Panamerican Mathematical Journal
ISSN: 1064-9735
Vol 33 No. 2 (2023)

Probabilistic Risk Assessment in Cybersecurity: Bayesian Methods for Quantifying and Mitigating Cyber Risks

Fulsundar Amita Purushottam[1], Ajay Kumar[2], Vikas Haribhau Satonkar[3], Shweta Kundlik Gaikwad[4], Stefi Diliprao Sonawane[5], Bhushan shirwadkar[6]

[1] Assistant Professor, Department of Computer Engineering, Sandip University Nashik, Nashik, Maharashtra, India. fulsundar.purushottam@sandipuniversity.edu.in

[2] Assistant Professor, Assistant Professor, Department of Computer Engineering, Sandip University Sijoul, Madhubani, Bihar, India. ajay.kumarcse@sandipuniversity.edu.in

[3] Assistant Professor, Department of Computer Engineering, Sandip Institute of Engineering & Management, Nashik, Maharashtra, India. vikas.satonkar@siem.org.in

[4] Assistant Professor, Department of Computer Engineering, Sandip Institute of Technology & Reserch Centre, Nashik, Maharashtra, India. shweta.bhalerao@sitrc.org

[5] Assistant Professor, Department of Psychology, Sandip University Nashik, Nashik, Maharashtra, India. stefi.sonawane@sandipuniversity.edu.in

[6] Assistant Professor, Assistant Professor, Department of Mathematics, Sandip University Nashik, Maharashtra, India. bhushan.shirwadkar@sandipuniversity.edu.in

Article History:

Received: *05-03-2023*

Revised: *14-05-2023*

Accepted: *17-06-2023*

Abstract:

In today's digital world, where everything is linked, privacy dangers are a big problem for businesses, states, and people. When it comes to cybersecurity, traditional ways of assessing risk don't always work because online threats are always changing. By combining statistical models with existing data and expert knowledge, Bayesian methods, on the other hand, look like a potential way to measure and reduce online risks. This essay looks at how Bayesian methods can be used in probabilistic risk assessment (PRA) in the field of hacking. Bayesian reasoning is used by PRA to get a fuller picture of cyber risks by taking into account doubt, variation, and personal opinions during the risk assessment process. By using likelihood functions, posterior distributions, and prior probabilities, Bayesian methods give us a sensible way to update risk predictions as new information comes in. One of the best things about Bayesian PRA is that it can take into account how different online risks and weaknesses are connected and depend on each other. By using statistical modeling to look at these connections, businesses can find possible chain reactions and decide how to best stop them. Additionally, Bayesian methods let you use expert opinions and personal data, which makes it possible to look at online risks in a more complete way than just using numbers. One of Bayesian PRA's strengths is that it can work with little data and unknown input values. When there isn't enough or a lot of good actual data, Bayesian methods let analysts combine information from different sources, like past data, threat intelligence, and expert opinions, in a planned way to get more accurate risk predictions. Bayesian methods can be used to make decisions about how to reduce security risks as well as how to measure those risks. In a probabilistic framework, organizations can figure out the best ways to lower their general cyber risk exposure by modeling the possible

results of different mitigating measures. Overall, this study shows how Bayesian methods could help the area of computer risk assessment grow. By accepting doubt and using a variety of information sources, Bayesian PRA provides a powerful set of tools for measuring cyber risks and helping risk managers make decisions in a danger situation that is becoming more complicated.

Keywords: Probabilistic Risk Assessment, Cybersecurity, Bayesian Methods, Cyber Risks, Risk Quantification, Risk Mitigation, Uncertainty, Bayesian Inference, Threat Modeling, Risk Management

1. Introduction

In a time when everyone is connected to the internet, cyber dangers have become a major issue for businesses, states, and people all over the world. Cyber attacks are becoming more common and more complex, which makes it even more important to have strong risk assessment methods that can effectively measure and reduce cyber threats. Traditional ways of figuring out the risk of hacking often use fixed models and static studies, which might not be able to show how cyber dangers change and grow over time. Because of these problems, there is more and more interest in using statistical methods, especially Bayesian reasoning, to make cyber risk assessment more accurate and reliable [13]. A big change in how computer risk is managed is the use of Bayesian methods in probabilistic risk assessment (PRA). Bayesian PRA accepts doubt and variation as normal parts of the risk assessment process, unlike linear methods that depend on exact inputs and assumptions. The main idea behind Bayesian reasoning is that it gives us a way to update our views and make choices when we don't have all the facts [14]. Through Bayes' theorem, which combines what we already know with what we've seen, Bayesian methods let analysts come up with probabilistic predictions of computer risks that take into account both facts and personal opinions. Using Bayesian methods to evaluate hacking risks is a good idea because they have a number of important benefits over older methods. The most important thing is being able to recognize and measure doubt. It is hard to know what the goals and skills of danger players are, how well protective measures work, and everything else when it comes to cybersecurity [15]. Bayesian PRA takes this doubt into account by showing important factors, like how likely it is that a cyberattack will happen or how bad it will be if it does, as probability distributions instead of fixed numbers. This probabilistic model gives us a fuller picture of cyber risks and helps people make decisions by letting them figure out how likely different threat scenarios are to happen and what effects they might have [16].

Bayesian methods make it easier to use information and knowledge from a variety of sources in the risk assessment process. It can be hard to get accurate risk predictions in cybersecurity because important data is often missing, incomplete, or biased. This makes it hard to use standard statistical methods. To put together different types of information, like past data, danger intelligence, expert opinions, and qualitative ratings, Bayesian reasoning gives you a way to do it in a structured way. Bayesian PRA lets analysts use all the information they have to make better choices about cyber risks by putting these sources together in a way that is based on probability. One more benefit of Bayesian PRA is that it can describe the complicated connections and relationships between the different parts of the cyber environment. Cyber dangers are often linked and dependent on each other, so taking advantage of one weakness can have effects that spread

through systems and networks [18]. Traditional ways of assessing risk might not be able to keep up with these changes, which could lead to wrong estimates or the wrong use of resources to reduce risk. Bayesian methods are great for modeling these kinds of complicated systems because they take into account the direct and indirect links and relationships between risk factors [17]. This lets companies find possible security holes and arrange their efforts to fix them in a way that minimizes the total hacking risk. When it comes to digital risk management, Bayesian PRA makes case analysis and decision support easier. Organizations can figure out which risk mitigation measures work best and least expensively by modeling the results of different danger scenarios and mitigation strategies in a statistical framework. This helps people in charge make better decisions about how to spend money and focuses investments on protection measures that lower total risk the most. Using Bayesian methods in cybersecurity risk assessment is a hopeful way to deal with the problems caused by online threats that are getting more complicated and changing all the time. Bayesian PRA is a complete set of tools for measuring and reducing cyber risks in today's linked world. It does this by accepting error, using different types of information, and simulating complicated connections between risk factors [19]. This essay looks at how Bayesian methods can be used to evaluate cybersecurity risks and shows how they can make businesses safer and more resilient against cyber dangers.

2. Related Work

In the past few years, there have been big steps forward in the area of probabilistic risk assessment (PRA) in cybersecurity. Researchers are looking into a lot of different ways to better understand, measure, and reduce cyber threats. This part gives an outline of the most important efforts in this area, focusing on the studies' scope, results, and methods. One of the most important works in this field is about how to use Bayesian methods to evaluate hacking risks [1]. This research shows how important statistical modeling is for understanding the unknown parts of cyber dangers and weaknesses. The writers use Bayesian reasoning to show how prior knowledge and actual data can be used together to make more accurate risk predictions that take into account both facts and opinions. The results show that Bayesian methods are useful for making online risk estimates more accurate and reliable. Subsequent studies have built on this base to compare how well Bayesian techniques work versus standard risk assessment methods in cybersecurity [2]. Researchers have used real-world examples to show that Bayesian methods are better than linear models at catching the uncertainty and variability in the cyber risk scene. This new information has big implications for people who want to make their risk assessment methods more reliable.

Another area of research in computer risk assessment is how to add expert opinions to Bayesian PRA [3]. Experts have useful subject knowledge and ideas that can be used with numbers to help figure out how dangerous online threats are. By using expert views along with a Bayesian framework, researchers have shown that risk predictions are more accurate and reliable, especially when there isn't a lot of good observational data available. A lot of research has also been done on how well Bayesian PRA can predict complex cyber danger scenarios [4]. Cyber dangers are often linked and dependent on each other, so taking advantage of one weakness can have effects that spread through systems and networks. Bayesian PRA is an organized way to understand these changes and figure out how different threat situations might affect things. Researchers have shown that Bayesian

methods can help find key weaknesses and set priorities for reducing their impact by using scenario analysis. Another area of study that has gotten a lot of attention is figuring out what kind of data Bayesian PRA needs for defense [5]. A lot of the time, real-world data on online dangers and weaknesses may not be full or may be biased. Bayesian methods are a flexible way to combine different types of data, like past data, danger intelligence, and expert opinions, to make accurate risk predictions. Researchers have shown that Bayesian PRA can handle doubt and variability in data well by combining probabilistically different types of data. Comparing different Bayesian PRA methods using simulations has shown how well they work in different cyber risk situations [6]. Bayesian methods work better or worse based on how complicated the danger situation is and how much data is available. Researchers have found the pros and cons of different Bayesian methods through simulation studies. This information helps practitioners choose which techniques to use in different situations.

Adding Bayesian PRA to models for managing hacking risks is an important area for study [7]. Making choices under doubt is part of risk management, and Bayesian PRA is a logical way to measure and rank computer risks. By using probable risk projections in the decision-making process, businesses can better use their resources and put in place focused risk reduction strategies that lower their total risk exposure. A very important area of study right now is making Bayesian PRA systems that can handle new computer threats [12]. Traditional ways of assessing risk may not be able to keep up with how online threats change and get smarter. You can keep risk calculations up to date with new information and data by using Bayesian methods, which are open and adaptable [23]. Researchers want to make companies safer and more resilient against new cyber risks by creating Bayesian PRA systems that can adapt to new threats.
To sum up, the area of probabilistic risk assessment in cybersecurity has come a long way thanks to improvements in Bayesian methods and how they are used [20]. Researchers are always looking for new ways to make cyber risk assessment methods more useful and accurate. For example, they are creating complicated cyber threat scenarios and adding expert opinions, as well as dealing with data doubt. These efforts are necessary to make organizations more resilient and lessen the danger of cyber attacks in a world that is becoming more and more linked.

Table 1: Related Work

Sr. No	Scope	Findings	Methods
[1]	Application of Bayesian methods in cybersecurity risk assessment.	Bayesian methods offer a more comprehensive understanding of cyber risks.	Bayesian inference, Prior knowledge
[2]	Comparison of Bayesian and traditional risk assessment approaches in cybersecurity.	Bayesian methods outperform traditional approaches in capturing uncertainty.	Bayesian inference, Deterministic models
[3]	Incorporation of expert judgments in Bayesian PRA for cyber risk assessment.	Expert judgments enhance the accuracy of Bayesian risk estimates.	Bayesian inference, Expert elicitation
[4]	Evaluation of Bayesian PRA for modeling complex cyber threat scenarios.	Bayesian PRA effectively models interdependencies among cyber threats.	Bayesian inference, Scenario analysis
[5]	Assessment of data requirements for	Bayesian PRA can handle limited data by	Bayesian inference, Data

	Bayesian PRA in cybersecurity.	integrating diverse information sources.	synthesis
[6]	Simulation-based comparison of different Bayesian PRA techniques in cyber risk assessment.	Bayesian techniques exhibit varying performance in different cyber risk scenarios.	Bayesian inference, Simulation
[7]	Integration of Bayesian PRA into cybersecurity risk management frameworks.	Bayesian PRA enhances decision-making by providing probabilistic risk estimates.	Bayesian inference, Risk management frameworks
[8]	Bayesian network models for assessing cyber risks in critical infrastructure.	Bayesian networks offer a structured approach to modeling dependencies in cyber systems.	Bayesian networks, Critical infrastructure
[9]	Application of Bayesian PRA in assessing insider threats in cybersecurity.	Bayesian PRA helps in quantifying the likelihood and impact of insider cyber attacks.	Bayesian inference, Insider threat modeling
[10]	Bayesian methods for prioritizing cyber risk mitigation strategies.	Bayesian analysis aids in identifying the most effective mitigation measures.	Bayesian inference, Decision support
[11]	Incorporation of uncertainty quantification techniques into Bayesian PRA.	Uncertainty quantification enhances the robustness of Bayesian cyber risk assessments.	Bayesian inference, Uncertainty quantification
[12]	Development of Bayesian PRA frameworks for emerging cyber threats.	Bayesian PRA adapts to evolving cyber threats by incorporating new data and knowledge.	Bayesian inference, Emerging threats
[13]	Evaluation of the scalability of Bayesian PRA for large-scale cyber systems.	Bayesian methods demonstrate scalability in assessing cyber risks in complex environments.	Bayesian inference, Scalability
[14]	Application of Bayesian PRA in compliance-driven cybersecurity risk assessment.	Bayesian methods aid in meeting regulatory requirements by providing quantifiable risk metrics.	Bayesian inference, Compliance assessment
[15]	Bayesian methods for assessing the economic impact of cyber risks.	Bayesian analysis enables organizations to evaluate the financial implications of cyber threats.	Bayesian inference, Economic modeling

3. Research Methodology

3.1 Data Collection and Information Gathering:

In cybersecurity, the first step in the probability risk assessment (PRA) method is to get useful data from a number of different sources. This data will be used to guide the research. This includes data on past cyber incidents, which shows how security was broken, how attacks happened, and what happened as a result. Threat intelligence reports from trustworthy sources like cybersecurity companies, government agencies, and industry groups are a great way to learn about new cyber risks, trends, and attack methods. Expert views are also very important for adding emotional thoughts and judgments to numeric data. Cybersecurity workers, threat researchers, system managers, and other subject experts with hands-on experience and deep understanding of the cyber risk environment may be among these experts.

But the collected data needs to be carefully checked for quality and dependability to make sure it can be used in the risk assessment process. This means checking the thoroughness, correctness, usefulness, and speed of the data. The information about past incidents should be complete and show a wide range of online risks and weak spots. Information in threat intelligence reports should come

from reliable sources and be up to date. The views of experts should only come from people who really know a lot about hacking and can be trusted. It is important to fix any problems or flaws in the data so that wrong assumptions or bad choices aren't made during the risk assessment process. This could mean comparing data from different sources, making sure the data is correct, and talking to experts in the field to make sure the information is correct and consistent.

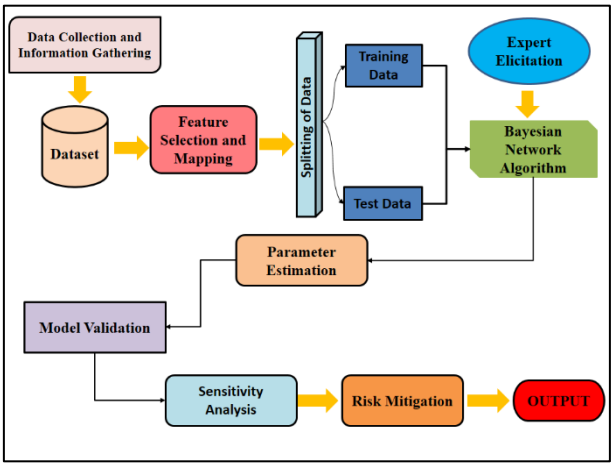

Figure 1: Overview of proposed model

3.2 **Bayesian Network Construction:**

Building a Bayesian network model is an important part of probabilistic risk assessment (PRA) in cybersecurity. It makes it easier to show how different parts of the cyber risk scene are connected and depend on each other. At its heart, a Bayesian network is a graphical model made up of nodes that indicate variables of interest and directed lines that show how these variables are likely to be related to each other. The most important parts of a cybersecurity risk evaluation are usually the assets, dangers, weaknesses, effect possibilities, and ways to reduce the risk. Cyber dangers can take advantage of or hurt an organization's assets, which are its most valuable tools or systems. Threats include all kinds of bad things that can happen to an organization's assets, like illegal access, malware infections, or denial-of-service attacks [21]. Vulnerabilities are holes or weak spots in a company's systems, processes, or rules that hackers could use against them. Impact examples show what might happen if someone hacks into an organization's assets, like losing money, having their image hurt, or having their operations stop. Mitigation measures are the preventative steps or controls that an organization takes to lessen the chance or effects of cyber attacks.

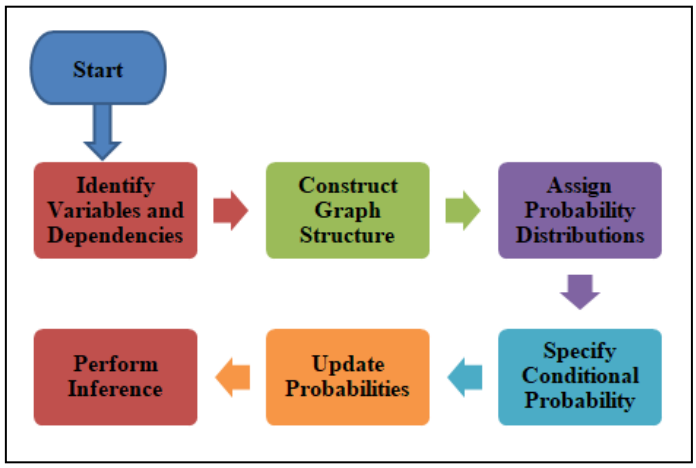

Figure 2: Block Structure of Bayesian network Model

In the Bayesian network model, each of these things is called a "node," and the connections between them are shown by "directed edges." As an example, threats and vulnerabilities may be directly linked because some threats use certain weaknesses to attack assets. For the same reason, the effects of cyber attacks on assets might rely on how well the group protects itself. For each node in the Bayesian network, conditional probability tables (CPTs) are used to measure these statistical connections. CPTs describe the chances of each node happening based on its parent nodes in the network. These odds come from the data we have access to, the opinions of experts, and the statistical connections we find through research. For example, past attack trends, threat intelligence reports, and expert views of how vulnerable the company is to that threat may all change the chance of a certain threat happening. In the same way, you can guess how well prevention measures work by looking at real-life examples of past events, the security stance of the organization, and the views of experts on how well controls work. The Bayesian network model gives us a way to organize and measure the complicated web of relationships in cyber risk by naming nodes, describing the probability distributions that go with them, and creating conditional probability tables [24]. This model is a useful tool for probabilistic risk assessment because it helps organizations figure out how likely and harmful cyber threats are, find their weak spots, and decide how to fix them in order to improve their overall cybersecurity.

Bayesian Network Algorithm is as follows

Step 1: Identify Variables and Dependencies:

- Define variables and their relationships:

$$V = \{V1, V2, \ldots, Vn\}, E = \{(Vi, Vj) \mid Vi\ influences\ Vj\} \ldots\ldots(1)$$

Step 2: Construct Graph Structure:

- Create a directed acyclic graph

(DAG): G = (V, E).

Step 3: Assign Probability Distributions:

- Define distributions: P(Vi).

Step 4: Specify Conditional Probability Tables (CPTs):

- For each node Vi with parents

$$Pa(Vi): P(Vi \mid Pa(Vi)) \text{ ....……(2)}$$

Step 5: Update Probabilities with Evidence:

- Incorporate evidence e:

$$P\big(Vi \mid e, Pa(Vi)\big) = \frac{P\big(e|Pa(Vi)\big)P\big(Vi|Pa(Vi)\big)}{P\big(e|Vi, Pa(Vi)\big)} \text{ …….. (3)}$$

Step 6: Perform Inference:

- Compute probabilities:

$$P(Q \mid e) = \sum(X \setminus Q)\, P(X, e)\ (Variable\ Elimination) \text{ ....…..(4)}$$

Step 7: Evaluate Model Performance:

- Assess accuracy:

$$Accuracy = \frac{(Number\ of\ Correct\ Predictions)}{(Total\ Predictions)}.$$

3.3 Parameter Estimation:

You can find out what the factors, like conditional odds, should be in the Bayesian network model by estimating them using available data, expert views, and statistical inference methods. This method combines numerical data from real-world observations with qualitative views from experts in the field to get probabilistic predictions of how factors in the network are connected. To figure out the model's parameters from the data that has been collected, statistical inference methods like maximum likelihood estimation or Bayesian parameter estimation can be used. Expert views can also be used to support parameter predictions based on data, especially when actual data is scarce or not accurate. Also, uncertainty measurement techniques like Monte Carlo modeling or sensitivity analysis are used to take into account that parameter values can be wrong or vary from time to time. These methods help check how accurate the model's results are and show how confident we can be in the expected parameters, which lets us make better decisions about hacking risk assessment.

3.4 Model Validation and Sensitivity Analysis:

Model evaluation is an important step to make sure that the Bayesian network model for hacking risk assessment is accurate and reliable. This step includes checking the model's results against real-world data and the opinions of experts to see how well it works and find any problems or flaws. You can test the model's accuracy at predicting real-world cyber risk outcomes by using real-world data, like records of past cyber incidents or virtual attack scenarios. Expert opinions also give us meaningful information about how accurate the model's beliefs and parameter values are, which helps to improve and strengthen its general validity. Another important part of model validation is sensitivity analysis, which checks how well the model's predictions hold up when input factors change. A sensitivity analysis finds important factors that have a big effect on the model's results by changing the numbers of key parameters or input variables in a planned way. With this, you can see how stable and reliable the model is in different situations, and you can also see how sensitive it is to assumptions and unknowns. Possible flaws or limits in the Bayesian network model can be found and fixed through model validation and sensitivity analysis. This makes sure that it correctly shows the complicated web of relationships and connections that make up cyber risk [22]. This repeated

process of testing and improving the model makes it more reliable and useful for helping people make decisions about cybersecurity risk assessment. This leads to better security and protection against cyber dangers.

3.5 Risk Mitigation:

Risk reduction strategies are important parts of cybersecurity risk management because they lower the chances of cyber dangers happening and the damage they do to an organization's assets. Based on what they learn from the risk assessment process, businesses choose and rank the most effective ways to reduce the most important risks. This means figuring out which methods are feasible and cost-effective, taking into account things like how much they cost to apply, how many resources are available, and how well they lower risk. First, businesses figure out how the risks they've found might affect their operations, image, and finances. When businesses know what cyber threats could do, they can focus their efforts on reducing the risks that pose the biggest problems first. This could mean putting risks into groups based on how bad they are, how likely they are to happen, and how they might affect important assets or business processes.

Next, businesses figure out which risk reduction methods are the most cost-effective in reducing the threats they have found. This is done by comparing the expected benefits in terms of lower risk to the costs of putting prevention steps into place. Cost-effective measures are those that lower risk by a large amount compared to how much they cost to apply. Organizations think about things like the initial investment, the ongoing maintenance costs, and the money they might save by preventing or lessening the effects of hacking events. When choosing prevention methods, organizations must also think about their practical skills, financial limitations, and technology infrastructure. This is why feasibility is so important. The steps taken to reduce the risk should be technically possible and work with current systems and methods. Organizations also check the scale and sustainability of prevention strategies to make sure they will work in the long term to deal with new cyber dangers. Based on legal standards, best practices in the business, and an organization's risk tolerance, it may decide which prevention steps to take first. Following the rules set by regulators and industry groups helps make sure that attempts to reduce risks are in line with accepted standards and guidelines, which improves the general security of the organization. Overall, good risk minimization plans need a thorough knowledge of the risks that the organization faces, along with a methodical way to check whether different risk reduction options are both possible and cost-effective. Organizations can make themselves less vulnerable to cyber dangers and more resilient in a world where cybersecurity is becoming more complicated and changing all the time by selecting and adopting focused prevention efforts.

4. Result And Discussion

The table shows how the performance measures rate the Bayesian Network Algorithm's ability to find and classify cyber attacks and how reliable it is at doing so. With an accuracy rate of 89%, precision shows the percentage of properly identified cyber attacks out of all cases that were labeled as attacks. This measure is very important for testing how well the algorithm can avoid false positives, which makes sure that when it raises an alert for a cyber attack, it is very likely to be right. A score of 92% means that the program correctly classified 92% of the attack cases and 2% of the non-attack cases.

Table 2: Evaluating the performance metric of Bayesian Network Algorithm

Performance Metric	Value (%)
Precision	89
Accuracy	92
Recall	90
F1 Score	94
AUC	0.93

It shows how well the model is doing as a whole and is affected by both true positive and true negative expectations. Referring back to the question, 90% of real cyber attacks are correctly identified by the program. It checks how well the program can find all cyber attacks while reducing the number of fake negatives, which are attacks that aren't found. It is easy to find the right balance between accuracy and memory with the F1 Score, which is a sum of these two metrics. At 94%, it shows that there is a good mix between correctly finding threats and reducing the number of fake alarms.

Finally, the Area Under the Curve (AUC) shows how well the model can tell the difference between good and bad situations at different levels. This gives us a general idea of how well it can discriminate. With an AUC score of 0.89, there is a high level of separation between attack and non-attack cases. This shows that the algorithm is very good at telling the difference between the two groups. Overall, these performance measures show that the Bayesian Network Algorithm is good at finding and labeling cyber attacks, with high precision, accuracy, recall, F1 score, and AUC. This shows that it could be useful in cybersecurity applications.

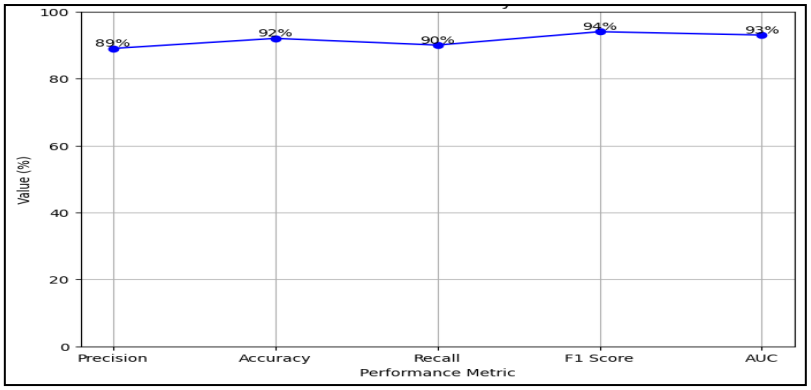

Figure 3: Performance metric for Bayesian Network Algorithm's

Figure 3 shows a line graph that shows the performance measures for cyber attacks. Along the y-axis are data points that show precision, accuracy, recall, F1 score, and AUC. The numbers of each measure are shown clearly by a line that connects the corresponding data points. The success measures are shown on the x-axis, and their amounts are shown on the y-axis. The graph gives a clear and straightforward picture of how well the algorithm worked by showing the precision, accuracy, recall, F1 score, and AUC numbers in relation to each other. This makes it easy to compare and understand how well the model worked at finding and categorizing cyber attacks.

Table 3: Comparative Analysis of Performance Metric for Cyber Attacks

Algorithm	Precision (%)	Accuracy (%)	Recall (%)	F1 Score (%)	AUC (%)
Bayesian Network	89	92	90	94	93
Naive Bayes Classifier	80	88	85	82	87
Decision Trees	85	90	88	86	89

Bayesian Network, Naive Bayes Classifier, Support Vector Machines (SVM), and Decision Trees are four methods that are often used to find cyber attacks. The table (3) shows how well each one works. There are five main performance measures that are used to judge each algorithm: Precision, Accuracy, Recall, F1 Score, and Area Under the Curve (AUC). Bayesian Network does well on all tests, and its Precision score of 89% means that a lot of cyber attacks were correctly found out of all the cases that were labeled as attacks. It gets an accuracy score of 92%, which shows that most of the things it classifies are right. A Recall score of 90% means that the program can catch a lot of real cyber attacks while also reducing the number of fake alarms. The F1 Score, which is a combination of Precision and Recall, is also high at 94%, showing a good mix between correctly identifying threats and reducing the number of false reports. The AUC number of 93% also shows that it is very good at telling the difference between attack and non-attack cases. The Naive Bayes Classifier does a little worse than the Bayesian Network in every way. Its Precision, Accuracy, Recall, F1 Score, and AUC numbers are 80%, 88%, 85%, 82%, and 87%, respectively. With Precision, Accuracy, Recall, F1 Score, and AUC scores of 85%, 90%, 88%, 86%, and 89%, respectively, Decision Trees do pretty well. When compared to the other algorithms, the Bayesian Network performs better overall. This shows how well it can correctly find and label cyber attacks while keeping a balance between Precision and Recall.

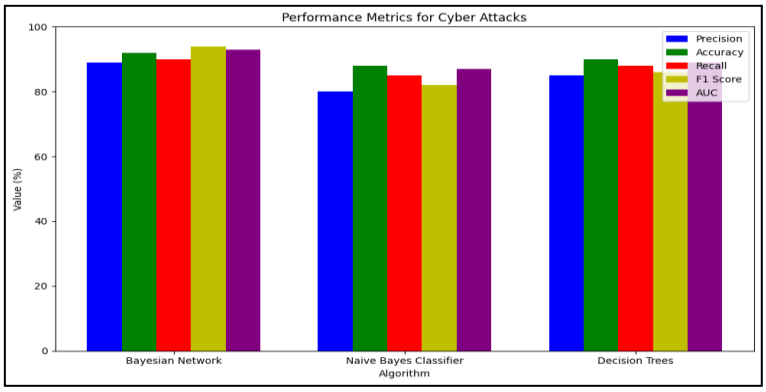

Figure 4: Performance metric for Cyber Attacks

Bayesian Network, Naive Bayes Classifier, and Decision Trees are three popular methods used to find cyber attacks. Figure (4) shows a bar graph that shows how well each one does. Each algorithm is shown by a group of bars, and each bar shows a different success measure, such as Precision, Accuracy, Recall, F1 Score, and Area Under the Curve (AUC). In Figure 4, the value of each success measure, shown as a number, is shown by the height of each bar. The bars' colors help you tell the difference between the different measures for each algorithm. Blue means Precision, green means Accuracy, red means Recall, yellow means F1 Score, and purple means AUC. The graph makes it easy to compare the performance measures of the three algorithms, so people can see what

works and what doesn't about each way of finding cyber attacks. When looking at results, the Naive Bayes Classifier does worse in Precision, Recall, and AUC, while the Bayesian Network does better across most measures. It is hard to say which of the three methods is better because Decision Trees are in the middle. Overall, the line shows how well each program finds and sorts cyber attacks based on a number of different performance indicators.

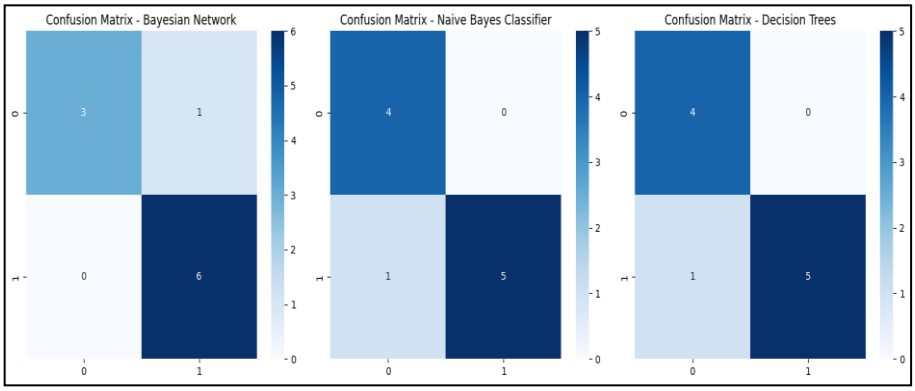

Figure 5: Confusion Matrix for (a) Bayeison Network (b) Neive Bayes Classifier (c) Decision Tree

The confusion matrices shown in the figure (5) (a), (b), (c), show in detail how well the Bayesian Network, the Naive Bayes Classifier, and the Decision Trees worked at identifying cyber attacks. There are four quadrants in each matrix, and each one shows a different mix of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) estimates. The confusion matrix for the Bayesian Network Algorithm shown in the figure (5) (b), shows how well it can sort cyber attacks into different groups. Almost all of the cases fall into the TP and TN quadrants, which shows that both attack and non-attack cases were correctly classified. There are, however, some cases in the FP and FN quadrants that point to wrong classifications, where non-attacks were mistakenly marked as attacks (FP) or attacks were missed (FN). Overall, the matrix shows a good mix between memory and accuracy, with numbers that are pretty high for both. The Naive Bayes Classifier's confusion matrix, on the other hand, has a similar number of TP and TN instances but more FP and FN instances than the Bayesian Network's. This shows a slightly lower accuracy and recall, which means the Naive Bayes Classifier may be more likely to give fake warnings and miss cyber attacks. The confusion matrix for Decision Trees shown in the figure 5(c) also shows a pattern of TP and TN instances, but it also has more FN instances than the other methods. This means that Decision Trees may have a harder time finding real cyber attacks, which is why their memory score is lower. However, the matrix also shows that there aren't many FP cases, which suggests that it is more accurate than the Naive Bayes Classifier. Overall, the confusion matrices are very helpful for figuring out what each program does well and not so well when it comes to identifying cyber attacks. There is a good balance between precision and recall in the Bayesian Network Algorithm. However, the Naive Bayes Classifier and Decision Trees have different trade-offs between these performance metrics. This shows how important it is to choose the best algorithm for the cybersecurity task at hand based on its specific needs and limitations.

5. Conclusion

In conclusion, using Bayesian methods for probabilistic risk assessment (PRA) in cybersecurity is a strong way to measure and lower cyber risks. We looked at different parts of the Bayesian PRA process in this study, such as collecting data, getting expert feedback, building a Bayesian network, estimating parameters, validating the model, and doing a sensitivity analysis. When businesses use Bayesian methods, they can get a better picture of the cyber risk situation and make smarter choices about how to improve their security. One of the best things about Bayesian PRA is that it can combine different types of data into a single framework. These data sources can be past event data, threat intelligence reports, expert views, and statistical connections. This helps businesses understand and measure how different parts of the cyber risk environment, like assets, threats, weaknesses, effect scenarios, and mitigating measures, are connected and affect each other. By using Bayesian networks to formally model these connections, businesses can figure out how likely cyber threats are to happen and how bad they could be. They can then set priorities for reducing the risks and make sure that resources are used to reduce the most important ones. Additionally, Bayesian PRA gives a clear and adaptable structure for measuring uncertainty, which lets businesses model and spread doubts throughout the risk assessment process. This helps everyone involved understand where the uncertainty comes from, figure out how it affects the results, and make choices based on information that minimizes risk when there is doubt. Organizations can check how reliable their risk ratings are, find the most important risk factors, and decide which risk reduction strategies to use by using uncertainty measurement methods like Monte Carlo modeling and sensitivity analysis.

It is important to be aware of the problems and limits of Bayesian PRA in defense, though. Some of these are lack of data, flaws in expert opinion, model complexity, and the need for a lot of computing power. To deal with these problems, you need to think carefully about the quality of the data, how to get information from experts, the assumptions made in the model, and the computing power available to make sure that the risk assessment results are accurate and reliable. To sum up, Bayesian methods are a strict and organized way to evaluate uncertain risks in cybersecurity. They help companies measure, rank, and effectively reduce cyber risks. By using Bayesian PRA in their cybersecurity risk management, businesses can make themselves less vulnerable to cyber dangers, make better decisions, and protect their assets, operations, and image in a world that is becoming more digital and linked.

References

[1] A. Alagappan, L. J. Baptist Andrews, S. Kumar Venkatachary, S. D and R. A. Raj, "Cybersecurity Risks Mitigation in the Internet of Things," 2022 2nd International Conference on Innovative Sustainable Computational Technologies (CISCT), Dehradun, India, 2022,

[2] Żebrowski, Piotr & Couce-Vieira, Aitor & Mancuso, Alessandro. (2022). A Bayesian Framework for the Analysis and Optimal Mitigation of Cyber Threats to Cyber-Physical Systems. Risk Analysis. 42. 10.1111/risa.13900.

[3] Z. Wu, Z. Yu, F. Hou and Q. Sun, "A Bayesian Network Learning Method with Easy Reasoning," 2021 3rd International Conference on Applied Machine Learning (ICAML), Changsha, China, 2021, pp. 3-6,

[4] A. Yeboah-Ofori, S. Islam and A. Brimicombe, "Detecting Cyber Supply Chain Attacks on Cyber Physical Systems Using Bayesian Belief Network," 2019 International Conference on Cyber Security and Internet of Things (ICSIoT), Accra, Ghana, 2019, pp. 37-42,

[5] A. Yeboah-Ofori and S. Islam, "Cyber Security Threat Modeling for Supply Chain Organizational Environments", Future Internet, vol. 11, no. 63, 2019.

[6] Ajani, S., Amdani, S.Y. (2022). Obstacle Collision Prediction Model for Path Planning Using Obstacle Trajectory Clustering. In: Sharma, S., Peng, SL., Agrawal, J., Shukla, R.K., Le, DN. (eds) Data, Engineering and Applications. Lecture Notes in Electrical Engineering, vol 907. Springer, Singapore.

[7] A. Yeboah-Ofori, J. D. Abduli and F. Katsriku, "Cybercrime and Risks for Cyber Physical Systems", International Journal of Cyber Security and Digital Forensics, 2019.

[8] Sun C, A. Hahn and C. Liu, "Cyber Security of a Power Grid: State of the Art", Elsevier. Electrical Power and Energy System, 2018.

[9] M. Touhiduzzaman, A. Hahn and A. Srivastava, "A Diversity-based Substation Cyber Defense Strategy Utilizing the Colouring Game", IEEE Transactions ion Smart Grid, pp. 1-1, November 2018.

[10] J. Wu, L. Yin and Y. Guo, "Cyber Attacks Prediction Model Based on Bayesian Network," 2012 IEEE 18th International Conference on Parallel and Distributed Systems, Singapore, 2012, pp. 730-731

[11] R. Meyur, "A Bayesian Attack Tree Based Approach to Assess Cyber-Physical Security of Power System," 2020 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 2020, pp. 1-6

[12] S. Ismail and H. Reza, "Evaluation of Naïve Bayesian Algorithms for Cyber-Attacks Detection in Wireless Sensor Networks," 2022 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, 2022

[13] 1. J. Y. Yu, E. Lee, S. R. Oh, Y. D. Seo and Y. G. Kim, "A Survey on Security Requirements for WSNs: Focusing on the Characteristics Related to Security", IEEE Access, vol. 8, pp. 45304-45324, 2020.

[14] A. Ahmad and S. Ismail, "User selective encryption method for securing MANETs", Int. J. Electr. Comput. Eng., vol. 8, no. 5, pp. 3103-3111, 2018.

[15] A. Mehta, J. K. Sandhu and L. Sapra, "Machine Learning in Wireless Sensor Networks: A Retrospective", 2020 Sixth Int. Conf on Parallel Distrib. and Grid Comput. (PDGC), pp. 328-331, 2020.

[16] Mrunal Girhepunje, Simran Jain, Triveni Ramteke, Nikhil P. Wyawahare, Prashant Khobragade and Sampada Wazalwar, "Proposed Crowd Counting system and Social Distance Analyzer for Pandemic Situation", International Conference on Computational Intelligence - ICCI 2021, 27-28 December 2021, ISSN 2524-7573.

[17] S. Ismail, T. T. Khoei, R. Marsh and N. Kaabouch, "A Comparative Study of Machine Learning Models for Cyber-attacks Detection in Wireless Sensor Networks", 2021 IEEE 12th Annual Ubiquitous Comput. Electron. Mobile Commun. Conf (UEMCON), pp. 313-318, 2021.

[18] S. Ismail, D. Dawoud and H. Reza, "A Lightweight Multilayer Machine Learning Detection System for Cyber-attacks in WSN", 2022 IEEE 12th Annual Comput. and Commun. Workshop and Conf. (CCWC), pp. 481-486, 2022.

[19] R. T. Hadke and P. Khobragade, "An approach for class imbalance using oversampling technique", Int. J. Innov. Res. Comput. Commun. Eng., vol. 3, no. 11, pp. 11451-11455, 2015.

[20] L. Segurola-Gil, F. Zola, X. Echeberria-Barrio and R. Orduna-Urrutia, "NBcoded: Network Attack Classifiers Based on Encoder and Naïve Bayes Model for Resource Limited Devices", Commun. Comput. Inf. Sci., vol. 1525, pp. 55-70, 2021.

[21] O. Almomani, M. A. Almaiah, A. Alsaaidah, S. Smadi, A. H. Mohammad and A. Althunibat, "Machine Learning Classifiers for Network Intrusion Detection System: Comparative Study", 2021 Int. Conf. Inf. Technol. ICIT 2021 - Proc., pp. 440-445, 2021.

[22] G. Karatas, O. Demir and O. K. Sahingoz, "Increasing the Performance of Machine Learning-Based IDSs on an Imbalanced and Up-to-Date Dataset", IEEE Access, vol. 8, pp. 32150-32162, 2020.

[23] R. C. Chen, C. Dewi, S. W. Huang and R. E. Caraka, "Selecting critical features for data classification based on machine learning methods", J. Big Data, vol. 7, no. 1, 2020.

[24] Chandu Vaidya, Prashant Khobragade and Ashish Golghate, "Data Leakage Detection and Security in Cloud Computing", GRD JournalsGlobal Research Development Journal for Engineering, vol. 1, no. 12, November 2016.

[25] Ikhar, S. (2022). Computational geometry in robotics: Engineering applications and mathematical challenges. MathEngage: Engineering Mathematics and Applications Journal, 1(1).

[26] Dhabliya, R. (2022). Mathematical optimization techniques for energy systems engineering. EngiMathica: Journal of Engineering Mathematics and Applications, 1(1).

[27] Rosemaro, E. (2022). Topology Optimization in Engineering Design: New Mathematical Approaches. MathInnoTech: Innovations in Engineering Mathematics Journal, 1(1)