A NOVEL METHOD FOR LARGE SCALE REQUIREMENT ELICITATION

[1]V. DHEEPA, [2]C. VIJAYALAKSHMI, [3]E. R. NAGANATHAN

School of Computing Science, Hindustan University, Chennai
Email: vdeepa@hindustanuniv.ac.in, vijiduai88@gmail.com & ernindia@gmail.com

Abstract: Requirements elicitation is the practice of collecting the requirements of a system from users, customers and other stakeholders. Requirement elicitation process will identify and prioritize requirements. This process difficult in large software projects with many stakeholders. Here will introduce a novel method that uses social networks and collaborative filtering to manage the requirement elicitation process in large scale. Collaborative filtering is done using an improved k-nearest neighbor algorithm where the existing system will be using k-nearest neighbor algorithm. In this method network formation process plays an important role, it identifies stakeholders and the stakeholder needs to recommend other stakeholders and stakeholder roles, by continuing this process we will obtain a social network where stakeholders acts as a nodes and their recommendations act as links. In the social network the project influence is determined by using variety of social network measures. After that every stakeholder in network needs to rate their initial set of requirements, then this method will recommends other relevant requirements to them using collaborative filtering, and prioritizes their requirements using their ratings and their project influence. So this novel method predicts stakeholder needs accurately, completely and accurately prioritized list of requirements compared to the existing methods, and this method will prioritize requirements within the time.

Keywords: Requirements specifications, elicitation methods, requirements prioritization, recommender systems, social network analysis, stakeholder analysis.

I. INTRODUCTION

The usage of software is increased. Because of this the length of the program and number of modules will be increased. Now a day the leading projects involve lot of stakeholder, individuals and groups.
The customers who spend for the system, developers who develop, users who use the system, and legislators who develop the rules for the system are included in stakeholder.

The large projects have lot of stakeholders all over the world and their requirements may conflict with each other.
Requirements elicitation is the practice of collecting the requirements of a system from users, customers and other stakeholders. Requirement elicitation process will identify and prioritize requirements. This process difficult in large software projects with many stakeholders.
Here will introduce a novel method that uses social networks and collaborative filtering to manage the requirement elicitation process in large scale. There are three basic problems in large scale requirement gathering information over load, inadequate stakeholder input, and biased prioritization of requirements.

Information overload is a big problem in big projects. These projects will always have many stakeholders and lot of requirements. Old methods for requirements elicitation require direct interactions with the stakeholders, like face-to-face meetings, brainstorming sessions and interviews.
These old methods will not manage the information elicited from the stakeholders. These methods fail when a project has hundreds, thousands, or even hundreds of thousands of stakeholders. And also stakeholders are omitted and their requirements overlooked. Users become frustrated when the software not meets their needs. Customers who pay for the system may need to pay for the mistakes.

1.2 Problem Statement

The large scale software projects with many stakeholders who must be a part, but are unable to meet them, possibly due to large numbers, overseas locations, or insufficient time. Most elicitation methods require one to one meetings with the stakeholders. This process is time consuming when there are many stakeholders.
Old methods for requirements elicitation require direct interactions with the stakeholders, like face-to-face meetings, brainstorming sessions and interviews.
These old methods will not manage the information elicited from the stakeholders. These methods fail when a project has hundreds, thousands, or even hundreds of thousands of stakeholders. And also stakeholders are omitted and their requirements overlooked.

1.3 Objective of the Work

In large scale software projects there will be a lot of clients who we can't able to meet often, because they may be present in some where around the country. Most of the elicitation methods need the customer interaction with the developer but this process needs more time and money. The existing method is lagging in elicitation of requirements from the stakeholders. Because of large number of customer many stakeholder requirements may be over looked. That

means it is not taken in to account. Because of requirement overlooked many important requirements are omitted in sea of information. So the main objective of the project is to give better user experience and to do requirement elicitation in a efficient way.

II. EXISTING SYSTEM

In large scale software projects there will be a lot of clients who we can't able to meet often, because they may be present in some where around the country. Most of the elicitation methods need the customer interaction with the developer but this process needs more time and money. The existing method is lagging in elicitation of requirements from the stakeholders. Because of large number of customer many stakeholder requirements may be over looked. That means it is not taken in to account. Because of requirement overlooked many important requirements are omitted in sea of information.

2.1 Demerits of Existing System

Information overload: Information overload is a big problem in big projects. These projects will always have many stakeholders and lot of requirements. Users become frustrated when the software not meets their needs. Customers who pay for the system may need to pay for the mistakes.

Biased prioritization of requirements: Occurs because current prioritization techniques depend on individual stakeholder who may not have a major part in large projects. As a result, important requirements known to only to some of the stakeholders can be lost in huge information.

Inadequate stakeholder input: Inadequate stakeholder input caused by selection of inadequate stakeholder omitting stakeholders is one of the common mistakes in requirement elicitation.

III. PROPOSED SYSTEM

The proposed system will introduce a novel method that uses social networks and collaborative filtering to manage the requirement elicitation process in large scale. Collaborative filtering is done using an improved k-nearest neighbor algorithm where the existing system will be using k-nearest neighbor algorithm. In this method network formation process plays an important role, it identifies stakeholders and the stakeholder needs to recommend other stakeholders and stakeholder roles, by continuing this process we will obtain a social network where stakeholders acts as a nodes and their recommendations act as links. In the social network the project influence is determined by using variety of social network measures. After that every stakeholder in network needs to rate their initial set of requirements, then this method will recommends other relevant requirements to them using collaborative filtering, and prioritizes their requirements using their ratings and their project influence. So this novel method predicts stakeholder needs accurately, completely and accurately prioritized list of requirements compared to the existing methods, and this method will prioritize requirements within the time.

3.1 Metrics of Proposed System

Information overload: To recommend relevant requirements [3] to stakeholders, and prioritizing the stakeholders and requirements.

Inadequate stakeholder: To recommend other stakeholders, and asking all stakeholders to provide initial list of requirements.

Biased prioritization: Using the stakeholders' ratings on the requirements and their position on the social network.

IV. PROPOSED SYSTEM IN DETAIL

4.1 Identify and Prioritize Stakeholder

The first step to identify and prioritize the network is social network analysis which is done by snowballing method. Snowball sampling begins with a set of actors each of these actors is asked to recommend other actors. Snowball sampling uses a small pool of initial informants to nominate, through their social networks, other participants who meet the eligibility criteria and could potentially contribute to a specific study. The term "snowball sampling" reflects an analogy to a snowball increasing in size as it rolls downhill.

A recommendation is a triple <User, User role, salience> Salience-is a number in ordinal scale (1-5) The process continues until no new actors are identified. Snowball Sampling is a method used to obtain research and knowledge, from extended associations, through previous acquaintances, "Snowball sampling uses recommendations to find people with the specific range of skills that has been determined as being useful." An individual or a group receives information from different places through a mutual intermediary. This is referred to metaphorically as snowball sampling because as more relationships are built through mutual association, more connections can be made through those new relationships and a plethora of information can be shared and collected, much like a snowball that rolls and increases in size as it collects more snow. Snowball sampling is a useful tool for building networks and increasing the number of participants. It applies various social network measures such to prioritize the stakeholders in the network. The social network measures produce a score for each stakeholder. The stakeholder roles are prioritized by the highest score and the place in social network.

4.2 Collect Profile

The stakeholders identified in previous stage will provide their preferences on the initial requirements [3]. A preference is a triple: <User; their requirement; rating> where rating is a number on an ordinal scale

(e.g., 0-5) reflecting the importance of the requirement to the stakeholder (e.g., 0 is unimportant and 5 is very important. Stakeholders can indicate requirements by X which they actively do not want.

As a result we will get profile of stakeholder who responded.

4.3 Predict Requirements

Collaborative filtering is process used to make atomic predictions. Collaborative filtering is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. Applications of collaborative filtering typically involve very large data sets. Collaborative filtering methods have been applied to many different kinds of data including: sensing and monitoring data, such as in mineral exploration, environmental sensing over large areas or multiple sensors; financial data, such as financial service institutions that integrate many financial sources; or in electronic commerce and web applications where the focus is on user data, etc. The remainder of this discussion focuses on collaborative filtering for user data, although some of the methods and approaches may apply to the other major applications as well.

By using improved Knn algorithm, we can predict the similar user intrest.

kNN Algorithm

In pattern recognition, the *k*-nearest neighbor algorithm (*k*-NN) is a method for classifying objects based on closest training examples in the feature space. *k*-NN is a type of instance-based learning, or lazy learning where the function is only approximated locally and all computation is deferred until classification. The *k*-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its *k* nearest neighbors (*k* is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of its nearest neighbor. The same method can be used for regression, by simply assigning the property value for the object to be the average of the values of its *k* nearest neighbors. It can be useful to weight the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones.

The neighbors are taken from a set of objects for which the correct classification is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. The *k*-nearest neighbor algorithm is sensitive to the local structure of the data. Nearest neighbor rules in effect compute the decision boundary in an implicit manner. It is also possible to compute the decision boundary itself explicitly, and to do so in an efficient manner so that the computational complexity is a function of the boundary complexity.

Collaborative Filtering Algorithm

Collaborative filtering (CF) is a technique used by some recommender systems. Collaborative filtering has two senses, a narrow one and a more general one. In general, collaborative filtering is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. Applications of collaborative filtering typically involve very large data sets. Collaborative filtering methods have been applied to many different kinds of data including: sensing and monitoring data, such as in mineral exploration, environmental sensing over large areas or multiple sensors; financial data, such as financial service institutions that integrate many financial sources; or in electronic commerce and web applications where the focus is on user data, etc. The remainder of this discussion focuses on collaborative filtering for user data, although some of the methods and approaches may apply to the other major applications as well.

In the newer, narrower sense, collaborative filtering is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating). The underlying assumption of the collaborative filtering approach is that if a person *A* has the same opinion as a person *B* on an issue, A is more likely to have B's opinion on a different issue *x* than to have the opinion on x of a person chosen randomly. For example, a collaborative filtering recommendation system for television tastes could make predictions about which television show a user should like given a partial list of that user's tastes (likes or dislikes). Note that these predictions are specific to the user, but use information gleaned from many users. This differs from the simpler approach of giving an average (non-specific) score for each item of interest, for example based on its number of votes.

4.4 Requirement Prioritization

Pair wise comparison:

Pair wise comparison is a kind of divide-and-conquer problem-solving method. It allows one to determine the relative order of a group of items. This is often used as part of a process of assigning weights to criteria in design concept development. This approach is used by the requirement engineer to find the most important requirement between two. Then the most important requirement is entered in the cell of the matrix. If both the requirement is equally important then both is entered in the matrix. By the way we can find the most important requirement by comparing all the cells in the matrix. For this we need $n*(n-1)/2$ comparisons.

BST: A binary search tree (BST), also known as an ordered binary tree, is a node-based data structure in which each node has no more than two child nodes. Each child must either be a leaf node or the root of another binary search tree. The left sub-tree contains only nodes with keys less than the parent node; the right sub-tree contains only nodes with keys greater

than the parent node. The most important requirement is assigned as a root node, then compares two requirements, put least important requirement as a left node and most important node as a right node then by repeating this process we will get a prioritized list of requirements.

V. SYSTEM ARCHITECTURE

The proposed system will introduce a novel method that uses social networks and collaborative filtering to manage the requirement elicitation process in large scale. In this method network formation process plays an important role, it identifies stakeholders and the stakeholder needs to recommend other stakeholders and stakeholder roles, by continuing this process we will obtain a social network where stakeholders acts as a nodes and their recommendations act as links. In the social network the project influence is determined by using variety of social network measures. After that every stakeholder in network needs to rate their initial set of requirements, then this method will recommends other relevant requirements to them using collaborative filtering, and prioritizes their requirements using their ratings and their project influence.

5.1 Step 1. Identify and Prioritize Stakeholders

The first step to identify and prioritize the network is social network analysis which is done by snowballing method. Snowball sampling begins with a set of actors each of these actors is asked to recommend other actors.

Snowball sampling uses a small pool of initial informants to nominate, through their social networks, other participants who meet the eligibility criteria and could potentially contribute to a specific study. The term “snowball sampling” reflects an analogy to a snowball increasing in size as it rolls downhill.

A recommendation is a triple <User, User role, salience> Salience-is a number in ordinal scale (1-5) the process continues until no new actors are identified. Snowball Sampling is a method used to obtain research and knowledge, from extended associations, through previous acquaintances, “Snowball sampling uses recommendations to find people with the specific range of skills that has been determined as being useful.” An individual or a group receives information from different places through a mutual intermediary. This is referred to metaphorically as snowball sampling because as more relationships are built through mutual association, more connections can be made through those new relationships and a plethora of information can be shared and collected, much like a snowball that rolls and increases in size as it collects more snow. Snowball sampling is a useful tool for building networks and increasing the number of participants. It applies various social network measures such to prioritize the stakeholders in the network. The social network measures produce a score for each stakeholder. The stakeholder roles are prioritized by the highest score and the place in social network.

5.2 Step 2. Collect Profile

The stakeholders identified in previous stage will provide their preferences on the initial requirements [3]. A preference is a triple: <User; their requirement; rating> where rating is a number on an ordinal scale (e.g., 0-5) reflecting the importance of the requirement to the stakeholder (e.g., 0 is unimportant and 5 are very important. Stakeholders can indicate requirements by X which they actively do not want.

As a result we will get profile of stakeholder who responded.

Requirements are in different level of abstraction, the higher level requirements are divided into several lower level requirements or specific requirements. If we are achieving all the low level requirements then the higher level requirements will be achieved obviously. If we are achieving all the high level requirements then the objective of the project is achieved.

5.3 Step 3. Predict Requirements

4.3 Predict Requirements

Collaborative filtering is process used to make atomic predictions.

collaborative filtering is the process of filtering for information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. Applications of collaborative filtering typically involve very large data sets. Collaborative filtering methods have been applied to many different kinds of data including: sensing and monitoring data, such as in mineral exploration, environmental sensing over large areas or multiple sensors; financial data, such as financial service institutions that integrate many financial sources; or in electronic commerce and web applications where the focus is on user data, etc. The remainder of this discussion focuses on collaborative filtering for user data, although some of the methods and approaches may apply to the other major applications as well.By using improved Knn algorithm, we can predict the similar user interest.

5.4 Step 4. Prioritize Requirements

By using pair wise comparison and Binary search tree we can prioritize the requirements which got from the previous stage.

Pair wise comparison is a kind of divide-and-conquer problem-solving method. It allows one to determine the relative order of a group of items.

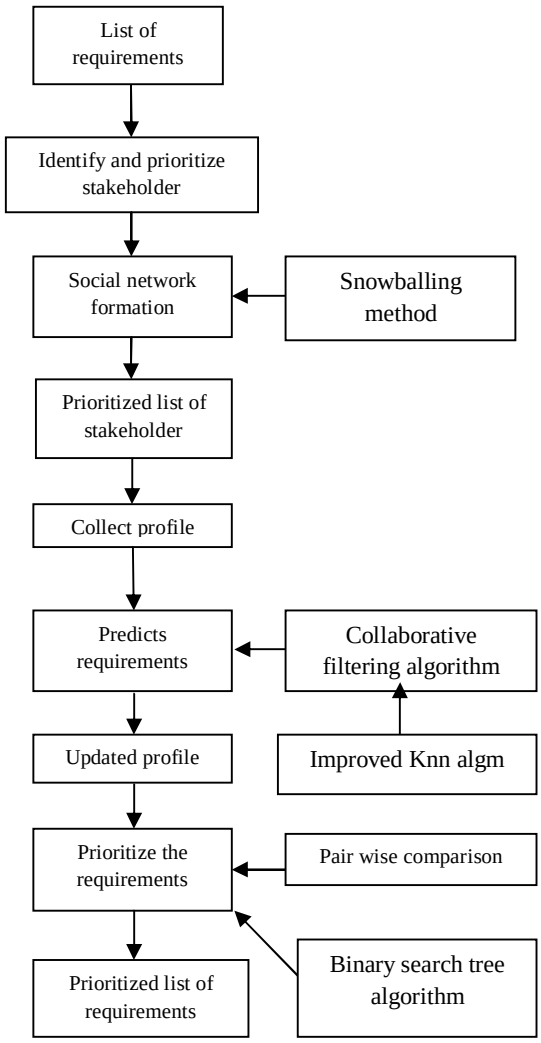

Figure 1. System architecture diagram

VI. RESULT

The problem in collecting requirements from stakeholder in large scale is addresses here and the solution for the same problem also defined here. With the help of collaborative filtering and social network formation the large scale requirements will be prioritized and the missing requirements also predicted as a result we got list of predicted and prioritized requirements with in a fraction of seconds.

CONCLUSION

In large scale software projects there will be a lot of clients who we can't able to meet often, because they may be present in some where around the country. Most of the elicitation methods need the customer interaction with the developer but this process needs more time and money. The existing method is lagging in elicitation of requirements from the stakeholders. Because of large number of customer many stakeholder requirements may be over looked. That means it is not taken in to account. Because of requirement overlooked many important requirements are omitted in sea of information. The proposed system will introduce a novel method that uses social networks and collaborative filtering to manage the requirement elicitation process in large scale. Collaborative filtering is done using an improved k-nearest neighbor algorithm where the existing system will be using k-nearest neighbor algorithm. In this method network formation process plays an important role, it identifies stakeholders and the stakeholder needs to recommend other stakeholders and stakeholder roles, by continuing this process we will obtain a social network where stakeholders acts as a nodes and their recommendations act as links. In the social network the project influence is determined by using variety of social network measures. After that every stakeholder in network needs to rate their initial set of requirements, then this method will recommends other relevant requirements to them using collaborative filtering, and prioritizes their requirements using their ratings and their project influence. So this novel method predicts stakeholder needs accurately, completely and accurately prioritized list of requirements compared to the existing methods, and this method will prioritize requirements within the time.

FUTURE WORK

For the future, the Stake rare method should be developed with the newer techniques to identify and prioritize stakeholders. This method using Knn algorithm to find the like minded users, for the future we can use improved and efficient Knn algorithm to find exact like minder user to improve the requirement prediction sate.

REFERENCES

[1] StakeRare: Using Social Networks and Collaborative Filtering for Large-Scale Requirements Elicitation. IEEE transaction on softare engineering, vol. 38, no. 3, may/ june 2012

[2] B. Nuseibeh and S. Easterbrook, "Requirements Engineering: A Roadmap," Proc. Conf. Future of Software Eng., pp. 35-46, 2000.

[3] H.Sharp, G.H. Galal, and A. Finkelstein, "Stakeholder Identification in the Requirements Engineering Process," Proc. Database and Expert System Applications Workshop, pp. 387-391, 1999.

[4] P. Zave, "Classification of Research Efforts in Requirements Engineering," ACM Computing Surveys, vol. 29, no. 4, pp. 315- 321, 1997

.

[5] J.Cleland-Huang and B. Mobasher, "Using Data Mining and Recommender Systems to Scale Up the Requirements Process," Proc. Second Int'l Workshop Ultra-Large-Scale Software-Intensive Systems, pp. 3-6, 2008.

[6] R.N. Charette, "Why Software Fails," IEEE Spectrum, vol. 42, no. 9, pp. 42-49, Sept. 2005.

[7] D.C. Gause and G.M. Weinberg, Exploring Requirements: Quality before Design. Dorset House Publishing Company, Inc., 1989.

[8] I. Alexander and S. Robertson, "Understanding Project Sociology by Modeling Stakeholders," IEEE Software, vol. 21, no. 1, pp. 23-27, Jan./Feb. 2004.

[9] I. Alexander, "A Taxonomy of Stakeholders: Human Roles in System Development," Int'l J. Technology and Human Interaction, vol. 1, no. 1, pp. 23-59, 2005.

[10] L. Lehtola, M. Kauppinen, and S. Kujala, "Requirements Prioritization Challenges in Practice," Proc. Int'l Conf. Product Focused Software Process Improvement, pp. 497-508, 2004.

★★★