Intrusion Tolerance for Networked Systems Through Two-Level Feedback Control
Supplementary Material

Author names withheld for double-blind reviewing.

Appendix A
Belief Computation

The belief state for each node $i \in \mathcal{N}_t$ can be computed recursively as

$$
\begin{aligned}
b_{i,t}(s_{i,t}) &\overset{(a)}{=} \mathbb{P}\big[s_{i,t} \mid \overbrace{o_{i,t}, a_{i,t-1}, o_{i,t-1}, \ldots, a_{i,1}, o_{i,1}, b_{i,1}}^{\mathbf{h}_{i,t}}\big] \\
&\overset{(b)}{=} \frac{\mathbb{P}\big[o_{i,t} \mid s_{i,t}, a_{i,t-1}, \mathbf{h}_{i,t-1}\big]\mathbb{P}\big[s_{i,t} \mid a_{i,t-1}, \mathbf{h}_{i,t-1}\big]}{\mathbb{P}\big[o_{i,t} \mid a_{i,t-1}, \mathbf{h}_{i,t-1}\big]} \\
&\overset{(c)}{=} \frac{Z(o_{i,t} \mid s_{i,t})\mathbb{P}\big[s_{i,t} \mid a_{i,t-1}, \mathbf{h}_{i,t-1}\big]}{\mathbb{P}\big[o_{i,t} \mid a_{i,t-1}\mathbf{h}_{i,t-1}\big]} \\
&\overset{(d)}{=} \frac{Z(o_{i,t} \mid s_{i,t}) \sum_{s \in \mathcal{S}} b_{i,t-1}(s) f_{\mathrm{N}}(s_{i,t} \mid s, a_{t-1})}{\mathbb{P}\big[o_{i,t} \mid a_{i,t-1}\mathbf{h}_{i,t-1}\big]} \\
&\overset{(e)}{=} \frac{Z(o_{i,t} \mid s_{i,t}) \sum_{s \in \mathcal{S}} b_{i,t-1}(s) f_{\mathrm{N}}(s_{i,t} \mid s, a_{t-1})}{\sum_{s', s \in \mathcal{S}} Z(o_{i,t} \mid s') f_{\mathrm{N}}(s' \mid s, a_{i,t-1}) b_{i,t-1}(s)}.
\end{aligned}
$$

(a) follows from the definition of $b_{i,t}$ (4); (b) is an expansion of the conditional probability using Bayes rule; (c) follows from the Markov property of Z (3); (d) follows from the Markov property of f_{N} (2); and (e) follows by definition of Z (3) and f_{N} (2). Computing the belief state through the expression in (e) requires $O(|\mathcal{S}|^2)$ scalar multiplications.

Appendix B
Proof of Theorem 1

Solving (6) corresponds to solving N_t finite, stationary, and constrained Partially Observed Markov Decision Processes (POMDPs) with bounded costs and the average cost optimality criterion. Since the POMDPs are equivalent it suffices to prove the statement for a single POMDP.

It follows from (2) and assumption A that $f_{\mathrm{N}}(s' \mid s, a) > 0$ for all t, s', s, a. Given this property and assumption D, we know that there exists a deterministic optimal strategy $\pi^{\star}_{i,t}$ for which the limit in (5) exists and which satisfies

$$
\pi^{\star}_{i,t}(b_{i,t}) \in \operatorname*{arg\,min}_{a \in \{\mathfrak{W}, \mathfrak{R}\}} \left[c_{\mathrm{N}}(b_{i,t}, a) + \sum_{o \in \mathcal{O}} \mathbb{P}\left[o | b_{i,t}, a\right] V^{\star}_{i,t}(b_{i,t+1}) \right] \tag{17}
$$

for all $b_{i,t}$ and $t \geq 1$ [1, Prop. 1], where $c_{\mathrm{N}}(b_{i,t}, a)$ is the expected immediate cost of a given $b_{i,t}$ and $V^{\star}_{i,t}$ is the value function [2, Thm. 7.4.1].

Each of the constrained POMDPs with infinite horizons defined in (6) can be converted into a sequence of unconstrained POMDPs $(\mathcal{M}_{i,k})_{k=1,2,\ldots}$ with horizon $T = \Delta_{\mathrm{R}}$, where $a_{i,T} = \mathfrak{R}$ ensures that (6b) is satisfied. This sequence is equivalent to the original POMDP because

$$
\begin{aligned}
&\operatorname*{arg\,min}_{\pi_{i,t}} \left[\lim_{T \to \infty} \mathbb{E}_{\pi_{i,t}} \left[\frac{1}{T} \sum_{t=1}^{T} C_{i,t} \mid B_{i,1} = p_{\mathrm{A}} \right] \right] \\
&\overset{(a)}{=} \operatorname*{arg\,min}_{\pi_{i,t}} \left[\lim_{T \to \infty} \frac{1}{T} \left(\mathbb{E}_{\pi_{i,t}} \left[\sum_{t=1}^{\Delta_{\mathrm{R}}} C_{i,t} \mid B_{i,1} = p_{\mathrm{A}} \right] + \right. \right. \\
&\qquad \left. \left. \mathbb{E}_{\pi_{i,t}} \left[\sum_{t=\Delta_{\mathrm{R}}}^{2\Delta_{\mathrm{R}}} C_{i,t} \mid B_{i,\Delta_{\mathrm{R}}} = p_{\mathrm{A}} \right] + \ldots \right) \right] \\
&\overset{(b)}{=} \operatorname*{arg\,min}_{\pi_{i,t}} \left[\lim_{T \to \Delta_{\mathrm{R}}} \mathbb{E}_{\pi_{i,t}} \left[\frac{1}{T} \sum_{t=1}^{\Delta_{\mathrm{R}}} C_{i,t} \mid B_{i,1} = p_{\mathrm{A}} \right] \right], \qquad (18)
\end{aligned}
$$

where $C_{i,t}$ is a random variable representing the cost of node i at time t; (a) follows from linearity of $\mathbb{E}$; (b) follows because all elements inside the parentheses are equivalent, which means that a strategy that minimizes one element minimizes the whole expression.

Consider the threshold structure in (7). We know that a strategy $\pi^{\star}_i$ that achieves the minimization in (18) induces a partition of $[0, 1]$ into two regions at each time t: a wait region $\mathcal{W}_t$ where $\pi^{\star}_i(b) = \mathfrak{W}$, and a recovery region $\mathcal{R}_t$ where $\pi^{\star}_i(b) = \mathfrak{R}$. The idea behind the proof of (7) is to show that $\mathcal{R}_t = [\alpha^{\star}_t, 1]$ for all t and some thresholds $(\alpha^{\star}_t)_{t=1,\ldots,T}$. Towards this deduction, note that $\mathcal{W}_t$ and $\mathcal{R}_t$ are connected sets [2, Thm. 12.3.4]. This follows because (*i*) the transition and observation matrices are TP-2 [2, Def. 10.2.1] (it is a consequence of assumptions A–C, and E); (*ii*) $c_{\mathrm{N}}(s_{i,t}, a_{i,t})$ (5) is submodular [2, Def. 12.3.2]; and (*iii*) $c_{\mathrm{N}}(s_{i,t}, a_{i,t})$ is weakly increasing in $s_{i,t}$ for each $a_{i,t}$.

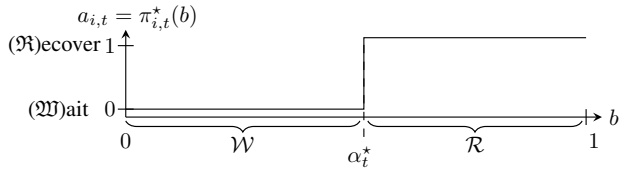
As $\mathcal{R}_t$ is connected, $1 \in \mathcal{R}_t \iff \mathcal{R}_t = [\alpha^{\star}_t, 1]$. Hence it suffices to show that $1 \in \mathcal{R}_t$. We obtain from (17) that

$$
\begin{aligned}
1 \in \mathcal{R}_t \iff &\mathbb{E}_{B_{i,t+1}} \left[V^{\star}_{i,t+1}(B_{i,t+1}) \mid A_{i,t} = \mathfrak{R}, B_{i,t} = 1 \right] \leq \\
&\mathbb{E}_{B_{i,t+1}} \left[V^{\star}_{i,t+1}(B_{i,t+1}) \mid A_{i,t} = \mathfrak{W}, B_{i,t} = 1 \right].
\end{aligned}
$$

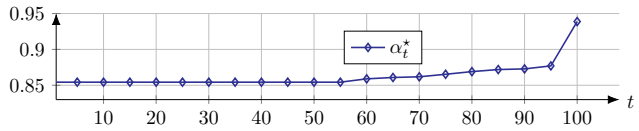
Clearly

$$
\begin{aligned}
&\mathbb{E}_{B_{i,t+1}} \left[B_{i,t+1} \mid A_{i,t} = \mathfrak{R}, B_{i,t} = 1 \right] \leq \\
&\mathbb{E}_{B_{i,t+1}} \left[B_{i,t+1} \mid A_{i,t} = \mathfrak{W}, B_{i,t} = 1 \right].
\end{aligned}
$$

Further $B' \leq B \implies V^{\star}_{i,t+1}(B') \leq V^{\star}_{i,t+1}(B)$ for all $i \in \mathcal{N}_t$ and $t \geq 1$ [2, Thm. 11.2.1], which implies that $1 \in \mathcal{R}_t$ for all t (see Fig. 14). □



(a) Structure of an optimal threshold recovery strategy $\pi^\star_{i,t}$; $\mathcal{W}$ and $\mathcal{R}$ denote the wait and recovery sets.



(b) Optimal recovery thresholds $\alpha^\star_t$, where $\Delta_R = 100$.

Fig. 14: Illustration of Thm. 1 and Cor. 1; the parameters for computing the figures are listed in App. E.

Appendix C
Proof of Corollary 1

When $\Delta_{\mathrm{R}} \to \infty$, (6) reduces to a set of unconstrained stationary and finite POMDPs, which means that there exists an optimal deterministic stationary strategy for each node [1, Prop. 1]. Such a strategy partitions the belief space into two time-independent regions $\mathcal{W}$ and $\mathcal{R}$, which means that the recovery threshold $\alpha^\star$ is time-independent (Thm. 1).

When $\Delta_{\mathrm{R}} < \infty$, it follows from (17) that it is optimal to recover node $i \in \mathcal{N}_t$ at time t iff

$$\begin{aligned}
&c_{\mathrm{N}}(b_{i,t}, \mathfrak{R}) + \mathbb{E}_{B_{i,t+1}}[V^\star_{i,t}(B_{i,t+1}) \mid a_t = \mathfrak{R}, b_{i,t}] \leq \\
&c_{\mathrm{N}}(b_{i,t}, \mathfrak{W}) + \mathbb{E}_{B_{i,t+1}}[V^\star_{i,t}(B_{i,t+1}) \mid a_t = \mathfrak{W}, b_{i,t}] \\
&\iff 1 \leq \eta b_{i,t} + W_{i,t}(b_{i,t}) \\
&\iff b_{i,t} \geq \underbrace{\frac{1 - W_{i,t}(b_{i,t})}{\eta}}_{\alpha^\star_t},
\end{aligned}$$

where $W_{i,t}(b_{i,t}) \triangleq \mathbb{E}_{B_{i,t+1}}[V^\star_{i,t+1}(B_{i,t+1}) \mid a_{i,t} = \mathfrak{W}, b_{i,t}] - \mathbb{E}_{B_{i,t+1}}[V^\star_{i,t+1}(B_{i,t+1}) \mid a_{i,t} = \mathfrak{R}, b_{i,t}]$

Hence $\alpha^\star_t \leq \alpha^\star_{t+1}$ iff $W_{i,t}$ is non-increasing in t for all $b_{i,t}$ and i. We prove this using mathematical induction on $k = T, T-2, \ldots, 1$. For $k = T-1$ we have $W_{i,T-1}(b_{i,T-1}) = 0$ for all b and i. Next, for $k = T-2$ we have

$$\begin{aligned}
&W_{i,T-2}(b_{T-2}) = \min\left[1 + \mathbb{E}_{B_{i,T}}[V^\star_{i,T}(B_{i,T}) \mid a_{T-1} = \mathfrak{R}],\right. \\
&\mathbb{E}_{B_{i,T-1},B_{i,T}}[\eta B_{i,T-1} + V^\star_{i,T}(B_{i,T}) \mid a_{T-2} = a_{T-1} = \\
&\qquad \left.\mathfrak{W}, b_{T-2}]\right] - \min\left[1 - \mathbb{E}_{B_{i,T}}[V^\star_{i,T}(B_{i,T}) \mid a_{T-1} = \mathfrak{R}],\right. \\
&\left.\mathbb{E}_{B_{i,T-1},B_{i,T}}[\eta B_{i,T-1} + V^\star_{i,T}(B_{i,T}) \right. \\
&\qquad \left. \mid a_{T-2} = \mathfrak{R}, a_{T-1} = \mathfrak{W}, b_{T-2}]\right] \\
&\overset{(a)}{\geq} 0 = W_{i,T-1}(b_{i,T-2}),
\end{aligned}$$

where (a) follows because $\mathbb{E}[B' \mid a_t = \mathfrak{W}, b] \geq \mathbb{E}[B' \mid a_t = \mathfrak{R}, b]$ by definition (2).

Assume by induction that $W_{i,k}(b) \geq W_{i,k+1}(b)$ for $k = t, t+1, \ldots, T-3$ and all b and i. We will show that this assumption implies $W_{i,k-1}(b) \geq W_{i,k}(b)$ for all b and i.

There are three cases to consider:

1) If $B_k \in \mathcal{R}$ both when $a_{i,k-1} = \mathfrak{W}$ and when $a_{i,k-1} = \mathfrak{R}$, then

$$\begin{aligned}
W_{i,k-1}(b_{i,k-1}) &= \mathbb{E}_{B_k}[V^\star_{i,k}(B_k) \mid a_{i,k-1} = \mathfrak{W}, b_{i,k-1}] - \\
&\quad \mathbb{E}_{B_k}[V^\star_{i,k}(B_k) \mid a_{i,k-1} = \mathfrak{R}, b_{i,k-1}] \\
&\overset{(a)}{=} \mathbb{E}_{B_{k+1}}[1 + V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{R}] - \\
&\quad \mathbb{E}_{B_k+1}[1 + V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{R}] \\
&\overset{(b)}{=} \mathbb{E}_{B_{k+1}}[V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{R}] - \\
&\quad \mathbb{E}_{B_{k+1}}[V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{R}] \overset{(c)}{=} W_{i,k}(b_{i,k-1}),
\end{aligned}$$

where (a) follows from (17).

2) If $B_k \in \mathcal{W}$ both when $a_{i,k-1} = \mathfrak{W}$ and when $a_{i,k-1} = \mathfrak{R}$, then

$$\begin{aligned}
W_{i,k-1}(b_{i,k-1}) &= \mathbb{E}_{B_k}[V^\star_{i,k}(B_k) \mid a_{i,k-1} = \mathfrak{W}, b_{i,k-1}] - \\
&\quad \mathbb{E}_{B_k}[V^\star_{i,k}(B_k) \mid a_{i,k-1} = \mathfrak{R}, b_{i,k-1}] \\
&\overset{(a)}{=} \mathbb{E}_{B_k,B_{k+1}}[\eta B_k + V^\star_{i,k+1}(B_{k+1}) \mid \\
&\qquad a_{i,k} = a_{i,k-1} = \mathfrak{W}, b_{i,k-1}] - \\
&\quad \mathbb{E}_{B_k,B_{k+1}}[\eta B_k + V^\star_{i,k+1}(B_{k+1}) \mid \\
&\qquad a_{i,k} = \mathfrak{W}, a_{i,k-1} = \mathfrak{R}, b_{i,k-1}] \\
&\overset{(b)}{\geq} \mathbb{E}_{B_{k+1}}[V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = a_{i,k-1} = \mathfrak{W}, b_{i,k-1}] - \\
&\quad \mathbb{E}_{B_{k+1}}[V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{W}, a_{i,k-1} = \mathfrak{R}, b_{i,k-1}] \\
&\overset{(c)}{=} W_{i,k}(b_{i,k-1}),
\end{aligned}$$

where (b) follows because $\mathbb{E}[B' \mid a = \mathfrak{W}, b] \geq \mathbb{E}[B' \mid a = \mathfrak{R}, b]$ by definition (2).

3) If $B_k \in \mathcal{R}$ when $a_{i,k-1} = \mathfrak{W}$, and $B_k \in \mathcal{W}$ when $a_{i,k-1} = \mathfrak{R}$, then

$$\begin{aligned}
W_{i,k-1}(b_{i,k-1}) &= \mathbb{E}_{B_k}[V^\star_{i,k}(B_k) \mid a_{i,k-1} = \mathfrak{W}, b_{i,k-1}] - \\
&\quad \mathbb{E}_{B_k}[V^\star_{i,k}(B_k) \mid a_{i,k-1} = \mathfrak{R}, b_{i,k-1}] \\
&\overset{(a)}{=} \mathbb{E}_{B_{k+1}}[1 + V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{R}] - \\
&\mathbb{E}_{B_k,B_{k+1}}[\eta B_k + V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{W}, a_{i,k-1} = \mathfrak{R}] \\
&\overset{(b)}{\geq} \mathbb{E}_{B_{k+1}}[1 + V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{R}] - \\
&\quad \mathbb{E}_{B_{k+1}}[1 + V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{R}] \\
&\overset{(c)}{\geq} \mathbb{E}_{B_{k+1}}[V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{R}] - \\
&\quad \mathbb{E}_{B_{k+1}}[V^\star_{i,k+1}(B_{k+1}) \mid a_{i,k} = \mathfrak{R}] \overset{(d)}{=} W_{i,k}(b_{i,k-1}),
\end{aligned}$$

where (b) follows from (17).

The case where $a_{i,k-1} = \mathfrak{R} \implies B_k \in \mathcal{R}$ and $a_{i,k-1} = \mathfrak{W} \implies B_k \in \mathcal{W}$ can be discarded due to Thm. 1 since $\mathbb{E}[B' \mid a = \mathfrak{W}, b] \geq \mathbb{E}[B' \mid a = \mathfrak{R}, b]$, which means that if $B_k \in \mathcal{R}$ when $a_{i,k-1} = \mathfrak{R}$, then $B_k \in \mathcal{R}$ also when $a_{i,k-1} = \mathfrak{W}$. It follows by induction that $W_{i,t}(b) \geq W_{i,t+1}(b)$ for all t, b, and i. □

Appendix D
Proof of Theorem 2

Solving (12) corresponds to solving a finite and stationary Constrained Markov Decision Process (CMDP) with bounded

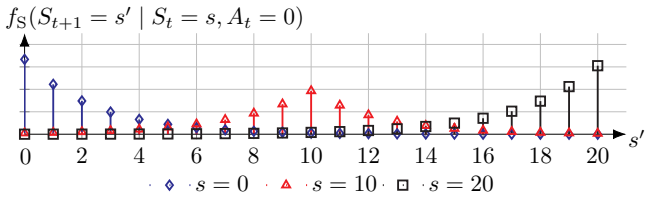


Fig. 15: Example transition function for Prob. 2.

costs and the average cost optimality criterion. Assumption A implies that the CMDP is feasible and assumption B implies that the CMDP is *unichain* [2, Def. 6.5.1], which means that there exists an optimal stationary strategy $\pi^\star$ for which the limit in (10) exists [3, Thm. 8.4.5][2, Thms. 6.5.2–6.5.4].

By introducing the Lagrange multiplier $\lambda \geq 0$ and defining the immediate cost to be $c_\lambda(s_t) = s_t + \lambda[\![s_t \geq f + 1]\!]$ we can reformulate the CMDP as an unconstrained MDP through Lagrangian relaxation [4, Thm. 3.7]. The optimal strategy in the unconstrained MDP satisfies

$$\pi^\star_\lambda(s_t) \in \arg\min_{a \in \{0,1\}} \left[c_\lambda(s_t) + \mathbb{E}_{S_{t+1}} \left[V^\star_\lambda(S_{t+1}) \right] \right], \tag{19}$$

where $V^\star_\lambda$ is the value function [4, Thm. 3.6].

Since $c_\lambda(s_t)$ is non-decreasing in s it follows from assumptions C and D that the MDP has an optimal threshold strategy for any λ [2, Thm. 9.3.1][3, Prop. 4.7.3]. Further, we know from Lagrangian dynamic programming theory that there exists an optimal strategy in the CMDP which is a randomized mixture of two optimal deterministic strategies of the MDP with different Lagrange multipliers λ_1 and λ_2 (15) [2, Thm. 6.6.2], [4, Thm. 12.7]. When combined, these two properties imply Thm. 2. □

Appendix E
Hyperparameters

Hyperparameters for the experimental results and figures reported in this paper are listed in Table 1.

Appendix F
Computation of MTTF and Reliability Functions

The MTTF and the reliability function $R(t)$ can be calculated using numerical methods for Markov chains. Specifically, the number of healthy nodes in the system can be modeled as a Markov chain with state space $\mathcal{S} \triangleq \{0, 1, \ldots, N\}$ and transition matrix $\mathbf{P} \in [0,1]^{|\mathcal{S}|^2}$. In this Markov chain, the subset of states $\mathcal{F} \triangleq \{0, 1, \ldots, f\} \subset \mathcal{S}$ represents the states where service is unavailable. (f is a fixed tolerance threshold and service is guaranteed if $S \geq f + 1$ (Prop. 1).) When calculating the MTTF, we assume that there are no recoveries, which means that $\mathcal{F}$ is absorbing. As a consequence, the mean time to failure (MTTF) can be defined as

$$\mathbb{E}[T^{(f)} \mid S_1 = s_1] \triangleq \mathbb{E}_{(S_t)_{t \geq 1}} \left[\inf \{ t \geq 1 \mid S_t \in \mathcal{F} \} \mid S_1 = s_1 \right],$$

i.e., the MTTF is the mean hitting time of $\mathcal{F}$ in the Markov chain starting at $s_1 \in \mathcal{S}$.

Intrusion recovery parameters	**Values**
Confidence levels	Confidence levels for all figures were computed based on the Student-t distribution
Fig. 5, Fig. 5, Fig. 12b, Fig. 6a, Fig. 6b	$p_{\mathrm{C}_1} = 10^{-5}, p_{\mathrm{C}_2} = 10^{-3}$, $k = 1$ $\eta = 2$, $\mathcal{O} = \{0, \ldots, 9\}$, $Z(\cdot \mid 0) = \mathrm{BetaBin}(n = 10, \alpha = 0.7, \beta = 3)$, $Z(\cdot \mid 1) = \mathrm{BetaBin}(n = 10, \alpha = 1, \beta = 0.7$
Figs. 5–6	no recoveries, $p_{\mathrm{U}} = 0$, $\Delta_{\mathrm{R}} = 100$, $p_{\mathrm{A}} = 0.1$, $k = 1$
Fig. 4, Fig. 12b	$p_{\mathrm{U}} = 2 \times 10^{-2}$, $\Delta_{\mathrm{R}} = 100$, $k = 1$
Figs. 7-8	$\eta = 2$, $p_{\mathrm{A}} = 0.1$, $p_{\mathrm{C}_1} = 10^{-5}$, $p_{\mathrm{C}_2} = 10^{-3}$, $p_{\mathrm{U}} = 2 \times 20^{-2}$, $k = 1$, $Z(\cdot \mid 0) = \mathrm{BetaBin}(n = 10, \alpha = 0.7, \beta = 3)$, $Z(\cdot \mid 1) = \mathrm{BetaBin}(n = 10, \alpha = 1, \beta = 0.7$,
Fig. 9	$\epsilon_{\mathrm{A}} = 0.9$, $N = 10$, $f = 3$, see Fig. 15 for f_{S}
Evaluation in §X	$p_{\mathrm{U}} = 2 \times 10^{-2}$, $p_{\mathrm{A}} = 10^{-1}$, $p_{\mathrm{C}_1} = 10^{-5}$, $p_{\mathrm{C}_2} = 10^{-3}$, $\Delta_{\mathrm{R}} = \infty$, $\epsilon_{\mathrm{A}} = 0.9$, $s_{\max} = 13$, $\eta = 2$, $N_1 = 3$, $f = \min[\frac{N_1 - 1}{2}, 2]$ f_{S} estimated from simulations of Prob 1, PO = CEM in Alg. 1
Fig. 13	$\eta = 2$, $p_{\mathrm{A}} = 0.1$, $p_{\mathrm{C}_1} = 10^{-5}$, $p_{\mathrm{C}_2} = 10^{-3}$, $p_{\mathrm{U}} = 2 \times 20^{-2}$, $k = 1$,PO = CEM in Alg. 1
MINBFT [5, §4.2] parameters	
USIG implementation	RSA with key lengths 1024 bits [6]
T_{exec}, T_{vc}, cp, L	30 seconds, 280 seconds, 10^2, 10^3
PPO [7, Alg. 1] parameters	
lr α, batch, # layers, # neurons, clip ϵ	10^{-5}, $4 \cdot 10^3 t$, 4, 64, 0.2,
GAE λ, ent-coef, activation	0.95, 10^{-4}, ReLU
SPSA parameters [8, Fig. 1]	
$c, \epsilon, \lambda, A, a, N, \delta$	10, 0.101, 0.602, 100, 1, 50, 0.2
M number of samples for each evaluation	50
Incremental pruning parameters [9, Fig. 4]	
Variation, ϵ	normal, 0
Cross-entropy method [10][11, Alg. 1]	
λ (fraction of samples to keep)	0.15, 100
K population size	100
M number of samples for each evaluation	50
Differential evolution [12, Fig. 3]	
Population size K, mutate step	10, 0.2
Recombination rate	0.7
M number of samples for each evaluation	50
Bayesian optimization [13, Alg. 1]	
Acquisition function	lower confidence bound [14, Alg. 1]
β, Kernel	2.5, Matern(2.5)
M number of samples for each evaluation	50
Linear Programming	
Solver	CBC [15]

TABLE 1: Hyperparameters.

By standard Markov chain calculations:

$$\mathbb{E}[T^{(f)} \mid S_1 = s_1] = \begin{cases} 0 & \text{if } s_1 \in \mathcal{F} \\ 1 + \sum_{s' \in \mathcal{S} \setminus \mathcal{F}} \mathbf{P}_{s_1, s'} \mathbb{E}[T^{(f)} \mid S_1 = s'] & \text{if } s_1 \notin \mathcal{F}, \end{cases}$$

which defines a set of $|\mathcal{S}|$ linear equations that can be solved using Gaussian elimination.

Similarly, since the reliability function is defined as $R(t) \triangleq \mathbb{P}[T^{(f)} > t] = \mathbb{P}[S_t > f]$, we have from the Chapman-Kolmogorov equation that

$$R(t) = \sum_{s \in \mathcal{S} \setminus \mathcal{F}} \left(\mathbf{e}_{s_1}^T \mathbf{P}^t \right)_s, \tag{20}$$

where $\mathbf{e}_{s_1}$ is the s_1-basis vector.

Appendix G
The MINBFT Consensus Protocol [5, §4.2]

TOLERANCE is based on a reconfigurable consensus protocol for the partially synchronous system model with hybrid failures, a reliable network, and authenticated communication links (see §IV and Prop. 1). Examples of such protocols include MINBFT [5, §4.2], MINZYZZYVA [5, §4.3], REMINBFT

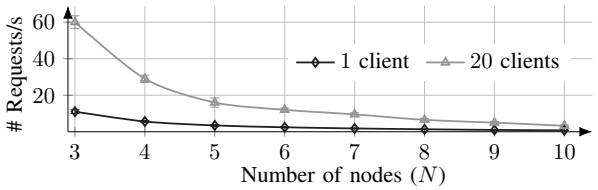


Fig. 16: Average throughput of our implementation of MINBFT during normal operation [5, §4.2]; the error bars indicate the 95% confidence interval based on 1000 samples.

[16, §5], and CHEAPBFT [17, §3]. Our implementation uses MINBFT. Correctness of MINBFT is proven in [5, App. A].

MINBFT is based on PBFT [18] with one crucial difference. While PBFT assumes Byzantine failures and tolerates at up to $f = \frac{N-1}{3}$ failures, MINBFT assumes hybrid failures and tolerates up to $f = \frac{N-1}{2}$ failures. The improved resilience of MINBFT is achieved by assuming access to a trusted component that provides certain functions for the protocol. In particular, MINBFT relies on a tamperproof service at each node that can assert whether a given sequence number was assigned to a message. This service allows MINBFT to prevent equivocation [19] and imposes a first-in-first-out (FIFO) order on requests issued by clients. In TOLERANCE, the tamperproof service is provided by the virtualization layer (see Fig. 2).

We extend the original MINBFT protocol [5, §4.2] to be reconfigurable, where the reconfiguration procedure is based on [20, §IV.B]. The different stages of the protocol are illustrated in Fig. 17 and the throughput is shown in Fig. 16. Our implementation is available in source files associated with this document and hyperparameters are listed in Table 1.

APPENDIX H
DISTRIBUTIONS OF SYSTEM METRICS

Our testbed implementation of TOLERANCE collects hundreds of metrics every time step. To measure the information that a metric provides for detecting intrusions, we calculate the Kullback-Leibler (KL) divergence $D_{\text{KL}}(\widehat{Z}_{O|0} \parallel \widehat{Z}_{O|s>0})$ between the distribution of the metric when no intrusion occurs $\widehat{Z}_{O|\mathbb{H}} \triangleq \widehat{Z}(\cdot \mid S_i = \mathbb{H})$ and during an intrusion $\widehat{Z}_{O|\mathbb{C}} \triangleq \widehat{Z}(\cdot \mid S_i = \mathbb{C})$:

$$D_{\text{KL}}(\widehat{Z}_{O|\mathbb{H}} \parallel \widehat{Z}_{O|\mathbb{C}}) = \sum_{o \in \mathcal{O}} \widehat{Z}_{O|\mathbb{H}} \log \left(\frac{\widehat{Z}_{O|\mathbb{H}}}{\widehat{Z}_{O|\mathbb{C}}} \right).$$

Here $o \in \mathcal{O}$ realizes the random variable O (3), which represents the value of the metric. ($\mathcal{O}$ is the domain of O.)

Figure 18 shows empirical distributions of the collected metrics with the largest KL divergence. We see that the IDS alerts have the largest KL divergence and thus provide the most information for detecting the type of intrusions that we consider in this paper (see Table 6).

REFERENCES

[1] Y. Xiong, N. Chen, X. Gao, and X. Zhou, "Sublinear regret for learning pomdps," 2022.

[2] V. Krishnamurthy, *Partially Observed Markov Decision Processes: From Filtering to Controlled Sensing*. Cambridge University Press, 2016.

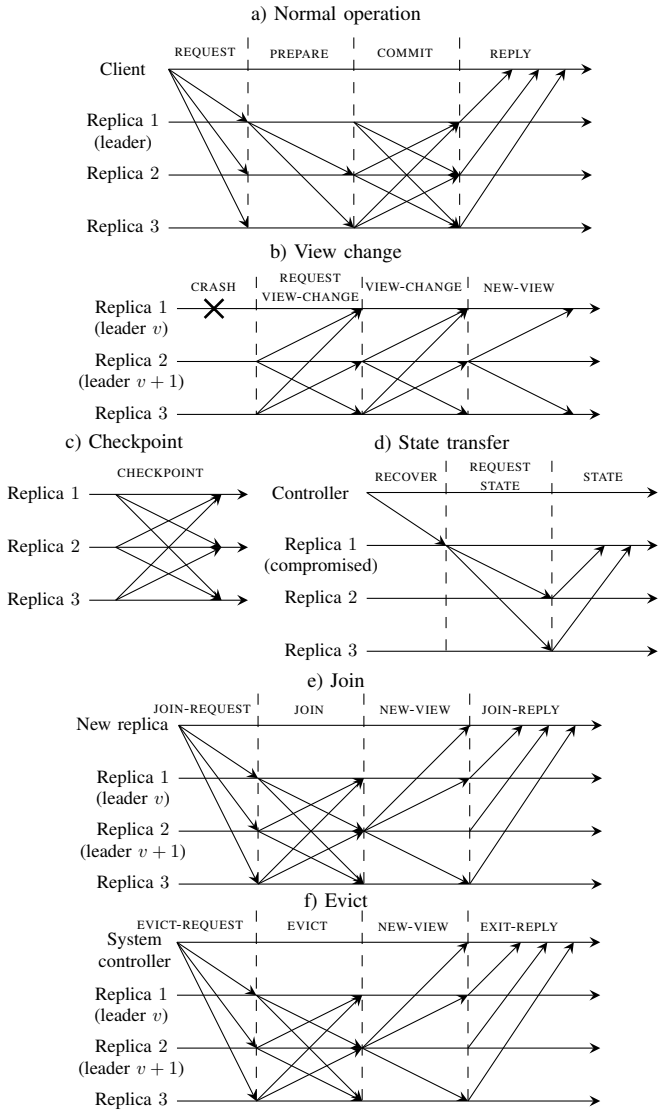


Fig. 17: Time-space diagrams illustrating the message patterns of the MINBFT consensus protocol [5, §4.2].

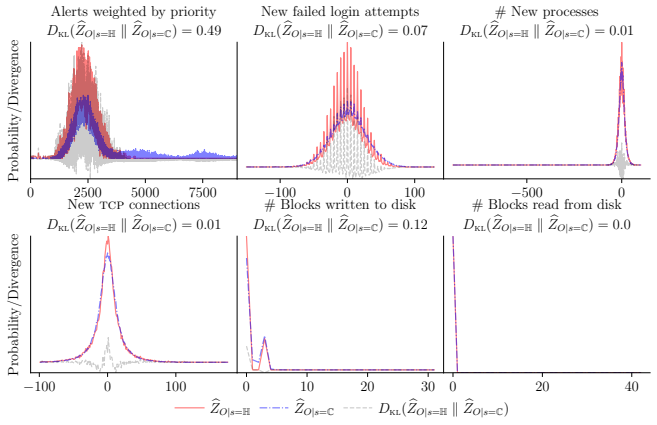


Fig. 18: Empirical distributions of selected infrastructure metrics; the red and blue lines show the distributions when no intrusion occurs and during an intrusion, respectively.

[3] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, 1st ed., USA, 1994.

[4] E. Altman, *Constrained Markov Decision Processes*. Chapman and Hall, 1999.

[5] G. T. dos Santos Veronese, "Intrusion tolerance in large scale networks," Ph.D. dissertation, 2010.

[6] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol. 21, no. 2, p. 120–126, feb 1978.

[7] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *CoRR*, 2017, http://arxiv.org/abs/1707.06347.

[8] J. Spall, "Implementation of the simultaneous perturbation algorithm for stochastic optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 3, pp. 817–823, 1998.

[9] A. Cassandra, M. L. Littman, and N. L. Zhang, "Incremental pruning: A simple, fast, exact method for partially observable markov decision processes," in *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*, ser. UAI'97. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1997, p. 54–61.

[10] R. Rubinstein, "The cross-entropy method for combinatorial and continuous optimization," *Methodology And Computing In Applied Probability*, vol. 1, no. 2, pp. 127–190, Sep 1999.

[11] R. J. Moss, "Cross-entropy method variants for optimization," 2020.

[12] R. Storn and K. Price, "Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, Dec 1997.

[13] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of bayesian optimization," *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148–175, 2016.

[14] N. Srinivas, A. Krause, S. Kakade, and M. Seeger, "Gaussian process optimization in the bandit setting: No regret and experimental design," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ser. ICML'10. Madison, WI, USA: Omnipress, 2010, p. 1015–1022.

[15] J. Forrest and R. Lougee-Heimer, "Cbc user guide," in *Emerging theory, methods, and applications*. INFORMS, 2005, pp. 257–277.

[16] T. Distler, C. Cachin, and R. Kapitza, "Resource-efficient byzantine fault tolerance," *IEEE Transactions on Computers*, vol. 65, no. 9, pp. 2807–2819, 2016.

[17] R. Kapitza, J. Behl, C. Cachin, T. Distler, S. Kuhnle, S. V. Mohammadi, W. Schröder-Preikschat, and K. Stengel, "Cheapbft: Resource-efficient byzantine fault tolerance," in *Proceedings of the 7th ACM European Conference on Computer Systems*, ser. EuroSys '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 295–308.

[18] M. Castro and B. Liskov, "Practical byzantine fault tolerance and proactive recovery," *ACM Trans. Comput. Syst.*, vol. 20, no. 4, p. 398–461, nov 2002.

[19] B.-G. Chun, P. Maniatis, S. Shenker, and J. Kubiatowicz, "Attested append-only memory: Making adversaries stick to their word," in *Proceedings of Twenty-First ACM SIGOPS Symposium on Operating Systems Principles*, ser. SOSP '07. New York, NY, USA: Association for Computing Machinery, 2007, p. 189–204.

[20] X. Hao, L. Yu, L. Zhiqiang, L. Zhen, and G. Dawu, "Dynamic practical byzantine fault tolerance," in *2018 IEEE Conference on Communications and Network Security (CNS)*, 2018, pp. 1–8.