I.J. Mathematical Sciences and Computing, **2017, 1, 17-27**
Published Online January 2017 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijmsc.2017.01.02

***Available online at* http://www.mecs-press.net/ijmsc**

Fuzzy Expert System Based Test Cases Prioritization from UML State Machine Diagram using Risk Information

Wasiur Rhmann [a*], Vipin Saxena [b]

[*a] *Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Raebareli Road, Lucknow, India*
[b] *Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Raebareli Road, Lucknow, India*

Abstract

Regression testing is used to check that changes in the some functionality of the software to not affect its old behaviours. Test case prioritization is essential for reducing the cost of regression testing. In this paper a test cases prioritization model based on fuzzy logic is presented. State machine diagram is used to capture the behaviour of the system. Risk information is associated with the states. After change in the functionality of the system new state machine diagram is designed. This new state machine diagram is converted into Weighted Extended Finite State Machine (WEFSM). Weights are assigned to nodes and edges based on change and risk exposure. Risk exposure and change information of each test case is used as input to fuzzy model. Test cases are categorized in retestable, reusable and obsolete.

Index Terms: Regression testing, Test case, State machine diagram, Fuzzy logic.

1. Introduction

Regression testing is performed to check whether system after modification is working correctly according to its specification. Regression testing involves testing the features that could be affected by the modification in the software system. Modification or changes in some functionality of the system may introduce bugs in the software which may cause software failures. It is not feasible to retest software with all test cases due to time and resource constraints. Various researchers have proposed different techniques for regression test cases prioritization [1, 2, 3]. These techniques offer flexibility in selection of test cases to handle time and resources problems. Most of the regression testing techniques are code based. Code based regression testing suffers from complexity of large size software.

* Corresponding author.
E-mail address: wasiurrhmann786@gmail.com

Unified modeling language is standard modeling language for object oriented systems. It consists of different diagrams to represent the system at different level of abstraction. Information contained in UML diagram is used for testing of the software system. In the present work, UML state machine diagram is used to represent the functionality of the system. State diagram represents the different states in which objects goes during its lifetime [4].

Regression test case prioritization schedules the execution of test cases in such order which can increase the effectiveness of test cases. Quality of the software product or outcome of the project may be affected by risk. Risk is a probabilistic event which may occur or not. If risk occurs then there may be loss. Use of risk information will be beneficial for safety critical system and risk based testing advocates the allocation of larger part of testing time and resources for risky components [5].

In the present work risk exposure and change in system is considered for prioritization of test cases. Risk exposure and change information is used as input and output is prioritized test paths. These test paths are categorized in three categories namely reusable, retestable and obsolete test paths.

2. Related work

Dahiya et al. [6] used UML diagrams for selection of regression test cases. Authors considered change in semantic operation, change in control path and deletion or addition of the functionality. They created a tool that categorized test cases into reusable, retestable, obsolete and newly generated test cases. Malz et al. [7] presented an automated approach of test cases prioritization using fuzzy logic and software agents. Proposed approach increases the effectiveness of test and fault detection rate. Hettiarachchi et al. [8] used the software requirements as risk indicators and fuzzy expert system is used to estimate the requirement risks. Proposed risk estimation process is less subjective and results indicated the effectiveness of the proposed approach for fault detection. Sapna and Balakrishnan [9] proposed a technique of regression test cases prioritization using UML. Authors converted the control flow graph into Steiner tree and modified the nodes of graph based on changes. Minimized test cases are generated from the created Steiner tree. Stallbaum et al. [10] proposed a technique of automatically test case scenarios generation from UML activity diagram. Authors named their technique RiteDAP(Risk Based Test Case derivation and prioritization). Generated test cases are prioritized using fault information. For quantification of risk probability of failure of actions and estimated damage caused by failure is used. Risk information is used to exercise the code with risk prone area [11] by executing the test cases as early as possible. More important test cases are those which are more likely to cover defects associated with system’s risk.

3. Background

Fuzzy logic system [12] is three steps process which are as follows:

3.1. Fuzzification

Fuzzification is the process of converting crisp values into fuzzy input values. Membership functions are defined for each input and output variables.

Linguistic variables have values from natural language and are used for input and output. In our case Linguistics input variables are Risk and Change and output variable is Test path.

Membership function is used for evaluation of degree of truth. In the present work, triangular membership function is used.

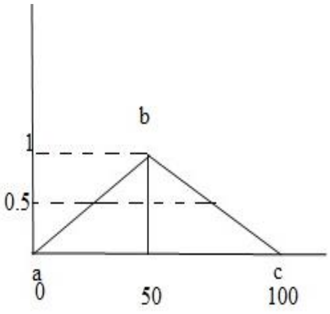

Fig.1. Triangular Membership Function

Triangular membership function is specified by three parameters (a, b, c) and defined as follows:

$$\text{Triangular membership function (x, a, b, c)}=\begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ \frac{c-x}{c-b} & b \leq x \leq c \\ 0 & c \leq x \end{cases} \quad (1)$$

Membership value of an element x is defined as $\mu_A(x)$ for conventional set A.

$\mu_A(x)= 1$ if x is element of A
and $\mu_A(x)= 0$ if x is not element of A

3.2. Rule base definition and evaluation

Once input and output variables are defined and membership functions are selected then rules are defined. Rules are written in IF <antecedent> THEN <conclusions> form. Results from these rules give fuzzy output values. A simple fuzzy rule is defined as follows:

IF m is A. **THEN** n is B.
Where m and n are linguistic variables and A, B are values of linguistic variables m, n respectively.

A takes value between 0 and 1. There may be more than one linguistic variable and value combined with the fuzzy logic operators or there can be more than one part of antecedent of a rule and all parts are calculated simultaneously using logical operators. Antecedent takes value between 0 and 1.

AND represents as intersection and its value is minimum of the membership function value. AND is represented by equation 2:

$$\mu_{A\cap B}(x)=\mathrm{Min}(\mu_{A,}\ \mu_B) \quad (2)$$

OR represents as union and its value is maximum of the membership function value. OR is represented by equation3:

$$\mu_{A\cup B}(x)=\mathrm{Max}(\mu_{A,}\ \mu_B) \quad (3)$$

3.3. Defuzzification

Process of conversion of fuzzy outputs to crisp outputs is called defuzzification. Various methods are used for defuzzification process. Defuzzification process is optional.

Fuzzy expert systems have been applied in various fields like risk assessment in construction [13], in medical field it is used for diagnosis of diseases [14-15], software effort prediction [16] and cost of software estimation [17].

4. Proposed methodology

Proposed technique of prioritized test paths generation is defined in following steps:

1. Draw state diagram of the modified system with risk information attached to the states. Risk information attached to different states of the system are quantified with risk exposure values and risk exposure values are calculated from the given formula;

Risk Exposure.= Probability of occurrence of risk(P) * Cost of damage of risk occurred(C); (4)

2. Convert the state machine diagram into weighted extended finite state machine (WEFSM) and generate different test paths from modified state diagram by traversing in depth first manner. WEFSM is drawn is such a way that there is a node corresponding state, decision node and start and end node of the state diagram;
3. Risk[18], change and Test paths are categorized as follows:

   Risk={Annoying, Hindering, catastrophic , Damaging}

   Increasing order of risk severity is from annoying to damaging. Damaging type of risk is of highest severity. Risk is categorized in no change in the system and change in the system.

   Change={NoChange, Change}

   Output test paths have three categories which are given as follows:

   Test paths={Obsolete test path, Reusable test path, Retestable test path}

   Obsolete test paths are least critical and there is no need to test these paths during regression testing, reusable test paths may be tested as they cover some risk and change information while retestable paths are necessary to test the system as they cover both risk and change components of the system.

4. For each generated test path calculate the maximum risk exposure value and maximum change value;
5. Maximum risk exposure value and change values are input to the fuzzy expert system which gives output test paths as obsolete, reusable and retestable.
6. Fuzzy rules are described as follows:

   If Risk is annoying **and** change is nochange
   Then obsolete test path;
   If Risk is damaging **and** change is change
   Then retestable test path;
   If Risk is damaging **and** change is nochange
   Then reusable test path;
   If Risk is catastrophic **and** change is change
   Then retesatble test path;

7. Retestable test paths are of highest priority then reusable test paths. Obsolete test paths are discarded.

5. Case Study

In the Fig. 2 a state diagram for ATM is presented and risk information is attached with different states. Risk categorization is based on ATM study of [19]. Here user ATM display welcome message and when user insert ATM card in machine then machine ask for Pin number. If Pin number is correct then machine displays main menu and user can select for cash withdrawal or can enquiry for balance. Risk of least severity will be when machine is not performing balance inquiry and it will come under the category of annoying risk. Risk of highest severity will be attached with the transaction process of cash withdrawal which will cause highest damage so it comes in damaging risk category.

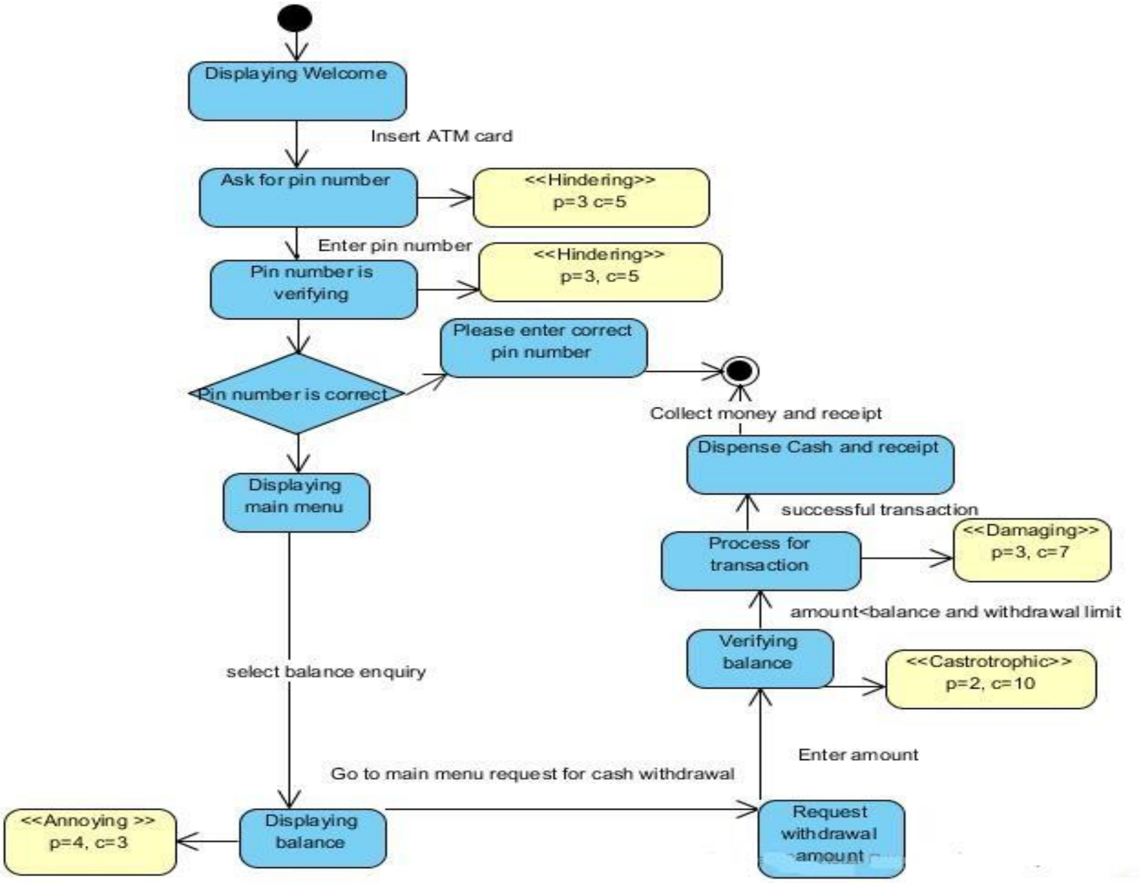

Fig.2. State Machine Diagram for ATM

In Fig. 3 new states are added after enhancement of ATM presented in Fig. 2. After verifying the pin number ATM machine ask for fingerprint. Then user enters his fingerprint mark on fingerprint scanner. This fingerprint is verifyied from adhaar number information. Aadhaar number is unique identy provided to citizens of India by government of India. It is based on bimetric data of citizen[20]. Aadhaar number is 12 digit number issued by government. It is linked with biometric information and demographic data of resident such as fingerprints, iris, photographs and these informations are stored in a central database. It can detect fake and duplicate identities online. Genuineness of aadhaar number can be verified by any service provider by using Aadhaar Verification Service(AVS) available on its website.

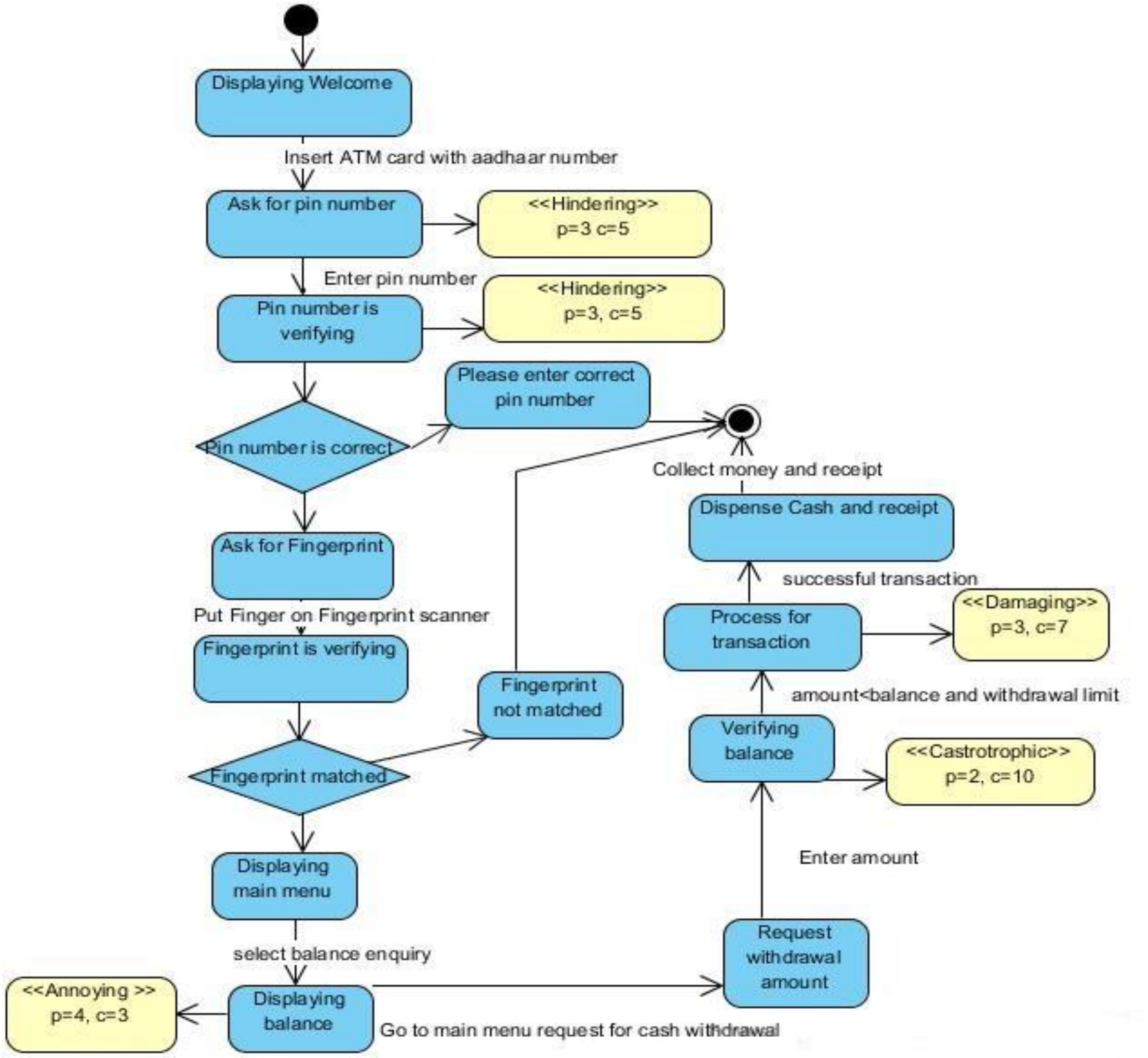

Fig.3. State machine diagram for ATM with Aadhaar number

Weighted Extended Finite state machine (WEFSM) is created from state machine diagram of Fig. 4. For each state, decision node and start and end node there is a node in WEFSM. Change node of WEFSM are represented as rectangle and weight of 30 is assigned each change node. Risk exposure values attached with the states are assigned to corresponding edges of WEFSM.

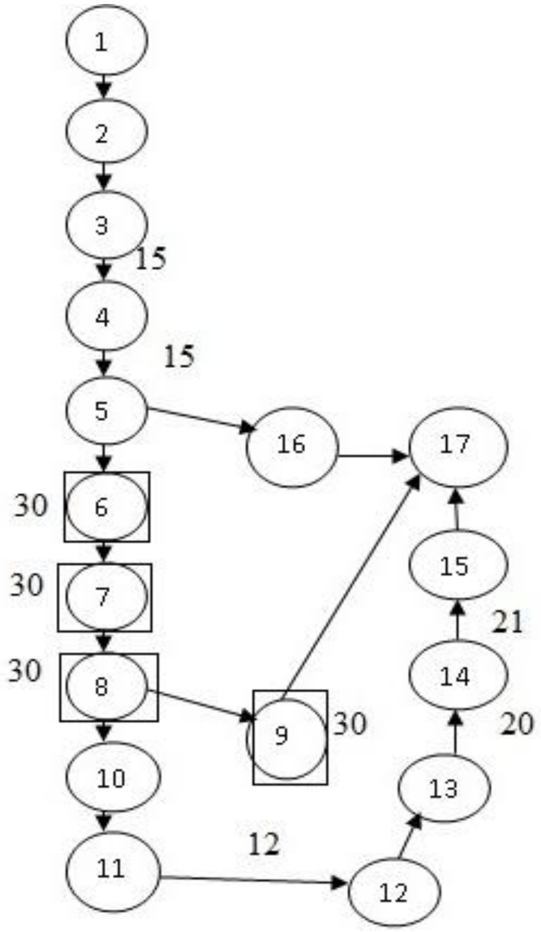

Fig.4. Weighted Extended Finite State Machine (WEFSM)

Test paths are derived by traversing the WEFSM in DFS manner. Risk Exposure and ChangeValue are maximum risk exposure and change value. In our case there are three test paths. Test paths, risk exposure and change values are calculated in Table 1.

Table. 1 Test Paths and their Risk Exposure and Change Value

Test Path	Risk Exposure	Change Value
1-2-3-4-5-16-17	15	0
1-2-3-4-5-6-7-8-9-17	15	30
1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-1714-15-17	21	30

Fig. 5, Fig. 6 and Fig. 7 represents the membership functions for risk, change and Test suite. Risk from 0 to 25 is categorized in Annoying, from 25 to 50 as hindering, from 50 to 75 as catastrophic and from 75 to 100 as damaging which is of highest severity.

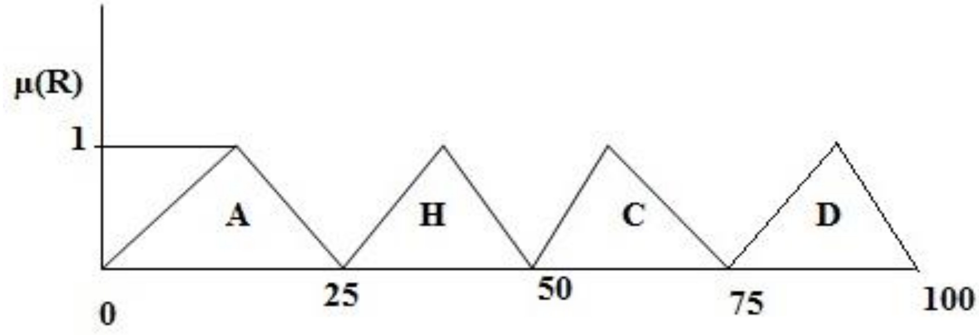

Fig.5. Membership function for Risk= {Annoying, Hindering, Catastrophic, Damaging}

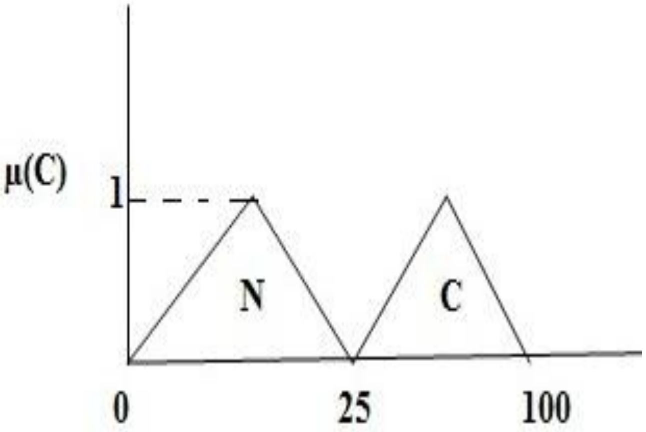

Fig.6. Membership function for Change= {No Change, Change}

Test paths from 0 to 25 values are categorized as obsolete, from 25 to 30 are as reusable and from 30 to 100 as retestable. Retestable test path are essential to exercise which covers highest change and risky components. After that Reusable test paths are needed for detection of faults and obsolete test paths are discarded as they cover no risk and change.

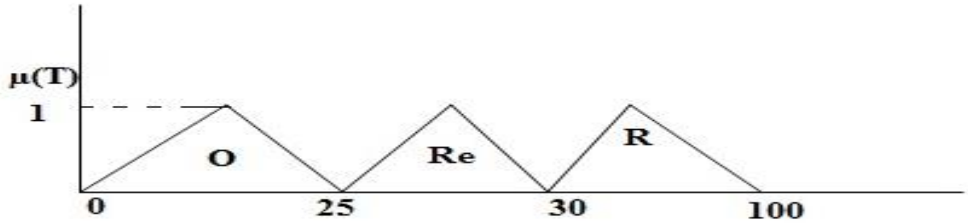

Fig.7. Membership function for Test paths={ Obsolete Test path, Reusable Test path, Retestable Test path}

Table 2. Conversion of Crisp Input of Risk and Change into Fuzzy Values

Risk	Change	μ(R)	μ(C)
15	0	0.4	0
15	30	0.4	0.153
21	30	0.725	0.153

Table 3 represents the test paths with their fuzzy values for risk information and change information. These values are AND and minimum is taken as Test path value by applying rules given in step6 of Section 4. Output corresponding paths is also given by applying step6.

Table 3. Conversion of Fuzzy Input Values into Fuzzy Output Values using Fuzzy Rules Defined in Step6

Test path	μ(R)	μ(C)	μ(T)	Output
TP1	0.4	0	0	Obsolete test path
TP2	0.4	0.153	0.153	Retestable test path
TP3	0.725	0.153	0.153	Retestable test path

6. Conclusions

Presented model considered the risk information and change so this can be suitable for regression testing of safety critical systems and financial system. This model minimizes the test cases by discarding the test cases for obsolete test paths. Test cases can be selected from reusable and retestable test paths. As Proposed regression testing technique considered risk information with change information it can be useful for system which have critical information like health care system, heart pacemaker, insulin pump etc. Identification of risky components from UML may help in better planning of project.

References

[1] Harrold MJ, Rosenblum D, Rothermal G and Weyuker E. Empirical Studies of a Prediction Model for Regression Test Selection, IEEE Trans. Software Eng. , 2011; 27(3): 248-263.

[2] Rothermal G, Harrold MJ. Analyzing Regression Test Selection Techniques, IEEE Trans. Software Eng., 1996; 22(8): 529-551.

[3] Rothermal G, Untech R, Chu C and Harrold MJ. Prioritizing Test Cases for Regression Testing, IEEE Transaction on Software Eng., 2001; 27(10): 929-948.

[4] Hamilton K and Miles R. Learning UML 2.0. O Reilly. 2006.

[5] Wenland MF, Karanz M.and Schieferdecker I. A Systematic Approach to Risk-Based Testing Using Risk-annoted Requiremts Models, International Conference on Software Engineering Advances, Lisbon, 2012; pp. 636-642.

[6] Dahiya S, Bhatia RK and Rattan D. Regression Test Selection using Class, Sequence and Activity Diagram. IET Software. 2016; 10**,** (3)**:** 72-80**.**

[7] Malz C, Jazdi N and Göhner P. Prioritization of Test Cases using Software Agents and Fuzzy Logic. IEEE Conference, Washington, USA, 2012; pp. 283-286.

[8] Hettiarachchi C, Do H, and Choi B. Risk Based Test Case Prioritization using Fuzzy Expert System**,** Information and Software Technology**,** 2016; 69: 1-15**.**

[9] Sapna PG and Balakrishnan A. An Approach of Generating Minimal Test Cases for Regression Testing**,** Procedia Computer Science, 2015; 47: 188-196.

[10] Stallbaum H, Mtzger A and Pohl K. An Automate Detection Technique for Risk Based Test Case Generation and Prioritization, Proceedings of the 3rd International Workshop on Automation of Software Test, May 2008, pp. 67-70.

[11] Yoon M, Lee E, Song M and Choi B. A Test Case Prioritization through Correlation of Requirement and Risk, Journal of Software Eng. Appl., 2012; 10(5): 823-835.

[12] Jamshidi M, Vadiee N and Ros TJ. Fuzzy Logic and Control, Prentice Hall, 1993.

[13] Carr V, Tah JHM. A Fuzzy Approach to Construction Project Risk Assessment and Analysis: Construction Project Management, Adv. Eng. Softw. 2001; 32(10): 847-857.

[14] Adeli A, Neshat M. A Fuzzy Expert System for Heart Disease Diagnosis, Proceedings of International Conference, 2010; 1: 1-7.

[15] Kadhim MA, Alam MA, and Kaur H. Design and Implementation of Fuzzy Expert System for Back pain Diagnosis, International Journal of Innov. Technol. Creat Eng, 2011; 1: 16-22.

[16] Ahmed MA, Ssliu MO and Ghamdi JA. Adaptive Fuzzy Logic Based Framework for Software Development Effort Prediction, Inf. Software. Technol., 2005; 47(1): 31-48.

[17] Kazemifard A, Zaeri A, Ghasem-Aghaee N and Nematbaksh MA, Mardudkhi F. Fuzzy Emotional COCOMOII Software Cost Estimation using Multiagent Systems, Appl. Soft. Comp., 2011; 11(2): 2260-2270.

[18] Schaefer H, Risk Based Testing, strategies for prioritizing tests against deadlines, Software Test Consulting://home.c2i.net/shaefer/testing.html, 2005.

[19] Jogersen PC. Software Testing a Craftsman’s Approach, 4th edition, CRC Press. 2014.

[20] https://en.wikipedia.org/wiki/Unique_Identification_Authority_of_India.

Authors’ Profiles

Wasiur Rhmann, research scholar in Department of Computer Science, Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow, India. He has completed B. Sc. (Hons.) in Physics and M. C. A. from Aligarh Muslim University India. His research interest is Software testing, UML modeling, Software Engineering.

Vipin Saxena, Professor in Department of Computer science, Babasaheb Bhimrao Ambedkar University, Lucknow India. He is also Dean of School of Information Sciences and Director of UIET Engineering college of University. He is M. Sc. and Ph. D. from I. I. T. Roorkee. His research interest includes Software engineering, UML modeling, Distributed networking. He has about 100 international and national journals of repute.

How to cite this paper: Wasiur Rhmann, Vipin Saxena,"Fuzzy Expert System Based Test Cases Prioritization from UML State Machine Diagram using Risk Information", International Journal of Mathematical Sciences and Computing(IJMSC), Vol.3, No.1, pp.17-27, 2017.DOI: 10.5815/ijmsc.2017.01.02