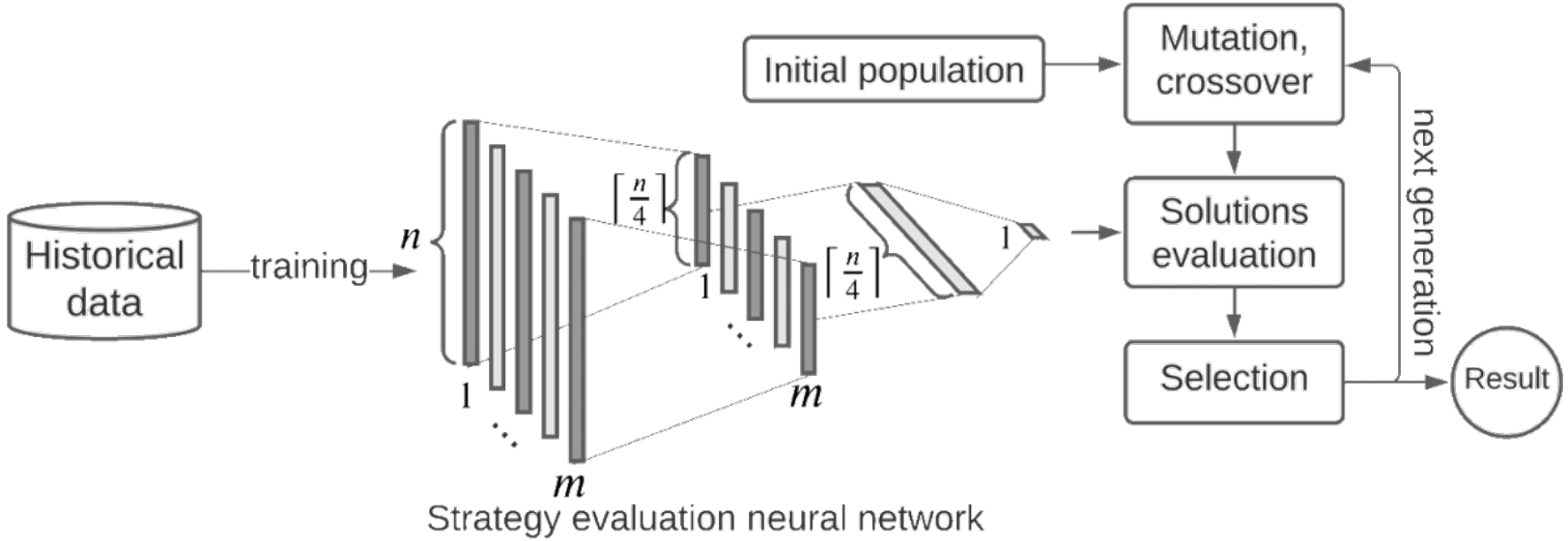
hitherto approaches	problem	our approach
assume particular BR model and compute optimal defender's strategy according to that model	• theoretical model does not reflect real-life scenario • in practice there is no knowledge about opponent's cognitive capabilities	no assumption about any particular BR model → the **model is learnt** from historical data

STRATEGY EVALUATION NEURAL NETWORK

- strategy evaluation neural network (SENN) for evaluating defender's strategy based on historical data (previous gameplays)
- *n*m* inputs: target's coverage - a probability that at least one defender's unit is allocated to the target *t* in each time step
- output: defender's payoff when playing a given strategy presented in the input



NEUROEVOLUTIONARY APPROACH



- SENN is incorporated into Evolutionary Approach for Security Games (EASG)
  [A. Żychowski, J. Mańdziuk, *Evolution of Strategies in Sequential Security Games,* 20th International Conference on Autonomous Agents and Multiagent Systems (**AAMAS 2021**), London]
- no prior knowledge about the attacker's target preferences or payoffs distribution
- suitable for bounded rationality scenario (no particular assumptions about the model)

NESG EVALUATION

- 90 randomly generated game instances inspired by real-world cybersecurity scenario
- number of time steps: 1, 2, or 3
- number of targets: 4, 8, 16, 32, 64, 128
- between 1/4 and 3/4 of all targets can be effectively protected
- 3 popular bounded rationality models (Anchoring Theory, Quantal Response, Prospect Theory)

- 5000 training examples per game

SENN ERROR ON TEST DATA

	Anchoring Theory			Quantal Response			Prospect Theory		
targets	1 step	2 steps	3 steps	1 step	2 steps	3 steps	1 step	2 steps	3 steps
4	0.006	0.006	0.006	0.004	0.005	0.005	0.010	0.010	0.010
8	0.011	0.012	0.014	0.008	0.008	0.008	0.018	0.019	0.020
16	0.024	0.026	0.028	0.019	0.021	0.022	0.046	0.049	0.051
32	0.043	0.045	0.048	0.031	0.033	0.035	0.075	0.080	0.084
64	0.080	0.081	0.086	0.064	0.065	0.069	0.132	0.142	0.145
128	0.119	0.125	0.131	0.104	0.110	0.121	0.232	0.241	0.251

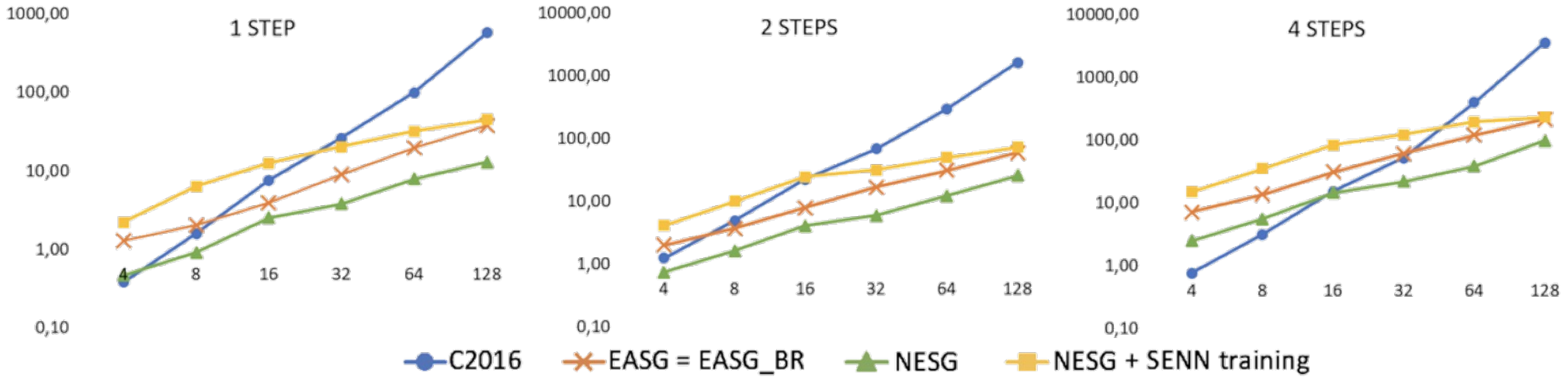
- a neural network can accurately approximate defender's payoff despite having no direct knowledge about game utilities
- network error increases with the number of targets and/or steps
- better accuracy for Anchoring Theory and Quantal Response than for Prospect Theory

NESG PAYOFF RESULTS

1 step	Anchoring Theory				Quantal Response				Prospect Theory			
targets	C2016	EASG	EASG_AT	NESG	C2016	EASG	EASG_QR	NESG	C2016	EASG	EASG_PT	NESG
4	-0.470	-0.472	-0.468	-0.469	-0.406	-0.408	-0.404	-0.405	-0.419	-0.420	-0.417	-0.418
8	-0.456	-0.457	-0.440	-0.440	-0.418	-0.422	-0.386	-0.388	-0.422	-0.423	-0.407	-0.407
16	-0.387	-0.391	-0.371	-0.371	-0.377	-0.378	-0.336	-0.338	-0.329	-0.335	-0.315	-0.318
32	-0.411	-0.412	-0.393	-0.397	-0.428	-0.429	-0.390	-0.394	-0.397	-0.404	-0.367	-0.370
64	-0.579	-0.586	-0.567	-0.568	-0.582	-0.584	-0.536	-0.537	-0.560	-0.564	-0.483	-0.486
128	-0.397	-0.405	-0.369	-0.372	-0.578	-0.578	-0.526	-0.529	-0.462	-0.463	-0.345	-0.347
2 steps	Anchoring Theory				Quantal Response				Prospect Theory			
targets	C2016	EASG	EASG_AT	NESG	C2016	EASG	EASG_QR	NESG	C2016	EASG	EASG_PT	NESG
4	-0.566	-0.566	-0.563	-0.564	-0.540	-0.541	-0.534	-0.535	-0.548	-0.549	-0.547	-0.547
8	-0.568	-0.572	-0.553	-0.555	-0.526	-0.528	-0.510	-0.512	-0.556	-0.556	-0.517	-0.518
16	-0.327	-0.331	-0.314	-0.317	-0.326	-0.331	-0.301	-0.302	-0.326	-0.331	-0.291	-0.294
32	-0.499	-0.500	-0.475	-0.479	-0.487	-0.487	-0.435	-0.435	-0.501	-0.502	-0.454	-0.457
64	-0.457	-0.463	-0.427	-0.427	-0.421	-0.424	-0.403	-0.408	-0.466	-0.471	-0.407	-0.410
128	-0.607	-0.614	-0.563	-0.567	-0.601	-0.604	-0.540	-0.544	-0.593	-0.595	-0.566	-0.571
4 steps	Anchoring Theory				Quantal Response				Prospect Theory			
targets	C2016	EASG	EASG_AT	NESG	C2016	EASG	EASG_QR	NESG	C2016	EASG	EASG_PT	NESG
4	-0.479	-0.481	-0.478	-0.479	-0.487	-0.489	-0.485	-0.486	-0.511	-0.512	-0.508	-0.510
8	-0.497	-0.500	-0.466	-0.467	-0.509	-0.513	-0.455	-0.456	-0.517	-0.519	-0.496	-0.499
16	-0.545	-0.547	-0.525	-0.525	-0.531	-0.534	-0.502	-0.503	-0.570	-0.574	-0.535	-0.538
32	-0.478	-0.484	-0.460	-0.464	-0.500	-0.505	-0.468	-0.470	-0.525	-0.531	-0.492	-0.496
64	-0.563	-0.568	-0.547	-0.551	-0.587	-0.593	-0.553	-0.555	-0.600	-0.600	-0.561	-0.563
128	-0.531	-0.536	-0.493	-0.497	-0.545	-0.549	-0.503	-0.505	-0.553	-0.555	-0.512	-0.512

NESG obtains better results than the literature methods with no BR consideration (C2016, EASG) and close to the method that is aware of the exact BR model (EASG_XX)

NESG TIME SCALABILITY



- near linear time scalability
- NESG outperforms baseline version of evolutionary algorithm (EASG) thanks to strategy evaluation procedure optimization

SUMMARY

- a novel method (NESG) for calculating defender's payoff in Stackelberg Security Games that uses strategy evaluation neural network is proposed
- the setting reflects real-world scenario: no explicit knowledge about the opponent's payoff distribution or bounded rationality model is available; only historical data (results of previous games) is available
- NESG does not need to assume perfect rationality of the attacker and is able to infer the actual attacker's cognitive decision model through learning
- high quality results with low computation cost (time scalability)

Thank you