Security Requirement Elicitation Techniques: The Comparison of Misuse Cases and Issue Based Information Systems

Naveed Ikram[1], Surayya Siddiqui[2], Naurin Farooq Khan[1]
Ibn-e-Sina Empirical Software Engineering Lab
[1]Riphah International University
Islamabad, Pakistan
[2]International Islamic University
Islamabad, Pakistan

***Abstract*—There are myriads of security elicitation techniques reported in the literature, but their industrial adoption is inadequate. Furthermore there is a shortage of empirical and comparative evaluations which can aid the software industry in this respect. This paper compares two security elicitation techniques – Misuse cases (MUC) and Issue based information systems (IBIS) by carrying out controlled experiments. A 2*2 factorial design was used with 30 undergraduate students selected randomly who solved security goal identification tasks on an individual basis using the two techniques. Two dependent variables chosen were; *effectiveness of the techniques* in terms of number of security goals identified and *coverage of the techniques* in terms of number of types of security goals, time taken to learn, execute and interpret results by each technique in three different situations. The main finding was that in a situation of low level of detail, the time taken to interpret results was lower in IBIS while in medium and high level of detail MUC is more effective for finding security goals and provides better coverage by taking less learning time. The generality of the results is limited due to the fact that undergraduate students participated in the experiment. The study provides guideline for the software industry about the choice of security elicitation technique in three different situations. The study can be extended by adding multiple techniques for comparison and a framework can be developed.**

***Index Terms*—Security requirements engineering, Misuse cases, Issue Based Information Systems, Experiments.**

I. Introduction

The acceleration of security threats and vulnerabilities during the last decade has shifted the focus of software developers away from considering security in the context of security policies or architectural mechanisms. Now it is widely undisputed to consider security at requirement level because it serves as a basis for system development life cycle. The academia has to offer myriads of security elicitation techniques to choose from, however there is not much evidence in the literature that can help pin down the appropriate security elicitation technique in different situation [1]. Besides, no guidelines are present in the literature which can help in the selection these techniques in different situation. [2][3][4][5][6][7][8][9]. This lack of knowledge coupled with the lack of awareness of people [10] writing the security requirements becomes the rationale of our research work. We present an experimental evaluation of two security elicitation techniques: Misuse cases (MUC) and Issue based Information systems (IBIS) in three different situational characteristics. The goal of this study is to pave a way for the industrial take up of these and similar techniques in general. On the basis of the results of the evaluation, proper guidelines are developed for the practitioners. This can help not only to outline the philosophy of different security requirement elicitation techniques but also to make decision in their selection in a particular situation. As a result security engineers can appreciate the strengths and counter the weaknesses of security requirements elicitation techniques in terms of a given situation. The authors admit that in order to establish a concrete empirical foundation of the selection of techniques, similar experiments should be performed in an industrial set up with practitioner as subjects involving other security requirements methods.

II. Related Work

There is a dearth of empirical evaluation of security elicitation techniques [1]. However there are a few comparative researches regarding security elicitation which exist in literature with a focus to compare their characteristics. Common Criteria, MUC and Attack Trees have been comparatively evaluated on a case study based on the characteristics of each technique as in [2]. The authors suggest using all of the techniques as a combined methodology, and lack to provide any guidance as to which technique to choose in different situations. L. Chung, F. Hung, E. Hough, and D. Ojoko-Adams undertook evaluation of MUC technique and eight requirement elicitation techniques, (SSM, QFD, CORE, IBIS, JAD, FODA, CDA, and ARM) out of which IBIS, JAD and ARM were selected based on predefined criteria [3]. The context was of SQUARE evaluation. However the evaluation criteria were not specific to security domain. Eight security requirement engineering processes were compared with a focus on different phases of SDLC [4]. The findings from the study showed that none of the techniques was suitable for security elicitation at requirement engineering (RE) level. CLASP and Microsoft's SDL were theoretically compared, with a major

978-1-4799-6336-2/14/$31.00 © 2014 IEEE

focus on strengths and weaknesses of the two processes [5]. The evaluation however remains silent about how each process contribute to the requirement phase of SDLC. A. Herrmann, D. Kerkow, and J. Doerr provided a case study based evaluation of IESE-NFR and MOQARE – requirement elicitation techniques [6]. The evaluation used two different benchmarks, former being quality attributes while later being security. A comparative survey of security engineering process and security elicitation techniques was undertaken in order to acknowledge their strengths and weaknesses [11]. The authors divided each security requirement processes into five phases namely security elicitation, analysis, specification, management and support for later stages of SDLC and established an evaluation criteria to measure performance level of each process in each phase. A theoretical evaluation of RE processes (SQUARE, XP Oriented approach, CLASP, Microsoft SDL, Secure Software Development Process, Security Engineering Processes and Security Requirement Engineering Framework) was performed [7]. The authors endeavored to develop a new security requirement engineering process which is easy to use and provides step wise guidance. A thorough literature review of security related concepts and security requirement elicitation and engineering processes was undertaken by B. Fabian, S. Gürses, M. Heisel, T. Santen, and H. Schmidt [8]. The authors categorized the security processes into four multilateral approaches namely UML based, goal oriented, problem frame based, risk analysis based and Common Criteria based approaches. Although the outcome of their research was a conceptual framework on the basis of which certain security requirement engineering methods were evaluated, the selection of these methods is a question mark. Also the study only considers the conceptual properties of the selected methods and does not add any situational perspective to it. A theoretical comparison of security development lifecycles was presented [9]. The authors took the activities as a basis of comparison with a focus on security in general and their selection in different situational characteristics. An experimental evaluation of Textual and Diagrammatical MUC provides guidelines about their performance in terms of number of failure mode, ease of use and intention of use etc [12]. Similarly MUC and Attack Trees were evaluated in an experimental set up during two different situations [1]. The experimental findings suggested that Attack Trees performed better in terms of number and types of threats identified, perceived ease of use, usefulness and intention of use.

Our work is different from previously reported comparison in sense that we do not take into account the techniques' characteristics. We will extend [1] work by moving one step further and will add three situational characteristics.

III. Research Method

A. Motivation

The literature above shows that despite the plethora of security elicitation techniques present in the academia, very little knowledge exists that can guide the security elicitation engineer to select appropriate security elicitation technique. This lack of knowledge further accentuates the academia-industry divide and hinders the industrial take up of these security elicitation techniques. Our comparative evaluation could potentially contribute towards promoting the industrial take up of these security elicitation techniques.

B. Selection of Techniques

We have decided to evaluate MUC and IBIS due to the following reasons

1. Both MUC and IBIS have visual support.
2. Both techniques explicitly demand representation of all stakeholders' point of view in order to elicit complete set of requirements.
3. MUC is easy to learn and use and it provides complete solution [11] while IBIS has the flexibility to be learned and is used in different problem domains [13][14].

Although IBIS is used to elicit functional requirements, it has the ability to be modeled according to the nature of the given problem. IBIS has been used in a comparative evaluation of security elicitation techniques on three case studies in the context of SQUARE evaluation as in [3]. Furthermore the findings suggested that it generated discussion between stakeholders and raised security issues that has not been addressed otherwise.

1) Misuse Cases: Misuse cases (MUC) are an extension of use cases [15][16][17]. They are used to describe the behavior of an entity or a system that should not occur hence the name negative use cases. MUC are used to identify threats by capturing the behavior of an actor that is hostile to the system and hence are used to elicit security requirements [15][18]. MUC in addition to traditional use cases, have two more entities namely *misuse case* and *misuser*. *Misuse case* is the ordering of actions carried out by any entity to harm a system, while *misuser* is the entity who is responsible for starting the *misuse case*. Apart from that MUC introduces *mitigates* and *threatens* relationships. The former shows that a use case can mitigate the success chances of a *misuse case* while the later shows hindrance of a use case by a *misuse case*.

2) Security Requirement Elicitation by Issue Based Information System: IBIS is used for the elicitation of functional requirements in complex multilatory systems where multiple stake holders are involved [13]. IBIS comprises of three major elements; issues, positions and arguments. Issues that must be addressed, positions are the responses/ideas while arguments are the pros-cons of an issue that respond to the idea. It endorses greater participation in early phases to form common consensus when it comes to eliciting requirements. The participation by different stake holders also gives the opportunity that difficulties of the proposed solution which were left out by the designer can be picked up others. It can also trace back the decision making process. Another ability of IBIS is that it can be modeled with respect to the problem's nature [19] as it has been used to model the security requirements [3]. It makes use of questions, discussions among groups, model/map theories and arguments to settle contradictions. The conflict resolution is managed by group discussions. The specification patterns – map tables are used to define issues, arguments, associated positions and conflicting issues.

C. Selection of Situational Characteristics

The security requirement elicitation process is based on the availability of requirement engineering level artifacts such as business goals, functional requirements and stakeholder hierarchy etc. These artifacts provide system's understanding and are used as a building driver of security elicitation process. We will take into account the availability of such different type of system artifacts at RE level and set them as different situational characteristics. The situation of low level of detail presents the minimum set of artifacts, while the situation of high level of detail presents the maximum set of artifacts in our experimental setting. Higher the number of artifacts presents the maximum information to the subjects and vice versa. The Table 2 describes the three situational characteristics in detail.

D. Research Approach

We did a two session experimental comparison of MUC and IBIS. In each session the subjects used the two techniques on project documentation of online shopping mall in 3 different situations: low level of detail, medium level of detail and high level of detail. A 2*2 factorial experimental design was chosen in order to control the order in which the techniques were used and activities were solved. The subjects were chosen randomly from final year students of Software Engineering discipline. This random sampling helped remove any prejudices about the intellectual abilities of the students. The participants were divided into groups of 3, each dedicated for a particular situation. Every group was further divided into sub-groups of two - each used both the techniques but in different orders. Time stamps for each activity were also recorded. Dependent variables, used as outcome variable to measure the effectiveness and coverage of MUC and IBIS are described in Table 1. Effectiveness is measured by the number of security goals while the coverage is measured as number of security goal types and time taken to interpret, exercise & analyze the results by each technique.

E. Hypothesis Development

This section describes hypothesis statements for situation of low level of detail. We formulated null hypotheses for our design measure, i.e. there is no significance difference in effectiveness and coverage of the techniques tried first versus last by each subject in situation of low level of detail. Same hypothesis can be repeated for each independent variable in context of all dependent variables. In order to gauge the effectiveness and coverage of the techniques we devised main hypothesis. Tables 3-7 show the hypothesis for comparing MUC and IBIS in situation of low level of detail. Hypothesis concerning the medium and high level of detail are similar.

F. Experimental Design

Since there was a small sample, experimental design chosen for the evaluation of the two techniques was repeated measure design. This design measure experiences unsystematic practice effect and boredom effects [20]. To overcome such precincts we used random assignment of subjects and counter balancing approach. The random selection and division of participants was done in two ways. First, thirty participants were randomly selected and were divided into three groups of ten each, A first come first served approach was chosen to assign subject to groups. Each group was assigned to a particular situation. Second, participants of the groups dedicated to each situation were further randomly assigned into two sub-groups of five each. The two sub-groups in each situation performed MUC and IBIS however the order was opposite against one another (Table 8). For example in the situation of low level of detail, Group A performed MUC first followed by IBIS where as Group B performed IBIS followed by MUC. This reverse ordering helped to cater for the counterbalancing.

TABLE 1. DEPENDENT VARIABLES

Variables	Measured by
Effeciveness of the Technique	Number of security goals
Coverage of the Technique	Number of types of security goals
	Time taken to understand and interpret the technique
	Time taken to exercise the technique
	Time taken to analyze and translate the resutls in unambiguous and user understandable documents

TABLE 2. SITUATIONAL CHARACTERISTICS

Situation Name	Artifact
Low Level	Problem statement, position statement, project goals, scope and use case description
Medium Level	Problem statement, position statement, project goals, scope and use case description, user hierarchy, use case description, use case diagram, over all description of the responsibilities of system users
High Level	Problem statement, position statement, project goals, scope and use case description, user hierarchy, use case description, use case diagram, over all description of the responsibilities of system users, action sequence and deployment diagram

TABLE 3. HYPOTHESIS REGARDING NO OF GOALS IN SITUATION OF LOW LEVEL OF DETAIL

H0	There is no significance difference in number of goals identified using MUC and IBIS, in situation of low level of detail
H1	There is significance difference in number of goals identified using MUC and IBIS, in situation of low level of detail
H2	MUC is greater than IBIS, in terms of number of goals identified in situation of low level of detail
H3	IBIS is greater than MUC, in terms of number of goals identified in situation of low level of detail

TABLE 4. HYPOTHESIS REGARDING NO OF GOAL TYPES IN SITUATION OF LOW LEVEL OF DETAIL

H0	There is no significance difference in number of types of goals identified using MUC and IBIS, in situation of low level of detail
H1	There is significance difference in number of types of goals identified using MUC and IBIS, in situation of low level of detail
H2	MUC is greater than IBIS, in terms of number of types of goals identified in situation of low level of detail
H3	IBIS is greater than MUC, in terms of number of types of goals identified in situation of low level of detail

TABLE 5. HYPOTHESIS REGARDING LEARNING TIME UTILIZATION IN SITUATION OF LOW LEVEL OF DETAIL

H0	There is no significance difference in learning time using MUC and IBIS, in situation of low level of detail
H1	There is significance difference in learning time using MUC and IBIS, in situation of low level of detail
H2	MUC is greater than IBIS, in terms of taking less time in situation of low level of detail
H3	IBIS is greater than MUC, in terms of taking less time in situation of low level of detail

TABLE 6. HYPOTHESIS REGARDING EXECUTION TIME UTILIZATION IN SITUATION OF LOW LEVEL OF DETAIL

H0	There is no significance difference in technique execution time using MUC and IBIS, in situation of low level of detail
H1	There is significance difference in technique execution time using MUC and IBIS, in situation of low level of detail
H2	MUC is greater than IBIS, in terms of taking less execution time in situation of low level of detail
H3	IBIS is greater than MUC, in terms of taking less execution time in situation of low level of detail

TABLE 7. HYPOTHESIS REGARDING RESULT INTERPRETATION TIME IN SITUATION OF LOW LEVEL OF DETAIL

H0	There is no significance difference in technique result interpretation time using MUC and IBIS, in situation of low level of detail
H1	There is significance difference in technique result interpretation time using MUC and IBIS, in situation of low level of detail
H2	MUC is greater than IBIS, in terms of taking less result interpretation time in situation of low level of detail
H3	IBIS is greater than MUC, in terms of taking less result interpretation time in situation of low level of detail

TABLE 8. EXPERIMENTAL DESIGN

Situation /Technique	MUC	IBIS
Low level of detail	Group A Group B	Group B Group A
Medium level of detail	Group C Group D	Group D Group C
High level of detail	Group E Group F	Group F Group E

G. Experimental Procedure

The experiment took place in two consecutive sessions on 10 Feb 2011, venue university lab no 25. Each session was of four hours. Each activity of the session was marked with time stamps. The first session started at 8: 45 and comprised of fifteen subjects, randomly assigned to three groups of five subjects each. The participants were seated at a reasonable distance and two invigilators were engaged to curb the peeking and cheating prospects among them.

1. A fifteen minutes tutorial as a power point presentation on security requirements was presented to the three groups. The tutorials and other related material can be found on https://docs.google.com/file/d/0Bz-8NAQloxv3RzIydHlZdEMwSDQ/edit?usp=sharing
2. Group 1 was assigned a scenario of low level of detail, Group 2 was assigned a scenario of medium level of detail and Group 3 was assigned a scenario of high level of detail.
3. A thirteen minutes multimedia presentation on introduction of MUC was presented to the three groups.
4. A learning task description was presented and they were asked to read and produce a diagram of the presented technique.
5. At the end of the learning activity the three groups were asked to develop diagram of the given technique. After which the participants were required to analyze the diagram and identify security goals without any time constraint.
6. Steps 3-6 were repeated for IBIS with the same groups.

The second session took place at 1:15 with fifteen participants divided into three groups of five subjects each. The experimental procedure was the same as described above with only difference of technique order i.e. IBIS was executed prior to MUC.

IV. ANALYSIS AND RESULTS

A. Data Preparation

Measurement of both the techniques for effectiveness and coverage in each situation was coded in order to develop a database of outcome variables for analysis. Individual responses of each participant for each outcome variable were coded and rechecked for inaccuracy by the author. It involved coding of total number of security goals identified by the two techniques in each situation. Time stamps were also coded as technique learning time, technique execution time and technique interpretation time and were rechecked for accuracy.

B. Data Accuracy

The obtained data was screened using normality tests-Kolmogorov-smirnov and Shapiro-Wilk in SPSS, thereby indicating that data of three situations is normally distributed and meet the assumptions of parametric statistical testing class. Since the experimental design was of dependent nature, t-test was used to analyze the difference between the two techniques in given situations. Significance value of 0.05 was selected to assure validity of results as we are prepared to accept 5% of anything happening by chance.

C. Number of Security Goals Identified Using MUC and IBIS in Situation of Low, Medium and High Level of Detail

Figure 1 shows the graphical summary of the effectiveness of the two techniques. In majority of the cases more security goals have been identified using MUC as compared to IBIS. In a situation of low level of detail, 6 cases identify more security goals by using MUC, in 3 cases IBIS score is higher and in one case it is a tie. In situation of medium level of detail, score of 10 cases is higher using MUC as compared to IBIS while in situation of high level of detail, 8 cases' scores are higher using MUC as compared to IBIS. Descriptive statistics shows the similar results with mean differences of MUC and IBIS as (25.7 and 21.8), (29.9 and 18.8) and (34.8 and 18.4) in situation of low, medium and high level of detail respectively.

Statistical findings of number of security goals reveal that in situation of low level of detail, no significance difference was observed (t(9) = 1.56, p>0.05) between two techniques in terms of no of security goals. The calculated value of t(1.56) falls in acceptance region of tabulated value of z±2.26 and value of p is greater than 0.05%. This means that difference between MUC and IBIS is not significantly proven and performance of MUC is equal to that of IBIS when low level of project's detail is available, therefore we accept (H0) null hypothesis. In situation of medium level of detail, significance difference is indicated by (t(9) = 4.65, p<0.05) where calculated value of t (4.65) falls in rejection region of tabulated value of z±2.26 and p is less than 0.05 so we reject (H0) null hypothesis and accept H1 in situation of medium level of detail. Furthermore we also accept H2 as on average, use of MUC identified significantly greater security goals (M =29.90, SE=2.17) as compared to IBIS (M=18.80, SE=1.54, t(9)=4.659, p<0.05). The findings suggest that with medium level of documentation detail, MUC provides high effectiveness in terms of identifying more security goals as compared to IBIS. In situation of high level of project's detail, there is a significance difference between MUC and IBIS as calculated value of (t (9)=2.62, p<0.05) falls in rejection region of tabulated value of z±2.26 and p<0.05. We reject null hypothesis (H0) and accept H1 in situation of high level of detail as use of MUC identified significantly greater security goals (M=34.80, SE=5.42) as compared to IBIS (M=18.40, SE=2.02, t(9) = 2.625, P<0.05). These statistics at high level of detail advocates the selection of MUC as it provides greater effectiveness as compared to IBIS.

D. Number of Types of Security Goals Identified Using MUC and IBIS

A trend of mixed scores is seen in terms of variety of security goals identified in Fig. 2. In a situation of low level of detail 4 cases have higher scores by using IBIS, 3 cases have higher scores using MUC where as 3 cases' scores are a tie. In situation of medium level of detail, 6 cases performed better using MUC, 3 using IBIS and 1 case is a tie. Similarly 3 cases performance is higher using IBIS, 5 cases using MUC and 2 cases remain a tie in a situation of high level of detail. The descriptive statistics shows an inconspicuous difference between MUC and IBIS with a mean difference of (7.5000 and 8.1000), (7.4000 and 6.7000) and (7.000 and 7.5000) in situation of low, medium and high level of detail respectively.

The statistical calculation of t-test (t(9)=0.651, p>0.05) for number of types of security goals shows no significance different between the two techniques in situation of low level of detail. As calculated value of t(-0.65) falls in acceptance region of tabulated value of z±2.26 and value of p = 0.531, we accept (H0) null hypothesis in situation of low level of detail. Similarly we also accept (H0) null hypothesis in situation of medium level of detail where calculated value of t(0.843) falls in acceptance region of tabulated value of p =0.421. Further more in situation of high level of detail, calculated value of t(-0.6) also falls in acceptance region of tabulated value of z±2.26 with p=0.563 so we also accept (H0) null hypothesis. We conclude that no significance difference is found in both techniques regarding the identification of types of security goals.

E. Learning Time Utilization Using MUC and IBIS in Situation of Low, Medium and High Level of Detail

The learning time of the two techniques in situation of low level of detail illustrates that 7 cases have higher learning time using IBIS while 2 cases using MUC (Fig. 3). In situation of medium level of detail 8 cases' show more time consumption for IBIS, 1 for MUC while 1 case shows the same time consumption for MUC and IBIS. In a situation of high level of detail, 9 cases take more time to learn IBIS while 1 case takes more time to learn MUC.

Statistical findings of learning time consumed by both the techniques highlight that no significance difference (t(9)=-1.711,p>0.05) is found in situation of low level of detail. As calculated value of t(-1.71) falls in acceptance region of tabulated value of z±2.26 and p>0.05, we accept (H0) null hypothesis in situation of low level of detail. With the values of t-test (t(9)=-4.672, p<0.05), there is a significance difference between the two techniques in terms of time taken to learn each technique in situation of medium level of detail. So we reject (H0) null hypothesis in favor of H1 in this situation. Furthermore we also accept H2 that concludes that MUC (M=29.500,SE=1.00) is better than IBIS (M=33.00, SE=0.51, t(9)=-4.672, p=0.001) regarding learning time in situation of medium level of detail. Similar findings are observed in situation of high level of detail where statistical difference (t(9)=-4.482,p<0.05) is found between MUC and IBIS. As value of t-test (-40482) falls in rejection region and p=0.05 we reject (H0) null hypothesis in favor of H1, in a situation of high level of detail. We also accept H2 since on average MUC takes significantly less time to be learned (M=29.30, SE=0.84) than IBIS (M=34.30, SE=0.84, t(9)=-4.672, p=0.002). Over all, findings of learning time utilization show that both have an equal chance of selection with output of t-test value of (t(9)=-1.177 and p>0.05)) with (MUC(m=31 & IBIS(m=32.9)). In situations of medium level of detail statistical figures are (M=29.500, SE=1.00) than IBIS (M =33.00, SE=0.51, t(9)=-4.672, p=0.001) while the statistic in situation of high level of detail (M=29.30, SE=0.84) than IBIS (M=34.30, SE=0.84, t(9)=-4.672, p=0.002) indicating that performance of MUC and IBIS have significant difference as situation of project changes from low level of detail to medium or high level of detail . It also shows that MUC is more effective in terms of learning time consumption as compared to IBIS in situation of medium and high level of detail.

F. Analysis of Execution Time Utilization Using MUC and IBIS in Situation of Low, Medium and High Level of Detail

The graphical summary given in Fig. 4 of the execution time in the situation of low, medium and high level of detail does not provide a clear estimation of the comparison. Descriptive statistics shows mean of MUC (21.9) is higher than IBIS (19.9) in situation of low level of detail, mean of MUC (25.2) is higher than IBIS (19.2) in situation of medium level of detail and mean of MUC (23) is also higher than IBIS (19.9) in situation of high level of detail. The analysis of t-test reflects the same trend with the value of (t(9) = 0.436, p>0.05) in low level of detail, (t(9) = 1.87, p>0.05) in the medium level of detail and (t(9) = 0.75, p>0.05) in high level of detail. In conclusion we accept null hypothesis (H0) with the values of (t(0.436), p=0.673), (t (1.87), p=0.094) and (t(0.075, p=0.472)) in the situation of low, medium and high level of detail respectively.

G. Analysis of Result Interpretation Time Utilization by Using MUC and IBIS in Situation of Low, Medium and High Level of Detail

Figure 5 reveals that in the situation of low level of detail MUC takes more interpretation time in 9 cases while one case is a tie. Similarly the means of MUC in descriptive statistics is higher (13.4) than mean of IBIS (8.7). In situation of medium level of detail MUC score is higher on 8 cases while one case scores more in IBIS than in MUC. One case is a tie. In situation of medium and high level of detail we see similar results as mean difference for MUC and IBIS is (12.6 and 8.1) and (12.3 and 7.8) respectively. It is noted that seven cases scores higher using MUC, two cases using IBIS where as one case both techniques have same score in situation of high level of detail.

Statistical findings in a situation of low level of detail show that the value of t(4.92) falls in the rejection region with p<0.05% so we reject null hypothesis (H0) in the favor of H1. We also accept H3 since MUC takes significantly more time to be interpreted (M = 13.4, SE = 0.858) than IBIS (M = 8.7, SE = 0.53, t(9) = 4.921, p<0.001). The results show that IBIS seems to be a better choice than MUC as it takes less time to be interpreted. The medium level of detail shows significant difference between the two techniques as value of t(3.28) falls in the rejection region with p<0.05 %. We reject null hypothesis (H0) and accept H1 in this situation. Furthermore we also retain H3 since MUC takes significantly more time to be interpreted (M = 12.6, SE = 0.63) than IBIS (M = 8.1, SE = 1.1, t(9) =3.28, p=0.009). Statistical findings in situation of high level of detail reveal that value of t(2.63) falls in the rejection region with p<0.05% and MUC takes more time (M=12.30, SE = 1.1) than IBIS (M=7.8, SE = 0.92, t(9) = 2.63, p=0.027) so we reject null hypothesis (H0) in favor of H1and also accept H3 in situation of high level of detail. In both medium and high level of detail IBIS is considered to be a better choice of selection as it takes less time to be interpreted.

V. DISCUSSION

The primary contribution of this research is to comparatively evaluate MUC and IBIS in terms of coverage and effectiveness under three situations. This comparative evaluation can serve as a guideline to establish whether a respected technique (MUC or IBIS) is pertinent to a given situation of low, medium and high level of detail and hence can help a security analyst in selection of the proper technique. The Table 9 shows the summary of results of our findings, grey area depicts no statistical difference between the two techniques while ✓ means the corresponding technique is greater as compared to the other technique in terms of performance under relative outcome variables. It can help to serve as guidelines for security analyst. Analysis of the Table 9 reveals that difference does exist between performances of both techniques on the scale of outcome variables in three different situations. MUC takes less learning time and captures more number of security goals as compared to IBIS in situations of medium and high level of detail while IBIS takes less time to be interpreted in all the three situations of low, medium and high level of detail. This motivates us to develop guidelines for the security analysts which are as follows:

1. Consider security requirement elicitation at requirement engineering level.
2. Consider two techniques: MUC and IBIS.
3. Review Table 2 to recognize the situational characteristics that will help in selection of appropriate security requirement elicitation technique.
4. On the basis of step 3, identify project specific situational characteristics by using Table 2 as guideline whether it's low, medium or high level of detail.
5. Select appropriate technique according to the technique's capability in given situation using Table 9.

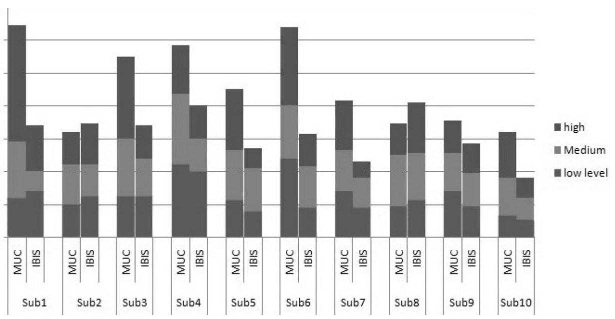

Fig. 1. Graphical summary of no of security goals identified

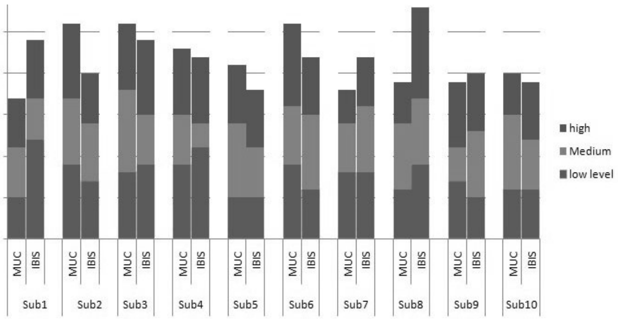

Fig. 2. Graphical summary of no of types of security goals

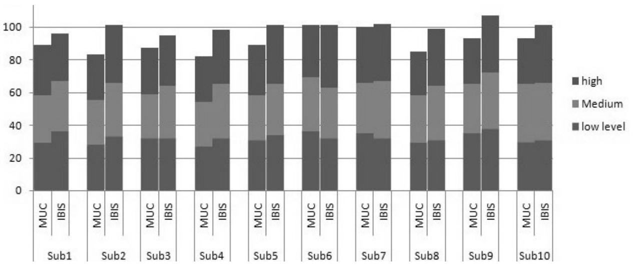

Fig. 3. Graphical summary of learning time utilization

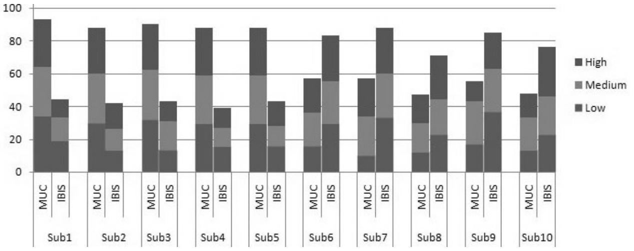

Fig. 4. Graphical summary of execution time utilization

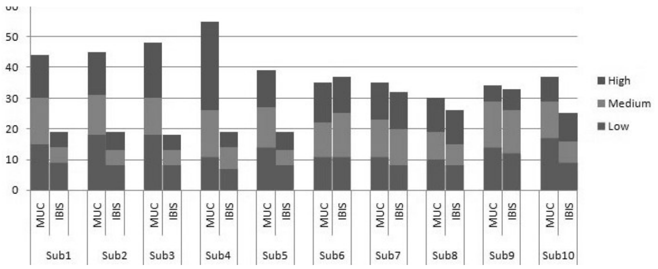

Fig. 5. Graphical summary of result interpretation time

TABLE 9. RESULTS

Dependent Variables	**Low Level**		**Medium Level**		**High Level**	
	MUC	IBIS	MUC	IBIS	MUC	IBIS
No of security goals			✓		✓	
No of types of security goals						
Learning time			✓		✓	
Execution time						
Result interpretation time		✓		✓		✓

VI. THREATS TO VALIDITY

A. Internal Validity

Internal validity appraises whether the observed outcomes were due to the treatment or due to some other factors [1]. Since 100% control of all threats to internal validity is not possible, we take into account several known threats. Selection bias is mitigated by random selection and random assignment of participants, thereby reducing the prospect that the participants in one group were better than the ones in the other on intellectual grounds. Another potential threat, learning effect was controlled by counterbalancing. Furthermore diffusion of treatments or compromised participant independence was taken into account by engaging two invigilators and diminishing the communication among participants. Another internal validity problem is history or previous knowledge of the participants that can modify the results. The participants were studying in the final years of their four years software engineering bachelor degree program. They had taken Requirement Engineering course in their first year but had not taken any courses specific to security which could familiarize them with MUC or IBIS. So we can safely anticipate that they were at the same level as far as their subject knowledge was concerned. The content bias was controlled by providing the same material of the technique introduction and task description. Since the total time of each session was 4-5 hours so we can safely claim that out internal validity is not influenced by time pressures.

B. External Validity

External validity is the degree to which the conclusions in the study would hold for other person in other places and other time [21]. Using repeated measure design we divided the sample in 2 groups for each situation and conducted the experiment with different subjects in each group. All the groups in different situations followed the same activity procedure; furthermore counterbalancing tool was used to eliminate order and learning effects. Participants were asked to reference each identified security goal to the relevant point of the diagram. This strategy helped in making sure the participants have genuinely identified the security goal rather than just general text book based security goals based on their intuition. The participant's motivation was kept fueled with promising advent of refreshments at the end of the session. Also we conducted the experiment during the semester break and kept the participants convenience in mind by choosing the time and date of the experiment such that they are free from their studies and other exam workload.

C. Construct Validity

Our experiment which was undertaken by giving situation based scenarios and technique introduction tutorials cannot mimic a real project or conceptual framework hence, it could not possibly explore the full breadth of either technique. But this could apply to both the technique and balance out any effects on the observation made on effectiveness of the technique. Apart from that much care was taken while delivering the technique introduction tutorial and queries were entertained during the presentation and in the following activities of the session. Session time was set around 4-5 hours to provide flexible time frame so that participants can perform by considering all factors of the situations. The contents of the situation based scenarios were developed while keeping important factors in mind so that a complete version of the relevant situation can be provided.

D. Conclusion Validity

"*A threat to conclusion validity is a factor that can lead one to reach an incorrect conclusion about a relationship in ones observations*" [21]. The pilot study that we conducted highlighted the need for inclusion of example diagrams of both techniques and for a system from a simpler domain. As a result of which we added examples of the two techniques in the tutorial and provided them in a manner that was easily read and understood by the participants. In order to remove "poor reliability of treatment implementations" as identified by [21] we followed a predefined experimental procedure and accurately recorded time stamps for each activity. We provided the participants with a comfortable environment to work in so that we could circumvent "*relevant irrelevancies in the settings*"[21]. Furthermore, personal possessions of the subjects were controlled by selecting participants randomly, from final year students of 4 years degree program. Credibility of results was also assured by considering 0.05 significance value to decide whether results are significantly proven or not.

VII. Conclusion and Future Work

This research work has undertaken an experimental evaluation of two security elicitation techniques; MUC and IBIS with the intention of developing guidelines for current practitioners and future security analysts. The evaluation was based on coverage and effectiveness of each technique in the situations of low, medium and high level of detail. The experimental findings disclose that in situation of low level of detail, MUC and IBIS showed no significance difference in terms of effectiveness which is measured by the number of security goals found. However in coverage of the technique IBIS showed better performance in one attribute i.e. in interpretation time consumption while no significance difference was found in the other two attributes. In the situation of medium and high level of detail it is observed that there is no significant difference in terms of technique execution time consumption. While in terms of number of security goals found, MUC's performance surpasses that of IBIS. Also it is found that result interpretation time for MUC is higher compared to IBIS, thereby making it the preferred technique to be selected in the respected situation. Current practitioners and future analysts may take these findings as guideline regarding selection of MUC and IBIS. This also suggests some future pathways in order to bring about firm empirical grounds for the selection of best techniques according to the situation at hand. Earlier evaluations fall short on considering the situational characteristics and do not attempt to provide guidelines for the industry. More experiments and empirical evaluations should be carried out using different techniques in an industrial set up using practitioners as subjects.

References

[1] A. L. Opdahl and G. Sindre, "Experimental comparison of attack trees and misuse cases for security threat identification," *Inf. Softw. Technol.*, vol. 51, no. 5, pp. 916–932, 2009.

[2] M. H. Diallo, J. Romero-Mariona, S. E. Sim, and D. J. Richardson, "A comparative evaluation of three approaches to specifying security requirements," in *12th Working Conference on Requirements Engineering: Foundation for Software Quality, Luxembourg*, 2006.

[3] L. Chung, F. Hung, E. Hough, and D. Ojoko-Adams, "Security quality requirements engineering (SQUARE): case study phase III," 2006.

[4] D. Mellado, E. Fernández-Medina, and M. Piattini, "A comparative study of proposals for establishing security requirements for the development of secure information systems," *Comput. Sci. Its Appl.-ICCSA 2006*, pp. 1044–1053, 2006.

[5] J. Gregoire, K. Buyens, B. D. Win, R. Scandariato, and W. Joosen, "On the secure software development process: CLASP and SDL compared," in *Proceedings of the Third International Workshop on Software Engineering for Secure Systems*, 2007, p. 1.

[6] A. Herrmann, D. Kerkow, and J. Doerr, "Exploring the Characteristics of NFR Methods–a Dialogue about two Approaches," *Requir. Eng. Found. Softw. Qual.*, pp. 320–334, 2007.

[7] I. A. Tondel, M. G. Jaatun, and P. H. Meland, "Security requirements for the rest of us: A survey," *Softw. IEEE*, vol. 25, no. 1, pp. 20–27, 2008.

[8] B. Fabian, S. Gürses, M. Heisel, T. Santen, and H. Schmidt, "A comparison of security requirements engineering methods," *Requir. Eng.*, vol. 15, no. 1, pp. 7–40, 2010.

[9] M. U. A. Khan and M. Zulkernine, "On selecting appropriate development processes and requirements engineering methods for secure software," in *Computer Software and Applications Conference, 2009. COMPSAC'09. 33rd Annual IEEE International*, 2009, vol. 2, pp. 353–358.

[10] R. Araujo, "Security requirements engineering: A road map," *Secur. Feature July*, pp. 1067–7, 2007.

[11] J. Romero-Mariona, H. Ziv, and D. Richardson, "Security Requirements Engineering: A Survey," Technical Report UCI-ISR-08-2. University of California, Irvine, 2008.

[12] T. Stålhane and G. Sindre, "Safety Hazard Identification by Misuse Cases: Experimental Comparison of Text and Diagrams," *Model Driven Eng. Lang. Syst.*, pp. 721–735, 2008.

[13] "The what and whence of issue-based information systems," *Eight to Late*. [Online]. Available: http://eight2late.wordpress.com/2009/07/08/the-what-and-whence-of-issue-based-information-systems/. [Accessed: 17-Nov-2012].

[14] J. Conklin, "Growing a Global Issue Base: An Issue-based Approach to Policy Deliberation," *Dir. Implic. Adv. Comput.*, 2008.

[15] G. Sindre and A. L. Opdahl, "Eliciting security requirements with misuse cases," *Requir. Eng.*, vol. 10, no. 1, pp. 34–44, 2005.

[16] I. Alexander, "Modelling the interplay of conflicting goals with use and misuse cases," in *Proceedings of 8th International Workshop on Requirements Engineering: Foundation for Software Quality (REFSQ'02)*, 2002, pp. 145–152.

[17] I. Alexander, "Misuse cases: Use cases with hostile intent," *Softw. IEEE*, vol. 20, no. 1, pp. 58–66, 2003.

[18] L. Røstad, "An extended misuse case notation: Including vulnerabilities and the insider threat," in *XII Working Conference on Requirements Engineering: Foundation for Software Quality, Luxembourg*, 2006.

[19] J. Conklin, "Dialog mapping: Reflections on an industrial strength case study," in *Visualizing argumentation*, Springer, 2003, pp. 117–136.

[20] A. Field, *Discovering statistics using SPSS*. Sage Publications Limited, 2009.

[21] Trochim, William M, "Research Methods Knowledge Base." [Online]. Available: http://socialresearchmethods.net/kb/index.php. [Accessed: 17-Nov-2012].