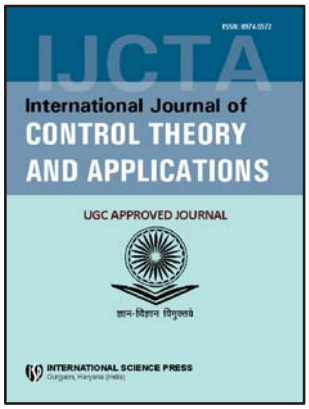

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 33 • 2017

Failure Analysis Based on Relaibility Testing using Genetic Algorithm

R. Bhargavi[a] and Mattal Sneha Abraham[a]

[a]*Computer Science Department, VIT University, Chennai, India*

E-mail: bhargavi.r@vit.ac.in, mattalsneha.abraham2015@vit.ac.in, snehaabraham146@gmail.com

Abstract: Silicon wafer is an essential component in the Integrated Circuit chips. Wafers are usually tested for identification of failures. To identify the failures, there is a need for testing the wafer. A reliable failure analysis of the wafer improves the reliability of the ICs. This paper proposes a reliable testing platform (RELIS) for automated testing, failure analysis and management. RELIS provides a clean user interface for providing the user inputs for development of software that is used for testing. This paper also proposes a genetic algorithm based classification technique for classifying the faults present and for predicting the failures. This Algorithm with fitness function is used for generation of various mining classification rules. The proposed algorithm is being tested on failure analysis dataset. The proposed reliability testing platform is implemented using Extended JavaScript and SQL Developer.

Keywords: *ARIMA Model; Classification; Data Mining;Failure analysis; Genetic Algorithm;Reliability testing; Regression; RELIS.*

1. INTRODUCTION

A wafer, also called a slice or substrate, it is a thin slice of semiconductor material, such as a crystalline silicon, used in electronics for the fabrication of integrated circuits. Most of the modern cars use several ICs in engines of the vehicles. These ICS undergo lot of testing before they are used in the cars. Testing is of two types namely Functional testing, and Regressiontesting. Functional testing is a testing technique that is used to test the functionality of software. For example,functional testing is used for automated driving. Regression testing is done to check the IC's workingunder various abnormal conditions. Regression testing is used for checking whether the chip is subjected to any damage when engine is over heated, whether it will work in high pressure, and any other conditions which do not occur in the normal scenarios. The objective of this paper is to propose a reliable testing platform (RELIS) for automated testing, failure analysis and management.It is good for identifying failures at the initial stages rather than analyzing at the last stage. Here by using Genetic Algorithm we are going to analyze failures and classification of faults based on that. Failure Analysis is used to collect and analyze the data to determine in case of failure occurred. Reliability Testing (RELIS) is used to check the reliability for development of software for wafer testing.They can be used for generation of wafer which can be used for chip development. It is used for highly modeled vehicles. Genetic Algorithm is a method for selection based on a natural selection process. This Algorithm uses flexible technique that can be strongly used to solve

various complex tasks and other optimization problems. This Algorithm is tested using failure analysis dataset. Analysis is used rather than regression models because it is time dependent as well as it is used to analyze based on increasing or decreasing trend. It is software used for testing the reliability of the system that can be used for testing of wafer. Wafer will undergo several test procedures and is used for chip design. So it can be used for highly modeled vehicles. It can be subjected to various stages such as Trial, Sample, Test, Steps. Reliability of the system as well as lot management is tested. In Lot Management, the criteria of the wafer are being tested. Failure can occur in the wafer due to the conditions present in failure mode. The user interface consists of the test, lot number present which test it is linked to, what are the standard conditions, whether it is confirmed failure or what, types of defect such as ageing defect or time zero defect, failure mode and mech code and such all other conditions. Then it will analyze the failure and it will help in predicting various failures. Failure Mode means the causes that can lead to failure. The causes of failure are various electrical defect, wrong value characterization, poor knowledge of application, wrong marking, failed due to incomplete metal i.e. high current, electrolyte inverted, battery failure or due to bad wire bond stitch failure, failed functional test due to silicon residue, particle induced gate oxide breakdown, marquage efface, baking defect, failure in high ibat current, wire corrosion, metallization problems, crack after chemical opening, ficcurepuse, attack of the passivation, over corrosion, catering, crack, silicon melting, solder crack, leakage over top side, broken wire over the ball, metal corrosion, leakage, solder removal, defect in bvdss, silicon cleavage, cracked die, solder non uniform thickness, humidity in cavity, resin delimination, current spike, thermal mismatch, crack split die, break due to thermal mismatch, broken stitch bonding process, wrong setting in die fab, thermal fatigue, cross section and various other causes of failure are present in the application. Reliability testing is done based on failures present in application on various conditions. It is used to manage reliability tests in reliability laboratories. The data to be processed mainly concern reliability trials that can be performed in several cases like: periodic tests performed to check production reliability levels, qualification of a process i.e. site of a customer, systematic reliability tests performed on production lots of a customer. Quality Acceptance environment (QA) is based on training and user acceptance before a new release. Depending on failure in the system clusters are being generated after prediction based on values.Data mining technique is used toanalyze the data of failure. Thus we can analyze results based on particular value of the historic data. It is a task based on data mining and a technique used for statistical analysis. It can be used as a common technique for pattern recognition, information retrieval, data compression as well as other techniques.

Here we are testing the failures present in the software for development of wafer. So that quality of the software is being taken into consideration. Thus trials are generated before testing of the system. Testing is done in the software before it goes to the quality environment.

TRIAL : Trial search is based on the following 3 criteri.

Trial criteria: Criteria related to Trial definition/information and lot of information. Product/Manufacturing criteria: criteria related to Product and Manufacturing information.

Test criteria: Criteria related to Tests information.

TESTING : Validate a trial to failure quantities. Display evaluation at trial and samples level. Result collection setup at trial level and samples level is taken place. New data is saved based on user's requirement.

Features of the system:

1. **Definition of the Trials to be Performed :** Create, Update and Copy Trials are the three functions that are used in definition of trials.
2. **Trial Scheduling:** It is used to manage and calculate the Scheduled start date and end date. It is based on theoretical steps duration. The function is used to change the scheduling of a trial in one change.
3. **Monitoring of the Tests in Progress:** It is used to start-up the tests and update the current status of the tests that in progress. It can collect the test results step by step and failure analysis results.

4. **Documents Attachment :** External documents are saved in trial and can be displayed through standard tools. Documents are classified by different category and can be attached at any level of trial.
5. **Results Validation (Confirmation of the Observed Failures):** Validation is done of test and trial. Evaluation is done for reopening a test or trial.
6. **Results Evaluation:** The global test results can be evaluated by sample evaluation by operators and final evaluation can be done at trial level by laboratory's responsible.
7. **Lot Management:** Used to display product information from any process in trial management and product composition information that are transferred from PRIS.
8. **Definition and Maintenance of 'Parametric Data' (Parameter Management Section):** Those data include data from the attributes for the device that is under test, downloaded from PRIS database. Lots are created for tests linked to Traceability systems (TIS).
9. **Reporting:** To view trials and it has the possibility of users to define their own reports using QSRS reporting tool.
10. **Application Support :** Remedy Solutions are provided.

2. LITERATURE SURVEY

Various studies based on failure analysis have been done yet now. But in this application, we are going to analyze data based on the failureand predict the data with the help of graph. This study can help inanalyzing the failure data and it also helps users to classify data based on particular faults present. This can help users to analyze where does most failure occur and be cautious or develop wafer in such a way that it avoids this type of failure. Users can find solution for a fault that is common if it is clustered so that it is easy for a user to analyze the fault.

"Clustering and Classification of Maintenance Logs using Text Data" (Brett Edwards, Michael Zatorsky&RichiNayak, 2008) This paper is used to analyze the data based on 12 years of data and it is used to classify the data based on supervised and unsupervised learning. [1]

"A New Approach For time Series Forecasting Based on Genetic Algorithm" (Mahesh S. Khadka, Benjamin Popp, K. M. George & N. Park, 2010) is proposing a new fusion approach is based on Concordance and Genetic Algorithm. [2]

"Drop Impact Reliability Testing Lead free Chip Scale Packages " (San Luis Obispo,) is defining Chip Scale Packages is used for handheld devices such as mobile phones, PDAs etc. and how failure present in this device.[3]

"Analysis and detection of timing failures in an experimental test chip" (Piero Franco, Siyad Ma, Jonathan Chang, Yi-Chin Chu, Sanjay Wattal, E. J. McCluskey Robert L. Stokes, and William D. Farwell,)is defining clock speeds and clocking modes that can be used to characterize defective parts. [4]

Andrzej J. Strojwas and Stephen W. Director in this paper "A pattern recognition based method for IC failure analysis" is defining set of algorithms that can be used for various IC failure modes and how it can be used for various processes. [5]

Michael **Stevenson**in this "Journal of Failure Analysis and Prevention" is providing practical definitions regarding the causes of failure as well as how to eliminate future failures. [6]

Zhaofeng Wang and Sr. FA Engineer in this paper "Failure Mechanisms and Failure Analysis Process Flow" has described about various method of analyzing the failures and how the failure workflow process. [7]

Jim Glancey "Failure Analysis Methods" has analyzed the reasons/causes of failure and why failure occurs in the system. [8]

M. Steven Ferrier "The Failure Analysis Process" gives a proper flow showing path to each actual prior cause through observed secondary effects, and elimination of incorrect possible prior causes. [9]

Rosalynn MacGrego in this paper "Assess Hazards with Process Flow Failure Modes Analysis" is used to identify potentially hazardous scenarios. [10]

"A Genetic Algorithm for Classification" (Raul Robo, Stefan Holman, 2009) is used for defining how genetic algorithm is applied for classification of data. [11]

3. PROPOSED METHODOLOGY

Genetic Algorithm uses data classification technique over here. The dataset is divide into training and test data. Chromosome represents classification rule, gene represents the value of an attribute and rule (chromosome) is coded by a set of array.

The fitness of each function is calculated by:

$$\text{Fitness} = w1 * PA + w2 * CPH + w3 * S$$

Where PA denotes predictive accuracy where

$$PA = tp/(tp + fp),$$

CPH is comprehensibility $CPH = MNC - ANC,$

MNC is maximum number of conditions,

ANC is actual number of conditions,

S is sensitivity, $S = tp/(tp + fn)$

Architecture

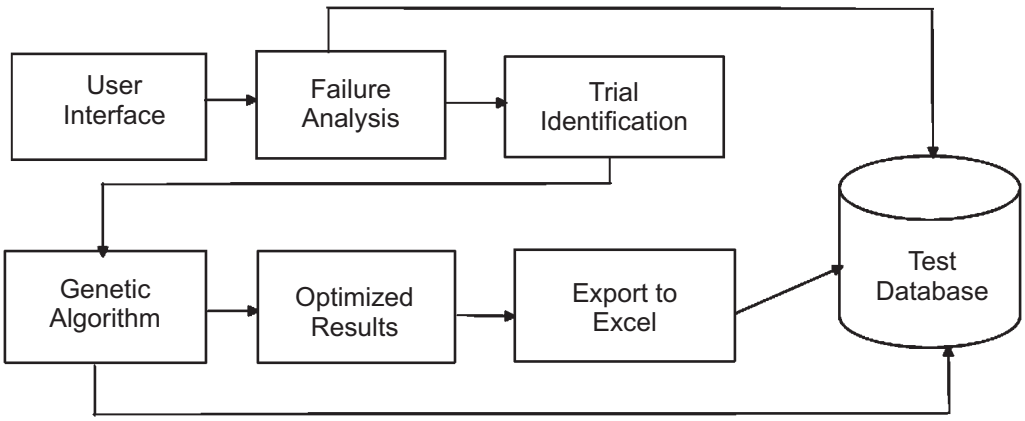

Figure 1: Architecture Diagram

The input to the model is failures in the reliability system and output is the classification of failure data in a particular group. The working of module is as below:

1. The first page is the initial user interface related to the reliability system.
2. The failures are being analyzed and they are being stored in database. And after complete failure analysis this is stored as excel document for user convenience.
3. Trial Identification page gives results based on the user interface.

4. The population module generates chromosomes for initial population of genetic algorithm. The chromosome is represented in the following way where wti= weight of the ith similarity measure. We have used a real-value chromosome because it is more natural representation for our problem and it decreases dramatically the number of genes required to specify a design, thus making the solution space easier to search.

5. Fitness module is used to find the fitness of the solution. Fitness function is a performance measure that is used to evaluate how good each solution is. Given a chromosome, the fitness function must return a numerical value that represents the chromosomes utility. Previous work has been done in GA considering fitness functions based on recall and precision only. However it has been observed that a better fitness function can be obtained if we also consider the order of the retrieval of documents. Order-Based Fitness function takes into account the number of relevant and irrelevant documents along with the order of their appearance.

6. Genetic Algorithm module applies the standard GA functions (selection, crossover, mutation) to generate new population from the old population, which is discussed as follows: Selection: Selection embodies the principle of survival of the fittest. Chromosomes having higher fitness are selected for crossover. The roulette wheel reproduction process was used to select individuals. Crossover: Crossover is the genetic operator that combines two chromosomes together to form new chromosome. Mutation: -We need to add a little bit randomness into our populations' genetics otherwise every combination of solutions we can create would be in our initial population.

7. Optimized results are stored in excel for user relevancy. This is stored in database for ease of users.

3.1. Fitness Functions for classification

Genetic Algorithm is used for analyzing the failure and predicting the class of the failures based on the data collected. To generate a fitness functions, we require rules such as If X then Y, so X is the antecedent of the rule and conjunction of the conditions is formed whereas Y is the consequent of the rule.

The factors used for classifying fitness functions are as follows:

True positive (*tp*) : If actual class is Y, predicted class too is Y.

False positive (*fp*) : If actual class is Y, predicted class is not Y.

True negative (*tn*) : If actual class is not Y, predicted class too is not Y.

False negative (*fn*) : If actual class is not Y, predicted class is Y.

Info-Gain (Ai)is information gain by attribute *i* and |dom G| is domain cardinality of class attribute G.

$$\text{Fitness} = w_1 x\ (\text{CF}\ x\ \text{Comp}) + w_2 x\ \text{Simp}$$

The first parameter denotes prediction accuracy and second parameter denotes rules comprehensibility.

$$\text{CF} = tp\ /\ tp + fp \text{ is confidence factor and}$$

$$\text{Comp} = tp\ /\ (tp + fn) \text{ is the completeness factor}$$

Simp represents the rule's simplicity and w_1 & w_2 are weights defined by the user.

Fitness function is determined by: $\text{Fitness} = tp\ /\ (tp + \text{A}.\ fn) * tn\ /\ (tn + \text{B}.\ fp)$

Where A & B are parameters that are defined by users that are determined according to user's request.

$$\text{Fitness} = w_1 x(|\text{A}\ \&\ \text{C}| - 1/2)\ /\ |\text{A}| + w_2 x\ \text{CM}$$

Here |A & C| determines he antecedent and consequent rule and |A| determines the number of instances that satisfy the antecedent.

Comprehensibility is determined by using the formula:

$$CM = (L - n) / (L - 1)$$

L is the number of conditions of the rule and n is the length of the rule.

Fitness function not only determines prediction accuracy and clarity of the rulebut it also determines interestingness of the rule that is presented in the below relation.

$$\text{Fitness} = \frac{w_1 C(R) + w_2 \text{ Pr } ed \text{ ACC} + w_3 \text{ RI}nt}{w_1 + w_2 + w_3}$$

Here C(R) denotes comprehensibility, Pr *ed* ACC denotes predictive accuracy, RInt denotes interestingness of the rule and M is the maximum no of conditions in the rule.

$$C(R) = M - \text{No of conditions}(R).$$

Pr *ed* ACC is calculated by $\text{Pr } ed \text{ ACC} = \frac{|A \ \& C|}{|A|}$

Where A, B & C are parameters defined by

$$\text{RI}nt = \frac{\sum_{i=1}^{n-1} \text{Info Gain } (A_i)}{\log_2(|\text{dom}(G)|)}$$

3.2. Algorithm

1. Start with initial population *i.e* failures of the data.
2. While termination condition not satisfied DO
3. BEGIN
4. Compute fitness value based on each polynomial.
5. Select individuals to reproduce based on their fitness.
6. Create offspring by crossover mutation.
7. Using mutation create some individuals.
8. Replace population by using new generation of individuals.
9. END

3.3. Data used for Failure Analysis

T_FAILURE_ANALYSISAttributes used are:

Trial Code: This is used to evaluate trial based on a particular test.

Sample Code: This means how many samples are passed.

Step Count: It calculates the no of steps required for failure in this application.

Failure_Mode_Code : A particular code is assigned to every failure steps.

Failure_Mech_Code : It represents the code related to wafer.

Failure_Nb : It is not mandatory for wafer generation.

Failure Comment: If failure occurs we can give a comment based on particular failure.

Physical_Code_1: It is collected list from user's database.

Physical_Code_2 : It is collected list from user's database.

Physical_Code_3 : It is collected list from user's database.

Ewip_analysis_req_id: This is Analysis ID in EWIP Tool.

Datalog_Upload: Here we going to create upload document for failure testing.

Reject_quantity: How much quantity of data will be rejected?

Go test: This is a test available if required only.

Ageing defect : This is a type of defect present in wafer related to the material's age.

Confirmed failure : It is used to check whether it is confirmed or unconfirmed failure.

Part_Identification_No : Part No is given in this part for wafer.

Time_Zero_Defect : This is another type of defect present in wafer.

Material: The material which is used for wafer generation is being defined over here.

Created by : This has been created by which extension is being defined.

Created_date : Creation date is defined over here.

Modified_by: This has been modified by which extension.

Modified date: Modified date is defined over here.

Test_id: Id of every test is being defined over here.

Macro_Failure_Mode: This is used for macro failure mode classification.

Macro_Failure_Mode_Detail : This is used for retrieving results from macro failure mode detail classification.

4. RESULTS

Here in this paper we are trying to analyze the values based on Genetic Algorithm. So by this algorithm, we can decrease the risk of faults in this application as well as increase accuracy and efficiency of the system. However the predicted value does not have a lot of difference. But we will take into account data of past six months and with that we can plot data for the next months. So that failure is decreased efficiently. This means if actual values of failure increases or decreases compared to the previous month data and if predicted values also increases.

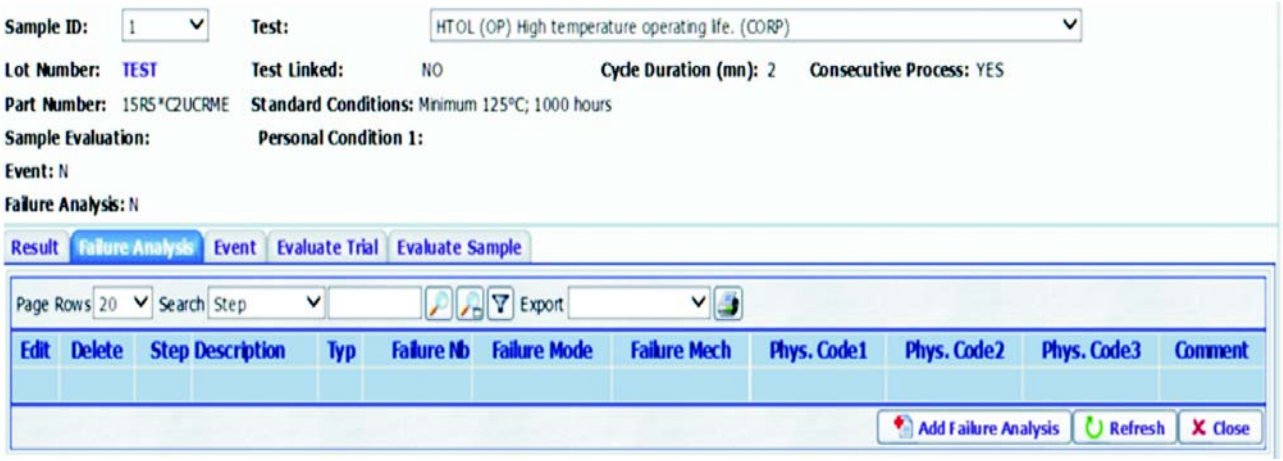

Figure 1: Failure Analysis Screen

Fig 1 is used to depict Failure Analysis Screen. This failure analysis is used for user interface. UI is used to keep track of failures that may occur in reliability testing. Here in this failure analysis screen, on click of Add Failure Analysis all minute information is being obtained. These are the parameters that will decide whether failure occurs or not. If failure occurs then in what manner and in what amount.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	TRIAL_C	SAMPLE_	STEP_C	FAILURE_	FAILURE_	FAILURE_	FAILURE_	PHYSICAL	PHYSICAL	PHYSICAL_ CODE_3	CREATED	CREATED_DATE	MODIFIED_BY	MODIFIED_DATES
2	CSBL9819	1	070	DACJ		1	en ices>> , bonne apres etuvage(150¿c,8h).				EXT191	08-APR-08	EXT191	08-APR-08
3	CSBL9819	3	090	DACJ		1	en bvdss< : bonne apres etuvage(150¿c,8h).				EXT191	08-APR-08	EXT191	08-APR-08
4	CSBL9819	1	070	DACJ	1HAC	1	en isolement< : crack/ceramic				EXT191	08-APR-08	EXT191	08-APR-08
5	CSBL9820	1	070	DAAE	1HAB	1	Presence des ebrechures.				EXT191	08-APR-08	EXT191	08-APR-08
6	CSBL9820	2	090	DACJ	1HDB	1	Restauration apr¿s etuvage(150¿c,8h).				EXT191	08-APR-08	EXT191	08-APR-08
7	CCAM944	1	010	DA01	1AAA	2	Various electrical defect				EXT191	08-APR-08	EXT191	08-APR-08
8	CCAM960	1	050	DA01		1	Various electrical defect				EXT191	08-APR-08	EXT191	08-APR-08
9	CCAM961	1	020	DA01	1AAA	1	Various electrical defect				EXT191	08-APR-08	EXT191	08-APR-08
10	CCAM964	1	010	EBAB	1CCB	1	Wrong Value (Boards)/Insufficient characterization				EXT191	08-APR-08	EXT191	08-APR-08
11	CCAM970	1	010	EBAA		1	Electrolytic Inverted (Boards)/				EXT191	08-APR-08	EXT191	08-APR-08
12	CCAM970	1	010	DA01	1AAA	1	Various electrical defect/Poor knowledge of application				EXT191	08-APR-08	EXT191	08-APR-08
13	CCAM970	1	010	DB0A	1AAA	1	Vcc min.				EXT191	08-APR-08	EXT191	08-APR-08
14	CCAM973	1	110	ECBA	1GBB	1	Mech. between Components (Boar/wrong marking				EXT191	08-APR-08	EXT191	08-APR-08
15	CCAM974	1	040	DDAA	1BAA	3					EXT191	08-APR-08	EXT191	08-APR-08
16	CRLN0023	1	110	DDFC	1EGA	1	Failed due to incomplete metal1 etch causing high Ibat current.				EXT191	08-APR-08	EXT191	08-APR-08
17	CRLN0023	2	110	DA01	1FDA	1	Failed continuity on the battery + end tab due to bad wire bond stitch weld.				EXT191	08-APR-08	EXT191	08-APR-08
18	CRLN0026	1	050	DAEA	3FAA	1	Neg. bat lead detached from battery.Corrective actions implemented from Rayovac				EXT191	08-APR-08	EXT191	08-APR-08
19	CRLN0026	2	110	DAEA	1FAA	1	Unit over welded. First samples from weld conversion. Corrections implemented.				EXT191	08-APR-08	EXT191	08-APR-08
20	CRLN0026	2	170	DAEA	1FAA	2	Units over welded.First samples from weld conversion.Corrections implemented.				EXT191	08-APR-08	EXT191	08-APR-08
21	CRLN0026	2	210	DAEA	1FAA	3	Unit over welded.First samples from weld conversion.Corrections implemented.				EXT191	08-APR-08	EXT191	08-APR-08
22	CRLN0026	1	110	DACA	1EAA	2	Failed functional test due to silicide residue along a word line gate.				EXT191	08-APR-08	EXT191	08-APR-08
23	CRLN0026	1	190	DACA	1EAA	1	Fail inp Ikg pin1,due to random metal etch def. which was not caused by new proc				EXT191	08-APR-08	EXT191	08-APR-08

Figure 2: Dataset for failure

Fig 2 is used as failure data. It is used to analyze the attributes of failure. These are the terms which are useful for users for analyzing the failures. Algorithm is started with a set of solutions called population. Genetic Algorithm module applies the standard GA functions (selection, crossover, mutation) to generate new population from the old population, which is discussed as follows. Selection: – Selection gives the principle of 'survival of the fittest'. Chromosomes having higher fitness value are selected for crossover. The roulette wheel reproduction process isused for individual selection. Crossover: – Crossover is the genetic operator that takes together two chromosomes together to form new chromosome.Mutation: -We need to add a little bit randomness into our populations genetics otherwise every combination of solutions we can create would be in our initial population.

5. CONCLUSION

Data-Mining is used to extract knowledge and information from the database. Classification is a data mining technique that helps in generating training dataset and model is being generated based on the data which too helps in prediction. Genetic Algorithm is a search technique that is used to generate optimized results and database is used as a large search space. This paper proposed a reliable testing platform (RELIS) for automated testing, failure analysis and management. RELIS is usedanalyzing the various causes failures andrepresents the data as a graph. This improves the interpretability and understanding of the data. Also, the data queried by the user using different queries is downloadable. This helps the users to analyze the failure data offline. This paper is used to discover various classification rules and genetic algorithm is used to generate optimized result. Genetic Algorithm and fitness function were tested and implemented on failure analysis dataset. Fitness function generated gives good prediction accuracy and comprehensibility.

REFERENCES

[1] Mining Brett Edwards Michael ZatorskyRichiNayak"Clustering and Classification of Maintenance Logs using Text Data"(2008).

[2] Mahesh S. Khadka, Benjamin Popp, K. M. George and N. Park "A New Approach For time Series Forecasting Based on Genetic Algorithm"

[3] San Luis Obispo "Drop Impact Reliability Testing Lead free Chip Scale Packages" (2008)

[4] Piero Franco, Siyad Ma, Jonathan Chang, Yi-Chin Chu, Sanjay Wattal, E. J. McCluskey Robert L. Stokes, William D. Farwell "Analysis and detection of timing failures in an experimental test chip"

[5] Andrzej J. Strojwas, Stephen W. Director "A pattern recognition based method for IC failure analysis" (1982)

[6] Based on "Relis Version 2.0.0".

[7] Zhaofeng Wang, Sr FA Engineer "Failure Mechanisms and Failure Analysis Process Flow"

[8] Wikipedia.com

[9] Jim Glancey "Failure Analysis Methods" (2006)

[10] Michael **Stevenson** "Journal of Failure Analysis and Prevention"

[11] M. Steven Ferrier "The Failure Analysis Process"

[12] Rosalynn MacGrego "Assess Hazards with Process Flow Failure Modes Analysis"

[13] Raul Robo, Stefan Holman "A Genetic Algorithm for Classification" (2009)