MANTECH Publications

Journal of Software Engineering & Software Testing
Volume 7 Issue 2
ISSN: 2457-0516 (Online)

Prediction of Software Errors: An Initial Exploratory Study in SMEs

Dr. Upendranath Nanda[1], Manoj Kumar[2]

Professor[1], Student[2]

Department of Computer Science Engineering

Pemiya Rishikesh Institute of Technology

Corresponding Authors Email: - *kumarmanoj222@gmail.com*

Abstract

The technique by which classification systems are meant to forecast faulty code snippets, to use software archive historical facts including code sophistication, and to alter records for software fault metrics is known as software defect prediction. We attempted to understand the links between numerous variables that are crucial for IT SME's in this study piece. The study employs a well-structured questionnaire and the IBM SPSS programme for data analysis and interpretation.

Keywords: *SME's, defect-prevention methods, IBM SPSS*

INTRODUCTION

Deferring remedies until testing and operational stages may result in higher expenses, and it may be too late to fundamentally alter the system. Continuous research in the field of PC programme immovable quality has been composed toward the identification of software modules that are presumably going to be issue slanted, in view of item, process-related estimations, before the testing stage, with the goal that early distinct verification of weakness slanted modules in the life-cycle can help in occupying programme testing and affirmation tries in the advantageous way. Perlis has been concerned with software estimates, which address quantitative portrayal of programme qualities and the key role they play in visualising the idea of the software. That is, there is a direct relationship between some multidimensional nature estimations and the number of adjustments ascribed to

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

difficulties later discovered in testing and approval. Crawford believes that numerous variable models are essential for obtaining large estimations regardless of programme size.

In this way, examining the relationship between the quantity of defects in programmes and software comprehensive design estimations dismantles to the advantage of future researchers. Multifaceted design and size estimations have been utilised in an attempt to predict the number of flaws that a structure would disclose during operation or testing. Generally, attempts have mostly focused on the following three topic perspectives:

- Estimating the number of flaws in the structure.
- Estimating the structure's tenacity in the face of time and disillusionment.
- Recognizing how structure and testing types affect flaw counts and discontent density.

The consistency improvement programme is a well-organized technique for identifying immovable quality issues by testing, distinguishing such defects, and implementing corrective steps to slow the rate of occurrence. Some of the notable ideal circumstances of the lasting quality improvement programme incorporate achievement assessments and anticipating the item constancy designs. Regardless, the fundamental layout of a steadfastness advancement programme incorporates three primary elements. These are the administrators, testing, and frustration uncovering, evaluation, and beneficial movement structure.

Unwavering quality development models are commonly referred to as Probabilistic and Statistical models. The data obtained throughout the procedure cannot be combined because to cloud factors associated with probabilistic constancy advancement models. Cloud characteristics are associated with genuine resolute quality progression models.

Time self-ruling reliability advancement models-the quantity of disappointments or repairs in unambiguous time between time are not dependent on time. Time subservient unwavering quality advancement models - a trustworthiness advancement model has a time restriction.

These are temporal models for continuous reliability improvement. Discrete reliability improvement models are appropriate for unrecoverable articles because they have two discrete states: a consistent quality working condition and a

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

mistake. Classically unwavering quality improvement models- logical rigging is probability theory. Duane's unwavering quality improvement model and its modifications are endless, time-consuming, and factual models. Stochastic lasting quality advancement models- a non-stationary stochastic approach for improving dependability. Quality advancement models based on Bayes and semi-Bayes. Unconventional lasting quality advancement models- There are all reliability advancement models for which there is no current belief to arrange all together classes.

Principles of Defect Prevention

How does a gadget that removes faults work? A defect reduction mechanism is required for the technique (Figure 1.). A specification analysis kicks off the critical portion of the fault avoidance cycle, which is turning customer preferences into product specifications without creating any further errors. In order to analyse failures, we establish programme systems, perform code inspection, and controls in line with defect monitoring and reporting.

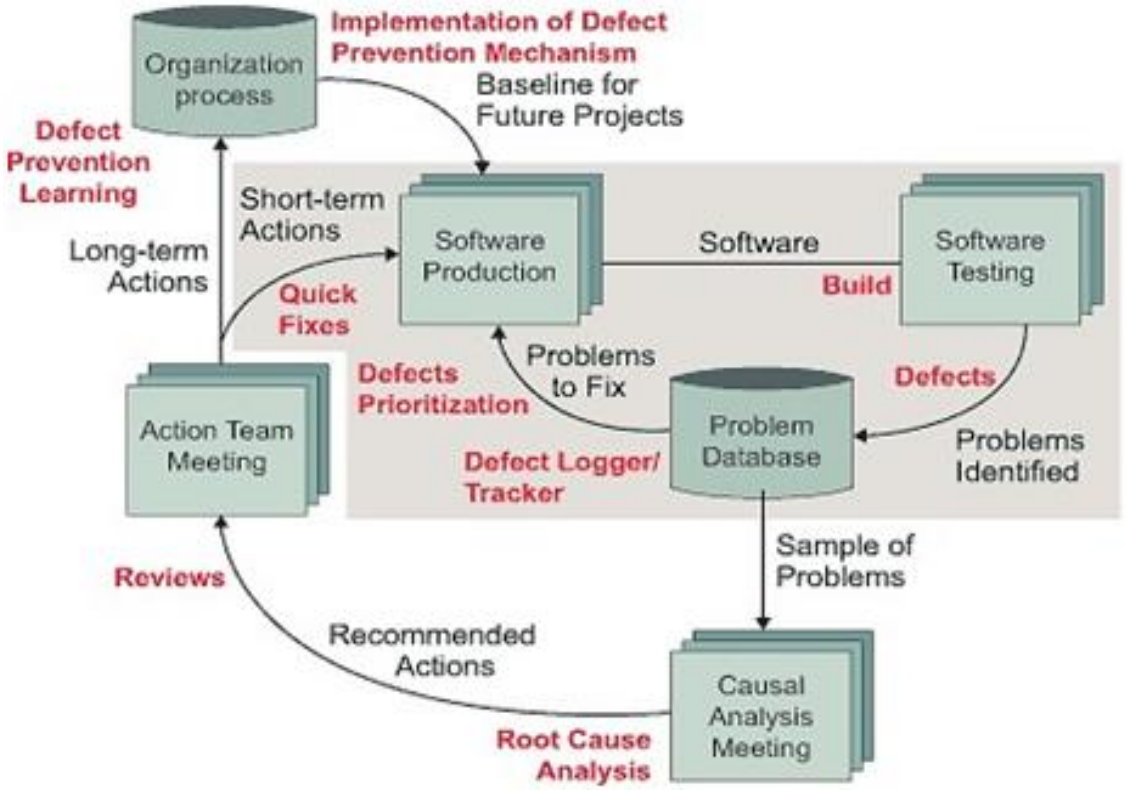


Fig.1:-Defect Prevention Cycle (Source: 1998 IEEE Software Productivity Consortium)

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

The methods and processes in the gray-colored frame illustrate the management of faults in many of the software industry's established paradigms - detecting, tracking/documenting, and assessing errors in order to obtain quick, short-term solutions. Strategies that are key components of defect-prevention methods are depicted on the white background. The essential phase in the defect prevention method is to examine problems in order to determine their source, provide a speedy remedy, and avoid future defects. The organisation is offered as a model for future initiatives, with the backing and confidence of its team leaders.

REVIEW OF LITERATURE

Software Defect Prediction (SDP) is one of the most important operations in the SDLC Testing field. It detects modules that have slanted square distortion and require further testing. As a result, the testing resources will be used productively while without jeopardising the needs. Despite the well-known fact that SDP is useful in testing, it is not always easy to predict the faulty modules. There were several unexpected challenges that hampered the smooth execution despite the usage of Defect Prediction models. This paper noticed a part of the distinct challenges with SDP and pondered what has been done so far to control them. Imperfectness Prediction in Bundle is viewed as the primary beneficial and esteem talented movement. According to bundle specialists, it is a critical element on which the character of the product being made is dependent. It has played an important role in lowering the expenditures on the bundle business, of not being able to pass on the necessities within budget and on time. Aside from that, the clients' reaction to the product quality has exhibited a significant shift from perplexing to tolerable. Because of the part of revamping the bundle quality, there is a demand for bundle distortion. It leads the United States to quantify the longer term, for example, to differentiate the modules that square measure likely to have troubles. This assists the bundling venture higher-up team in dealing with those regions inside the project on partner favourable terms and with adequate work. This evaluation area has existed since the 1990s. With nearly twenty-four years of existence, this location still need subsidence two or three difficulties. This study has appeared and examined what has been done thus far and what should be done in the future. Six issues were discussed in total: the consideration of attributes as related to issue, the gathering activity of customary measures for

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

execution evaluation, issues with cross undertaking distortion desire, cash parts of bundle imperfectness figure and modernity cumbersomeness issue, and the absence of any wide structure for the bundle disfigurement estimate.

The bundle business is profitable if it can attract the general public's attention. This is feasible provided the organisation can produce high-quality goods. To distinguish an object as being of high quality, it should be free of flaws and capable of delivering anticipated outcomes. It should be sent in partner expected value, time, and be practicable with the least amount of labour. Imperfectness impedance is the most obvious, but frequently overlooked, component of any task's bundle quality assertion. When used at all stages of package development, it will decrease the time, money, and resources required to style a high-quality item. A minor increase in the compensatory activity live may usually result in a large drop in overall quality cost. However, the main purpose of value examination isn't to make certain that the value invested square measure the finest potential very cost, which magnifies the advantage obtained from that venture. Because of the value examination, the emphasis has shifted to defect blockage. Over a period of time, it has been found in a huge number of organisations that when the business execution improves, bundle quality improves, and the cost of value decreases owing to the utilisation of powerful defect identification and impedance actions. It is undeniably undeniable that multivariate analysis has been used to develop a prediction model for framework testing failures. The investigation will authenticate the definitions and significance of measures from interest, style, and cryptography current forecasting flaws for framework testing by using applied math approaches such as multivariate examination. Furthermore, it has been revealed that in order to have a decent model, the prediction should fall within a sketched out least and most shift, making it possible to integrate and execute defect prediction as a piece of bundle improvement approach, particularly examine technique. A forecast of faults misuse whole range is not recommended since it requires extremely strong facts and rigorous data gathering to be used for constructing such a model. The exercises were subjected to several limits while effecting the test. Due to the limited area of data collection, the inquiry initially produced just one generic model. Second, the information obtained is limited to package progression routes where their measurements square measure

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

carefully gathered and followed. Comes that were not concerned with measuring variety are no longer a part of the data acquired. The third limitation is that this research is solely focused on the shaped advancement model because that is the method model used by the association first class for this inquiry. Fourth, the data set used in this study might be a combination of measurements from electronic and part-based bundles. In this way, the findings of the inquiry square measure a conclusive result of misuse measures from all programme types.

RESULTS & ANALYSIS

Data was obtained from software professionals using a questionnaire survey and analysed using IBM SPSS. The stratifies sampling approach was used, with a 50% response distribution, a 5% standard error rate, and a 95% confidence level.

The table below shows the Anova findings used to investigate the correlations between the various factors examined for our investigation.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
The business objectives for the project and the product were well defined at the start of the product itself.	Between Groups	85.928	2	42.964	109.155	.000
	Within Groups	144.453	367	.394		
	Total	230.381	369			
The business objectives for the project and product were documented at the start of theproject.	Between Groups	123.121	2	61.560	158.062	.000
	Within Groups	142.936	367	.389		
	Total	266.057	369			
The business objectives for the product and the	Between Groups	20.855	2	10.427	10.704	.000
	Within	357.526	367	.974		

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

project were understood by the development team at the start of the project.	Groups					
	Total	378.381	369			
The business objectives of the project and the product were understood by the customer at the start of the project.	Between Groups	1.416	2	.708	.564	.569
	Within Groups	460.857	367	1.256		
	Total	462.273	369			
There were one or more persons at the customer site who was/were clearly responsible and available for customer decision making.	Between Groups	6.650	2	3.325	4.024	.019
	Within Groups	303.231	367	.826		
	Total	309.881	369			
Were there several customer departments involved in the project?	Between Groups	18.622	2	9.311	13.541	.000
	Within Groups	252.353	367	.688		
	Total	270.976	369			
The customer departments had conflicting interests which had to be resolved.	Between Groups	253.364	2	126.682	412.757	.000
	Within Groups	112.638	367	.307		
	Total	366.003	369			
At the start of the	Between	25.632	2	12.816	32.142	.000

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

project, the familiarity with and comprehension of the application domain of the key people on theproject.	Groups					
	Within Groups	140.356	352	.399		
	Total	165.989	354			
At the start of the project, the familiarity with and comprehension of the platform to be used (e.g., programming language(s), Operating System, database management systems) of the key.	Between Groups	59.409	2	29.704	73.337	.000
	Within Groups	148.648	367	.405		
	Total	208.057	369			
At the start of the project, the familiarity with the type of system architecture used (e.g., client- server, Internet JAVA applications) of the key people on the project.	Between Groups	56.844	2	28.422	83.126	.000
	Within Groups	125.483	367	.342		
	Total	182.327	369			
At the start of the project, the familiarity with and comprehension of	Between Groups	56.844	2	28.422	83.126	.000
	Within Groups	125.483	367	.342		

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

the software development environment (e.g., compiler, code generator, CASE tools) of the key people on the project.	Total	182.327	369			
At the start of the project, the ability to communicate easily and clearly with the others (e.g., good interviewing skills and other information gathering techniques, good verbal communication skills,ability to lead people) of the key people on the proje	Between Groups	26.806	2	13.403	34.324	.000
	Within Groups	143.305	367	.390		
	Total	170.111	369			
At the start of the project, the knowledge and experience of the software development process andtechniques to be used during the project (e.g., functional and/or object modeling techniques,	Between Groups	85.144	2	42.572	58.773	.000
	Within Groups	265.832	367	.724		
	Total	350.976	369			

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

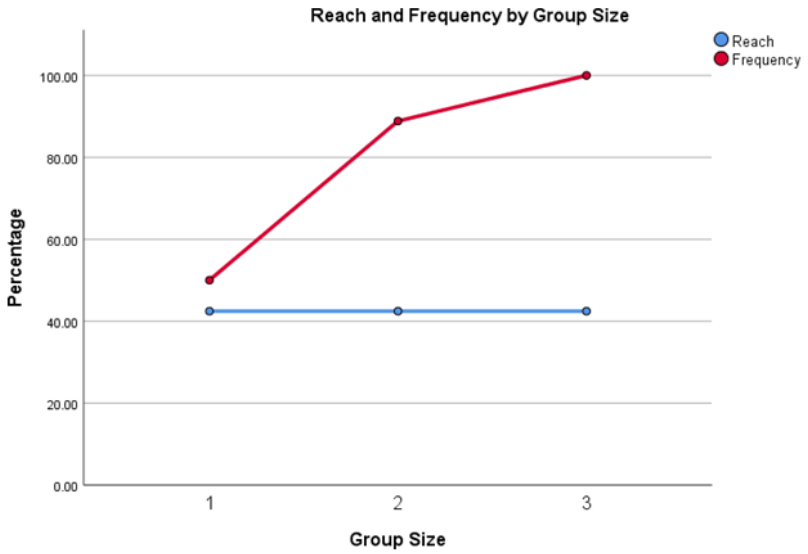
testing techniques, and conducting a cost/benefits analysis) of						
At the start of the project, the knowledge and experience of the documentation standards to be used during the project (e.g., modeling notation, and requirements document structure and content) of the key people of the project.	Between Groups	144.864	2	72.432	109.889	.000
	Within Groups	241.904	367	.659		
	Total	386.768	369			

From the Anova table we can understand the existing relationships type of companies and various factors considered , Factors related to business objectives and planning phases were administered, F test > F critical value and Sig< 0.05 indicating strong relationships between the variables.

Heat: 0. Maximum Group Size: 3. Reach and Frequency.				
Variables	Statistics			
	Reach	Pct of Cases	Frequency	Pct of Responses
Q12, Q17, Q18	157	42.4	314	100.0
Q17, Q18	157	42.4	279	88.9
Q12, Q18	157	42.4	192	61.1
Variables: Q12, Q17, Q18				

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

Information	
	Count
Set unions to be calculated	4.000
Sets of variables analyzed(unions plus one-variable sets)	7.000
Assumes no variables are forced	



Model Description		
Model Name		MOD_1
Analysis Type		Univariate
Series Name	1	1. Please indicate the type of company
	2	The business objectives for the project and the product were well defined at the start of the product itself.
	3	The business objectives for the project and product were documented at the start ofthe project.
	4	The business objectives for the product and the project were understood by the development team at the start of the project.
	5	The business objectives of the project and the product were understood by the customer at the start of the project.
	6	There were one or more persons at the customer site who was/were clearly responsible and available for customer

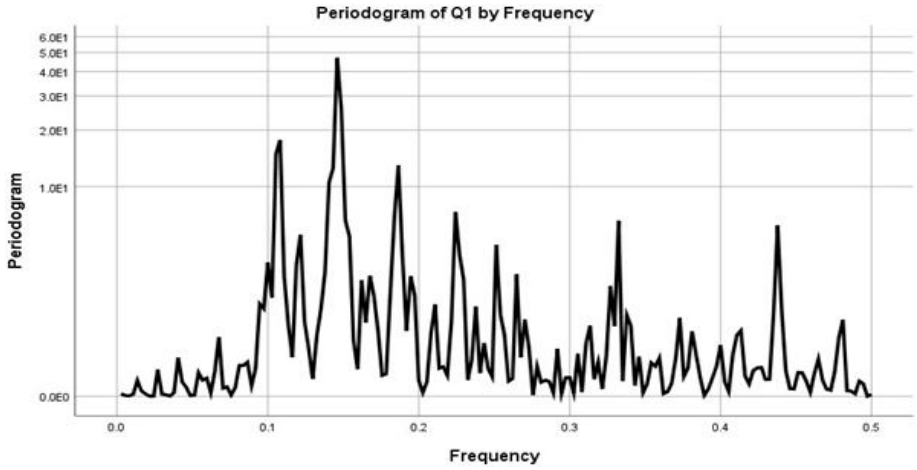
Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

		decision making.
	7	The customer departments had conflicting interests which had to be resolved.
	8	At the start of the project, the familiarity with and comprehension of the application domain of the key people on the project.
	9	At the start of the project, the familiarity with and comprehension of the platform

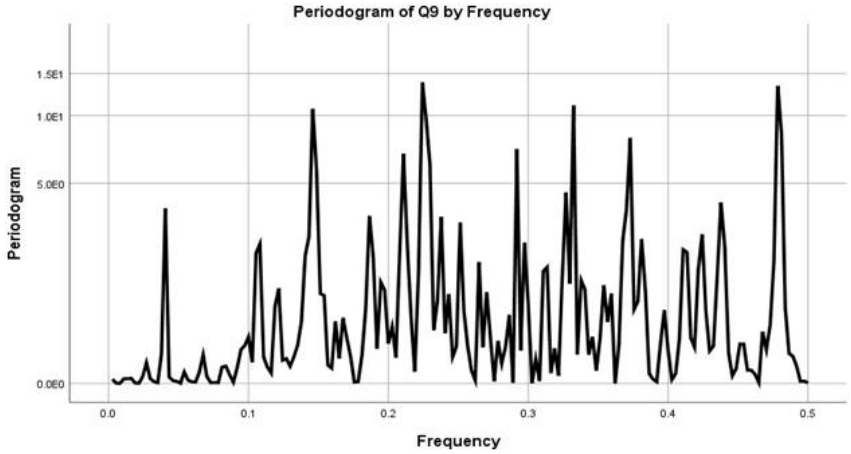
			to be used (e.g., programming language(s), Operating System, database management systems) of the key.
	10		At the start of the project, the familiarity with the type of system architecture used (e.g., client- server, Internet JAVA applications) of the key people on the project.
	11		At the start of the project, the familiarity with and comprehension of the software development environment (e.g., compiler, code generator, CASE tools) of the keypeople on the project.
Range of Values			Reduced by Centering at Zero
Periodogram Smoothing	Spectral Window		Tukey-Hamming
	Window Span		5
	Weight Value	W(-2)	2.239
		W(-1)	2.240
		W(0)	2.240
		W(1)	2.240
		W(2)	2.239
Applying the model specifications from MOD_1			

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

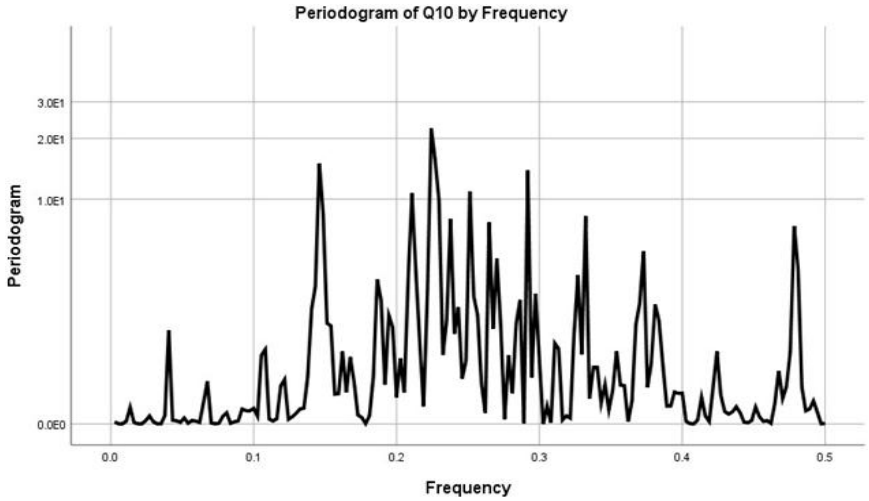
Please indicate the type of company



The business objectives for the project and the product were well defined at the start of the product itself.

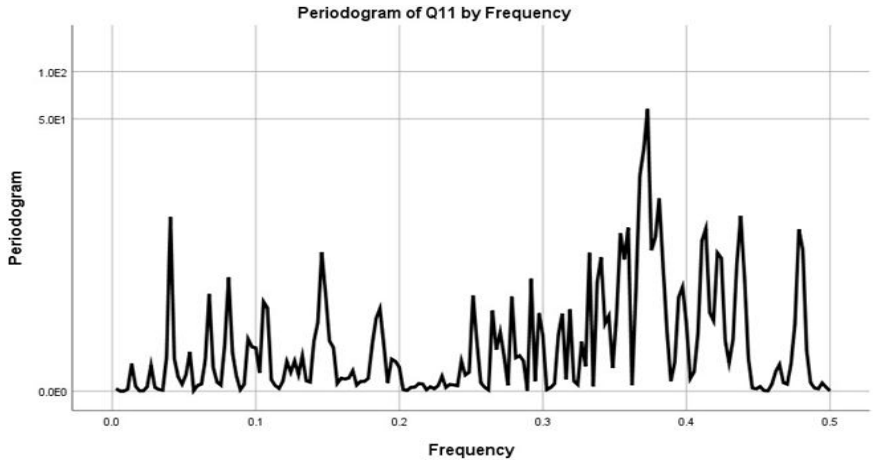


The business objectives for the project and product were documented at the start of the project.

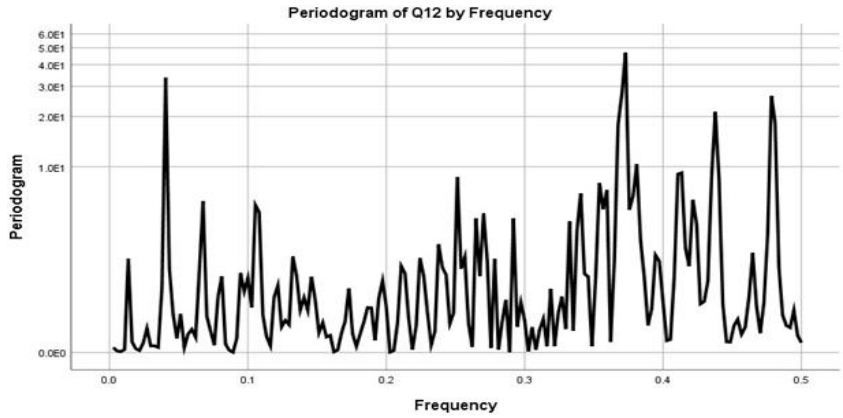


Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

The business objectives for the product and the project were understood by the development team at the start of the project.



The business objectives of the project and the product were understood by the customer at the start of the project.



CONCLUSION

The Anova test on variables and the Spectral analysis performed show a stronger relationship between the requirements engineering process and the planning phases in relation to defect prediction being used by IT SME's. However SME's need to adopt more techniques to make their projects more secure and reliable in terms of lead time reduction.

REFERENCES

1. Márquez, G., Astudillo, H., & Taramasco, C. (2020). Security in Telehealth Systems From a Software Engineering Viewpoint: A Systematic Mapping Study. *IEEE Access*, *8*, 10933-10950.

2. Eom, T., Hong, J. B., An, S., Park, J. S., & Kim, D. S. (2019). A Systematic Approach to Threat

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

Modeling and Security Analysis for Software Defined Networking. *Ieee Access*, *7*, 137432-137445.

3. Huang, Y., Bian, Y., Li, R., Zhao, J. L., & Shi, P. (2019). Smart contract security: A software lifecycle perspective. *IEEE Access*, *7*, 150184-150202.

4. Shan, C., Chen, B., Hu, C., Xue, J., & Li, N. (2014). Software defect prediction model based on LLE and SVM.

5. Moyo, S., & Mnkandla, E. (2020). A Novel Lightweight Solo Software Development Methodology With Optimum Security Practices. IEEE Access, 8, 33735-33747.

6. Siboni, S., Sachidananda, V., Meidan, Y., Bohadana, M., Mathov, Y., Bhairav, S., ... & Elovici, Y. (2019). Security testbed for internet-of-things devices. IEEE Transactions on Reliability, 68(1), 23-44. software behavior modeling based on extended finite state automata.

7. Reason, J. T., Carthey, J., & de Leval,

8. M. R. (2001). Diagnosing ―vulnerable system syndrome‖: an essential prerequisite to effective risk management. BMJ Quality & Safety, 10(suppl 2), ii21-ii25.

9. NIST Big Data Public Working Group. NIST Big Data Interoperability Framework: Volume 6, Reference Architecture. Available online: https://bigdatawg.nist.gov/_uploadfile s/NIST.SP.1500-6r1.pdf (accessed on 19 January 2019).

10. Apache Software Foundation. Apache Mahout: Scalable Machine Learning and Data Mining. Available online: https://mahout.apache.org (accessed on 19 January 2019).

11. Apache Software Foundation. Apache Spark. Available online: https://spark.apache.org (accessed on 19 January 2019).

12. Paradigm4. SciDB—The Array Database Management System. Available online:

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved

https://www.paradigm4.com/technolo gy/Open_Source (accessed on 19 January 2019).

13. Pareek, P. K., Nandikolmath, T. V., & Gowda, P. (2012). FMEA implementation in a foundry in bangalore to improve quality and reliability. Mechanical Engineering and Robotics Search, 1(2), 82-87.

Page 48-63 © MANTECH PUBLICATIONS 2022. All Rights Reserved