Karafan

Quarterly Scientific Journal of Technical and Vocational University

Spring 2022, Vol. 19, No. 1, p. 495-521

Journal Homepage: https://karafan.tvu.ac.ir/?lang=en

doi:10.48301/KSSA.2022.320681.1909

TVU

A Method for Assessing the Security Risk in Cyber-physical Systems with Incomplete Information Using Bayesian Game Theory

Hamed Sepehrzadeh[1*]

[1]Assistant Professor, Department of Computer Engineering, Technical and Vocational University (TVU), Tehran, Iran.

ARTICLE INFO

Article Type:
Original Research

Received: 12.20.2021
Revised: 01.19.2022
Accepted: 02.08.2022

Keyword:
Cyber-physical Systems (CPSs)
Security Risk
Game Theory
Attacker
Attack Detection

***Corresponding Author:**
Hamed Sepehrzadeh
Email: *hsepehrzadeh@tvu.ac.ir*

ABSTRACT

In recent years, with the development and advancement of various aspects of information, we have witnessed the introduction of new technologies in various sectors of life and industry. The industrial sector has been most impacted with many critical infrastructures based on new technologies. On the other hand, the increasing complexity of these sectors has made the task of managing and maintaining safety much more difficult than before, so that in recent years the issue of security in industrial systems, particularly in critical and complex infrastructure, has become one of the major challenges. An attack on these systems can have adverse physical effects and consequences on equipment, products, service outages, and even human health. In this paper, a method for modeling and assessing the security risk in cyber-physical systems is presented. In this method, the interaction between the system and the attacker was modeled as a Bayesian game with incomplete information. The considered security parameters were divided into two categories of attack and defensive parameters and the attacker and the system behavior predicted using the proposed model. The inputs of this model were control components, process model, system and attack parameters, and its outputs were quantitative values for security risk metric.

E-ISSN: 2538-4430
ISSN: 2382-9796

EXTENDED ABSTRACT

Introduction

Different Cyber-physical systems (CPSs) are the combination and integration of computing and communication systems with physical processes. Although this combination and integration has increased the efficiency and reliability of the systems, it has also exposed these systems to sabotage attacks.

The main goal of this paper was to provide a method for modelling and quantitative assessment of security risk in CPSs with specific and functional components so that the appropriate estimate of the security risk in attacking these systems can be estimated. The proposed method was based on game theory and is a two-player non-zero-sum game between the system and the attacker.

Methodology

In this method, game theory was used to study the behavior of the system and the attacker. The proposed model is a two-player game model between the attacker and the system, non-zero-sum and with incomplete information. Formally, the proposed game model was defined as a multiple $G = <P, Si, U_{ij}>$, where P is the set of players, S_i is the set of actions or strategies of player i, and U_{ij} is the probability of player i choosing strategy j. A game in which some players do not have the benefit of knowledge of one another is called a game with incomplete information. In fact, most of the time, the system and the attacker do not have complete information of one another. Therefore, the best option to model the conflict between them is to use the Bayesian game-based modeling method with incomplete information.

Attacker's strategy was considered either to attack or not to attack: $SA = \{A, NA\}$. Furthermore, the strategy of the system was considered in the form of detection and non-detection of the attack and disorder caused by the attack: $SS = \{D, ND\}$.

In this model, it was assumed that the system did not have full knowledge of the attacker's knowledge, and for this reason, it estimated the attacker's knowledge to be lower than *k'* with pk probability and higher than *k'* with *pk-1* probability. As a result, it can be assumed that two types of attackers are being dealt with: type one, with a low level of knowledge, and type two, attacker with a high level of knowledge.

In order to model this game which was solved using the conventional game solving method, the standard form of the game was divided into two models. Tables 1 and 2 show the standard form of the game model for type one and two attackers, respectively. The first line in each house of the table shows the profit of the attacker and the second line shows the profit of the system.

The parameters of the desired game model are as follows:

- *K* represents the attacker's level of knowledge about the target's physical process and its failure conditions.
- *R* shows the attacker's profit from physical disruption to the system. This parameter also shows the damage caused to the system due to physical disturbance.
- *E* represents the necessary cost for the system in order to invest in security in order to detect intrusion and anomaly, in order to prevent physical disruption by attackers to the system.

- *W* shows the benefit of the system from detecting an attack.
- *B* is the benefit and reward of the system from the discovery of an attack.
- *Pa* shows the probability of an attack in order to cause physical disruption to the system.
- *Pd* is the probability of detecting an attack.

Table 1. The game model for the attacker type 1.

		System	
		No-Detection (ND)	**Detection (D)**
Attacker	Attack (A)	$K \times R_0 - C_0$ $-K \times R_0 - W_0$	$-F_0 - C_0$ $B_0 - K \times E_0$
	No-Attack (NA)	0 0	0 $-K \times E_0$

Table 2. The game model for the attacker type 2.

		System	
		No-Detection (ND)	**Detection (D)**
Attacker	Attack (A)	$K \times R_1 - C_1$ $-K \times R_1 - W_1$	$-C_1$ $B_1 - K \times E_1$
	No-Attack (NA)	0 0	0 $-K \times E_1$

The combined Nash equilibrium of the game model for the attacker type 1 and the system are as follows:

$$\begin{aligned} &U_{attacker}(A) = U_{attacker}(NA) \rightarrow \\ &(1 - P_{d0}) \times (K \times R_0 - C_0) + P_{d0} \times (-C_0) = 0 \\ &\rightarrow P_{d0} = (K \times R_0 - C_0)/(K \times R_0) = 1 - \frac{C_0}{K \times R_0} \end{aligned} \qquad \begin{aligned} &U_{system}(NI) = U_{system}(I) \rightarrow \\ &(P_{a0}) \times (-K \times R_0 - W_0) + (1 - P_{a0}) \times 0 = \\ &(P_{a0}) \times (B_0 - K \times E_0) + (1 - P_{a0}) \times (-K \times E_0) \\ &\rightarrow P_{a0} = \frac{K \times E_0}{B_0 + (K \times R_0) + W_0} \end{aligned} \quad (1)$$

Similarly, the combined Nash equilibrium for the attacker type 2 can be obtained as follows:

$$\begin{aligned} &U_{attacker}(A) = U_{attacker}(NA) \rightarrow \\ &(1 - P_{d1}) \times (K \times R_1 - C_1) + P_{d1} \times (-C_1) = 0 \\ &\rightarrow P_{d1} = (K \times R_1 - C_1)/(K \times R_1) = 1 - \frac{C_1}{K \times R_1} \end{aligned} \qquad \begin{aligned} &U_{system}(NI) = U_{system}(I) \rightarrow \\ &(P_{a1}) \times (-K \times R_1 - W_1) + (1 - P_{a1}) \times 0 = \\ &(P_{a1}) \times (B_1 - K \times E_1) + (1 - P_{a1}) \times (-K \times E_1) \\ &\rightarrow P_{a1} = \frac{K \times E_1}{B_1 + (K \times R_1) + W_1} \end{aligned} \quad (2)$$

Results and discussion

The considered system as an example is a laboratory system that consists of two tanks that are placed at different heights relative to each other, three valves and one level sensor. Figure 1(a) shows the probability of attack and attack detection for different values of attacker knowledge and Figures 1 and 2 show the risk of attacks performed for different values of attacker knowledge.

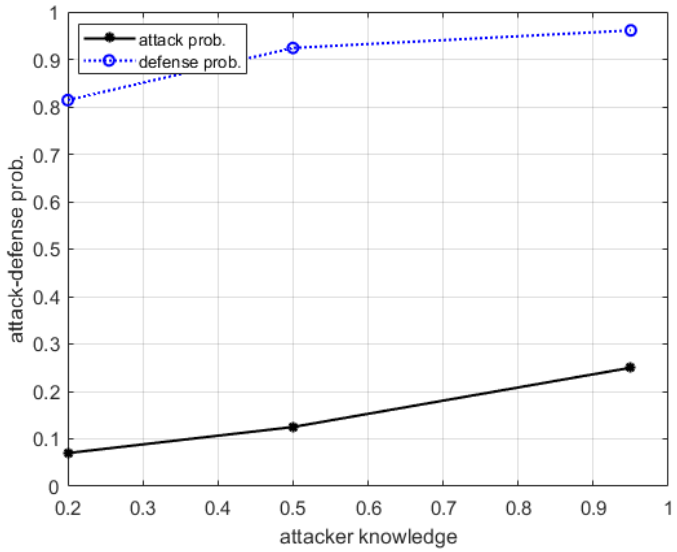


Figure 1. The attack and attack detection probability vs. attacker knowledge.

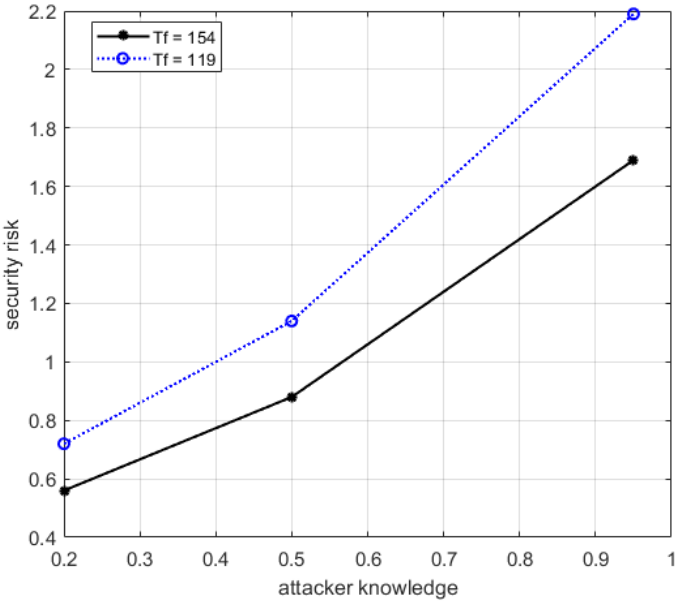


Figure 2. Risk of attack vs. attacker knowledge.

As can be observed, as the attacker's knowledge increased, the probability of attack and the probability of detection also increased. The risk of attacks increased due to the increase in the attacker's knowledge and, as a result, the increase in the probability of an attack.

Conclusion

In this paper, a method for modeling and assessing the security risk of CPSs using the game theory with incomplete information was presented. The presented method modeled the confrontation between the system and the attacker as a Bayesian game with incomplete information. Since the information of the system and the attacker did not have complete information of one another, this method was very suitable for modeling the conflict between the behavior of the system and the attacker. The input of the model comprised of the game components: the cost of the attacker's attack, the amount of damage caused to the system by a successful attack, the penalty for not detecting the attack, the system's profit and reward from detecting the attack, and the attacker's knowledge about the target system. The output of the model comprised of the relationships that allowed calculating the probability of attack and the probability of detecting the attack. In addition, a method to evaluate the security risk using the probability of attack and its detection, the damage amount of the attack on the system and the time to system failure was provided.

کارافن
فصلنامه علمی دانشگاه فنی و حرفه‌ای
بهار ۱۴۰۱، دوره ۱۹، شماره ۱، ۵۲۱-۴۹۵
آدرس نشریه: https://karafan.tvu.ac.ir/
doi:10.48301/KSSA.2022.320681.1909

روشی برای ارزیابی مخاطره امنیتی در سیستم‌های سایبر- فیزیکی با اطلاعات ناقص با استفاده از نظریه بازی بیزی

حامد سپهرزاده[1*]

۱- استادیار، گروه مهندسی کامپیوتر، دانشگاه فنی و حرفه‌ای، تهران، ایران.

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

دریافت مقاله: ۱۴۰۰/۰۹/۲۹
بازنگری مقاله: ۱۴۰۰/۱۰/۲۹
پذیرش مقاله: ۱۴۰۰/۱۱/۱۹

کلید واژگان:
سیستم‌های سایبر- فیزیکی
مخاطره امنیتی
نظریه بازی
مهاجم
کشف حمله

***نویسنده مسئول:** حامد سپهرزاده
پست الکترونیکی:
hsepehrzadeh@tvu.ac.ir

چکیده

در سال‌های اخیر با توسعه و پیشرفت جنبه‌های مختلفی از دانش، شاهد ورود فناوری‌های جدید در بخش‌های مختلفی از زندگی و صنعت شده‌ایم. در این بین، بخش صنعت بیش‌ترین تأثیر را پذیرفته است به‌طوری که بسیاری از زیرساخت‌های حیاتی مبتنی بر فناوری‌های جدید شده است. از طرفی، افزایش پیچیدگی در این بخش‌ها، مدیریت و حفظ ایمنی را بسیار سخت‌تر از قبل کرده است، به‌طوری که در سال‌های اخیر موضوع امنیت در سیستم‌های صنعتی و به‌خصوص زیرساخت‌های حیاتی و پیچیده به یکی از معضلات اساسی تبدیل شده است. حمله در این سیستم‌ها می‌تواند تأثیرات و پیامدهای فیزیکی ناگواری بر تجهیزات، تولیدات، قطعی سرویس و حتی سلامت افراد داشته باشد. در این مقاله، روشی برای مدل‌سازی و ارزیابی مخاطره امنیتی در سیستم‌های سایبر- فیزیکی ارائه شده است. در این روش، تقابل بین سیستم و مهاجم به‌صورت یک بازی بیزی با اطلاعات ناقص مدل شده است. مؤلفه‌های امنیتی در نظر گرفته شده به دو دسته مؤلفه‌های دفاعی و مؤلفه‌های هجومی تقسیم‌بندی شده‌اند و رفتار مهاجم و سیستم با استفاده از مدل ارائه‌شده پیش‌بینی شده است. ورودی‌های این مدل، مؤلفه‌های کنترلی، مدل فرایند، مؤلفه‌های سیستمی و هجومی هستند و خروجی آن مقادیر کمّی برای سنجه مخاطره امنیتی هستند.

شاپای الکترونیکی: ۲۵۳۸-۴۴۳۰
شاپای چاپی: ۲۳۸۲-۹۷۹۶

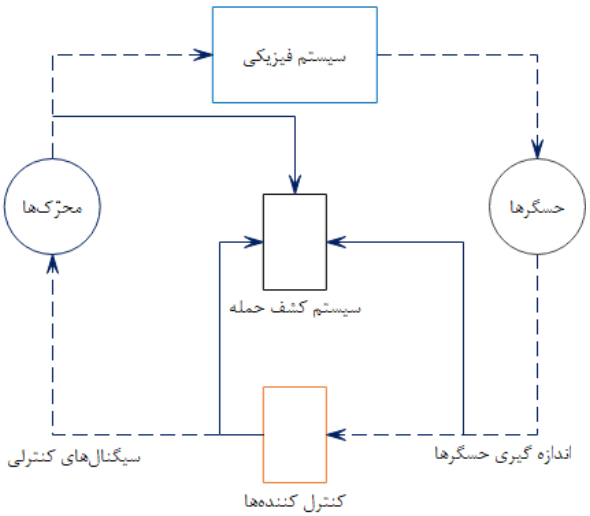
مقدمه

سیستم‌های سایبر- فیزیکی، ترکیب و یکپارچه‌سازی سیستم‌های رایانشی و ارتباطی با فرایندهای فیزیکی هستند [۱]. در این سیستم‌ها، فرایندهای فیزیکی توسط بخش سایبری نظارت و کنترل می‌شوند. سیستم‌های سایبر- فیزیکی در متون مختلف به‌عنوان سیستم‌های کنترل شبکه‌ای یا سیستم‌های کنترل صنعتی نیز شناخته می‌شوند [۲]. این سیستم‌ها در زیرساخت‌های حیاتی مانند شبکه برق هوشمند، خطوط توزیع آب، گاز و سوخت، اتومبیل‌های پیشرفته، صنایع شیمیایی، صنایع هوایی، حمل‌ونقل، سلامت و پزشکی قابل مشاهده هستند [۲].

شکل ۱، انتزاعی از یک سیستم سایبر- فیزیکی را نشان می‌دهد. در این سیستم‌ها مجموعه‌ای از حسگرها پدیده‌های فیزیکی مانند سرعت، دما، رطوبت و فشار را اندازه‌گیری می‌کنند. سپس، مشاهدات به کنترل‌کننده‌ها ارسال می‌شوند. این کار معمولاً با نوشتن اطلاعات دریافت شده در بافرهای ورودی کنترل‌کننده‌ها انجام می‌شود [۳]. کنترل‌کننده‌ها از آخرین مقدار ذخیره شده در بافرهای ورودی استفاده می‌کند و تصمیم کنترلی مناسب را بر آن اساس می‌گیرند. در هر لحظه وضعیت سیستم به کنسول اپراتور در ایستگاه واسط انسان- ماشین ارسال می‌شود تا کاربر سیستم در جریان وضعیت سیستم باشد [۳].

ارتباطات بین حسگرها و کنترل‌کننده‌ها در این سیستم‌ها به سه دسته تقسیم می‌شوند [۴]: ۱) ارتباطات حسگر به حسگر برای جمع‌آوری مشاهدات دریافت شده، ۲) ارتباطات حسگر به کنترل‌کننده برای ایجاد تصمیمات کنترلی ۳) کنترل‌کننده به کنترل‌کننده برای اخذ تصمیم کنترلی مناسب.

لازم است سیستم‌های کشف حمله (نفوذ و بی‌نظمی) به‌طور مداوم وجود هرگونه نفوذ امنیتی در حسگرها، محرّک‌ها و کنترل‌کننده‌ها را بررسی کنند [۵]. در حقیقت سیستم‌های کشف نفوذ و بی‌نظمی بر تشخیص انحرافات موجود از رفتار عادی سیستم متمرکز هستند. به همین منظور، مشاهدات حسگرها و تصمیمات کنترل‌کننده‌ها برای این سیستم‌ها نیز ارسال می‌شوند تا با استفاده از مدل فرایند فیزیکی تحت کنترل، وجود حمله را تشخیص دهند.



شکل ۱. نمایشی از سیستم‌های سایبر- فیزیکی [۶].

این ترکیب و یکپارچه‌سازی اگرچه باعث افزایش کارایی و قابلیت اطمینان سیستم‌ها شده است اما از طرف دیگر این سیستم‌ها را در معرض حملات خراب‌کارانه نیز قرار داده است. حمله به این سیستم‌ها ممکن است منجر به ورود خسارت به تجهیزات، تولیدات، اختلال فیزیکی در کارکرد سیستم، از دسترس خارج شدن سیستم برای مدتی، یا تهدید ایمنی برای افراد شود [۷].

از طرف دیگر، مهاجمان به‌منظور حمله به این سیستم‌ها نیاز است که در سطح کنترل سیستم متمرکز شوند. بنابراین باید دانش کافی در مورد فرایند فیزیکی سیستم، اصول کنترلی، پردازش سیگنال و نحوه آسیب دیدن سیستم فیزیکی تحت کنترل داشته باشند. بدون این سطح از دانش، نتیجه حمله به‌جز اختلال ناچیز مورد قابل‌توجهی نخواهد بود [۷].

هدف اصلی این مقاله، ارائه روشی برای مدل‌سازی و ارزیابی کمّی مخاطره[1] امنیتی در سیستم‌های سایبر- فیزیکی با مؤلفه‌های خاص و کاربردی است تا بتوان تخمین مناسبی از مخاطره امنیتی در حمله به این سیستم‌ها را برآورد کرد. روش پیشنهادی، مبتنی بر نظریه بازی [۸; ۹] است و یک بازی دونفره غیرمجموع صفر بین سیستم و مهاجم است. نظریه بازی تلاش می‌کند تا رفتار ریاضی حاکم بر یک موقعیت راهبردی (تضاد منافع) را مدل‌سازی کند. این موقعیت زمانی پدید می‌آید که موفقیت یک فرد وابسته به راهبردهایی است که دیگران انتخاب می‌کنند. هدف نهایی یافتن راهبرد بهینه برای بازیکنان است. از این رو نظریه بازی انتخاب مناسبی برای مدل‌سازی تقابل سیستم و مهاجم خواهد بود.

فرض رایج در بازی‌ها معمولاً این است که بازیکنان از منفعت یکدیگر مطلع هستند و این تنها برای ساده‌سازی است و بعضاً برای بعضی از کاربردها می‌تواند قابل‌قبول باشد [۸]. در اغلب مواقع سیستم و مهاجم، اطلاعات کاملی از یکدیگر ندارند؛ بنابراین بهترین گزینه برای مدل کردن تقابل بین آن‌ها استفاده از روش مدل‌سازی مبتنی بر بازی بیزی با اطلاعات ناقص است [۱۰; ۱۱]. به بازی‌ای که حداقل یکی از بازیکنان در آن در مورد بعضی از ویژگی‌های بازی که بر تصمیم‌گیری‌های او مؤثر است اطّلاعات کاملی ندارد بازی بیزی گفته می‌شود [۱۱]. به همین دلیل در این مقاله از بازی بیزی برای مدل‌سازی تقابل بین مهاجم و سیستم استفاده شده است.

ساختار ادامه این مقاله به این شرح است: در بخش ۲، بعضی از کارهای مرتبط در حوزه مطالعه امنیت سیستم‌های سایبر- فیزیکی بررسی شده است. در بخش ۳، روش پیشنهادی توصیف شده است. در بخش ۴ مثالی از کاربرد مدل پیشنهادی در قالب یک مطالعه موردی ارائه شده و در بخش ۵ نتایج حاصل از مقاله بیان شده است.

کارهای مرتبط

از آنجا که حمله به سیستم‌های سایبر- فیزیکی اثراتی متمایز با سیستم‌های سایبری خواهد داشت؛ پرداختن به امنیت این سیستم‌ها بسیار اهمیت دارد. کارهای زیادی در این زمینه انجام شده است از جمله نویسندگان در [۱۲] به امنیت سایبر- فیزیکی سیستم‌های نظارت، محافظت و کنترل مناطق وسیع، از دیدگاه یک حمله سایبری هدایت شده پرداخته‌اند و یک روش مبتنی بر نظریه بازی برای در نظر گرفتن این موضوع ارائه کرده‌اند. در این روش، تقابل بین مهاجم و مدافع در سیستم‌های سایبر- فیزیکی با درنظر گرفتن هزینه‌های مربوط در یک چارچوب با استفاده از نظریه بازی، در نظر گرفته شده است. هزینه‌های در نظر گرفته شده عبارتند از: اقدامات مهاجم در لایه سایبری، اثرات حمله

[1] Risk

از لایه سایبری به اثرات بر روی سیستم فیزیکی، اقدامات مدافع در لایه سایبری بر حسب تقویت امنیت، اقدامات مدافع در لایه فیزیکی بر حسب فنون عملیاتی جدید. همچنین نویسندگان ذکر کرده‌اند که چگونه محیط‌های آزمایش سایبر-فیزیکی می‌توانند برای ارزیابی تحقیق امنیتی، و انجام مطالعه واقعی دفاع و حمله برای محیط‌های شبکه برق هوشمند مورد استفاده قرار گیرند.

در [۱۳] یک روش مبتنی بر نظریه بازی برای مطالعه حمله و دفاع در سیستم‌های سایبر-فیزیکی ارائه شده است. در این روش تعدادی فرمول برای جنبه‌های حمله و دفاع سیستم‌های سایبر-فیزیکی، تحت توابع سود و هزینه و بودجه‌های مختلف مهاجم و مدافع ارائه شده است. نتیجه بازی ارائه شده، با استفاده از تابع سود خطی، نمایی منفی و S-Shaped بررسی شده و نتایج با یک تعادل نش مشخص شده‌اند. در این سیستم فرض شده است که سیستم سایبر- فیزیکی دارای تعداد مشخصی از هر منبع سایبری و فیزیکی است که حداقلی از هریک برای کارکرد سیستم با استفاده از این منابع تعیین شده است.

در [۱۴] هم یک بازی امنیتی برای سیستم‌های سایبر- فیزیکی ارائه شده است. در این مدل، هزینه‌هایی برای حملات و اقدامات متقابل در برابر آن‌ها در نظر گرفته شده است. همچنین ماهیت محیط و فرایند کشف حمله به‌صورت تصادفی بیان شده‌اند. در نهایت یک حل‌کننده برای محاسبه راهبردها و هزینه آن‌ها، به روش برنامه‌ریزی خطی ارائه شده است. بازی ارائه‌شده، یک بازی max-min است که مهاجم و مدافع دو طرف بازی در نظر گرفته شده‌اند. مؤلفه‌های در نظر گفته شده در این بازی احتمال کشف، احتمال شکست حمله، احتمال موفقیت حمله و هزینه حمله است.

در [۱۵] از نظریه بازی برای مدل‌سازی احتمالات حملات موفق در فضای سایبری و فیزیکی به‌صورت تابعی از تعداد منابعی که مورد حمله قرار می‌گیرند یا تحت دفاع قرار دارند بیان شده است. این روش همچنین دیدی را نسبت به بازیابی زیرساخت‌های حیاتی و تخصیص منبع بهینه تحت مقادیر هزینه و هدف مختلف که ممکن است بازیکنان داشته باشند، فراهم می‌آورد.

در [۱۶] نویسندگان بر زمان انجام حملات متمرکز شده‌اند و نشان داده‌اند که زمان انجام حملات امنیتی (مانند حمله جلوگیری از سرویس در برابر سیگنال‌های حسگرها و کنترل‌کننده‌ها) تأثیر به‌سزایی در مؤثر بودن حمله و بروز آسیب به فرایند فیزیکی دارد. در این مقاله، رفتار زمانی حملات بررسی شده و نشان داده شده است که اگر حمله در زمان صحیح انجام شود (مثلاً بسته به آنکه آخرین داده حسگر چه بوده است)، سیستم ممکن است به حالت ناامن وارد شود. همچنین در صورتی که حمله در زمان مناسب آغاز نشود نمی‌تواند تأثیرات مخربی بر سیستم داشته باشد.

در [۱۷] یک زبان برای توصیف حملات به سیستم‌های سایبر- فیزیکی ارائه شده است. هسته این زبان یک طبقه‌بندی از حمله‌ها به سیستم‌های سایبر- فیزیکی است. رده‌بندی، جنبه‌های منطقی مجزایی از حمله‌ها به سیستم‌های سایبر- فیزیکی که باید توصیف شوند را مشخص می‌کند. این زبان می‌تواند علاوه بر حملات سایبری مرسوم، حملات بین حوزه‌ای را نیز توصیف کند. این زبان ساختاری را برای حملات مختلف به سیستم‌های سایبر- فیزیکی فراهم می‌کند که پیش‌نیاز مهمی برای تحلیل کیفی و کمّی حملات به سیستم‌های سایبر- فیزیکی است. مبنای این زبان یک رده‌بندی شش بعدی است. هدف اصلی این رده‌بندی، توصیف ویژگی‌هایی است که باید در توصیف حملات به سیستم‌های سایبر- فیزیکی در نظر گرفته شوند. همچنین ساختار مشابهی نیز می‌تواند برای توصیف اقدامات متقابل در برابر حملات سایبر- فیزیکی استفاده شود. معیارهای حمله و دفاع برحسب یک کنش (برای

مثال اجرای یک روش) توصیف شده‌اند که موفقیت آن‌ها به برقرار بودن یک یا چند پیش‌شرط بستگی دارد. اجرای یک روش، یک یا چند تغییر را در سیستم ایجاد می‌کند. تغییرات می‌توانند تغییرات آنی[1] باشند (که به گروه سبب[2] تعلّق دارند) یا تغییرات دنباله‌دار باشند (که با گروه اثر[3] تعلّق دارند). به دلیل وابستگی پیچیده متقابل در سیستم‌های سایبر- فیزیکی هر تغییر در سیستم‌های سایبر- فیزیکی یک تغییر آنی و دنباله‌ای از اثرات و پیامدهای بعدی خواهد داشت. نویسندگان حمله سایبری، استاکس‌نت را به‌عنوان مطالعه موردی با این زبان توصیف کرده‌اند.

در [۱۸] نویسندگان روشی را برای مدل کردن حملات و راهبردهای دفاعی سیستم‌های سایبر- فیزیکی ارائه کرده‌اند. آن‌ها سه نوع خرابی را برای این سیستم‌ها در نظر گرفتند که عبارتند از: ساییدگی، خرابی و خروج از دور[4]. روش ارائه شده برمبنای شبکه‌های پتری تصادفی است و خروجی آن تعیین مقدار بهینه مؤلفه‌های طراحی سیستم مانند سطح بهینه میزان افزونگی مؤلفه‌هاست.

در [۱۹] نویسندگان روشی را برای ارزیابی مخاطره امنیتی سیستم‌های سایبر- فیزیکی با استفاده از یک بستر آزمایشی با کنترل‌کننده‌های صنعتی در دنیای واقعی و پروتکل‌های ارتباطی ارائه کرده‌اند. آزمایش‌های آن‌ها نشان می‌دهد که همه حملات نمی‌توانند باعث آسیب فیزیکی شوند و توسعه آن زمان می‌برد و ممکن است بتوان پیامد حمله را خنثی کرد.

در [۲۰] نویسندگان برخی از تحقیقات موجود در مورد امنیت سیستم‌های سایبر- فیزیکی را مطالعه و طبقه‌بندی کرده‌اند. آن‌ها به‌طور کلی بر روی مدل‌سازی و شناسایی حملات، مشکلات اصلی موجود در برآورد پیامدهای حملات علیه این سیستم‌ها و راه‌حل‌های آن‌ها و توسعه معماری امنیتی تمرکز کردند.

در جدول ۱ تعدادی از کارهای پیشین که در مطالعه امنیت سیستم‌های سایبر- فیزیکی انجام داده‌ایم و تحقیقات سایرین در سال اخیر بررسی شده است.

در مقایسه با روش‌های پیشین، روش پیشنهادی، اطلاعات ناکامل مهاجم و سیستم از یکدیگر را در نظر دارد و براساس آن مدل بازی دو نفره غیر- مجموع صفر به‌صورت مدل بازی بیزی ارائه می‌دهد. با این روش نبودقطعیت در رفتار سیستم و مهاجم بهتر مدل می‌شود. با توجه به نوع بازی با اطلاعات ناقص، روش پیش‌بینی رفتار مهاجم، متفاوت خواهد بود. مؤلفه‌های در نظر گرفته‌شده در مدل بازی، دانش مهاجم، هزینه حمله، هزینه سرمایه‌گذاری امنیت سیستم، خسارت سیستم از حمله موفق، سود مهاجم از حمله موفق و منفعت سیستم از کشف حمله است. همچنین خروجی مدل سنجه‌های[5] کمّی مخاطره امنیتی با یک رویکرد جدید، احتمال حمله و کشف حمله هستند که در مقایسه با کارهای انجام‌شده متمایز است.

[1] Immediate
[2] Cause
[3] Effect
[4] Exfiltration
[5] Metrics

جدول ۱. مرور شماری از مطالعات اخیر در حوزه مطالعه امنیت سیستم‌های سایبر- فیزیکی.

شماره مرجع	شرح کار انجام شده
[۶]	ارائه روشی برای مدل‌سازی و ارزیابی امنیت سیستم‌های سایبر- فیزیکی با استفاده از شبکه پتری تصادفی. با استفاده از این روش، تأثیر مؤلفه‌های دفاعی و هجومی مانند بازه زمانی کشف حمله، زمان تا بروز اختلال فیزیکی و احتمال نادرست- مثبت سیستم‌های کشف نفوذ قابل‌تخمین خواهند بود.
[۲۱]	ارائه یک روش مدل‌سازی برای پیش‌بینی رفتار مهاجم به سیستم‌های سایبر- فیزیکی و ارزیابی پیامد حملات با استفاده از درخت حمله فازی.
[۲۲]	ارائه یک روش مدل‌سازی برای ارزیابی تأثیر حملات بر سیستم‌های سایبر- فیزیکی با استفاده از زنجیره مارکوف تصادفی و نظریه بازی‌ها. از نظریه بازی‌ها برای مطالعه رفتار مهاجم در مراحل نفوذ و ایجاد اختلال فیزیکی استفاده شده است.
[۲۳]	ارائه یک روش مدل‌سازی مبتنی بر نظریه‌بازی‌ها برای ارزیابی امنیت سیستم‌های سایبر- فیزیکی. در نظر گرفتن مؤلفه‌های جدید امنیتی از جمله بازه زمانی کشف حمله، دانش مهاجم و جریمه مهاجم
[۷]	ارائه روشی برای ارزیابی انتشار پیامد حملات امنیتی به سیستم‌های سایبر- فیزیکی. امکان اولویت‌بندی حملات انجام شده از نظر خطر امنیتی و نوع پیامدها.
[۲۴]	ارائه یک متدولوژی زمان طراحی برای ارائه رویکردی برای برآورد امنیت سیستم‌های سایبر- فیزیکی به‌صورت کیفی و کمّی با استفاده از شبکه‌های پتری تصادفی.
[۲۵]	ارائه یک مدل تخمین مخاطره احتمالی برای سیستم‌های سایبر- فیزیکی. مدل ارائه شده بر این واقعیت بنا شده است که حملات چند مرحله‌ای، دنباله‌ای از حملات اولیه هستند که در آن حریف سطح دسترسی لازم را برای انجام آن‌ها از طریق حملات قبلی در زنجیره به‌دست می‌آورد.
[۲۶]	توسعه یک روش ساده و کاربردی برای ارزیابی سطح امنیت سیستم‌ها و تجهیزات سایبر- فیزیکی که یک تأسیسات الکتریکی را تشکیل می‌دهند. روش پیشنهادی به‌عنوان نقطه شروعی برای ارائه مجموعه‌ای از آزمون‌های انطباق، برای ارزیابی تجهیزاتی که قرار است در یک شبکه برق نصب یا اصلاح شوند، قابل استفاده است.
[۲۷]	انجام تجزیه‌وتحلیل امنیتی سیستم‌های سایبر- فیزیکی با استفاده از یادگیری ماشین. دید مناسبی از تحلیل امنیت با استفاده از یادگیری ماشین ارائه شده است.
[۲۸]	ارائه یک مدل تزریق داده نادرست جدید در کانال ارتباطی بین کنترل‌کننده به محرّک، با اثرات ناهمگن برای سیستم‌های سایبر- فیزیکی.
[۲۹]	ارائه چارچوبی برای مدل‌سازی امنیت و تصدیق رسمی آن با تمرکز بر ویژگی بقاپذیری سیستم‌های سایبر- فیزیکی. این چارچوب در نهایت به‌عنوان مطالعه موردی بر یک خودرو به‌عنوان یک سیستم سایبر- فیزیکی اعمال شده است.
[۳۰]	ارائه روشی برای ارزیابی امنیت سیستم‌های سایبر- فیزیکی با استفاده از شبکه پتری وزن‌دار رنگی و نظریه بازی بیزی. مدل بازی ارائه شده یک بازی مجموع صفر و تنها شامل مؤلفه‌های هزینه و خسارت است. سنجه موردنظر، احتمال حمله به هر گره از این سیستم‌ها با حل مدل شبکه پتری وزن‌دار رنگی است.

روش پیشنهادی

در این بخش، به تفصیل روش پیشنهادی را توصیف می‌کنیم. ابتدا به تشریح مدل بازی پیشنهادی، ساختار آن و مؤلفه‌های آن می‌پردازیم، سپس در مورد حل مدل بازی پیشنهادی خواهیم پرداخت و در نهایت سنجه‌های کمّی موردنظر را بیان خواهیم کرد.

توصیف روش پیشنهادی

در این روش، به‌منظور مطالعه رفتار سیستم و مهاجم از نظریه بازی استفاده شده است. مدل پیشنهادی، یک مدل بازی دونفره بین مهاجم و سیستم، غیرمجموع صفر و با اطلاعات ناقص است. به‌طور صوری، مدل بازی پیشنهادی به‌صورت یک چندتایی $G = <P, S_i, U_{ij}>$ تعریف می‌شود که P در آن مجموعه بازیکنان، S_i مجموعه اقدامات یا راهبردهای بازیکن i، و U_{ij} سودمندی بازیکن i از انتخاب راهبرد j است. به بازی‌ای که بعضی از بازیکنان از منفعت سایرین اطلاعی ندارند بازی با اطلاعات ناقص گفته می‌شود. در حقیقت، در اغلب مواقع سیستم و مهاجم اطلاعات کاملی از یکدیگر ندارند؛ بنابراین بهترین گزینه برای مدل کردن تقابل بین آن‌ها استفاده از روش مدل‌سازی مبتنی بر بازی بیزی با اطلاعات ناقص است.

به زبان ریاضی، مجموعه بازیکنان P به‌صورت $P = \{A, S\}$ تعریف می‌شود که A مجموعه انواع مهاجمان و S مجموعه انواع سیستم است. با فرض اینکه Ta_i نوع مهاجم i باشد، انواع مهاجمان A_i توزیع احتمال $A=\{a_1, a_2, \ldots, a_n\}$ را خواهند داشت. به‌طور مشابه برای انواع سیستم S_i هم توزیع احتمال $S=\{s_1, s_2, \ldots, s_n\}$ وجود خواهند داشت.

در تعادل نش [۹] فرض می‌شود که هر بازیکن، ترجیحات سایر بازیکنان را به‌درستی می‌داند در حالی که در بسیاری از مواقع، بازیکنان اطّلاعات کاملی در مورد تصمیم و انتخاب سایر بازیکنان ندارند. راه پیشنهادی برای مدل‌سازی بازی‌های با اطّلاعات ناقص، در نظر گرفتن طبیعت به‌عنوان یک بازیگر بازی است که پیش از تصمیم بازیکنان نوع آن‌ها را تعیین می‌کند. با اعمال این تغییر، اطّلاعات ناقص تعدادی از بازیکنان در مورد منفعت سایرین به اطّلاعات ناقص در مورد حرکت طبیعت در انتخاب نوع بازیکنان تبدیل می‌شود. با این روش، بازی‌های با اطلاعات ناقص هم به همان روش بازی‌های مرسوم مدل‌سازی می‌شوند.

در بازی بیزی راهبرد بازیکنان قبل از آن که طبیعت نوع بازیکن را مشخّص کند تعیین می‌شود در واقع هر بازیکن پیش از انتخاب نوع تصمیم می‌گیرد که به ازای هر نوع موجود، چه عملی را انجام دهد. تعادل نش بیزی به‌عنوان یک نمایه راهبرد تعریف می‌شود که با توجه به نوع هر بازیکن و با توجه به راهبردهایی که سایر بازیکنان انجام می‌دهند، سود موردانتظار را برای هر بازیکن به حداکثر می‌رساند. به این معنا که نمایه راهبرد x یک تعادل نش بیزی است اگر و فقط اگر برای هر بازیکن i، و با فرض ثابت بودن راهبرد هر بازیکن دیگر، راهبرد x_i سود مورد انتظار بازیکن i را به حداکثر برساند. در مورد نقاط تعادل بازی در بازی بیزی، در صورتی که بازیکن یک، از نوع بازیکن دوم مطلع نباشد، انتخاب a_1 او بهینه است در صورتی که رابطه زیر برقرار باشد:

$$p \times u_1(a_1, a_2(t_1)) + (1-p) \times u_1(a_1, a_2(t_2)) \geq p \times u_1(a_1', a_2(t_1)) + (1-p) \times u_1(a_1', a_2(t_2)); \forall a_1' \quad (۱)$$

که p احتمال انتخاب بازیگر نوع یک، u_1 تابع سودمندی بازیگر یک، $a_2(t_1)$ انتخاب بازیکن دو اگر بازیکن دو نوع یک باشد و $a_2(t_2)$ انتخاب بازیکن دو اگر بازیکن دوم نوع دو باشد.

در صورتی که بازیکن دوم از نوع بازیکن اول مطلّع باشد، برای تعادل بازی خواهیم داشت:

$$u_2(a_1, a_2(t_1), t_1) \geq u_2(a_1, a_2'(t_1), t_1); \forall a_2' \quad (۲)$$

که t_1 و t_2 نوع بازیکن دوم، u_2 تابع سودمندی بازیکن دوم و $a_2(t_1)$ انتخاب بازیکن دو است اگر بازیکن دوم از نوع یک باشد. این رابطه برای حالتی که بازیکن دوم نوع دو t_2 باشد نیز برقرار است. در این حالت چون بازیکن نوع خود را می‌شناسد، احتمال p اهمیتی برای او ندارد. راهبرد مهاجم به‌صورت انجام حمله یا نبودن حمله در نظر گرفته می‌شود: $S_A = \{A, NA\}$. همچنین، راهبرد سیستم به‌صورت کشف و کشف‌نشدن حمله و بی‌نظمی ناشی از حمله در نظر گرفته می‌شود: $S_S = \{D, ND\}$.

در این مدل فرض شده است که سیستم از میزان دانش مهاجم، اطلاع کاملی ندارد و به همین دلیل برآورد می‌کند که میزان دانش مهاجم با احتمال p_k کمتر از k' و با احتمال $1-p_k$ بالاتر از k' خواهد بود. از طرفی، در صورتی حمله مهاجم، تأثیر خراب‌کارانه بر سیستم خواهد داشت که سطح دانش او بالاتر از k' در مورد سیستم، فرایند فیزیکی، پردازش سیگنال، کنترل، و دانش سایبری باشد. براساس این موضوع، میزان سود مهاجم از حمله به سیستم، هزینه سرمایه‌گذاری سیستم در ارتقای امنیت متفاوت خواهد بود. در نتیجه می‌توان فرض کرد که با دو نوع مهاجم روبه‌رو هستیم: مهاجم نوع یک با سطح دانش پایین که حمله او جز ایجاد تأثیر جزئی احتمالی پیامد قابل‌توجهی نخواهد داشت و مهاجم نوع دو با سطح دانش بالا که اقدام او در صورت موفقیت می‌تواند منجر به بروز اختلال شدید یا خرابی فیزیکی گردد. با توجه به توضیحات ذکر شده، بازیکنان اطلاعات کاملی درباره یکدیگر ندارند؛ بنابراین بازی، یک بازی بیزی با اطلاعات ناقص است.

یک راه تعریف کردن بازی، فهرست کردن واضح تمام راهبردهای ممکن و سودهای بازیکنان است. توصیف بازی به این روش با تابع سودمندی یا هزینه، شکل استاندارد یا ماتریس بازی نامیده می‌شود. در واقع ساختار اصلی نظریه بازی در بیشتر تحلیل‌ها شامل ماتریسی چند بعدی است که در هر بعد مجموعه‌ای از گزینه‌ها قرار گرفته‌اند و درایه‌های این ماتریس سود کسب شده برای بازیکنان به‌ازای ترکیب‌های مختلف از گزینه‌های موردانتظار است.

به‌منظور مدل‌سازی این بازی که قابل حل با استفاده از روش حل بازی‌های مرسوم باشد، شکل استاندارد بازی به‌صورت دو مدل تجزیه شده است. جدول ۲ و ۳، به‌ترتیب، شکل استاندارد مدل بازی برای مهاجم نوع یک و دو را نشان می‌دهند. سطر اول در هر خانه جدول، سود مهاجم و سطر دوم سود سیستم را نشان می‌دهد.

مؤلفه‌های مدل بازی موردنظر به‌صورت زیر هستند که برای بازی با مهاجم نوع یک و دو به‌ترتیب با اندیس صفر و یک نمایش داده شده‌اند:

- مؤلفه K نشان‌دهنده سطح دانش مهاجم در مورد فرایند فیزیکی هدف و شرایط خرابی آن است.
- مؤلفه R میزان سود مهاجم از بروز اختلال فیزیکی به سیستم را نشان می‌دهد. این مؤلفه همچنین ضرر واردشده به سیستم بر اثر بروز اختلال فیزیکی را نشان می‌دهد.
- مؤلفه E نشان‌دهنده هزینه لازم برای سیستم به‌منظور سرمایه‌گذاری در امنیت با هدف کشف نفوذ و بی‌نظمی، به‌منظور جلوگیری از بروز اختلال فیزیکی توسط مهاجمین به سیستم است.
- مؤلفه W سود سیستم ازکشف حمله مهاجم را نشان می‌دهد.
- مؤلفه B منفعت و جایزه سیستم از کشف حمله مهاجم به قصد بروز اختلال فیزیکی در سیستم است.
- مؤلفه P_a احتمال اقدام به حمله مهاجم، به منظور ایجاد اختلال فیزیکی به سیستم را نشان می‌دهد.
- مؤلفه P_d احتمال کشف حمله سیستم است.

جدول ۲. شکل استاندارد مدل بازی برای مهاجم نوع یک.

		سیستم	
		کشف نشدن (*ND*)	کشف (*D*)
مهاجم	حمله (*A*)	$K \times R_0 - C_0$ $-K \times R_0 - W_0$	$-F_0 - C_0$ $B_0 - K \times E_0$
	عدم حمله (*NA*)	۰ ۰	۰ $-K \times E_0$

جدول ۳. شکل استاندارد مدل بازی برای مهاجم نوع دو.

		سیستم	
		کشف نشدن (*ND*)	کشف (*D*)
مهاجم	حمله (*A*)	$K \times R_1 - C_1$ $-K \times R_1 - W_1$	$-C_1$ $B_1 - K \times E_1$
	عدم حمله (*NA*)	۰ ۰	۰ $-K \times E_1$

اکنون به توضیح نحوه محاسبه سود بازیکنان می‌پردازیم. ابتدا نمایه راهبرد (*A, ND*) را در نظر می‌گیریم. در این حالت، مهاجم پاداش R_0 را در صورت حمله موفقیت‌آمیز و پاداش $K \times R_0$ را به دلیل ایجاد اختلال فیزیکی در سیستم کسب کرده است. با توجه به اینکه $0 \leq K \leq 1$ فرض شده است میزان خسارت وارد شده بر سیستم با میزان دانش مهاجم از سیستم و فرایند فیزیکی تحت کنترل متناسب است. مهاجم همچنین مقدار C_0 را به‌منظور حمله به سیستم هزینه کرده است. سیستم هم ضرر $K \times R_0$ را به دلیل خسارت واردشده متحمل خواهد شد و به اندازه W_0 به دلیل کشف نشدن حمله، جریمه خواهد شد. بنابراین برای سود مهاجم در این نمایه راهبرد ($U_A(A, ND)$) خواهیم داشت:

$$U_A(A, ND) = K \times R_0 - C_0 \quad (۳)$$

همچنین برای سود سیستم خواهیم داشت:

$$U_S(A, ND) = -K \times R_0 - W_0 \quad (۴)$$

حالت دوم حالتی است که مهاجم و سیستم نمایه راهبرد (*A, D*) را انتخاب کرده‌اند. در این صورت مهاجم مقدار C_0 را به‌منظور حمله به سیستم هزینه کرده است. در این نمایه راهبرد، سیستم به میزان B_0 به دلیل کشف موفقیت‌آمیز بهره خواهد برد و از طرفی هزینه‌ای به میزان $K \times E_0$ به دلیل سرمایه‌گذاری در امنیت به منظور کشف حمله، متحمل خواهد شد. بنابراین، برای سود مهاجم خواهیم داشت:

$$U_A(A, I) = -C_0 \quad (۵)$$

همچنین برای سود سیستم رابطه زیر به‌دست می‌آید:

$$U_S(A,\ I) = B_0 - K \times E_0 \quad (۶)$$

برای نمایه راهبرد $(NA,\ ND)$ از آنجایی که حمله و کشفی رخ نداده است، سود مهاجم و سیستم صفر خواهد بود:

$$U_S(NA, ND) = U_A(NA, ND) = 0 \quad (۷)$$

در نهایت برای نمایه راهبرد $(NA,\ D)$ سود مهاجم به دلیل تصمیم به نبود حمله صفر و سود سیستم به دلیل سرمایه‌گذاری به‌منظور کشف حمله برابر است با:

$$U_S(NA, I) = -K \times E_0 \quad (۸)$$

از آنجایی که مقادیر مدل بازی برای مهاجم، نوع دو متفاوت است، مؤلفه‌های میزان خسارت واردشده به سیستم R، هزینه حمله مهاجم C، میزان سرمایه‌گذاری سیستم در کشف حمله E، جریمه سیستم از کشف نشدن W و سود سیستم از کشف حمله B برای دو مدل بازی متفاوت خواهد بود و با اندیس صفر و یک به‌ترتیب برای مدل بازی مهاجم نوع یک و نوع دو در نظر گرفته شده است. همچنین احتمال انتخاب نوع مهاجم یک توسط طبیعت هم p_k و احتمال انتخاب نوع مهاجم دو توسط طبیعت هم $1\text{-}p_k$ خواهد بود.

حل مدل بازی پیشنهادی

حل یک بازی به معنی یافتن راهبردهایی برای رسیدن به نقاط تعادل در بازی است. این راهبردها اصولاً از قواعد عقلانی به نتیجه می‌رسند. مشهورترین تعادل‌ها، تعادل نش است. براساس نظریه تعادل نش، اگر فرض کنیم در هر بازی با راهبرد ترکیبی، بازیکنان به طریق منطقی و معقول راهبردهای خود را انتخاب کنند و به دنبال حداکثر سود در بازی هستند، دست کم یک راهبرد برای به دست آوردن بهترین نتیجه برای هر بازیکن قابل انتخاب است و چنانچه بازیکن راه‌کار دیگری به غیر از آن را انتخاب کند، نتیجه بهتری به دست نخواهد آورد.

در ادامه به حل بازی و یافتن نقاط تعادل آن خواهیم پرداخت. ابتدا تعادل نش بازی را بررسی خواهیم کرد و در ادامه تعادل نش ترکیبی را مورد مطالعه قرار خواهیم داد. از آنجایی که پارامترهای دو مدل بازی در اندیس تنها متفاوت هستند، حل بازی با حذف اندیس انجام شده است تا برای هر دو مدل قابل استفاده باشد. به‌منظور استفاده از این حل، حتماً مقادیر متناظر با هر مدل در نظر گرفته خواهد شد.

تعادل نش: ابتدا تعادل نش از نظر مهاجم را در نظر می‌گیریم. از آنجایی که او از حالت خودآگاه است می‌تواند نوع خود را مشخص کند؛ بنابراین صرف نظر از اندیس مشخص‌کننده نوع، برای او خواهیم داشت:

۱- اگر $K > max(C/R\ ,(B+W)/(E-R))$، در آن صورت نمایه راهبرد $(A,\ ND)$ یک نقطه تعادل از نظر مهاجم خواهد بود.

برای آنکه بازی در نمایه راهبرد $(A,\ ND)$ دارای نقطه تعادل باشد، طبق تعریف باید داشته باشیم:

$$K \times R - C > 0 \to K > C/R, -K \times R - W > B - K \times C \to K > \frac{B+W}{E-R} \quad (۹)$$

۲- در صورتی که $K < C / R$ باشد، در آن صورت نمایه راهبرد (NA, ND) یک نقطه تعادل از نظر مهاجم خواهد بود.

برای آنکه بازی در نمایه راهبرد (NA, ND) دارای نقطه تعادل باشد، طبق تعریف باید داشته باشیم:

$$K \times R - C < 0, -K \times E < 0 \rightarrow K < C/R \quad (۱۰)$$

به‌طور مشابه، بررسی سایر نمایه پروفایل‌ها نشان می‌دهد که سیستم در آن‌ها دارای نقطه تعادل نخواهد بود.

اکنون به تعادل نش از نظر سیستم می‌پردازیم. از آنجایی که سیستم نوع مهاجم را نمی‌داند بنابراین تعادل او باید از رابطه (۱) محاسبه گردد:

۳- در صورتی که رابطه (۱۱) و (۱۲) بر قرار باشند، در آن صورت نمایه راهبرد (A, ND) یک نقطه تعادل از نظر سیستم خواهد بود.

برای آنکه بازی در نمایه راهبرد (A, ND) دارای نقطه تعادل از نظر سیستم باشد، طبق تعریف باید داشته باشیم:

$$p \times (K \times R_0 - C_0) + (1 - p) \times (K \times R_1 - C_1) \geq 0 \quad (۱۱)$$

همچنین

$$p \times (-K \times R_0 - W_0) + (1 - p) \times (-K \times R_1 - W_1) \geq p \times (B_0 - K \times E_0) + (1 - p) \times (B_1 - K \times E_1) \quad (۱۲)$$

۴- در صورتی که رابطه (۱۳) برقرار باشد، نمایه راهبرد (NA, ND) یک نقطه تعادل از نظر سیستم خواهد بود:

برای آنکه بازی در نمایه راهبرد (NA, ND) دارای نقطه تعادل از نظر سیستم باشد، طبق تعریف باید داشته باشیم:

$$p \times (K \times R_0 - C_0) + (1 - p) \times (K \times R_1 - C_1) \leq 0 \quad (۱۳)$$

تعادل نش ترکیبی: تعادل نش ترکیبی مدل بازی ارائه شده برای مهاجم به صورت زیر خواهد بود:

$$U_{attacker}(A) = U_{attacker}(NA) \rightarrow (1 - P_{d0}) \times (K \times R_0 - C_0) + P_{d0} \times (-C_0) = 0 \rightarrow P_{d0} = (K \times R_0 - C_0)/(K \times R_0) = 1 - \frac{C_0}{K \times R_0} \quad (۱۴)$$

برای تعادل نش ترکیبی سیستم خواهیم داشت:

$$U_{system}(NI) = U_{system}(I) \rightarrow (P_{a0}) \times (-K \times R_0 - W_0) + (1 - P_{a0}) \times 0 = (P_{a0}) \times (B_0 - K \times E_0) + (1 - P_{a0}) \times (-K \times E_0) \rightarrow P_{a0} = \frac{K \times E_0}{B_0 + (K \times R_0) + W_0} \quad (۱۵)$$

به این شکل، احتمال کشف حمله سیستم در مدل بازی با مهاجم نوع یک با استفاده از رابطه ۱۴ و احتمال حمله مهاجم با استفاده از رابطه ۱۵ قابل محاسبه است.

اکنون به بررسی تعادل نش ترکیبی مدل بازی با مهاجم نوع دو می‌پردازیم. ابتدا، برای مهاجم خواهیم داشت:

$$U_{attacker}(A) = U_{attacker}(NA) \rightarrow \\ (1 - P_{d1}) \times (K \times R_1 - C_1) + P_{d1} \times (-C_1) = 0 \\ \rightarrow P_{d1} = (K \times R_1 - C_1)/(K \times R_1) = 1 - \frac{C_1}{K \times R_1} \quad (۱۶)$$

همچنین برای سیستم خواهیم داشت:

$$U_{system}(NI) = U_{system}(I) \rightarrow \\ (P_{a1}) \times (-K \times R_1 - W_1) + (1 - P_{a1}) \times 0 = \\ (P_{a1}) \times (B_1 - K \times E_1) + (1 - P_{a1}) \times (-K \times E_1) \\ \rightarrow P_{a1} = \frac{K \times E_1}{B_1 + (K \times R_1) + W_1} \quad (۱۷)$$

به این ترتیب، احتمال کشف حمله سیستم در مدل بازی با مهاجم نوع دو P_{d1} با استفاده از رابطه ۱۶ و احتمال حمله مهاجم P_{a1} با استفاده از رابطه ۱۷ قابل محاسبه است.

اکنون به محاسبه احتمال حمله مهاجم و احتمال کشف سیستم در حالت سراسری و با در نظر گرفتن دو مدل بازی ارائه شده می‌پردازیم. همان‌طور که بیان شد احتمال انتخاب نوع مهاجم یک توسط طبیعت p_k و احتمال انتخاب نوع مهاجم دو توسط طبیعت $1\text{-}p_k$ خواهد بود.

بنابراین احتمال حمله مهاجم P_a با استفاده از رابطه زیر محاسبه می‌شود:

$$P_a = P_k \times P_{a0} + (1 - P_k) \times P_{a1} \quad (۱۸)$$

همچنین برای احتمال کشف حمله توسط سیستم P_d در مدل بازی پیشنهادی مطابق با رابطه زیر خواهیم داشت:

$$P_d = P_k \times P_{d0} + (1 - P_k) \times P_{d1} \quad (۱۹)$$

به‌طور مشابه با تعادل نش، از آنجا که فرض شده است مهاجم می‌داند در چه حالتی قرار دارد و از چه سطحی از دانش در مورد سیستم برخوردار است، در این حالت براساس سطح دانش خود، احتمال حمله و کشف را به‌ترتیب از رابطه‌های (۱۴) و (۱۵) یا (۱۶) و (۱۷) بسته به نوع خود محاسبه می‌کند و نیازی به استفاده از رابطه‌های (۱۸) و (۱۹) ندارد.

سنجه‌های کمّی

اکنون به توصیف سنجه‌های کمّی در نظر گرفته شده در این روش می‌پردازیم. سنجه‌های امنیتی که با استفاده از روش ارائه‌شده ارزیابی شدند به‌صورت زیر هستند:

- **احتمال حمله مهاجم**: این سنجه نشان می‌دهد که مهاجم با چه احتمالی حمله خواهد کرد. این سنجه با استفاده از مدل نظریه بازی پیشنهادی محاسبه می‌شود که نتیجه به‌دست‌آمده براساس مؤلفه‌های دفاعی و هجومی است.
- **احتمال کشف حمله**: این سنجه نشان می‌دهد که سیستم با سرمایه‌گذاری‌هایی که انجام می‌دهد چه احتمال کشفی را باید با توجه به مؤلفه‌های مختلف بازی داشته باشد. این سنجه با استفاده از حل مدل بازی پیشنهادی محاسبه می‌شود.
- **مخاطره امنیتی**: با توجه به اهمیت زمان تا خرابی سیستم بعد از حمله موفق، از رابطه زیر برای محاسبه مخاطره r استفاده می‌شود:

$$r = \frac{P_a}{T_f \times P_d} \times M \qquad (۲۰)$$

که P_a احتمال حمله مهاجم، T_f زمان تا ورود آسیب به سیستم بعد از حمله، P_d احتمال کشف حمله توسط سیستم و M میزان خسارت واردشده به سیستم در اثر موفقیت‌آمیز بودن حمله است. در واقع هرچه احتمال حمله کم و زمان تا خرابی سیستم زیاد باشد، مخاطره حمله کاهش خواهد یافت. از طرفی هرچه احتمال کشف حمله کاهش یابد و خسارت واردشده به سیستم بیشتر شود با مخاطره بالاتری مواجه خواهیم بود.

مطالعه موردی

سیستم موردمطالعه، یک سیستم آزمایشگاهی است که از دو مخزن که در ارتفاع متفاوتی نسبت به یکدیگر قرار گرفته‌اند تشکیل شده است [۳۱]. شکل ۲، سیستم در نظر گرفته شده را به تصویر کشیده است. ارتفاع هر دو مخزن ۱۰۰ سانتی‌متر است و قطر آن‌ها به‌ترتیب ۱۲ و ۵ سانتی‌متر است. این سیستم دارای سه شیر است: شیر آب ورودی (V_{in})، شیر میانی (V_1) و شیر خروجی (V_2). همچنین یک حسگر سطح آب (y_1) دارد که در مخزن دوم قرار گرفته است: $Y = \{ y_1 \}$. یک کنترل‌کننده نیز کارکرد سیستم را کنترل می‌کند و وضعیت شیرها را براساس داده‌های دریافتی از حسگر تعیین می‌کند.

فرض شده است که شیر ورودی تنها دو وضعیت باز و بسته دارد و جریان ورودی آن در وضعیت باز به‌صورت زیر تعریف می‌شود:

$$\dot{V}_{in} = 400 l/h \qquad (۲۱)$$

همچنین فرض شده است که دو شیر دیگر می‌توانند وضعیتی بین ۰ تا ۸۰ داشته باشند که صفر نشان‌دهنده کاملاً باز و ۸۰ نشان‌دهنده کاملاً بسته است. علاوه بر این، سرعت باز و بسته شدن شیرهای ۱ و ۲، ۸۰ ثانیه طول می‌کشد.

اکنون به توصیف روابط مدل فرایند سیستم می‌پردازیم. پویایی سطح آب در دو مخزن به‌صورت زیر تعریف می‌شود:

$$\dot{h}_1 = \frac{\dot{V}_{in} - \dot{V}_1}{A_1},\ \dot{h}_2 = \frac{\dot{V}_1 - \dot{V}_2}{A_2} \quad (۲۲)$$

که A_1 و A_2 به‌ترتیب مساحت سطح مخزن یک و دو هستند. رفتار پویایی شیر یک (V_1) به‌صورت زیر تعریف می‌شود:

$$\dot{V}_1 = \begin{cases} K_1.\sqrt{h_1 - (h_2 - H)} & if\ h_2 > H \\ K_1.\sqrt{h_1} & if\ h_2 \le H \end{cases} \quad (۲۳)$$

به‌طوری‌که H حد آستانه مشخص شده را نشان می‌دهد و اختلاف ارتفاع دو مخزن است.
همچنین رفتار پویای شیر دوم (V_2) از رابطه زیر قابل محاسبه است:

$$\dot{V}_2 = K_2.\sqrt{h_2} \quad (۲۴)$$

در این رابطه، K_1 و K_2 ضریب هستند و وابستگی بین جریان و وضعیت شیر را نشان می‌دهند. این ضرایب از رابطه‌های زیر قابل محاسبه هستند:

$$K_1 = \begin{cases} 1.85 \times 10^{-4} \times e^{-6\times10^{-6}\times P_1^3} \dfrac{m^{\frac{5}{2}}}{s} & if\ 0 \le P_1 < 80 \\ 0\ \dfrac{m^{\frac{5}{2}}}{s} & if\ P_1 = 80 \end{cases} \quad (۲۵)$$

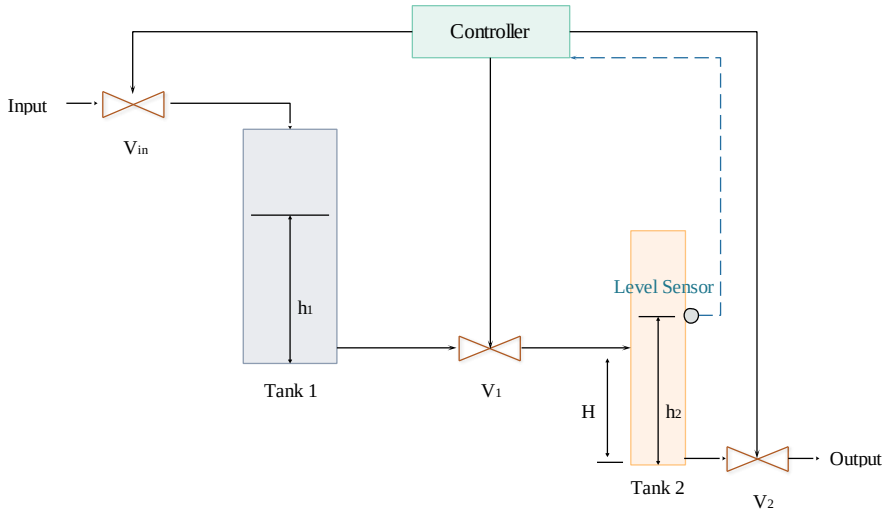
و

$$K_2 = \begin{cases} 2.26 \times 10^{-4} \times e^{-5.7\times10^{-6}\times P_2^3} \dfrac{m^{\frac{5}{2}}}{s} & if\ 0 \le P_2 < 80 \\ 0\ \dfrac{m^{\frac{5}{2}}}{s} & if\ P_2 = 80 \end{cases} \quad (۲۶)$$

سیستم دو مرحله عملیاتی دارد: مرحله شروع و مرحله عملیات پایدار. مرحله شروع از دو گام تشکیل می‌شود. در گام اول، شیر ورودی باز است و شیر یک بسته است. بعد از مدت‌زمان t_1، گام دوم شروع می‌شود. در این گام، شیر یک باز است و جریان به داخل مخزن دوم برقرار است. بعد از مدت‌زمان t_2، مرحله عملیات پایدار آغاز می‌شود. در این مرحله بسته به شرایط سیستم در گام سوم یا چهارم قرار دارد. در طول گام سوم، شیر دوم (خروجی) باز است. تا زمانی‌که سطح آب داخل مخزن دوم بالای حد آستانه L_{min} قرار داشته باشد، کنترل‌کننده در مرحله سوم باقی می‌ماند.
به محض اینکه سطح آب داخل مخزن دوم از حد آستانه تعیین‌شده پایین‌تر رود، کنترل‌کننده به گام چهارم وارد می‌شود تا سطح آب از حد آستانه L_{max} بالاتر رود. در گام چهارم، شیر خروجی بسته است. شکل ۳، رفتار سیستم در حالت عادی را نشان می‌دهد.

دو نیازمندی کلی برای این سیستم تعیین شده است. اول اینکه سرریزی اتفاق نیفتد و دوم اینکه سطح آب داخل مخزن دوم پایدار باشد و به‌طور دوره‌ای بین دو سطح L_{min} و L_{max} متغیر نباشد.

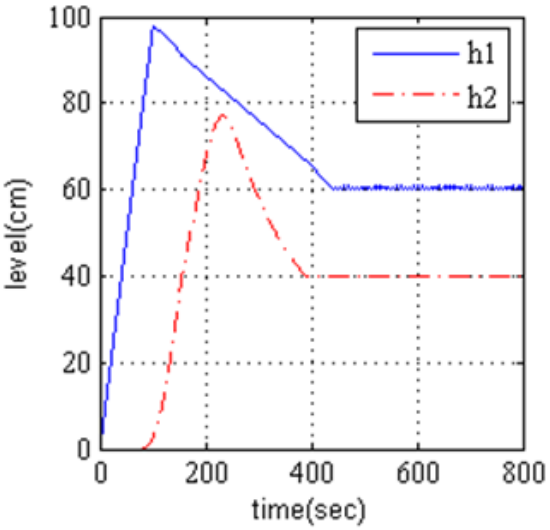
به‌منظور ارزیابی مدل ارائه‌شده برای سیستم آزمایشگاهی، ابتدا باید مدل‌های بازی ارائه‌شده حل شود. جدول ۲، مقادیر مؤلفه‌های بازی را ارائه می‌کند. با توجه به تعریف تعادل نش محض، با استفاده از داده‌های تعیین‌شده برای مدل بازی، هیچ تعادل نش محضی در بازی وجود ندارد؛ بنابراین به حل مدل‌های بازی ارائه‌شده با استفاده از رابطه‌های به‌دست‌آمده از مدل بازی می‌پردازیم.



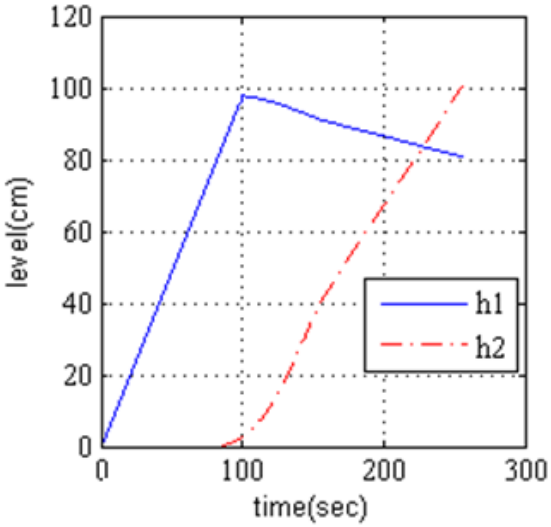
شکل ۲. سیستم آزمایشگاهی در نظر گرفته شده در مطالعه موردی [۳۱].

مهاجمی را درنظر می‌گیریم که دستورات بدخواهانه‌ای را به شیر خروجی ارسال می‌کند تا آن را در حالت بسته نگه دارد. دو حمله مختلف که هریک در مرحله‌های مختلف انجام شده‌اند را بررسی می‌کنیم. به‌عنوان حمله اول، فرض می‌کنیم حمله در مرحله آغازین کار سیستم و پیش از رسیدن به مرحله کارکرد پایدار انجام شده است. برای این کار، مهاجم سعی می‌کند فرایند را در حالت چهار نگه دارد. حمله در لحظه $T = ۱۰۱$ ثانیه آغاز می‌شود و سطح آب داخل مخزن در لحظه $T = ۲۵۵$ ثانیه به حالت نامطلوب (سرریز) می‌رسد. در این حالت ۱۵۴ ثانیه بعد از آغاز، حمله به نتیجه رسیده است (شکل ۴). به‌عنوان حمله دوم فرض می‌کنیم حمله در لحظه $T = ۶۰۰$ ثانیه، هنگامی که سیستم در مرحله کارکرد پایدار است انجام شود. در این مورد، سطح آب داخل مخزن دوم پس از ۱۱۹ ثانیه به بالای ۱۰۰ سانتی‌متر خواهد رسید (شکل ۵).

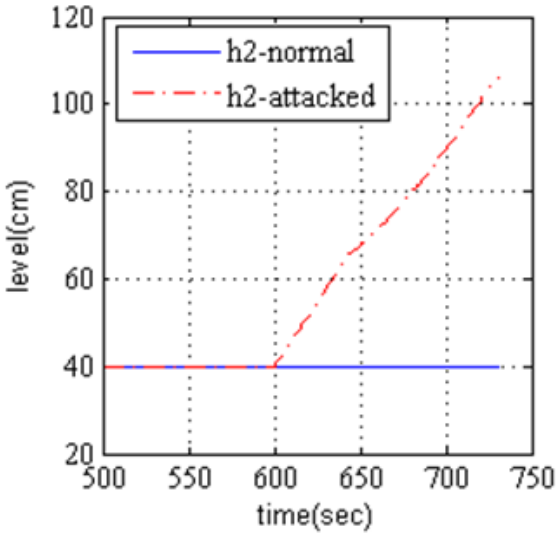
جدول ۴ و ۵ به‌ترتیب مقادیر مؤلفه‌های بازی در حالت مهاجم نوع یک و مهاجم نوع دو را نشان می‌دهد. فرض شده است در صورتی که سطح دانش مهاجم از ۰.۷۵ بیشتر باشد، مهاجم نوع دو شناخته می‌شود. این مؤلفه‌ها برای سیستم‌های مختلف، متفاوت است و برحسب ارزش دارایی که مهاجم می‌تواند هدف قرار دهد قابل محاسبه است.



شکل ۳. ارتفاع دو مخزن سیستم موردمطالعه در حالت عادی.



شکل ۴. ارتفاع دو مخزن سیستم موردمطالعه در حالت حمله اول.



شکل ۵. ارتفاع مخزن دوم سیستم موردمطالعه در حالت حمله دوم.

جدول ۴. مقادیر مؤلفه‌های بازی مدل مهاجم نوع یک.

مؤلفه	W_0	C_0	B_0	E_0	R_0
مقدار	KR_0	۰.۰۰۱	۰.۰۲	۰.۰۱	۰.۰۲

جدول ۵. مقادیر مؤلفه‌های بازی مدل مهاجم نوع دو.

مؤلفه	W_1	C_1	B_1	E_1	R_1
مقدار	KR_1	۰.۰۰۵	۰.۱	۰.۱	۰.۱۵

برای میزان خسارت واردشده به سیستم از آنجا که هدف هر دو حمله یکسان است M = ۱۰۰۰ و برای احتمال اینکه سیستم با مهاجم با دانش کم مواجه باشد p_k = ۰.۲ در نظر گرفته شده است. در هر آزمایش، تأثیر یک مؤلفه را بررسی می‌کنیم و برای سایر مؤلفه‌ها همان مقدار پیش‌فرض را در نظر می‌گیریم.

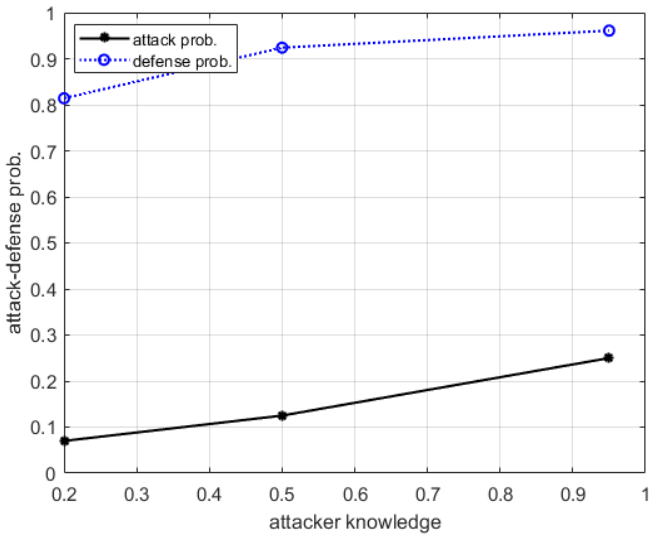
اکنون مقادیر مختلف مؤلفه دانش مهاجم را بررسی می‌کنیم. با در نظر گرفتن دانش مهاجم برابر با K = ۰.۲، برای احتمال حمله مهاجم و احتمال کشف سیستم برای مدل بازی مهاجم نوع یک با استفاده از رابطه‌های (۱۴) و (۱۵) به‌ترتیب خواهیم داشت: P_{a0} = ۰.۰۷ و P_{d0} = ۰.۷۵. همچنین برای احتمال حمله مهاجم و احتمال کشف سیستم برای مدل بازی مهاجم نوع دو با استفاده از رابطه‌های (۱۶) و (۱۷) به‌ترتیب خواهیم داشت: P_{a1} = ۰.۱۴ و P_{d1} = ۰.۸۳. بدین ترتیب احتمال حمله و کشف سراسری طبق رابطه (۱۸) و (۱۹) برابر خواهد بود با: P_a = ۰.۰۷ و P_d = ۰.۸۱۴. در این صورت مخاطره حمله برای حمله اول طبق رابطه (۲۰)، برابر r =(۰.۰۷*۱۰۰۰) / (۰.۸۱۴*۱۱۹) = ۰.۵۶ خواهد بود و مخاطره حمله برای حمله دوم، برابر r = (۰.۰۷*۱۰۰۰) / (۰.۸۱۴*۱۵۴) = ۰.۵۶

۰.۷۲ خواهد بود. از آنجا که میزان دانش مهاجم از ۰.۷۵ کمتر است، در این حالت احتمال حمله مهاجم از رابطه (۱۵) در نظر گفته شده است.

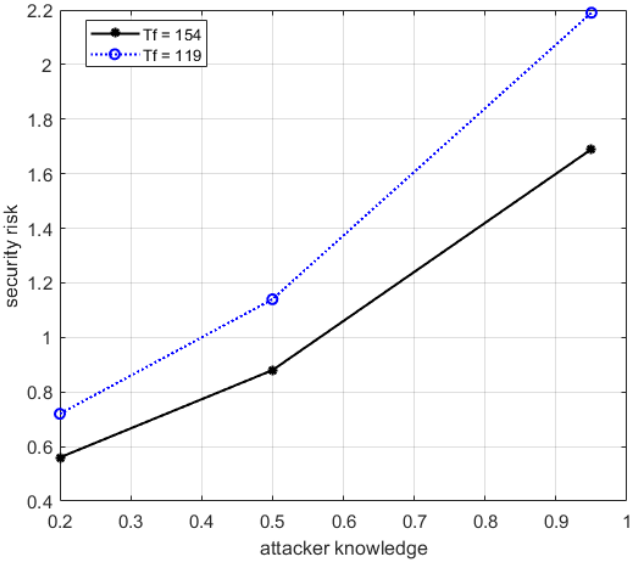
با در نظر گرفتن میزان دانش مهاجم برابر با K = ۰.۵، برای احتمال حمله مهاجم و احتمال کشف سیستم برای مدل بازی مهاجم نوع یک به‌ترتیب خواهیم داشت: P_{a0} = ۰.۱۲۵ و P_{d0} = ۰.۹. همچنین برای احتمال حمله مهاجم و احتمال کشف سیستم برای مدل بازی مهاجم نوع دو به‌ترتیب خواهیم داشت: P_{a1} = ۰.۲۵ و P_{d1} = ۰.۹۳. بدین ترتیب احتمال حمله و کشف سراسری برابر خواهد بود با: P_a = ۰.۱۲۵ و P_d = ۰.۹۲۴. در این صورت مخاطره حمله برای حمله اول، برابر ۰.۸۸ = (۱۵۴*۰.۹۲۴)/ (۱۰۰۰*۰.۱۲۵)= r خواهد بود و مخاطره حمله برای حمله دوم، برابر r = (۱۱۹*۰.۹۲۴)/ (۱۰۰۰*۰.۱۲۵) = ۱.۱۴ خواهد بود.

به طور مشابه، با در نظر گرفتن دانش مهاجم برابر با K = ۰.۹۵، برای احتمال حمله و کشف سراسری برابر خواهد بود با: P_a = ۰.۲۵ و P_d = ۰.۹۶۱. در این صورت مخاطره حمله برای حمله اول، برابر r =(۱۰۰۰*۰.۲۵)/ (۱۵۴*۰.۹۶۱) = ۱.۶۹ خواهد بود و مخاطره حمله برای حمله دوم، برابر r = (۱۰۰۰*۰.۲۵)/ (۱۱۹*۰.۹۶۱) = ۲.۱۹ خواهد بود. از آنجا که میزان دانش مهاجم از ۰.۷۵ بیشتر است، در این حالت احتمال حمله مهاجم از رابطه (۱۷) محاسبه شده است.

شکل ۶، احتمال حمله و کشف حمله برای مقادیر مختلف دانش مهاجم و شکل ۷، میزان مخاطره حملات انجام شده را برای مقادیر مختلف دانش مهاجم نشان می‌دهد.



شکل ۶. احتمال کشف و احتمال حمله بر حسب میزان دانش مهاجم.



شکل ۷. مخاطره امنیت برحسب میزان دانش مهاجم برای دو حمله انجام شده.

همان‌طور که مشاهده می‌شود با افزایش میزان دانش مهاجم، احتمال حمله افزایش و احتمال کشف نیز افزایش خواهد یافت. مخاطره حملات نیز با توجه به افزایش دانش مهاجم و در نتیجه افزایش احتمال حمله افزایش یافته است. از طرفی همانطور که مشاهده می‌شود، در همه مطالعات مخاطره حمله دوم که زمان تا خرابی کمتری دارد، بالاتر است.

نتیجه‌گیری

در این مقاله، روشی برای مدل‌سازی و ارزیابی مخاطره امنیتی سیستم‌های سایبر- فیزیکی با استفاده از نظریه بازی‌ها با اطلاعات ناقص ارائه شده است. روش ارائه‌شده، تقابل بین سیستم و مهاجم را به‌صورت یک بازی بیزی با اطلاعات ناقص مدل می‌کند. از آنجا که اطلاعات سیستم و مهاجم از یکدیگر اطلاعات کاملی نیست، این روش بسیار برای مدل‌سازی تقابل بین رفتار سیستم و مهاجم مناسب است. ورودی مدل، مؤلفه‌های بازی مانند هزینه حمله مهاجم، میزان خسارت واردشده به سیستم بر اثر حمله موفق، جریمه کشف نشدن حمله، سود و جایزه سیستم از کشف حمله و دانش مهاجم درمورد سیستم موردهدف هستند. خروجی مدل، رابطه‌هایی هستند که امکان محاسبه احتمال حمله و احتمال کشف حمله به صورت کمّی را فراهم می‌کنند. همچنین روشی برای ارزیابی مخاطره امنیتی با استفاده از احتمال حمله و کشف آن، میزان خسارت حمله بر سیستم و زمان تا بروز خسارت به سیستم ارائه شده است. کاربردپذیری روش ارائه شده با انجام یک مطالعه موردی بر روی یک سیستم صنعتی آزمایشگاهی نشان داده شده است. همان‌طور که بررسی‌ها نشان می‌دهد موفقیت حمله مهاجم به مؤلفه‌های مختلفی ازجمله میزان دانش او در مورد سیستم فیزیکی تحت کنترل و شرایط خرابی سیستم بستگی دارد.

در کارهای آینده مدل ارائه‌شده را به‌منظور ارزیابی امنیت سیستم‌های سایبر فیزیکی با مدل‌های مبتنی بر حالت به کار خواهیم گرفت. همچنین یکی دیگر از جنبه‌های موردانتظار واسط انسان- ماشین، در سیستم‌های سایبر-

فیزیکی است. جایی که مهاجم ممکن است اثر حملات را از کاربر سیستم مخفی کند تا اثرات مخربی را برجای گذارد. پرداختن به این نوع حملات، یکی از اولویت‌های بعدی در مطالعه، امنیت سیستم‌های سایبر- فیزیکی است.

References

[1] Hu, F., Lu, Y., Vasilakos, A. V., Hao, Q., Ma, R., Patil, Y., Zhang, T., Lu, J., Li, X., & Xiong, N. N. (2016). Robust Cyber–Physical Systems: Concept, models, and implementation. *Future Generation Computer Systems*, *56*(1), 449-475. https://doi.org/10.1016/j.future.2015.06.006

[2] Krotofil, M., & Larsen, J. (2014). Are You Threatening My Hazards? In *Advances in Information and Computer Security: 9th International Workshop on Security, IWSEC 2014, Hirosaki, Japan, August 27-29, 2014. Proceedings*. Springer, Cham. https://doi.org/10.1007/978-3-319-09843-2_2

[3] Kopetz, H., & Sytems, R-T. (2011). *Real-Time Systems: Design Principles for Distributed Embedded Applications* (2 ed.). Springer New York. https://doi.org/10.1007/978-1-4419-8237-7

[4] Li, H., Lai, L., & Poor, H. V. (2012). Multicast Routing for Decentralized Control of Cyber Physical Systems with an Application in Smart Grid. *IEEE Journal on Selected Areas in Communications*, *30*(6), 1097-1107. https://doi.org/10.1109/JSAC.2012.120708

[5] Jagtap ,S. S., VS, S. S., & Subramaniyaswamy, V. (2021). A hypergraph based Kohonen map for detecting intrusions over cyber–physical systems traffic. *Future Generation Computer Systems*, *119*, 84-109. https://doi.org/10.1016/j.future.2021.02.001

[6] Orojloo, H., & Abdollahi Azgomi, M. (2019). Modelling and evaluation of the security of cyber-physical systems using stochastic Petri nets. *The Institution of Engineering and Technology Cyber-Physical Systems: Theory & Applications*, *4*(1), 50-57. https//:doi.org/10.1049/iet-cps.2018.0008

[7] Orojloo, H., & Abdollahi Azgomi, M. (2017). A method for evaluating the consequence propagation of security attacks in cyber–physical systems. *Future Generation Computer Systems*, *67*, 57-71. https://doi.org/10.1016/j.future.2016.07.016

[8] Cui, Y., Quddus, N., & Mashuga, C. V. (2020). Bayesian network and game theory risk assessment model for third-party damage to oil and gas pipelines. *Process Safety and Environmental Protection*, *134*, 178-188. https://doi.org/10.1016/j.psep.2019.11.038

[9] Liang, X., & Xiao, Y. (2013). Game Theory for Network Security. *IEEE Communications Surveys & Tutorials*, *15*(1), 472-486. https://doi.org/10.1109/SURV.2012.062612.00056

[10] Abapour, S., Mohammadi-Ivatloo, B., & Tarafdar Hagh, M .(2020) .A Bayesian game theoretic based bidding strategy for demand response aggregators in electricity markets. *Sustainable Cities and Society*, *54*, 101787. https://doi.org/10.1016/j.scs.2019.101787

[11] Dahiya, A., & Gupta, B. B. (2021). A reputation score policy and Bayesian game theory based incentivized mechanism for DDoS attacks mitigation and cyber defense. *Future Generation Computer Systems*, *117*, 193-204. https://doi.org/10.1016/j.future.2020.11.027

[12] Ashok, A., Hahn, A., & Govindarasu, M. (2014) .Cyber-physical security of Wide-Area Monitoring, Protection and Control in a smart grid environment. *Journal of Advanced Research*, *5*(4), 481-489. https://doi.org/10.1016/j.jare.2013.12.005

[13] Ma, C. Y. T., Rao, N. S. V., & Yau, D. K. Y. (2011, April 10-15). *A game theoretic study of attack and defense in cyber-physical systems.* 2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Shanghai, China https://doi.org/10.1109/INFCOMW.2011.5928904

[14] Vigo, R., Bruni, A., & Yüksel ,E. (2013). Security Games for Cyber-Physical Systems. In *Secure IT Systems: 18th Nordic Conference, NordSec 2013, Ilulissat, Greenland, October 18-21, 2013, Proceedings*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-41488-6_2

[15] He, F., Zhuang, J., & Rao, N. S. (2012, May 19-23). *Game-theoretic analysis of attack and defense in cyber-physical network infrastructures.* IIE Annual Conference. Proceedings, Hilton Bonnet Creek, Orlando, Florida United States. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.719.2652&rep=rep1&type=pdf

[16] Krotofil, M., Cárdenas, A., Larsen, J., & Gollmann, D. (2014). Vulnerabilities of cyber-physical systems to stale data—Determining the optimal time to launch attacks. *International Journal of Critical Infrastructure Protection*, *7*(4), 213-232. https://doi.org/10.1016/j.ijcip.2014.10.003

[17] Yampolskiy, M., Horváth, P., Koutsoukos, X. D., Xue, Y., & Sztipanovits, J. (2015). A language for describing attacks on cyber-physical systems. *International Journal of Critical Infrastructure Protection*, *8*, 40-52. https://doi.org/10.1016/j.ijcip.2014.09.003

[18] Mitchell, R., & Chen, I. (2016). Modeling and Analysis of Attacks and Counter Defense Mechanisms for Cyber Physical Systems. *IEEE Transactions on Reliability*, *65*(1), 350-358. https://doi.org/10.1109/TR.2015.2406860

[19] Tantawy, A., Abdelwahed, S., Erradi, A., & Shaban, K. (2020). Model-based risk assessment for cyber physical systems security. *Computers & Security*, *96*, 101864. https://doi.org/10.1016/j.cose.2020.101864

[20] Alguliyev, R., Imamverdiyev, Y., & Sukhostat, L. (2018). Cyber-physical systems and their security issues. *Computers in Industry*, *100*(1), 212-223. https://doi.org/10.1016/j.compind.2018.04.017

[21] Orojloo, H., & Abdollahi Azgomi, M. (2016). Predicting the behavior of attackers and the consequences of attacks against cyber-physical systems. *Security and Communication Networks*, *9*(18), 6111-6136. https://doi.org/10.1002/sec.1761

[22] Orojloo, H., & Abdollahi Azgomi, M .(2018) .A Stochastic Game Model for Evaluating the Impacts of Security Attacks Against Cyber-Physical Systems. *Journal of Network and Systems Management*, *26*(4), 929-965. https://doi.org/10.1007/s10922-018-9449-0

[23] Orojloo, H., & Abdollahi Azgomi, M .(2017) .A game-theoretic approach to model and quantify the security of cyber-physical systems. *Computers in Industry*, *88*, 44-57. https://doi.org/10.1016/j.compind.2017.03.007

[24] Tripathi, D., Singh, L. K., Tripathi, A. K., & Chaturvedi, A. (2021). Model based security verification of Cyber-Physical System based on Petrinet: A case study of Nuclear power plant. *Annals of Nuclear Energy*, *159*, 108306. https://doi.org/10.1016/j.anucene.2021.108306

[25] Kholidy, H. A. (2021). Autonomous mitigation of cyber risks in the Cyber–Physical Systems. *Future Generation Computer Systems*, *115*(10), 171-187. https://doi.org/10.1016/j.future.2020.09.002

[26] Sánchez Rodríguez, M. Á., Bermejo Higuera, J., Bermejo Higuera, J. R., Sicilia Montalvo, J. A., & González Crespo, R .(2021) .A systematic approach to analysis for assessing the security level of cyber-physical systems in the electricity sector.

Microprocessors and Microsystems, *87*(1), 104352. https://doi.org/10.1016/j.micpro.2021.104352

[27] Ahmed Jamal, A., Mustafa Majid, A-A., Konev, A., Kosachenko, T., & Shelupanov, A. (2021). A review on security analysis of cyber physical systems using Machine learning. *Materials Today: Proceedings*, -(-), -. https://doi.org/10.1016/j.matpr.2021.06.320

[28] Dong, L., Xu, H., Wei ,X., & Hu, X. (2022). Security correction control of stochastic cyber–physical systems subject to false data injection attacks with heterogeneous effects. *International Society of Automation Transactions*, *123*, 1-13. https://doi.org/10.1016/j.isatra.2021.05.015

[29] Bernardi, S., Gentile, U., Marrone, S., Merseguer, J., & Nardone, R. (2020). Security modelling and formal verification of survivability properties: Application to cyber-physical systems. *Journal of Systems and Software*, *171*, 110746. https://doi.org/10.1016/j.jss.2020.110746

[30] Liu, X., Zhang, J., Zhu, P., Tan, Q., & Yin, W. (2021). Quantitative cyber-physical security analysis methodology for industrial control systems based on incomplete information Bayesian game. *Computers & Security*, *102* ,102138 .https://doi.org/10.1016/j.cose.2020.102138

[31] Kowalewski, S., Stursberg, O., Fritz, M., Graf, H., Hoffmann, I., Preußig, J., Remelhe, M., Simon, S., & Treseler, H. (1999). A Case Study in Tool-Aided Analysis of Discretely Controlled Continuous Systems: The Two Tanks Problem. In *Hybrid Systems V*. Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-49163-5_9