CHAPTER I

INTRODUCTION

1.1. Rationale

During 1990s, CRM introduce himself focuses mainly on implementing and automating processes of customer data [1]. In Internet era, CRM was transforming to Web 2.0 concept. The concept called Social Customer Relationship Management, as know as SCRM. Greenberg [1]defined SCRM as "a philosophy and a business strategy, supported by a technology platform, business rules, processes and social characteristics, designed to engage the customer in a collaborative conversation in order to provide mutually beneficial value in a trusted & transparent business environment. It is the company's response to the customer's ownership of the conversation". Mix and match between customer and business is the core of SCRM Strategy. The SCRM Strategy is not only managing customer data but also make customer engagement.

To make customer engagement, an integrated system between a company and customer activity is needed. The Company needs relationships data such as friendships, social status, and others meaningful relationships between customers on Social CRM. The system also needs a way to get the social activities customer data. One solution to obtain social activity data from customers is by interacts directly with them. However, the company has limited quality interaction with all customer; therefore, the social media is one many tools to build interactions between a company and its customers. Customers are already in social networks and form a community themselves based on their interest or hobbies. The company must have engagement to this community rather than to every user. In the network community, the customers themselves control the community. The Company has to manage the community to control them as the primary goal of Social CRM. One form of managing over the community is by taking the influential people on the network community.

Social Network Analysis (SNA) was used to identifying influential people on the social network [2]. The method of SNA calculation dividing into three major ways of calculation or searches, i.e. Connection, Distributions, and Segmentation. SNA Connection used to look for similarities and differences in content, the amount of content, familiarity and closeness tendencies. SNA Distribution lead to the inside cover mathematical calculations between points and reconnection is in the graph. Moreover, the last is SNA Segmentation used to gathering nodes then identified a group.

Focusing on SNA Distributions theory calculation. The method consists of six pieces sub methods, the bridge, i.e. centrality, density, distance, structural holes, and tie strength. Influential people identification uses centrality sub methods. It used centrality because it is a special method to categorize and quantify the importance and influence point or user node relative to other user nodes that are connected to the network, either directly or indirectly. Calculation methods in centrality consist of degree, betweeness and closeness that is suitable method applied to this thesis research.The result can be categorized into other definition of influential people. Lisa Barone [3] identified influential people as the social butterflies, the thought Leaders, the trendsetters, the reporters, and the everyday customer. So, the result can be categorized in the definition above by looking the object content or profile manually.

1.2. Statement of the Problem

Based on rationale above the problem based on question: "How to identify influential people on social networks with activity relationships attributes?" The problem of the research is influential people identification as parts of Social CRM Input. The influential people engaging with the company to attempt high quality relationships with the company as the community. Influential people identification proposed with many classifications. The classification is using subjective parameters; they do not have exactly parameter to calculate them automatically by math or calculation algorithm, they are based on the user profile surveyed.

1.3. Theoretical Framework

Nowadays, since W3C proposing new concept of web communication, semantic web has become popular. Ontology chosen as knowledge based semantic description for basic formalization based on ontology as knowledge based to communicate between website in Internet. Ontologies give more power on conceptual with domain and relation in the concept. Ontology is modeling language for formalization design in software engineering. The modelling language communicating between remote team member design [1] to avoid miss-configuration of software and mediating mechanism for accessing heterogeneous data and information sources [2]. Ontology is a solution for software engineering knowledge representation [1]. Ontologies are classified into software development life cycle phase [3]. Ontology presenting metadata schemas, offering a controlled vocabulary of concept richer than UML class/object model [4].

Ontology can improve SNA measurement [5]and software engineering knowledge [6]. Within graph theory and network analysis, centrality is used to determine the relative importance of a node within the graph [7]. Centrality researchs in different aspects [8] to get influential people. Beside for political issues [9] and terrorism information [10] on social network [11], in online market business with higher values for three degrees the most popular (degree, betweeness, closeness) of centrality represent users who can targeted as productive user to selling product [12].

Semantic description on ontology to identify influential people will conduct in this thesis. Ontologies are an effective tool for modeling, sharing and reusing knowledge [6]. Facilitates open model to identify influential people for better description in future works.

This thesis will apply semantic description on ontology to identify influential people. Ontologies are an effective tool for modeling, sharing and reusing knowledge (Hilera & Fernández-Sanz, 2010). Ontology serving model to identify influential people for better description in future works. Ontologies are closely related to modern object oriented software engineering with richer information model compared to UML and as a shared conceptualization, naturally easy to develop in future work [4].

Within the scope of graph theory and network analysis [13] [14] by using centrality [15]. A person can be considered influential from how many listed friend he has, his ability to connect two networks of friends, or a person who known a lot of others influential person. Degree, betweeness

and closeness centrality as popular centrality concept is adapted in measuring graph. A graph consist a node and an edge. A node is define person, edge define relation between nodes. Influential generates by friendship graph, which only has node of people connect each other with friend status relation as edge. Activity graph has user with other object node in social media considered as the one of source measuring influential people.

Influential people is classified by activity data like group, works, education and the others data. Several classification is proposed to identify type of influential people. Lisa Barone categorized influential people in Five aspect. There are the social butterflies, the thought leaders, the trendsetter, the reporters and the everyday customer. Klout.com Categorized influential people in Sixteen type of them. Lastly, Richard Millington classifies influential people in four aspects from their activities. i.e. expertise, fame, authority, relationship.

1.4. Conceptual Framework / Paradigm

The existence or relation of people seen from Internet especially on the social media. However, the influential factor of a person to his population cannot be measured. Knowing who is having influential factor on the community can make easy step market infiltration. The search engine showing only about attributes data of user, and it cannot see the most influential people on the community. Therefore, developing influential people engine that is able to search and find the influential people on the community is required. Influential people ranks from customer of a public service company data or ranks table from user communities or population can be generated.

1.5. Objectives and Hypotheses

The objective of this research is to look for influential people by building user data collector and Calculation engine to get influential people ranks by some calculation. The data collector engine collects the attributes, relationships and activities of user data. The calculation engine calculates the user data to generate the ranks table from the data collector output. On the other hand, Ontology builds the model on semantic based. Data collector interacts with source of the user data, which is the data on the Internet especially on the social media.

Based on theoretical framework, on graph theory there is a node in the graph as important node relatively with the graph. Social Network have same behavioral with graph will be use same method to find the important user in the network. Hypothesis of this research is the influential people can be identify from relationship and activities data with same method to get influential nodes in the graph theory.

1.6. Assumption

There are some variables that considered in general problem becausehe result needs people data taken from social network :

1. Subjective evaluation for influential people is conducted by checking user profile who has one of the aspects [17]; fame, authority, relationship and expertise and then the results are compared to thode of the others.
2. Social media have a user with non-personal account.
3. User always use real names in social network because some people have a same name or even their full name. It is quite difficult to identify each of them.
4. To obtain full personal data, user privilege is required.
5. Not every people shared their personal data.

1.7. Scope and Limitation

The scope and limitation of this study are:

1. Semantic description in this thesis designed for the purpose of software engineering, with ontology to describing the user relationship scheme in social network based software. Intelligent system or data mining fields is not include.
2. This thesis select Facebook as social media because it has a required model with the data collector. The social media must have relationships user data.
3. Subjective verification will be used to confirm user profile based from the calculation process.
4. Data population are collected from community by means of data collector application. The application integrated into community registration portal.

1.8. Importance of the Study

The study of the research has some important advantages in the real world, such as

1. It can be used to develop software for influential people identification.
2. In the field of social customer relationship management, it helps to identify influential people.