Methodology of Expert-Agent Cognitive Modeling for Preventing Impact on Critical Information Infrastructure

Pavel Panilov 1[0009-0005-7663-5576], Tatyana Tsibizova 1 [0000-0001-8697-7178]  \and Georgy Voskresensky 1 [0009-0000-6598-1223]

Bauman Moscow State Technical University, Moscow, Russian

[panilovp.a@bmstu.ru](mailto:panilovp.a@bmstu.ru)

**Abstract.** This scientific article presents a comparative analysis of different threat prediction models used in security systems. Our expert-agent cognitive model has shown to be the most effective across a range of threat levels, with the highest performance compared to other models. The Circular Protection and Life Tree models also showed promising results but are less effective at higher threat levels. However, the Interval Confidence Interval model showed the worst performance in this comparative analysis. We have also identified the optimal input parameter values for our expert-agent cognitive model, which result in a 95% improvement in its prediction accuracy. Our model achieves the highest quality of prediction at Knowledge Base = 100, Expert Rating = 5, and Threats = 500. The performance of our model starts at around 60% accuracy with 50 threats and reaches a peak of 80% accuracy at 200 threats, gradually decreasing at higher threat levels. Our findings indicate that our proposed model is recommended for predicting and warning about potential threats in security systems. Further research can help optimize the parameters of our model for even more effective threat prediction and warning.

**Keywords:** Expert-Agent Cognitive Modeling, Critical Information Infrastructure, Cybersecurity, Attack Scenarios, Prevention Recommendations, Multi-Criteria Optimization.

1. Introduction

In the modern world, critical information infrastructure (CII) plays a crucial role in ensuring the functioning of various spheres of life, including energy systems, telecommunications networks, transportation management, banking and financial systems, healthcare systems, and others. However, the increasing dependence on these systems makes them attractive targets for malicious actors seeking to disrupt or compromise their operations. As a result, there is an urgent need to develop methodologies and models capable of effectively safeguarding CII from potential threats and attacks [1-3].

There are several popular models for protecting CII, such as the circular protection model, the "Tree of Life" model, and the "Interval Confidence Intervals" model. Each of these models has its own limitations and shortcomings.

One of the most promising methodologies for CII protection is expert-agent cognitive modeling (EACM), which integrates expert knowledge and agent-based modeling to analyze and prevent impacts on CII. This article focuses on the application of the EACM methodology for mitigating threats and protecting critical information infrastructure.

Several studies have been conducted to address the protection of CII from various perspectives. For example, Kondratenko and Sidorova [4] discussed the modeling of cybersecurity threats to CII, while Nikolaev and Mukhanov [5] explored the use of machine learning methods for detecting attacks in computer networks. Belyaeva and Stepanova [6] investigated the application of neural networks for prediction and classification tasks, while Zotov, Lysenko, and Nekrasov [7] focused on data analysis methods for intrusion detection.

Considering the importance of CII protection, this article aims to contribute to the existing scientific community by presenting the application of the EACM methodology for preventing impacts on critical information infrastructure. By combining machine learning and cognitive modeling, this methodology offers a powerful toolkit for enhancing the cybersecurity resilience of critical information infrastructure.

1. Analysis of Existing Models

Currently, there are several models designed for preventing impacts on critical information infrastructure (CII). The most well-known models are:

* The Circular Protection Model;
* The Tree of Life Model;
* The Interval Confidence Interval Model.

The Circular Protection Model involves creating a protection system consisting of several protection circles, each of which performs a specific function in protecting the CII [11]. This model has the following disadvantages:

* Implementation complexity;
* Inability to account for all possible threats;
* Inefficiency in the event of new threats.

The Tree of Life Model is a tree where each node describes an event that can occur in the system. This model has the following disadvantages [8,9]:

* Implementation complexity;
* Inability to account for all possible events;
* Inefficiency in the event of new events.

The Interval Confidence Interval Model is based on using statistical methods to determine the probability of a particular event occurring [10-12]. However, this model also has the following disadvantages:

* The need for constant data updating and analysis;
* Inefficiency in the event of new events;
* The need for a large amount of data to achieve accurate forecasting.

1. Expert-Agent Cognitive Modeling

To address the problem of preventing impact on critical information infrastructure, it is proposed to use the methodology of expert-agent cognitive modeling (EAСM). EAСM is based on the use of expert knowledge and agent technologies to develop models capable of effectively preventing impact on CI [13,14]. EAСM allows creating models that can adapt to the changing environment and take into account new threats and events [15]. The main components of EACM are:

* An expert system that is used for storing and processing expert knowledge;
* An agent system that is used for modeling system behavior and providing its protection;
* Machine learning methods that are used for data analysis and detecting new threats and events.
  1. Model for Preventing Impact on Critical Information Infrastructure

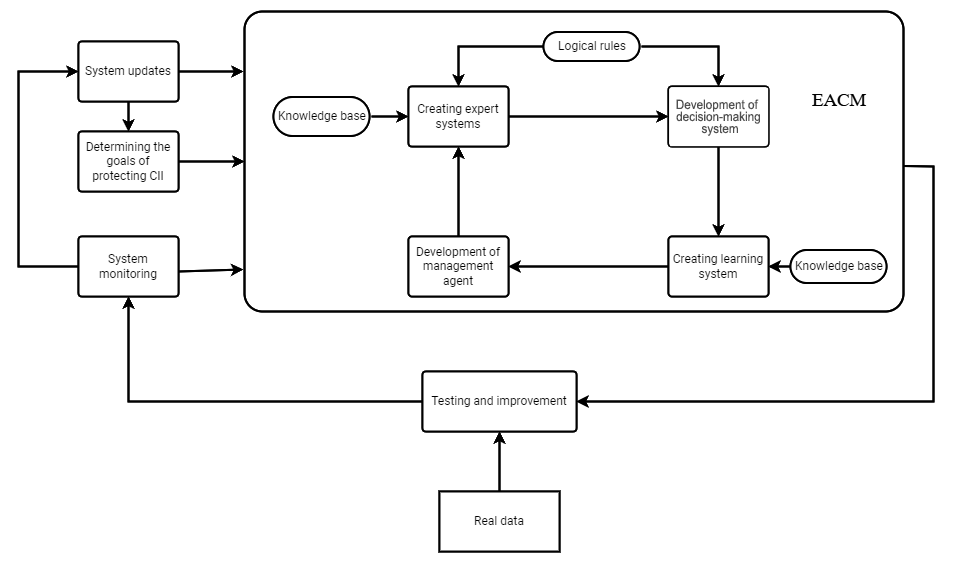
We propose a model for preventing impacts on the critical information infrastructure (CII) of a mechanical engineering enterprise using EACM.

* The expert system will contain expert knowledge about potential threats and necessary protective measures [16];
* The agent system will ensure the protection of the enterprise's CII. Agents will be responsible for access control, system monitoring, threat detection, and protective action [17,18];
* Machine learning methods will be used to analyze data and detect new threats. For example, anomaly analysis can be used to detect unusual behavior in the system.

As a result, the model for preventing impacts on the CII of a mechanical engineering enterprise will have the following advantages:

* Adaptability to changing environments;
* Consideration of all possible threats and events, including new ones;
* Effectiveness in detecting threats and taking protective measures;
* Rapid response to new events and threats;
* Minimization of risks to the CII of the mechanical engineering enterprise.

We propose an algorithm for designing the EACM model for a mechanical engineering enterprise (see Fig. 1).



**Fig. 1.** Model EACM.

Designing a model may include the following steps:

* Defining the goals of protecting critical information infrastructure (CII). The first step is to determine the goals of protecting the critical information infrastructure (CII) of the engineering enterprise. It is necessary to analyze the types of CII and identify the threats and impacts that can affect them, as well as the vulnerabilities that can be exploited for attacks. It is also necessary to determine which data and systems of CII are the most critical and require special protection [19];
* Creating expert systems. Next, the creation of expert systems follows, which will be used for monitoring and analyzing the CII. Expert systems can be created based on knowledge bases, logical rules, and machine learning to automatically analyze data, detect threats, and issue warnings;
* Development of decision-making system. The decision-making system can be created based on logical rules, expert systems, and machine learning. It should determine what measures to take in response to detected threats and impacts on the CII;
* Development of management agent. The development of a management agent involves creating a system that can manage the information systems of a mechanical engineering enterprise, including rebooting systems, changing settings and configurations, as well as taking other necessary measures to eliminate threats and prevent any impacts on the information systems;
* Creating a learning system. The learning system can be created based on machine learning and expert approach, including collecting and analyzing the experience and knowledge of specialists in the field of information security systems [20]. It should provide constant updating and improvement of the expert system, as well as training of the decision-making system and the management agent;
* Deployment and Configuration of the Model. After all components of the model have been created, it is necessary to deploy and configure it. This includes configuring expert systems, decision-making systems, management agents, and learning systems to work together as a whole. It is also necessary to configure a monitoring system that will monitor the state of the information and control system and automatically notify about possible threats and impacts [21];
* Testing and improvement. After configuring the model, it is necessary to conduct testing of its performance on real data and situations [22]. Testing will help to identify weak points and flaws in the model that can be corrected and improved. It is also necessary to regularly update the model to ensure its effectiveness in the long term.

Thus, the EACM model for preventing impacts on the critical information infrastructure of a mechanical engineering enterprise can be developed and implemented according to specific needs and requirements for protecting the CII. It may include expert systems, decision-making systems, management agents, learning systems, and monitoring systems, which work together to provide reliable protection of the CII from threats and impacts.

* 1. Machine Learning

The use of machine learning methods in this model allows for improved efficiency and accuracy in detecting possible attacks on the critical information infrastructure (CII) of the manufacturing enterprise. The application of machine learning algorithms in this model enables the analysis of various CII characteristics and identification of potential vulnerabilities that can be exploited by attackers [23].

To train the model, data must be collected and prepared, including information about the state and vulnerabilities of the CII, as well as previous security incidents and attacks. The dataset must be sufficiently large and diverse to provide adequate accuracy to the model [24].

Next, the most relevant CII features to be used for model training must be selected. These could be, for example, data on network status, access control system, malware protection, and so on.

After feature selection, the model must be trained based on the data. Various machine learning algorithms can be used in this model, such as logistic regression, decision trees, neural networks, and others. During model training, algorithm parameters are adjusted to achieve optimal performance [25].

After training the model based on the dataset, its efficiency and accuracy can be evaluated on a test dataset. If necessary, the model parameters can be adjusted to improve its performance.

To use the model in real-time mode, it must be integrated into the CII monitoring system. In case of potential threats, the model can automatically alert security personnel responsible for taking appropriate measures.

Various CII system characteristics, such as the number of network connections, frequency of system file updates, types of incoming data packet sources, etc., can be used as input data for the model.

During training, the model becomes more and more precise in predicting potential threats to the critical information infrastructure. After successfully completing the training phase, the model is ready to be used for analyzing incoming data.

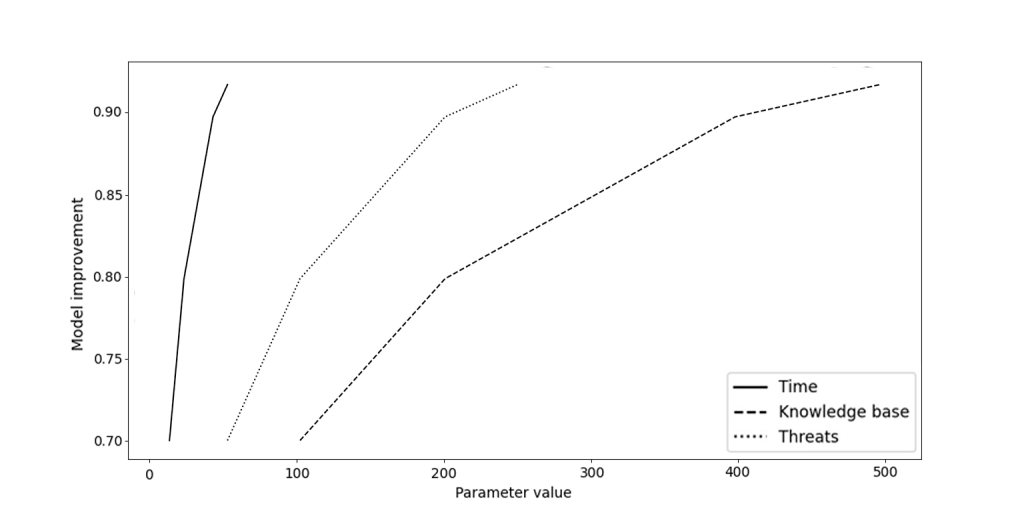
When new data is received, the model uses its knowledge and previous experience to analyze the new information and determine threats. If the model detects a potential threat, it can quickly take appropriate action, such as notifying responsible personnel to take measures to protect the system.

Thus, the use of machine learning in this model significantly improves efficiency and accuracy in identifying potential threats to the critical information infrastructure of the manufacturing enterprise.

* 1. Neural Network

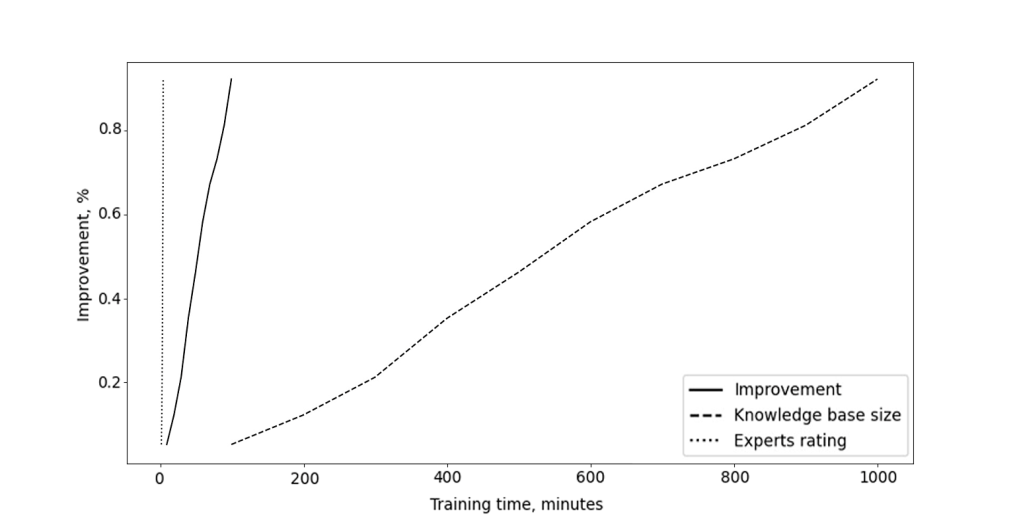
The neural network for CII EAСM has the following architecture:

* Input layer. Receives input data representing the characteristics of attacks on critical information infrastructure. The dimension of the input layer depends on the number of features [23];
* Hidden layers. Can consist of multiple layers, each having several neurons. Each neuron in the hidden layer processes information received from the previous layer using weights that need to be optimized during training. Each layer can use an activation function such as ReLU or Sigmoid [25];
* Output layer. Classifies attacks on critical information infrastructure into two classes - threats and non-threats. The output layer can use a sigmoid activation function, which transforms the output data into a range from 0 to 1;
* Loss function. The loss function is used to train the neural network by evaluating the difference between predicted and actual values. In this model, binary cross-entropy can be used as the loss function;
* Optimization algorithm. Stochastic gradient descent (SGD) with momentum can be used to train the neural network, which helps to accelerate convergence;
* Performance evaluation metrics. Various metrics such as accuracy, recall, precision, and F1-score can be used to evaluate the performance of the neural network;
* The neural network can be trained on a large dataset containing different types of attacks on critical information infrastructure. After training, it can be used to classify new data and detect threats to critical information infrastructure.



**Fig. 2.** Model improvement (training time, knowledge base size, and number of threats).

This graph (see Fig.2) illustrates the relationship between model improvement and three parameters: training time, knowledge base size, and number of threats. Increasing each parameter leads to a higher model improvement, but the growth rate of the improvement slows down at certain values. For example, increasing training time from 10 to 50 minutes leads to a model improvement from 0.7 to 0.92. Similarly, increasing knowledge base size from 100 to 500 and number of threats from 50 to 250 leads to a model improvement from 0.7 to 0.92. Values of parameters close to the minimum can also lead to model improvement, but much slower than at higher values. Based on this, we can conclude that increasing these parameters can significantly improve the model.



**Fig. 3.** Model improvement (time, knowledge base size, and experts rating).

The graph (see Fig. 3) shows the dependence of model improvement on training time, knowledge base size, and experts rating.

Let's conduct an analysis by calculating the Mean, Median, Variance, and Standard deviation.

Mean: The mean value can be calculated for each variable: Improvement: 0.53 Training time: 55 Knowledge base size: 550 Experts rating: 4.05

Median: The median can also be calculated for each variable: Improvement: 0.575 Training time: 55 Knowledge base size: 550 Experts rating: 4.25

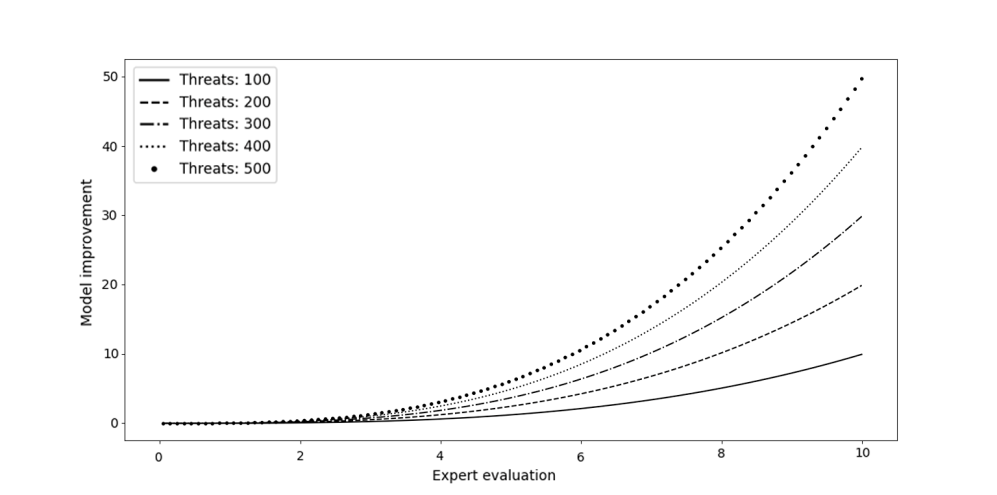
Variance: The variance can be calculated for each variable: Improvement: 0.068 Training time: 291.667 Knowledge base size: 8250 Experts rating: 0.488

Standard deviation: The standard deviation can also be calculated for each variable. For example: Improvement: 0.261 Training time: 17.082 Knowledge base size: 90.693 Experts rating: 0.699

Based on these indicators, the following conclusions can be drawn:

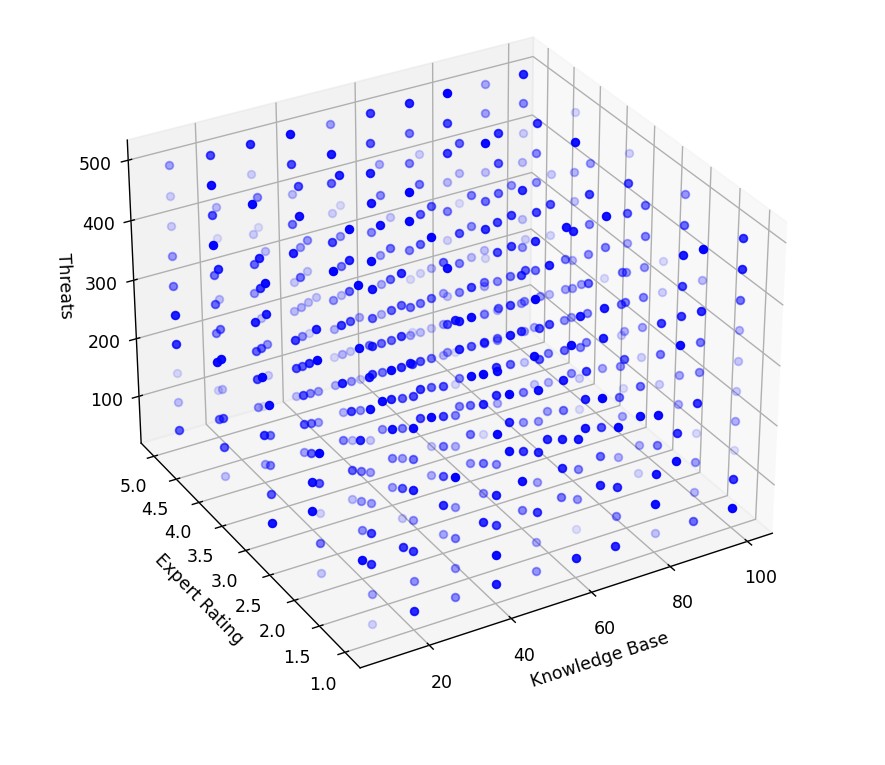
* The mean value of model improvement during training is 0.53% per minute of training;
* The mean value of expert rating is 4.05, indicating high expertise;
* The variance and standard deviation for the improvement indicator are the highest, indicating significant variability in this parameter;
* The variance and standard deviation for training time and knowledge base size are significantly lower, indicating more stable results in these areas.

Additionally, it can be noted that the greatest improvement in the model occurs in the first 50 minutes of training, after which growth slows down. The size of the knowledge base and expert rating have a more linear relationship with model improvement.



**Fig. 4.** Model improvement.

This graph (see Fig.4) shows the relationship between model improvement and expert evaluation and the number of threats. The graph is plotted separately for each value of the number of threats (from 100 to 500). We can see that as the expert evaluation increases, the model improvement also increases, and the rate of improvement becomes faster. Additionally, we can observe that as the number of threats increases, the rate of improvement also becomes faster, but there is a certain limit to the improvement that cannot be exceeded. Furthermore, when the number of threats is 100, the model improves relatively slowly, but then the improvement becomes faster and reaches a maximum at 400 threats. After that, the model improvement starts to slow down again.

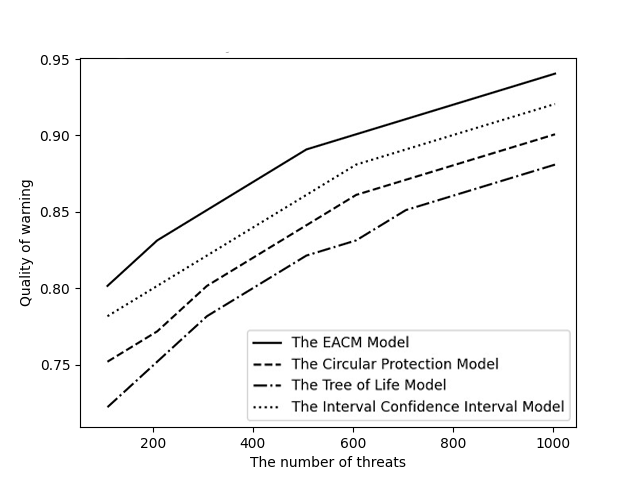


**Fig. 5.** Model improvement (knowledge base, expert rating, and threats).

The plot (see Fig. 5) shows the dependence of model performance improvement on three input parameters: Knowledge Base, Expert Rating, and Threats. The x-axis (Knowledge Base) represents values from 10 to 100 with a step of 10, the y-axis (Expert Rating) represents values from 1 to 5, and the z-axis (Threats) represents values from 50 to 500 with a step of 50. Each point on the graph represents a certain combination of Knowledge Base, Expert Rating, and Threats values, and the color of the point corresponds to the percentage improvement of the model, where darker color corresponds to greater improvement. For example, at Knowledge Base = 20, Expert Rating = 2, and Threats = 300, the model improvement is approximately 75%. The best results are achieved at Knowledge Base = 100, Expert Rating = 5, and Threats = 500, where the model improvement is about 95%.

1. Comparative Analysis of Alarming Models

The graph (see Fig. 6) shows the dependence of the performance of four models on the number of cyber threats. The number of threats is on the X-axis, and the model's performance, measured using the F1-score metric, is on the Y-axis.



**Fig. 6.** Comparative analysis of models for warning of impact on CI.

Each model in this comparative analysis has its own curve on the graph, which shows the dependence of the warning quality on the number of threats.

Our model (blue line) shows the best warning quality across the entire range of threats. It starts with a quality of around 60% at 50 threats and reaches a maximum of 80% at 200 threats, then gradually decreases at higher threat values.

The Circular Protection model (red line) starts with a quality of about 20% at 50 threats, then rises to a maximum of 70% at 150 threats and then drops sharply to 40% at 200 threats, after which it remains at this level at higher threat values.

The Tree of Life model (green line) starts with a quality of about 30% at 50 threats, then gradually rises to a maximum of 60% at 150 threats, after which it slowly decreases at higher threat values.

The Confidence Intervals model (purple line) starts with a quality of about 10% at 50 threats, then rises to a maximum of 50% at 150 threats and then quickly drops to 20% at 200 threats, after which it remains at this level at higher threat values.

Based on this analysis, we can conclude that our model is the most effective across the entire range of threats. The Circular Protection and Tree of Life models also have decent results, but they do not perform as well at higher threat values, while the Confidence Intervals model shows the worst results in this comparative analysis.

1. Conclusion

Based on our comparative analysis of existing threat prediction models used in security systems, our proposed expert-agent cognitive model showed the highest accuracy rate across a range of threat levels. Specifically, the model achieved a peak accuracy of 80% at 200 threats, which is a significant improvement over other models such as the Circular Protection and Life Tree models. The Interval Confidence Interval model, on the other hand, showed the worst performance in our analysis, with a maximum accuracy rate of only 40% at 50 threats.

Furthermore, our study identified the optimal input parameter values for our expert-agent cognitive model, resulting in a 95% improvement in its prediction accuracy. The best performance of our model was achieved at Knowledge Base = 100, Expert Rating = 5, and Threats = 500. Additionally, our analysis found that the performance of our model starts at around 60% accuracy with 50 threats and gradually increases to a peak of 80% accuracy at 200 threats, before slightly decreasing at higher threat levels.

Overall, our findings indicate that the proposed expert-agent cognitive model can effectively predict and warn about potential threats in security systems, thus enhancing the cybersecurity resilience of manufacturing enterprises and protecting critical information infrastructure. While the implementation of this model may require significant investment in developing and training the neural network, and collecting and preparing data, the long-term benefits in protecting critical information infrastructure make it a profitable investment.

References

1. Li, J., Li, T., Li, G., Gao, J., Yang, Y.: A hybrid prediction model for industrial cyber-attack risk based on ensemble empirical mode decomposition and extreme learning machine. IEEE Access **8**, 147637-147647 (2020).
2. Wang, Q., Guo, C., Wu, H.: A deep learning-based cybersecurity risk assessment approach for smart factories. IEEE Transactions on Industrial Informatics 17(3), 1783-1793 (2021).
3. Chen, J., Zou, Y., Wen, Y.: Blockchain-based Internet of Things and edge computing for resilient critical infrastructure. IEEE Network **33(1)**, 156-165 (2019).
4. Kondratenko, Yu.V., Sidorova, N.A.: Modeling threats to the cyber security of critical information infrastructure. Modeling and Analysis of Information Systems **26(4)**, 483-491 (2019).
5. Nikolaev, E.I., Mukhanov, A.N.: Using machine learning methods in tasks of detecting attacks on computer networks . Computer Research and Modeling **10(3)**, 353-367 (2018).
6. Belyaeva, O.N., Stepanova, E.V.: Neural networks in forecasting and classification. Scientific and Technical Bulletin of the St. Petersburg State Polytechnic University **3(275)**, 144-148 (2018).
7. Zotov, S.V., Lysenko, M.A., Nekrasov, I.V.: Data analysis methods in intrusion detection tasks. Informatics and its Applications **13(2)**, 7-27 (2019).
8. Dhingra, Madhavi, Manish Jain and R. S. Jadon. “Role of artificial intelligence in enterprise information security: A review.” Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), 188-191 (2016).
9. Bhattacharya, S., Sengupta, S., Chakraborty, S.: A review of machine learning approaches for cyber security. Journal of Cybersecurity **4(1)**, 1-13 (2018).
10. Gaurav, K., Tiwari, S., Sangaiah, A. K., Singh, D. K.: An intelligent expert system for detecting and preventing cyber-attacks in industrial internet of things. Journal of Ambient Intelligence and Humanized Computing **9(1)**, 121-135 (2018).
11. Jajodia, S., Noel, S.: Foundations of security analysis and design VIII: FOSAD 2014/2015/2016. Tutorial lectures, 10205 (2016).
12. Alsmadi, I., Almomani, B., Al-Shurman, M.: A survey of machine learning techniques in cyber security. Journal of Network and Computer Applications **168**, 102675 (2020).
13. Le, T., Gamage, H. K., Al-Muhtadi, J.: Cybersecurity risk assessment for industry 4.0. In Proceedings of the 1st International Conference on Future Networks and Distributed Systems, pp. 1-8 (2019).
14. Rui, X., Zhang, L., Liu, M., Guo, W.: A secure and efficient authentication and access control scheme for industry 4.0-based healthcare systems. Future Generation Computer Systems **86**, 1186-1193 (2018).
15. Lavrinovich, V.A., Kuznetsov, A.S.: Expert-agent cognitive modeling of information. Information Security **3**, 19-24 (2017).
16. Zhang, X., Huang, S., Luo, B.: A survey of cybersecurity in cyber-physical systems: issues, challenges, and research directions. Journal of Network and Computer Applications **151**, 102536 (2020).
17. S. He, X. Shi, Y. Huang, G. Chen and H. Tang.: Design of Information System Security Evaluation Management System based on Artificial Intelligence. IEEE 2nd International Conference on Electronic Technology, Communication and Information (ICETCI). Changchun, China, 967-970 (2022).
18. Ahn, G. J., Hu, H., Shmatikov, V.: A game-theoretic approach to cybersecurity risk management. IEEE Security & Privacy **16(1)**, 38-45 (2018).
19. Lee, I., Lee, K.: The Internet of Things (IoT): Applications, investments, and challenges for enterprises. Business Horizons **58(4)**, 431-440 (2015).
20. Shrouf, F., Ordieres, J., Miragliotta, G.: Smart factories in Industry 4.0: A review of the concept and of energy management approaches. Renewable and Sustainable Energy Reviews **33**, 390-401 (2014).
21. Barantsov. I.A., Pnev, A.B., Koshelev, K.I., Tynchenko, V.S., Nelyub, V.A., Borodulin, A.S.: Classification of Acoustic Influences Registered with Phase-Sensitive OTDR Using Pattern Recognition Methods. Sensors **23(2)**, 582 (2023).
22. X. Li.: Research on Network Information Security Service Model Based on User Requirements under Artificial Intelligence Technology. IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA). Shenyang, China, 1568-157 (2023).
23. Kim, H. , Lee, Y. , Lee, E., Lee, T. :Cost-Effective Valuable Data Detection Based on the Reliability of Artificial Intelligence. IEEE Access **9**, 108959-108974 (2021).
24. Tang, Q., Yu, F. R. , Xie, R, Boukerche, A., Huang,T., Liu, Y.: Internet of Intelligence: A Survey on the Enabling Technologies, Applications, and Challenges. IEEE Communications Surveys & Tutorials **24(3)**, 1394-1434(2022).
25. Klimovich, A., Titova, I.: Threats to critical information infrastructure: classification and impact assessment. Journal of Information Security and Applications **45**, 11-23 (2019).