Optimal Causal Cyber Defence Agents via Simulation

Alex Andrew[1], Sam Spillard[2], Joshua Collyer[1], Neil Dhir[2]

Outcomes

- We apply recent causal optimization algorithms (namely Causal Bayesian Optimisation (CBO) and Dynamic Causal Bayesian Optimisation (DCBO)) in a cybersecurity setting by simulating a system in a novel cyber security simulator.
- We propose that CBO & DCBO can act as a Blue Agent when provided with a view of a simulated network and a causal model of how a Red Agent spreads within that network.
- We submit the PISCHAT directed acyclic graph as that causal model, describing how a Red Agent compromise spreads within a computer network.
- We demonstrate a complete cyber-simulation system and use it to generate observational data for these algorithms.
- We show that these algorithms converge optimally on the intervention that best prevents cyber intrusions propagating through a network, performing more efficiently than traditional Bayesian Optimisation.
- Our results demonstrate the utility of using causal models in a cybersecurity setting, where agents are often required to take optimal sequential decisions.

Setup & Numeric Results

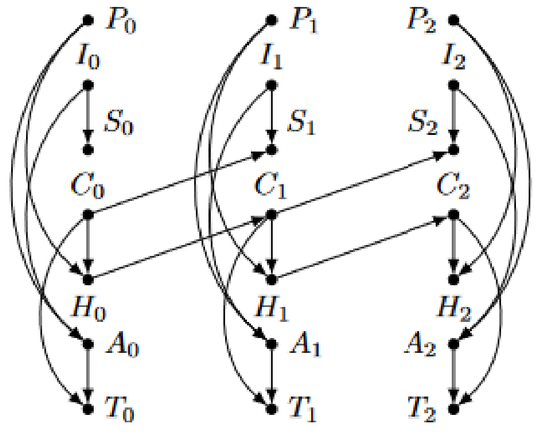

Figure 2: Causal diagram $\mathcal{G}$ shown for the first three time-steps. The target variable is T with manipulative variables given by $\mathbf{X}_t = \{P_t, I_t\}$. The rest are non-manipulative variables. The within slice topology repeats for 25 time-steps in our experimental setup and the variable connectivity repeats as shown for the same length of time. For variable descriptions see table 1.

- We use the YAWNING TITAN simulator to generate observational data of a computer network.
- We allow BO, CBO and DCBO to intervene on the manipulative variables, to minimize the total cost in the final 3 time steps of the simulation.
- Causal model improves convergence in comparison to Bayesian Optimisation:

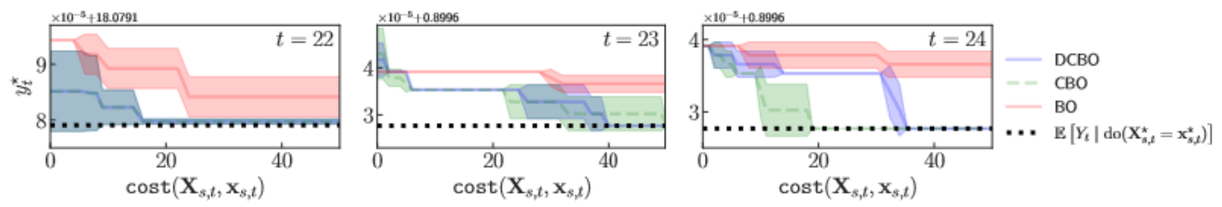

Figure 3: Experimental results of different optimisation methods applied to YT data, showing convergence of DCBO and competing methods (CBO and BO) across five replicates. The black dotted line shows the optimal response value $\{y_t^* \mid t = 22, 23, 24\}$. Shaded areas are $\pm$ one standard deviation. The x-axis shows the total *cumulative cost* of the intervention set over 50 trials - effectively how many times the optimisation had to probe the underlying function in order to build up an approximation of the causal relationship.

Future Work

- Online agent evaluation, allowing the causal agent to make decisions in real-time
- Comparison with traditional reinforcement learning approaches to this problem
- Deploy causal agents only when necessary, using optimal stopping

Contact

{ndhir, sam.s}@turing.ac.uk

[1]Dstl

[2]The Alan Turing Institute