Stochastic Modeling and Applications
Vol.26 No. 1 (January-June, 2022)
ISSN: 0972-3641

CYBER THREADS PREDICTION USING STRUCTURAL EQUATION MODELING (SEM)

SANKALP RAI[1] BHAVANA NARAIN[2] J. DURGA PRASAD RAO[3] AND ANJUL RAI[4]

Abstract. Arithmetical approach of SEM is widely employed designed for assessing and estimating valuable associations between analytical and complex statistical data. It investigates the structural links between measured and hidden components. It can be tailored to meet the demands of law enforcement agencies in identifying and discouraging criminals. In this paper, fundamentals of structural equation modeling (SEM) for cyber-crime prediction is covered.

KEYWORDS: Cyber-crime Prediction, Statistical Data, Structural Equation Modeling.

1. Introduction

In calculating and testing the network of a relationship between variables, SEM is a significant method and tactic. Confirmatory Factor Analysis (CFA), Confirmatory Composite Analysis (CCA), Path Analysis (PA), Partial least Square (PLS) regression, Path Modeling, and Latent Growth Modeling are some of the statistical analysis and modeling techniques used in SEM (LGM). SEM is a collection of matrix equations that is based on linear statistical models. This method employs a grouping of feature psychoanalysis plus Multiple Regression Analysis (MRL). Variables that are measured throughout the data collecting process, while latent variables are variables that are accessed via connecting to observed variables because they cannot be measured directly. The structural model and the measurement model are the two essential components of structural equation modeling. The method of structural equation modeling is similar to that of linear regression analysis. Confirmatory modeling and exploratory modeling are both used in SEM. The majority of the time, confirmatory modeling starts with it. SEM provides every possible generic and practical framework for statistical analysis. Factor analysis, multilinker regression analysis, discriminate analysis and path analysis, seemingly unrelated regression, errors-in-variable models, random-effects models, and other traditional multivariate pressures are included. A graphical path diagram is frequently used to visualize structural equations, and it also illustrates a set of Matrix equations [1]. Many common data analysis methods are generalized by SEM [2].To detect aberrant actions, a concept space algorithm is used. These behaviors can readily be linked to the next time they occur. If cyber forensic and physical evidence is present, it is simple to connect the crimes, categories them, and link them to a single perpetrator. When such evidence is not available, an examination of criminal offence performance can be used to identify a connected series of offences [3][4].

2. Details of The Structural Equation Model

- Latent variables have been depicted as a circle.
- Variables that have been manifested or calculated have been displayed as squares.
- Variance was depicted as double-direction arrows into an object, and residual persisted.
- To give the model scale, the dormant IQ variable has been set to 1.

2.1 Latent and Observed Variables

The observed variable (manifest variable) is the measured variable in the data collection process; latent variables are variables that are measured via connecting to seen variables since they cannot be assessed directly. Latent variables must be represented by numerous observable variables since they reflect abstract concepts. The number of observable variables linked with a latent variable should be at least three in structural equality models. Latent variables in a research model express speculative conception. It crucial to pay attention to the direction of the arrows connecting the observed and hidden variables.

2.2 Reimbursement of Using Structural Equation Modeling (SEM)

Algebraic Modeling - Any statistical representation is arithmetical constraint so as to encloses a pot of numerical assumption about how sample data evolves. The algebraic representation depicts an idealized shape and the process of data evolution. Adopting a statistical model serves three reasons [4][8][9][10].

- Prediction
- Information uprooting
- Representation of stochastic structure

2.3 Conditional Expectation

In conditional expectation, the value is on standard, over a randomly large number of occurrences of events given that a specific set of criteria is known to occur. When the unconstraint variable relies on a set of values, it evaluates the dependent projection of the responsive variable.

2.4 Regression Analysis

There are two main uses for regression analysis. The first prediction and second forecasting, regression investigation is extensively employed in the fields of Machine learning for large overlap, but scenario regression psychoanalysis can be used for basic linkages between the independent and dependent variables. Time series and growth curves are used in regression analysis. Curves, pictures, graphs, and complex compounds are all predicted by regression analysis [9][10].

2.5 Extrapolation

Extrapolation is a method of evaluation of variable's value based on its coalition with another variable outside of the original consideration range. Extrapolation is an area under discussion to superior indecisively and has a higher risk of producing meaningless outcomes [6]. It results in estimation between known observations, but it is an area under discussion to superior indecisively and has a higher risk of producing pointless outcomes.

2.6 Specific Software

The OpenMx R software package generates an unconstrained starting point as well as a rebuilt Mix function version. Second, open-source R has a SEM package. It functions as a software unit within the R library, containing responsibilities and optimizers that enable the SEM model to be executed and evaluated quickly and easily. Based on raw data on the correlation matrix, the model can be predictive. The SEM model can handle both continuous and ordinal data [7][8].

3 Structural Equation Modeling (SEM) For Criminal Behavior

Computer Vision attempts to replicate human vision by evaluating digital picture inputs in the same way that humans do. Detecting an emotion will not be a tough task for a human, but it will be difficult for any machine to do so because they are unfamiliar of human behavior. We employed Euclidian distance calculation to predict human mood. Appropriate behavior is predicted based on the various sequences of emotion. This study studied the links between port cyber security hygiene and cyber risks by conceptualizing and developing three aspects of port cyber security hygiene (i.e., human, infrastructure, and procedure elements) [5][6][7].

4 Methodology

The research methodology used to test the proposed hypotheses is described below. The investigation will be divided into three stages: a pilot study, statistical analysis, and structural equation modeling (SEM). The illustration is shown in the diagram below.

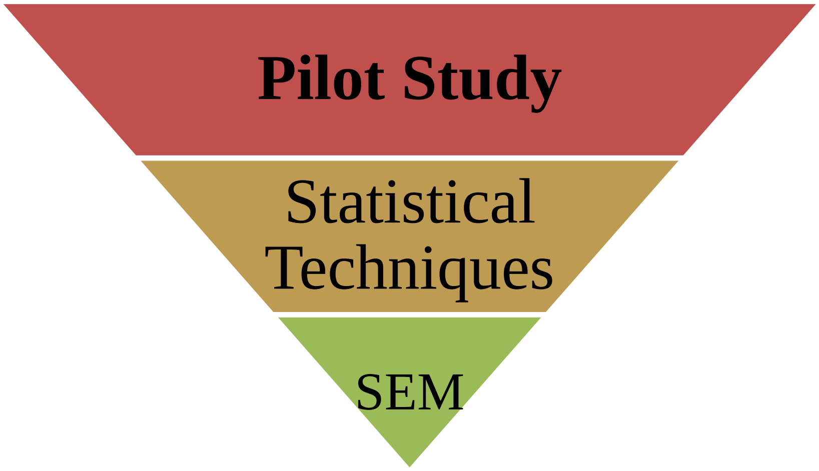

Figure 1.1: Block Diagram of Framework

Phase One:

The first phase of the research will focus on instrument preparation, with a panel of cyber security experts examining the survey instruments' construct validity. Following that, a pilot study of the survey instrument will be undertaken on a data sample to determine its suitability.

Phase two:

The validity of the element integrated in phase one will be the emphasis of phase two. Statistical and modeling tools will be used in the analysis.

Phase three:

In this phase, the focus is on determining the model's validity. SEM techniques will be performed to examine the procedure. SEM attempts to determine the connections between latent variables and the data's fit to a theoretically constructed model.

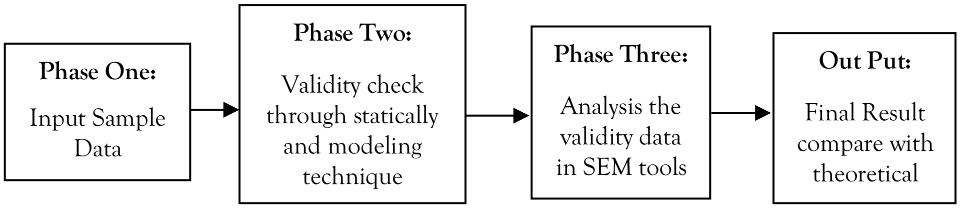
Phase One:
Input Sample Data
Phase Two:
Validity check through statically and modeling technique
Phase Three:
Analysis the validity data in SEM tools
Out Put:
Final Result compare with theoretical

Fig. 1. 2: Flow Diagram of our Proposed Work

To differentiate primary activities handled by the framework, a layer-by-layer technique is used, with each succeeding layer taking input from the previous layer. The proposed CIT framework's input layer is responsible for data collecting and storage. Layer 2 keeps an eye on the threats that come in through Layer 1. Validity check through statically and modeling technique are used as the data starts coming in. Over fitting can be avoided by using cross-validation. In layer 2, phase three SEM tool used for experiment and analysis. Threat management, reporting procedures, and remedies are the emphasis of Layer 3 of the proposed structure.

5.Results

Pilot study, statistical analysis and structural equation modeling done with the help of python. The extracted results have been used to identify human criminal behavior. At the end, the study is extended to identify statistical value and test values. Analysis has been done in the above chapter’s gives us background to develop this research method for getting results.

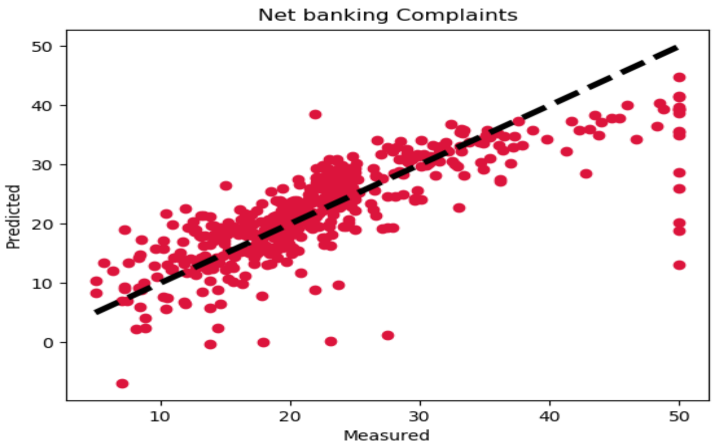

Fig. 7.5 out Put of Net Banking Complaints

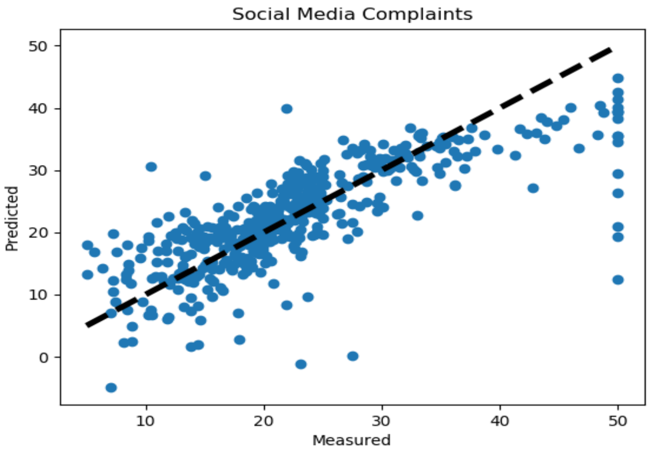

Fig. 7.4out Put of Social Media Complaints

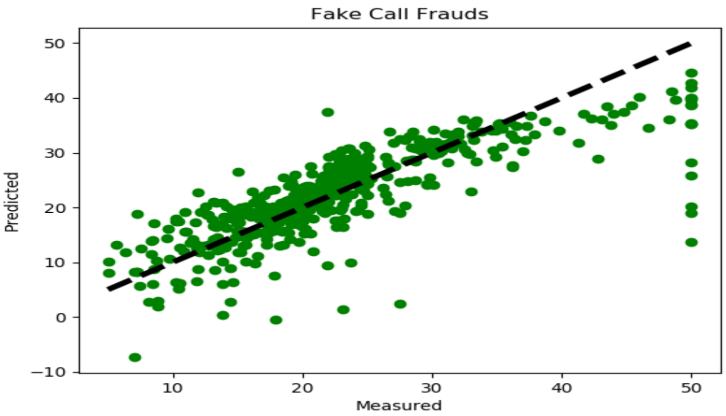

Fig. 7.6 out Put of Fake Calls Complaints

6. Conclusion

Structural equation modeling contains a verity of power full analysis techniques. SEM software and inexpensive computers have made it increasingly easy to apply structural equation models to all sorts of data. For standard analyses one can use any package. Model specification has become much easier with recent versions of the software Structural equation modeling contains a variety of powerful analysis techniques.

Reference

1. Ahmet Okutan , Gordon Werner 1, Shanchieh Jay Yang1 and Katie McConky "Forecasting cyberattacks with incomplete, imbalanced, and insignificant data, Cybersecur 1, 15 (12/ 2018).
2. 2. Big Data Mining Algorithms for Predicting Dynamic Product Price by Online Analysis* M Nayak, B Narain - Computational Intelligence in Data Mining, 2020.
3. C. Machado, A.A. Medeiros Frohlich, IoT Data Integrity Verification for CyberPhysical Systems Using Blockchain, 2018 IEEE 21st International Symposium on Real-Time Distributed Computing (ISORC), 2018.
4. Data encryption standard algorithm in multimodal biometric image, SS More, B Narain,BT Jadhav - Int. J. Comp. Sci. Eng, 2018.
5. G. Settanni, F. Skopik, A. Karaj, M. Wurzenberger, R. Fiedler, Protecting cyber physical production systems using anomaly detection to enable selfadaptation, 2018 IEEE Industrial Cyber-Physical Systems (ICPS), 2018.
6. Impact of digital image processing on research and educationB Narain, AS Zadgaonkar, S Kumar Natl Semin Work, 2013
7. J. Wang, W. Tu, L.C.K. Hui, S.M. Yiu, E.K. Wang, Detecting Time Synchronization Attacks in Cyber-Physical Systems with Machine Learning Techniques, 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS) , 2017.
8. K.-K. Choo, M.M. Kermani, R. Azarderakhsh, M. Govindarasu, Emerging embedded and cyber physical system security challenges and innovations, IEEE Trans. Dependable Secure Comput. 14 (3) (2017) 235–236.
9. M.S. Mahmoud, M.M. Hamdan, U.A. Baroudi, Modeling and control of CyberPhysical Systems subject to cyber attacks: a survey of recent advances and challenges, Neurocomputing. 338 (2019) 101–115.
10. More S.S., Narain B., Jadhav B.T. (2021) Advanced Encryption Standard Algorithm in Multimodal Biometric Image. In: Rizvanov A.A., Singh B.K., Ganasala P. (eds) Advances in Biomedical Engineering and Technology. Lecture Notes in Bioengineering. Springer, Singapore. https://doi.org/10.1007/978-981-15-6329-4_7

Sankalp Rai: Sub Inspector of Police, Crime Investigation Department (C.I.D.), Police Headquarter, Atal Nagar, Naya Raipur,492002, Chhattisgarh India
E-MAIL: sankalp.rai@gov.in, sankalprai@gmail.com

Bhavana Narain: Faculty of Computer Science and Application, MATS University Raipur, Chhattisgarh, India
E-MAIL: dr.bhavana@matsuniversity.ac.in

J. Durga Prasad Rao: Faculty of Computer Science and Application, Shri Shankaracharya Mahavidyalaya, j unvain, Bhilai, Chhattisgarh, India
E-MAIL: j.durga.prasad.rao@gmail.com

Anjul Rai: Faculty of Mechanical Engineering Department, Bhilai Institute of Technology, Durg,491001, Bhilai, Chhattisgarh, India
E-MAIL: anjul.rai@bitdurg.ac.in