Cedric: A Collaborative DDoS Defense System Using Credit

Extended Abstract

Jiawei Li
INSC, Tsinghua University
Beijing, China
li-jw19@mails.tsinghua.edu.cn

Hui Wang[*]
INSC, Tsinghua University
ZGC Lab
Beijing, China
jessiewang@tsinghua.edu.cn

Jilong Wang
INSC, Tsinghua University
ZGC Lab
Beijing, China
wjl@cernet.edu.cn

ABSTRACT

Distributed denial of service (DDoS) is one of the most common and damaging cyber attacks, and its impact grows rapidly with the massive use of Internet. Collaborative DDoS defense across countries enables faster and more efficient DDoS attack mitigation. Collaboration requires countries that are not target victims to help detect and block the malicious flow, but selfish countries may refuse to do so because lacking individual gain compared with individual cost. In this paper, we model a stochastic game where selfish countries interact repeatedly and form coalitions to defend DDoS attacks. We design a multi-agent system, *Cedric*, to simulate and solve this complex stochastic game. Each agent adopts Q-learning to find their long-term optimal strategies, and credits are used to encourage efficient collaboration. The Shapley Value based reward assignment of Cedric satisfies several desired properties about fairness and stability. Simulations with trace data of over 7 years' global DDoS attacks support the superiority of Cedric empirically.

KEYWORDS

DDoS Attack; Coaliton Game; Stochastic Game; Credit Assignment

ACM Reference Format:
Jiawei Li, Hui Wang, and Jilong Wang. 2023. Cedric: A Collaborative DDoS Defense System Using Credit: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

A Distributed Denial of Service (DDoS) attack is an attempt to interrupt the functions of a server by flooding it with an enormous number of requests from multiple sources. Although this kind of attack is ancient and well-known, it is never been fundamentally resolved. With more and more critical services utilizing the Internet, DDoS attacks are causing more damage and have bigger impacts on national security businesses and daily life[1][2]. Collaborative DDoS defense requires multiple countries (or domains/ Internet service providers) to detect and block the malicious traffic together to only allow the legitimate traffic flows, which makes the DDoS attack more efficient to mitigate, faster to recover, and causes less social damage[3][6][7]. Since the costs for human resources (an engineer to block the malicious IP address) and software/hardware deployment costs to conduct DDoS defenses might be cheaper in some countries, and social gain can be improved if the defense is conducted in upstream countries, collaborative DDoS defense can bring more social welfare globally than each country do it themselves. However, selfish countries may refuse to do so because lacking individual gain compared with individual cost.

We focus on the incentive issue in collaborative DDoS defense for selfish countries from a game-theoretic perspective and make the following contributions.

- First, we analyze the characteristics and selfish country incentives in a typical DDoS attack and formulate it as a collaborative DDoS defense game.
- Then, since the formulation is a complex stochastic game, we design a multi-agent system, called Cedric, to simulate and solve the collaborative DDoS defense game.
- We propose to use credits to foster efficient cooperation. Especially, our reward assignment mechanism based on Shapley Value satisfies several nice properties of fairness and stability theoretically.
- We conduct simulations on global Internet topology and over 7 years' DDoS attack trace data. The equilibrium that the game achieves in Cedric supports the nice properties empirically.

2 GAME FORMULATION

We formulate the collaborative DDoS defense across countries as an infinite horizon stochastic game.

DDoS Attack. The Internet is simplified and represented as a graph $G = (\mathcal{N}, \mathcal{E})$. A DDoS attack is denoted as a tuple $ddos = (t, src, dst)$, where t is the attack start time, $src = (< c_1, b_1 >, < c_2, b_2 >, ..., < c_n, b_n >)$ is a list of <country, bandwidth> pairs that each pair represent a source country where a zombie locates and the corresponding malicious flow bandwidth. Each malicious flow b_i starts from the source country c_i, and finally reaches the destination country dst. The set of links from c_i to dst is called a path, denoted as ϕ_i.

Collaborative Defense. At the start of a DDoS attack, the victim country dst requests other countries for help. Upon receiving the request, each country $i \in \mathcal{N}$ decides whether to help detect and block the malicious flow. Each country i's payoff in round t could be written as

$$u_{it} = u_i(a_t) = V_{dst}^t \cdot \mathbb{I}(e) + \sum_{j \in \mathcal{N}} V(\hat{b}_{ijt}) - c_{it} \cdot a_{it} \tag{1}$$

where $a_t \in \times_{i \in N} \mathcal{A}_i$ is the joint action profile of all countries at t. V_{dst}^t is a constant payoff of recovering the victim application in dst. $\mathbb{I}(\cdot)$ is an indicator function that returns 1 if the condition inside holds and 0 otherwise. e denotes the event that i is dst and

[*] corresponding author

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom. © 2023 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

the DDoS attack in round t is successfully defended. $\hat{b}_{ijt}$ denotes the maximum blocked malicious flow bandwidth in the country i by country j in the defense coalition, which can potentially be used by legitimate flow and $V(\hat{b}_{ijt})$ is the corresponding social gain generated by country i.

Repeated Interaction. The DDoS attack events occur at different time with different src and dst. A state in the stochastic game is a tuple $\chi = (ddos, history)$ including the current DDoS attack event and the whole history of past DDoS events and country actions. At each state, the action of a country is to either join the defense coalition (1) or not join (0). Each country i's strategy is a function $s_i(\chi)$ that maps each possible state χ to an action. Each country aims to find an optimal strategy that maximizes its total payoff in the long run,

$$s_i^* = \underset{s_i}{argmax} \sum_{t=1}^{\infty} u_{it}(\tilde{a}_{-it}, s_i(\chi_{it})) \tag{2}$$

where $\tilde{a}_{-it} \in \times_{j \in \mathcal{N}, j \neq i} \tilde{\mathcal{A}}_j$ is the joint action profile of all other countries except for i. χ_{it} is the state of i at time t. The state-action space as well as the complexity to solve the game grows exponentially with the number of countries in $\mathcal{N}$.

3 CEDRIC DESIGN

We design a multi-agent system, *Cedric*, to simulate and solve the DDoS defense game. By proper design of the credit assignment mechanism, we give incentives for agents to cooperate efficiently and achieve the most number of successful DDoS defenses with the most social gain and the least cost, or equivalently, the most social welfare.

Definition 3.1. (Social Welfare) The social welfare of a DDoS defense game $G = (\mathcal{N}, \mathcal{E})$ is defined as

$$\begin{aligned} SW(G) &= \sum_{t=1}^{\infty} \sum_{i \in \mathcal{N}} u_{it} \\ &= \sum_{t=1}^{\infty} V_{dst}^t \cdot \mathbb{I}(e) + \sum_{t=1}^{\infty} \sum_{i \in \mathcal{N}} (V(\hat{b}_{it}) - c_{it} \cdot a_{it}) \end{aligned} \tag{3}$$

where

$$V(\hat{b}_{it}) = \sum_{j \in \mathcal{N}} V(\hat{b}_{ijt}) \tag{4}$$

Credit Mechanism. Agents have 2 roles in the DDoS defense game: victim and helper. As a victim, an agent earns an immediate positive individual payoff in a one-shot DDoS defense. As a helper, an agent earns an immediate negative individual payoff in a one-shot DDoS defense. This kind of sometimes positive sometimes negative immediate feedback, like stock, is hard to give useful insights to guide agent behaviors and makes finding optimal strategy complex. We propose to use credits to mediate the positive and negative immediate individual payoff, i.e. let the victim compensate some credits κ to helpers to make individual payoffs of them positive in each one-shot DDoS defense. The immediate reward an agent receives in a one-shot DDoS defense is defined as its individual payoff plus earned credits

$$r_{it} = u_{it} + \kappa_{it} = V_{dst}^t \cdot \mathbb{I}(e) + V(\hat{b}_{it}) - c_{it} \cdot a_{it} + \kappa_{it} \tag{5}$$

For a victim, $\kappa_{it} < 0$. For helpers, $\kappa_{it} > 0$.

Credits would not appear or disappear out of thin air, but can only be traded in the multi-agent system. The total rewards equal the social welfare in each one-shot DDoS defense and the total cumulative rewards equal the social welfare of the whole game G.

$$\begin{aligned} \sum_{t=1}^{\infty} \sum_{i \in \mathcal{N}} r_{it} &= \sum_{t=1}^{\infty} \sum_{i \in \mathcal{N}} (u_{it} + \kappa_{it}) \\ &= \sum_{t=1}^{\infty} (\sum_{i \in \mathcal{N}} u_{it} + \sum_{i \in \mathcal{N}} \kappa_{it}) \\ &= \sum_{t=1}^{\infty} \sum_{i \in \mathcal{N}} u_{it} = SW(G) \end{aligned} \tag{6}$$

In an ideal agent strategy under the credit assignment mechanism, the cumulative rewards and cumulative individual payoff are equal.

$$\sum_{t=1}^{\infty} r_{it} \approx \sum_{t=1}^{\infty} u_{it} \tag{7}$$

Reward Assignment. Cedric's reward assignment is designed based on the Shapley Value. The assignment is feasible and holds nice properties about stability and fairness.

Definition 3.2. (Reward Assignment) In Cedric, the reward assignment for each agent i in the defense coalition $\mathcal{G}$ at time t is

$$\begin{aligned} r_{it} = & \sum_{C \subseteq \mathcal{G} \setminus \{i\}} \frac{|C|!(|\mathcal{G}| - |C| - 1)!}{|\mathcal{G}|!} \\ & \cdot (SW(\mathcal{N}, C \cup \{i\}, t) - SW(\mathcal{N}, C, t)) \end{aligned} \tag{8}$$

Theorem 3.3. *In Cedric, the formed defense coalition and the corresponding reward assignment is feasible and fair in each DDoS attack event. It is stable if the defense coalition is convex.*

4 EXPERIMENTS

We model agents' sequential decision processes as a multi-agent reinforcement learning (MARL) process with event-triggered interactions. For each agent, we adopt an independent Q-learning framework, which learns the Q-value and optimal action at each state. Different agents have different Q-tables and different strategies. Especially, since the complexity of computing the Shapley Value based reward assignment in equation (8) in each round of DDoS defense can be a disaster in practice[4], we propose a sample-based approach to approximate the Shapley Value based reward following [8] and [5]. We implement a custom DDoS defense environment in OpenAI Gym and conduct simulations based on Internet topology and public DDoS attack records. The code is available at this github repo.

5 CONCLUSION

In this paper, we formulate a stochastic game among countries to analyze incentives in global collaborative DDoS defense. We design a multi-agent system, Cedric, to simulate selfish countries' strategies and propose to use credit to achieve efficient cooperation. Cedric achieves nice fairness and stability properties theoretically and reaches desired equilibrium empirically on over 7 years' of DDoS attack trace data.

ACKNOWLEDGMENTS

This work is conducted with the support of the National Key Research and Development Program of China under Grant 2020YFE0200500. We thank Shuying Zhuang, Yi Qiao, Siwei Wang for useful discussions.

REFERENCES

[1] 2016. Hackers Attacked Clinton's and Trump's Websites Before Election Day. https://www.theatlantic.com/technology/archive/2016/11/campaign-websites-under-attack/506942/. Accessed: 2022-10-23.

[2] 2017. Arranging a black market DDoS attack can cost as little as $7 per hour. https://www.ciodive.com/news/arranging-a-black-market-ddos-attack-can-cost-as-little-as-7-per-hour/438844/. Accessed: 2022-10-23.

[3] 2021. Warding off DDoS Attacks with Anti-DDoS — Part 4: Global DDoS Collaborative Protection and GameShield. https://alibaba-cloud.medium.com/warding-off-ddos-attacks-with-anti-ddos-part-4-global-ddos-collaborative-protection-and-6688f4990ce3. Accessed: 2022-10-23.

[4] Jiahui Li, Kun Kuang, Baoxiang Wang, Furui Liu, Long Chen, Fei Wu, and Jun Xiao. 2021. Shapley counterfactual credits for multi-agent reinforcement learning. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. 934–942.

[5] Duc Thien Nguyen, Akshat Kumar, and Hoong Chuin Lau. 2018. Credit assignment for collective multiagent RL with global rewards. *Advances in neural information processing systems* 31 (2018).

[6] Bahman Rashidi and Carol Fung. 2016. CoFence: A collaborative DDoS defence using network function virtualization. In *2016 12th International Conference on Network and Service Management (CNSM)*. 160–166. https://doi.org/10.1109/CNSM.2016.7818412

[7] Bahman Rashidi, Carol Fung, and Elisa Bertino. 2017. A Collaborative DDoS Defence Framework Using Network Function Virtualization. *IEEE Transactions on Information Forensics and Security* 12, 10 (2017), 2483–2497. https://doi.org/10.1109/TIFS.2017.2708693

[8] Jianhong Wang, Yuan Zhang, Tae-Kyun Kim, and Yunjie Gu. 2020. Shapley Q-Value: A Local Reward Approach to Solve Global Reward Games. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*. AAAI Press, 7285–7292. https://ojs.aaai.org/index.php/AAAI/article/view/6220