MINISTRY OF EDUCATION AND RESEARCH

Technical University of Moldova
Faculty of Computers, Informatics and Microelectronics
Department of Software Engineering and Automatics

Admitted to defence
Department Head:
Ion Fiodorov, PhD, associate professor

____

«____» ________________ 20__

QUALITATIVE DATA COLLECTION WITH PROBABILISTIC RISK ASSESSMENT

Master Thesis

Master Student:	**Alexandr Matrohin, SI-201M**
Supervisor:	**Rodica Bulai,** **university assistant**

Chisinau, 2022

EXECUTIVE SUMMARY

This master thesis focuses on the main current problem of cyber risk quantitative assessment methodology using an innovative method of combination triangular distribution and Monte Carlo simulation.

The structure of the thesis consists of 4 main chapters, conclusions from the calculated examples and bibliographical references.

The first chapter will be based on analyses of existing methods of cyber risk assessments, current problems and difficulties, understanding business requirements and needs, describe the confident interval methodology for subject matter expert calibration.

Chapter 2 will be based on the research of setting up 90% confident interval training for subject matter experts to understand effectiveness, applicability, problems and difficulties which may rise in calibration process.

Chapter 3 will describe existing probabilistic models, data analytics and analyses, security metric maturity model, short description and overview.

Chapter 4 will represent research of evaluation of probabilistic models based on business criteria and requirements. I identified the most suitable one which is monte Carlo simulation and verified it based on real data examples. The next step will be feather improvement of Monte Carlo simulation classic model to do simulations based on minimum most likely and maximum values. To achieve this goal I choose triangular distribution and combine it with Monte Carlo simulation. The combination of these two methods gives me possibility to do simulation with all tree input data. I did comparison between combined method and classic Monte Carlo simulation with same input data. Finally, I analyzed the output results of combined method based on different inputs with extreme values to ensure persistent and correctness of this method.

REZUMAT

Teza de master se focusează pe o problemă actuală privind metodologia de evaluare cantitativă a riscului informational, utilizând o metodă inovatoare de combinare a distribuției triunghiulare și simulare Monte Carlo.

Structura tezei este formată din introducere, 4 capitole, concluzii bazate pe exemplele prezentate și referințe bibliografice.

Primul capitol prezintă analiza metodelor existente de evaluare a riscului informațional, problemele și dificultățile actuale privind înțelegerea cerințelor și nevoilor de afaceri, precum și descrierea metodologiei intervalului de încredere pentru calibrarea experților în materie.

Capitolul 2 se bazează pe cercetarea stabilirii instruirii privind obținerea intervalului de încredere de 90% pentru experții în domeniu pentru a înțelege eficacitatea, aplicabilitatea, problemele și dificultățile care pot apărea în procesul de calibrare.

Capitolul 3 descrie modelele probabilistice existente, analiza datelor, modelul de maturitate a unei metrici de securitate.

Capitolul 4 prezintă rezultatele cercetării de evaluare a modelelor probabilistice bazate pe criterii și cerințe de afaceri. A fost identificată ceea mai potrivită, și anume simularea Monte Carlo și s-a făcut verificarea pe baza exemplelor de date reale. În baza rezultatelor obținute a fost îmbunătățit semnificativ modelul clasic de simulare Monte Carlo pentru a face simulări bazate pe valorile minime, cele mai probabile și maxime. În acest scop, a fost aleasă distribuția triunghiulară combinată cu simularea Monte Carlo. Combinația acestor două metode au dat posibilitatea de a face simulare cu toate cele trei datele de intrare. S-a făcut o comparație între metoda combinată și simularea clasică Monte Carlo cu aceleași date de intrare și au fost analizate rezultatele metodei combinate bazate pe diferite intrări cu valori extreme pentru a se demonstra persistența și corectitudinea acestei metode.

CONTENTS

INTRODUCTION

The process of risk assessment and treatment is fundamental to the implementation of an effective cyber security program and plays a crucial role for the national and international regulations in the field of data protection. A complete understanding of cyber risks is necessary, in order to ensure that the security controls an organization has in place are sufficient to provide an appropriate level of protection against cyber threats. However, defining reliable models for the cyber risk exposure is still an open problem. Cyber risk evaluation and the study of its related impact are performed mostly in qualitative ways, which are usually affected by errors and misrepresentations of the risk. They also have several disadvantages, such as the approximate nature of the achieved results and the difficulty of performing a cost-benefits analysis. Quantitative approaches, in their turn, are usually based on scoring systems that associate a certain score to a technological and organizational context.

The qualitative method is commonly implemented does not give a realistic measure of the cyber risk and the related impact. Reliable models for the measure of the cyber risk are not available or have significant limitations, like the lack of generalization and the fact that most works consider only the analysis of past data to derive probabilistic models, while it is not clear how to obtain reliable estimates about future events. Some quantitative approaches the well-known HTMA (how to measure anything) and the FAIR methods rely on a subjective evaluation of the likelihood of an event (in particular, of the probability of a successful attack due to a certain threat) given by a team of experts. These kinds of probabilities usually show some level of inaccuracy and should be replaced by more objective models. The impact of the set of considered threats is then measured in terms of economic loss, which is also subjectively estimated. Based on these premises, the need to improve the quantitative evaluation of the cyber risk of an organization, through a dynamic monitoring of the attacks and vulnerabilities the organization is experiencing, clearly emerges [14, 15, 16].

One way to reduce the uncertainty in this scenario is the one of relying on opinions of experts. This kind of approach is followed in HTMA, where a set of threats is characterized by a likelihood value and the corresponding impact. Basically, the likelihood is the probability of successful attack due to each threat, while the impact expresses the subsequent economic loss. In particular, in the HTMA method, the impacts are estimated through interviews to experts that, for each threat, are asked with the 90% confidence range for possible economic losses. Their answers are then used to define the random variable associated to the impact.

Assuming experts' opinions only as a starting point to be progressively and continuously improved through the acquisition of new and updated information on the organization's behaviour against cyber threats.

The first goal is to improve/calibrate the experts' opinion, to identify the problems, difficulties, efforts and results of calibrating process.

The second goal is to define a methodology for fitting a probabilistic model into a real case study. This approach should to be clear enough for a company's CEO and experts to accept it and transform the risk assessment technics from qualitative to quantitative method.

The third goal is to adopt the methodology to be practical in used for SME (small and medium enterprise), be easy to understand, affordable, precise, scientific and persistent.

It the thesis the survey research is defined as the process of conducting research using surveys that researcher send to survey respondents. The data collected from surveys is then statistically analysed to draw meaningful research conclusions regarding to calibration method. Another part of the thesis based on comparative research of popular existing probabilistic methods to identify the most suitable one for thesis targets.

The thesis structure consists of introduction, four chapters and conclusion.

The first chapter describes existing methodology for risk matrix, qualitative method problems and difficulties, confidential interval and security risks.

The second chapter describes calibration technics, research and analysis. Methods to improve probability calibration and exercise results. Conceptual obstacles to calibration and conclusion.

The third chapter describes existing probabilistic models, data analytics and analysis. Understanding of security metrics and maturity model, Bayesian, Regression Model Predicting Judge Estimates, PERT, Beta distribution and Monte Carlo simulation models.

The fourth chapter describes the main criteria for probabilistic models should have to be used in risk assessment practice for SMEs and example of computing a risk based on Monte Carlo simulation combined with triangular distribution.

BIBLIOGRAPHY

1. W. HUBBARD, Douglas, SEIERSEN, Richard. How to Measure Anything in Cybersecurity Risk. USA : John Wiley & Sons, Inc., Hoboken, 2016. 280 p. ISBN 978-1-119-08529-4
2. HAYDEN, Lance. IT Security Metrics A Practical Framework for Measuring Security & Protecting Data. USA : The McGraw-Hill Companies, 2010. 368 p. ISBN: 978-0-07-171341-2
3. BURTESCU, Emil. Decision Assistance in Risk Assessment – Monte Carlo Simulations. Pitesti Romania : Informatica Economică, 2012 vol. 16, no. 4/2012 86-92 p. ISSN 1453-1305
4. CARLSSON, Elin, MATTSSON, Moa. The MaRiQ Model: A quantitative approach to risk management in cybersecurity. Uppsala : Uppsala Universitet, 2019. 97 p. ISSN: 1650-8319, UPTEC STS 19017
5. Puza, Borek. Bayesian methods for statistical analysis. Australia : ANU eView, 2015. 679 p. ISBN: 9781921934254
6. ÇELIK, Şenol, KORKMAZ Mehmet. Beta distribution and inferences about the beta functions. Turkey : Asian Journal of Science and Technology, 2016. Vol. 07, Issue, 05, pp.2960-2970. ISSN: 0976-3376
7. M. HAILPERN, Susan, F. ViISINTAINER, Paul. Odds ratios and logistic regression: further examples of their use and interpretation USA : The Stata Journal, 2003. Number 3, pp. 213–225 st0041
8. W. HOSMER, David, LEMESHOW, Jr.Stanley, X. STURDIVANT, Rodney . Applied Logistic Regression. USA : John Wiley & Sons, Inc., Hoboken, 2013. 500 p. ISBN 978-0-470-58247-3
9. Adetoye Aribisala, 1Adegboyega Otenaike, 1Olusegun Balogun and 2Lizzy Ofusori . Analysis of an Engineering Project Using Program Evaluation and Review Technique. Nigeria : FUOYE Journal of Engineering and Technology, 2017. Vol. 2, Issue 1, ISSN: 2579-0625
10. Y. RUBINSTEIN, Reuven, P. KROESE, Dirk. Simulation and the Monte Carlo Method. USA: John Wiley & Sons, Inc., 2016. 414 p. ISBN:9781118631980
11. WYROZEBSKI, Pawel, WYROZEBSKA, Agnieszka. Benefits of Monte Carlo simulation as the extension to the Programme Evaluation and Review Technique Poland : The 2nd Electronic International Interdisciplinary Conference September, 2. - 6. 2013
12. MOJTABA HOSSEINI BAMAKAN, Seyed, DEHGHANIMOHAMMADABADI, Mohammad. A Weighted Monte Carlo Simulation Approach to Risk Assessment of Information Security Management System International, Journal of Enterprise Information Systems, 11(4), 63-78, Oct.-Dec. 2015 Access date 15.09.2021 https://www.researchgate.net/publication/289556184
13. THE OPEN GROUP. Risk Management – The Open Group Guide. Zaltbommel, Netherlands: Van

Haren Publishing, 2011. 120 p. ISBN 978-90-8753-663-3.

14. HUBBARD Douglas W., SEIERSEN Richard. Measuring and Managing Information Risk. A FAIR Approach. Hoboken, New Jersey, USA: John Wiley & Sons, Inc., 2016. 280 p. ISBN: 978-1-119-08529-4.

15. FREUND, Jack, JONES, Jack. Measuring and Managing Information Risk. A FAIR Approach. Oxford, UK: Butterworth-Heinemann, 2015. 408 p. ISBN: 978-0-12-420231-3

16. THE OPEN GROUP. Open FAIR - ISO/IEC 27005 Cookbook. Technical Guide. The Open Group, 2010. 44 p. ISBN 1-931624-87-9.