International Journal of Research in Computer Science
eISSN 2249-8265 Volume 2 Issue 2 (2011) pp. 7-14

www.ijorcs.org

UNIFICATION OF RANDOMIZED ANOMALY IN DECEPTION DETECTION USING FUZZY LOGIC UNDER UNCERTAINTY

S.Rajkumar[1], V.Narayani[2], Dr.S.P.Victor[3]

[1]*Research Scholar/Bharathiar University, Asst.Prof & Head / CSE, NIET,Coimbatore, India.*
[2]*Research Scholar, Dept. Of Computer Science, St.Xavier's College, Tirunelveli, India.*
[3]*Associate Professor & Head, Dept. Of Computer Science, St.Xavier's College, Tirunelveli, India.*

Abstract: *In the recent era of computer electronic communication we are currently facing the critical impact of Deception which plays its vital role in the mode of affecting efficient information sharing system. Identifying Deception in any mode of communication is a tedious process without using the proper tool for detecting those vulnerabilities. This paper deals with the efficient tools of Deception detection in which combined application implementation is our main focus rather than with its individuality. We propose a research model which comprises Fuzzy logic, Uncertainty and Randomization. This paper deals with an experiment which implements the scenario of mixture application with its revealed results. We also discuss the combined approach rather than with its individual performance.*

Keywords: *Deception, Detection, Uncertainty, Fuzzy logic, Randomness*

I. INTRODUCTION

Detection of Deception is useful for managers, employers, and for anyone to use in everyday situations where telling the truth from a lie can help prevent you from being a victim of fraud/scams and other deceptions [1].

A. Identifying the Deception

Deception detection between relational partners is extremely difficult, unless a partner tells a blatant or obvious lie or contradicts something the other partner knows to be true [5].

B. Fuzzy logic

Fuzzy logic is the part of artificial intelligence or machine learning which interprets a human's actions. Computers can interpret only true or false values but a human being can reason the degree of truth or degree of falseness. Fuzzy models interpret the human actions and are also called intelligent systems [7]. Fuzzification is the process of changing a real scalar value into a fuzzy value. This is achieved with the different types of fuzzifiers. Fuzzification of a real-valued variable is done with intuition, experience and analysis of the set of rules and conditions associated with the input data variables. There is no fixed set of procedures for the fuzzification [6].

C. Uncertainty

Uncertainty must be taken in a sense radically distinct from the familiar notion of risk, from which it has never been properly separated. Although the terms are used in various ways among the general public, many specialists in decision theory, statistics and other quantitative fields have defined uncertainty, risk, and their measurement as follows:

1. Uncertainty: A state of having limited knowledge where it is impossible to exactly describe existing state or future outcome, more than one possible outcome.
2. Measurement of Uncertainty: A set of possible states or outcomes where probabilities are assigned to each possible state or outcome.
3. Risk: A state of uncertainty where some possible outcomes have an undesired effect or significant loss.
4. Measurement of Risk: A set of measured uncertainties where some possible outcomes are losses, and the magnitudes of those losses variables [3].

D. Randomness:

The Dictionary of Oxford defines 'random' as "Having no definite aim or purpose; not sent or guided in a particular direction; made, done, occurring, etc., without method or conscious choice; haphazard." This concept of randomness suggests a non-order or non-coherence in a sequence of steps or symbols, such that there is no intelligible pattern or combination [8].

II. PROPOSED RESEARCH MODEL

The following figures show the basic and its expanded form for the proposed model.

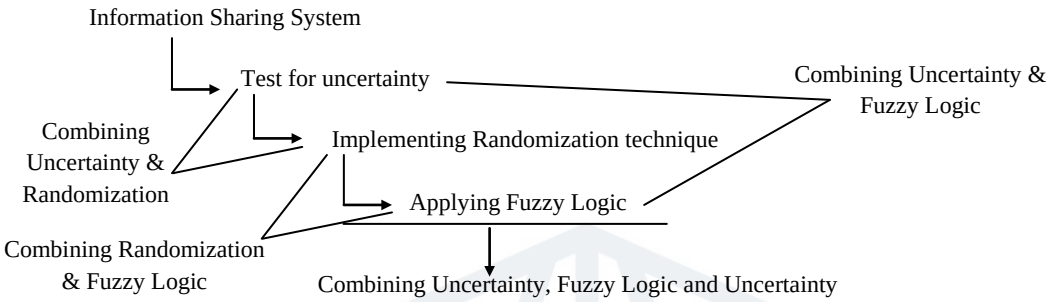


Figure1: Basic Proposed Model

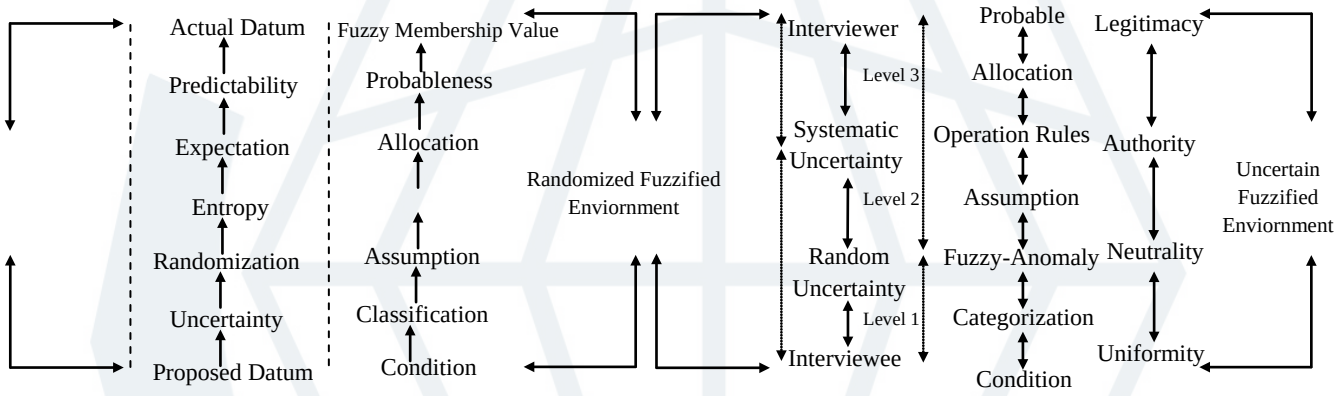


Figure2: Expanded Proposed Model

III. RESEARCH METHODOLOGY

A. Fuzzified Anomalies for Our Proposed Research Model:

The interception of fuzzified anomaly in the field of Recruiters selection process can be analyzed as,

i. Specify the range of conditions

$$0 \leq C_{Ans}(x) = \mu_t(x) \leq 1$$

Candidate Answer at the time‘t’ holds the membership function.

ii. Classification and categorization

Table I: Membership value assignments

Factor-X	Membership value μ_t (x)
Fully knowledged*	0.900 to 1.000
Maximized knowledge	0.800 to 0.899
Desired knowledge	0.700 to 0.799
Sufficient knowledge	0.600 to 0.699
Average knowledge	0.500 to 0.599
Partial knowledge	0.400 to 0.499
Show-off knowledge	0.300 to 0.399
Minimized knowledge	0.200 to 0.299
Poor knowledge	0.100 to 0.199
Null knowledge*	0.000 to 0.099

* Null and fully knowledge of values 0.000 & 1.000 are subject to constraints of Ideal machine.

iii. Probing the assumptions

It is a critical thought of identifying the associations based on assumptions towards a competitor by the corresponding recruiter.

For example

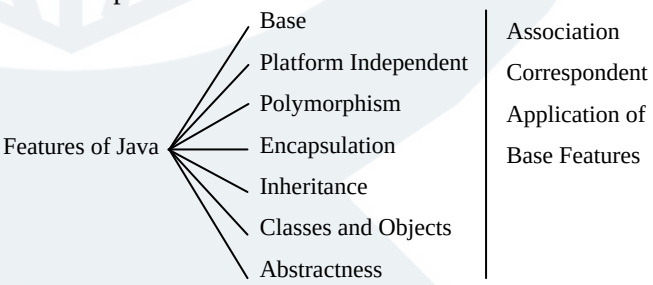


Figure3: Association Rules sample

Recruiter Selection

Assumption ⟶ Deceiver

=>Association of the following

- cues identification (verbal and non verbal)
- Test mode –self explanation
- Critical questions
- Concentration on each counter output
- Usage of Ranking / comparison

iv. Operational rules

If (More Quantified Data) Then

If (Gestural Deception) Then

If (Verbal DD) Then

If (Non-verbal/modal DD) Then
If (Contradictory Results) Then
Deception Detection= true

v. Allocation of Boolean sets

$$Alloc(x) = \pi \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\alpha_i(x)\beta_i(x)\right) + \pi \sum_{i=1}^{N} \sum_{k=1}^{N} \left(\alpha_i(x)\gamma_i(x)\right)/2N$$

N = Number of testing components/ Questions
α_i = Assumption for an candidate with an initial setting of $\alpha_1(x) = 1$ as a deceiver
β_j = Non verbal communication
γ_k = verbal communication
$0 <= Alloc(x) = \mu_t(x) <= 1$

Where Alloc(x) =1 represents deceiver and Alloc(x) = 0 represents non deceiver.

vi. Statistical probability

Deceivers most probably use the recurrence strategic tokens during their responses. Let us consider the collection of sentences CR(s) consisting of a sequence of N words such as (r1, r2, …, rN), then the probability for the occurrence of CR(s) can be computed as

$$P(C_R(s) = \pi_{i=1}^{N} P\ (r_i/r_{i-n+1}, \ldots., r_{i-1})$$

where $P(r_i/r_{i-n+1}, \ldots., r_{i-1})$ = frequency $(r_{i-n+1}, \ldots., r_i)$ / frequency $(r_{i-n+1}, \ldots., r_{i-1})$

B. Randomness-Entropy for Our Proposed Research Model:

Shannon denoted the entropy H of a discrete random variable X with possible values {x1, ..., xn} as,

$$H(X) = E(I(X))$$

Here E is the expected value, and I is the Information content of X. I(X) is itself a random variable. If p denotes the probability mass function of X then the entropy can explicitly be written as.

$$H(x) = \sum_{i=1}^{n} p(x_i) I(x_i) = \sum_{i=1}^{n} p(x_i) \log_b \frac{1}{p(x_i)} = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i)$$

where b is the base of the logarithm used. Common values of b are 2, Euler's number e, and 10, and the unit of entropy is bit for b = 2, nat for b = e, and dit (or digit) for b = 10.[3]

In the case of pi = 0 for some i, the value of the corresponding summand 0 logb 0 is taken to be 0, which is consistent with the limit:

$$\lim_{p \to 0} p \log p = 0$$

The proof of this limit can be quickly obtained applying LHospital's rule.

$$H(x) = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i)$$

C. Randomness-Entropy for Our Proposed Research Model:

ii) Las vegas Algorithm

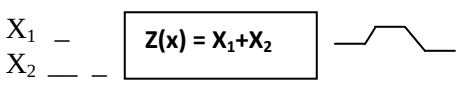
Las Vegas algorithm is a randomized algorithm that always gives correct results; that is, it always produces the correct result or it informs about the failure. The usual definition of a Las Vegas algorithm includes the restriction that the expected run time always be finite, when the expectation is carried out over the space of random information, or entropy, used in the algorithm. The complexity class of decision problems that have Las Vegas algorithms with expected polynomial runtime is ZPP.(Zero-error Probabilistic Polynomial Time) It turns out that

$$ZPP = RP \bigcap RP^{-1}$$

class. RP-Randomized Polynomial complexity class and its inverse as CO-(RP) or RP-1 which is intimately connected with the way Las Vegas algorithms are sometimes constructed. Namely the class RP is randomized polynomial time consists of all decision problems for which a randomized polynomial-time algorithm exists that always answers correctly when the correct answer is "no", but is allowed to be wrong with a certain probability bounded away from one when the answer is "yes". Thus Las vegas plays its vital role in decision making.

D. Random Uncertainty Evaluation for Our Proposed Research Model

The uncertainty has a probabilistic basis and reflects incomplete knowledge of the quantity. All measurements are subject to uncertainty and a measured value is only complete if it is accompanied by a statement of the associated uncertainty .The output quantity denoted by Z is often related to input quantities denoted by X1, X2,…,XN in which the true values of X1, X2,…,XN are unknown. Then the uncertainty measurement function Z(x) = f(X1, X2, …, XN) Consider estimates X1, X2, …, XN respectively towards X1, X2,…, XN based on certificates, reports, references, alarms and assumptions. Each Xi ~ prob. Distribution



The standard uncertainty value for Z(xi) can be approximated as standard deviation for prob(xi)

Table 2: Probability Rating

Interval	Knowledge Rating for a candidate	Prob.
0 – 10	NULL-A	0.0 to 0.1
11 – 20	POOR-B	0.1 to 0.2
21 - 30	MINIMIZED-C	0.2 to 0.3
31 - 40	SHOW-OFF-D	0.3 to 0.4
41 - 50	PARTIAL-E	0.4 to 0.5
51 – 60	AVERAGE-F	0.5 to 0.6
61 – 70	SUFFICIENT-G	0.6 to 0.7
71 – 80	DESIRED-H	0.7 to 0.8
81 – 90	MAXIMIZED-I	0.8 to 0.9
91 – 100	FULLY-J	0.9 to 1.0

i. Standard / Critical Questionnaire

Expert–.25,I–.5,H-.75,G–.99,F to A – 1.0

Above AVG–I-.25,H-.5,G-.75,F–.99,E to A – 1.0

Average–H-.25,G–.5,F-0.75, E-0.99, Dto A – 1.0

Below AVG–G-.25,F-.5,E-.75,D-.99,Cto A – 1.0

Dissatisfied– F-.25,E–.5,D-.75,C-.99,Bto A – 1.0

Nullified–E-0.25, D-0.5, C-0.75 ,B – 0.99 ,A-1.0

ii. Optimal / Normal Questionnaire

Expert: J-0.5 ,I-0.75, H-0.99, G to A -> 1.0

Above AVG: I-0.5,H-0.75,G-0.99, F to A->1.0

Average:H-0.5,G-0.75,F-0.99, E to A->1.0

Below AVG: G-0.5,F-0.75,E-0.99,D to A->1.0

Dissatisfied: F-0.5,E-0.75,D-0.99,C to A->1.0

Nullified: E -0.5,D-0.75,C-0.99,B to A ->1.0

iii. Explicit / Easier Questionnaire

Expert: J-0.75 I -0.99 H-A -> 1.0

Above average: I-0.75,H-0.99,G-A->1.0

Average:H-0.75,G-0.99,F-A->1.0

Below average:G-0.75,F-0.99,E-A->1.0

Dissatisfied: F-0.75,E-0.99,D-A->1.0

Nullified: E-0.75,D-0.99,C-A->1.0

IV. EXPERIMENT

The research methodology is a combination or a fusion of Uncertainty, Fuzziness and Randomized nature. We want to test the integrity with various application domains for evaluation and comparison. The Domain needed for the focusing are as follows,

1. Matrimonial Centre/Site.
2. Spam Mail
3. Jobsite.
4. Social Network.
5. SMS system.
6. Advertisements.

DOMAIN 1: Matrimonial Centre / Site

Referring a qwer centre at Tirunelveli district, Tamilnadu, India .A sample of 60 profiles is taken and the results are in Table 3.

Fuzzy classification implementation derives several components Such as Name, Parent, Image, DOB,Job Description, Salary, Marital status, qualification, extra curricular activities etc. Based on Uncertainty measurements we focus on the key factors as follows,

Table-3: Matrimonial Profile Assessment

Deception	Deception Possibilities (Expected/ Warnedby owner)	Using Uncertainty Evaluation	Using Fuzzification Evaluation	Using Randomness Evaluation
Image	>80 %	49/60	9/11	2/2
DOB	>70%	38/60	18/22	4/4
Job Description	>80%	39/60	18/21	3/3
Salary	>90%	52/60	6/8	2/2
Marital status	>20%	52/60	8/8	Nil
Qualification	>50%	25/60	30/35	5/5

DOMAIN 2: Spam Mail

We took a sample of 100 mails. Fuzzy classification implementation derives several components Such as Attraction, Affection, Intimation, Online commercial, subscription, Entertainment, Sympathy, donation, softwares, marketing etc. Based on Uncertainty measurements and randomized datum analysis we focuses on the key factors as follows,

Table-4: SPAM MAIL Assessment

Deception component	Deception Possibilities (Expected)	Using Uncertainty Evaluation	Using Fuzzification Evaluation	Using Randomness Evaluation
Prize/Lottery/Travel trip	11	1/11	0/11	-
Friend Invitation	13	1/13	0/13	-
Love/Marriage/Sex	12	2/12	0/12	-
E-shopping	8	3/8	0/8	-
Magazines/Club/Subscription	4	1/4	0/4	-
Movie/Songs/Video/File downloads	6	1/6	0/6	-
Help self/Others/Sympathy	3	1/3	1/3	0/3
Charity/Welfare/Disaster donation	3	2/3	1/3	0/3
Games/Play or Download	6	5/6	0/6	-
Advertisements/Marketing	21	10/21	5/21	0/21

Genuine Mails: 13/100

DOMAIN 3: Job Site

We collected some data from ABCD Softech ltd coimbatore-software company where we used HR manager datum for our research purpose. Fuzzy classification implementation derives several components Such as Qualification, Experience, Past salary, Expertise, skills, reasons for quit the past job, organizing capability etc. Based on Uncertainty measurements and randomized datum analysis we focuses on the key factors and implementing the Fuzzy, Uncertainty and randomness evaluation as follows,

Table-5: JOB SITE Profile Assessment

Deception component	Deception Possibilities (Expected / Warned by owner)	Using Uncertainty Evaluation	Using Fuzzification Evaluation	Using Randomness Evaluation
Qualification	>60%	63/100	32/63	32/32
Experience	>90%	91/100	15/91	15/15
Drawn salary	>80%	86/100	40/86	30/40
Expertise	>75%	72/100	21/72	21/21
Reasons for quit prior Job	>90%	95/100	80/95	60/80

Genuineness in jobsite datum is of 5 % in all the aspects.

DOMAIN 4: Social Network

Here the datum are organized from 10 of my well known friends circle,10 of my third party relation circle,10 of random sample of asdf engg college students with initial awareness of our research concept. Fuzzy classification implementation derives several components such as Qualification, Experience, Age, Sex, Location etc. Based on Uncertainty measurements and randomized datum analysis we focuses on the key factors and implementing the Fuzzy, Uncertainty and randomness evaluation as follows,

Table-6: Social Network Profile Assessment

Deception component	Deception Possibilities (Expected)	Using Uncertainty Evaluation	Using Fuzzification Evaluation	Using Randomness Evaluation
Age	>90 %	17/30	9/17	6/9
Sex	>75%	5/30	3/5	3/3
Location	>90 %	6/30	3/6	3/3
Job& Qualification	>90 %	13/30	4/13	4/4
Name	>95 %	28/30	14/28	14/14

DOMAIN 5: SMS-Short Messaging Service

Here the datum are organized from our mobile, well known friends circle, third party relation circle, random sample of asdf engg college students with initial awareness of our research concept. Fuzzy classification implementation derives several components Such as Message length, frequency and type etc. Based on Uncertainty measurements and randomized datum analysis we focuses on the key factors and implementing the Fuzzy, Uncertainty and randomness evaluation as follows,

Table-7: SMS Strategies Assessment

Deception component	Deception Possibilities (Expected/Warned by owner)	Using Uncertainty Evaluation	Using Fuzzification Evaluation	Using Randomness Evaluation
Message Frequency	Often/Rare/Normal	OK	OK	OK
Message length	Short maximum	----	OK	OK
Message Sender Age/Sex/Location	Forbidden	OK	OK	----
Message Type	Interruption/Interception/ Modification/ Fabrication	OK	OK	OK
Message Motive	Fallacy/Attraction/Threat/ Trap/ Emotional/Sensitive/ Informative	---	OK	OK
Message Multimedia content	Audio/ Image/ Video/ Text	OK	OK	------
Message Language	Regional/Good English/Lazy typist	OK	OK	OK
Message Time	Day/Night/Midday/Midnight	OK	----	OK
Message Model/Prototype	Standard Format/ Predefined/Customized	OK	OK	OK
Message Mobile Network	Internal/External	OK	OK	OK

OK- represents deception identification possibility.

DOMAIN 6: Advertisements

In this domain the datum are organized from our TV, Internet, and Newspapers etc. Fuzzy classification implementation derives several components Such as Advt type, Mode, Pitch, Motive, categorization etc. Based on Uncertainty measurements and randomized datum analysis we focuses on the key factors and implementing the Fuzzy, Uncertainty and randomness evaluation as follows,

Table-8: Advertisement Assessment

Deception component	Deception Possibilities (Expected/Warned by owner)	Using Uncertainty Evaluation	Using Fuzzification Evaluation	Using Randomness Evaluation
Events/Exhibition/Park	20 %	5 %	10%	5 %
Consumer-Products	30 %	5%	20%	5%
Medicines & Cosmetics	70 %	30%	30%	10%
Land/Real-estate related	90 %	20%	60%	10%
Medicine-Private/Secret disease cures	90 %	15%	65%	10%
Travel & tourism	20 %	5%	10%	5%
Education related	50 %	10%	30%	10%
Food Related	20 %	4%	10%	6%
Cloth Related	30 %	10%	12%	8%
Government sectors Related	3 %	1%	1%	1%

V. RESULTS AND DISCUSSIONS

Our experiment comprises three levels and seven stages which are revealed as follows,

At Level-1 we applied the concept of Fuzziness, Uncertainty and Randomness towards several domains; we faced lot of diverging factors which leads us to unpredictable characteristics for identifying the deception detection. At first we applied the Randomness towards the Matrimonial site, spam mail, jobsite, social network, and short messaging service and advertisement domains. It identifies the least deception detections but filtering the fair sided datum from our spool of research items. Second we applied the Uncertainty tools for identifying the deception detection towards Matrimonial site, spam mail, jobsite,

social network, and short messaging service and advertisement domains. It provides us tolerable detections of deception; third we applied the Fuzziness concept for identifying the deception detection towards Matrimonial site, Spam mail, jobsite, social network, and short messaging service and advertisement domains.

Then at Level -2 we implement the mixture of two components at a time such that the output filtering results of first component will be the input for the second component, taking Randomness & Uncertainty, Randomness & Fuzziness and Fuzziness & Uncertainty are the three stages for this level of implementation to towards the Matrimonial site, spam mail, jobsite, social network, short messaging service and advertisement domains. Here Randomness & uncertainty produces 50% to 60 % efficiency, Randomness & Fuzziness produces 61% to 75 % Fuzziness & uncertainty produces 76% to 90% for detecting deceptions.

Then at Level-3 we implement the Full-fledged combination of Randomness, Uncertainty and Fuzziness to towards the Matrimonial site, spam mail, jobsite, social network, and short messaging service and advertisement domains. It produces more than 95 % efficiency in detecting deceptions successfully.

Table-9: Performance Assessment

Level	Stage	Success Rate	Matrimonial Site	Spam Mail	Jobsite	Social Network	SMS	Advertise-ment
1	1	Randomness	20 %	18 %	26 %	31 %	24 %	18 %
	2	Uncertainty	35 %	36 %	32 %	36 %	34 %	30 %
	3	Fuzziness	45 %	44 %	48 %	46 %	43 %	40 %
2	1	Randomness & Uncertainty	56 %	53 %	53 %	52 %	57 %	56 %
	2	Randomness & Fuzziness	68 %	71 %	66 %	67 %	63 %	66 %
	3	Fuzziness & Uncertainty	79 %	86 %	83 %	87 %	86 %	90 %
3	1	Randomness, Uncertainty and Fuzziness	93 %	92 %	94 %	93 %	96 %	96 %

VI. CONCLUSION

Deception detection is an art or a fashion which itself act as a separate course in our new trends of life. We consider three major concepts of decision making components such as fuzzy logic, Uncertainty and Randomness for our implementation strategies to towards the Matrimonial site, spam mail, jobsite, social network, and short messaging service and advertisement domains. We implement the strategies individually, then taking two at time and finally with all the three components and identifies several results. All the components mixture produces excellent results than taking two at a time moreover taking two at a time is better than applying in its individual nature. In This research paper we acquire several information concepts, modality which leads us to vast science of deception detection. So we include the facts which relates to our concept are identified for processing in our research.

Media plays a vital role in detecting the deceptions. Direct communication mode can be analyzed with the gestures feeling the waves of opponent in an exact/accurate mode, whereas video conference can be handled with proper care [4].The repetitive plays varying the speed of presentation analysis is an additional skill present in video conference while audio chat focuses on the pitch stress and pause time gaps of communication response as its primary factors [2]. SMS or Email is blind folded in detecting deceptions. Subjects in our case ABCD students are unaware of few technically advanced psychometric deceptive keywords. In future we will try with experts in this same area. We suggested this technique for the police department to use it in their deep core investigations for deception detection.

Deception is always availabe as an vital part in our daily life. But individuals and organizations need not to be powerless to detect its use against them. This paper reflects our computational effort in identifying Deception in its deep core. In future we will implement the mixture of Neuro fuzzy, Genetic algorithm,Automata theory and NP-Completeness with its deep impact.

VII. REFERENCES

[1] Alan Ryan -Professional liars - Truth-Telling, Lying and Self-Deception Social Research, Fall, 1996.

[2] Bond,c.,F. “A world of lies: the global deception research team”, JournalofCrossculturePsychology 2006Vol.37(1)60-74.

[3] Burgoon, J.K., and Qin,T. “The Dynamic Nature of Deceptive Verbal Communication”. Journal of Language and Social Psychology, 2006, vol25(1), 1-22. doi:10.1177/0261927X05284482

[4] Gupta, Bina (1995). Perceiving in Advaita Vedanta: Epistemological Analysis and Interpretation . Delhi: Motilal Banarsidass. pp. 197. ISBN 81-208-1296-4.

[5] Pennebaker,J.W,Mehl,M.R.&Niederhoffer,K. ”Psychological aspects of natural language use: our words, ourselves”. Annual Review of Psychology, 2003, 54,547-577

[6] Steve Woznaik,L.Simon, 2002. “The art of deception: controlling the human element of security”. Wiley; 1 edition.

[7] Whissell,C., Fournier,M.,Pelland,R., Weir, D.,& Makaree,K. ”A comparison of Classfiifcation methods for predicting deception in computer-mediated communication”. Journal of Management Information systems, 2004,139-165.

[8] Zuckerman, M.,DePaulo, B.M. and Rosenthal, R.”Verbal and Nonverbal Communication of Deception”. In L.Berkowitz(Ed)(1981). doi:10.1016/S0065-2601(08)60369-X

How to cite

S.Rajkumar, V.Narayani, Dr.S.P.Victor, "Unification of Randomized Anomaly in Deception Detection using Fuzzy Logic under Uncertainty". *International Journal of Research in Computer Science*, 2 (2): pp. 7-14, February 2012. doi:10.7815/ijorcs.22.2012.016