Mechanism Design for Cooperation in Cybersecurity *

Alistair Simmons[1]

Duke Kunshan University, Kunshan, Jiangsu 215316, China
your email
LinkedIn

Abstract. This paper proposes a novel mechanism design aimed at increasing cooperation among companies in joint cybersecurity efforts, introducing an unprecedented Automated Indicator Sharing (AIS) system rooted in economic and game-theoretic principles. By incorporating structured incentives, blockchain technology for security and transparency, and the strategic utilization of trusted third-party mediators, this mechanism not only addresses traditional barriers such as the fear of exposing sensitive information and trust deficits but also pioneers a model for seamless and secure data exchange. The AIS system designed here stands out in the cybersecurity literature for its automated, transparent, and incentive-compatible approach to threat intelligence sharing across industries, promoting a higher level of collaboration and mutual benefit in cybersecurity defenses.

Notes: In submission Final Project of COMPSCI/ECON 206 Computational Microeconomics, 2024 Spring Term (Seven Week - Second) instructed by Prof. Luyao Zhang at Duke Kunshan University.

Keywords: computational economics · game theory · innovative education · cybersecurity

1 Introduction

Cyber threats continue to evolve in complexity and scale, posing significant risks across various industries. Despite the clear benefits of collaborative cybersecurity, numerous barriers—including fear of exposing sensitive information and lack of trust—hamper effective inter-company cooperation. This paper introduces a mechanism design that addresses these challenges through structured incentives, enhanced by the immutable and transparent nature of blockchain technology, and the strategic use of trusted mediators.

* **Acknowledgments**: I would like to thank Professor Luyao Zhang for her incredible insights and guidance in learning about game theory and structuring this research. I would also like to thank Shiran for his helpful insights and edits.

2 Background

The the Stag Hunt exemplifies situations requiring mutual trust and cooperation to achieve optimal collective outcomes, such as when two hunters must decide independently whether to collaborate on hunting a stag for a higher reward or to pursue individual hares for a guaranteed but lower payoff. This concept is particularly relevant to cybersecurity, where similar cooperation is facilitated by initiatives like Information Sharing and Analysis Centers (ISACs). [1] ISACs help companies within an industry share crucial data on threats and vulnerabilities, enhancing collective defenses much like hunters in a Stag Hunt deciding to pursue the stag together, thereby applying these theoretical models to real-world challenges in cybersecurity to foster a cooperative environment where information sharing is both encouraged and strategically beneficial. [2]

A Scholarly Intervention

Previous research on game theory and cybersecurity information sharing established an equilibrium, showcasing how the effectiveness of information sharing directly influences security investments and vulnerability to cyber attacks. [3] The model demonstrates a linear correlation between the effectiveness of information sharing (parameterized by y) and the reduction in individual security investment needs, thus highlighting how improved information sharing can lead to more efficient resource allocation in cybersecurity defenses. However, this advantage is balanced against the risks of information leakage, represented by leakage cost coefficients (ϕ), underscoring the critical need for meticulously designed mechanisms that not only capitalize on the benefits of information sharing but also safeguard against its inherent risks. This model supports foundational assumptions of the proposed mechanism design: improving cybersecurity information sharing capabilities benefits all entities involved. This research focuses on the exchange of cybersecurity information between specific firms, and the proposed model builds on the this equilibrium to propose an unprecedented industry-wide initiative: [3]

$$t_i = \frac{hV(1-\alpha)+f_i-\alpha f_j}{h(1-\alpha)+f_i-\alpha f_j}\left[\frac{f_j-\alpha f_i}{h(1-\alpha)+f_i-\alpha f_j}-\alpha\frac{f_i-\alpha f_j}{h(1-\alpha)+f_i-\alpha f_j}\right] - \gamma^2\left(\frac{2\phi_1 f_i+\phi_3 f_j}{4\phi_1-\phi_2-\phi_3}\right)$$

B Constructing Mechanism Design

B.1 Theoretical Concept

The proposed mechanism design leverages a contribution-based pricing model to incentivize companies to share critical cybersecurity information, with rewards scaled according to the severity and scope of disclosed vulnerabilities [4].

This mechanism would create an industry partnership for cybersecurity information sharing that companies can join by contributing information or funding. The compensation structure differentiates between critical industries, core businesses, and small businesses, offering higher rewards for root vulnerabilities in critical sectors and lower rewards for specific vulnerabilities in smaller businesses [5]. Implementation involves an independent panel that verifies and values each reported vulnerability to ensure fair compensation, supported by a dynamic pricing strategy that adjusts based on the cybersecurity landscape [6]. Beyond monetary rewards, contributors gain reputation benefits and access to a collective intelligence pool [7]. The reputation benefits would come from credentials of membership for being a trustworthy industry partnership to promote cybersecurity information sharing. This model not only enhances overall cybersecurity by encouraging the sharing of vital information but also targets incentives to maximize prevention of cyberattacks, fostering a proactive, collaborative security environment [8].

B.2 Market Value

The global cost of cybercrime is expected to reach $10.5 trillion annually by 2025, suggesting a significant market for cybersecurity information sharing services that can mitigate these costs [9]. Further, with the average cost of a data breach reported as $3.86 million [10], the potential for cost savings through proactive information sharing is substantial.

C Application

The cyber information sharing system utilizes blockchain technology to create a decentralized, immutable ledger, enhancing trust by preventing data tampering and ensuring transparency. Smart contracts automate the verification and compensation processes, providing automatic and transparent payments to contributors when predefined criteria are met, thus improving efficiency and reducing administrative burdens. Smart contracts can also be used by the company sharing the cybersecurity information to specify what entities can access the information. For example, smart contracts can be used to specify sharing cybersecurity data only within a particular critical industry, such as water processing plants. Furthermore, smart contracts can be used to stipulate that recipients of the information do not use it to damage the reputation of or exploit the company who shared the data. Legal repercussions from contract law can also be introduced to promote accountability, and further research on the overlap between smart contracts and contract law is necessary.

These technologies also secure data transactions against unauthorized access and enable a complete, auditable trail of all information exchanges, crucial for integrity and compliance checks. This integration of blockchain and smart contracts fundamentally transforms sharing into a more reliable, efficient, and secure

Fig. 1. Cybersecurity Informat Sharing Model Concept

practice, guaranteeing fair compensation and bolstering trust among participants. Using the blockchain enables for third party oversight of a decentralized information sharing network.

C.1 Pricing and Utility of Information

A value-based pricing model for cybersecurity information can strategically align pricing with the criticality of the data shared. This model evaluates three core criteria: the importance of the industry, the severity of the vulnerability, and the scope of the information shared. Industries are categorized into critical, core, and regular businesses, reflecting their significance to national security and economic stability. Vulnerabilities are classified based on their potential impact, ranging from root vulnerabilities that pose extensive threats to access-point vulnerabilities that affect localized areas. The scope of information, including the volume of data and its reach across organizations, further refines the value assessment. For example, critical industry data detailing root vulnerabilities shared widely warrants the highest pricing due to its potential to prevent significant security incidents, whereas information about less severe vulnerabilities in less critical sectors commands lower prices.

Implementing this model ensures that the prices charged for cybersecurity information proportionally reflect its value in preventing and mitigating cyber threats. This method not only encourages the sharing of vital information by aligning the price with the potential return on investment but also enhances the overall security landscape. According to [11], aligning economic incentives with

cybersecurity practices promotes the sharing of crucial data and the adoption of robust cybersecurity measures across industries. By pricing information based on its strategic importance, organizations are more likely to invest in and benefit from comprehensive cybersecurity intelligence, leading to a more resilient digital infrastructure.

C.2 Modeling The Benefits

Stag and Hare Model I am using the Stag Hunt game to model the economic benefits of creating incentives for exchanging cybersecurity information, transforming partnerships from noncooperative to cooperative. This approach helps demonstrate how tailored incentives can effectively alter the payoff landscape, encouraging entities to collaborate in sharing vital cybersecurity data. By adjusting the risk-reward equation to favor cooperation, my model predicts a shift in strategic behaviors, leading to enhanced mutual gains and stronger collective defenses against cyber threats. This transformation is crucial for developing resilient cybersecurity networks where shared information leads to improved situational awareness and more robust responses to emerging threats.

Stag and Hunt Without Incentives This matrix captures the decision-making scenario where two companies, particularly Company 1 with three times the cybersecurity knowledge of Company 2, consider whether to share their information. The benefit for sharing is gaining access to the other company's cybersecurity knowledge. However, sharing entails a risk of -0.5, which could result from competitors exploiting the shared information or it leading to a devaluation of the company's market standing. This model represents the basic cooperative dynamics without additional incentives, reflecting the natural concerns companies might have about sharing sensitive information.

Table 1. Matrix 1: No Incentives

	C	D
C	(3.5, 3.5)	(0.5, 4)
D	(4, 0.5)	(1, 1)

Stag and Hunt With Incentives In this upgraded scenario, companies are enticed to share information through proportional incentives valued at 0.2 percent of their overall cybersecurity knowledge. Additionally, risks associated with sharing are mitigated to -0.3 due to the implementation of blockchain and smart contract technologies, which provide secure, transparent mechanisms for sharing and reduce the potential for misuse or devaluation. This matrix demonstrates how the introduction of incentives and secure sharing infrastructures can potentially alter the decision calculus of the companies, making the option to cooperate more appealing by offsetting inherent risks.

Table 2. Matrix 1: No Incentives

	C	D
C	(3.9, 4.3)	(0.9, 3)
D	(4, 0.9)	(1, 1)

C.3 Comparative Analysis

Using Nash.py and QuantEcon, the equilibrium for the payoff matrices were calculated to be as follows:

No Incentives Model (Matrix 1) The Nash Equilibrium calculations result in three outcomes: one where both companies choose to cooperate (C, C), one where both choose not to cooperate (D, D), and a mixed-strategy equilibrium where they randomize their strategies with certain probabilities. This indicates that without additional incentives, the company with more information does not have a clear incentive to share, as the risk of sharing may outweigh the benefits.

Incentives Model (Matrix 2) The Nash Equilibrium is consistently at (C, C), where both companies choose to cooperate. This suggests that with the incentives in place, it becomes in the interest of both companies to share information. The provision of incentives aligns their strategies towards mutual cooperation, illustrating the effectiveness of the incentives in promoting information sharing.

C.4 Usefulness and Application

The Nash Equilibrium obtained from the payoff matrices reveal a significant shift in strategy due to the mechanism design, highlighting its influence on cybersecurity information sharing and economic utility. In the no incentives model, the existence of multiple equilibrium, including a mixed-strategy one, indicates a state of indecision and risk aversion among companies, particularly for the one with more information. This aligns with Akerlof's concept of information asymmetry, where the risk of sharing valuable data can lead to a non-cooperative equilibrium [8].

D Conclusion

In conclusion, this paper presents a comprehensive mechanism design to address the escalating challenges posed by cyber threats across various industries. By integrating structured incentives with the robustness of blockchain technology and the strategic deployment of trusted mediators, the proposed design fosters intercompany cooperation despite the prevailing concerns of information exposure

and trust deficits. The employment of game-theoretic frameworks, particularly the Stag Hunt analogy, underscores the necessity for mutual trust and collaboration in achieving superior collective security outcomes. Furthermore, the research substantiates how enhanced information sharing not only mitigates individual security investment needs but also curtails potential security breaches effectively, balancing these benefits against the risks of information leakage. This mechanism design is poised to transform traditional cybersecurity approaches by cultivating a cooperative environment that encourages strategic information exchange and fortifies collective cyber defenses, ultimately enhancing the resilience of entire industries against the specter of cyber threats.

Bibliography

[1] "Information sharing and analysis centers (isacs)," https://www.nationalisacs.org/, 2004, accessed: 2024.

[2] B. Skyrms, *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press, 2004.

[3] X. Gao, W. Zhong, and S. Mei, "A game-theoretic analysis of information sharing and security investment for complementary firms," *The Journal of the Operational Research Society*, vol. 65, no. 11, pp. 1682–1691, 2014. [Online]. Available: http://www.jstor.org/stable/24505080

[4] R. Anderson, *Why Information Security is Hard-An Economic Perspective*. ACM, 2001.

[5] J. Pawlick, S. Farhang, and Q. Zhu, "A game-theoretic flipit model with multiple defenders," *International Conference on Decision and Game Theory for Security*, pp. 135–154, 2015.

[6] H. R. Varian, "System reliability and free riding," *Economics of Information Security*, vol. 12, pp. 1–15, 2004.

[7] B. Schneier, *Liars and Outliers: Enabling the Trust that Society Needs to Thrive*. John Wiley & Sons, 2012.

[8] G. A. Akerlof, *The Market for Lemons: Quality Uncertainty and the Market Mechanism*. Oxford University Press, 1970, vol. 84, no. 3.

[9] S. Morgan, "Cybercrime to cost the world $10.5 trillion annually by 2025," *Cybercrime Magazine*, 2020.

[10] IBM, "Cost of a data breach report 2020," https://www.ibm.com/security/digital-assets/cost-data-breach-report/, 2020.

[11] S. J. Shackelford and S. Myers, "Economic incentives for cybersecurity: Using economics to design technologies ready for the marketplace," *Telecommunications Policy*, vol. 38, no. 11, pp. 992–1001, 2014.