Framework for calculating Return On Security Investment (ROSI) for security-Oriented organizations

Tahreem Yaqoob, Azka Arshad, Haider Abbas, Muhammad Faisal Amjad, Narmeen Shafqat

PII: S0167-739X(18)31208-1
DOI: https://doi.org/10.1016/j.future.2018.12.033
Reference: FUTURE 4654

To appear in: *Future Generation Computer Systems*

Received date : 16 May 2018
Revised date : 22 October 2018
Accepted date : 15 December 2018

Please cite this article as: T. Yaqoob, A. Arshad, H. Abbas et al., Framework for calculating Return On Security Investment (ROSI) for security-Oriented organizations, *Future Generation Computer Systems* (2019), https://doi.org/10.1016/j.future.2018.12.033

Framework for Calculating Return on Security Investment (ROSI) for Security-Oriented Organizations

Tahreem Yaqoob[a], Azka Arshad[a], Haider Abbas[a], Muhammad Faisal Amjad[a], Narmeen Shafqat[a]

[a]*National University of Sciences and Technology, Islamabad 44000, Pakistan*

Abstract

Today's business environment is extremely dynamic and reliant on innovative Information Technology (IT). Such dependence upon technology leads to an increased rate of successful cyber-attacks whose impact is greater than ever. Due to the exponential increase in security breaches, companies should secure their IT systems by adopting appropriate risk management framework. Organizations have to make justified investments in cyber security. However, it is quite challenging to convince higher management to invest in security measures, since such investments cannot be exactly translated into profits. The Return on Security Investment (ROSI) holds great importance to justify such security investments. A large number of ROSI solutions have already been proposed. However, these solutions do not provide any approach to analyze the impact of single security investment upon whole infrastructure. Furthermore, uncertainty of security incident emerges as another important challenge.The existing ROSI frameworks work on approximations, which can be influenced by employees' exposure and experience, resulting in wrong estimation. The objective of this research is to propose a comprehensive framework to measure ROSI effectively by overcoming gaps in the traditional approaches. The framework has been validated with the help of Common Vulnerability Security System (CVSS) attack dataset. The results show that the annual loss in the absence of security mechanisms is very high i.e. 585,553. However, by following the proposed systematic approach to determine ROSI, it can be reduced to 146,388 which is comparatively low. As a result, organization can save its resources, time, money, trust, and reputation in the market.

Keywords:
ROSI calculations, CVSS, Information security, Annual loss, Investment decisions, Bayesian theorem.

1. Introduction

In today's technology-driven world, companies are largely dependent on IT for the fundamental business procedures. IT forms the base of key business processes and is not only a rationalization tool any more [1]. Such technological dependence has increased the number of successful security breaches to a great extent. British Insurance Company, Lloyds, in a report states that cyber security breaches result in 400 billion dollars loss per year [2]. Moreover, according to 2018 HISCOX Small Business Cyber Risk report [3], 47% of the small organizations experienced at least one cyber-attack during last year whereas 44% of them had two to four attacks. (Gartner), a global research company, estimates that the business organizations around the globe are expected to invest almost 170 billion dollars on cyber security in the coming five years [4], it also predicts that the total spending on information security will reach 93 billion dollars in 2018 [5]. According to 2017 Cost of Cybercrime study [6], the average cost of cybercrimes has increased by 62% during the past five years. At the moment, the business organizations need proactive security measures to prevent the business losses. Therefore, the significance of information security is immensely enhanced over the last couple of years.

Despite the increasing need of cyber security, convincing executive management to invest in security measures is the biggest challenge faced by the security managers today. Before spending money, decision-makers want to know whether that investment can be financially justified in terms of profits or not. The executive management really does not care that Intrusion Detection System (IDS) or firewall protect servers of the organization. Instead, they are more concerned about knowing the impact of such security measures upon the bottom line. It is important to consider that security investments cannot be directly translated into monetary profits, but they can prevent business losses considerably [7]. Therefore, to describe the significance of security investment, it is essential to demonstrate the impact of lack of security mechanism upon productivity. It is significant for the security managers to explain the severity of security breach with respect to a potential loss for the organization. ROSI is an effective approach to justify such investments as it helps in identifying

- Cost-effective solution
- The right amount of money to invest in security
- Impact of security investment on productivity [8]

Numerous methodologies of ROSI exist to help decision makers but they pose great challenges in the domain of cyber security[8]. These frameworks lack some important inputs especially required for cyber security investment. One of the limitations is that these frameworks do not calculate the likelihood of a particular threat mathematically rather the probability of attack occurrence is usually determined by experiences and exposure of employees. Therefore, reliable approximation of risk by using same approach is challenging and organizations usually end up getting different results even under the same conditions [7]. Another important limitation of existing frameworks is that they do not consider the impact of single security investment on the whole architecture of organization. In almost all the cases, a security investment prevents multiple cyber-threats. Traditional approaches allow security managers to observe the benefit of an investment against single problem domain. In contrast, investments in cyber security impact whole infrastructure. One security investment can protect multiple assets against different threats and vice versa. It is, therefore, essential to have a holistic picture of all security investments against multiple threats upon the whole architecture.

This paper proposes a framework for calculating ROSI by overcoming gaps in the existing literature. Our proposed methodology calculates the impact of an attack on the whole business by considering its effect upon all critical assets. For determining likelihood of cyber-attack, we have used the Bayesian theorem [9] [10]that will overcome uncertainty to a reasonable extent. This method is well established and tested. Its inclusion in estimation and uncertainty will be very useful and productive if applied properly [11]. The rest of this research is organized into five sections. Section II is literature review, which discusses existing ROSI mechanisms along with their strengths and weaknesses. Section III proposes a framework for calculating ROSI that can assist managers to invest in security. Section IV is about the evaluation of the proposed framework by comparing it with existing ROSI frameworks. Section V concludes the research along with some future directions.

2. Literature Review

The following section studies the ROSI framework proposed by different researchers that are followed worldwide. These frameworks are studied and compared for identifying best practices in order to develop an improved ROSI framework.

The research [1] mentions that cost estimation and detailed scenario analysis can assist in precise ROSI calculations. The ENISA report [7] also introduced a ROSI metric based on components of risk. ROSI can be computed by considering the annual loss in the absence of security controls. However, the report has some limitations. Firstly, the annual rate of attack occurrence is estimated from intuition and no mathematical formula is proposed for its calculation. The second limitation is related to calculating single loss expectancy (SLE). The report mentions SLE as a loss that will occur if one threat materializes on one asset. However, there can be a scenario where one attack can affect multiple assets. The research in [9] presents a model in which researchers merged the attack graphs and Bayesian networks to explain the ambiguity in cyber security. Practical results depict that by making use of the Bayesian networks better security analysis can be done.

Security Attribute Evaluation Method (SAEM) presented by [12] incorporates cost-benefit approach to choose optimal investment by considering several possible investments. It takes care of the annual losses with respect to the probability of threat occurrence, however, it fails to deliver stand-alone quantitative cost estimates. The Hoover model[13] makes use of Net Present Value (NPV) technique to show profits. It tries to calculate profit from security investment from net benefit. Magnusson et. al proposed a quantitative methodology that uses regression analysis to prevent risk by estimating favorable level of security. However, it does not provide an approach to calculate the monetary value that is to be actually invested[13]. According to [14], to have better evaluation of security investment, return on investment (ROI) should be coupled with a return on attack (ROA) index. However this paper does not provide guidelines to calculate such measures. Another model, System Dynamics Model, [15] was designed to interpret the investment strategies based on policies and scenarios. However, the study lacks financial calculations.

The research [16] presented a model which implements a Fibonacci sequence and moving averages to calculate ROSI. All security incidents were collected in a database. However, the approach to rate these incidents with respect to their occurrence probability has not been covered.Aguiar Rodriguez et. al, [17] highlighted the significance of the cost benefit model for ROSI calculations with the help of comprehensive survey. Although, guidelines for quantifying risk and cost are proposed but mathematical formulas to compute them in the real-world scenarios are lacking. In paper [18] a model based on game theory has been mentioned. However it is a complex approach for mid-sized organizations to adopt and can cause overhead as well.

Mathew and Elisabeth [19] in their research have proposed a Probabilistic Risk Analysis (PRA) approach that assists stakeholders to take decisions regarding security investments proactively. It is based on redevelopment of the Multiarmed Bandits (MAB) problem in order to allocate cyber defense teams to nodes present in the smart grid network. These teams proactively protect the network from intruders by gathering necessary information and taking defensive approaches. This model considers the likelihood of attack and the rate at which nodes could be attacked. Furthermore, the model quantifies the net cyber risk and net financial benefit to assist stakeholders to take effective security investment decisions. This approach is effective as it protects smart grid network before threat materialization and assists management in investing in security, thereby, preventing IT assets against targeted cyber-attacks. The limitation of this research is its complexity for small and medium sized business organizations.

Paper [20] make use of game theory to show the level of security investment to be made against targeted and opportunistic attacks in the interconnected firms. Authors discuss that security threats don't hold equal importance and deserve different level of security. The paper mentions that security investment should increase proportionally as potential loss increases and that joint investment can increase the level of security and decrease the total cost. Authors mention liability and information sharing as two reasons to persuade the firms to make security in-

Table 1
Comparison of Existing ROSI Framework

Title	Methodology	Analysis
Introduction to Return on Security Investment [7]	Risk assessment	Absence of guidelines or mathematical formulas for calculating impact and likelihood of an attack appropriately.
SAEM [12]	Cost-benefit analysis	Approach doesn't provide cost estimates and baseline financial data.
ROA [14]	Return on attack	No guidelines provided for estimation of ease of attack from attacker's perspective.
Return On Security Investment (ROSI) A Practical Quantitative Model [15]	Cost benefit model	Guidelines for quantifying risk and cost are available but no mathematical formulas have been derived.
Forecasting for ROSI Institute for technological research of Sao Paulo [16]	Fibonacci sequence	Does not include Risk management frame work.
Calculating Security Return on Investment [18]	Game theory	It is a complex approach for midsized organizations to adopt and causes overhead.
Calculating Security Return on Investment [18]	savings/cost	This model does not work well for security organizations.
Calculating Security Return on Investment [18]	Returns Expected - Investment Cost/ Investment Cost.	This method does not deal with trends and historical information to calculate attack occurrence probability.
Cyber Risk Analysis for a smart grid [19]	PRA and MAB	Complex for small and medium-sized business organizations.
Optimal information security investment in a healthcare information exchange [21]	Classical economic analysis	Mathematical models use assumptions, not real data.
Decision support approaches for cyber security investment [25]	Game Theory, Combinatorial Optimization and Hybrid of two	Implementation of methodology becomes complex in larger organizations.
Evaluation of complex security scenarios using defense trees and economic indexes [29]	Attack trees	Limitation of this study is that it lacks financial calculations like cost-benefit analysis, Annual Loss Expectancy (ALE) and ROSI.

vestments.Huang et. al [21] examines the relationship between security investment and risks in Health Information Exchange (HIE) by using classical economic analysis. The study uses a network-based approach to analyze the finances of information security investments. The proposed model uses priori principle to check the profits of security investment. Kumar [22] make use of NPV model to propose a methodology to calculate the benefits of information security investments.

The paper [23] proposes a novel Risk Assessment and Optimization Model (RAOM) to solve the security investment decision problem. A Multi-Objective Tabu Search (MOTS) algorithm has been developed to justify the security investments made by an organization. Sawik [24] proposed a security investment model using bi-objective trade-off model. He states that the selected security measures clearly depend on chosen confidence level and cost-risk preference of the decision maker. In paper [25] authors have used three methods to solve the challenges faced by decision makers while making decision regarding security investments. Game theory, combinatorial optimization, and a hybrid of the two methodologies have been considered in this paper. The proposed framework evaluates the effectiveness of controls regarding different assets in the presence of threats. The game has two players, first is the security manager and second is the attacker. In the proposed methodology the direct costs are combined with a Knapsack algorithm to calculate the optimum level of security investment. The combinatorial optimization method uses multi-objective multiple choice Knapsack based strategy. Authors built and validate the decision support tool based on hybrid methodology.

In paper[26] authors have used Game Theory to study the relation between risks and security requirements. In study [27], a methodology has been proposed using Game Theory for estimating security investments. Later, authors have also compared Game Theory and decision theory approaches on parameters like vulnerabilities, investments,and profits from investments. Ashish and Anand [28] proposed a system to evaluate security investments by relating bypass rate with each of the security counter-measures adopted by the organization, table 1 shows a comparison of the aforementioned existing ROSI frameworks. This comparison will further help us in formulating an improved ROSI framework by covering all gaps in the existing literature.

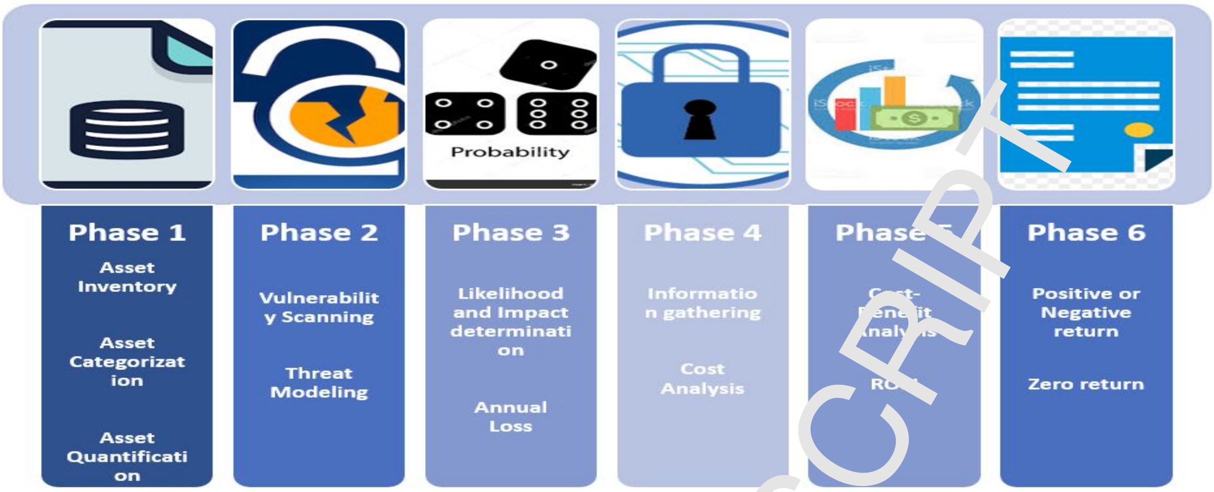

Fig. 1. Proposed ROSI Framework

3. Proposed Return On security Investment Framework

To justify security investments systematically, an improved ROSI framework based on traditional ROSI approaches has been proposed in this section. The proposed framework comprises of six important phases as illustrated in fig. 1. The first phase is about asset categorization and analysis. It assists in developing asset inventory of the organization. The criticality of the asset's operation determines its worth and helps in defining all critical assets of an organization. In the second step, the framework makes use of vulnerability scanning and threat modeling approaches to enlist vulnerabilities of critical assets. The third phase is concerned with calculating likelihood of particular vulnerability/threat as identified in previous phase by using predictive statistical Bayesian theorem [9]. The attack data from CVSS, outcomes of vulnerability scan and threat modeling results will be used as an input to estimate probability of a particular threat. This phase will also determine annual loss of an organization in case vulnerability is exploited and threat is materialized. The fourth phase is about enlisting possible countermeasures that can mitigate threat materialization by covering loopholes present in the assets. It aids in analyzing effectiveness of these countermeasures with respect to business operations of an organization and the loss it covers. The next phase is concerned with calculating ROSI. For this, cost-benefit analysis is performed. The last phase of this framework suggests some recommendations regarding ROSI.

For validation, the proposed framework uses dataset from CVSS results of the vulnerability scan and threat modeling as input to predict occurrence of attack on all critical assets of an organization. The framework specifically applies guidelines provided in [30] to perform threat analysis, identify their occurrence and strategies to combat cyber-attacks. For better comprehension of the phases of the proposed framework, an example scenario of an e-commerce business, as depicted in fig. 2, has been selected. The central sales medium of an internet-based clothes retailer is its online web store. The web-based application of this store is linked with a database. Before placing an order the customers are required to provide all payment and shipping related information. After analyzing and verifying all the details, the web app stores and processes the orders. This scenario will be analyzed for all its critical assets, related vulnerabilities and threats. Based on this analysis, the counter measures will be analyzed in order to compute ROSI. The details of each phase are described in subsequent sections.

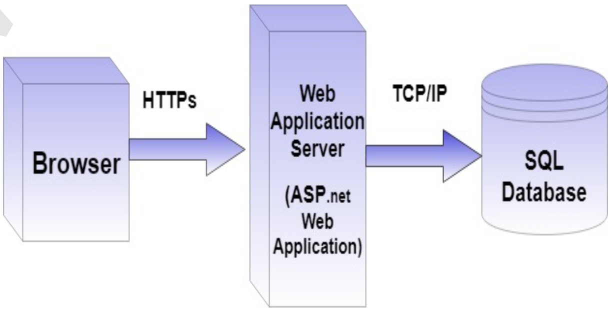

Fig. 2. Example scenario of E-commerce business

3.1. *Phase I: Asset Identification and Analysis*

This phase is all about identifying assets involved in performing business processes. The phase is important as it provides an opportunity to have an insight of all critical assets whose compromise could lead to a huge financial loss. ISO 27001 guidelines [31] are used to identify, categorize and prioritize assets. In our example, we recognized the web-based application, the database, pertinent servers, customer information and credit card details as critical assets as shown in table 2.
This phase is further divided into several important sub-phases.

3.1.1. *Develop an asset inventory*

To develop the asset inventory, it is significant to list all important assets by interviewing the head of each department.The asset inventory of our example scenario is as shown in table 2.

Table 2
Asset Inventory

Assets	C	I	A	Asset Criticality=C+I+A
Web Server	4	3	5	12
Database Server	5	5	4	14
Application Server	3	3	3	9
Workstations	3	3	1	7
Proxy Server	3	3	2	8
Email Server	4	4	2	10

3.1.2. *Prioritization of assets*

This step identifies all critical assets with respect to confidentiality, integrity and availability (CIA) of the stored information. The criticality value can be computed by using eq. (1).

$$Criticality = C + I + A \tag{1}$$

Where C, I and A can have values from 1 to 5, and thus criticality can range from 3 to 15. The higher the value of criticality, the more crucial is the protection of that asset.

3.1.3. *Asset value quantification*

This step is about assigning monetary value to an asset based on its criticality. The table 2 shows that the web server and database server are most critical assets. Asset value (monetary value) can be determined by using eq. (2).

$$Asset\ value = Physical\ Cost * Criticality\ Value \tag{2}$$

3.2. *Phase II: Vulnerability and Threat Identification*

This phase is about identifying crucial security events i.e. threats that exploit known vulnerabilities [29]. These events can be identified based on publicly available data, security expert's knowledge and usage of security scanning tools. Security experts actually make use of their experiences, implicit knowledge as well as explicit data including a list of vulnerabilities for the risk identification[6] . In order to identify vulnerabilities and threats, this framework utilizes CVSS database, penetration testing and scanning tools. This phase is further divided into several important steps to identify threats and vulnerabilities.

3.2.1. *Vulnerability scanning*

This phase determines all loopholes present in the operation, design and security procedures of a system that might lead to successful security breaches. Complete analysis of the system exposes the threat of SQL injection and Cross Site Scripting attack on data base and web server.

3.2.2. *Threat modeling*

The second important step of this phase is threat modeling. Threat Modeling refers to a logical analysis of the system design in order to identify and correct security related issues at initial design-level. Such an approach provides detailed information regarding methodology to attack a particular application by identifying its worth or criticality of information, attack prone areas and a list of applicable threats etc. The vulnerabilities and their entry points in the given scenario are depicted with the help of fig. 3. The web application vulnerabilities are present in the system that makes the system vulnerable to SQL injection and XSS attacks as depicted in table 3.

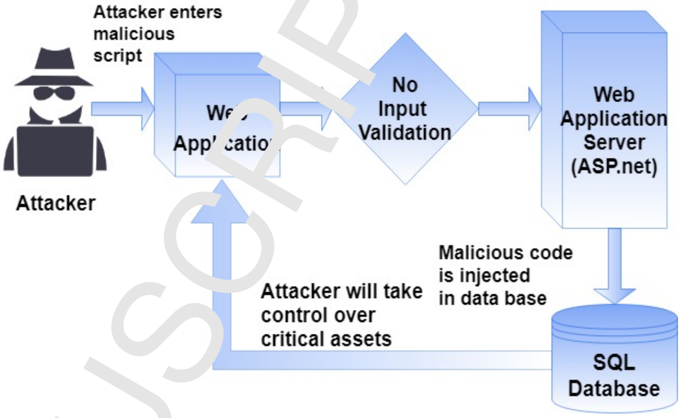
Attacker enters malicious script
Attacker
Web Application
No Input Validation
Web Application Server (ASP.net)
Malicious code is injected in data base
Attacker will take control over critical assets
SQL Database

Fig. 3. Entry Points of Threat Actors

The vulnerability scans of example scenario show that our system has web application vulnerabilities that include code execution and privilege escalation that accounts up to 5% and 15% respectively. Scans also show that web application lacks input validation mechanisms that can lead to attack and it accounts up to 80%. Threats can be materialized by appending SQL commands or executing scripts in the input fields of web application like login, search, register and contact us fields. The materialization of threat can lead to serious loss in terms of money and reputation as confidential information of customers can be easily compromised.

Table 3
Threat Description

Threat Vectors	SQL injection and XSS attack
Threat Target	Data access component
Attack techniques	Attacker appends SQL commands and execute malicious scripts in the input fields of web applications like login, search, register and contact us fields
Countermeasures	Input validation, user error and exception handling

3.3. *Phase III: Likelihood and Impact Determination*

This phase is concerned with estimating the likelihood and impact of a successful security breach with regard to the criticality of the asset. Such value could be computed based on the previous stats. The datasets from the CVSS will be used as a baseline to predict the probability of threat materialization. The limitation of existing approaches is that the probability estimates of threat materialization are unable to provide an accurate view as they are based on employee experiences. Therefore, by using same formula to calculate ROSI, managers end up getting different results even in similar conditions. The proposed framework overcomes such limitation by adopting strong predictive statistical model that is based on the Bayesian Theorem in order to calculate the probability of threat materialization systematically. The dataset to be used covers following important aspects with respect to the example scenario, that later helped in predicting the threat materialization.

- Number of unpatched and known exploits for a specific version of web server,
- Criticality of exploits is determinable in terms of exploits rating,
- The rate of vulnerabilities exposure,
- The ratio of successful vulnerability exploitation.

The Bayesian theorem will consider all these aspects to predict the rate of occurrence.

3.3.1. Likelihood Determination

The Bayesian theorem [9] is a formula that describes methodology to update the probabilities of hypotheses when an evidence is given. This method is time tested and well accepted and its use in estimation and uncertainty can be very advantageous and helpful if applied accurately. The Bayesian approach specifies eq. (3) to calculate the probability of threat materialization.

$$P(X|Y) = \frac{P(X) * P(X|Y)}{P(Y) + P(\tilde{X}|Y) * P(\tilde{X})} \tag{3}$$

- P(X) is the prior probability that is probability of event X happening
- P(X|Y) is the posterior probability i.e. conditional probability of event X in condition that event Y occurs. It assumes that event X is dependent on event Y
- P(Y) is the probability of event Y happening
- P($\tilde{X}$) is the probability of event X not happening
- $P(\tilde{X}|Y)$is the conditional probability of event X in condition that event Y does not occur.

The Bayesian approach is effective as it is not based on a single estimate. Instead, it considers comprehensive posterior distribution for all parameters of each model [32]. As it collects more information, the ambiguity about the parameters of the model will minimize and it progressively narrows posterior distribution. There are several more advantages of such method that include the capability to take account of former knowledge and utilize it systematically in order to accomplish outcome with higher accuracy [11]. The reliable predictive model like the Bayesian theorem overcomes uncertainty in the predictions to a reasonable extent [33].

The likelihood in the given scenario is calculated with the help of some prior statistics. In the sample, SQL injection attack is due to input validation, code execution and privilege escalation vulnerabilities. The attack data from CVSS shows that 45% systems were successfully exploited using input validation vulnerability while 30% systems were exploited due to code execution problems and 25% were attacked due to privilege escalation problem. Similarly, the CVSS data set for XSS attack shows 40.9% systems were successfully exploited using input validation vulnerability while 50.76% systems were exploited due to code execution problems and 8.30% were attacked due to privilege escalation problem. As already mentioned in section 2, the results obtained from the system scanning show that probabilities of assets having input validation, code execution and privilege escalation vulnerability are 80%, 5% and 15% respectively. What is the probability that these vulnerabilities will cause SQL injection and XSS attack in our system? To compute this, first we will compute the probability of SQL injection and then likelihood of XSS materialization will be calculated using eq. (4).

Firstly the probability of SQL injection materialization due to lack of input validation is calculates using eq. (4).

- Let P(Q) = Likelihood of lack of input validation mechanism in our system
- P(R) = Likelihood of code execution problem present in the system
- P(S) = Likelihood of privilege escalation problem in the system
- P(Q|T) = Likelihood of SQL injection attack due to lack of input validation
- P(T|Q) = Likelihood of SQL injection if input validation mechanism is missing
- P(T|R) = Likelihood of SQL injection if unauthenticated remote code is executed
- P(T|S) = Likelihood of SQL injection if privilege escalation problem exists in the system

$$P(Q|T) = \frac{P(Q)*P(T|Q)}{P(Q)*P(T|Q)+P(R)*P(T|R)+P(S)*P(T|S)+P(\tilde{T}|Q)*P(\tilde{Q})} \tag{4}$$

$$= \frac{0.45*0.8}{0.45*0.8+0+0.30*0.05+0.25*0.15+0.2*0.55}$$

$$=0.68$$

Similarly, the probability of SQL injection materialization due to code execution vulnerability is calculated by using eq. (4).

- Let P(Q) = Likelihood of lack of input validation mechanism in our system
- P(R) = Likelihood of code execution problem present in the system
- P(S) = Likelihood of privilege escalation problem in the system
- P(Q|T) = Likelihood of SQL injection attack due to lack of input validation

- P(T|Q) = Likelihood of SQL injection if input validation mechanism is missing
- P(T|R) = Likelihood of SQL injection if unauthenticated remote code is executed
- P(T|S) = Likelihood of SQL injection if privilege escalation problem exists in the system

$$= \frac{0.30*0.05}{0.30*0.05+0.45*0.8+0.25*0.15+0.70*0.95}$$

$$=0.013$$

Likewise, the probability of occurrence of SQL injection because of privilege escalation is computed using steps mentioned above.

- Let P(Q) = Likelihood of lack of input validation mechanism in our system
- P(R) = Likelihood of code execution problem present in the system
- P(S) = Likelihood of privilege escalation problem in the system
- P(Q|T) = Likelihood of SQL injection attack due to lack of input validation
- P(T|Q) = Likelihood of SQL injection if input validation mechanism is missing
- P(T|R) = Likelihood of SQL injection if unauthenticated remote code is executed
- P(T|S) = Likelihood of SQL injection if privilege escalation problem exists in the system

$$= \frac{0.25*0.15}{0.25*0.15+0.45*0.8+0.30*0.15+0.75*0.95}$$

$$=0.03$$

The probability of SQL injection in our system will be computed by taking sum of individual vulnerabilities that lead to the successful materialization of attack.

$$\text{Likelihood of attack} = 0.68+0.01+0.03$$

$$= 0.72$$

Now we will calculate the probability of successful XSS in our system due to lack of input validation mechanism.

- Let P(Q) = Likelihood of lack of input validation mechanism in our system
- P(R) = Likelihood of code execution problem present in the system
- P(S) = Likelihood of privilege escalation problem in the system
- P(Q|T) = Likelihood of XSS injection attack due to lack of input validation
- P(T|Q) = Likelihood of XSS injection if input validation is missing
- P(T|R) = Likelihood of XSS injection if unauthenticated remote code is executed
- P(T|S) = Likelihood of XSS injection if privilege escalation problem exists in the system

$$P(Q|T) = \frac{P(Q)*P(T|Q)}{P(Q)*P(T|Q)+P(R)*P(T|R)+P(S)*P(T|S)+P(\tilde{T}|Q)*P(\tilde{Q})} \tag{5}$$

$$= \frac{0.409*0.8}{0.409*0.8+0.05*0.504+0.15*0.083+0.591*0.2}$$

$$=0.68$$

Similarly, the probability of XSS attack materialization due to code execution vulnerability is calculated by using eq. (5).

- Let P(Q) = Likelihood of lack of input validation mechanism in our system
- P(R) = Likelihood of code execution problem present in the system
- P(S) = Likelihood of privilege escalation problem in the system
- P(Q|T) = Likelihood of XSS injection attack due to lack of input validation
- P(T|Q) = Likelihood of XSS injection if input validation mechanism is missing
- P(T|R) = Likelihood of XSS injection if unauthenticated remote code is executed
- P(T|S) = Likelihood of XSS injection if privilege escalation problem exists in the system

$$= \frac{0.15*0.083}{0.15*0.083+0.409*0.8+0.05*0.54+0.85*0.917}$$

$$=0.01$$

Likewise, the probability of occurrence of XSS attack because of privilege escalation is computed using steps mentioned above.

- Let P(Q) = Likelihood of lack of input validation mechanism in our system
- P(R) = Likelihood of code execution problem present in the system
- P(S) = Likelihood of privilege escalation problem in the system
- P(Q|T) = Likelihood of XSS injection attack due to lack of input validation
- P(T|Q) = Likelihood of XSS injection if input validation mechanism is missing
- P(T|R) = Likelihood of XSS injection if unauthenticated remote code is executed

- P(T|S) = Likelihood of XSS injection if privilege escalation problem exists in the system

$$= \frac{0.05*0.504}{0.05*0.504+0.409*0.8+0.15*0.08+0.95*0.496}$$

$$=0.02$$

The probability of XSS attack in our system will be computed by taking sum of individual vulnerabilities that lead to the successful materialization of attack.

$$\text{Likelihood of attack} = 0.68+0.01+0.02$$

$$= 0.71$$

3.3.2. *Impact determination*

The Impact determination deals with identifying potential loss due to the materialization of threat upon critical assets. The impact of threat materialization upon n assets can be computed by using eq. (6) [15].

$$Impact = \sum_{i=1}^{n} Exposure factor_i * Asset\ value_i + rec\_cost_i (6)$$

- Where i is number of assets
- rec_cost_i is recovery cost that would be incurred in bringing asset back to its normal state. It includes cost of repairing asset, cost of labor trying to recover asset, reputation loss and loss of working hours of employees due to the threat materialization
- Asset value is actual worth of asset in terms of tangible cost as well the criticality of information stored in these assets.
- Exposure factor is the amount of asset being exposed due to the materialization of threat.

In the given scenario, loss due to SQL attack is calculated with the help of table 4 by using the above formula where number of critical assets is 2 (i.e. the database server and web server).

Table 4
Impact Calculation for SQL attack

Asset	Exposure Factor	Asset Value	Recovery Cost
Database Server	45/100	350,000	20,000
Web Server	40/100	300,000	20,000

$$Impact = \sum_{j=1}^{2} 45/100 * 350,000 + 20,000$$

$$+40/100 * 300,000 + 20,000$$

$$=177,500 + 140,000$$

$$=317,500$$

Moreover, the annual loss can be computed by using eq. (7) [15].

$$Annual\ Loss = Impact * Likelihood \quad (7)$$

- The impact is a cumulative loss of all assets due to the materialization of threat (Covered in Sec 3.3.2)
- The likelihood is the probability of the threat materialization (Covered in Sec 3.3.1)

In given case scenario, the annual loss due SQL injection is:

$$\text{Annual Loss} = 317,500 * 0.72 = 228,600$$

Similarly loss due to XSS attack is calculated with the help of table 5 using eq. (6) and eq. (7).

Table 5
Impact Calculation for XSS attack

Asset	Exposure Factor	Asset Value	Recovery Cost
Database Server	70/100	350,000	20,000
Web Server	75/100	300,000	25,000

$$Impact = \sum_{j=1}^{2} 70/100 * 350,000 + 20,000$$

$$+75/100 * 300,000 + 25,000$$

$$=259,000 + 243,750$$

$$=502,750$$

Moreover, the annual loss due to XSS attack can be computed by using eq. (7) as:

$$\text{Annual Loss} = 502,750 * 0.71 = 356,953$$

Now we will calculate overall annual loss due to the materialization of applicable attacks on our system by using eq. (8).

$$Total\ Annual\ Loss = \sum_{k=1}^{n} Loss_k * Liklihood_k \quad (8)$$

- Where k is a list of threats
- $Loss_k$ is loss of all assets due to k
- $Likelihood_k$ is the occurrence of k threats

In order to calculate total the annual loss, we will make use of eq. (8) as given below.

$$\text{Annual Loss} = 356,953 + 228,600 = 585,553$$

$$\text{Total Annual Loss} = 585,553$$

3.4. Phase IV: Countermeasure Analysis

This phase helps in listing all possible countermeasures for further tradeoff analysis. Cost is the most important factor to consider in this phase. Since, in the domain of cyber security, one countermeasure can control the threat materialization in multiple assets or domain. Unlike, traditional ALE approaches, our proposed framework considers this limitation and investigates the impact of one investment on the entire architecture of business process.

$$\begin{aligned} Total\ investment = \sum_{j=1}^{n} Implementation\ cost_j \\ +Maintenance\ cost_j + Installation\ cost_j \\ +Training\ cost_j \end{aligned}$$

Overall security investment can be computed with the help of above equation.

- Where j is number of countermeasures used to control the occurrence of the threat

After studying specifications and reviews of different intrusion detection systems, managers decided to choose Adaptive Intelligent Intrusion Prevention System (AIIPA-SQL) as it can protect threat landscape of business. Maintenance and training cost was calculated after discussion with the security manager and brainstorming sessions. The overall cost of investment is calculated below:

Total investment= 280,000 + 5000 + 5000 + 15000 = 300,000

3.5. Phase V: ROSI Calculation

This phase is concerned with calculating ROSI. For this, it is significant to perform cost-benefit analysis. In the previous phases, we have selected countermeasures and here we will simulate a particular control in our network to have an insight of the countermeasure's effectiveness. After the countermeasure's implementation, the probability of threat materialization should be relatively low. In the given scenario, the probability of threat materialization reduces to 0.25 [34]which means that the countermeasure is beneficial for business.

3.5.1. Cost benefit analysis

Cost benefit analysis assists in evaluating the effectiveness of countermeasure as it compares the annual loss before and after deployment of countermeasure. In the given scenario, we can observe a clear difference in both annual losses. After countermeasure, the loss becomes 585,553 * 0.25 = 146,388 from 585,553 which is very low.

3.5.2. ROSI

ROSI provides executive management a fair idea about security investment and its benefits in terms of controlling the business loss. ROSI in our framework will be calculated using eq. (9)

$$\sum_{i,j,k=1}^{n} \frac{Annual\ loss_{i,k} - Modifiedannual\ loss_{i,k}(j) - Cost_j}{Cost_j} * 100 \qquad (9)$$

- Where $Annual\ Loss_{i,k}$ = Total annual loss due to the materialization of k attacks on i critical asset(s)
- $Annual\ Loss_{i,k}\ (j)$ = Total annual loss due to the materialization of k attacks on asset i after control j
- $Cost_j$ = Total investment cost of j countermeasures (in this case 1 control)

3.6. Phase VI: Recommendation

ROSI states whether investment to be made is beneficial for an organization or not. Positive return exhibits that company should make such investment as it will be beneficial for an organization in coming future. On the other hand, negative return indicates that investment will not be beneficial. Zero return, however; demonstrates that investment will neither be profitable nor will result in any loss.

4. Evaluation and Analysis

The proposed methodology provides greater precision than existing methods. Our method makes use of actual data and expert knowledge in calculating probability of an attack rather than assessments. Traditional ROSI methodologies are based on approximations. Likewise, probability of a cyber-attack is just estimated through historical data or personal experience without any mathematical calculations that might result in wrong estimation. Our proposed framework provides detailed guidelines for asset identification, threat modeling and impact analysis. The probability of an attack is calculated using Bayesian Theorem which reduces the uncertainty to a reasonable extent. Our proposed framework makes use of ISO-27001 guidelines to develop an asset inventory and uses a mathematical formula (eq. (1) and eq. (2)) to quantify and prioritize assets. Asset categorization and prioritization is an initial and significant step to identify assets on which threat materialization can lead to serious loss. Further, it provides a baseline by pointing out the areas where security investment is much needed to prevent financial or reputational loss. However, in other ROSI frameworks, guidelines or mathematical formula for identifying critical assets are missing as shown in table 5. Similarly, the proposed framework provides a systematic approach to identify vulnerabilities of critical assets and suggests vulnerability scanning and threat modeling as vital approaches. In contrast, traditional approaches do not provide any approach to identify vulnerabilities and applicable threats as described in table 5. The ROSI framework uses strong predictive statistical Bayesian theorem (eq. (3)) to determine probability of threat materialization. It takes CVSS attack dataset, vulnerability scan outcomes and threat modeling results as an input to estimate likelihood of threat occurrence in an organization. This theorem provides actual insight of threat occurrence according to which organizations can foresee the loss due to certain security loopholes and can prevent it before any unforeseen attack occurs [9]. However, traditional ROSI frameworks rely on estimation, prior knowledge, employees' experience and assumption that can be misleading. Similarly, mathematical formula (eq. (6)) to

Table 6
Analysis Summary

Specification	**Proposed Method**	**A Practical Quantitative Model [15]**	**ROA [14]**	**Calculating Security Return on Investment [18]**	**ROSI British Computer Society [18]**	**SAEM [12]**
Mathematical formula for asset categorization and prioritization	✓Mathematical formula i.e. criticality = C+I+A and guidelines are provided	×	×	×	×	×
Guidelines for vulnerability and threat identification	✓Vulnerability scanning and threat modeling	✓Approach to determine vulnerabilities of critical assets along with applicable threats is lacking	×	×	×	×
Systematic approach to determine likelihood	✓Bayesian theorem eq. (3)is used to calculate likelihood with the help of CVSS attack data, vulnerability scan outcomes, and threat modeling results	✓Probability i.e.Annual rate of occurrence (ARO) is calculated based on prior knowledge, interviews and assumptions	×	×	×	✓Multi-attribute approach is used to determine estimated, low and high probabilities based on assumptions, prior knowledge and interview
Mathematical formula for calculating impact	✓The framework provides a formula eq. (6)to estimate impact due to cyber-attacks	✓Guidelines are provided	×	×	×	✓Impact is determined based on assumptions, prior knowledge and experience
Mathematical formula to calculate annual loss	✓It provides mathematical formula eq. (7) that takes exposure factor, asset value, recover cost and likelihood to compute annual loss of overall organization	✓Formula SLE*ARO is given but guidelines or formula to determine ARO is based on assumptions	✓To calculate loss, difference between two ALEs are taken but is based on assumptions	×	×	×
Holistic approach to calculate overall loss and security investment	✓It determines threats applicable on critical assets of an organization. Furthermore, it assists in analyzing the impact of single security investment upon whole infrastructure	×	×	×	×	×

calculate single loss is also provided in the proposed framework while other approaches do not provide such formula or guidelines. Annual loss is also determined with the help of probability computed by Bayesian theorem (eq. (7) and eq. (8)), which is based on actual facts and attacks in case of proposed framework. On the other side, traditional frameworks calculate loss based on assumptions that vary owing to employee's experiences, knowledge and influences. Finally, the proposed framework provides a holistic approach (section 3.4) to determine threats applicable on critical assets of an organization. This assists in analyzing the impact of single security investment upon whole infrastructure which is missing in traditional approaches. The summary of comparison between traditional and proposed approaches is discussed at length in table 5.

5. Conclusion and Future Work

Cyber security is becoming top-level priority in many organizations, who are now looking for approaches to protect their critical assets from cyber-attacks using the limited allocated budget. In order to serve this need, we have proposed a framework for designing cost-effective security strategies, which in turn makes the business profitable. The paper also investigates and compares some of the existing ROSI models and proposes an extension to those frameworks. Several methodologies related to ROSI have been presented in the literature. Traditional models are beneficial for cost-effective and optimal investment in general. However, in information security domain, it is very challenging to analyze the impact of one investment on other parts of network or organization. Furthermore, uncertainty is present in current knowledge of threat occurrence and there is a need to overpower such uncertainty to a reasonable extent. We have proposed a framework and guidelines for calculating ROSI and showed how the important factors(likelihood,annual loss and total annual loss) could be obtained mathematically. Unique methodologies to calculate the likelihood and impact of an attack have also been mentioned. The validation of framework is done by using a dataset from CVSS and outcomes of vulnerability scan and threat modeling. The evaluation metric shows that our framework considers uncertainty in an effective manner. In future, it is important to perform detailed experimental analysis of the proposed ROSI approach in real time organizational environments and automate the proposed methodology.

References

[1] C. Locher, Methodologies for Evaluating Information Security Investments - What Basel II Can Change in the Financial Industry,in: ECIS 2005 Proceedings, 2005, p. 122.

[2] S. Gandel, Lloyd's CEO: Cyber attacks cost companies 400 billion every year, (2015). http://fortune.com/2015/01/23/cyber-attack-insurance-lloyds/(accessed 03 October 2018).

[3] HISCOX Small Business Cyber Risk report, (2018). https://www.hiscox.com/documents/2018-Hiscox-Small-Business-Cyber-Risk-Report.pdf (accessed 03 October 2018).

[4] Gartner Forecasts Worldwide Information Security Spending to Exceed 124 Billion in 2019,(2018). https://www.gartner.com/en/newsroom/press-releases/2018-08-15-gartner-forecasts-worldwide-information-security-spending-to-exceed-124-billion-in-2019 (accessed 03 October 2018).

[5] Gartner Inc., Gartner Says Worldwide Information Security Spending Will Grow 7 Percent to Reach $86.4 Billion in 2017, (2018). https://www.gartner.com/newsroom/id/3784965 (accessed 14 March 2018).

[6] Ponemon institute. 2017 Cost of Cyber Crime: Insights on Security Investment that Make A Difference (2017).

[7] Investing in Security for ROI ENISA, (2018). https://www.enisa.europa.eu/news/enisa-news/investing-in-security-for-roi (accessed 14 March 2018).

[8] D. Schatz,B. Rabih, Economic valuation for information security investment: a systematic literature review, Information Systems Frontiers. 19(5)(2017)1205-1228.

[9] P. Xie, L. Jason, O. Xinming, L. Peng, L. Renato, Using Bayesian Networks for Cyber Security Analysis, in:Dependable Systems and Networks (DSN), IEEE/IFIP international conference, 2010, pp. 211-220.

[10] P.Truccoa, E.Cagnoa, F. Ruggerib, O.Grandea, A Bayesian Belief Network modeling of organizational factors in risk analysis: A case study in maritime transportation, https://doi.org/10.1016/j.ress.2007.03.035

[11] Wiecki, T. Predicting future returns of trading algorithms: Bayesian cone - Quantopian Blog. [online] Quantopian Blog, 2015, Available: https://blog.quantopian.com/bayesian-cone/. [Accessed 5 Feb. 2018].

[12] A. Butler, Security attributes evaluation method: a cost-benefit approach, in: Proceedings of the 24th international conference on Software engineering,ACM, 2002, pp. 232-240.

[13] C. Magnusson, J. MoIvidsson, S. Zetterqvist, Value creation and Return On Security Investments (ROSI, in: IFIP International Federation for Information Processing, 2007, p. 232.

[14] M. Cremonini, P. Martini, Evaluating Information Security Investments from Attackers Perspective: the Return-On-Attack (ROA),in:WEIS, 2005.

[15] W. Sonnenreich, Return On Security Investment (ROSI): A Practical Quantitative Model, Journal of Research and practice in Information Technology. 38(1)(2006)p.45.

[16] E. Pontes, A Comprehensive Risk Management Framework for Approaching the Return on Security Investment (ROSI), in:Risk Management in Environment, Production and Economy. InTech, 2011.

[17] A. Rodrgue, Understanding the dynamics of Information Security Investments. A Simulation-Based Approach, Masters thesis, 2017.

[18] O. Don, Calculating Security Return on Investment, Software Engineering Carnegie. Mellon University Institute and US Department of Homeland Security, 2007.

[19] S. Matthew David, and M. Elisabeth Pat-Cornell. Cyber Risk Analysis for a Smart Grid: How Smart is Smart Enough? A Multiarmed Bandit Approach to Cyber Security Investment, IEEE Transactions on Engineering Management, 2018.

[20] W. Yong, G. Feng, N. Wang, H. Liang, Game of information security investment: Impact of attack types and network vulnerability. 42. 10.1016/j.eswa.2015.03.033(2015).

[21] C. Derrick Huang , S. Behara, J. Goo, Optimal information security investment in a healthcare information exchange: An economic analysis. Decision Support Systems, 61, 111(2014)

[22] R. Kumar, S. Park, C. Subramniam, Understanding the value of countermeasure portfolios in information security, Journal of Management Information Systems 25(20)241279(2008).

[23] V. Viduto, C. Maple, W. Huang, D. Lopez-Perez, A novel risk assessment and optimisation model for a multi-objective network security countermeasure selection problem, 53, 599610(2012)

[24] T. Sawik, Selection of optimal countermeasure portfolio in IT security planning, Decision Support Systems (2013), http://dx.doi.org/10.1016/j.dss.2013.01.001

[25] A. Fieldera, E. Panaousisb, P. Malacariac, C. Hankina, F.Smeraldic, Decision support approaches for cyber security investment,https://doi.org/10.1016/j.dss.2016.02.012(2016)

[26] K. Hui, W. Hui, T. Yue, Information Security Outsourcing with System Interdependency and Mandatory Security Requirement,https://doi.org/10.2753/MIS0742-1222290304(2014)

[27] H. Cavusoglu, S. Raghunathan, T. Yue, Decision-Theoretic and Game-Theoretic Approaches to IT Security Investment,https://doi.org/10.2753/MIS0742-1222250211 (2014)

[28] A. Arora, A. Nandkumar, Cash-out or flame out! opportunity cost and strategy: theory and evidence from the information security industry,

http://www.nber.org/papers/w15532(2009)

[29] S. Bistarelli, Evaluation of complex security scenarios using defense trees and economic indexes, Journal of Experimental and Theoretical Artificial Intelligence. 24(2)(2012)pp.161-192.

[30] I. Halpern, E. Kevin, T. Randall, A. Osbourne, A, Information security threat identification, analysis, and management, U.S. Patent Application 14/926,596, filed February 18, 2016.

[31] S. Al-Dhahri, M. Al-Sarti, A. Aziz, Information Security Management System, International Journal of Computer Applications. 158(7)(2017).

[32] N. Feng, W. Harry, L. Minqiang, A security risk analysis model for information systems: Causal relationships of risk factors and vulnerability propagation analysis, Information sciences. 256(2014)pp. 57-73.

[33] H. Hemantha, C. Tejaswini, Investments in information security: A real options perspective with Bayesian postaudit, Journal of Management Information Systems. 25(3)(2008)pp. 337-375.

[34] C. Pinzn, J. Bajo, . Herrero, E. Corchado, AIIDA-SQL: An Adaptive Intelligent Intrusion Detector Agent for Detecting SQL Injection Attacks, 10th International Conference on Hybrid Intelligent Systems (2010)

Tahreem Yaqoob has received her B.S. degree in Computer Science with emphasis in Security of cloud network from Fatima Jinnah Women University, Pakistan in 2014. Currently, she is perusing M.S.in Information Security from National University of Sciences and Technology, Islamabad, Pakistan. Her research interests include security issues in healthcare environment and medical devices.

Azka Arshad has received her B.S. degree in Computer Engineering with emphasis in GSM technology from National University of Sciences and Technology, Pakistan in 2015. Currently, she is perusing M.S.in Information Security from National University of Sciences and Technology, Islamabad, Pakistan. Her research interests include security investment management.

Haider Abbas is a Senior Member IEEE and Cyber Security Professional who took professional trainings and certifications from Massachusetts Institute of Technology (MIT), USA, Stockholm University, Sweden, IBM and EC-Council. He received his M.S. in Engineering and Management of Information Systems (2006) and Ph.D. in Information Security (2010) from KTH, Sweden. He is an associate editor or on the editorial board of a number of international journals including IEEE Journal of Biomedical and Health Informatics, Journal of Network and Computer Applications, Electronic Commerce Research, IEEE Access and Cluster Computing. Dr. Abbas also won many awards and received several research grants for ICT related projects from various research funding authorities and working on scientific projects in US, EU, KSA and Pakistan. He is the principal advisor for several graduate and doctoral students at King Saud University, KSA, and National University of Sciences and Technology, Pakistan.

Muhammad Faisal Amjad is a senior member of the IEEE and an Assistant Professor in the Department of Electrical Engineering, National University of Sciences and Technology Pakistan. He received his PhD degree in Computer Science from the University of Central Florida USA in 2015. His current research focuses on network security, digital forensics and malware analysis. He specializes in dynamic spectrum access and defence against security vulnerabilities in Cognitive Radio Networks as well as wireless sensor and ad hoc networks, game theory and multi-agent systems.

Narmeen Shafqat received her Bachelor's degree in Electrical (Telecommunication) in 2013 and MS in Information Security in 2016 from NUST- MCS with gold medal. Narmeen acquired several information Security trainings from USA including Information Design Assurance Red Team (IDART) Training from Sandia National Laboratories, USA, Training on Forensics and Malware Analysis from Sandia National Laboratories, USA, and Cyber Executive Training from Monterey Institute of International Studies, MIIS, USA. Her research is focused on but not limited to, global cyber laws and strategies, vulnerability assessment, network security, OS security and digital forensics. Narmeen is certified secure computer user (CSCU certified.) Narmeen started her professional career as Research Associate/ Team Lead with a national RnD organization. She has also served as Project Manager and Senior Information Security Officer at a US-based telemedicine company. Currently, Narmeen is associated with MCS - National University of Sciences & Technology (NUST), Pakistan as Lecturer in the Department of Information Security.

Tehreem Yaqoob

Azka Arshad

Haider Abbas

Muhammad Faisal Amjad

Narmeen Shafqat

Highlights

- Justifying security investments always remained a challenge for managers since such investments cannot be exactly translated into profits.
- Research work focuses on identifying limitations present in traditional approaches in order to propose a framework for calculating ROSI based on the Bayesian theorem that overcomes uncertainty to a reasonable extent.
- This research also highlights future recommendations for automating the proposed methodology.