Return On Security Investment (ROSI): A Practical Quantitative Model

Ran Liu
Whiting School of Engineering Security Informatics
Johns Hopkins University
Baltimore, United States

1 Introduction

Yahoo! user information leakage event, that 1 billion accounts were stolen by hackers[1], highlights the necessity of information security investment. Information security investment is different from other investments since investment in security is to prevent events that have not occurred. So how do companies invest in security? How to measure whether the investment in security is effective? What aspects should be invested in? Those are problems which should be considered when an enterprise makes investment decisions. This paper is organized in 3 sections. Section 2 contains the cons and pros of ROSI. Finally, conclusions are discussed in section 3.

2 A RETURN ON INVESTMENT MODEL FOR SECURITY

In order to solve the problems above, the concept of Return of Security Investment (ROSI) is introduced. Generally speaking:

[1] Yahoo: Hackers Stole Data On Another Billion Accounts, https://www.forbes.com/sites/thomasbrewster/2016/12/14/yahoo-admits-another-billion-user-accounts-were-leaked-in-2013/#667f6669d5c2

$$ROSI = \frac{(\text{Risk Exposure} * \text{ Risk Mitigated}) - \text{Solution Cost}}{\text{Solution Cost}}$$[2]

For example, hypothetically, Company ALPHA estimates that it will encounter 4 times information security incidents in a year and each incident results in a loss of $ 20,000. After the introduction of the security program, in the ideal situation, it will only suffer 1 security incident per year, whose cost is $ 20,000, so:

Risk Exposure: $20,000 * 4 = $ 80,000 per year

Risk Mitigated: 75%

Solution Cost: $20,000

$$ROSI = \frac{(\$80{,}000 * \ 75\%) - \$20{,}000}{\$20{,}000} = 200\%$$

So this security program is worth the investment. But if only one security incident will be prevented per year, the ROSI will decrease to 0, which means this security program is not worth the investment.

2.1 CONS AND PROS OF THE ROSI MODEL

There are many advantages of ROSI model.

Firstly, through the ROSI model, the ratio of investment and returns can be quantitatively described, conducive for the organization to make investment decisions clearly. The ROSI makes information security investment easily to understand. Enterprises can use this model to adjust their input resources to get the best output.

[2] Return On Security Investment (ROSI): A Practical Quantitative Model, by Wes Sonnenreich, http://www.infosecwriters.com/text_resources/pdf/ROSI-Practical_Model.pdf

For example, enterprises can choose the information security level based on their requirement, such as financial situations.

Secondly, the data used to ROSI model generally refers to the company's historical information security reports and records and expert knowledge database (FBI and CSI). Thus, the data source is objective in some degree. Some ROSI models are optimized to better meet the requirement of customers such as a model introduced by Intel, which uses accident occurrence data, business impact data and business loss and cost data in the database to get the accurate ROSI of security programs. [3] Those ROSI models are beneficial for the organization to make the right investment decisions.

Thirdly, by comparing the ROSI of different programs, enterprises can make more effective investment decisions. For example, in hypothetical situation, Company ALPHA can choose program A, which will cost $20,000 and the ROSI is 200%. He can also choose program B, which will cost another $20,000 but it can prevent all the security incidents. Company ALPHA can choose the program based on their financial situations and operation strategic to get the best result.

Finally, the model calculates the difference before and after the implementation of the program[4], which can help to evaluate the effect of security program. If not, enterprise can adjust the solution to choose more efficient one.

[3] Measuring the Return on IT Security Investments, Intel, https://communities.intel.com/docs/DOC-1279

[4] Project Risk Management for Hardware Development: How to Calculate Risk Exposure (Part 10), Eric Graves, https://www.playbookhq.co/blog/project-risk-management-how-to-calculate-risk-exposure

However, it is difficult to accurately calculate the returns of security investment due to the following reasons:

1. Enterprise assets especially intangible assets cannot be very specific and clear, and are not easy to be determined;
2. Various security threats are unpredictable and the magnitude of their risks cannot be accurately assessed;
3. Not all security vulnerabilities can be found and the magnitude of their risks cannot be accurately assessed;
4. Not all security vulnerabilities in the enterprise network can be found and new security vulnerabilities can be found every day;
5. Loss is usually caused by multiple vulnerabilities, so it is difficult to determine one-to-one correspondence between loss and vulnerability;

These reasons will be illustrated in detail below.

2.2 It is difficult to determine risk exposure

Generally speaking, one can use single loss expectancy (SLE) * annual rate of occurrence (ARO) to get the Risk Exposure[5]. For example, hypothetically, the security incident ALPHA will lead to that 1,000 employees lose labor productivity and each employee's salary average is $ 10 per hour, so one can use 1000 * $ 10 = $10,000 as the expected minimum loss per hour because of incident ALPHA. It will take 100 hours to solve this problem, so SLE = 1000 * $ 10 * 100h = $ 1,000,000. Expected security incident ALPHA will occur five times a year, so ARO = 5. Thus, the ALE (annualized loss

[5] An introduction to return on security investment – RoSI, https://www.infosecurity-magazine.com/news/an-introduction-to-return-on-security-investment/

expectancy) = ARO * SLE = $1,000,000 *5 =$5,000,000. In general, the data above can be obtained in 2 ways:

1. Obtaining data through the company's historical information security reports and records, including the losses and occurrence frequency of previous threats that occurred;
2. Obtaining data through open resource of CSI, FBI, and Symantec agencies;

However, obtaining data from above ways is not accurate. For the first way that obtaining data from historical information security reports and records, which was first introduced by Intel, they try to predict interim trends in security incident occurrence to derive the financial impact of technology adoption for security programs by analyzing historical cyber-attack incident trending data from similar environments and then extrapolating[6]. However, this method has following problems:

Firstly, most enterprises do not arrange specific staff to be responsible for recording and counting the losses and costs of security incidents. So one cannot get any historical information security data in such enterprises.

Secondly, a single security incident is usually caused by a combination of multiple vulnerabilities, so it is difficult to find one-to-one correspondence between incident and vulnerability. Besides, even though some security incidents looks alike, the reasons behind them may be quite vary which makes the losses are difficult to be assessed.

[6] Measuring the Return on IT Security Investments, Intel, https://communities.intel.com/docs/DOC-1279

Thirdly, after the security incidents happening, some enterprises are busy with resolving security vulnerabilities to reduce losses. [7]Thus, some enterprises are too busy to record every vulnerabilities. Even more, sometimes, they just resolve every vulnerability they find. They will not record statistics data until problems solved, so such data is not very accurate.

Finally, the loss of intangible assets such as reputation is difficult to be estimated.

The second way that obtain data from some agencies such as FBI, CSI, is a common way for most information security cases[8]. However, these data can't be generalized. Taking the example mentioned by Professor Kociemba in class, if a professor's laptop is stolen, the loss may be $ 5,000, which is the tag price of the laptop. However, if there are mid-term test exams in this laptop, the loss shall include the cost of resetting the question of the exams to prevent test exams leakage. The loss may increase to $ 10,000. That is, under different circumstances the costs are different.

Moreover, the loss of productivity is difficult to be determined. In the example above, we use the employee hourly salary multiply by the amount of time to resolve the incident to get the amount of loss. However, the information security incident does not necessarily mean that employees will completely lose their productivity. For example, network breakdown renders employees unable to work with internet. However, some tasks don't need internet. So employees can switch to such kind of work until network fixed. Some kind of incidents like this won't result in losing

[7] Return On Security Investment (ROSI): A Practical Quantitative Model, by Wes Sonnenreich, http://www.infosecwriters.com/text_resources/pdf/ROSI-Practical_Model.pdf

[8] Return on Security Investment (ROSI), 1 June 2011, https://www.finance.nsw.gov.au/policy-document/return-security-investment-rosi

productivity completely. In addition, the resolving time takes 100 hours, perhaps 50 hours of it are non-working hours when employees do not work. To reduce the negative impact on business operation, enterprise prefers to solve problems on weekends. So the so-calculated SLE is not accurate. Someone suggests to use questionnaire survey to get SLE[9]. However, this method still have a great disadvantage. It has high-quality requirements of questionnaire. A question like "How much impact do you think this incident has on your job?" does not help to calculate SLE at all.

2.3 It is difficult to determine Risk Mitigated:

To determine Risk Mitigated is just as difficult as to determine the Risk Exposure for the following reasons.

Firstly, it is difficult to determine whether Risk Mitigated is due to a security solution or not. [10]Taking the previous example of Company ALPHA, after applying the security program into practice, information security incident occurrence frequency decreases from four times to one time per year. It seems the solution works to prevent 3 of 4 information security incidents from happening. However, the decline of 3 incidents may be just due to no security incidents happened within this year by chance. Because the investment to information security is to prevent incidents that have not happened yet, it's hard to determine the information security solution really works or not.

[9] Calculated Risk: Return on Security Investment, Scott Berinato, http://www.csoonline.com/article/2113094/metrics-budgets/calculated-risk--return-on-security-investment.html

[10] Return On Security Investment (ROSI): A Practical Quantitative Model, by Wes Sonnenreich, http://www.infosecwriters.com/text_resources/pdf/ROSI-Practical_Model.pdf

Secondly, Security solution don't work independently. Usually, there are many different information security solutions working together to protect enterprises' assets. For example, to protect sensitive information, access permissions should be set to prevent unauthorized access. Besides, firewall should be updated to provide high level security. What's more, enterprises should prepare backup power and network system in case security system loses power. So we cannot simply determine which solution works and which solution don't work, which makes it difficult to evaluate each security solution separately.

Thirdly, SLE is not exactly the same. Taking the example of Company ALPHA, one calculate the ROSI of solution by assuming each incident would result in the same loss of $20,000. However, it hardly happens in the real world. As discussed before, even though some security incidents looks alike, the reasons behind them may be quite vary, which makes the losses are difficult to be assessed. Besides, same incident does not necessarily mean that SLE is exactly the same. SLE is highly impacted by the circumstance. Considering about this, in the example of Company ALPHA, each SLE may be $ 10,000, $ 20,000, $ 20,000 and $ 30,000 respectively. The sum of each SLE - ALE is still $ 80,000 and their average is still $ 20,000. But what if the loss of the incident which is not prevented is $ 10,000, then ROSI will increase to 250%. And if the loss of the incident which is not prevented is $ 30,000, then ROSI will decrease to 150%. So Risk Mitigate of ROSI can't be accurately determined.

Some research group offer an opinion if one can use a survey to get data, the Risk will be more accurately.[11] However, this method need the high-quality of questionnaire

[11] Project Risk Management for Hardware Development: How to Calculate Risk Exposure (Part 10), Eric Graves, https://www.playbookhq.co/blog/project-risk-management-how-to-calculate-risk-exposure

as discussed before. Other teams suggest that one can use assessment to get “score” based on some algorithm[12]. This score can represent the amount of risk currently being mitigated. This method seems provide more accurate data comparing with other method. However, it’s difficult to design the assessment to get objective data. That is, to get useful data not only require to carefully design the questions on the assessment, but also need carefully choose appropriate algorithm to deal with data. For example, what’s kind of questions should be put on the assessment? How to determine the weight of each question? Should they have the same score or not?

2.4 It is difficult to determine Solution cost

The other factor making ROSI difficult to calculate is that it’s hard to determine the real solution cost. Solution cost is not just a quoted price listed on the contract, but also includes other costs. For example, Company ALPHA purchases the token system to prevent the business information leakage, which is worth $20,000. To let every employee know how to use token system, Company ALPHA has to spend $10,000 to train employees. What’s more, because employees can’t work during the training, the loss of productivities should be calculated into the solution cost, which can be $20,000. So the real solution cost in this example shall include the cost of staff training and the cost of lost productivity. That is $20,000 + $10,000 + $20,000 =$50,000, which is 2.5 times of the purchasing price. But with time going on, it turns out that token system not only can protect sensitive information from leaking, but also can improve the employees’ working efficiency. They found the time and staffs used to protect sensitive data can be used to other work. So the productivity increases which is worth $100,000. In this case, the cost should deduct the earnings brought by increased

[12] Return On Security Investment (ROSI): A Practical Quantitative Model, by Wes Sonnenreich, http://www.infosecwriters.com/text_resources/pdf/ROSI-Practical_Model.pdf

efficiency. So now the solution cost is $20,000 + $10,000 + $20,000 - $100,000 = -$50,000. Other cost should be considered is the maintenance cost. [13]When information security solution adapted, it required specific staff to be responsible for the operation of security system. Besides, the electric power fees and the database to support the operation of the security system and other fees shall be included in the solution cost. Therefore, it's difficult to determine the real cost of solutions. In some research papers, author suggests to take a long-term view because of this reason. The long-term view not only means that to calculate ROSI in a long term[14], but also mean to invest more wisely. For example, enterprise can choose to pay the cost of information security solution for one time, which is worth $10,000. Or enterprise can choose to pay $5,000 per year for 2 years. [15]In Dr.Sonnenreich's opinion, this method can save money to invest other businesses. And the earnings brought by other businesses can decrease the cost of information security solutions. This method is better than simply pay the cost for one time in most cases. However, investment has a risk, which means one might increase cost because of the investment failure. For the reason discussed above, this method has a great disadvantage.

3 Conclusion

In conclusion, the ROSI approach provides a quantitative factor impacting information security to guide decisions of information security investment. To summarize, ROSI model use the data from the previous security disaster and expert knowledge

[13] How to calculate ROSI, http://www.cbdio.com/BigData/2016-09/27/content_5292774.htm

[14] Return On Security Investment (ROSI): A Practical Quantitative Model, by Wes Sonnenreich, http://www.infosecwriters.com/text_resources/pdf/ROSI-Practical_Model.pdf

[15] Return On Security Investment (ROSI): A Practical Quantitative Model, by Wes Sonnenreich, http://www.infosecwriters.com/text_resources/pdf/ROSI-Practical_Model.pdf

database (FBI, CSI), then ROSI model applies the ratio of investment and returns to be beneficial for the organization to make the right investment decisions. Meanwhile, by comparing ROSI of different programs, the pros and cons of programs can be directly compared. Finally, the model calculates the difference before and after the implementation of the solution, which can help to evaluate the effect of security program. However, since ROSI data is difficult to be accurately determined, ROSI model also has many problems. Firstly, it is difficult to determine risk exposure because obtaining data from present ways that discussed before is not accurate. That is, such data cannot be obtained smoothly and the requirement of data is vary from business to business. Secondly, It is difficult to determine Risk Mitigated because one cannot determine information security solutions work or not by exiting methods. Finally, the real cost of solution is difficult to determine because the cost of solution includes: the price of solution, the training cost, maintenance cost and other intangible cost. In a summary, ROSI provides effective method to help the security investment decision. However, when implying ROSI into practice, one should pay special attention to the accurate of data and other cons discussed before.