Supplementary Material for Regstar: Efficient Strategy Synthesis for Adversarial Patrolling Games (Supplementary material)

David Klaška[1] **Antonín Kučera**[1] **Vít Musil**[1] **Vojtěch Řehák**[1]

[1]Faculty of Informatics, Masaryk University, Brno, Czech Republic

A PROOF OF THEOREM 1

In this section, we show that Theorem 1 holds even for a special type of *deterministic-update* regular strategies where, for all $(v, m) \in \widehat{V}$ and $v' \in V$, there is at most one m' such that $\sigma((v, m), (v', m')) > 0$.

For the rest of this section, we fix a patrolling graph $G = (V, T, E, time, d, \alpha, \beta)$. We say that a strategy γ is *optimal* if $\mathrm{Val}_G(\gamma) = \mathrm{Val}_G$. The existence of optimal strategies in patrolling games has been proven in Brázdil et al. [2015].

For a strategy γ and $o = v_1, \ldots, v_n, v_n \to v_{n+1} \in \Omega$, let $\gamma[o]$ be a strategy that starts in v_n by selecting the edge $v_n \to v_{n+1}$ with probability one, and for every finite path of the form $v_n, v_{n+1}, \ldots, v_{n+k+1}$ where $k \geq 1$ we have that $\gamma[o](v_n, v_{n+1}, \ldots, v_{n+k+1}) = \gamma(v_1, \ldots, v_{n+k+1})$.

First, we need the following lemma.

Lemma 1. *Let γ be an optimal strategy and $o \in \Omega$ such that* $\mathrm{Prob}^\gamma(o) > 0$. *Then $\gamma[o]$ is optimal.*

Proof. For every $\ell \geq 1$, let $\Omega(\ell)$ be the set of all observations $o = v_1, \ldots, v_\ell, v_\ell \to v_{\ell+1}$ such that $\mathrm{Prob}^\gamma(o) > 0$. Clearly, for every fixed $\ell \geq 1$ we have that

$$\sum_{o \in \Omega(\ell)} \mathrm{Prob}^\gamma(o) \cdot \mathrm{Val}_G(\gamma[o]) \leq \mathrm{Val}_G$$

because $\mathrm{Val}_G(\gamma[o]) \leq \mathrm{Val}_G$ and $\sum_{o \in \Omega(\ell)} \mathrm{Prob}^\gamma(o) = 1$. We show that

$$\mathrm{Val}_G \leq \sum_{o \in \Omega(\ell)} \mathrm{Prob}^\gamma(o) \cdot \mathrm{Val}_G(\gamma[o]) \qquad (1)$$

which implies $\mathrm{Val}_G = \sum_{o \in \Omega(\ell)} \mathrm{Prob}^\gamma(o) \cdot \mathrm{Val}_G(\gamma[o])$, and hence $\mathrm{Val}_G(\gamma[o]) = \mathrm{Val}_G$ for every $o \in \Omega(\ell)$.

It remains to prove (1). Since $\mathrm{Val}_G = \mathrm{Val}_G(\gamma)$, it suffices to show that, for an arbitrarily small $\varepsilon > 0$,

$$\mathrm{Val}_G(\gamma) \leq \varepsilon + \sum_{o \in \Omega(\ell)} \mathrm{Prob}^\gamma(o) \cdot \mathrm{Val}_G(\gamma[o])\,.$$

For every $o \in \Omega(\ell)$, let π_o be an Attacker's strategy such that $\mathbb{E}U_D(\gamma[o], \pi_o) \leq \mathrm{Val}_G(\gamma[o]) + \varepsilon$. Consider another Attacker's strategy $\widehat{\pi}$ waiting for the first ℓ moves and then "switching" to an appropriate π_o according to the corresponding observation. Then,

$$\begin{aligned}
\mathrm{Val}_G(\gamma) &\leq \mathbb{E}U_D(\gamma, \widehat{\pi}) \\
&\leq \sum_{o \in \Omega(\ell)} \mathrm{Prob}^\gamma(o) \cdot \mathbb{E}U_D(\gamma[o], \pi_o) \\
&\leq \sum_{o \in \Omega(\ell)} \mathrm{Prob}^\gamma(o) \cdot (\mathrm{Val}_G(\gamma[o]) + \varepsilon) \\
&= \varepsilon + \sum_{o \in \Omega(\ell)} \mathrm{Prob}^\gamma(o) \cdot \mathrm{Val}_G(\gamma[o])\,.
\end{aligned}$$

This completes the proof of Lemma 1. □

Proof of Theorem 1. We show that for an arbitrarily small $\varepsilon > 0$, there exists a deterministic-update regular strategy σ_ε such that $\mathrm{Val}_G(\sigma_\varepsilon) \geq \mathrm{Val}_G - \varepsilon$.

Let $d_{\max} = \max_{t \in T} d(t)$, and $\alpha_{\max} = \max_{t \in T} \alpha(r)$ and let an optimal Defender's strategy γ be fixed.

We say that two non-empty finite paths $h, h' \in \mathcal{H}$ are δ-*similar*, where $\delta > 0$, if the following conditions are satisfied:

- h and h' end with the same vertex v,
- $\mathrm{Prob}^\gamma(h) > 0$, $\mathrm{Prob}^\gamma(h') > 0$,
- for every target t and every $\ell \in \{1, \ldots, d_{\max}\}$, the probabilities that γ successfully detects an ongoing attack at t in at most ℓ time units after executing the histories h and h' differ at most by δ.

Note that there are only *finitely many* pairwise non-δ-similar histories. More precisely, their total number is bounded from above by $|V| \cdot \left(\lceil \delta^{-1} \rceil\right)^{d_{\max} \cdot |T|}$.

Let us fix an arbitrarily small $\varepsilon > 0$, and let $\delta = \varepsilon / \alpha_{\max}$. Furthermore, let $\kappa = |V| \cdot \left(\lceil \delta^{-1} \rceil\right)^{d_{\max} \cdot |T|}$. We construct a regular deterministic-update strategy σ_ε as follows:

- Let H_δ be the set of all finite paths h of length at most $\kappa \cdot d_{\max}$ such that $\mathrm{Prob}^\gamma(h) > 0$ and for all proper prefixes h', h'' of h whose length is a multiple of $d_{\max}$ we have that if h', h'' are δ-similar, then $h' = h''$. For notation simplification, from now on we identify memory elements with such finite paths.
- For every eligible pair (v, h), the distribution $\sigma_\varepsilon(v, h)$ is determined in the following way:
  - If the length of h is a multiple of $d_{\max}$ and there is a proper prefix h' of h where the length of h' is also a multiple of $d_{\max}$ and the histories h, h' are δ-similar, then $\sigma_\varepsilon(v, h) = \sigma_\varepsilon(v, h')$ (since h' is shorter than h, we may assume that $\sigma_\varepsilon(v, h')$ has already been defined).
  - Otherwise, $\sigma_\varepsilon(v, h)$ is a distribution $\mu \in \mathit{Dist}(V \times H_\delta)$ such that $\mu(v', hv') = \mathrm{Prob}^\gamma(hv')/\mathrm{Prob}^\gamma(h)$ for every vertex v' such that $hv' \in H_\delta$. For the other pairs of $V \times H_\delta$, the distribution μ returns zero.
- The initial distribution assigns $\gamma(\lambda)(v)$ to every $(v, v) \in V \times H_\delta$. For the other pairs of $V \times H_\delta$, the initial distribution returns zero.

Intuitively, the strategy σ_ε mimics the optimal strategy γ, but at appropriate moments "cuts" the length of the history stored in its memory and starts to behave like γ for this shorter history. These intermediate "switches" may lower the overall protection, but since the shorter history is δ-similar to the original one, the impact of these "switches" is very small.

More precisely, we show that, for an arbitrary Attacker's strategy π, $\mathbb{E}U_D(\sigma_\varepsilon, \pi) \geq \mathbb{E}U_D(\gamma, \pi) - \varepsilon$. Since γ is optimal, we obtain $\mathbb{E}U_D(\sigma_\varepsilon, \pi) \geq \mathrm{Val}_G - \varepsilon$, hence $\mathrm{Val}_G(\sigma_\varepsilon) \geq \mathrm{Val}_G - \varepsilon$ as required. For the rest of this proof, we fix an Attacker's strategy π. For every target τ, let π_τ be an Attacker's strategy such that $\pi_\tau(u \to v) = \mathit{attack}_\tau$ for every edge $u \to v$, i.e., π_τ attacks τ immediately. Furthermore, let $\mathit{Att}(\pi, \tau)$ be the set of all observations o such that $\mathrm{Prob}^{\sigma_\varepsilon}(o) > 0$ and $\pi(o) = \mathit{attack}_\tau$. We have the following:

$$\begin{aligned}
&\mathbb{E}U_A(\sigma_\varepsilon, \pi)\\
&\quad = \sum_{\tau \in T} \sum_{o \in \mathit{Att}(\pi,\tau)} \mathrm{Prob}^{\sigma_\varepsilon}(o) \cdot (\alpha(\tau) - \mathbf{P}^{\sigma_\varepsilon}(\tau \mid o))\\
&\quad = \sum_{\tau \in T} \sum_{o \in \mathit{Att}(\pi,\tau)} \mathrm{Prob}^{\sigma_\varepsilon}(o) \cdot \mathbb{E}U_A(\sigma_\varepsilon[o], \pi_\tau)
\end{aligned}$$

Here, $\sigma_\varepsilon[o]$, where $o = v_1, \ldots, v_n, v_n \to v_{n+1}$, is a strategy that starts in v_n by executing the edge $v_n \to v_{n+1}$, and then behaves identically as σ_ε after the history o (since σ_ε is deterministic-update, the associated memory elements are determined uniquely by o).

Now, realize that for every $o \in \mathit{Att}(\pi, \tau)$, there exists an observation o' (stored in the finite memory of σ_ε) such that $\mathrm{Prob}^\gamma(o') > 0$ and the strategy $\sigma_\varepsilon[o]$ "mimics" the strategy $\gamma[o']$ until the finite path stored in the memory of σ_ε is "cut" into a shorter path in the way described above. Since at most one such "cut" is performed during the first $d_{\max}$ steps and the shorter path obtained by the cut is δ-similar to the original one, we obtain that the difference between $\mathbb{E}U_A(\sigma_\varepsilon[o], \pi_\tau)$ and $\mathbb{E}U_A(\gamma[o'], \pi_\tau)$ is at most ε.

By Lemma 1, we obtain $\mathbb{E}U_D(\gamma[o'], \pi_\tau) \geq \mathrm{Val}_G$, hence $\mathbb{E}U_A(\gamma[o'], \pi_\tau) \leq \alpha_{\max} - \mathrm{Val}_G$ and $\mathbb{E}U_A(\sigma_\varepsilon[o], \pi_\tau) \leq \alpha_{\max} - \mathrm{Val}_G + \varepsilon$. This gives

$$\begin{aligned}
&\mathbb{E}U_A(\sigma_\varepsilon, \pi)\\
&\quad \leq \sum_{\tau \in T} \sum_{o \in \mathit{Att}(\pi,\tau)} \mathrm{Prob}^{\sigma_\varepsilon}(o) \cdot (\alpha_{\max} - \mathrm{Val}_G + \varepsilon)\\
&\quad = (\alpha_{\max} - \mathrm{Val}_G + \varepsilon) \cdot \sum_{\tau \in T} \sum_{o \in \mathit{Att}(\pi,\tau)} \mathrm{Prob}^{\sigma_\varepsilon}(o)\\
&\quad \leq \alpha_{\max} - \mathrm{Val}_G + \varepsilon
\end{aligned}$$

since the sum is equal to the probability that π attacks at all against σ_ε, which is at most 1. Hence, $\mathbb{E}U_D(\sigma_\varepsilon, \pi) \geq \mathrm{Val}_G - \varepsilon$ and we are done. □

References

T. Brázdil, P. Hliněný, A. Kučera, V. Řehák, and M. Abaffy. Strategy synthesis in adversarial patrolling games. *CoRR*, abs/1507.03407, 2015.