WeDRisk Approach for Management of Web and Distributed Software Risks: Evaluation Experiment (Part 1)

Ayad Ali Keshlaf
Faculty of Engineering,Computer Engineering Dept.
Sabrata University, Sabrata, Libya,
ayadali2005@gmail.com

Abstract

WeDRisk is an approach designed to manage risks of web and distributed software development. The WeDRisk approach aims to address the weaknesses of existing approaches to risk management, which are less able to deal with the specific challenges of web and distributed development. A key part of the approach is flexibility, to deal with the rapid evolution which is typical of such developments. This paper presents the approach and describes the design, execution and analysis of a controlled experiment that has been used to evaluate some central parts of the approach, in terms of their usefulness, ease of understanding and usability. The paper focuses on the evaluation of estimation module part. Other parts will be introduced in next papers. The experiment result illustrated how the estimation module is useful, understandable, flexible easy to use, and considers web and distributed development risks factors.

KeywordsWeb and Distributed software risks, Risks Clustering, Risks Estimation, Evaluation Experiment.

1. INTRODUCTION

SOFTWARE industry is continuously evolving and it becomes more and more vulnerable to new challenges and risks. Web and Distributed (WD) software development is an example of that. Due to the development environment and technologies acceleration WD development faces a new set of challenges and risks (e.g. evolving, lack of face to face meetings, time zone difference, cultural differences and communications failures [1], [9], [4] and [5]. These risks and challenges are difficult to be managed with traditional management approaches, which already they have many weaknesses such as; they are not flexible as they offer only one type of Risk Management (RM), and they do not consider WD factors that affect WD risks (i.e. dependencies, sites distribution and communication) [2]. Indeed, there is a need for developing new flexible and evolvable methods, approaches and tools to accommodate the risk management needs for the WD development. WeDRisk [1], [3] is an approach that has been designed for this purpose. It is proposed in order to tackle some weaknesses of the existing software risk management approaches in managing the WD development risks. While the approach is particularly aimed towards WD development, it should be applicable to the modern software developments in general. The approach is designed to be flexible, customizable and able to evolve. Moreover, it considers the risks from three perspectives (project, process, and product).

In additional to the introduction section the paper structure consists of five other sections which include; Section II introduces WeDRisk approach structure and the module which is targeted for the evaluation. In Section III the evaluation experiment scope, questions, hypothesis background and related work, method (e.g. apparatus, material, and subjects), injected situations, validation, controlling measures design, result and analysis are presented. Paper discussion is in Section IV, limitations are in Section V and finally, Section VI introduces the conclusion and future work.

2. WEDRISK STRUCTURE

Figure 1 shows the main structure of the WeDRisk approach, which, consists of three phases namely; RM Establishment phase, RM Implementation phase and RM Evaluation and Evolution phase. It is also provided with a communication channel. The phases consist of modules, which contain components, steps, techniques and guidelines [1], [3].

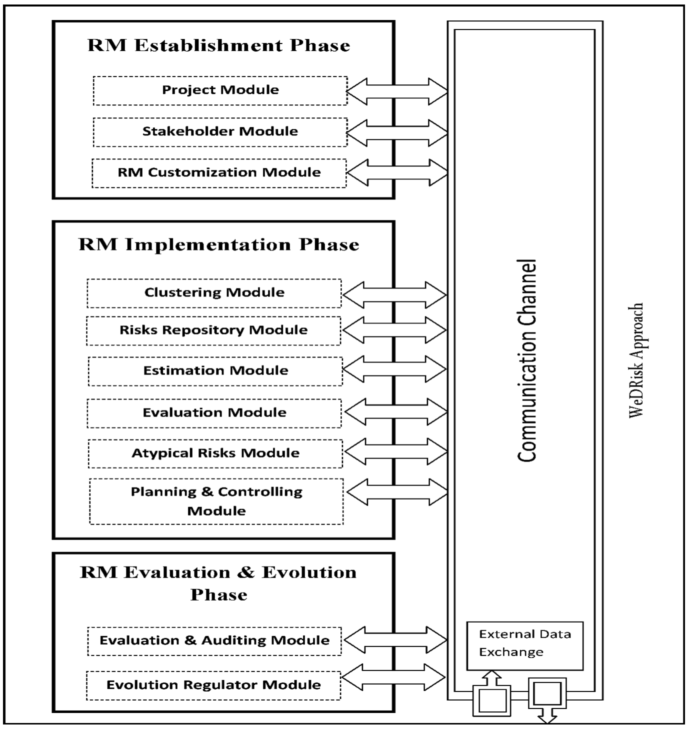

Figure 1:WeDRisk Approach Main Structure [1], [3]

Parts of this paper appear in a PhD thesis [1]. The paper reports a controlled experiment which has been conducted at School of Computing Science - Newcastle University- UK to evaluate some WeDRisk approach modules. As this paper focuses on the estimation module, therefore, the only this module is described in the following section:

A. Estimation Module

The estimation module in the WeDRisk approach offers two options for risk estimation. First option uses the ordinary Risk Exposure (RE) equation (1), whereas, the second option uses an improved equation named as Total Risk Estimation Value (TREV) equation (2), which, is intended to include Web and Distributed Factors (WDF). The module uses a special estimation matrix to estimate the total value of WDF. Following section describes the two equations and the WDF estimation matrix.

1) Risk Exposure (RE): RE is a famous equation (1) that has been used for many years to estimate software risks. It depends on the estimation of the probability of the risk occurring (RiskProb) and magnitude of the losses if the risk is occurred (RiskMag).

$$RE = RiskProb * RiskMag \quad (1)$$

There are different ways (qualitative and quantitative methods) to estimate the probability and magnitude of the risks. In our estimation module we tried to mix between the qualitative and quantitative estimations. For this purpose, a ranked line as shown in Figure 2 is used by the WeDRisk estimation module.

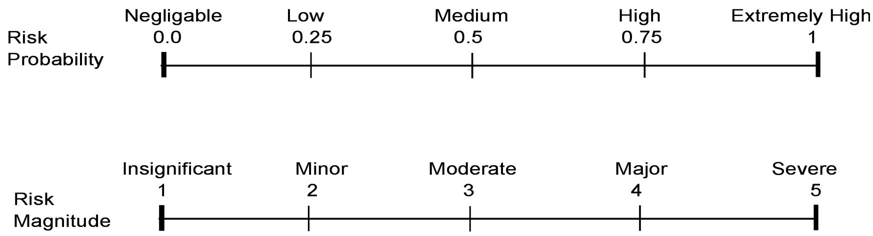

Figure 2: Probability and Magnitude Estimation Line [1]

As can be seen in Figure 2 the probability is ranged from Negligible to Extremely High with values from 0 to 1. On the other hand, the values for risk magnitude are ranged from 1 when the risk associated loss is Insignificant to 5 when it is Severe. This technique is used for its simplicity of use and at the same time it represents the ascending sequence of the probability and magnitude of the risk with quantitative values.

RE equation was used for the assessment of collocated software development risks since the late of 80s [6], [7]. However, the software industry is evolutionary industry and rapidly growth especially with the new phenomenon of WD software industry, which involve a new set of factors that could have an effect on the risks and need to be considered by the estimation equation. Therefore, we believe that the RE equation could be improved by including the WD factors. The TREV is an attempt to produce an improved equation for this purpose.

2) Total Risk Estimation Value (TREV): In additional to the two main aspects of the risk (Probability and Magnitude) the TREV equation (2) includes the WDF as a part of the question [1]. In fact, this equation is an improved equation based on the RE equation to estimate WD development risks. A special matrix is designed to estimate the total WDF.

Both the TREV equation and the WDF estimation matrix are described below.

$$TREV = RE * \sum WDF \qquad (2)$$

Where,RE is the Risk Exposure and WDF is the Web and Distributed Factor

3) WDF Estimation Matrix: The WDF estimation matrix (see Table I) estimates three WD factors (Sites Dependency, Sites Distribution and Communication Availability), which, could have considerable and changeable affect on risk exposure. For instance, the importance of a risk could be changed if we considered the dependency level on other sites or risks. The WD factors and used ranking system in the matrix are described below [1]:

I) Sites Dependency Level

In distributed development, one site progress could depend on other site progress. This means any delay (e.g. due to a risk) in a site affects other dependent sites. This will have worst affect when there is a large number of sites depending on each other or there is a cross dependency between them. Usually, the dependency is not considerable issue during the estimation of the probabilities and magnitudes of the risks, as the developers do not see the big picture of the relations between the sites and they just deal with the risks individually.

II) Sites Distribution

The number and the distribution of development sites have an important influence on the risks in terms of type, number and significance. Multisite projects, which have sites in different countries are much more vulnerable to the distribution risks (e.g. time zone and cultural differences) than the ones, which are multisite but in one country. Therefore, we need to consider this factor during the estimation of risks. See how the WDF estimation matrix Table I ranks that.

III) Communication Availability

Communication plays a vital role in web and distribution development. Occasionally, availability and reliability of communication are different from situation to another and from time to another. Therefore, the effect of communication on the risks needs to be considered as a part of risks estimations.

The above three WD factors have been resulted from reviewing of challenges and risks of the WD development. Summary, these factors are selected and considered due to the following reasons:

- We believe that the consideration of these factors could change the importance of risks priorities.
- These factors are changeable from risk to risk, from situation to situation and from time to time.
- Developers/Managers should not just deal with the WD development risks individually; they should have a pig picture about their relations and dependencies.

IV) Matrix Ranking Technique

The ranking values in the matrix increase gradually starting from 1 and ending with 5. One means the factor has a very low or negligible affect where the value five means that it has the highest negative effect on the risks. The given values are assigned based on the importance of each level and its affect on the risks. Each factor might have different values in different times. In the case when there is no any change in any of the factors the ranking value will take a default value, which is one. At the end the total of the assigned values will be added to each other to obtain the total WDF value for the desired risk (see Table I).

The aspects that are targeted for the evaluation in this study include; consideration and coverage of WD factors, usability and usefulness of WDF estimation matrix and the TREV equation.

3. THE EXPERIMENT

A. Problem Being Solved

This experiment aims to evaluate some novel aspects of the WeDRisk approach, mainly, the consideration and estimation of web and distributed factors. The experiment is divided into three sections. Each section tests different hypothesis(es) to evaluate one of the WeDRisk modules. The idea behind this experiment is to present (inject) some distribution and none distribution risky situations and then asking the subjects to deal with these risky situations. Based on the evaluation needs the subjects could be divided into control and experimental groups. Experimental group subjects use the WeDRisk modules to deal with the injected risky situations; whereas, control group subjects depend on their knowledge and experience [1].

The experiment methodology was chosen to evaluate the WeDRisk modules due to the following reasons [15], [12]:

- To emulate the same working environment for all the subjects.
- Many observations, measures and support works need to be done or provided during the evaluation of the modules and the experiment is the best option for that.
- It is good for focusing on specific variables, measures, and the relationships between them with extra flexibility in asking questions.
- Novelty of the evaluated modules, so that we could not find suitable data set form the previous approaches that can be used to evaluate these modules. Also due to the imposed restrictions on data by the developers. However, there are some limitations associated with the use of the experiment technique, which, are pointed out under limitations section.

B. Experiment Scope

This experiment is mainly designed to evaluate three WeDRisk modules (estimation, customization and atypical).

C. Questions and Hypothesises

The experiment is mainly designed to answer the following three questions [1]:

Q1: What is the coverage and consideration of WD factors by the WeDRisk approach?
Q2: How easy is it to understand and use the WeDRisk approach?
Q3: How usable and helpful are the evaluated modules?

In order to answer these questions, five hypothesises were tested in this experiment to evaluate the three WeDRisk modules (estimation, customization and atypical). Hypothesises H1 and H2 were used to evaluate the estimation module, which are:

H1: The TREV equation is an ideal option to estimate the WD development risks and to consider the WD factors compared with the RE equation.
H2: The WDF estimation matrix is useful, understandable and helpful to estimate and consider WD factors.

TABLE I : WDF ESTIMATION MATRIX [1]

	Factors Levels				
WD Factor	***1**	***2**	***3**	***4**	***5**
Sites Dependency Level n=1	NO Dependency (D)	Low D. Affects One Node	Medium D. Affects One Node + It is Cross D.	High D. Affects Multi Nodes	Very High D. Affects Multi Nodes + it is Cross D.
Sites Distribution n=2	1 site	>1 site but in the same city	Sites are in different cities but at the same country	Sites are in different countries but at the same continent	Sites are in different cities, countries and different continent
Communication Availability n=3	Excellent 24/7/12 available, excellent history and infrastructure	Good Go od History and infrastructure. Very rare to face problems	Acceptable The history and infrastructure are fine but there is a very small chance for problems	Bad Faces problems from time to time and either the history or infrastructure are bad	Totally Unavailable Currently not available and both the history and infrastructure are very bad
Sub Totals = (No. of Ticks * Factor Level)					
WDF =	$\sum SubTotals$				

D. Related Work

WeDRisk approach has been developed to manage WD risks and to tackle some of the existing approaches weaknesses. A number of evaluation cycles including case studies, experts and experiments [8] were conducted to evaluate the usefulness and usability of WeDRisk approach and its coverage to the WD development RM needs. The design (e.g. hypothesis, subjects selection, data collection, biases avoiding, validation, procedures tasks and instructions) of this experiment is inspired by the design of some software engineering experiment and empirical studies and evaluation techniques in [9], [11], [12], [13], [14].

E. Method

The method of this experiment can be summarized as follows:

1) Apparatus: The apparatus that were used in this experiment include normal stationary, hard copies of the experiment material, data collection forms and sport watch. A computer was used for data saving and Minitab tool was used for statistical analysis purpose.
2) Materials: Hard copies of all required information and forms were prepared and provided on equality base for all the subjects.
3) Subjects: We recruited by email about 35 subjects to participate in this experiment. Of these the researcher selected 24 subjects to participate in this experiment as they met the conditions (experience and knowledge in software development). The subjects were researchers, PhD student at School of Computing Science/Newcastle University, some visitor students and researchers at Center for Software Reliability - Newcastle University. The subjects either they were involved in software projects or had at least attended software engineering courses.In order to improve the design of the experiment a pilot study was carried out before starting the real experiment sessions. The pilot study was conducted using another group of participants. A number of issues were addressed and improved for the real experiment including; tasks required time estimation, tasks sequence, instructions to subjects, data collection procedure and risky situation design.
4) Data Confidentiality: All subjects' bibliographic data and the collected data are confidential to the experimenter and his supervisor. The only use of the data is for the research purpose.All subjects were given reference names and their actual data is just used for providing the 10 Amazon vouchers as compensating for their time.
5) Variables: The dependent variable of this experiment was the injected situations whereas; the independent variable was the subjects' reactions.

6) Measurement Units: Table II defines the measurement units, which are used in this experiment.

TABLE II: EXPERIMENT MEASUREMENT UNITS

Unit	Definition	How it is measured
Used Time	A duration of time that spent to implement as specific task	Difference between starting time and ending time of the task implementation
Effort	The exertion spent to implement a specific task or achieve the goal	Observations, used time comparison, asking specific questions, feedbacks and tries to implement the task

F. Injected Risky Situations

Emulating the management of WD development risks in this controlled experiment was not easy task. In this regard, we tried carefully to design a number of risky situations to inject them during the running of the experiment phases with consideration to:

- The situation should cover the evaluation aspects of the modules.
- It can be reused to evaluate more than one module.
- Working independently (standalone).
- It should reflect the real risky situation as possible as can.
- It should be simple, self-explanation, consistent and understandable.
- It must be short to avoid any boring and wasting time.

For these purposes the designed situations were tested many times and improved based on the comments and notifications taking into the account the reading time. Generally, the evaluation of the modules is almost independent operation. Only the injected situations were shared (to reduce the reading and understanding time) without any considerable affect on the evaluation operation. Therefore, in order to avoid any biases the situations were injected randomly with a time gap between them. The injected situations covered collocated development and WD development and they were given reference numbers for the use during the experiment stages.

G. Generalization and Threats Validity

1) Generalization: A number of measures have been taken in order to make the experiment sample reflect the real population of real WD development. Mainly, the selected sample was concentrated on a set of subjects who are working or worked in the field of software development or have attended software engineering courses. In fact, all of the recruited subjects were either software researchers who are working in WD software development projects or PhD students who are involved also in software researches and had experiences in related projects. The selected sample of subjects is intended to reflect the real software projects. However, it is costly and almost impossible to cover all software development population in this experiment. Different evaluation techniques were used to evaluate the WeDRisk approach including case studies, other experiment and expert evaluations in order generalize as much as can.
2) Validation: The experiment validity is an important issue to insure the quality and generalization of findings. Two types of validity are involved in this experiment. Internal validity, which is concerned with the study supports to the findings and the external validity which is concerned with the generalization of the results [9], [10]. The threats to internal

and external validity are addressed and taken into account as follows:

– Internal Validity

Selection: The assigning of the subjects to the experimental and controlled groups were done on random bases, as well as the risky situation were also injected randomly.

History: The subjects were selected from the same place, therefore, in order to reduce the influence affect we recruited and contacted them individually and they performed the tasks individually in different times.

Motivation: As the implementation of the tasks does not take a long time, so that there was not much concern about boredom or losing of enthusiasm during the experiment.

Time: It is expected that the subjects might perform the tasks in harry, as they afraid that the time is not enough which could affect the taken decisions and the results. Thus, in the pilot study we estimated the required time for the experiment and we informed the subjects about the estimated time when they were recruited to consider that. Moreover, during the experiment, they have been told to take enough time when they perform the tasks and also they can stop if they are not willing to continue.

Training: Before performing the tasks in the experiment, a brief description is given to the subjects. Moreover, enough clarification and training were provided before each section of the experiment. Also the subjects have been told that they have the right to ask any questions during the experiment.

- External Validity

Subjects: The difficulties of generalizing from students to professionals have been taken into account. Therefore, the subjects were mixed (students and researchers) who work in software projects. In fact, the use of students as subjects in this experiment may not have that threatens since most of the students were PhD students who had experience with software projects and have almost some sort of professional abilities.

Tasks: We tried to reduce the number of tasks as much as possible. Environment: The experiment environment tried to emulate the real development environment (developer, place, project and risks).

H. Controlling Measures

The controlling measures in the experiment were taken in order to reduce the involved biases and keeping everything the same except the tested variable. They are also to make the experiment ready for any replication. The controlling measures include the following aspects:

Environment: The environment was controlled because we want to record the time consumption individually, avoid the affect of subjects on each other, provide the help on equality bases, observe subjects progress during the tasks implementation, to give the same chances of time and support to all of the subjects.

Injected Situations: The same injected situations were used by all of the subjects but in random order and using reference numbers. The injection of the situations is controlled by the experimenter to avoid any biases.

Tasks Sequence: The sequence of tasks is maintained by the experimenter during all experiment stages, but if any subject refused to continue at any stage then there won't be any affect on the other stages and the data from the implemented stages can be considered as a part of the experiment result.

Provided Support: All of the essential information, support and training material were provided on equality bases.
Control Group: The control group were used in this experiment as a part of controlling measures and for results comparisons.

I. Experimenter and Subjects Tasks
At the beginning of the experiment, the experimenter gave a brief description about the experiment stages and assigned tasks, and then the subjects read, filled and singed the consent from. After that the experimenter requested the subjects to understand and perform the assigned tasks (e.g. reading the injected risky situation and using the provided modules) and also asked them to be accurate as they can. The subjects were also told that they have the right to ask any related questions to get the required clarification and they have the right to stop at any time if they are not comfortable. At the end of each section the subjects were asked to give their feedbacks and provide any suggestions or comments that could improve the desired module. Generally, during the experiment the experimenter has to do other tasks which include:

- Providing the training, related material and required support on equality base
- Injecting the risky situation on random base.
- Managing tasks sequence during the experiment.
- Collecting the data, observing experiment progress, assigning the tasks and recording the time. Distributing the subjects into control and experimental groups (on random base).

Subject tasks: In additional to the common subjects tasks (e.g. reading and signing the consent forms), the subjects have to perform some specific tasks to evaluate the modules, which include: specifying a suitable estimation equation (RE or TREV) for each injected risky situation and then using that equation to estimate the involved risk.

J. Biases Avoiding

Biases are always expected in any experiment, following are the measures that were taken to avoid the potential biases during this experiment:

- Randomizing the subjects grouping and the risky situations injection.
- Providing the required information, material and support on equality base.
- In order to avoid any subjective answers, the subjects were implemented the tasks without any time pressure.
- All of the subjects were able to clarify and ask any related questions and get the needed answers.
- The subjects and injected situations were given reference numbers to avoid the consequence of remembering them.
- The subjects have freely expressed their comments and suggestion at the end of each module evaluation.

K. Procedures

This section describes the experiment design and the procedures to evaluate the estimation module. All the tasks were implemented individually by the subjects (one subject in each session).

Three main tasks that were performed by the subject in this stage (specifying a suitable estimation equation, estimating the involved risk using that equation and giving a feedback). These tasks are described in details below:

I) Specifying a Suitable Estimation EquationFour prepared risky situations were randomly and individually injected by the experimenter. Then the subjects were asked to specify which estimation equation (RE or TREV) is suitable to deal with each situation. The subjects read the injected situations and based on their understanding and experience indicated a suitable estimation equation to estimate the involved risks. Regardless, the selected equations, the subjects were also requested to justify their decisions by writing down the reasons behind their selections. The experimenter monitored the implementation progress of this task and recorded the time that used to take each decision. After the implementation of this task the subjects were asked to declare any comments or suggestions about the equations.

II) Estimating Using the Specified EquationSince the estimation equation is specified for each risk, the next step is the risk estimation using the specified equation (RE or TREV). Thus, if the decision is the RE equation, then the subject needs firstly to estimate the probability and magnitude of the risk and then apply the ordinary RE equation to estimate the risk. On the other hand if the decision is the TREV equation, then the subject estimates the total WDF value for the risk before using the TREV equation to estimate the risk. The implementation details for the both equations (RE and TREV) are provided below:

- RE Equation (Eq. 1): In order to estimate the RE value (Eq. 1), the subject needs to read the injected risky situation first and then uses the line estimation technique (Figure 2) to estimate the probability and magnitude for the involved risk. Based on his ranking on the lines he can obtain the associated values for the probability and magnitude. Subsequently, RE is obtained by multiplying the probability and magnitude values (see Eq. 1). The minimum value of RE is zero, which happens only when there is no any chance for the risk to be occurred and the maximum is five when the risk is certain and its magnitude is severe.

- TREV Equation (Eq. 2): Before applying the TREV equation the subject needs to estimate the RE value using the same above technique and then estimates the second part of TREV equation, which is the ∑WDF (see Eq. 3) (the second part of TREV equation) using the WD estimation matrix (See Table I on page 3). Subsequently, the subject applies the TREV equation to get the TREV estimation for the desired risk.

$$\sum WDF = \sum_{n=1}^{3} (ColTicksNo * Factor - Level\ n) \qquad (3)$$

Where, n is the number of WD factors
ColTicksNo is the number of ticks in each column
Factor-level is the rank of the WD factor during the estimation time

Subsequently, the subject applies the TREV equation to get the TREV estimation for the desired risk.

III) General Feedback

Once the estimation task is completed for the four injected risky situations, then the subject is requested to give his feedback about the estimation operation in general. Four guided questions are designed for this purpose (see Table V). The subject can also write any comments or suggestions that might help to improve the estimation module.

L. Result and Analysis

Hypothesis H1 is used to evaluate the estimation module as following:

Hypothesis H1: The TREV equation is an ideal option to estimate the WD development risks and to consider the WD factors compared with the RE equation.

Three tasks were implemented by the subjects to test hypothesis H1, which are; specifying a suitable estimation equation (RE or TREV), justifying their decision and giving a feedback regarding the suitability of TREV or RE for the WD development. As can be seen in Table III, the result demonstrated that there is a consensuses in the taken decisions (e.g. for the situation number 6004 there were 17 subjects agreed to use the RE against only 6 for TREV and for the situation referenced 0901 there were 22 subject agreed to use TREV against only 1 for RE). Chi-square is a test that used to determine whether there is a significant difference between two or more frequencies in one or more categories. Therefore, Chi-Square (x2) test is used and as a result there is an evidence to support that the proportion of subjects who selected RE equation in situations numbered 6004 and 1072 is significant higher than the subjects who selected TREV equation in the same situation with Chi-Square (x2) of 23 with p-value of 0.001 (<0.05). In contrast, it is the opposite in situations numbered 0901 and 8033, the number of subjects who selected TREV is significant higher than the subjects who selected RE. In fact, the result complies with what we have expected regarding the suitability of RE or TREV equations for each situation. Generally, the time that was used to decide what type of estimation is suitable for the situations was slightly higher in some situations due to the differences in text long. The justifications of the subjects' decisions (see Table IV) are concentrated on the number of sites, involvement of WD factors or complexity of the situation. They justified the selection of RE with; it is a single site or less involvement of WD factors or

simplicity of the situation and the opposite is correct for the TREV equation.
The subjects were asked (To what extant do you agree that TREV is more suitable than RE for WD risks estimation?). As can be seen in Table V the result showed that most of the subjects were agreed or strongly agreed with the suitability of the TREV equation for the WD development risks. Meanwhile, the subjects are also agreed that WD factors should be considered at the estimation of the WD development risks.

Hypothesis H2: The WDF estimation matrix is useful, understandable and helpful to estimate and consider WD factors.
The consensuses of the estimated WDF values, subjects feedbacks regarding the usefulness, helpfulness and ease of use of the matrix,

TABLE III: SUMMARY OF ESTIMATION DECISION (RE/TREV) [1]

Situation No.:	**6004**	**0901**	**1072**	**8033**
RE **TREV**	17 6	1 22	5 18	19 4
Avg. Used Time	22.0	27.9	32.8	28.0

TABLE IV: RE/TREV DECISION JUSTIFICATIONS SUMMARY [1]

Justifications for RE	Justifications for TREV
Single site and no distribution **N**o involvement of WD factors **S**imple and RE is enough	**M**ulti Sites and WD factors are involved **C**omplexity of the situation and involved factors **T**REV gives more abilities

confusing or a need for support and experimenter observation parameters were used to test this hypothesis. The statistics calculations for the WDF estimated values for situation number 0901 illustrated that the mean is 12.455 and the median is 12.00, the standard deviation is 1.057 and the P-value is less than 0.05. This indicates that the data is very close together and the data clusters around the mean. Meanwhile, looking at the distribution of WDF for situation number 1072 as illustrated by statistical graphically summary (see Figure 3), we see that the mean is 10.611 and the median is 10.50. This indicates that the data is very close together, also the standard deviation is 1.145 and P-value is less than 0.05 indicating that the data clusters around the mean. Therefore, the WDF estimation for both situations (numbered 0901 and 1072) have demonstrated a highly consensus taking in the consideration that the values of WDF in the matrix are ranged from 3 to 15 (minimum of WDF is (3*1 = 3) when all of the selected options are at the first level, and the maximum of WDF is (3*5 = 15) when all of them are at fifth level).

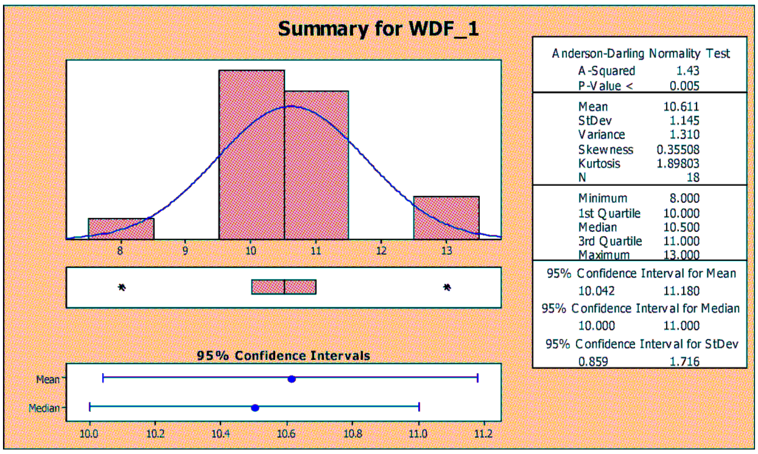

Figure 3: Statistics Graphical Summary for (WDF situation 1072) [1]

The estimated values of WDF (sites dependency, sites distribution and communication availability) for the two situations numbered 0901 and 1072 have also shown a highly consensus as statistics calculations exhibited. Figure 4 illustrates an example for statistics graphical summary for the sites distribution factor estimation of the situation number 1072.

Generally, the average time used for the estimation using the TREV equation was 3.069 minutes which is acceptable and was not so high compared with the average time of using RE equation, which is 1.375, taking into the account that in TREV case, the estimation includes additional WD factors estimation and extra calculations.

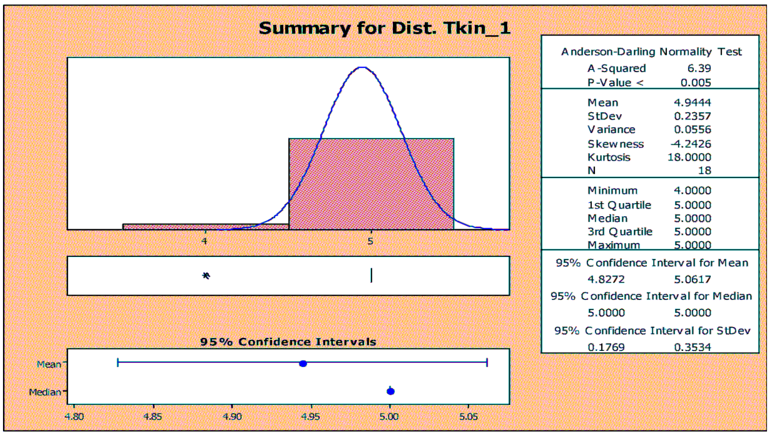

Figure 4: Statistics Graphical Summary for (Sites Distribution 1072) [1]

At the end of this evaluation stage, the subjects were asked to give a general feedback about the estimation module by answering some guided questions and also providing their suggestions and comments to improve the module. Table V summarizes the answers and the feedbacks regarding the WD factors, TREV equation and the support matrix. The subjects were asked to rate the usefulness of the estimation matrix, 17 subjects were responded. Their answers as showed in Table V were ranged from Useful to Very Useful with percentage rate (29.4%) for Useful, (52.9%) for Very Useful, (17.6%) for Strongly Useful and the other options (Not Useful and Somewhat Useful) have got (0%) of ticks. The next question was about the ease of use of the matrix to estimate the WDF which was answered by 23 subjects. As Table V illustrated the answers were asserted on the ease of use of the matrix. The values as shown on the table were rated as (0%) ticks to Difficult and Somewhat Easy, (13%) for Moderate, (52.1%) for Easy and (34.7%) for Very Easy. The subjects were agreed with including of the WDF at the estimation of WD development risks with percentage rate (59.0%) for Strongly Agree and (40.9%) for the Agree option and (0%) for the rest of the options (Neutral, Disagree and Strongly Disagree). The subjects has showed their support to the included factors as they did not criticize the three factors; however, some of them suggested to include other factors as can be seen in Table V.

4. DISCUSSION

The estimation module was evaluated in this controlled experiment. For the simplicity purpose the experiment was designed with a very minimum of dependency so that the evaluation of each module can be conducted individually and without any affect on the evaluation of other modules. The evaluation aspects include; ease of understand and use, usability and the usefulness of the modules. The result of the experiment was used to explore what sort of improvements that need to be made on the WeDRisk modules.

Hypothesises H1 and H2 were tested to evaluate the estimation module. Experiment result regarding these hypothesises is discussed below:

Hypothesis H1: The TREV equation is an ideal option to estimate the WD development risks and to consider the WD factors compared with the RE equation.

TABLE V: EVALUATION FEEDBACK FOR THE ESTIMATION MODULE

To what extent do you rate the usefulness of WD Factors estimation matrix?				
Not Useful	Somewhat Useful	Useful	Very Useful	Strongly Useful
0	0	5	9	3
How easy the use of the WD factors estimation matrix?				
Difficult	Somewhat Easy	Moderate	Easy	Very Easy
0	0	3	12	8
To what extant do you agree that TREV is more suitable than RE for WD risks estimation?				
Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
8	12	2	0	0
To what extant do you agree that WD factors should be included at W-D risks estimation?				
Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
13	9	0	0	0
Any other WD factors that should be considered: - Reliability of the sites - Sites local communications - Used methods and technologies compatibility across the sites - Geographical or environmental phenomena				
Any comments / suggestions: - It is easy to learn and use, clear and gives a quick result - Lack of finance information consideration - Estimate sites risks individually and then combine the overall sites estimations - The formula (WD/TREV) works well depending on how much information provided for each scenario - Use different guidance for TREV (big company risks different from small company risks)				

The result of the experiment regarding this hypothesis implies the following two points:

I) Experiment result exhibited a consensus in subjects' decisions regarding the suitability of RE and TREV for each injected situation and it came as what the experimenter has expected. The subjects justified their decisions (see Table IV) with the simplicity, not involvement of WD factor or the situation is single site for the RE selection and they justified the TREV selection with complexity, multisite or involvement of the WD factors. From the result, it can be understood that the subjects have decided to use RE equation for none WD development situations and TREV equation for the WD situations.

II) The subjects were agreed or strongly agreed that the TREV equation is more suitable than RE equation to estimate the WD development risks.

Therefore, based on the above result it can be concluded that hypothesis H1 has got a very strong support from the experiment result.

Hypothesis H2: The WDF estimation matrix is useful, understandable and helpful to estimate and consider WD factors.
The evaluated matrix involves three WD factors (sites dependency, sites distributions and communication), which could have an effect on the WD development risks. These factors are not final and one of the evaluation aims is to explore if they need to be modified or if there is any other factors need to be considered. In this regard, as the result showed in the previous section the matrix has helped the subjects to estimate the risks with consideration to the involved WD factors. The matrix was easily used with negligible confusing and with highly understanding and it gave consensuses result. Some subjects were argued to add some other factors to the matrix (e.g. reliability of the

sites; sites local communications; used methods; technologies compatibility across the sites and geographical or environmental phenomena). However, the factors that should be included in this matrix are the ones which have an effect on the risks, but not the risks themselves. Therefore, the above suggested factors and others could be revised and added to the matrix if they comply with the WDF factor definition. Changing and modification of the matrix factors is easy task and the matrix could be used in other types of software development with some changes.

Based on the consensus in the estimated WDF values as it is demonstrated with situations 0901 and 1072, the positive feedback regarding the usefulness of the matrix, ease of use of the matrix, suggestions and comments regarding the matrix as described above and also the experimenter observations it can be concluded that the hypothesis H2 has got a strong support in this experiment. Meanwhile, observation, suggestions and discussions regarding the TREV equation during the experiment stages exhibited that there is an ability to use the TREV equation for other types of software development risks, if the factors (the ones which affect the risks) of those types are identified and considered. A good issue with this matrix is that if any of the WD factors is not included in desired development, a default value " 1" will be given to that factor, which has no effect on the result. Finally, from the experiment result, it is obviously there is a chance to generalize the TREV equation and upgrading the matrix by including different factors when there is a need.

5. LIMITATIONS

Although in this controlled experiment we tried to emulate real risky situations in order to evaluate the WeDRisk approach, but there is a number of limitations associated with this experiment. Some of these limitations are general and related to the use of experiments method and some others are specific to this experiment. Hereafter, the involved limitations:

- The experiment running cost is high so it limited the number of subjects; therefore, the experimenter tried to tackle that by focusing on involved subjects type.
- Sometimes the samples in the experiments do not reflect the real population, but in this experiment most of the subjects have an experience in software projects via their work, study or research backgrounds.
- Usually experiments are affected by biases. In this experiment a number of measures were taken to avoid and reduce the involved biases.
- It is preferred if the subjects were from distributed software industry, but due to time limitation and imposed data restriction we could not.
- Due to the experimenter ability limitation at the experiment running time, the subjects were participated individually (one by one) in order to ensure providing needed support, observing tasks implementation and recording tasks implementation used time.

6. CONCLUSION AND FUTURE WORK

The rapid and evolved web and distributed development face many specific challenges and risks. The existing software risk management approaches are not able to accommodate these challenges and risks.

Existing approaches either they are designed for the collocated development type or they have many weaknesses. In this paper we presented an approach called WeDRisk, which is designed mainly to manage the web and distributed software development risks. The WeDRisk approach aims to address the weaknesses of existing approaches to risk management. It is flexible approach and able to deal with the rapid evolution nature of the web and distributed developments. This paper also described the design, execution and analysis of a controlled experiment that has been used to evaluate some central parts of the approach (Estimation module is one of them), in terms of their usefulness, ease of understanding and usability. Many reasons lead to conduct this controlled experiment such as:

- The novelty of the evaluated modules, so that we could not find suitable data form pervious approaches that can be used to evaluate the WeDRisk approach modules.
- Due to developers data restrictions, we could not find a suitable case study to evaluate the WeDRisk modules.

Finally, as described and discussed above it can be concluded that estimation module has successfully and effectively dealt with the injected risky situations. The experiment result illustrated how this module covers and considers the web and distributed factors and it is useful, understandable and easy to use. The experiment result exhibited that there are some improvement could be done on the modules which came in terms of suggestions, observations or form result analysis (e.g. TREV equation can be generalized and used to estimate risks in other types of software developments.

ACKNOWLEDGMENT

Since this work is a part of a PhD research at School of Computing Science at Newcastle University-UK, the author would like to thank Dr. Steve Riddle and the staff of this school for their inspiration and support.

REFERENCES

[1] A. A. Keshlaf WeDRisk: An Approach to Managing Web and istributed Software Development Risks, PHD Thesis at University of Newcastle upon Tyne. Computing Science, ,publisher: University of Newcastle upon Tyne, 2013, url:http://books.google.com.ly/books?id=sIZlngEACAAJ.

[2] A.A., Keshlaf, and S., Riddle, Risk Management for Web and Distributed Software Development Projects,The Fifth International Conference on Internet Monitoring and Protection (ICIMP10),Barcelona-Spain, IEEE Computer Scocity, Pages 22 -28, 2010.

[3] A.A., Keshlaf, and S., Riddle, Web and Distributed Software Development Risks Management: WeDRisk Approach, International Journal on Advances in Software, vol. 3 no. 3 - 4, Pages 447-460, 2010.

[4] A., Lamersdorf, J., Munch, J. and A. and Viso Torre, A., A Risk-Driven Model for Work Allocation in Global Software Development Projects, The Sixth IEEE International Conference on Global Software Engineering (ICGSE11), IEEE, Pages 15-24, 2011.

[5] R., Rocha, C.,Costa and R., Prikladnicki, Collaboration Models in Distributed Software Development: a Systematic Review,Clei Electronic Journal, vol. 14 no 2, Page 1, Agust 2011.

[6] K.,Lyytinen, L.,Mathiassen and J.,Ropponen Attention Shaping and Software RiskA Categorical Analysis of Four Classical Risk Management Approaches,Journal of Information Systems Research , Vol. 9 no.3, Pages233255 March 1998.

[7] B.W., Boehm, Software risk management: principles and practices, Software IEEE,Vol. 8 no.1, Pages 32 -41 March 1991.

[8] A.A., Keshlaf, and S., Riddle, An Empirical Study of Web and Distributed Software Risks from Three Perspectives: Project,

Process and Product,ICIMP 2011 The Sixth International Conference on Internet Monitoring and Protection(ICIMP11), March 20-25, 2011 - St. Maarten, The Netherlands Antilles pages Pages 38-44, 2011.

[9] M. Nordio, H.-C. Estler, B. Meyer, J. Tschannen, C. Ghezzi, and E. D. Nitto Nitto, How do Distribution and Time Zones Afect Software Development? A Case Study on Communication, In 6th IEEE International Conference on Global Software Engineering (ICGSE11), August 15-18 2011 Finland, IEEE, Pages 176 - 184, 2011.

[10] D. I., Sjoberg, T. Dyb and M. Jrgensen, The Future of Empirical Methods in Software Engineering Research, Future of Software Engineering(FOSE'07), , IEEE Computer Society, Pages 358-378, 2007.

[11] M. V. Zelkowitz, D. R. Wallace and D. W. Binkley, Experimental validation of new software technology, In Lecture notes on empirical software engineering, Natalia Juristo and Ana M. Moreno (Eds.). Software Engineering And Knowledge Engineering, Vol. 12. World Scientific Publishing Co., Inc., River Edge, NJ,2003, USA Pages 229-263, 2003.

[12] B. Kitchenham, DESMET: A method for evaluating Software Engineering methods and tools, Technical Report TR96-0,Department of Computer Science, University of Keele UK, August 1999.

[13] B. Kitchenham, H. Al-Khilidar,M. A. Babar, M. Berry, K. Cox,J. Keung, F. Kurniawati,M. Staples,H. Zhang and L. Zhu, Evaluating guidelines for reporting empirical software engineering studies,Springer Science, 2008.

[14] R. A. Maxion and K. S. Killourhy, Keystroke Biometrics with Number-Pad Input, IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), PP 201-210, IEEE 2010.

[15] V. R. Basili, The Role of Controlled Experiments in Software Engineering Research,Empirical Software Engineering Issues. Critical Assessment and Future Directions. V. Basili, D. Rombach, K. Schneideret al, Springer Berlin / Heidelberg, pp, 2007.