

---

# SCALABLE EDGE BLOCKING ALGORITHMS FOR DEFENDING ACTIVE DIRECTORY STYLE ATTACK GRAPHS

---

A PREPRINT

**Mingyu Guo, Max Ward, Aneta Neumann, Frank Neumann, Hung Nguyen**

1. School of Computer Science, University of Adelaide, Australia  
{mingyu.guo, aneta.neumann, frank.neumann, hung.nguyen}@adelaide.edu.au

2. School of Physics, Maths and Computing, Computer Science and Software Engineering  
University of Western Australia, Australia

and

Department of Molecular and Cellular Biology  
Harvard University, Cambridge, Massachusetts, USA  
{max.ward}@uwa.edu.au

November 23, 2022

## ABSTRACT

Active Directory (AD) is the default security management system for Windows domain networks. An AD environment naturally describes an attack graph where nodes represent computers/accounts/security groups, and edges represent existing accesses/known exploits that allow the attacker to gain access from one node to another. Motivated by practical AD use cases, we study a Stackelberg game between one attacker and one defender. There are multiple entry nodes for the attacker to choose from and there is a single target (Domain Admin). Every edge has a failure rate. The attacker chooses the attack path with the maximum success rate. The defender can block a limited number of edges (i.e., revoke accesses) from a set of blockable edges, limited by budget. The defender’s aim is to minimize the attacker’s success rate.

We exploit the tree-likeness of practical AD graphs to design scalable algorithms. We propose two novel methods that combine theoretical fixed parameter analysis and practical optimisation techniques.

For graphs with small tree widths, we propose a tree decomposition based dynamic program. We then propose a general method for converting tree decomposition based dynamic programs to reinforcement learning environments, which leads to an anytime algorithm that scales better, but loses the optimality guarantee.

For graphs with small numbers of non-splitting paths (a parameter we invent specifically for AD graphs), we propose a kernelization technique that significantly downsizes the model, which is then solved via mixed-integer programming.

Experimentally, our algorithms scale to handle synthetic AD graphs with tens of thousands of nodes.

## 1 Introduction

Active Directory (AD) is the *default* system for managing access and security in Windows domain networks. Given its prevalence among large and small organisations worldwide, Active Directory has become a major target by cyber attackers.<sup>1</sup> An AD environment naturally describes a *cyber attack graph* — a conceptual model for describing the

---

<sup>1</sup>Enterprise Management Associates EMA [2021] found that 50% of organizations surveyed had experienced an AD attack since 2019.

causal relationship of cyber events. In an AD graph, the nodes are computers/user accounts/security groups. An edge from node A to node B represents that an attacker can gain access from A to B via an existing access or a known exploit. Unlike many attack graph models that are of theoretical interest only,<sup>2</sup> AD attack graphs are actively being used by real attackers and IT admins. Several software tools (including both open source and commercial software) have been developed for scanning, visualizing and analyzing AD graphs. Among them, one prominent tool is called BLOODHOUND, which models the *identity snowball attack*. In such an attack, the attacker starts from a low-privilege account, which is called the attacker’s entry node (i.e., obtained via *phishing* emails). The attacker then travels from one node to another, where the end goal is to reach the highest-privilege account called the DOMAIN ADMIN (DA).

Given an entry node, BLOODHOUND generates a *shortest attack path* to DA, where a path’s distance is equal to the number of hops (fewer hops implies less chance of failure/being detected). Before the invention of BLOODHOUND, attackers used personal experience and heuristics to explore the attack graph, hoping to reach DA by chance. BLOODHOUND makes it easier to attack Active Directory.

Besides the attackers, defenders also study Active Directory attack graphs. The original paper that motivated BLOODHOUND Dunagan et al. [2009] proposed a heuristic for blocking edges of Active Directory attack graphs. The goal is to cut the attack graph into multiple disconnected regions, which would prevent the attacker from reaching DA. In an AD environment, edge blocking is achieved by revoking accesses or introducing monitoring. *Not all edges are blockable*. Some accesses are required for the organisation’s normal operations. *Blocking is also costly* (i.e., auditing may be needed before blocking an edge).

We study how to *optimally* block a limited number of edges in order to minimize a strategic attacker’s success rate (chance of success) for reaching DA. In our model, we assume that different edges have different *failure rates*. The defender can block a *blockable* edge to increase the failure rate of that edge (from its original failure rate) to 100%. The defender can block at most  $b$  edges, where  $b$  is the defensive budget. For example, if it takes 1 hour for an IT admin to block an edge (auditing, reporting, implementation) and one eight-hour day is dedicated to AD cleanup, then  $b = 8$ .

We study both *pure* and *mixed* strategy blocking. A pure strategy blocks  $b$  edges deterministically. A mixed strategy specifies multiple sets of  $b$  edges, and a distribution over the sets. We follow the standard Stackelberg game model by assuming that the attacker can observe the defender’s strategy and play a best response. For mixed strategy blocking, the attacker can only observe the probabilities, not the actual realizations. There is a set of entry nodes for the attacker to choose from. The attacker’s strategy specifies an entry node and from it an attack path to DA. The attacker’s goal is to maximize the success rate by choosing the best path.

The pure strategy version of our model can be reduced to the *single-source single-destination shortest path edge interdiction* problem, which is known to be NP-hard Bar-noy et al. [1995]. However, NP-hardness on general graphs does not rule out efficient algorithms for practical AD graphs. *Active Directory style attack graphs exhibit special graph structures and we can exploit these structures to derive scalable algorithms*. We adopt *fixed-parameter analysis*. Formally, given an NP-hard problem with problem size  $n$ , let the easy-to-solve instances be characterized by special parameters  $k_1, \dots, k_c$ . If we are able to derive an algorithm that solves these instances in  $O(f(k_1, \dots, k_c) \text{POLY}(n))$ , then we claim that the problem is *fixed-parameter tractable (FPT)* with respect to the  $k_i$ s. Here,  $f$  is an arbitrary function that is allowed to be exponential. We do require that the running time be polynomial in the input size  $n$ . That is, an FPT problem is practically solvable for large inputs, as long as the special parameters are small (i.e., they indeed describe *easy* instances).

*It should be noted that this paper’s focus is not to push the frontier of theoretical fixed-parameter analysis. Instead, fixed-parameter analysis is our means to design scalable algorithms for our specific application on AD graphs.* As a matter of fact, our approaches combine theoretical fixed-parameter analysis and practical optimisation techniques.

We observe that practical AD graphs have two noticeable structural features.<sup>3</sup> We first observe that *the attack paths tend to be short*. Note that we are not claiming that long paths do not exist (i.e., there are cycles in AD graphs). The phrase “attack paths” refer to shortest paths that the attacker would actually use. The BLOODHOUND team uses the phrase “six degrees of domain admin” to draw the analogy to the famous “six degree of separation” idea from the small-world problem Milgram [1967] (i.e., all people in this world are on average six or fewer social connections away from each other). That is, in an organisation, it is expected that it takes only a few hops to travel from an intern’s account to the CEO’s account. Similar to the “small-world” hypothesis, attack paths being short is an *unproven observation* that we expect to hold true for practical purposes.

<sup>2</sup>Lallie et al. [2020] surveyed over 180 attack graphs/trees from academic literatures on cyber security.

<sup>3</sup>We have included an example synthetic AD graph in the appendix, which is generated using BLOODHOUND’s AD graph generator DBCREATOR. The AD environment of an organisation is considered sensitive, so in this paper, we only reference AD graphs generated using DBCREATOR and ADSIMULATOR, which is another open source tool for generating synthetic AD graphs.

The second structural feature is that AD graphs are very similar to trees. The tree-like structure comes from the fact that it is considered a best practise for the AD environment to follow the organisation chart. For example, human resources would form one tree branch while marketing would form another tree branch. However, an Active Directory attack graph is almost never exactly a tree, because there could be valid reasons for an account in human resources to access data on a computer that belongs to marketing. We could interpret Active Directory attack graphs as *trees with extra non-tree edges that represent security exceptions*.

Our aim is to design *practically scalable* algorithms for optimal pure and mixed strategy edge blocking. For organisations with thousands of computers in internal networks, the AD graphs generally involve *tens of thousands of nodes*. We manage to scale to such magnitude by exploiting the aforementioned structural features of practical AD graphs.

We first show that having short attack paths alone is not enough to derive efficient algorithms. Even if the maximum attack path length is a constant, both pure and mixed strategy blocking are NP-hard. We then focus on exploring the tree-likeness of practical AD graphs.

Our first approach focuses on pure strategy blocking only. For graphs with small tree widths, we propose a tree decomposition based dynamic program, which scales better than existing algorithms from Guo et al. [2022]. We then propose a general method for converting tree decomposition based dynamic programs to reinforcement learning environments. When tree widths are small, the derived reinforcement learning environments’ observation and action spaces are both small. This leads to an anytime algorithm that scales better than dynamic program, but loses the optimality guarantee.

Our second approach handles both pure and mixed strategy blocking. We invent a non-standard fixed parameter specifically for our application on AD graphs. A typical attack path describes a *privilege escalation* pathway. It is rare for a node to have more than one *privilege escalating* out-going edges (as such edges often represent security exceptions or misconfigurations). We observe that practical AD graphs consist of *non-splitting paths* (paths where every node has one out-going edge). For graphs with small numbers of non-splitting paths, we propose a kernelization technique that significantly downsizes the model, which is then solved via mixed-integer programming. We experimentally verify that this approach scales *exceptionally well* on synthetic AD graphs generated by two open source AD graph generators (DBCATOR and ADSIMULATOR).

## 2 Related Research

Guo et al. [2022] studied edge interdiction for AD graphs, where the attacker is given an entry node by nature. In Guo et al. [2022], the defensive goal is to maximize the attacker’s *expected* attack path length (i.e., number of hops). In this paper, the defensive goal is to minimize the attacker’s *worst-case* success rate, based on the assumption that different edges may have different failure rates. The authors proposed a tree decomposition based dynamic program, but it only applies to acyclic AD graphs. Practical AD graphs do contain cycles so this algorithm does not apply. The authors resorted to graph convolutional neural network as a heuristic to handle large AD graphs with cycles. *Our proposed algorithms can handle cycles, scale better, and produce the optimal results (instead of being mere heuristics)*. Furthermore, Guo et al. [2022] only studied pure strategy blocking.

Goel et al. [2022] studied a different model on edge interdiction for AD graphs. Under the authors’ model, both the attacker’s and the defender’s problem are #P-hard, and the authors proposed a defensive heuristic based on combining neural networks and diversity evolutionary computation.

The model studied in this paper is similar to the *bounded length cut* problem studied in Golovach and Thilikos [2011] and Dvořák and Knop [2018], where the goal is to remove some edges so that the minimum path length between a source and a destination meets a minimum threshold. Dvořák and Knop [2018] proposed a tree decomposition based dynamic program for the bounded length cut problem. The authors’ algorithms require that all source and destination nodes be added to every bag in the tree decomposition. *This is fine for showing the theoretical existence of FPT algorithms, but it is practically not scalable*. Furthermore, for bounded length cut, if a path is shorter than the threshold, then it *must be cut* and if a path is longer than the threshold, then it *can be safely ignored*. This is a much clearer picture than our model where we need to judge the relative importance of edges and spend the budget on the most vital ones.

Jain and Korzhyk [2011] proposed a double-oracle algorithm for equilibrium calculation on attack graphs, whose model is defined differently. Their approach is designed for general graphs so it only scales to a few hundred nodes and therefore is not suitable for practical AD graphs. Aziz et al. [2018, 2017] studied node interdiction for minimizing inverse geodesic length. Durkota et al. [2019] and Milani et al. [2020] studied deception based defense on cyber attack graphs.

### 3 Formal Model Description

We use a directed graph  $G = (V, E)$  to describe the Active Directory environment. Every edge  $e$  has a failure rate  $f(e)$ . There is one destination node DA (Domain Admin). There are  $s$  entry nodes. The attacker can start from any entry node and take any route. The attacker’s goal is to maximize the success rate to reach DA, by picking an optimal entry node and an optimal attack path. The defender picks  $b$  edges to block from a set of blockable edges  $E_b \subseteq E$ , where  $b$  is the defensive budget. The aim of the defender is to minimize the attacker’s success rate.

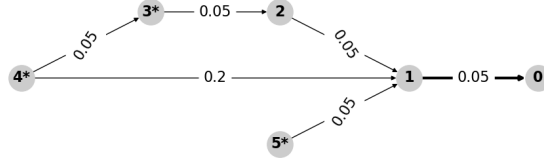


Figure 1: Example attack graph. Node 0 is DA. Node 4, 3, 5 are entry nodes (marked using \*). Edge labels represent the edges’ failure rates. Thick edges (i.e.,  $1 \rightarrow 0$ ) are not blockable. In the appendix, we illustrate how to optimally block 2 edges from this graph for both pure and mixed strategy.

We study both pure and mixed strategy blocking. We use  $B(e)$  to describe the probability that a blockable edge  $e \in E_b$  is blocked. For pure strategy blocking,  $B(e)$  equals either 0 or 1. For mixed strategy blocking,  $B(e)$  is in  $[0, 1]$ . The budget constraint is  $\sum_{e \in E_b} B(e) \leq b$ . We adopt the standard Stackelberg game model by assuming that the attacker can observe the defensive strategy and then plays a best response. For a mixed strategy defense, we assume that the attacker can observe  $B(e)$ ’s probabilistic values, but not the realisations.

Given  $B$ , the attacker’s optimal attack path can be found via  $\max_{p \in P} \left\{ \prod_{e \in p} (1 - f(e))(1 - B(e)) \right\}$ , where  $P$  is the set of all attack paths from all entry nodes. This maximization problem is equivalent to  $\min_{p \in P} \left\{ \sum_{e \in p} (-\ln(1 - f(e)) - \ln(1 - B(e))) \right\}$ . By applying natural log to convert from product to sum, we treat an edge’s “distance” as  $-\ln(1 - f(e)) - \ln(1 - B(e))$  (nonnegative). The attacker’s optimal attack path can be solved using Dijkstra’s shortest path algorithm Dijkstra [1959]. Let  $SR(B)$  be the success rate of the attacker facing blocking policy  $B$ . The defender’s problem is  $\min_B SR(B)$ .

Earlier we mentioned that our plan is to exploit the special structural features of practical AD graphs. Our first result is a negative result, which shows that having short attack paths alone is not enough to derive efficient algorithms. That is, we do need to consider the tree-like features.

**Theorem 1.** *Both pure and mixed strategy blocking are NP-hard for constant maximum attack path length.*

Proof is deferred to the appendix due to space constraint.

### 4 Tree Decomposition based Dynamic Program for Pure Strategy Blocking

*Tree decomposition* is a process that converts a general graph to a tree, where every tree node is a bag (set) of graph vertices. The maximum bag size minus one is called the *tree width*. A small tree width indicates that the graph is close to a tree. Many NP-hard combinatorial problems for general graphs become tractable if we focus on graphs with small tree widths. We show that this is also true for our model. For the rest of this section, we assume the readers are already familiar with tree decomposition related terminologies. We have included all relevant background information regarding tree decomposition in the appendix, including all algorithms/definitions/terminologies referenced in this paper (i.e., **nice** tree decomposition, **introduce/forget/join** node). We have also included a running example of our algorithm, pseudocode and relevant proofs in the appendix.

Throughout the discussion, we use *nodes* to refer to tree nodes in the tree decomposition and *vertices* to refer to vertices in AD graphs.

Besides assuming a small tree width, another key assumption of our dynamic program is that we assume a path’s success rate is from a small set of at most  $H$  values. A path’s success rate is the attacker’s success rate for going through it without any blocking. If all edges have the same failure rate, then  $H$  is just the *maximum attack path length*  $l$  plus 2 (0 to  $l$  hops, plus “no path”). In general, if the number of edge types is a small constant  $k$ , then  $H \in O(l^k)$ . In our experiments, we assume that there are two types of edges (high-failure-rate and low-failure-rate edges), which corresponds to  $H \in O(l^2)$ . It should be noted that  $H$  is only used for *worst-case* complexity analysis. In experiments,

given a specific graph, a path’s number of possible success rates is often significantly less (i.e., if the path is not blockable altogether, then there is only one possible success rate).

We call our DP TDCYCLE (tree decomposition with cycles<sup>4</sup>). The first step is to treat the attack graph as an undirected graph and then generate a tree decomposition. It should be noted that the optimal tree decomposition with the minimum tree width is NP-hard to compute Arnborg et al. [1987]. In our experiments, we adopt the *vertex elimination* heuristic for generating tree decomposition Bodlaender et al. [2006]. We then convert the resulting tree decomposition into a *nice tree decomposition* Cygan et al. [2015], where the root node is a bag containing DA only and all the leaf nodes are bags of size one. Please refer to the appendix for the process details. We use TD to denote the resulting nice tree decomposition. TD has  $O(wn)$  nodes where  $w$  is the tree width.

**Lemma 1.** *Let  $(u, v)$  be an arbitrary edge from the original AD graph. Under TD, there exists one and only one forget node  $X$ , whose child is denoted as  $X'$ , where  $\{u, v\} \subseteq X'$  and  $X' \setminus X$  is either  $\{u\}$  or  $\{v\}$ .*

The above lemma basically says that every edge can be “assigned” to exactly one forget node. For forget node  $(x_2, x_3, \dots, x_k)$  with child  $(x_1, x_2, \dots, x_k)$  (i.e.,  $x_1$  is forgotten), we assign all edges between  $x_1$  and one of  $x_2, \dots, x_k$  to this forget node. The high-level process of our dynamic program is that we first remove all edges from the graph. We then go through TD **bottom up**. At forget node  $X$ , we examine all the edges assigned to  $X$ . If an edge is not blockable or we decide not to block it, then we put it back into the graph. Otherwise, we do not put it back. After we finish the whole tree (while ensuring that the budget spent is at most  $b \iff$  we have put back at least  $|E| - b$  edges), we end up with a complete blocking policy.

Let  $X = (x_1, x_2, \dots, x_k)$  be a tree node ( $k \leq w + 1$ ). Let  $St(X)$  be the subtree of TD rooted at  $X$ . Let  $Ch(X)$  be the set of all graph vertices referenced in  $St(X)$ . Let  $Ch(X)' = Ch(X) \setminus X$ .  $Ch(X)'$  is then the set of vertices already forgotten after we process  $St(X)$  in the bottom-up fashion. A known property of tree decomposition is that the vertices in  $Ch(X)'$  cannot directly reach any vertex in  $V \setminus Ch(X)$ . That is, any attack path from an entry vertex in  $Ch(X)'$  to DA must pass through some  $x_i$  in  $X$ . Also, any attack path (not necessarily originating from  $Ch(X)'$ ) may involve vertices in  $Ch(X)'$  by entering the graph region from  $Ch(X)'$  via a certain  $x_i$  and then exit the region via a different node  $x_j$ . An attack path may “enter and exit” the region multiple times but all entries and exits must be via the vertices in  $X$ .

Suppose we have spent  $b'$  units of budget on (forget nodes of)  $St(X)$ . We do not need to keep track of the specifics of which edges have been blocked. We only need to track the total budget spending and the following “distance” matrix:

$$M = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1k} \\ \dots & & & \\ d_{k1} & d_{k2} & \dots & d_{kk} \end{bmatrix}$$

$d_{ij}$  represents the minimum path distance<sup>5</sup> between  $x_i$  and  $x_j$ , where the intermediate edges used are the edges we have already put back after processing  $St(X)$ . Diagonal element  $d_{ii}$  represents the minimum path distance from any entry vertex (among  $Ch(X)$ ) to  $x_i$ . We say the tuple  $(M, b')$  is *possible* at  $X$  if and only if it is possible to spend  $b'$  ( $b' \leq b$ ) on  $St(X)$  to achieve the distance matrix  $M$ .

Every tree node of TD corresponds to a DP subproblem and there are  $O(wn)$  subproblems. The subproblem corresponding to  $X$  is denoted as  $DP(X)$ .  $DP(X)$  simply returns the collection of all possible tuples at node  $X$ .

**Base cases:** For a leaf node  $X = \{x\}$ , if  $x$  is an entry vertex, then  $DP(X)$  contains one tuple, which is  $([0], 0)$ . Otherwise, the only possible tuple in  $DP(X)$  is  $([\infty], 0)$ .

**Original problem:** The root of TD is  $\{DA\}$ . The original problem is then  $DP(\{DA\})$ , which returns the collection of all possible tuples at the root. Every tuple from the collection has the form  $([d_{DA,DA}^*], b')$ , which represents that it is possible to spend  $b'$  to ensure that the attacker’s distance from DA is  $d_{DA,DA}^*$ . The maximum  $d_{DA,DA}^*$  in  $DP(\{DA\})$  corresponds to the attacker’s success rate facing optimal blocking.

We then present the recursive relationship for our DP:

**Introduce node:** Let  $X = (x_1, \dots, x_k, y)$  be an introduce node, whose child is  $X' = (x_1, \dots, x_k)$ . Given a possible tuple in  $DP(X')$ , we generate a new tuple as follows, which should belong to  $DP(X)$ .  $d_{yy}$  is 0 if  $y$  is an entry vertex

<sup>4</sup>This is to differentiate from the dynamic program proposed in Guo et al. [2022], which cannot handle cycles and does not guarantee correctness for practical AD graphs as they do contain cycles.

<sup>5</sup>Recall that given an edge with failure rate  $f(e)$ , we treat the edge’s “distance” as  $-\ln(1 - f(e))$  when it is not blocked.

and it is  $\infty$  otherwise (when  $y$  is introduced, all its edges have not been put back yet so it is disconnected from the  $x_i$ ).

$$\left( \begin{bmatrix} d_{11} & \dots & d_{1k} \\ \dots & & \\ d_{k1} & \dots & d_{kk} \end{bmatrix}, b' \right) \rightarrow \left( \begin{bmatrix} d_{11} & \dots & d_{1k} & \infty \\ \dots & & & \\ d_{k1} & \dots & d_{kk} & \infty \\ \infty & \dots & \infty & d_{yy} \end{bmatrix}, b' \right)$$

**Forget node:** Let  $X = (x_2, \dots, x_k)$  be a forget node, whose child is  $X' = (x_1, \dots, x_k)$ . At  $X$ , we need to determine how to block edges connecting  $x_1$  and the rest  $x_2, \dots, x_k$ . There are at most  $k - 1$  edges to block so we simply go over at most  $2^{k-1}$  blocking options. For each specific blocking option (corresponding to a spending of  $b''$ ), we convert a tuple in  $DP(X')$  to a tuple in  $DP(X)$  as follows (the new tuple is discarded if  $b' + b'' > b$ ):

$$\left( \begin{bmatrix} d_{11} & \dots & d_{1k} \\ \dots & & \\ d_{k1} & \dots & d_{kk} \end{bmatrix}, b' \right) \rightarrow \left( \begin{bmatrix} d'_{22} & \dots & d'_{2k} \\ \dots & & \\ d'_{k2} & \dots & d'_{kk} \end{bmatrix}, b' + b'' \right)$$

The  $d'_{ij}$  are updated distances considering the newly put back edges. We need to run an *all-pair shortest path* algorithm with complexity  $O(k^3)$  for this update.

**Join node:** Let  $X$  be a join node with two children  $X_1$  and  $X_2$ . For  $(M_1, b_1) \in DP(X_1)$  and  $(M_2, b_2) \in DP(X_2)$ , we label  $(M', b_1 + b_2)$  as a possible tuple in  $DP(X)$  if  $b_1 + b_2 \leq b$ .  $M'$  is the element-wise minimum between  $M_1$  and  $M_2$ .

**Theorem 2.** TDCYCLE's complexity is  $O(H^{2w^2} b^2 w^2 n)$ .

Proof is deferred to the appendix due to space constraint.

To summarize our dynamic program, we follow a bottom-up order (from leaf nodes of nice tree decomposition TD to the root). We propagate the set of all possible tuples  $(M, b')$  as we process the nodes. At introduce/join nodes, we follow a pre-determined propagation rule and do not make any blocking decisions. At forget nodes, we decide which edges to block from at most  $w$  edges. Given a specific AD graph and its corresponding tree decomposition TD, we can convert our dynamic program to a reinforcement learning environment as follows, which leads to an anytime algorithm that scales better than dynamic program (i.e., RL can always produce a solution, which may or may not be optimal, and generally improves over time; on the other hand, dynamic program cannot scale to handle slightly larger tree widths). *Our conversion technique can potentially be applied to tree decomposition based dynamic programs for other combinatorial optimisation problems.*

- We use post-order traversal on TD to create an ordering of the nodes (children are processed before parents).
- Instead of propagating *all possible* tuples  $(M, b')$ , we only propagate the *best tuple* (we have found so far during RL training). For example, consider a forget node  $X$  with child  $X'$ . The best tuple at node  $X'$  is passed on to  $X$  as *observation*. The *action* for this observation is then to decide which edges to block at  $X$ .
- Specifically to our model, for forget node  $X$ , if there are  $k$  ( $k \leq w$ ) edges to block, then we treat it as  $k$  separate steps in our reinforcement learning environment. That is, every *step* makes a blocking decision on a single edge and the action space is always binary.
- For introduce/join nodes, since we do not need to make any decisions, our reinforcement learning environment automatically processes these nodes (between *steps*).
- We set a final *reward* that is equal to the solution quality.

After we convert a specific AD graph into a reinforcement learning environment, we can then apply standard RL algorithms to search for the optimal blocking policy for the AD graph under discussion.

Following our conversion, both the observation and the action spaces are small when tree widths are small. Unfortunately, one downside of the above conversion technique is that there is no guarantee on having a small *episode length*. We could argue that it is impossible to guarantee small observation space, small action space, and small episode length at the same time, unless the AD graph is relatively small in scale. *After all, we are solving NP-hard problems.* Experimentally, for smaller AD graphs, we are able to achieve near-optimal performances as the episode lengths are manageable. For larger AD graphs, the episode lengths are too long. We introduce the following heuristic for limiting the episode length. Let  $T$  be the target episode length. Our idea is to hand-pick  $T$  *relatively important* edges and set the unpicked edges not blockable. In our experiments, we first calculate the *min cut* that separates the entry nodes from DA (unblockable edges' capacities are set to be large). Let the number of blockable edges in the min cut be  $C$ . If  $C \geq T$ , then we simply

treat these  $C$  edges as important and set the episode length to  $C$ . If  $C < T$ , then we add in  $T - C$  blockable edges that are closest to DA.

In the appendix, we include an experiment showing that our RL-based approach is not merely performing “random searching” via exploration. It is indeed capable of “learning” to react to the given observation. Under our original approach, the observation contains the distance matrix  $M$ , the budget spent  $b'$ , and also the current step index. We show that if we replace  $M$  by the zero matrix or by a random matrix, then the training results significantly downgrade.

## 5 Kernelization

*Kernelization* is a commonly used fixed-parameter analysis technique that preprocesses a given problem instance and converts it to a much smaller *equivalent* problem, called the *kernel*. We require that the kernel’s size be bounded by the special parameters (and not depend on  $n$ ).

As mentioned in the introduction, for practical AD graphs, most nodes have at most one out-going edge. If an edge is not useful for the attacker, then we can remove it without loss of generality. If an edge is useful for the attacker, then generally, it is *privilege escalating*. It is rare for a node to have two separate *privilege escalating* out-going edges (as they often correspond to security exceptions or misconfigurations). We use SPLIT to denote the set of all splitting nodes (nodes with multiple out-going edges). We use SPLIT+DA to denote SPLIT with DA added. We use ENTRY to denote the set of all entry nodes. We invent a new parameter called the number of *non-splitting paths*. Experimentally, this parameter leads to algorithms that scale exceptionally well for synthetic AD graphs generated using two different open source AD graph generators.

**Definition 1** (Non-splitting path). Given node  $u$ , let  $v$  be one of  $u$ ’s successors. The non-splitting path  $\text{NSP}(u, v)$  is defined recursively:

- If  $v \in \text{SPLIT+DA}$ , then  $\text{NSP}(u, v)$  is  $u \rightarrow v$ .
- Otherwise,  $v$  must have a unique successor  $v'$ .  $\text{NSP}(u, v)$  is the path that combines  $u \rightarrow v$  and  $\text{NSP}(v, v')$ .

In words,  $\text{NSP}(u, v)$  is the path that goes from  $u$  to  $v$ , then repeatedly moves onto the only successor of  $v$  if  $v$  has only one successor, until we reach either a splitting node or DA. We use  $\text{DEST}(u, v)$  to denote the ending node of  $\text{NSP}(u, v)$ . We have  $\text{DEST}(u, v) \in \text{SPLIT+DA}$ . A non-splitting path is called *blockable* if at least one of its edges is blockable.

An AD graph can be viewed as the **joint** of the following set of non-splitting paths. Our parameter (the number of non-splitting paths  $\#\text{NSP}$ ) is the size of this set.

$$\{\text{NSP}(u, v) \mid u \in \text{SPLIT} \cup \text{ENTRY}, v \in \text{SUCCESSORS}(u)\}$$

The blockable edge furthest away from  $u$  on the path  $\text{NSP}(u, v)$  is denoted as  $\text{BW}(u, v)$ . BW stands for *block-worthy* due to the following lemma.

**Lemma 2.** *For both pure and mixed strategy blocking, we never need to spend more than one unit of budget on a non-splitting path. For any problem instance, there exists an optimal defense that blocks only edges from the following set:*

$$\text{BW} = \{\text{BW}(u, v) \mid u \in \text{SPLIT} \cup \text{ENTRY}, v \in \text{SUCCESSORS}(u)\}$$

We present how to formulate our model as a nonlinear program, based on the aforementioned non-splitting path interpretation. The nonlinear program can then be converted to MIPs and be efficiently solved using state-of-the-art MIP solvers. We use  $B_e$  to denote the unit of budget spent on edge  $e$ .  $B_e$  is binary for pure strategy blocking and is between 0 and 1 for mixed strategy blocking. As mentioned earlier,  $B_e \geq 0$  for  $e \in \text{BW}$  and  $B_e = 0$  for  $e \notin \text{BW}$ . We use  $r_u$  to denote the success rate of node  $u$ . The success rate of a node is the success rate of the optimal attack path starting from this node, under the current defense (i.e., the  $B_e$ ). We use  $c_{u,v}$  to denote the success rate of the non-splitting path  $\text{NSP}(u, v)$  when no blocking is applied.  $c_{u,v} = \prod_{e \in \text{NSP}(u,v)} (1 - f(e))$ . The  $c_{u,v}$  are constants. We use  $r^*$  to represent the attacker’s optimal success rate.

$$C = \{(u, v) \mid u \in \text{SPLIT} \cup \text{ENTRY}, v \in \text{SUCCESSORS}(u)\}$$

$$C^+ = \{(u, v) \mid (u, v) \in C, \text{NSP}(u, v) \text{ blockable}\}$$

$$C^- = \{(u, v) \mid (u, v) \in C, \text{NSP}(u, v) \text{ not blockable}\}$$

We have the following nonlinear program:

$$\begin{aligned}
 \min \quad & r^* \\
 r^* \geq & r_u & \forall u \in \text{ENTRY} \\
 r_u \geq & r_{\text{DEST}(u,v)} \cdot c_{u,v} \cdot (1 - B_{\text{BW}(u,v)}) & \forall (u,v) \in C^+ \\
 r_u \geq & r_{\text{DEST}(u,v)} \cdot c_{u,v} & \forall (u,v) \in C^- \\
 b \geq & \sum_{e \in \text{BW}} B_e \\
 r_{\text{DA}} = & 1 \\
 r_u, r^* \in & [0, 1] \\
 B_e \in & \{0, 1\} \text{ or } [0, 1]
 \end{aligned}$$

The above program has at most  $O(\#\text{NSP})$  variables and at most  $O(\#\text{NSP})$  constraints (both do not depend on  $n$ ).

**Integer program for pure strategy blocking:** For pure strategy blocking, the above program can be converted to an IP by rewriting  $r_u \geq r_{\text{DEST}(u,v)} \cdot c_{u,v} \cdot (1 - B_{\text{BW}(u,v)})$  into an equivalent linear form  $r_u \geq r_{\text{DEST}(u,v)} \cdot c_{u,v} - B_{\text{BW}(u,v)}$ .

## 6 Mixed Strategy Blocking

We can convert the nonlinear program from the previous section into a program that is *almost* linear. We first observe that if under the optimal mixed strategy defense, the attacker's success rate is at least  $\epsilon$ , then that means we never want to block any edge with a probability that is *strictly* more than  $1 - \epsilon$ . Due to this observation, we artificially enforce that we do not block an edge with more than  $1 - \epsilon$  probability and solve for the optimal defense under this restriction. If in the end result, the attacker's success rate is at least  $\epsilon$ , then our restriction is not actually a restriction at all. If our solution says that the attacker's success rate is less than  $\epsilon$ , then we have an almost optimal defense anyway. In this paper, we set  $\epsilon = 0.01$ . That is, for  $e \in \text{BW}$ , we require that  $0 \leq B_e \leq 0.99$  instead of  $0 \leq B_e \leq 1$ .

With the above observation, we can convert the nonlinear program, which involves multiplication, to an *almost* linear program as follows. Our trick is to replace  $r_u$  by  $r'_u = -\ln(r_u)$ , replace  $r^*$  by  $r'^* = -\ln(r^*)$ , replace  $B_e$  by  $B'_e = -\ln(1 - B_e)$ , and finally replace  $c_{u,v}$  by  $c'_{u,v} = -\ln(c_{u,v})$ . Our variables are now  $r'^*$ , the  $r'_u$  and the  $B'_e$ . Due to monotonicity of natural log, we can rewrite the earlier nonlinear program as:

$$\begin{aligned}
 \max \quad & r'^* \\
 r'^* \leq & r'_u & \forall u \in \text{ENTRY} \\
 r'_u \leq & r'_{\text{DEST}(u,v)} + c'_{u,v} + B'_{\text{BW}(u,v)} & \forall (u,v) \in C^+ \\
 r'_u \leq & r'_{\text{DEST}(u,v)} + c'_{u,v} & \forall (u,v) \in C^- \\
 b \geq & \sum_{e \in \text{BW}} (1 - e^{-B'_e}) \\
 r'_{\text{DA}} = & 0 \\
 r'_u, r'^* \in & [0, \infty) \\
 B'_e \in & [0, -\ln(0.01)]
 \end{aligned}$$

Unfortunately, the above program is not linear as the budget constraint is not linear. Furthermore, the above program is not even convex: Since  $1 - e^{-x}$  is concave, the average of two feasible solutions may violate the budget constraint.

### 6.1 Iterative LP based approximation heuristic

We note that  $1 - e^{-x}$  is increasing. That is, as long as we push down the total of  $B'_e$ , we eventually will reach a situation where the budget constraint is satisfied. That is, we rewrite the budget constraint as  $b' \geq \sum_{e \in \text{BW}} B'_e$ , which is a linear constraint. We **guess** a value for  $b'$  and then solve the corresponding LP. We verify whether the LP solution also satisfies the original nonlinear budget constraint. If it does, then that means we could increase our guess of  $b'$  (increasing  $b'$  means spending more budget). If our LP solution violates the original nonlinear budget constraint, then it means we should decrease our guess of  $b'$ . A good guess of  $b'$  can be obtained via a binary search from 0 to  $-\lvert\text{BW}\rvert \ln(0.01)$ .

### 6.2 MIP based approximation

Another way to address the nonlinear budget constraint is to add the  $B_e$  back into the model ( $0 \leq B_e \leq 0.99$  for  $e \in \text{BW}$ ). The budget constraint  $\sum_{e \in \text{BW}} B_e \leq b$  is now back to linear. Of course, this cannot be the end of the story, since we also need to link the  $B_e$  and the  $B'_e$  together. We essentially have introduced a new set of nonlinear constraints, which are  $B'_e = -\ln(1 - B_e)$  for  $e \in \text{BW}$ .

The function  $-\ln(1 - x)$  is close to a straight line if we focus on a small interval. If  $x \in [a, b]$ , the straight line connecting  $(a, -\ln(1 - a))$  and  $(b, -\ln(1 - b))$  is an upper bound of  $-\ln(1 - x)$ . A straight line representing a lower



bound of  $-\ln(1-x)$  is the tangent line at the interval’s mid point  $\frac{a+b}{2}$ . An illustration of  $-\ln(1-x)$ , bounded above and below by straight lines is provided in the appendix.

Given a specific  $B_e$ , we could divide  $B_e$ ’s region  $[0, 0.99]$  into multiple smaller intervals. For each region, we have two straight lines that represent the lower and upper bounds on  $-\ln(1-x)$ . We could join the regions and compose  $g_L(x)$  and  $g_U(x)$ , which are the *piece-wise linear* lower and upper bounds on  $-\ln(1-x)$ . We replace  $B'_e = -\ln(1-B_e)$  by  $B'_e = g_L(B_e)$  or  $B'_e = g_U(B_e)$ , respectively, to create two different MIP programs. We use the *multiple choice model* presented in Croxton et al. [2003] to implement piece-wise linear constraints with the help of auxiliary binary variables. The program with  $B'_e = g_L(B_e)$  underestimates the  $B'_e$ , which results in an overestimation of the budget spending and therefore an overestimation of the attacker’s success rate. This program results in **an achieved feasible defense**. The program with  $B'_e = g_U(B_e)$  on the other hand results in **a lower bound on the attacker’s success rate**. When the intervals are fine enough, experimentally, the achieved feasible defense is close to the lower bound (therefore close to optimality).

## 7 Experiments

All our experiments are carried out on a desktop with i7-12700 CPU and NVIDIA GeForce RTX 3070 GPU. Our MIP solver is Gurobi 9.5.1. For pure strategy blocking, we proposed two algorithms: TDCYCLE and IP (integer program based on kernelization). We also include a third algorithm GREEDY to serve as a baseline. GREEDY spends one unit of budget in each round for a total of  $b$  rounds. In each round, it greedily blocks one edge to maximally decrease the attacker’s success rate. For mixed strategy blocking, we have an iterative LP based heuristic ITERLP, a mixed integer program MIP-F(EASIBLE) for generating a feasible defense and a mixed integer program MIP-LB for generating a lower bound on the attacker’s success rate.

We evaluate our algorithms using two attack graphs generated using BLOODHOUND team’s synthetic graph generator DBCREATOR. We call the attack graph R2000 and R4000, which are obtained by setting the number of computers in the AD environment to 2000 and 4000. R2000 contains 5997 nodes and 18795 edges and R4000 contains 12001 nodes and 45780 edges. We also generate a third attack graph using a different open source synthetic graph generator ADSIMULATOR. ADSIMULATOR by default generates a trivially small graph. We increase all its default parameters by a factor of 10 and create an attack graph called ADS10. ADS10 contains 3015 nodes and 12775 edges. Even though ADS10 contains less nodes, experimentally it is actually more expensive to work with compared to R2000 and R4000, as it is further away from a tree.

We only consider three edge types: ADMINTO, MEMBEROF, and HASSESSION. These are a representative sample of edges types used in BloodHound. We set the failure rates of all edges of type HASSESSION to 0.2 (requiring a session between an account and a computer, therefore more likely to fail) and set the failure rates of all edges of the other two types to 0.05. We set the number of entry nodes to 20. We select 40 nodes that are furthest away from DA (in terms of the number of hops to reach DA) and randomly draw 20 nodes among them to be the entry nodes. We define  $HOP(e)$  to be the minimum number of hops between an edge  $e$  and DA. We set MAXHOP to be the maximum value for  $HOP(e)$ .

An edge is set to be blockable with probability  $\frac{HOP(e)}{MAXHOP}$ . That is, edges further away from DA are set to be more likely to be blockable. Generally speaking, edges further away from DA tend to be about individual employees’ accesses instead of accesses between servers and admins. We set the budget to 5 and 10. All experiments are repeated 10 times, with different random draws of the entry nodes and the blockable edges. The numbers in the table are the attacker’s average success rates over 10 trials. The numbers in the parenthesis are the average running time.

budget=5	R2000	R4000	ADS10
GREEDY	0.521 (0.04s)	0.376 (0.34s)	0.448 (4.57s)
TDCYCLE	0.480 (0.10s)	0.373 (15366s)	-
IP	0.480 (0.01s)	0.373 (0.06s)	0.409 (0.09s)
ITERLP	0.337 (0.37s)	0.180 (0.57s)	0.300 (1.90s)
MIP-F	0.335 (0.11s)	0.179 (3.54s)	0.303 (4.75s)
MIP-LB	0.333 (0.11s)	0.176 (2.61s)	0.297 (4.98s)

budget=10	R2000	R4000	ADS10
GREEDY	0.499 (0.07s)	0.376 (0.65s)	0.448 (8.97s)
TDCYCLE	0.263 (0.18s)	-	-
IP	0.263 (0.01s)	0.117 (0.03s)	0.315 (0.10s)
ITERLP	0.266 (0.39s)	0.033 (0.60s)	0.190 (1.92s)
MIP-F	0.270 (0.02s)	0.023 (0.54s)	0.191 (11.27s)
MIP-LB	0.263 (0.02s)	0.014 (0.24s)	0.183 (3.62s)

**Interpretation of Results:** For pure strategy blocking, TDCYCLE and IP are both expected to produce the optimal results. As expected, they perform better than GREEDY. TDCYCLE doesn’t scale for 3 out of 6 settings. On the other hand, IP scales exceptionally well. As mentioned earlier, IP scales better since it is based on a parameter that we invent specifically for describing AD graphs. For graphs with large number of non-splitting paths and small tree widths, we expect TDCYCLE to scale better, *but such graphs may not be AD graphs*. For mixed strategy blocking, the attacker’s success rates under both ITERLP and MIP-F(EASIBLE) are close to MIP-LB (lower bound on the attacker’s success rate), which indicates that both heuristics are near-optimal.

**Results on scaling TDCYCLE via reinforcement learning:** We present results on two settings: R2000 with  $b = 5$  and ADS10 with  $b = 10$ . These two are the *cheapest* and the *most expensive* among our six experimental settings. All our experimental setups are the same as before. We directly apply *Proximal Policy Optimization* PPO Schulman et al. [2017].<sup>6</sup> For each environment, we use training seed 0 to 9 and record the best result.

	Opt	Greedy	RL	=Opt	Time
R2000, $b = 5$	0.480	0.521	0.480	10/10	1hr
ADS10, $b = 10$	0.315	0.448	0.319	9/10	4hr

OPT, GREEDY, RL each represents the average performance of these three different approaches (average over 10 trials with random draws of the entry nodes and the blockable edges). “Time” refers to training time. For R2000 with  $b = 5$ , we obtain the optimal result in 10 out of 10 trials without limiting the episode length. The maximum episode length is 27 over 10 trials. For ADS10 with  $b = 10$ , we set an episode length of 15 and obtain the optimal result in 9 out of 10 trials. We recall that, for this setting, TDCYCLE doesn’t scale at all, but we manage to achieve near-optimal results via reinforcement learning.

<sup>6</sup>We list our hyper-parameters in the appendix.

## Acknowledgements

Frank Neumann has been supported by the Australian Research Council through grant FT200100536. Hung Nguyen is supported by the Next Generation Technology Fund (NGTF-Cyber) grant MyIP 10614 the Australian Research Council (ARC-NISDRG) grant NI210100139. This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide.

## References

- Enterprise Management Associates EMA. The Rise of Active Directory Exploits: Is it Time to Sound the Alarm? Technical report, 2021.
- Harjinder Singh Lallie, Kurt Debattista, and Jay Bal. A review of attack graph and attack tree visual syntax in cyber security. *Computer Science Review*, 35:100219, February 2020. ISSN 15740137. doi:10.1016/j.cosrev.2019.100219.
- John Dunagan, Alice X. Zheng, and Daniel R. Simon. Heat-ray: Combating identity snowball attacks using machine-learning, combinatorial optimization and attack graphs. In *Proceedings of the ACM SIGOPS 22nd Symposium on Operating Systems Principles - SOSP '09*, page 305, Big Sky, Montana, USA, 2009. ACM Press. ISBN 978-1-60558-752-3. doi:10.1145/1629575.1629605.
- Amotz Bar-noy, Samir Khuller, and Baruch Schieber. The Complexity of Finding Most Vital Arcs and Nodes. Technical report, University of Maryland, 1995.
- Stanley Milgram. The Small-World Problem. *Psychology Today*, 1:61–67, 1967.
- Mingyu Guo, Jialiang Li, Aneta Neumann, Frank Neumann, and Hung Nguyen. Practical fixed-parameter algorithms for defending active directory style attack graphs. In *The 36th AAAI Conference on Artificial Intelligence (AAAI), Vancouver, Canada, 2022*, 2022. URL <https://arxiv.org/abs/2112.13175>.
- Diksha Goel, Max Hector Ward-Graham, Aneta Neumann, Frank Neumann, Hung Nguyen, and Mingyu Guo. Defending active directory by combining neural network based dynamic program and evolutionary diversity optimisation. In *GECCO '22: Genetic and Evolutionary Computation Conference, 2022*, 2022. URL <http://arxiv.org/abs/2204.03397>.
- Petr A. Golovach and Dimitrios M. Thilikos. Paths of bounded length and their cuts: Parameterized complexity and algorithms. *Discrete Optimization*, 8(1):72–86, February 2011. ISSN 15725286. doi:10.1016/j.disopt.2010.09.009.
- Pavel Dvořák and Dušan Knop. Parameterized Complexity of Length-bounded Cuts and Multicuts. *Algorithmica*, 80(12):3597–3617, December 2018. ISSN 1432-0541. doi:10.1007/s00453-018-0408-7.
- Manish Jain and Dmytro Korzhuk. A Double Oracle Algorithm for Zero-Sum Security Games on Graphs. In *10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011), Taipei, Taiwan, May 2-6, 2011, Volume 1-3*, pages 327–334, 2011.
- Haris Aziz, Serge Gaspers, Edward J. Lee, and Kamran Najeebullah. Defender stackelberg game with inverse geodesic length as utility metric. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems, AAMAS '18*, pages 694–702, Richland, SC, 2018. International Foundation for Autonomous Agents and Multiagent Systems.
- Haris Aziz, Serge Gaspers, and Kamran Najeebullah. Weakening Covert Networks by Minimizing Inverse Geodesic Length. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, pages 779–785, Melbourne, Australia, August 2017. International Joint Conferences on Artificial Intelligence Organization. ISBN 978-0-9992411-0-3. doi:10.24963/ijcai.2017/108.
- Karel Durkota, Viliam Lisý, Branislav Božanský, Christopher Kiekintveld, and Michal Pěchouček. Hardening networks against strategic attackers using attack graph games. *Computers & Security*, 87:101578, November 2019. ISSN 01674048. doi:10.1016/j.cose.2019.101578.
- Stephanie Milani, Weiran Shen, Kevin S. Chan, Sridhar Venkatesan, Nandi O. Leslie, Charles Kamhoua, and Fei Fang. Harnessing the power of deception in attack graph-based security games. In Quanyan Zhu, John S. Baras, Radha Poovendran, and Juntao Chen, editors, *Decision and Game Theory for Security*, pages 147–167, Cham, 2020. Springer International Publishing. ISBN 978-3-030-64793-3.
- Edsger W Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1:269–271, 1959.
- Stefan Arnborg, Derek G. Corneil, and Andrzej Proskurowski. Complexity of finding embeddings in a k-tree. *Siam Journal of Discrete Mathematics*, 8(2):277–284, 1987.

- Hans L. Bodlaender, Fedor V. Fomin, Arie M. C. A. Koster, Dieter Kratsch, and Dimitrios M. Thilikos. On Exact Algorithms for Treewidth. In Yossi Azar and Thomas Erlebach, editors, *Algorithms – ESA 2006*, Lecture Notes in Computer Science, pages 672–683, Berlin, Heidelberg, 2006. Springer. ISBN 978-3-540-38876-0. doi:10.1007/11841036\_60.
- Marek Cygan, Fedor V. Fomin, Łukasz Kowalik, Daniel Lokshtanov, Daniel Marx, Marcin Pilipczuk, Michał Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer International Publishing, 2015. ISBN 978-3-319-21274-6. doi:10.1007/978-3-319-21275-3.
- Keely L. Croxton, Bernard Gendron, and Thomas L. Magnanti. A Comparison of Mixed-Integer Programming Models for Nonconvex Piecewise Linear Cost Minimization Problems. *Management Science*, 49(9):1268–1273, 2003. ISSN 0025-1909.
- John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms, 2017. URL <https://arxiv.org/abs/1707.06347>.
- Georg Baier, Thomas Erlebach, Alexander Hall, Ekkehard Köhler, and Heiko Schilling. Length-Bounded Cuts and Flows. In *In Proc. 33rd International Colloquium on Automata, Languages and Programming (ICALP)*, pages 679–690, 2006.
- Irit Dinur and Samuel Safra. On the Hardness of Approximating Minimum Vertex Cover. *Annals of Mathematics*, 162(1):439–485, 2005. ISSN 0003-486X.

## Appendix

### 7.1 An example synthetic AD graph

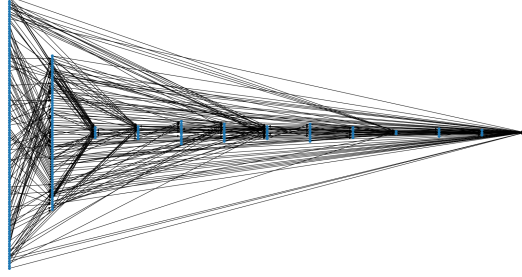


Figure 2: An example Active Directory attack graph generated using BLOODHOUND team’s synthetic graph generator DBCREATOR (setting the number of computers in the network to 1000). Only nodes reachable to DA (rightmost node) are shown. Every node only needs several hops to reach DA and the whole graph is similar to a tree.

### 7.2 How to optimally block 2 edges from the example attack graph described in Figure 1

- Optimal pure strategy defense: One optimal pure strategy defense is to block  $2 \rightarrow 1$  and  $5 \rightarrow 1$ . After these two edges are blocked, the attacker’s optimal attack path is  $4 \rightarrow 1 \rightarrow 0$ , with a success rate  $(1 - 0.2)(1 - 0.05) = 0.76$ .
- Optimal mixed strategy defense: The optimal mixed strategy defense is to block  $2 \rightarrow 1$  with probability 0.675, block  $4 \rightarrow 1$  with probability 0.634, and block  $5 \rightarrow 1$  with probability 0.691. Note that  $0.675 + 0.634 + 0.691 = 2$ . Under this defense, there are three optimal attack paths for the attacker, which are  $3 \rightarrow 2 \rightarrow 1 \rightarrow 0$ ,  $4 \rightarrow 1 \rightarrow 0$ , and  $5 \rightarrow 1 \rightarrow 0$ . All three optimal attack paths’ success rates are 0.279. For example, for attack path  $3 \rightarrow 2 \rightarrow 1 \rightarrow 0$ , the success rate equals  $(1 - 0.05)(1 - 0.05)(1 - 0.675)(1 - 0.05) = 0.279$ .

### 7.3 Proof of Theorem 1

We only present the proof for mixed strategy blocking. For pure strategy, we can derive a similar proof or directly apply the result from Baier et al. [2006], where the authors show that for directed graphs, the *single-source single-destination 4-length-bounded edge cut* problem is NP-hard. 4-length-bounded edge cut is to calculate the minimum number of edges to block so that the distance from the source to the destination is at least 5 (equivalently, exactly 5).

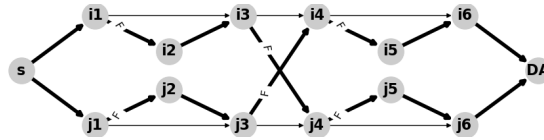


Figure 3: Proof gadget:  $i1$  to  $i6$  and  $j1$  to  $j6$  represent *connected* node  $i$  and  $j$  from a vertex cover instance

*Proof for mixed strategy only.* Dinur and Safra [2005] showed that the minimum vertex cover problem is NP-hard to approximate to within a factor of 1.3606. We show that optimal mixed strategy blocking is at least as hard as approximating vertex cover to within a factor of 1.3606 by reducing vertex cover to mixed strategy blocking.

Let  $G(V', E')$  be a vertex cover instance.  $G'$  is an undirected graph. Let  $n' = |V'|$  and let  $k^*$  be the minimum vertex cover number. We construct an attack graph as follows. We first construct a single entry node  $s$  and a single destination  $DA$ . For every node  $i \in V'$ , we construct 6 nodes and 9 edges in the attack graph. For  $i \in V'$ , the 6 nodes are  $i1$  to  $i6$ , and the 9 edges are  $s \rightarrow i1$ ,  $i6 \rightarrow DA$ , and the 7 edges among  $i1$  to  $i6$  (as shown in Figure 3). For every edge  $(i, j) \in E'$ , we construct 2 edges in the attack graph. If  $i$  and  $j$  are connected by an edge in  $G'$ , we construct  $i3 \rightarrow j4$  and  $j3 \rightarrow i4$ . The thick edges are not blockable and the thin edges are blockable. Edges marked with the label  $F$  are called *broken edges*. Broken edges have a failure rate of  $1 - \epsilon^4$ , where  $\epsilon = \frac{1}{n'}$ . Every other edge has a failure rate of 0.

We define  $b^*$  to be the minimum budget where under the optimal mixed strategy defense, the attacker's success rate is at most  $\epsilon^8$ . We first show that  $b^* \leq n' + k^*$ . Assuming that an oracle provides us with a minimum vertex cover solution with  $k^*$  vertices, we can convert this vertex cover solution to a pure strategy defense that uses a budget of  $n' + k^*$  and can ensure that the attacker's success rate is at most  $\epsilon^8$ . Suppose  $i$  and  $j$  are connected and only  $i$  belongs to the minimum vertex cover, then for  $i$  we spend two units of budget (block  $i1 \rightarrow i3$  and  $i4 \rightarrow i6$ ) and for  $j$  we spend one unit of budget (only block  $j3 \rightarrow j4$ ). With the above blocking, all attack paths must pass through at least two broken edges. Therefore, by blocking  $2 \cdot k^* + 1 \cdot (n' - k^*) = n' + k^*$  edges, the attacker's success rate is at most  $\epsilon^8$ . Mixed strategy is more general than pure strategy. By the definition of  $b^*$ , we must have  $b^* \leq n' + k^*$ .

Now let us consider the optimal mixed strategy blocking solution corresponding to budget  $b^*$ . We show that we can convert this optimal mixed strategy solution to a pure strategy solution by increasing the budget spending by at most  $\frac{1}{1-\epsilon^2}$ . The resulting pure strategy blocking solution corresponds to a valid vertex cover solution that is arbitrarily close to optimality. This essentially means that if we can solve the mixed strategy blocking problem, then we can approximate vertex cover to within a factor that is arbitrarily close to 1, which shows optimal mixed strategy blocking is at least as hard as the known NP-hard problem of approximating vertex cover to within a factor of 1.3606.

Under the optimal mixed strategy defense with budget  $b^*$ , we call an edge *nearly blocked* if the defender blocks it with at least  $1 - \epsilon^2$  probability. For node  $i$ , we define  $L(i)$  to be 1 if the *left* edge  $i1 \rightarrow i3$  is nearly blocked.  $L(i) = 0$  otherwise. Similarly, we define  $M(i)$  to be 1 if the *middle* edge  $i3 \rightarrow i4$  is nearly blocked and  $R(i)$  to be 1 if the *right* edge  $i4 \rightarrow i6$  is nearly blocked. For two connected nodes  $i$  and  $j$  in  $G'$ , we can prove six inequalities: 1)  $L(i) + R(j) \geq 1$ ; 2)  $L(j) + R(i) \geq 1$ ; 3)  $L(i) + M(i) \geq 1$ ; 4)  $M(i) + R(i) \geq 1$ ; 5)  $L(j) + M(j) \geq 1$ ; 6)  $M(j) + R(j) \geq 1$ . Let us take  $L(i) + R(j) \geq 1$  as an example. If neither  $L(i)$  nor  $R(j)$  is nearly blocked, then the success rate by going through  $s \rightarrow i1 \rightarrow i3 \rightarrow j4 \rightarrow j6 \rightarrow DA$  is more than  $\epsilon^8$ . The above six inequalities imply two things. First,  $L(i) + M(i) + R(i) + L(j) + M(j) + R(j) \geq 3$ . That is, for  $i$  and  $j$  that are connected in  $G'$ , at least one of  $i$  and  $j$  have at least two edges nearly blocked. If node  $i$  has at least two edges that are nearly blocked, then we can modify the optimal mixed strategy and change to fully block  $i1 \rightarrow i3$  and  $i4 \rightarrow i6$  instead (budget on other edges reduced to 0). This way, the attacker can never gain a success rate that is strictly higher than  $\epsilon^8$  by using any of  $i$ 's edges. By changing from nearly blocked to fully blocked, we increase the budget spending by at most  $\frac{1}{1-\epsilon^2}$ . For node  $j$ , due to  $L(j) + M(j) \geq 1$  and  $M(j) + R(j) \geq 1$ , we have that at least one edge corresponding to  $j$  is nearly blocked and if there is exactly one edge nearly blocked, then that edge must be the middle one  $j3 \rightarrow j4$ . We again can increase our budget spending by at most  $\frac{1}{1-\epsilon^2}$  by changing the middle edge from nearly blocked to fully blocked (and reduce the spending on the other two edges to 0). Essentially, given the optimal mixed strategy defense, we can increase the budget spending by at most  $\frac{1}{1-\epsilon^2}$  in order to convert to a pure strategy defense that corresponds to a feasible vertex cover solution. Thus, we have  $\frac{b^*}{1-\epsilon^2} \geq n' + k^*$  and  $\frac{b^*}{1-\epsilon^2} - n'$  is an achieved vertex cover number. Earlier we proved

that  $b^* - n'$  is a lower bound of the vertex cover number. The approximation ratio is at most  $\frac{\frac{b^*}{1-\epsilon^2} - n'}{b^* - n'}$ , which becomes strictly less than 1.3606 when  $n'$  reaches a constant threshold. The mathematical details are presented below. For small  $n'$  values that are below the constant threshold, we could brute force to calculate the optimal vertex cover solution, which is also within a factor of 1.3606. That is, optimal mixed strategy blocking is at least as hard as the known NP-hard problem of approximating vertex cover to within a factor of 1.3606.

Below we show the details on the constant threshold:

$$\begin{aligned} \frac{\frac{b^*}{1-\epsilon^2} - n'}{b^* - n'} &= \frac{\frac{b^*}{1-\frac{1}{n'^2}} - n'}{b^* - n'} = \frac{\frac{b^* n'^2}{n'^2-1} - n'}{b^* - n'} \\ &= \frac{b^* n'^2 - n'(n'^2 - 1)}{(b^* - n')(n'^2 - 1)} = \frac{b^* n'^2 - n'^3 + n'}{b^* n'^2 - n'^3 + n' - b^*} \\ &= 1 + \frac{b^*}{b^* n'^2 - n'^3 + n' - b^*} \\ &= 1 + \frac{b^*}{(b^* - n')(n'^2 - 1)} \end{aligned}$$

The above expression is monotone in  $b^*$  and is maximized when  $b^*$  takes the minimum possible value. We know that  $b^*$  is at least  $(1 - \epsilon^2)(n' + k^*) \geq (1 - \frac{1}{n'^2})(n' + 1)$ . ( $k^*$  is at least 1 as it is the vertex cover number.) The approximation ratio is then at most

$$1 + \frac{(1 - \frac{1}{n'^2})(n' + 1)}{((1 - \frac{1}{n'^2})(n' + 1) - n')(n'^2 - 1)}$$

$$\begin{aligned}
 &= 1 + \frac{(n'^2 - 1)(n' + 1)}{((n'^2 - 1)(n' + 1) - n'^3)(n'^2 - 1)} \\
 &= 1 + \frac{n' + 1}{(n'^2 - 1)(n' + 1) - n'^3} \\
 &= 1 + \frac{n' + 1}{n'^2 - n' - 1}
 \end{aligned}$$

The above expression approaches 1 as  $n'$  approaches infinity. There exists a constant threshold where the above is at most 1.3606 when  $n'$  is above the threshold. □

## 7.4 Tree decomposition background

**Definition 2** (Tree Decomposition and Tree Width). Let  $G(V, E)$  be an undirected graph. Let  $T$  be a tree with  $t$  tree nodes  $(T_1, T_2, \dots, T_t)$ , where every tree node is a bag of graph vertices (subset of  $V$ ).  $T$  is a *tree decomposition* of  $G$  iff

- The union of the  $T_i$  equals  $V$ .
- For every edge  $e \in E$ , there exists at least one  $T_i$  such that both ends of  $e$  are in  $T_i$ .
- For every vertex  $v \in V$ , let  $T(v)$  be the set of tree nodes containing  $v$ . The subgraph of  $T$  induced by  $T(v)$  must be connected (must form a tree).

The *tree width* of  $G$  (under the tree decomposition  $T$ ) equals the maximum size/cardinality among the  $T_i$ , then minus 1.

AD graphs are directed, but we treat AD graphs as undirected for the purpose of applying tree decomposition. For all AD graphs used in this paper (generated using DBCREATOR and ADSIMULATOR), between any pair of directly connected vertices  $u$  and  $v$ , either  $u \rightarrow v$  or  $v \rightarrow u$  exists (never both). Hence treating the graphs as undirected does cause any ambiguity or lose any information, as we can always refer back to the original AD graph to query the edge directions. (Even if both  $u \rightarrow v$  and  $v \rightarrow u$  exist, we simply need to add one auxiliary vertex in between to remove any ambiguity caused. For example, keep  $u \rightarrow v$  and add  $x$  in between  $v \rightarrow u$  so that it becomes  $v \rightarrow x \rightarrow u$ .)

The optimal tree decomposition with the minimum tree width is NP-hard to compute Arnborg et al. [1987]. In our experiments, we adopt the *vertex elimination* heuristic for generating tree decomposition Bodlaender et al. [2006]. This heuristic maps a permutation of the graph nodes into a tree decomposition. The pseudocode is included.

---

### Algorithm 1 \*

Vertex elimination heuristic for tree decomposition **Input:** An undirected graph  $G$  and an arbitrary permutation of the graph vertices

- 1: **for** vertex  $i$  in  $G$  ordered by the permutation **do**
  - 2:   let  $N(i)$  be  $i$ 's neighbours under the current graph
  - 3:   create tree node  $T_i = \{i\} \cup N(i)$
  - 4:   **for** every pair of vertices  $a, b \in N(i)$  **do**
  - 5:     add edge  $(a, b)$  to  $G$
  - 6:   remove  $i$  from  $G$
  - 7: **for**  $T_i$  from 1 to  $n$  **do**
  - 8:   let  $j = \max\{T_i/\{i\}\}$
  - 9:   connect tree node  $T_i$  to  $T_j$
- 

In our experiments, we use two heuristic orderings of the vertices. One is *minimum degree* (i.e., the next node is the node with the minimum degree in the current graph). The other is *minimum fill in* (i.e., the next node is the node that requires the minimum number of added edges). We try both heuristics and pick the tree decomposition with the smaller tree width. It should be noted that our algorithm works with any tree decomposition. It is just that in our experiments, we used the above two heuristics.

Figure 4 is an example attack graph, which will be used as a running example for TDCYCLE.

Figure 5 is a tree decomposition of Figure 4.

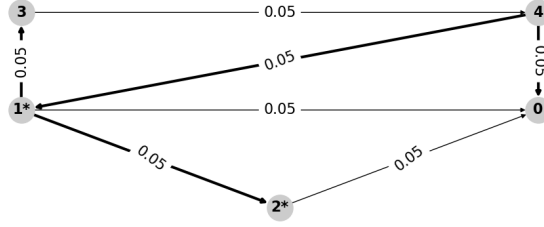


Figure 4: Example attack graph. Vertex 0 is DA. Vertex 1, 2 are entry vertices (marked using \*). Edge labels represent the edges’ failure rates. Thick edges (i.e.,  $1 \rightarrow 2$ ) are not blockable. We set a budget of 2.

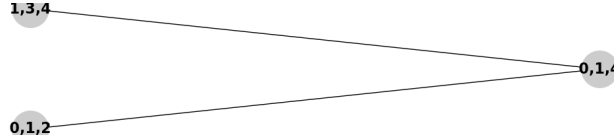


Figure 5: Tree decomposition of Figure 4. Both *minimum degree* and *minimum fill in* result in the same tree decomposition.

We then convert the resulting tree decomposition into a *nice tree decomposition* Cygan et al. [2015], as illustrated in Figure 6. The root node is a bag containing DA only and all the leaf nodes are bags of size one.<sup>7</sup> In a nice tree decomposition, there are only three types of nodes.

- **Introduce node:** An introduce node  $X$  has an only child  $X'$ .  $X' \subsetneq X$  and  $X \setminus X'$  has size 1. For example, in Figure 6,  $(3, 4)$  is an introduce node.
- **Forget node:** A forget node  $X$  has an only child  $X'$ .  $X \subsetneq X'$  and  $X' \setminus X$  has size 1. For example, in Figure 6,  $(1, 4)$  is a forget node.
- **Join node:** A join node  $X$  has two children.  $X$  must be identical to both children. For example, in Figure 6,  $(0, 1, 4)$  is a join node.

It is a trivial task to convert a tree decomposition into a nice tree decomposition. For example, in Figure 5,  $(0, 1, 4)$  splits into two branches. We just need to add two clones of it to make it a valid join node. As another example, in Figure 5,  $(1, 3, 4)$  is connected to  $(0, 1, 4)$  directly. By inserting  $(1, 4)$  in between,  $(0, 1, 4)$  (top-left one among the three clones) becomes an introduce node and  $(1, 4)$  becomes a forget node. A tree decomposition with tree width  $w$  can be converted to a nice tree decomposition with  $O(wn)$  nodes.

The whole point of nice tree decomposition is to divide the nodes into three categories, so that when we design DP, we just need to come up with three rules (one rule for each category of nodes).

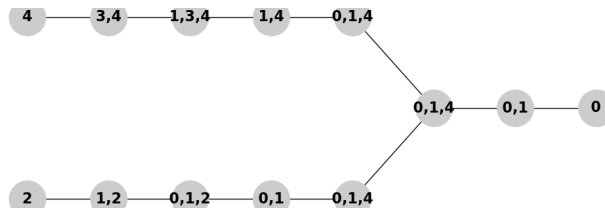


Figure 6: Nice tree decomposition built upon Figure 5.

## 7.5 TDCYCLE example based on Figure 6

Before discussing TDCYCLE in the context of Figure 6, we first prove Lemma 1.

<sup>7</sup>The standard nice tree decomposition definition involves empty bags. For example, a leaf node in our illustration contains one vertex, which can be interpreted as an introduce node by adding an auxiliary empty bag as its child. We ignore empty bags as they play no role in our DP.



*Proof of Lemma 1.* Let  $(u, v)$  be an edge in the attack graph. In any tree decomposition, there must be a bag  $X$  that contains both  $u$  and  $v$ . We check the parent of  $X$ . If the parent also contains both  $u$  and  $v$ , then we keep going up, until one of  $u$  or  $v$  is forgotten. (Both cannot be forgotten at the same node, since exactly one is forgotten at a forget node under a nice tree decomposition.) The forget node where either  $u$  or  $v$  is forgotten is the node we assign  $(u, v)$  to. The above forget node is unique for the following reasons. A known property of tree decomposition is that a node is forgotten at a unique forget node. The forget node for  $u$  and  $v$  must satisfy that either the forget node for  $u$  is an ancestor of the forget node for  $v$  (in which case  $(u, v)$  is assigned to the forget node for  $v$ ), or the reverse. This is because the subgraphs induced by  $u$  or  $v$  must be trees. That is, at the above unique forget node, we decide whether or not to block  $(u, v)$ .  $\square$

Recall that in our DP, for every node  $X$ , there is a corresponding subproblem  $DP(X)$ , which returns the collection of all possible tuples (distance matrix achieved and budget spent) after processing  $St(X)$ , where  $St(X)$  is the subtree rooted at  $X$ . In our example, all edges have the same failure rate 0.05, which translates to a distance of  $\delta = -\ln(1 - 0.05)$ .

- Base cases: There are only two leaf nodes (4) and (2). At leaf nodes, we have not put back any edges at all. 4 is not an entry vertex, so after processing (4), the only possible tuple is  $([\infty], 0)$ . That is,

$$DP((4)) = \{([\infty], 0)\}$$

Similarly, since 2 is an entry vertex, we have

$$DP((2)) = \{([0], 0)\}$$

- Original problem: We have set a budget of 2. The root node is (0). After processing all nodes, we have

$$DP((0)) = \{([3\delta], 2), ([\delta], 0)\}$$

That is, after processing all nodes in the nice tree decomposition, we end up with two possible tuples. We can spend 2 units of budget to ensure that the attacker's distance to 0 (DA) is  $3\delta$ , which translates to a success rate of  $(1 - 0.05)^3$ . We can also spend 0 units of budget to ensure that the attacker's distance to 0 (DA) is  $\delta$ , which translates to a success rate of  $1 - 0.05$ . Given a budget of 2, we conclude that the attacker's worst success rate is  $(1 - 0.05)^3$ .

In our implementation, we have a micro-optimisation step that filtered out  $([\delta], 1)$  as this tuple is worse off than  $([\delta], 0)$ . Given two different tuples  $(M_1, b_1)$  and  $(M_2, b_2)$ , if  $b_2 \geq b_1$  and  $M_1 \geq M_2$  (every element of  $M_1$  is larger or equal), then  $(M_2, b_2)$  is discarded.

- Introduce node: (3, 4) is an introduce node. 3 is introduced at this node, but its edges have not been put back yet, so 3 is disconnected. 3 is also not an entry vertex. Since  $DP((4)) = \{([\infty], 0)\}$ , we have

$$DP((3, 4)) = \left\{ \left( \begin{bmatrix} \infty & \infty \\ \infty & \infty \end{bmatrix}, 0 \right) \right\}$$

(1, 3, 4) is another introduce node. 1 is introduced at this node. Its edges have not been put back yet, but 1 is an entry vertex. So we have

$$DP((1, 3, 4)) = \left\{ \left( \begin{bmatrix} 0 & \infty & \infty \\ \infty & \infty & \infty \\ \infty & \infty & \infty \end{bmatrix}, 0 \right) \right\}$$

The 0 in the above matrix (top-left corner) indicates that the distance from 1 to an entry vertex is 0 (as 1 is an entry vertex itself).

- Forget node: (1, 4) is a forget node. 3 is forgotten here. That is, we need to consider putting back some edges. The two edges assigned to this forget node is  $1 \rightarrow 3$  and  $3 \rightarrow 4$ . We must put back  $1 \rightarrow 3$  as it is not blockable. For  $3 \rightarrow 4$ , we either block it or not. If we do not block  $3 \rightarrow 4$ , then we end up with the following possible tuple:

$$\left( \begin{bmatrix} 0 & 2\delta \\ \infty & 2\delta \end{bmatrix}, 0 \right)$$

The interpretation is that after putting back both  $1 \rightarrow 3$  and  $3 \rightarrow 4$ , the end result is as follows:

- top left 0: it takes 0 hops to go from an entry vertex to 1
- top right  $2\delta$ : it takes 2 hops to go from 1 to 4
- bottom left  $\infty$ : we cannot go from 4 to 1 (note: the edge  $4 \rightarrow 1$  has not been put back yet)
- bottom right  $2\delta$ : it takes 2 hops to go from an entry vertex to 4

If we block  $3 \rightarrow 4$ , then we end up with the following possible tuple:

$$\left( \begin{bmatrix} 0 & \infty \\ \infty & \infty \end{bmatrix}, 1 \right)$$

In summary, we have

$$DP((1, 4)) = \left\{ \left( \begin{bmatrix} 0 & 2\delta \\ \infty & 2\delta \end{bmatrix}, 0 \right), \left( \begin{bmatrix} 0 & \infty \\ \infty & \infty \end{bmatrix}, 1 \right) \right\}$$

- **Join node:** In Figure 6, there are three clones of  $(0, 1, 4)$ . We use  $(0, 1, 4)^R$ ,  $(0, 1, 4)^T$  and  $(0, 1, 4)^B$  to denote the one on the right, top and bottom, respectively.

We directly present the values of  $DP((0, 1, 4)^T)$  and  $DP((0, 1, 4)^B)$ .

$DP((0, 1, 4)^T)$  is

$$\left\{ \left( \begin{bmatrix} \infty & \infty & \infty \\ \infty & 0 & 2\delta \\ \infty & \infty & 2\delta \end{bmatrix}, 0 \right), \left( \begin{bmatrix} \infty & \infty & \infty \\ \infty & 0 & \infty \\ \infty & \infty & \infty \end{bmatrix}, 1 \right) \right\}$$

$DP((0, 1, 4)^B)$  is

$$\left\{ \left( \begin{bmatrix} \delta & \infty & \infty \\ 2\delta & 0 & \infty \\ \infty & \infty & \infty \end{bmatrix}, 0 \right), \left( \begin{bmatrix} \infty & \infty & \infty \\ \infty & 0 & \infty \\ \infty & \infty & \infty \end{bmatrix}, 1 \right) \right\}$$

$DP((0, 1, 4)^R)$  is then the aggregation of the above. We take one tuple  $(M^T, b^T)$  from  $DP((0, 1, 4)^T)$  and one tuple  $(M^B, b^B)$  from  $DP((0, 1, 4)^B)$ . We add  $(\min(M^T, M^B), b^T + b^B)$  into  $DP((0, 1, 4)^R)$ , provided that  $b^T + b^B \leq b$ .  $DP((0, 1, 4)^R)$  is

$$\left\{ \left( \begin{bmatrix} \delta & \infty & \infty \\ 2\delta & 0 & 2\delta \\ \infty & \infty & 2\delta \end{bmatrix}, 0 \right), \left( \begin{bmatrix} \delta & \infty & \infty \\ 2\delta & 0 & \infty \\ \infty & \infty & \infty \end{bmatrix}, 1 \right), \right. \\ \left. \left( \begin{bmatrix} \infty & \infty & \infty \\ \infty & 0 & 2\delta \\ \infty & \infty & 2\delta \end{bmatrix}, 1 \right), \left( \begin{bmatrix} \infty & \infty & \infty \\ \infty & 0 & \infty \\ \infty & \infty & \infty \end{bmatrix}, 2 \right) \right\}$$

## 7.6 Pseudocode of TDCYCLE

We present the pseudocode of TDCYCLE in this section.

---

### Algorithm 2 \*

TDCYCLE base cases setup **Input:** Nice tree decomposition TD

- 1: **for** leaf node  $X = \{x\}$  in TD **do**
  - 2:   **if**  $x$  is an entry vertex **then**
  - 3:      $DP(X) = \{([0], 0)\}$
  - 4:   **else**
  - 5:      $DP(X) = \{([\infty], 0)\}$
- 

---

### Algorithm 3 \*

TDCYCLE final step after finishing DP **Input:**  $DP((DA))$

- 1: **for**  $([d_{DA, DA}], b')$  in  $DP((DA))$  **do**
  - 2:   Keep track of the maximum  $d_{DA, DA}$
  - 3: **Return**  $1 - e^{-d_{DA, DA}}$
-

---

**Algorithm 4 \***


---

 TDCYCLE at introduce node **Input:** Nice tree decomposition TD

 Introduce node  $X = (x_1, \dots, x_k, y)$ 
 $X$ 's child  $X' = (x_1, \dots, x_k)$ 

- 1: **if**  $y$  is an entry vertex **then**
- 2:      $d_{yy} = 0$
- 3: **else**
- 4:      $d_{yy} = \infty$
- 5: **for** possible tuple in  $DP(X')$  **do**
- 6:     convert the tuple on the left (from  $DP(X')$ ) to the tuple on the right (added into  $DP(X)$ ):

$$\left( \begin{bmatrix} d_{11} & \dots & d_{1k} \\ \dots & & \\ d_{k1} & \dots & d_{kk} \end{bmatrix}, b' \right) \rightarrow \left( \begin{bmatrix} d_{11} & \dots & d_{1k} & \infty \\ \dots & & & \\ d_{k1} & \dots & d_{kk} & \infty \\ \infty & \dots & \infty & d_{yy} \end{bmatrix}, b' \right)$$


---

**Algorithm 5 \***


---

 TDCYCLE at forget node **Input:** Nice tree decomposition TD

 Forget node  $X = (x_2, \dots, x_k)$ 
 $X$ 's child  $X' = (x_1, \dots, x_k)$ 

- 1: generate all possible blocking policies for edges between  $x_1$  and one of  $x_2, \dots, x_k$  (at most  $2^{k-1}$  options)
- 2: **for** blocking policy with budget  $b''$  **do**
- 3:     **for** possible tuple  $(M, b')$  in  $DP(X')$  **do**
- 4:         **if**  $b' + b'' > b$  **then**
- 5:             continue
- 6:     convert the tuple on the left (from  $DP(X')$ ) to the tuple on the right (added into  $DP(X)$ ):

$$\left( \begin{bmatrix} d_{11} & \dots & d_{1k} \\ \dots & & \\ d_{k1} & \dots & d_{kk} \end{bmatrix}, b' \right) \rightarrow \left( \begin{bmatrix} d'_{22} & \dots & d'_{2k} \\ \dots & & \\ d'_{k2} & \dots & d'_{kk} \end{bmatrix}, b' + b'' \right)$$

- 7:     call *all-pair shortest path* routine to calculate the  $d'_{ij}$ , which are updated distances considering the newly put back edges under the current blocking policy.
- 

**Algorithm 6 \***


---

 TDCYCLE at join node **Input:** Nice tree decomposition TD

 Join node  $X$ 

 Two children  $X_1$  and  $X_2$ 

- 1: **for** possible tuple  $(M_1, b_1)$  in  $DP(X_1)$  **do**
  - 2:     **for** possible tuple  $(M_2, b_2)$  in  $DP(X_2)$  **do**
  - 3:         **if**  $b_1 + b_2 \leq b$  **then**
  - 4:             add  $(M', b_1 + b_2)$  into  $DP(X)$ , where  $M'$  is the element-wise minimum between  $M_1$  and  $M_2$
- 

## 7.7 Complexity of TDCYCLE

*Proof of Theorem 2.* There are  $O(wn)$  DP subproblems in total, which fall into three categories as follows. The return value of  $DP(X)$  is a collection. We use a binary array to represent this collection. The size of the array is  $H^{w^2}(b+1)$ , or simply  $O(H^{w^2}b)$ .<sup>8</sup>

- Introduce node: We need to go through the child node's output array (at most  $O(H^{w^2}b)$  tuples) and expand the distance matrix by one row and one column. The complexity is  $O(w^2)$  for each tuple. There are at most  $O(wn)$  introduce nodes. So the complexity for handling all introduce nodes is  $O(H^{w^2}bw^3n)$ .

---

<sup>8</sup>In experiments, we used Python's builtin set instead, because most tuples are not possible.

- Forget node: We need to go through the child node’s output array (at most  $O(H^{w^2}b)$  tuples) and shrink the distance matrix by one row and one column. The complexity is  $O(2^w w^3)$  for each tuple as we need to run an all pair shortest path for each blocking option and there are  $O(2^w)$  options. There are at most  $O(n)$  forget nodes (a vertex is forgotten once). So the complexity for handling all forget nodes is  $O(2^w H^{w^2} b w^3 n)$ .
- Join node: We need to go through both children nodes’ output arrays (at most  $O(H^{2w^2} b^2)$  pairs of tuples). We need to perform an element-wise min operation on the distance matrices, which has a complexity of  $O(w^2)$ . There are at most  $O(n)$  join nodes. So the complexity for handling all join nodes is  $O(H^{2w^2} b^2 w^2 n)$ .

We do not need to run an all-pair shortest path update at a join node. If the join node has a bag size of 1, then all-pair shortest path is not necessary. If the bag size is more than 1, then there will be forget nodes between the join node and root. We could wait until we reach a forget node to run the all-pair shortest path process.

Assuming  $H$  is at least 2, the overall complexity is then

$$\begin{aligned} & O(2^w H^{w^2} b w^3 n + H^{2w^2} b^2 w^2 n) \\ &= O((2^w w + H^{w^2} b) H^{w^2} b w^2 n) = O(H^{2w^2} b^2 w^2 n) \end{aligned}$$

□

It should be noted that the above is the worst case complexity. In experiments, the set size of  $DP(X)$  is often quite small (i.e., in the hundreds for R2000).

### 7.8 $-\ln(1-x)$ , when restricted to a small interval, can be approximated using two straight lines (one upper bound and one lower bound)

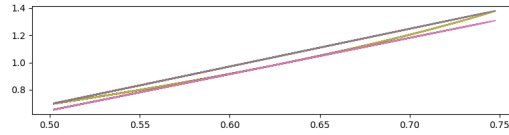


Figure 7:  $-\ln(1-x)$  on  $[0.5, 0.75]$ , bounded by straight lines

### 7.9 Reproducibility notes and preprocessing steps

We have included our source code as well as the synthetic AD graphs. R2000 and R4000 were generated using DBCREATOR by setting the numbers of computers to 2000 and 4000, respectively. ADS10 was generated using ADSIMULATOR with the following parameters: nComputers=1000, nUsers=1000, nOUs=200, nGroups=1000, nGPOs=200.

Both tools are open source and are available at:

<https://github.com/BloodHoundAD/BloodHound-Tools/tree/master/DBCcreator>

<https://github.com/nicolas-carolo/adsimulator>

In our experiments, we run 10 trials with different random setups (randomly picked blockable edges and entry nodes). The trials are using fixed random seeds from 0 to 9. For MIP-FEASIBLE and MIP-LB, we used 10 regions to approximate  $-\ln(1-x)$ :  $[0, 0.1]$ ,  $[0.1, 0.2]$ ,  $\dots$ ,  $[0.9, 0.99]$ .

Before running our algorithms, we perform the following preprocessing steps:

- In both R2000 and R4000, there are 7 admin nodes. In ADS10, there are 43 admin nodes (nodes whose ADMINCOUNT property is true). We merge them into a single destination node and set it as DA. We ignore all out-going edges of DA. (Once the attacker reaches DA, the attack has finished.)
- We delete all nodes that cannot reach DA.
- We delete all incoming edges to entry nodes, because the optimal attack path will never involve such edges.
- We delete non-entry nodes that have 0 in-degrees, because the attacker cannot start from these nodes or reach these nodes.

- If an edge is not block-worthy (Lemma 2), then we mark it as not blockable.
- For a splitting node  $u$ , if there are two non-splitting paths  $\text{NSP}(u, v_1)$  and  $\text{NSP}(u, v_2)$  that satisfy: 1) these two paths have the same end node ( $\text{DEST}(u, v_1) = \text{DEST}(u, v_2)$ ); 2) the first path has a higher success rate ( $c_{u, v_1} \geq c_{u, v_2}$ ); 3) The first path  $\text{NSP}(u, v_1)$  is not blockable, then the attacker always prefers the first path (facing any defense). We delete  $u \rightarrow v_2$  as it is never used.

### 7.10 Proximal Policy Optimization parameters and implementation details

We used the PPO implementation from the popular TIANSHOU library, with the following parameters. Our code is also included.

learning rate	0.001
discount factor	0.99
batch size	64
epoch	refer to experiment description
step per epoch	10000
step per collect	2000
repeat per collect	10
hidden sizes	[128, 128] for R2000 [512, 512] for others

The observation space is a tuple where every coordinate is between 0 and 1. An observation consists of the following:

- Distance matrix before blocking: The length is the square of the maximum tree decomposition bag size. Every matrix element is expressed in terms of probability so it is between 0 and 1. If the current bag size is less than the maximum bag size, then we only use the “top left” corner. The remaining slots are filled in with 0s (equivalent to adding auxiliary nodes that not connected to any other nodes).
- Distance matrix after blocking.
- Budget spent: if the total budget is 5 and 2 is left, then we append  $(1, 1, 0, 0, 0)$  to the observation.
- Current step index: if the total episode length is 5 and we have finished 2 steps, then we append  $(1, 1, 0, 0, 0)$  to the observation.

### 7.11 Experiments showing that our RL based approach is not merely performing random searching and the tree decomposition based distance matrix indeed provides a useful representation that facilitates learning

We conduct experiments by running reinforcement learning on the original observation space with the tree decomposition based distance matrix. We then repeat the experiments using the same setups and the same random seeds, but we replace the actual distance matrix by the **zero matrix** and by the **random matrix**. For a fair comparison, we do not modify the representation for the budget spent and the current step index. Our experiments show that learning significantly deteriorates after the replacements.

We design the following AD graph for our experiments. Node 0 is DA. Besides DA, we have 40 entry nodes from node 1 to node 40. Node  $i$  ( $1 \leq i \leq 40$ ) has one *blockable* edge pointing to DA. These are all the edges in the graph. We set a budget of 20. That is, we need to block half of the edges. For  $i$  that is **even**, the failure rate of edge  $i \rightarrow \text{DA}$  is  $\frac{40}{41}$ , which represents an edge that does not need to be blocked as the failure rate is high. For  $i$  that is **odd**, the failure rate of edge  $i \rightarrow \text{DA}$  is  $\frac{i}{41}$  ( $\frac{1}{41} \leq \frac{i}{41} \leq \frac{39}{41}$ ), which represents an edge that needs to be blocked considering that our budget is 20. The design motivation of this AD graph is as follows:

- The task is to pick out 20 “correct” edges to block out of 40 edges, which has a large search space (over  $10^{11}$ ).
- The attacker’s best success rate is  $\frac{40}{41}$ , which happens as long as  $1 \rightarrow \text{DA}$  is not blocked. That is, with this one mistake, we end up with the worst performance. The attacker’s worst success rate is  $\frac{1}{41}$ , which only happens when every edge blocked is “correct”. There is a wide range of possible results between  $\frac{1}{41}$  and  $\frac{40}{41}$  and it gets *increasingly difficult* to achieve better performance.
- We *intertwine* the block-worthy and not-block-worthy edges so that we do not accidentally learn toward policies such as “spend all budget immediately” (if all block-worthy edges have lower indices) or “spend all budget at the end” (if all not-block-worthy edges have lower indices). It should be noted that the edges’ indices are not included in the observation, so we will not accidentally learn toward “block odd edges”.

We use ORIGINAL to represent the original approach with the distance matrix intact. We use ZERO and RANDOM to represent the results after replacing the distance matrix using the zero matrix and the random matrix (every matrix element independently drawn from  $U(0, 1)$ ). We record the worst success rate for the attacker after 10 and 50 epochs. Experiments are repeated 10 times using training seed from 0 to 9. The table cells are the averages and the standard deviations over these 10 trials. Clearly, the tree decomposition based distance matrix provides a useful representation that facilitates learning (smaller numbers are better).

Epochs	ORIGINAL	ZERO	RANDOM
10	<b>0.298 ± 0.071</b>	0.492 ± 0.063	0.478 ± 0.048
50	<b>0.117 ± 0.024</b>	0.371 ± 0.039	0.415 ± 0.040

Lastly, we run experiments on R2000 with budget 5. We use the same experimental setup as before (as described in the experiment section). We randomly draw the entry nodes and the blockable edges using random seed 0 to 9. Experiments are repeated 10 times using training seed from 0 to 9. We record the worst success rate for the attacker after 5 and 10 epochs (due to time constraint – as we are running three experiments for 10 random graphs and for 10 training seeds, which in total equals 300 experiments). For seed 1, 4, 6, 7, 8, 9, every trial reaches optimality within 5 epochs, so we only present the result for seed 0, 2, 3, 5. The numbers do not differ much because 1) the gap between the optimal result and the greedy result is tiny to begin with; 2) the search space isn't large (given enough time, random searching often can find the optimal result for this graph). Still, we see that the original approach always produces the best result.

R2000 $b = 5$ 5 epochs	ORIGINAL	ZERO	RANDOM
Seed 0	<b>0.521 ± 0</b>	<b>0.521 ± 0</b>	<b>0.521 ± 0</b>
Seed 2	<b>0.521 ± 0</b>	<b>0.521 ± 0</b>	<b>0.521 ± 0</b>
Seed 3	<b>0.406 ± 0.030</b>	0.416 ± 0.040	0.506 ± 0.021
Seed 5	<b>0.392 ± 0.008</b>	0.394 ± 0.006	0.446 ± 0.050

R2000 $b = 5$ 10 epochs	ORIGINAL	ZERO	RANDOM
Seed 0	<b>0.521 ± 0</b>	<b>0.521 ± 0</b>	<b>0.521 ± 0</b>
Seed 2	<b>0.504 ± 0.043</b>	0.521 ± 0	0.521 ± 0
Seed 3	<b>0.396 ± 0</b>	<b>0.396 ± 0</b>	0.406 ± 0.030
Seed 5	<b>0.388 ± 0.010</b>	0.390 ± 0.009	0.396 ± 0

## 7.12 Description of the kernel

Due to space constraint, we did not explicitly mention what the kernel is in our section on kernelization. Also, for presentation purpose, we want to avoid confusion by bringing up a different graph. For completeness, the kernel is as follows: Only keep splitting nodes, entry nodes and DA. Delete all original edges. Replace every original non-splitting path by one edge. The edge's success rate is the unblocked success rate of the original non-splitting path. The edge is blockable if and only if the original non-splitting is blockable.

## 7.13 More on TDCYCLE vs IP

As shown in our experiments, IP scales extremely well and performs much better than TDCYCLE on all three graphs generated using two different open source AD graph generators. We expect TDCYCLE to perform relatively better if we are dealing with a graph with a small tree width and a large number of non-splitting paths. As an example, let us consider the following contrived attack graph. There is one entry node called ENTRY and one DA. We insert the following edges: ENTRY  $\rightarrow i \rightarrow$  DA for  $i$  from 1 to  $n - 2$  (that is, there are in total  $n$  nodes). All edges are set to be blockable and have the same failure rates. This attack graph has a tree width of 2. The maximum attack path length  $l$  is 2. The number of possible success rates for attack paths  $H$  is 4. With the above set up, TDCYCLE's complexity becomes  $O(b^2n)$ . On the contrary, for IP, we face an integer program with  $O(n)$  binary variables.

Lastly, TDCYCLE has two advantages over IP:

- TDCYCLE is *embarrassingly parallel*. That is, if we throw 100 CPU cores to it, we expect about 100 times speed up. For all three types of nodes (introduce/forget/join), the main calculation is a for loop that can be trivially made parallel. Given a node  $X$ , we just need to use a shared memory array to store the result for  $DP(X)$ . All parallel processes can write to this shared array without worrying about race condition.

On the contrary, IP is difficult to be made parallel.

- TDCYCLE does not require that we load the whole model into the memory. Memory paging is convenient under TDCYCLE. The dynamic program is in a bottom-up fashion. Node  $X$  in the tree decomposition requires a memory block to store its result  $DP(X)$ . For node  $X$ , we only need to allocate memory for it once we reach it. Also, after processing a node, we can safely delete the memory blocks for the children nodes.

On the contrary, memory paging is difficult for IP.