Epidemic Analysis of uncertainty in Deception Detection under Fuzzified Anomalies

S. Rajkumar, V. Narayani, S. P. Victor

Abstract - Nowadays in this competitive world of job seekers, the necessity of job makes many recruiters to provide more cautious on their selection process. The recruitment process is definitely a fuzzified anomaly for all the components available in the environment. The art of deception also changes its face with a modern artistic fashion. This paper deals with the uncertainty features which play the major role of Deception in a fuzzified environment of Recruitment process. We deal with the impacts of uncertainty in deception detections and also with the underlying environment of fuzzification. In this paper we proposed a Research Model which considers the linkage of fuzzification and uncertainty in Deception Detection. In this paper we implement our proposed model with an experiment which includes warning and lack of warning to the recruiters upon the competitors. Enumerated results and discussions mould the impact of uncertainty and fuzziness in Deception Detection.

Index terms **– Deception, Fuzzy logic, Randomization, Uncertainty.**

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.

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.

While it is difficult to deceive a partner over a long period of time, deception often occurs in day-to-day conversations between relational partners. Detecting deception is difficult because there are no known completely reliable indicators of deception. Deception, however, places a significant cognitive load on the deceiver. He or she must recall previous statements so that his or her story remains consistent and believable. As a result, deceivers often leak important information both verbally and nonverbally.

Manuscript received December 23,2011.

S.Rajkumar, Research Scholar, Bharathiar University, HOD/CSE, Nehru Institute of Engineering & Technology, Coimbatore, India, 9944451769, (rajkumar_proff@rediffmail.com).

V.Narayani, Dept. of Computer Science, St.Xavier's College, Tirunelveli, India, 9944451769, (narayaniv79@rediffmail.com).

Dr.S.P.Victor, Dept. of Computer Science, St.Xavier's College, Tirunelveli, India, 9486041594, (victorsp@rediffmail.com).

B. Fuzzy logic

Fuzzy sets have movable boundaries, *i.e.*, the elements of such sets not only represent true or false values but also represent the degree of truth or degree of falseness for each input. 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.

Fuzzy logic has mostly been applied to control systems. Fuzzy control systems interpret the expert human and replace them for performing certain tasks such as control of a power plant. Fuzzy controllers apply decision rules (if-then rules) by making use of critical variables to interpolate the output between the crisp boundaries. Some typical examples where fuzzy logic has been implemented are

1. Railway (Sendai Railways in Japan)
2. Automobile industries (transmission and braking)
3. Heating and cooling systems
4. Copy machines
5. Washing machines

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.

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.

II. RESEARCH DESIGN MODEL

Interviewing is more important in identifying deception than with simple observations in a possible environment. Our proposed research model N-sects the subject with its various criteria and then the functional components are implied for Deception detection.

A Proposed Model

In this proposed model we dealt with the entire scenario under uncertainty. Here all the components are covered with an fuzzified environment, each component can be accessed for the exact evaluation of any uncertain fuzzified anomalies.

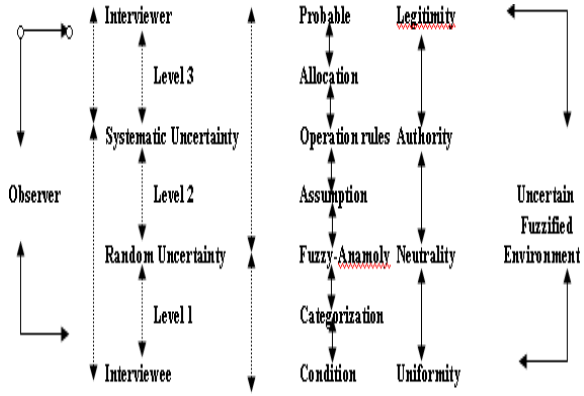


Fig. 1: Research Model

III. RESEARCH METHODOLOGY

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

a) Specify the range of conditions

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

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

b) Classification and categorization

Table I: Membership value assignments

Factor-X	Membership value $\mu_t(x)$
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.

c) Probing the assumptions

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

For example

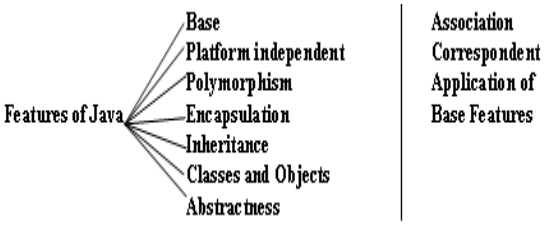


Fig.2: Sample Association

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

d) Operational rules

If (More Quantified Data) Then

If (Gestural Deception on Initial setup) Then

If (Verbal DD) Then

If (Non-verbal/modal DD) Then

If (Contradictory Results) Then

Deception Detection= true

e) Allocation of Boolean sets

$$Alloc(x) = \pi \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha_i(x)\, \beta_j(x)) + \pi \sum_{i=1}^{N} \sum_{k=1}^{N} (\alpha_i(x)\, \gamma_k(x)) / 2N$$

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.

f) Statistical probability

Deceivers most probably use the recurrence strategic tokens during their responses. Let us consider the collection of sentences $C_R(s)$ consisting of a sequence of N words such as $(r_1, r_2, \ldots, r_N)$, then the probability for the occurrence of $C_R(s)$ can be computed as

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

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

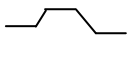
IV. UNCERTAINTY EVALUATION

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.

Fractional uncertainty is the measurement uncertainty divided by the measured value. The output quantity denoted by Z is often related to input quantities denoted by X_1, X_2,…,X_N in which the true values of X_1, X_2,…,X_N are unknown. Then the uncertainty measurement function Z(x) = f(X_1, X_2, …, X_N) Consider estimates X_1, X_2, …, X_N respectively towards X_1, X_2,…, X_N based on certificates, reports, references, alarms and assumptions.

Each X_i ~ prob. Distribution

X_1 –
X_2 — – **Z(x) = X_1+X_2**



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

Table II: Probability Rating

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

A) Standard / Critical Questionnaire
Expert – J – 0.25, I – 0.5,H-0.75, G – 0.99, F to A – 1.0
Above AVG– I -0.25,H- 0.5,G-0.75,F – 0.99,E to A – 1.0
Average– H-0.25, G – 0.5, F -0.75, E-0.99, Dto A – 1.0
Below AVG–G-0.25, F-0.5,E -0.75,D- 0.99,C to A – 1.0
Dissatisfied– F-0.25,E – 0.5,D-0.75,C- 0.99, B to A – 1.0
Nullified– E-0.25, D-0.5, C-0.75 ,B – 0.99 ,A-1.0
B)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
AverageH-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
C) 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
Unsatisfied ->F-0.75,E-0.99,D-A->1.0
Nullified ->E-0.75,D-0.99,C-A->1.0

V. EXPERIMENT

Sample space: Collecting toppers form each course and provide 10 days for preparation.
Problem: Identify a Single Java &.Net Expert
Subjects were students from an Engineering college at a large Indian university. They were selected from the different branches of Engineering for the participation in this study. Eligible subjects were told about the participation towards mock-up interviews for their placement training and also for research authorities. A total of 100 subjects took part in the study. They were all demanded with their resumes.

CSE subjects	- 53
ECE subjects	- 17
EEE subjects	- 20
Aero subjects	- 10
	100

Analysis process includes
1) Verification of Resumes
2) Written test Questionnaire
3) HR – Interview – Face to face

While analyzing the resumes, the requirement of Java for an IT professional can be manipulated in terms of C++ and VC++ and fake experience in software centers as part-time basis are all revealed by their certificate copies.

So we rejected 20 subjects after the confirmation about their resume due to the deceptiveness about their biodata.
Rejected Candidates

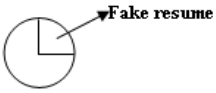


Fig.3: Fake Representation

CSE	- 03
ECE	- 07
EEE	- 05
Aero	- 05
	20

Now actual calculated subjects are,

CSE	- 50
ECE	- 10
EEE	- 15
Aero	- 05

80 subjects split into 2 parts of 40 each to hall A and hall B in which hall A is warning informed over Deceptiveness of subject but hall B is of lackness in warning.

Hall A	Hall B	
CSE- 25	CSE	- 25
ECE- 05	ECE	- 05
EEE- 08	EEE	- 07
Aero- 02	Aero	- 03

Now focussing on 40 subjects comparing the resumes and technical questiona the classification and categorization of subjects are as follows:
Knowledge can be assessed based on the academic performance,response to the Queries, Evaluation under technical schema etc.We can assess the subject performance based on their responses.The assessment values are,

(a)

Table III:Knowledge Assessment

Category	CSE	ECE	EEE	Aero	Total
Fully knowledged	2	1	0	0	3
Maximum knowledged	6	2	2	1	11
Desired knowledged	2	0	1	0	3
Sufficient knowledged	2	0	1	0	3
Average knowledged	6	0	1	0	7
Partial knowledged	4	0	1	0	5
Show-off knowledged	2	1	1	0	4
Minimized knowledged	1	1	1	0	3
Poor knowledged	0	0	0	1	1
Null knowledged	0	0	0	0	0
	25	5	8	2	40

Total number of suspicious Deceptive subjects => 5+4+3+1+0 = 13 subjects.

(b) Apply these input of 13 subjects to stage based on Probing the assumptions using the Association Informations

Cues - 3 Identified
Test mode - 4 Identified
Critical Ques. - 3 Identified
Concentration on each counter o/p – 2 Identified
Usage of Ranking/ Comparison – 1 Identified
Deceptions = 13 confirmed
Remaining subjects = 27

(c) Now applying the fuzzy operational rules we obtained the following results.

Quantitative data = 27 subjects
Irrelevant = 21 subjects
Deception Detection on initial = 20 subjects
Verbal Deception detection = 19 subjects
Non verbal DD filteration = 18 subjects
Contradictory Result = 16 subjects
Suspecting subjects = 16 subjects

(d) Applying these input of 16 subjects to stage 4 based on allocation of Boolean sets.

$¥_i$ = 1 to 16 $\alpha_i(x)$ =1

Let β be Non verbal & γ be Verbal communication memberships.

Table IV: Subject classification and Evaluation

Sub.	1	2	3	4	5	6	7	8
β_j	0.2	0.3	0.1	0.3	0.2	0.4	0.3	0.2
γ_j	0.1	0.25	0.1	0.2	0.3	0.3	0.2	0.1
Sub.	9	10	11	12	13	14	15	16
β_j	0.4	0.5	0.4	0.5	0.3	0.5	0.4	0.2
γ_j	0.3	0.4	0.3	0.4	0.1	0.3	0.2	0.1

subject 10,12 & 14 are suspetible in β remaining are rejected.But using γ_j 10, 12 & 14 are < 0.5 member value, So rejected => confirmed rejection of 16 subjects therefore Remaining subjects = 11,Applying these 11 subjects towards statistical probability stage.

The recurrence word collections are like
{And (r_1), Actually (r_2), It is …(r_3),
whereever(r_4), when we want to (r_5), It will
be (r_6), etc etc (r_7), once more please (r_8),
sorry (r_9), Mmmm (r_{10})}

Table V:Subjects & its Probability results

P(X)	S_1	S_2	$*S_3$	$*S_4$	S_5	S_6
$P(r_1)$	0.8	0.7	0.1	0	0.8	0.9
$P(r_2)$	0.7	0.7	0.2	0	0.7	0.6
$P(r_3)$	0.8	0.6	0	0	0.8	0.7
$P(r_4)$	0.6	0.7	0.1	0.1	0.6	0.6
$P(r_5)$	0.7	0.6	0.2	0.1	0.6	0.7
$P(r_6)$	0.6	0.5	0	0.1	0.8	0.9
$P(r_7)$	0.8	0.7	0	0.1	0.6	0.4
$P(r_8)$	0.7	0.3	0	0	0.3	0.8
$P(r_9)$	0.8	0.2	0	0.1	0.7	0.6
$P(r_{10})$	0.7	0.4	0.1	0	0.8	0.3

P(X)	$*S_7$	S_8	S_9	$*S_{10}$	S_{11}
$P(r_1)$	0	0.6	0.7	0.1	0.8
$P(r_2)$	0.1	0.8	0.7	0	0.8
$P(r_3)$	0	0.9	0.7	0.1	0.7
$P(r_4)$	0	0.8	0.6	0	0.6
$P(r_5)$	0.1	0.6	0.7	0	0.7
$P(r_6)$	0.1	0.8	0.7	0.1	0.8
$P(r_7)$	0	0.7	0.7	0.1	0.8
$P(r_8)$	0	0.6	0.8	0	0.8
$P(r_9)$	0	0.7	0.7	0.1	0.6
$P(r_{10})$	0	0.8	0.7	0.1	0.6

S_3, S_4, S_7 & S_{10} provide least probability for deception. Eliminating S_1, S_2, S_5 , S_6,S_8, S_9& S_{11} Total number of remaining subjects = 4.

Now we are in the situation of selecting a single expert in Java based on 4 options as subjects. Now these subjects leads us to uncertainty to select the best one among them.

Type A

By repeated measurement of X_i we can observe that it follows the Gaussian Distribution.

Type B

By using the above schematic approach we can observe that it follows the Rectangular Distribution

Using the type A – uncertainty measurement

Apply the previous results over Gaussian distribution.

S_3 0.6
S_4 0.5
S_7 0.3
S_{10} 0.6

S_3 & S_{10} will be rejected.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Remaining subjects = S_4 & S_7 = 2
Our assumption is S_7 is good but for
Confirmation now apply S_4 & S_7 for type B uncertainty measurement.
So apply the scientific records using
Rectangular distribution.

S_4 = 0.4 (wrong)
S_7 =0.2 (right)

Therefore final expert = S_7 = 1.
Now focussing the hall B consisting of 40 subjects we identitifed 10 deceptive subjects based on their resume and performance.

VI. RESULT

Comparing the performance in Hall A and Hall B we are having these following results

TableVI:Model Evaluation

Proposed Model	Hall A	Hall B
Fuzzy logic and uncertainty principle	1/40	30/40

In Hall A of 40 subjects

Table VII:Result Consolidation for Hall-A

S.No	Modular components	Deception Detected/ Actual	%
1	Classification & Assumption	13/40	32.5
2	Fuzzy oper. & allocation of Boolean logic	16/27	59.25
3	Statistical probability implementation	07-Nov	63.63
4	Type –A uncertainty	02-Apr	50
5	Type – B uncertainty	01-Feb	50

The remaining 30subjects of Hall B will now be analyzed with respect to the combination logic of fuzzy and uncertainty. Repeating the steps we finally collect a good subject from Hall B but with lower perfromance when compared with Hall A expert subject – S_7.

Hall B good candidate => 0.6 (wrong)
S_7 => 0.3 (right)

Not surprisingly the expert belongs to its
core course of engineering.

TableVIII:Result consolidation for Hall-B

S.No	Modular components	Deception Detected/Actual	%
1	Classification & Assumption	Oct-40	33
2	Fuzzy oper. & allocation of Boolean logic	Dec-20	60
3	Statistical probability implementation	05-Aug	62.5
4	Type –A uncertainty	01-Mar	66
5	Type – B uncertainty	01-Feb	50

VII. DISCUSSION

We achieve somewhat better results when combine more than on of individual component implementation of Fuzzy,Randomness and Unceratinity rather than with its individuality. We identified that warned reviewers were more successful at Detecting Deception in an interview rather than with unwarned reviewers.They found with twice efficiency rather than with the unwarned.These results definitely impact the recruiters or job providers or employers who used direct or computer based interviews for their selection process.Methods that accurately detect deception and generate few false alarms are preffered to methods that are not as discerning.

The selection of an expert from the set of toppers of Engineering college is different from collecting a set of good performers without deceivers. Each subject tries to deceive us in terms of their knowledge, practice and experience (exceptionals).

No. of Questions => 40 (passes – 2 qns only)
Exact Ans -> 10/10
Relevant Ans-> 9/10 upto 5/10 based on selection
Fabricated answers -> 2/10
Modificated answers -> 1/10
Irrelevant answers -> 0/10
Idiotic response -> 0/10

VIII. CONCLUSION

Detecting the deceptions in the process of recruitment selection is a tidious approach.The combination fuzzy logic and uncertainty plays a vital role in the selection process from a huge set of competitors. For selecting the best candidate the recruitment selection definitely avoid the manual prejudice process but focussing with recent trends and techniques of fuzzy logic and uncertainty leads them to achieve their goal in an efficient way. In near future we will combine randomization, fuzzy, uncertainty in deception detection.

Applying the Randomization techniques ,Fuzzylogic and Uncertainty towards identifying the deception is a critical process of complexity,but the results are more effective when compare it with implementing each phase individually.

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. 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. SMS or Email is blind folded in detecting deceptions.

Many automated systems are now required for deception detection handling in an optimized manner,we moreover try to implement the concept of artificial intelligence,neurofuzzy and Genitic algorithm combinations for detecting Deceptions in our recent data communication strategies and components.

REFERENCES

1. Steve Woznaik, Kevin D.Mitnick, Willaim L.Simon,2002. "The art of deception: controlling the human element of security". Wiley; 1 edition.
2. Zuckerman, M.,DePaulo, B.M. and Rosenthal, R."Verbal and Nonverbal Communication of Deception".In L.Berkowitz(Ed)(1981)
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.
4. Bond,c.,F. "A world of lies: the global deception research team", Journal of Cross-culture Psychology, 2006, Vol.37(1), 60-74.
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. 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, 2004,20(4),139-165.

AUTHORS PROFILE

Mr. S. Rajkumar completed his M.E–CSE at Sathyabama University, Chennai and currently doing his Ph.D in the area of Computational Science. He is a Research Scholar of Bharathiar University and working as a HOD/CSE at NIET Coimbatore.

Ms. V. Narayani completed her M.C.A in M.S University, Tirunelveli and M.Phil in Mother Teresa University, Kodaikanal. She submitted her Ph.D thesis in the area of Data Mining.

Dr. S. P. Victor earned his M.C.A. degree from Bharathidasan University, Tiruchirappalli. The M.S.University, Tirunelveli, awarded him Ph.D. degree in Computer Science for his research in Parallel Algorithms. He is the Head of the Department of Computer Science, and the Director of the Computer Science Research centre, St. Xavier's college (Autonomous), Palayamkottai, Tirunelveli. The M.S. University, Tirunelveli and Bharathiar University, Coimbatore have recognized him as a research guide. He has published research papers in international, national journals and conference proceedings. He has organized Conferences and Seminars at national and state level.