國立交通大學

資訊管理研究所

碩 士 論 文

社會網路服務推薦機制之研究

Recommendation of Social Network Based Services

研 究 生： 蕭涵文

指導教授： 李永銘 博士

中華民國 九十八 年 七 月

社會網路服務推薦機制之研究

學生：蕭涵文　　　　　　　　　　　　　　　　　　　指導教授：李永銘 博士

國立交通大學資訊管理研究所碩士班

摘　要

隨著社交網路的蓬勃發展，許多利基於社交網路上的服務，例如： 社交網站上的 application、部落格或個人入口首頁上的 widgets 和 gadgets 等等皆成長快速且多元。為了能有效地為社交網路上的使用者篩選出適合的服務，我們透過分析服務的熱門度與信譽、使用者個人喜好與其社交關係等三個面向，並利用倒傳遞類神經網路來模擬使用者的決策條件，建構出一個系統化的社會網路服務推薦機制。本實驗實作於全球著名的社交網路平台 Facebook 上；實驗結果顯示所提出的機制優於其他的方法，同時發現社交關係在社會網路服務的推薦上比使用者自身的喜好與服務的熱門度和信譽更為重要。

Recommendation of Social Network Based Services

Student: Han-Wen Hsiao Advisor: Dr. Yung-Ming Li

Institute of Information Management
National Chiao Tung University

ABSTRACT

Social network based services, such as applications on the social network websites, widgets on blogs, and gadgets on personal portals have grown dramatically in a tremendous amount. In order to efficiently recommend suitable and attractive social network based services to users, a systematical recommendation mechanism composing of service's popularity and reputation, user's preference and social relationship is proposed. A back-propagation neural network is applied to optimally model general users' decision making criteria of using social network based services. This recommender service is implemented to one of the most famous social network websites- Facebook. The experimental result shows that the proposed model outperforms than any other methodology, including Analytic Hierarchy Process. It is also found that social relationship plays the most important part in recommendation of social network based service, instead of user's preference or service's popularity and reputation.

致謝

終於走到這一步了。

兩年的碩士生涯於此畫下句點，回首來時路，真是充滿了酸甜苦辣。這兩年來的磨練，不只讓我學習到做研究的方針，也體悟到許多待人接物道理，使我從懵懵懂懂的大學生，蛻變成為與現實接軌的研究生，相信這段精彩的人生經驗將有助於我面對日後的職場生涯。這篇論文的產生，承蒙許多人的鼎力相助，若沒有你們不吝惜地教導、幫助、指引，就沒有這篇論文，在此誠摯地感謝你們，謝謝。

首先要感謝論文指導老師李永銘博士，兩年內的諄諄教誨和耳提面命開拓我學術上視野，不僅讓我親身學習到研究中的種種面向，也提供我獨立思考、自我成長和學習的機會。其次要感謝口試委員魏志平教授和劉敦仁教授，您所提出的精闢建議和指教補強了實驗不足之處，讓論文更臻完善。

交大資管所帶給我的不只是學術上的成長，更重要的是結交到一群彼此扶持、加油和打氣的摯友們，讓我獲得寶貴的人生資產。謝謝同舟共濟的碩班戰友連乃和小球，因為有你們的相濡以沫，幫助我克服許多研究上的困難，並給予我甚多心靈上的支持。多年後相信我會記得我們三人挑燈夜戰、徹夜未眠的畫面，以及我們如何幽默地自我解嘲和同仇敵愾的情形。博班學長們：Denny、易霖和無尾熊，感謝你們慷慨分享自身所學和經驗，給予我論文莫大的幫助，若沒有你們強力的支援，我可能沒辦法在這麼緊繃的時間壓力下完成論文，真的很感激你們。可愛的學弟妹：宗穎、阿雅和蕙如，你們的存在為 lab 增添許多歡樂，跟你們一起互相切磋成長，讓碩二這年更加充實有趣，也謝謝你們所開的食物和飲料團，這些可大大的調劑我苦悶的研究生活；謝謝李小弟常替我解答許多程式上的問題，謝謝阿雅姐成為家族接班人？謝謝家豪妹補給我許多山上現採食物和日劇精神食糧。很開心我能待在凝聚力這麼強的 lab，你們是我碩班生活中的寄託。最後要謝謝我的家人，感謝爸媽供應我念書，並不時地傾聽我的心聲和撫慰我的心靈。能成為你們的女兒是我修得的福氣。

蕭涵文　2009年6月 謹誌於 新竹 國立交通大學光復校區

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

CHAPTER 1 Introduction

1.1 Research Background and Motivation

The websites and services nowadays on the Internet are transforming profoundly to social related [19]. According to a survey done by Alexa.com (16 January, 2006), 15 out of the top 20 most popular websites are either social networking sites/services (SNS) or have embedded social networking functions. Also, it is reported that these SNS, or so called online social network (OSN), have attracted nearly half of web users [45]. SNS as a web site platform built for people to create and maintain social connections and share information and knowledge among individuals have emerged as an important medium for people to interact in the cyber world [27].The goal of these services originally lies in helping people establish online presence and social networks ;however would eventually shift to exploit the users base for commercial purpose [51]. Therefore, it is interesting to investigate whether the essential criteria or influential parameters have changed when users make their service decisions on social network website.

Diverse social network related services have been flourished and raised much popularity and attention in recent years, such as applications on social network website (e.g. Facebook and Myspace), Google gadgets and Yahoo widgets on blogs or personal portals, and Firefox add-ons in browsers. Among those, *Social Application*, which means the social network based application, might be the most representable one. Social application is different from traditional pc-based applications and digital goods (see in Table 1.1). A traditional application, such as calculator, Word, Excel, and so on emphasizes its functional ability to perform a specific job, while digital goods, for example: DVDs, CDs, MP3s, videos, provide contents, information, and knowledge to users. Social applications which have both functional abilities to interact with friends on social network platform and share peer-production contents to each other are sort of a combination that contains traits of applications and digital goods.

Table 1.1The different characteristics of application, digital goods and social application

	Example	*Functionality*	*Content*	*Social Interaction*
Applications	Word, Excel	○		
Digital Goods	DVDs, MP3s		○	
Social Applications	VisualBookShlef	○	○	○

With the efforts contributed by many third-party developers, application's functionality, style and purpose has gone diversity. Top social application issues, for example, are related to "casual communication", "gifting", "gestures", "meeting people" and so on [5]. By viewing the subjects, it appears that these subjects are highly social related. It's quite astonished that social application as a brand new service style is capable of growing rapidly in a great amount and fascinating tons of active users. The number of applications on Facebook are over fifty thousand [15]. Furthermore, 70% of Facebook users actively use at least one application per month and meanwhile 4200 applications are implemented by over ten thousand users per month [42]. This extensive usage rate of social application highly indicates that there is a grand opportunity lying in the application market. More and more companies notice about the potential business value of utilizing application since it is not only able to be as a channel to expose company's products and services to social networks, but also able to help company execute target marketing with a lower cost. Take Visual bookshelf [46], a book collection and review sharing application on Facebook, for instance. It plays as a role of bridge that connects users to Amazon.com by providing easily links to the users who are interesting in buy books shared from friends' collections or others' reviews. This business model is beneficial to users, application developers/providers, and cooperating sponsors. Because Visual bookshelf as an application service provider could get its commercial sponsor's fees through the linkages, and in the meanwhile, Amazon.com could exploit pull strategy to market products to interested users. Also, to the users with buying intention, the business corporation would accelerate their buying process by providing a convenient shopping procedure. Other than this, there are hundreds and thousands of applications that may perform this new electronic business model to make profits. According to [37],there have been more than 100 companies established based on OSN application development business and Facebook application based advertising campaigns have been surprisingly successful. Therefore, there are practical reasons lying behind for choosing social application as the research target.

1.2 Research Problem

The problem buried in the business model of social application could be described from two points of view. On one hand, users would face a problem of how to efficiently and appropriately choose interested applications from tons of them. On the other hand, for the third party of social application developers, who are dying to attract as many users as possible in the purpose of earning more sponsors from advertisements or corporate companies, would encounter a problem of discovering the users with strong interest in it. Consequently, one of the motivations of this paper is to deal with the visibility dilemma of social application by proposing a sophisticated recommender service based on users' social relationship and application tastes. It is hoped that with the help of recommender service, users might decrease searching cost and increase their application usage rate. Also, with the benefit of recommendation, developers could raise more founds along with the increasing used rate of application.

Recommending social application is different from recommending books, movies, music or any other digital goods, because social application is mainly made for making interactions with people. However, it seems like that little researches have been done on analyzing recommendation of social related service. Hence, for further understanding people's decision tendency and criterion when service essence is greatly social related, we aim to implement a systematic analysis based on empirical collected data.

1.3 Research Objectives

In this paper, we study the recommendation of social network based service based on combining both objective view of analyzing services' popularity and reputation situation ,and subjective view of investigating users' preference and his/her social similarity and interaction. It is supposed that with this recommendation mechanism, users would be informed of which social network based services are highly suitable, interested, and social attracted to them without searching tons of existing services by himself or herself. In the mean time, service platform providers could be able to utilize this recommendation method to improve their customer relationship management by providing proper social network based services to catch customers' attention and fit their needs. Moreover, we tend to discover users' concerns of using social network based services. If users' decision making criteria could be revealed, which means the importance and priorities of service attributes are uncovered, we could better understand how to provide appealing services to users that close to their desired.

To achieve this goal, we choose one of the most popular social network based service, i.e. social application, as the experimental target. A statistic survey is utilized to extract the appropriate weights of popularity and reputation from users' viewpoints. Data mining methodology is used to calculate user's social application preference based on user's historical data. To examine social attraction power contributed by user's friends who already use that social application, we need to evaluate user's social situations, such as social similarity and interactions frequency. These three analytical dimensions would be aggregated into a final result with the weights calculated from Artificial Neural Network to model human service acceptance decision. Also, user's feedbacks would be analyzed by one of the famous multi-criteria decision making method called Analytic Hierarchy Process (AHP) in order to discover the practical weighting of the three aspects in human decision.

1.4 Thesis Outline

The rest of the paper is organized as follows: Section 2 presents related works. Section 3 demonstrates the system framework of social network based service recommendation. Section 4 describes the experiments, along with data collection and data analysis, followed by the experimental discussion in section 5. Finally, the conclusion and future works are portrayed in section 6.

CHAPTER 2 Literature Reviews

This section will reviews related works including online social networks, recommendation mechanism, back propagation neural network and the Analytic Hierarchy Process. The three formers are associated with the research topic and methodology, and the last is applied for the use in benchmark.

2.1 Online Social Networks

With the speedy growth of online social networking website/service, researchers have put academic efforts in analyzing the characteristic of online communities and their social graph structures. Distinctive features of social network such as linkage, taste, and subgroup difference are studied in [20, 29, 35, 38] that reveal the reality of user behaviors and network features. Topology analysis of online social structure, including distribution of in-degree and out-degree, shortest path length, and page view are researched in [19, 22], and the popularity of user-generated content is described in[11]. In addition to empirical studies, simulation issues including methods for networking sampling [28] and the effect of missing data in social network [26] are also discussed. Still, privacy concerns of sharing information on social networks have raised scholars' interests [1, 13, 41].

Comparing to the researches of previous, studies on social network based service are apparently rare and new. Lately published studies focus on summarizing characteristics of Facebook application into a higher level [37], as well as analyzing the statistic data of the growths patterns[3] or the activities[39] of application. However, among those works, it is likely that little do papers perform approaches to systematically deal with application recommendation problem.

2.2 Recommendation Mechanism

The issue of recommendation has aroused much academic interests and been spotlighted for decades. The main purpose of recommendation is to deal with information overload problem by providing a recommender service that would present suitable items to targeted users based on collected or inference information[31]. Recommendation mechanism usually could be categorized into three types according to the sources of recommendation data[6]: 1) content-based mechanisms, which suggest items based on the similarity to users' previous preference profiles, 2) collaborative-based mechanisms, which recommend items based on general tastes of similar users' profile, and 3) hybrid mechanism that combines the previous two approaches. Regardless of the success of each mechanism in many research domains, there are still some drawbacks in these methods. For the content based approach, due to the syntactic nature of the similarity matrices employed to compare, the existing metrics would only be able to detect the similarity items with the same attributes or features, leading to an overspecialized problem of only including items very similar to those the users already know[9]. Meanwhile, it is also limited by the availability of specifically defined objects with features[47]. For collaborative based approach, since it is required to know many user profiles in order to elaborate accurate recommendation results, limitation would lie in practical concerns of difficultly collecting and deriving such a great amount of data for a given user. Still, sparsity problem, which occurs when available data are insufficient for identifying similar users, would limit its applicability and quality.

Therefore, in order to design a suitable recommendation mechanism which firstly fits to online social network environment and secondly makes up for some shortcomings in content and collaborative based methods, we tend to put social relationship into consideration. The aspect of social relationship, including trust, intimacy and social similarity, has been implemented in several academic researches, such as blog recommendation[18, 30] and social media recommendation[25].With the supported information of social relations, we could design a better hybrid social recommendation mechanism that integrates the advantages of these three dimensions.

2.3 Back-propagation Neural Network

Artificial neural network(ANN), composed of an interconnected group of artificial neurons, is a mathematical or computational model that is able to capture complex inputs and outputs' nonlinear data relationship by simulating the structure or functional aspects based on the concept of biological neural network model [4].Neurons in ANN are typically located in the input layer, in one or several hidden layer ,and in the output layer[17]. Each neuron connected to the others with an associated weight representing information utilized by the network to solve a given problem.

ANN can be classified into different categories according to supervised/unsupervised and feed-forward/feed-back recall structure[12]. Back-Propagation neural network (BPNN) is a famous artificial intelligence technique for supervised machine learning. It uses a generalized delta rule algorithm that performs a gradient descent in the error space to minimize the total error between the predicted data and the desired data [43] ,and consequently yield predictive output with high similarity to the desired output. The process of learning algorithm in BPNN is: firstly, the network would propagate the training input pattern, which is sent from the input layer, to the output layer. If the pattern derived is different from desired, an error would be calculated and then be propagated backwards through the network to the input layer. In the meantime, associated weights would also be modified. As the network converges, a pattern between desired and input data has learned. Testing data could be feed to the newly trained network to calculate the performance of the model. With the adaptively data driven advantage, neural network is suitable for many empirical data generating process. For example, ANN is applied in numerous fields, such as pattern recognition[34], financial management and stock market[14] and tourism demand[10]. BBPN is one of the most frequently used ANN for classification and prediction[49]. Researchers have proven that BPNN with the learning ability is appropriate to predict in nearly all kinds of domain[48]. In this paper, we want to leverage the advantage of BPNN to deal with the uncertain weighting problem of parameter combination.

CHAPTER 3 SYSTEM FRAMEWORK

To design the mechanism of social network based service recommendation, social application and its user's social relationship are two essential components that require comprehension and analysis. Application's attributes would influence users' using interests in different degrees. For instance, some users are subject-oriented that they pay more attention to what fits their preference, while some users maybe more likely to follow the current that use social applications as long as they are popular. However, some users are more social-related that they would consider whether to use based on their friends' usage situations or opinions. That is, users would probably make decisions according to how many friends are using or who is using. It's noted that this kind of person would think highly of friend's attraction than application's traits, since the essential of social application lies in users' interaction. Hence, it is necessary for the recommendation system to put users' social relationship into consideration. Besides, user's tendency of social application is hard to predict, for it might be a mixture of several factors with distinct weights. It is possible for a person who is both subject-oriented and social-oriented, but with different degree, or a person who simultaneously cares about subject, popularity, reputation, socially usage situation and so on. Therefore, the vision of this paper aims to propose an innovative social network based application recommendation system by considering both applications' objective aspects and users' subjective point of views. The whole recommendation system architecture is demonstrated in Figure 3.1.

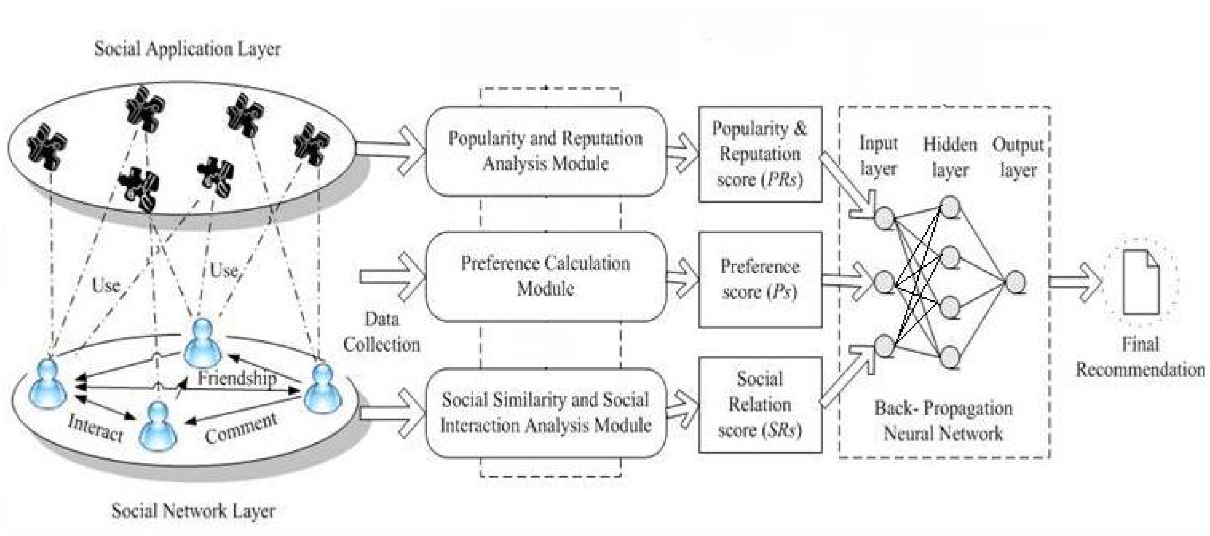


Figure 3.1 Architecture of social application recommendation mechanism

The systematical model includes three analytical modules which are "Popularity and Reputation Analysis Module", "Preference Analysis Module", and "Social Similarity and Social Interaction Analysis Module". In the popularity and reputation analysis module, a statistic survey is given to a mount of users in order to practically investigate how an application's public information, for example, number of users, number of fans and rating actually influence users' decision. After a general users' concern derived from the empirical study, application's population and reputation information could later be used to infer application's performance based on users' perceptions. In the preference analysis module, the target user's entire used application data is collected in the purpose of discovering user's category taste. In social similarity and social interaction analysis module, for the sake to evaluate an application's social attraction power, which mean the social invitation capability brought by the intimate friends who has used this application, the target user's social relationships and interactions among friends are extracted. These three modules would produce the corresponding scores (*PRs*, *Ps* and *SRs*), representing the degree of that dimension respectively. In order to optimally combine the scores to best represent users' points of view, an artificial intelligent method, back-propagation neural network, is applied. The details of calculations in the three modules are illustrated in the following sections.

3.1 Popularity and Reputation Analysis

Popularity and reputation of social network based application could be regarded as the clues for users to evaluate application's performance. It is reasonable to suppose that the higher popularity and reputation of an application, the more valued and attracted the social application may be. Common available popularity and reputation information online includes numbers of users, number of fans, number of friends using the application, and feedbacks such as reviews, discuss streams, and ratings. They are important materials for users to make their decisions on whether to use or not. However, for the concern of avoiding information overlap, we reduce the analysis dimensions to only three representative attributes, which are "No. of Users", "No. of Fans" and "Rating". Let's take review and ranking for example. Since it is ordinary for most websites to provide review mechanism along with rating, the comment of review could be highly positive related to ranking score. A review could be viewed as a subgroup of rating that expresses more thoughts and emotional manners than numerical rating number. Unfortunately, the opinion mining of review is beyond the research concept of this paper. Therefore, we tend to use only the rating of application to represent the overall reputation. Online rating, which could be traced back to 1990 [2], is omnipresent in books[32], movies[36] and news product items[8].Nevertheless, it is reported that current reputation mechanism leads to a disproportionately greater amount of positive feedback then negative or neural feedbacks [40].Thus, to eliminate the bias of relying on a specific parameter, we tend to combine more than one scope together. Fan is a newly used term in website that expresses a feeling of admiration or fondness of something. When users make themselves as fans of social network based applications, in some way, it represents users' stronger feelings toward it than others that are not chosen as fans. It is probably intuitive to regard "No. of Users" as a factor to evaluate popularity and "Rating" as a parameter to estimate reputation; however, "No. of Fans" seems like an undefined term which somehow might represent half popularity and half reputation. Therefore, in the research, popularity and reputation are merged together firstly to avoid the ambiguous problem of definition and secondly present a higher level of overall score. The three attributes that contribute to popularity and reputation are illustrated in Figure 3.2.

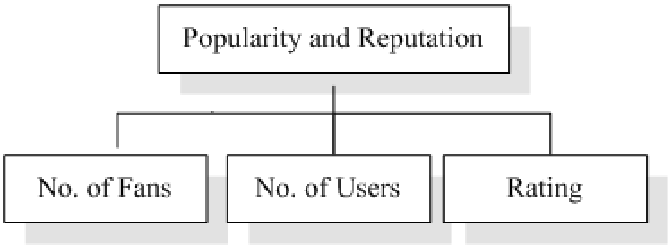


Figure3.2 Three attributes about popularity and reputation of social application

Although we could obtain the above numbers about the social application from websites, we still lack of information to decide whether they are meaningful to users or not when they come to the decisions. People might consider differently about the preferences or priorities of the popularity and reputation. For example, some people would stress more attention to popularity of application since they might want to expand their social boundaries by meeting new people while some others tend to be more affected by reputation because they think that the reputation could possibly reflect the truly using experiences. Still, there are others considering both factors without noticing the degrees. Because the preferences or priorities of popularity and reputation of people are diverse, a questionnaire survey is given to the active social network users in order to deduce the general weights of these three attributes. After the average weights are derived from empirical study, we utilized these weights to calculate the Popularity and Reputation score (PR_S) of each application in the following. Denote $D(A)$ as a decision matrix where $A=\{a_i, \text{for } i=1,\ldots,m\}$ is a set of alternatives representing available social network based applications and $a_i=\{a_{ij}$, for $j=1,2,3\}$ symbolizing the criteria values of "No. of Users", "No. of Fans" ,and "Rating" of that alternative respectively. The PR_S of each alternative is calculated as

$$PR_s(a_i)=\sum_{j=1}^{3} a_{ij} w_j \quad (4)$$

where a_{ij} is the element in the normalization decision matrix and w_j stands for the weight of criteria.

3.2 Users' Preference Analysis

Knowing users' preference perfectly is one of the key successful factors in the recommendation system. By analyzing users' social application using situation in the past, we can better understand users' preference and therefore could further recommend items with highly suitable and interest to them. Most of social network platform providers have categorized applications into several predefined categories in order to help users to search more easily by categories. Suppose under the circumstance of 1-1 relationship between the provided category and the application, user's preference of each category type could be inferred as follows. Denote $A=\{a_{1,...,}a_n\}$ means a set of social network based applications and $C=\{c_{1,...,}c_m\}$ represents a set of predefined categories on the social network platform. To consider the preference weight for the target user u_i to each category, we collect and exploit target user's application using histories including usage frequency. For $\forall c_j \in C$, $\mathrm{Sum}_{u_i}(c_j)$ is a summation of the usage frequency of the social application belonging to category c_j, the formula is defined as

$$Sum_{u_i}(c_j)=\sum_{a_k \in A(c_j)} fq_{(u_i)}(a_k), \quad (5)$$

where $A(c_j)=\{a_k \mid c_j \in C(a_k)\}$ denotes a set of applications which related to category c_j and $fq_{(u_i)}(a_i)$ indicates u_i's usage or participant frequency of a_i. The preference weight (PW) of category c_j for user u_i is formed as

$$PW_{u_i}(c_j)=\frac{\mathrm{Sum}_{u_i}(c_j)}{\sum_{j=1}^{m} Sum_{u_i}(c_j)}, \tag{6}$$

The preference weight of the corresponding category stands for the attractive strength of that kind of social application type. Thus we can predict other social applications that target users haven't noticed or used yet by evaluating their preference score (P_s) in the below function.

$$P_s(a_i)=PW_{u_i}(c_j), \tag{7}$$

3.3 Social Similarity and Social Interaction Analysis

Social similarity and social interaction are two essential factors utilized to analyze static and dynamic dimensions of human's social behaviors on social network platform respectively. Social similarity aims to consider implicit social behavior, such as "friends in common" and "application used in common", in order to recommend social applications with similar social circle and alike application taste. The more friends-in-common of the two people, the higher connection level and influence probability might lie between them. Because when two people have many mutual friends, it is highly likely that they are quite linked and closed in the social society and simultaneously has a greater possibility to be influenced through others' information flows in the same network. Additionally, "Mutual application" is another key element of similarity that refers to the idea of collaborative recommendation. Friends who have applied many social applications in common are more likely to have similar preference tastes. Thus their applied social applications could have tremendous possibility to be regarded as recommendation candidates with mighty subject attraction and high friend's participating rate (high social attraction).

Comparing with social similarity, social interaction is a more dynamic relation that contains all kinds of actions happening among people. For example, users might simply comment on others or apply the function ability of applications to send birthday cards to, give hugs to, or play games with friends. These social actions reveal not only the social intimacy level of the two people, but also the fondness level of interacting with each other through the function of application. It is assumed that people with high intimacy level would lead to high interaction frequency, and would accordingly further result in high social application usage rate, since most of interactions are supported by applications. Therefore, for people who often use social applications to interact with intimidate friends would be more attracted to and pay attention to those social applications that close friends are participating in. In other words, if the recommendation comes from a friend with strong social interaction strength, the fascinating power to the user would be greater than those come from friends with low social interaction level.

For the purpose of recommending social applications based on user's social behavior and relationship situation on the social network platform, a user and his/her friends' social similarity and social interaction information are retrieved to calculate an overall social relation score (*SRs*), representing the application's social attraction and appealing level for the recommended user.

3.3.1Social Relation score (*SRs*) calculation

To calculate the social relation score of social application, we firstly make two definitions (dofinition1&2) about the structural pattern of social network and social application according to the similar concepts in [24]. Furthermore, to be more specific for the target user that we focus to recommend for, an expending definition (definition 3) is applied.

Definition 1 (*Social network layer*) Social network layer, SNL, is defined as

$$SNL = \langle U, N_U, P \rangle, \tag{8}$$

where $U = \{u_1, ..., u_u\}$ is a set of users on the social network, and $N_U \subseteq U \times U$ is a set of friendship links and interaction links between users. $P = \{p_{u_1}, ..., p_{u_u}\}$ means a set of corresponding main pages on the social network website displaying posts, comments, and all kinds of interactions related to the corresponding user in U.

Definition 2 (*Social Application layer*). A social application layer, SAL, which is an abbreviation of social network based application layer, is defined as

$$SAL = \langle A, N_{U \times A} \rangle, \tag{9}$$

Where $A = \{a_{1,...,}a_n\}$ means a set of applications supplied on the social network website and $N_{U \times A} \subseteq U \times A$ is a link set between users and applications representing the usage relationship.

Definition 3 To further expand definition1 for a target user u_i on SNL and definition 2 for u_i's applications' using situation, we define u_i and his/her friends as

$$SNL_{u_i} = \langle F(u_i), N_{F(u_i)}, SR_{(u_i)}, p_{u_i} \rangle, \tag{10}$$

where $F(u_i) = \{f_1, ..., f_m\}$ is a set of u_i's friends on SNL. $N_{F(u_i)} \subseteq u_i \times F(u_i) \subseteq N_U$ is a set of interaction links between user u_i and his/her friends . $SR_{(u_i)} = \{SR_{(u_i f_1)}, ..., SR_{(u_i f_m)}\}$ indicates the corresponding weight of social relation in $N_{F(u_i)}$. It is noted that because from u_i's view point the most related users on social network are his/her friends; we emphasize a target users' social relationship on friendship. And the definition of u_i's application usage situation is defined as

$$SAL_{U_i} = \langle A(u_i),\ N_{u_i \times A(u_i)} \rangle , \tag{11}$$

where $A(u_i)$ means a set of applications that u_i uses and $N_{u_i \times A(u_i)} \subseteq u_i \times A(u_i) \subseteq U \times A$ is a link set between u_i and social applications, representing the usage relationship. The structural relationships of variables in definition 3 are demonstrated in Figure 3.3 .

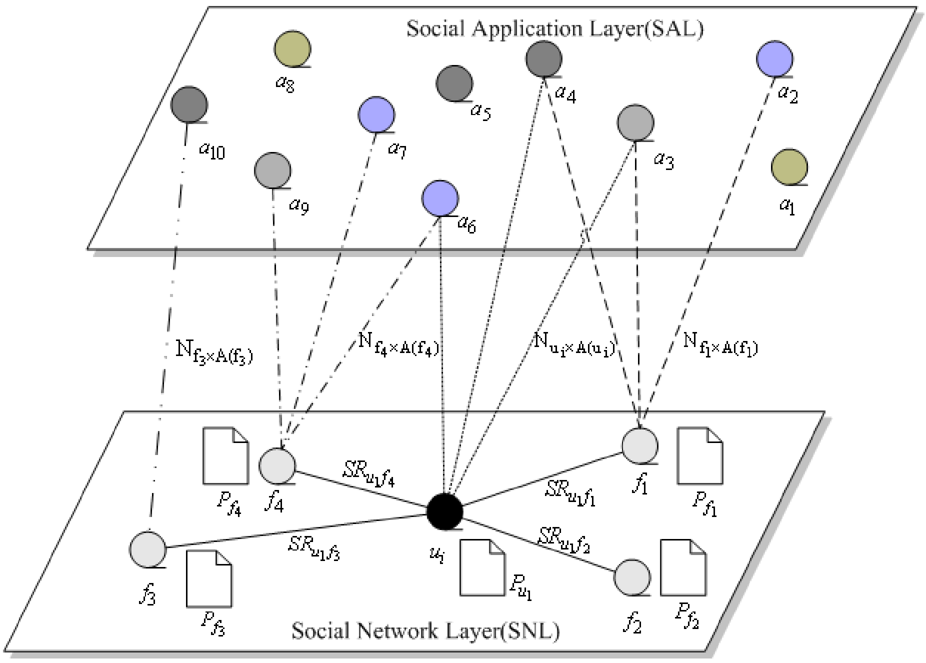


Figure 3.3 The structural relationship of *SNL* and *SAL*

The Social Relation (SR) between a user u_i and his/her friend f_j is composed of their Social Similarity (SS) value and Social Interaction (SI) value as:

$$SR_{(u_i, f_j)} = SS(u_i,\ f_j) + SI(u_i,\ f_j) , \tag{13}$$

The details of the formulation are deliberated as following.

Firstly, given a user u_i and one of his/her friend f_j , the Social Similarity (SS) between them is defined as

$$SS(u_i, f_j) = Sim_T\left(u_i, f_j\right) + Sim_{FC}\left(u_i, f_j\right), \tag{12}$$

where $Sim_T\left(u_i, f_j\right)$ represents the similarity level of their social application's tastes and $Sim_{FC}\left(u_i, f_j\right)$ indicates their similarity degree of friends in common, i.e., the friend consensus similarity.

The similarity of taste $Sim_T\left(u_i, f_j\right)$ is formulated by

$$Sim_T\left(u_i, f_j\right) = \frac{\left|\left\{A_k \middle| A_k \in A(u_i) \cap A(f_j)\right\}\right|}{A(u_i)}, \tag{13}$$

where $A(u_i) = \{a_k \mid \langle u_i, a_k \rangle \in N_{u_i \times A(u_i)}\}$ and $A(f_j) = \{a_k \mid \langle f_j, a_k \rangle \in N_{u_j \times A(u_j)}\}$ indicate a set of applications used by user u_i and f_j respectively. And the similarity of friend consensus $Sim_{FC}\left(u_i, f_j\right)$ is evaluated in the formulation of

$$Sim_{FC}\left(u_i, f_j\right) = \frac{\left|\left\{F_k \middle| F_k \in F(u_i) \cap F(f_j)\right\}\right|}{\text{Max}\left[F(u_i),\ F(F_j)\right]}, \tag{14}$$

where $F(u_i) = \left\{f_k \middle| \langle u_i, f_k \rangle \in N_{u_i \times A(u_i)}\right\}$ denotes a set of u_i 's friends. Secondly, the Social Interaction (SI) between u_i and f_j is denoted as

$$SI\left(u_i, f_j\right) = \frac{\left|Comment(f_j)\right| + \left|Interaction(f_j)\right|}{\text{Max}[Comment(P_{U_i}) + Interaction(P_{U_i})]}, \tag{15}$$

where $\left|Comment(f_j)\right|$ ($\left|Interaction(f_j)\right|$), which belongs to $N_{u_i \times A(u_i)}$, stands for the total number of comments (interactions) that f_j commends (interacts) to u_i and $\text{Max}[Comment(P_{U_i}) + Interaction(P_{U_i})]$ points out the maximum number of addition of comments and interactions from friends in $F(u_i)$. It is notable that interactions in this research contains all the actions supported by any application on the social network websites, such as giving and taking and sending and receiving actions.

After every Social Relation (SR) among target user and all of his/her friends are estimated, we can utilized them to finally calculate the Social Relation Score (SR_S) of applications in order to recommend target user something with high social attraction. The Social Relation Score (SR_S) is defined as

$$SR_S(a_i)=\sum_{f_k \in FUA(a_i)} SR_{(u_i,f_k)} \quad , \tag{16}$$

where $FUA(a_i)=\{ f_k | \langle f_k , a_i \rangle \in N_{X\times Z} \}$ means a set of u_i's friends using a_i. Notably, all of the value should be normalized before being computed in the formulation.

3.4 Neural Network-based Recommendation Mechanism

In this sector, we aim to combine popularity and reputation score (PR_S), preference score (P_S) and social relation score (SR_S) which are derived from the former analysis modules. The recommendation score $R_{u_i}(a_i)$ of the social network based application a_i for the user u_i is defined as following:

$$R_{u_i}(a_i) = \alpha PR_S{}^N(a_i) + \beta P_S{}^N(a_i) + \gamma SR_S{}^N(a_i) , \tag{17}$$

where the uppercase N of PR_S , P_S and SR_S stands for the scores after the process of normalization. Parameters α , β and γ which are individually between 0 to 1 are the system-set weights totally accumulated equal to 1.

In the purpose of modeling the optimal way of combining these three implicit related factors in order to significantly represent users' decision preference, a back-propagation neural network is adopted. A BPNN model is one of the most frequently used techniques for classification and prediction. Its' special abilities of accommodating complex and non-linear data and learning implicit relations lying behind the scene support us to deal with modeling and forecasting demands. The process of applying BPNN is described in the following. The initial recommendation list of top-k social network based applications is delivered to the target user by a web-based interface (see Appendix A).For each target user, he/she would review the recommendation results along with all the required information about these social applications, such as users number, fans number, rating, category, friends lists with social relation weightings individually and a hyperlink to the main page for further detailed description of that application. Users are required to make unbiased evaluations by scoring each application according to their own preference and conception. After users' feedbacks are obtained, they are put into BPNN model to systematically learn the weights of α, β and γ respectively through the neural network. Once the training is done, the trained BPNN model would be applied to compute the forecasted recommendation score $R_{u_i}^{F}(a_i)$ and generate the recommendation list.

CHAPTER 4 Experimental Study

So far, in the previous chapters, we have introduced the system framework and the corresponding modules of social network based application recommendation mechanism. In this section, to empirically examine the effectiveness and satisfaction of the proposed model, we select Facebook, which is the sixth most-trafficked website in the world[29] with over than 200 million active users and more than 52,000 currently available application [16], as the experiment platform.

The details of the experiment are organized as follows. Initially we describe the data collection process and analyze data characteristics, secondly illustrate the result of statistic survey which reveals general users' perceptions of the weighting relationships among application's attributes (No. users, No. Fans and Rating) when making their usage decision. Lastly, the recommendation experimental results and evaluations are demonstrated.

4.1 Data Descriptions

In the early April, 2009, we invite 44 active Facebook users aging between 23- 30 as our target users , and with the permission given from them, we start to crawl their personal profile pages, application pages (Appendix B), and friend lists to collect personal comments, interactions, application usage situations and tastes that happened within a past year. In addition, for every target user, the personal recommendation pool of social network based application is composed of the union of 1) 3000 applications with top popularity ranking provide in Facebook application category page (Appendix C) and 2) all applications which a target user's friends are using but he/her has not used. The detailed statistics information of this study is presented in the following. First of all, we overview the average distribution trend of the average numbers of 44 target users' "number of friends", "number of social applications used", "number of comments and interactions that happened between target user and friends" and "number of social applications in the first degree friendship", which means the number of social applications that friends are already using but the target user yet to use".

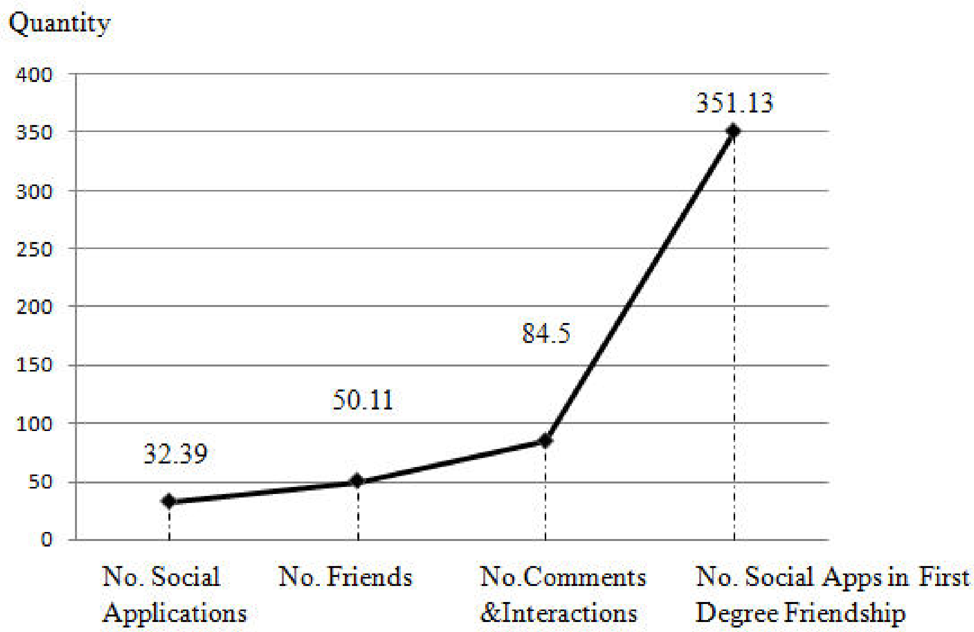


Figure 4.1 The average numbers of users' data attributes

As shown in Figure 4.1, the average number of social applications is 32.39, the average number of friends lies at 50.11, the average number of comments and interactions places at 84.5 and the average number of social applications in the first degree friendship is 351.13. It is amazed to discover that averagely there are over ten times of social applications lie in users' first degree friendship that are probably still un-noticed to the target users. This implies a great opportunity to leverage friends' usage social applications as recommendation candidates since they have more social attraction powers than others. A detail distribution of every target user's situation is present in Figure 4.2. The quantities of social applications are diverse majorly between one hundred and six hundreds.

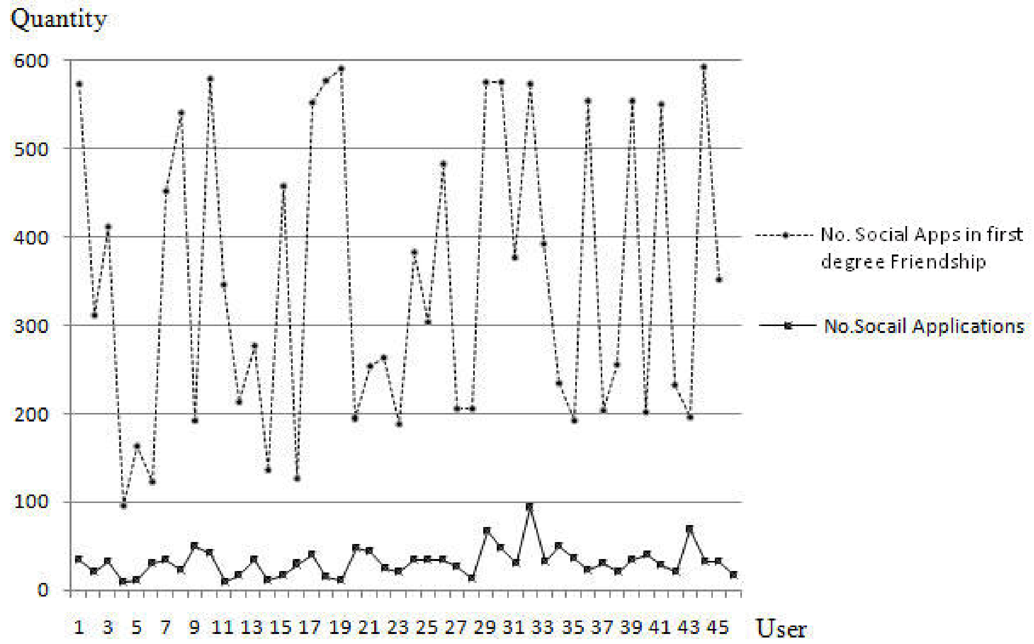


Figure 4.2 The distribution of No. social applications and No. applications in first degree friendship

Next, we zoon in to analyze the relationships among “number of friends”, “number of social applications”, and “number of comments and interactions” illustrated in Figure 4.3.

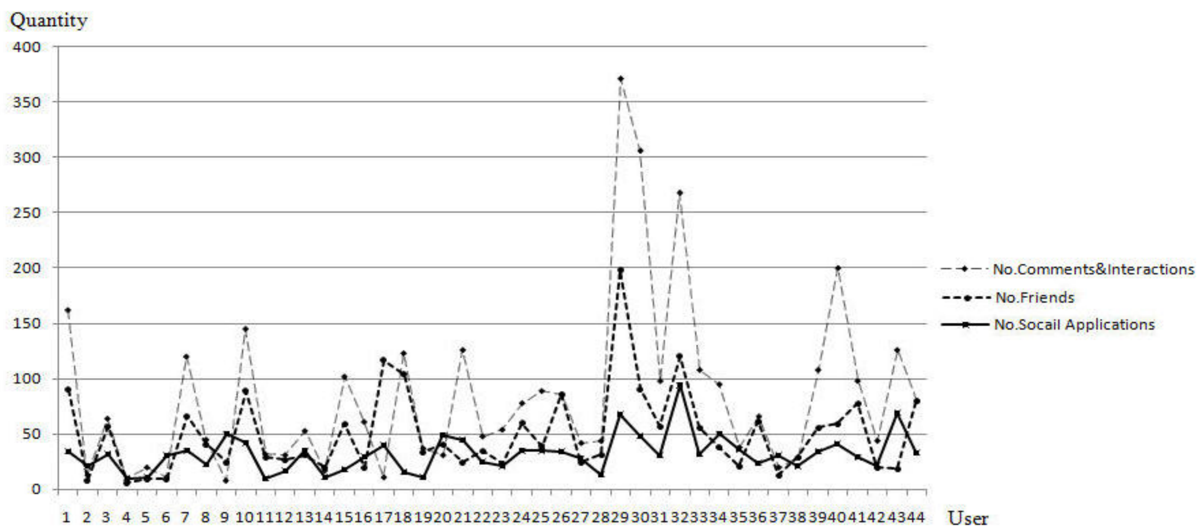


Figure 4.3 The distribution of target users’ No. Friend, No. Social Application and No. comments and interactions

The lines in Figure 4.3 fluctuate almost simultaneously upwards and downwards. There is likelihood that these three parameters may be associated. We further tend to unfold the assumption by performing a correlation test (see in Table 4.1).

Table 4.1 Correlation of user's

	Pearson Correlation		
	No. Social Application	No. Friends	No. Comts & Interactions
No. Social Application	**1**	**0.649****	**0.497****
No. Friends	**0.649****	**1**	**0.762****
No. Comts & Interactions	**0.497****	**0.762****	**1**

**. Correlation is significant at the 0.01 level (2-tailed)

As can been seen in Table 4.1, the three variables are significant correlated to each other. Therefore, based on the statistic evidence derived, we would go advance to infer the possible meaning lying behind. It is likely that the more friends a user has the higher possibility of him/her to be exposed in a condition of being invited or influenced by friends, thus having more chances to apply new social applications. Or it might be probable for a user to joined a social application due to its' social attraction brought from a friend with high comments and interactions level. Therefore, to recommend appropriate social applications to a target user, analyzing his/her social relationships would be a right way to go.

According to the collected data in the middle of May, 2009, social network based applications on Facebook are found to be categorized into eight types. Based upon thousands of collected social application data, the pie chart demonstrated in Figure 4.4 displays the quantity percentage for each application category.

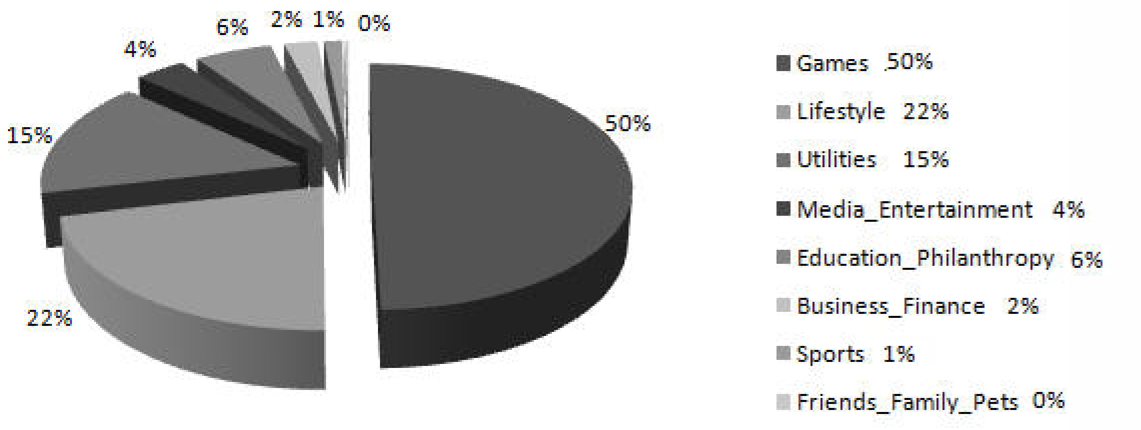


Figure 4.4 Types and distribution of Facebook category

In the pie chart, the "Game" category is in possession of one half amounts of numbers. Nearly one fourth of application is "lifestyle" and more than one eighth of numbers is "utilities". The application quantity of a specific category might reveals market tendency and users' needs for that kind of application. Evidence can be given by comparing it with the average distribution of 44 target users' preference weights, which are calculated according to users' usage histories. As shown in Figure 4.5, users' preference weights of categories appear to reflect the quantity of that application.

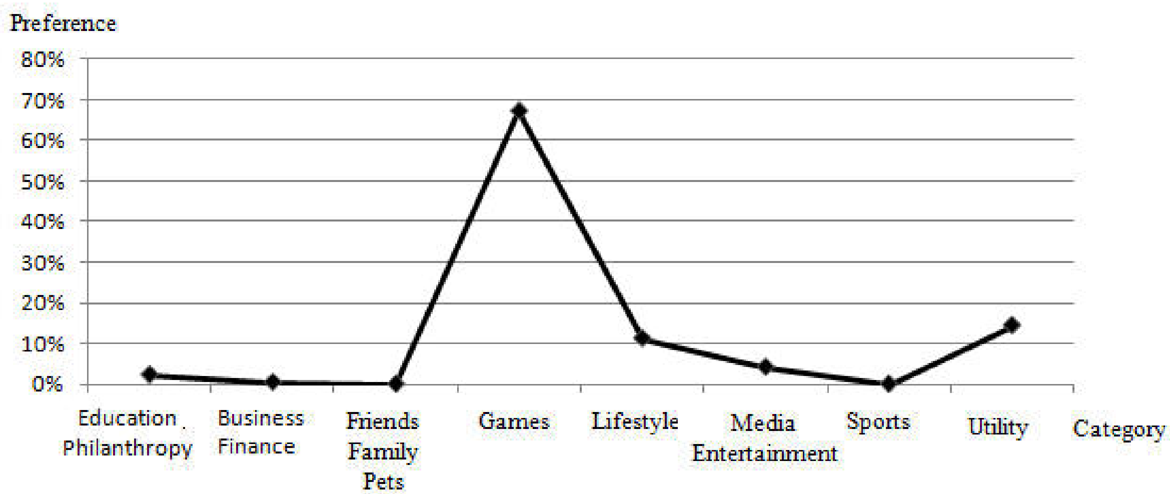


Figure 4.5 Users' average preference of application categories

4.2 Results of Empirical Survey for Popularity and Reputation.

"Number of users" and "rating", as clues to indicate popularity and reputation, are ordinarily available information provided on website for people to judge the possible quality and performance of a product or service. "Number of fans", which is newly used in the web, has gradually raised user's attention. However, we do not know their actually meaning and relative importance in user's decision making situation, especially in a brand new service, social network based application. Thus in our research, we invite 41 active social application users on Facebook to fill the relative weightings of the three attributes according to their decision experience and preference. The web based questionnaires can be referred to Appendix D. Based on our survey, we find out that people would think more highly about the information of No.Users with a degree of 0.48 and think almost equally of No. Fans and Rating scores as 0.25 and 0.27 respectively (see in Figure 4.6). It is notable that Facebook use "monthly active users" instead of the number of users.

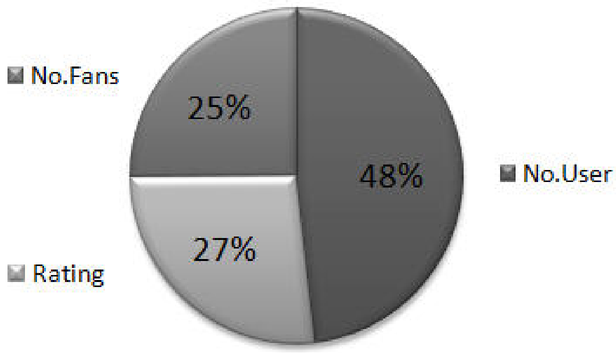


Figure 4.6 The weights of No. Users, Rating and No. Fans derived from empirical study

4.3 Recommendation Strategies

In this research, we design seven different recommendation strategies to evaluate the correctness of system design. The following are the strategies we use:

1) All + BPNN

This is the method proposed in the research that utilizes back-propagation neural network model to combine objective information (popularity and reputation scores) of social network based application and subjective information (preference score and social relationship score) of personal concerns.

2) All + AHP

Although Artificial Neural Network has been proved as an effective method to deal with unknown weighting problems, however, it is still a black box that reveals nothing about the actually value of weights. To make up for the drawback of BPNN, we experimentally attempt to fix the block box problem by using Analytic Hierarchy Process method to deliberate users' decisions making preference on social network based application. In this strategy, the weightings (α , β and γ) of *PRs*, *Ps* and *SRs* are derived from historical data, i.e., the initial feedbacks in this experiment.

3) All

All stands for the initial recommendation result that lacks of BPNN to learn the non-linear relationships between *PRs*, *Ps* and *SRs*. It can be regarded as a benchmark to prove whether it is necessary to adjust the weightings. In this study, we set $\alpha = 0.3$, $\beta = 0.3$ and $\gamma = 0.4$.

4) SRs +Ps

It would be interesting to exam the effect of only taking personal subjective information as the recommendation criteria. Therefore, in this strategy, we set $\alpha = 0$, $\beta = 0.5$ and $\gamma = 0.5$.

5) SRs

What if we only consider the influence of social relation? In this study we try to test the impact of social similarity and social intimacy, setting $\alpha = 0$, $\beta = 0$ and $\gamma = 1$.

6) PRs

PRs stands for the merge score of "No. User", "No. Fans" and "Rating". It is a brand new combination proposed in this paper therefore requires experimental validation. In this study, we set $\alpha = 1$, $\beta = 0$ and $\gamma = 0$.

7) Pop

Pop, which is the abbreviation of popularity, is widely used in most of the recommendation system. Facebook, for example, temporarily ranks their applications by the number of monthly active users. Thus, we select pop as the basic recommendation benchmark.

4.4 Experimental Design and Evaluation Results

According to the report that user often only accesses to the documents/articles lying on the first two page[21], which is usually 10 results per page, we design to split our recommendation strategies and benchmarks into two stages in order to fit the designed experiment process and keep the total number of list items close to 20s.

The experiment is divided into three stages. In the first stage, a recommendation list which unites each top 7 applications of strategy "*Pop*", strategy "*PRs*", strategy "*SRs*", strategy "*SRs +Ps*" and strategy "*All*" are delivered to the target users with a five point ranking scale, which starts from "strongly willing to use", "willing to use", "ok to use", "not willing to use", "strongly not willing to use". This five point ranking, which represents both the acceptance of the recommendation result and user's satisfaction, is used as an evaluation index. Among the 44 target users, we separate them into two groups, which are 30 and 14. The previous group is the main target users invited to evaluate all the seven recommendation strategies. The latter group is arranged for training the back-propagation neural network. In the second stage, based on 14 target users' feedbacks derived from first stage, we could train the neural networks to learn users' general implicit decision preference of the usage of social applications and therefore we could use the trained model to predict the recommendation scores ($R_{u_i}(a_i)$) of the social applications in the main 30 target users' data pools. Also, the initial feedbacks of all the target users in the two groups are all collected and put into Analytic Hierarchy Process model to calculate the relative weightings of the three factors and accordingly predict each social application's recommendation scores ($R_{u_i}(a_i)$) for the main 30 target users. In the third stage, we collect each top 7 recommendations from strategy "*All + BPNN*", strategy "*All + AHP*" and strategy according to individual's personal result, and present the union lists to target users.

4.4.1 Training Neural Network for Prediction

Since the BPNN model is one of the most widely used ANN models, general commercials ANN software package (e.g. NeuroSolution 5, NeuroShell 2, NeuFrame etc) can be applied even if users are beginners. In this part, we utilized the expert mode of NeuroSolution 5 to conduct BPNN. The expert mode provides three levels of neural network complexity, which are low, medium and high. In the low complexity level, we find out the network is composed of one input layer, one hidden layer, and one output layer; meanwhile, the medium and high level are both one input layer, two hidden layer and one output layer. By iteratively trying, we find out that two hidden layers' network outperforms one hidden layer. And under the two hidden layer network, we have tested (PEs_1, PEs_2) for pairs of (10, 5), (20, 10), (30, 15) and (40, 20), where PEs_1 stands for the processing elements in the first hidden layer and PEs_2 represents the second, and finally discover that (20, 10) has lowest Mean Squared Error for both testing data and cross validation testing data. The result is presented in Figure 4.7. It shows that the learning curve of Mean Squared Error (Testing) during 1000 epochs quickly drops down toward 0.07 and level off. The Mean Squared Error (Cross Validation) is vibrating between 0.8 and 0.9. The Mean Squared Error (MSE) is denoted as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2 , \tag{18}$$

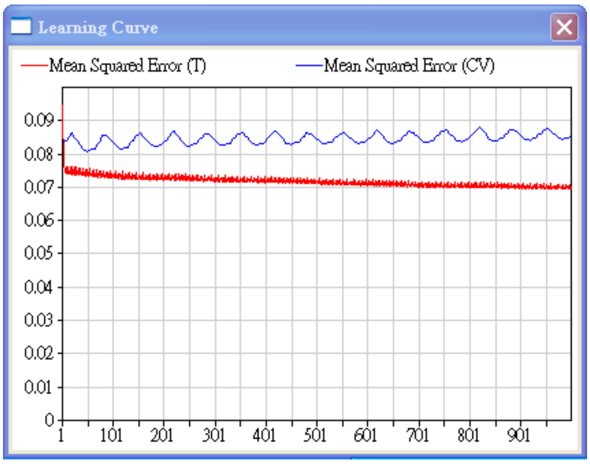


Figure 4.7 the MSE value of the trained neural network

4.4.2 Calculating Analytic Hierarchy Process

Analytic hierarchy process (AHP) [44] is one of well-known methods to treat multi-criteria decision making (MCDM) problems. By mathematical pair-wise comparison, AHP determinates relative importance or weights of criteria that supports human to make thorough decision. It has been applied in many research fields like product recommendation[33] and tourism recommendation [23, 50].

Generally, four stages are involved in using AHP for attribute weighting calculation. Firstly, a decision matrix including the value of each criterion for each alternative is constructed. Secondly, after the decision matrix was built, normalization is needed to eliminate the dimensional effect of multiple attributes in an alternative. Thirdly, instead of requiring expertise or users to personally define their preference weights of the criteria in the pair-wise comparison matrix, in this research, we tend to utilize users' histories to construct the initial values of the preference weights based on the calculation method described in [50]. Suppose $H(A)$ denotes as a set of target user's used application histories which contain elements a_{ij} in the normalized decision matrix and P_{xy} is the relative preference weight of x criterion to y criterion. The formulation of P_{xy} is expressed as

$$p_{xy} = \sum_{i=1}^{k} a_{iy} \Big/ \sum_{i=1}^{k} a_{ix} \, , \tag{1}$$

where $a_{ix}, a_{iy} \neq 0$, for i=1,2...,k is the number alternatives and j stands for the number of criterion. Based on the value of P_{xy} derived from the normalized decision matrix, the preference weight values of criteria could be next calculated from the pair-wise comparison matrix. Let comparison matrix B be a $n \times n$ matrix in which element b_{ij} denotes the relative preference weight of i criterion in terms of j criterion and formulates as

$$b_{ij} = \begin{cases} p_{ij} & \text{if } i > j \\ 1 & \text{if } i = j \\ p_{ij}^{-1} & \text{if } i < j \end{cases} , \tag{2}$$

Fourthly, to derive the relative weight of the criteria from the comparison matrix B, a geometric mean method is used as follows.

$$w_i = \left(\prod_{j=1}^{n} b_{ij} \right)^{1/n} , \tag{3}$$

where w_i is a relative weight value for an alternative i, and n represents the number of criterion. Notice that according to[7], it is recommended to use geometric mean method instead of eigenvector to avoid problems of left-right eigenvector asymmetry and dependent of relative measurements among alternatives.

In this paper, we tend to leverage AHP method to model the possible weightings of the three parameters, which are *PRs*, *Ps* and *SRs*, from the initial feedbacks. In order to confidently model users' decision criteria, we only select social network based applications from users' initial feedbacks that are equal or above three points. Based on the total 734 samples, the weights are derived in Table 4.2.

Table 4.2 Weights of *PRs*, *Ps*, *SRs*

Weight of *PRs* (α)	Weight of *Ps* (β)	Weight of *SRs* (γ)
0.314472340531904	**0.334456703979903**	**0.35107095548819**

As shown in Table 4.2, the three values are roughly one third individually, however, there are a point two degree difference among 1) *PRs* and *Ps*, 2) *Ps* and *SRs*, and a point four difference degree between *PRs* and *SRs*. From this AHP model, we might be able to know the truly relationship among the three factors and discover the fact that social relationship, including users' social intimacy and similarity, plays the most important part in the decision of using social network based service, followed by firstly user's self-preference and secondly social application's popularity and reputation.

4.4.3 Users' Evaluation Results and Discussion

In this section, we would present users' evaluations of the seven social application recommendation strategies. Based on the reviews given from main 30 target active Facebook users, the average rating scores of the each strategy are given in Figure 4.8. The average results are ordered from low to high, left to right. As can be seen, strategy "*All + BPNN*" receives the highest average rating, followed by strategy "*All + AHP*" and "*All*". Social related strategies such as "*SRs +Ps*" and "*SRs*" have better scores than popularity and reputation ("*PRs*") or popularity ("*Pop*") only.

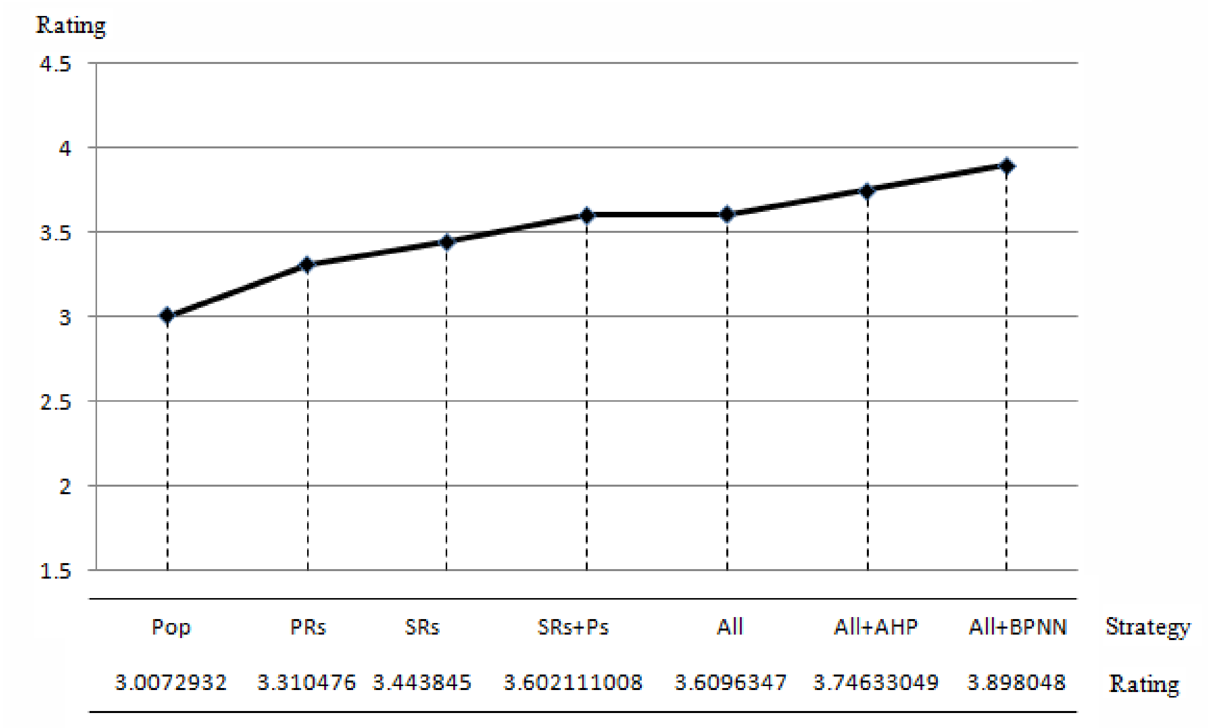


Figure 4.8 The average rating result of seven recommendation strategies

A statistical test (e.g. paired sample t-test) is used to further confirm the significant difference of the proposed recommendation result (see Table 4.3). At 95% significant level, all the test results show that strategy "*All+BPNN*" is significant under 0.05 in terms of the rests. Therefore, it proves that the proposed recommendation strategy is the best compared to other strategies.

Table 4.3The statistical verification results of "*All+BPNN*" versus the others

Paired Group		T-value	Sig.(2-taild)
All+BPNN &	***All+AHP***	8.019	0.000
	All	7.234	0.000
	SRs + Ps	5.326	0.000
	SRs	6.799	0.000
	PRs	8.206	0.000
	Pop	9.464	0.000

After briefly presenting the average rating score of the seven recommendation strategies, we furthermore focus on describing and comparing sets of strategies. Firstly, detail information and a statistic test about strategy "*Pop*" and "*PRs*" are presented.

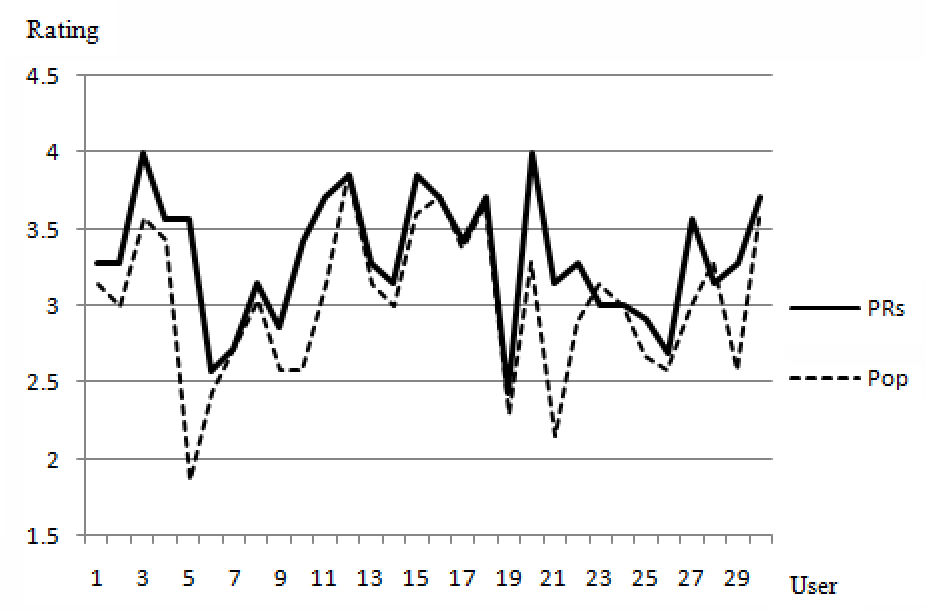


Figure 4.9 Rating distributions of *PRs* and *Pop*

Figure 4.9 illustrates the rating distribution of "*Pop*" and "*PRs*", clearly indicating that "*PRs*" nearly dominates "*Pop*" from every user's rating. By statistic testing, Table 4.4 demonstrates a significant difference between them.

Table 4.4 The statistical verification results of "*PRs*" versus "*Pop*"

	Mean	Std. Dev.	Pair T-Test	T-value	Sig.(2-tailed)
PRs	3.310	0.422	***PRs - Pop***	4.247	0.000
Pop	3.007	0.493			

It is important to verify that combining "No. Users", "No. Fans" and "rating" is a better way to evaluate application's overall popularity along with reputation.

Next, we aim at deliberate personal subjective factors to compare the impacts of social relations only and social relations added with application's preference. The following Figure displays the rating results of strategy "*SRs* + *Ps*" and "*SRs*".

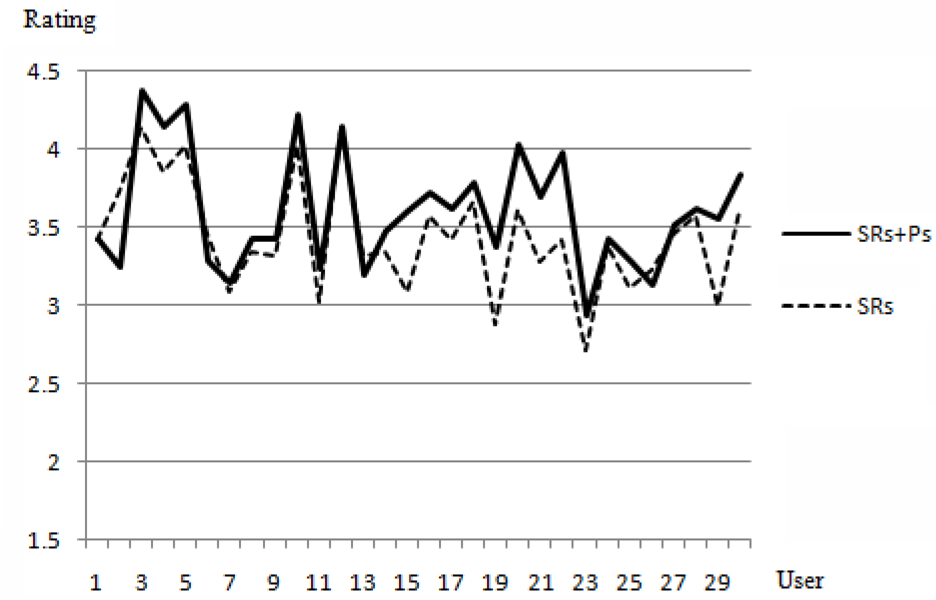


Figure 4.10 Rating distributions of "*SRs + Ps*" and "*SRs*"

Figure 4.10 shows that "*SRs + Ps*" is obviously greater than "*SRs*" in some area and lightly greater than "*SRs*" in the most area. To further analysis, a t-test is done and shown in Table 4.5.

Table 4.5 The statistical verification results of SRs + Ps versus SRs

	Mean	Std. Dev.	Pair T-Test	T-value	Sig.(2-tailed)
SRs + Ps	3.602	0.378	***SRs + Ps – SRs***	2.275	0.030
SRs	3.443	0.360			

As can been seen, a statistic proof has shown that "*SRs + Ps*" is significantly different from "*SRs*", and therefore we could confidently infer that "*SRs + Ps*" works better to fit users' personal concerns by taking both social relation and preference tastes into consideration.

Lastly, we intent to compare the influence of weighting modification process, and go beyond to analyze the consequence of two different weighting modification methods. Since there is only a slightly rating difference, which is 0.3%, between "*SRs + Ps*" and "*All*", it is definitely required a more sophisticated procedure to deal with the weighting issue. In this paper, we select a black box method and a white box method to exam the result respectively. Detail information of "*All + BPNN*", "*All + AHP*" and "*All*" is displayed in Figure 4.11 and the t-test is shown in Table 4.6.

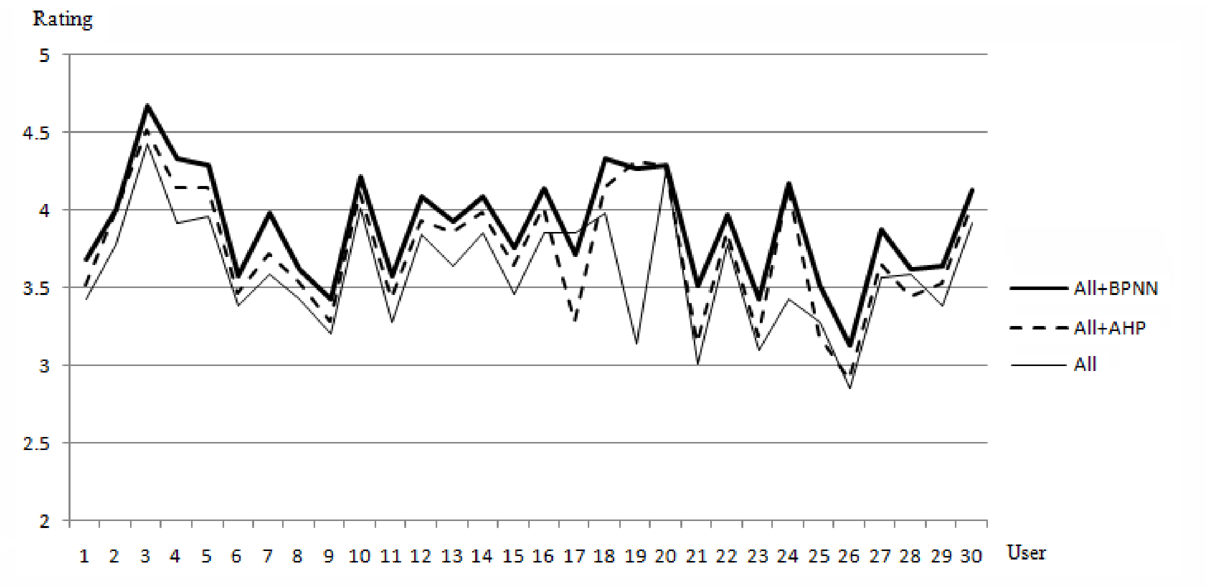


Figure 4.11 Users' rating of strategy "All + BPNN", "All + AHP" and "All"

Table 4.6 The statistical verification results of All + BPNN", "All + AHP" and "All"

	Mean	Std. Dev.	Pair T-Test	T-value	Sig.(2-taild)
All+BPNN	3.898	0.355	***All+BPNN - All+AHP***	8.019	0.000
All+AHP	3.746	0.408	***All+BPNN - All***	7.234	0.000
All	3.609	0.371	***All+AHP - All***	2.775	0.010

Based on the results shown above, both "*All + BPNN*" and "*All + AHP*" outperform "*All*". It suggests that under these two weighting modifications, the recommender service could better catch users' decision preferences and concerns, providing more suitable recommendation results. In addition, by comparing "*All + BPNN*" with "*All + AHP*", we could probably make a brief conclusion that BPNN as a more sophisticated mathematical modeling method, dose a better job than AHP that used in this paper. However, since "*All + AHP*" could help expose the truly weights of the three factors, it would give the decision maker a hint of how to deal with the social network based recommendation.

Although Analytic Hierarchy Process does not predict humans' decision making as well as Back-propagation neural network, however, it seems correctly uncover the decision strengths that could be partially proven in the outcomes of the average rating scores of the seven recommendation strategies (see in Figure 4.8). The rating scores of strategy "*SRs* + *Ps*" and "*SRs*" are higher than "*PRs*" or "*Pop*", exhibiting the similar importance calculated from AHP ("*SRs*" >"*PRs*").

In the end of this section, we would discuss and summarize the meaning of the recommendation results. Firstly, it is proved that the combination of the three parameters ("number of users", "rating" and "number of fans") is superior to the basic recommendation method, popularity, in extracting service's explicit performance. Users in their decision process would consider all the information that portrays the popularity and reputation together rather than popularity only. Secondly, social relation is justified as a more important factor to users than popularity and reputation. Both social related recommendation strategies are more outperformed. Thirdly, users would also think highly of service content. Whether the recommended service fits user's taste or not would play an essential role in recommendation service. Lastly, using back propagation neural network model to extract users' decision criteria weightings is more accurate than utilizing Analytic Hierarchy Process method. It might because that AHP is a one-time-calculation method that uses summarized numerical difference to infer the possible weightings, however, BPNN takes an iterative training process to model the relative weights. Nevertheless, AHP method does help us to uncover the mask of the possible weightings of "social application's popularity and reputation", "user's preference" and "social relations of applications", giving us an guiding principle for social network based services recommendation.

CHAPTER 5 Conclusion and Future Works

5.1 Conclusion

With the prosperity of social network based websites, more and more people have joined to use or plugged into developing the social network based services. This tremendous growth of social network based service leads to some dilemmas. For users' concern, it is hard and inconvenient to discover attractive services from tons of social services; for developers and cooperated sponsors' concern, it seems to lack of channels to promote themselves to extensive users that are interested in them. Therefore, in this paper, we propose a social network based service recommendation mechanism which combines aspects of services' popularity and reputation, users' service preferences and social relationships.

The contributions and interesting findings of this paper are summarized as follows. First, little researches have done by systematically studying the recommendation of social network based service. We are one of the pioneers to empirically perform an experiment of recommending social applications on the social network platform. Second, it is justified that the proposed mechanism outperforms other benchmarks by having greater rating results that satisfy users' expectations and requirements. Third, by empirically investigating users' decision perceptions regarding to the information of "rating", "number of fans", and "number of users" of a product or service, it is found that the information of "number of users" plays the most important part in users' decision making of whether to use or not. Also, by the calculation of Analytic Hierarchy Process, we uncover the priorities of "service's popularity and reputation", "user's preference" and "user's social relations". It is suggests that when a user comes to a decision, he/she would regard service' social attraction and relationship as the most important part, followed by "user's preference" and next "service's popularity and reputation". Both the proposed social network service recommendation mechanism and the derived decision preferences and priorities from users' points of view would help social network providers to well customized their users' recommendation lists and further improve their customer relationship management.

5.2 Limitation and Future Works

There are some limitations in our research. The methodology of calculating user's preference weights is restricted to one category per application due to the limitation of experimental platform. It is notable that the category design might be different among service platforms. To be more considerate, if the social network platform uses multi-categories to tag each application, presenting a many to one condition, clustering technique would require re-categorize. Moreover, in practical, due to the restriction of time, there is a little possibility for every target user to try every single category in the pass. Thus, to deal with the problem of empty preference weight of a specific category, a transmitting method might be suggested.

In spite of only depending on the service platform's pre-defined categories, it would be more accurate if ontology is constructed to describe social applications or services' attributes. Still, there are some other topics available for improving the research. To further analyze social application's reputation, a review mining would be needed to actually know the opinion of comments. It would be more complete and powerful to integrate popularity, reputation, and review mining together. Also, to recommend social application with high qualities, a QoS investigation and generalization might be required. To enforce the ability to model users' interests and tastes, we need to collect and analyze more personal information to rich users' profile, such as communities that user joined, and subjects that they collect or be fans of. Beside of studying one layer interaction between two people, user's social relationship analysis might expand to the user's communities for further putting the influence of network structural and other peers' interaction into consideration.

References

[1] Acquisti, A., and Gross, R., "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook " *Privacy Enhancing Technologies*, 2006, pp. 36-58.

[2] Adomavicius, G., and Tuzhilin, A., "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, 2005, pp. 734-749.

[3] "Adonomics," 2008; http://www.adonomics.com.

[4] "Artificial neural network," http://en.wikipedia.org/wiki/Artificial_neural_network.

[5] Atif, N., Saqib, R., and Chuah, C.-N., "Unveiling facebook: a measurement study of social network based applications," *Book Unveiling facebook: a measurement study of social network based applications*, Series Unveiling facebook: a measurement study of social network based applications, ed., Editor ed.^eds., ACM, 2008, pp.

[6] Balabanovic, M., and Shoham, Y., "content-based, collaborative recommendation," *Communications of ACM*, 1997, pp. 66-72.

[7] Barzilai, J., "Deriving weights from pairwise comparison matrices," *Journal of the Operational Research Society*, 1997, pp. 1226-1232.

[8] Billsus, D., Brunk, C.A., Evans, C., Gladish, B., and Pazzani, M., "Adaptive interfaces for ubiquitous Web access," *Communications of ACM*, 2002, pp. 34-38.

[9] Blanco-Fernadez, Y., Pazos-Arias, J.J., Gil-Solla, A., Ramos-Cabrer, M., Loepz-Nores, M., Garcia-Duque, J., Fernandez-Vilas, A., Diaz-Redondo, R.P., and Bermejo-Munoz, J., "A flexible semantic inference methodology to reason about user preferences in knowledge-based recommender systems," *Knowledge-Based Systems*, 2008, pp. 305-320.

[10] Burger, C.J.S.C., Dohnal, M., Kathrada, M., and Law, R., "A practitioners guide to time-series methods for tourism demand forecasting -- a case study of Durban, South Africa," *Tourism Management*, 2001, pp. 403-409.

[11] Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y., and Moon, S., "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system " *Proc. 7th ACM SIGCOMM conference on Internet measurement*, 2007.

[12] Deng, W.-J., Chen, W.-C., and Pei, W., "Back-propagation neural network based importance-performance analysis for determining critical service attributes," *Expert Systems with Applications*, 2008, pp. 1115-1125.

[13] Dwyer, C.A., and Hiltz, S.R., "Designing Privacy into Online Communities " *Social Science Research Network Working Paper Series*, 2008.

[14] Enke, D., and Thawornwong, S., "The use of data mining and neural networks for forecasting stock market returns," *Expert Systems with Applications*, 2005, pp. 927-940.

[15] "Facebok Application Dictionary," http://www.facebook.com/apps/.

[16] "FACEBOOK Statistics," 2009; http://www.facebook.com/press/info.php?statistics.

[17] Fausett, L., *Fundamentals of neural networks*, Prentice-Hall, 1994.

[18] Frank Edward, W., Stefano, B., and Frank, S., "A model of a trust-based recommendation system on a social network," *Autonomous Agents and Multi-Agent Systems*, 2008, pp. 57-74.

[19] Fu, F., Liu, L., and Wang, L., "Empirical analysis of online social networks in the age of Web 2.0," *Physica A: Statistical Mechanics and its Applications*, 2008, pp. 675-684.

[20] Granovetter, M., "Sociological Theory," 1983, pp. 201-233.

[21] Hölscher, C., and Strube, G., "Web Search Behavior of Internet Experts and Newbies," *Proc. 9th international World Wide Web conference on Computer networks : the international journal of computer and telecommunications netowrking.*, 2000.

[22] Holme, P., Edling, C.R., and Liljeros, F., "Structure and time evolution of an Internet dating community," *Social Networks*, 2004, pp. 155-174.

[23] Huang, Y., and Bian, L., "A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet," *Expert Systems with Applications*, 2009, pp. 933-943.

[24] Jung, J.J., "Social grid platform for collaborative online learning on blogosphere: A case study of eLearning@BlogGrid," *Expert Systems with Applications*, 2009, pp. 2177-2186.

[25] Katarzyna, M., Przemys, aw, K., and Tomasz, K., "Social Recommendations within the Multimedia Sharing Systems," *Proc. of the 1st world summit on The Knowledge Society: Emerging Technologies and Information Systems for the Knowledge Society*, Springer-Verlag, 2008.

[26] Kossinets, G., "Effects of missing data in social networks," *Social Networks*, 2006, pp. 247-268.

[27] Kwai Fun Ip, R., and Wagner, C., "Weblogging: A study of social computing and its impact on organizations," *Decision Support Systems*, 2008, pp. 242-250.

[28] Lee, S.H., Kim, P.J., and Jeong, H., "Statistical properties of sampled networks," *Physical Review E*, 2006.

[29] Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., and Christakis, N., "Tastes, ties, and time: A new social network dataset using Facebook.com," *Social Networks*, 2008, pp. 330-342.

[30] Li, Y.-M., and Chen, C.-W., "A synthetical approach for blog recommendation: Combining trust, social relation, and semantic analysis," *Expert Systems with Applications*, 2009, pp. 6536-6547.

[31] Li, Y., Lu, L., and Xuefeng, L., "A hybrid collaborative filtering method for multiple-interests and multiple-content recommendation in E-Commerce," *Expert Systems with Applications*, 2005, pp. 67-77.

[32] Linden, G., Smith, B., and York, J., "Amazon.com recommendations:item-to-item collaborative filtering," *IEEE Internet Computing.*, 2003.

[33] Liu, D.-R., and Shih, Y.-Y., "Integrating AHP and data mining for product recommendation based on customer lifetime value," *Information & Management*, 2005, pp. 387-400.

[34] Martin, S., Robner, R., and Vach, W., "Neural networks and logistic regression: Part

I,Computational Statistics & Data Analysis," *Computational Statistics & Data Analysis*, 1996, pp. 661-682.

[35] Milgram, S., "The Small World Problem," *Psychology Today*, 1967, pp. 60-67.

[36] Miller, B.N., Albert, I., Lam, S.K., Konstan, J.A., and Riedl, J., "MovieLens unplugged: experiences with an occasionally connected recommender system.," *Proc. the international conference of intelligent user interface*, 2003.

[37] Minas, G., Michael, S., Athina, M., and Xiaowei, Y., "Poking facebook: characterization of osn applications," *Proc. Proceedings of the first workshop on Online social networks*, ACM, 2008.

[38] Mislove, A., Marcon, M., Gummadi, K.P., Druschel, P., and Bhattacharjee, B., "Measurement and analysis of online social networks," *Proc. 7th ACM SIGCOMM conference on Internet measurement*, 2007.

[39] Nazir, A., Raza, S., and Chuah, C.N., "Unveiling Facebook: a measurement study of social network based applications," *Proc. 8th ACM SIGCOMM conference on Internet measurement*, 2008.

[40] Resnick, P., and Zeckhauser, R. eds., Trust among strangers in Internet transactions: empirical analysis of eBay's reputation system., Advances in applied microeconomics: The economics of the Internet and e-commerce., Amsterdam: Elsevier Science, 2002.

[41] Rosenblum, D., "What Anyone Can Know: The Privacy Risks of Social Networking Sites " *IEEE Security and Privacy*, pp. 40-49.

[42] Shepard, L., "Facebook Connect Now Live," 2008; http://developers.facebook.com/news.php?blog=1&story=174.

[43] Shu-Juan, L., Yan, L., Yong, L., Zhi Gang, L., and Jun, T., "Hybrid method of BPN and genetic algorithm for completion time prediction," *Proc. Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, 2005.

[44] T.L, S., "Decision Making with Dependence and Feedback:The analytic Network

Process," *Book Decision Making with Dependence and Feedback:The analytic Network Process*, Series Decision Making with Dependence and Feedback:The analytic Network Process, ed., Editor ed.^eds., PWS Publication, 2001, pp.

[45] "TechWeb," http://www.techweb.com.

[46] "Visual bookshelf," http://www.facebook.com/apps/application.php?id=2481647302.

[47] Wang, J.-C., and Chiu, C.-C., "Recommending trusted online auction sellers using social network analysis," *Expert Systems with Applications*, 2008, pp. 1666-1679.

[48] Weihua, S., Phoha, V.V., and Xin, X., "An adaptive recommendation trust model in multiagent system," *Proc. Intelligent Agent Technology, 2004. (IAT 2004). Proceedings. IEEE/WIC/ACM International Conference on*, 2004.

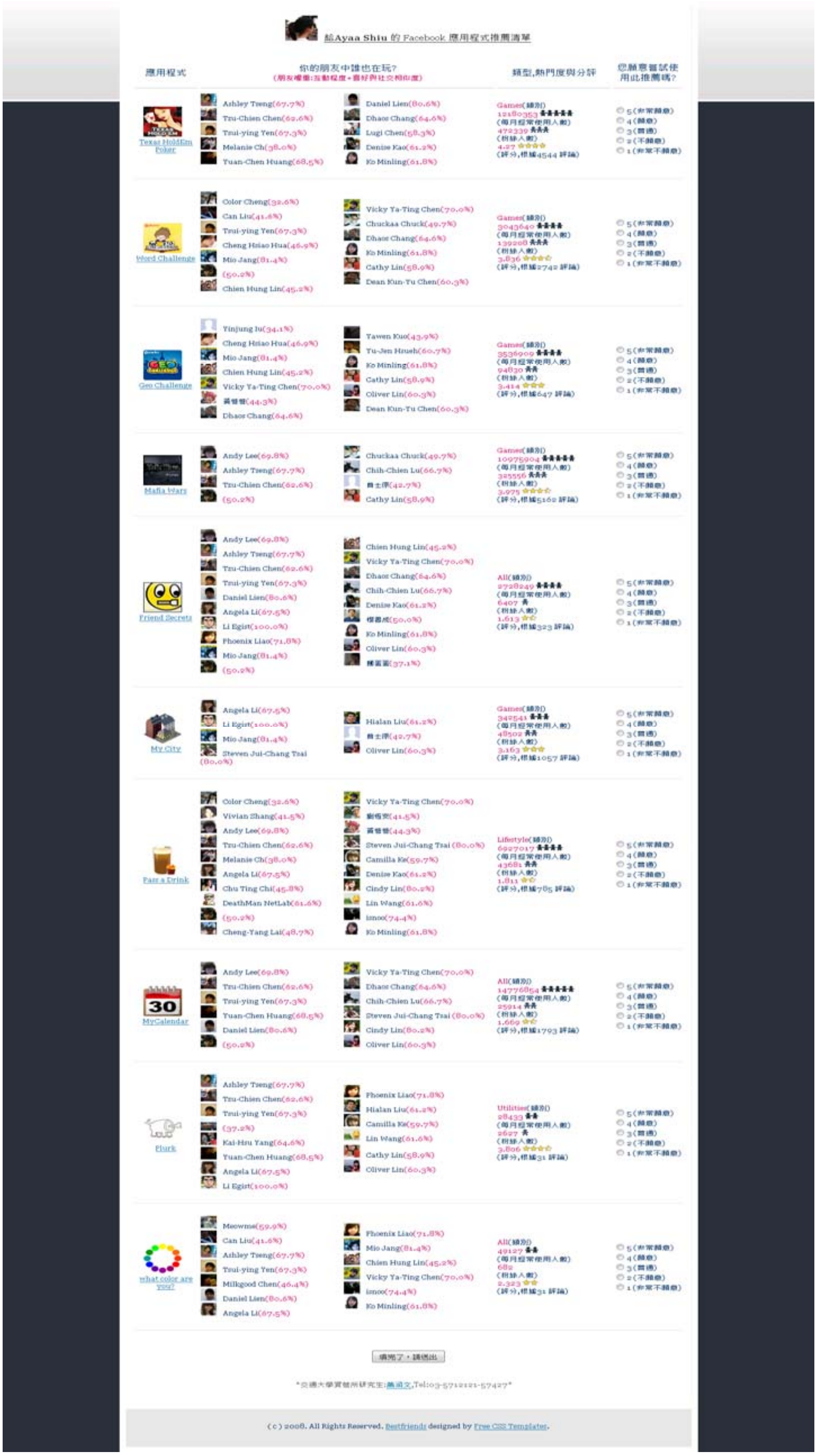
[49] Wu, D., Yang, Z., and Liang, L., "Using DEA-neural network approach to evaluate branch efficiency of a large Canadian bank," *Expert Systems with Applications*, 2006, pp. 108-115.

[50] Yim, H.S., Ahn, H.J., Kim, J.W., and Park, S.J., "Agent-based adaptive travel planning system in peak seasons," *Expert Systems with Applications*, 2004, pp. 211-222.

[51] Yong-Yeol, A., Seungyeop, H., Haewoon, K., Sue, M., and Hawoong, J., "Analysis of topological characteristics of huge online social networking services," *Proc. Proceedings of the 16th international conference on World Wide Web*, ACM, 2007.

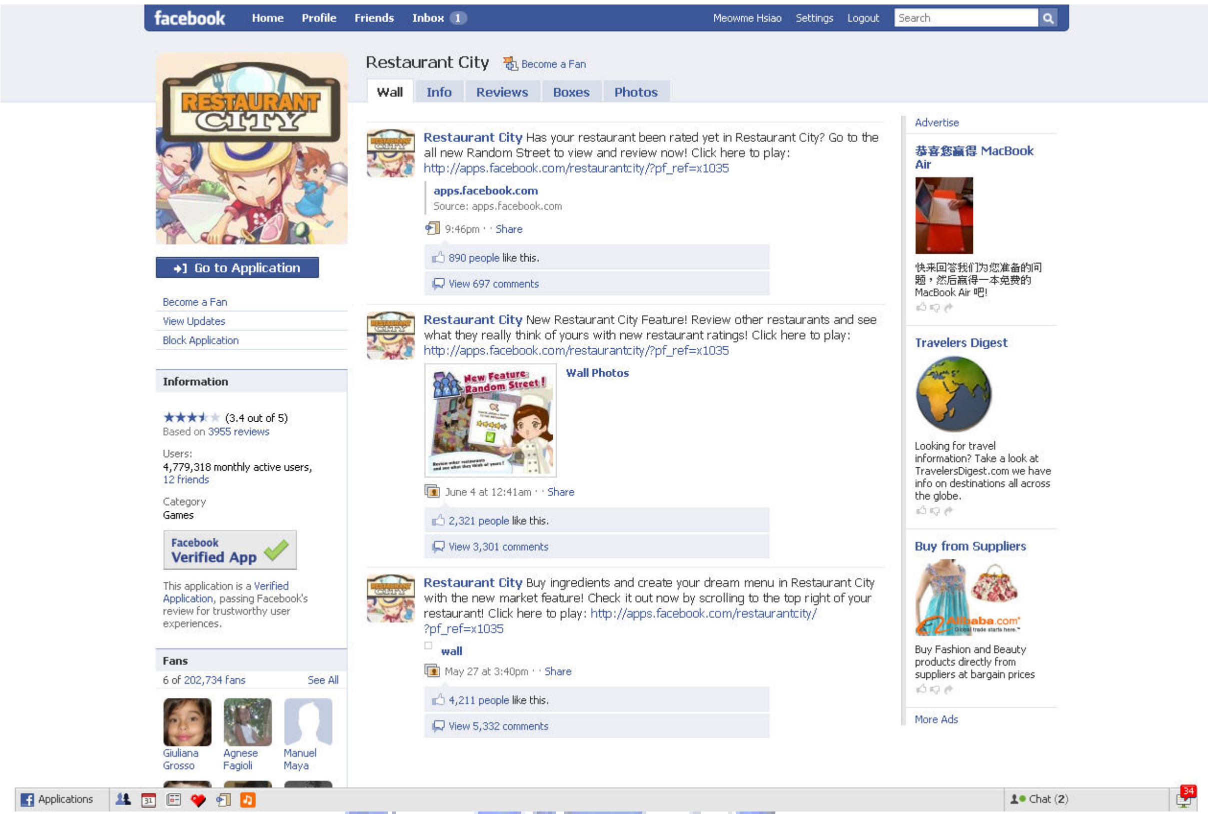
Appendix A

The User Interface of Evaluation From
for Social Network based Application recommendation



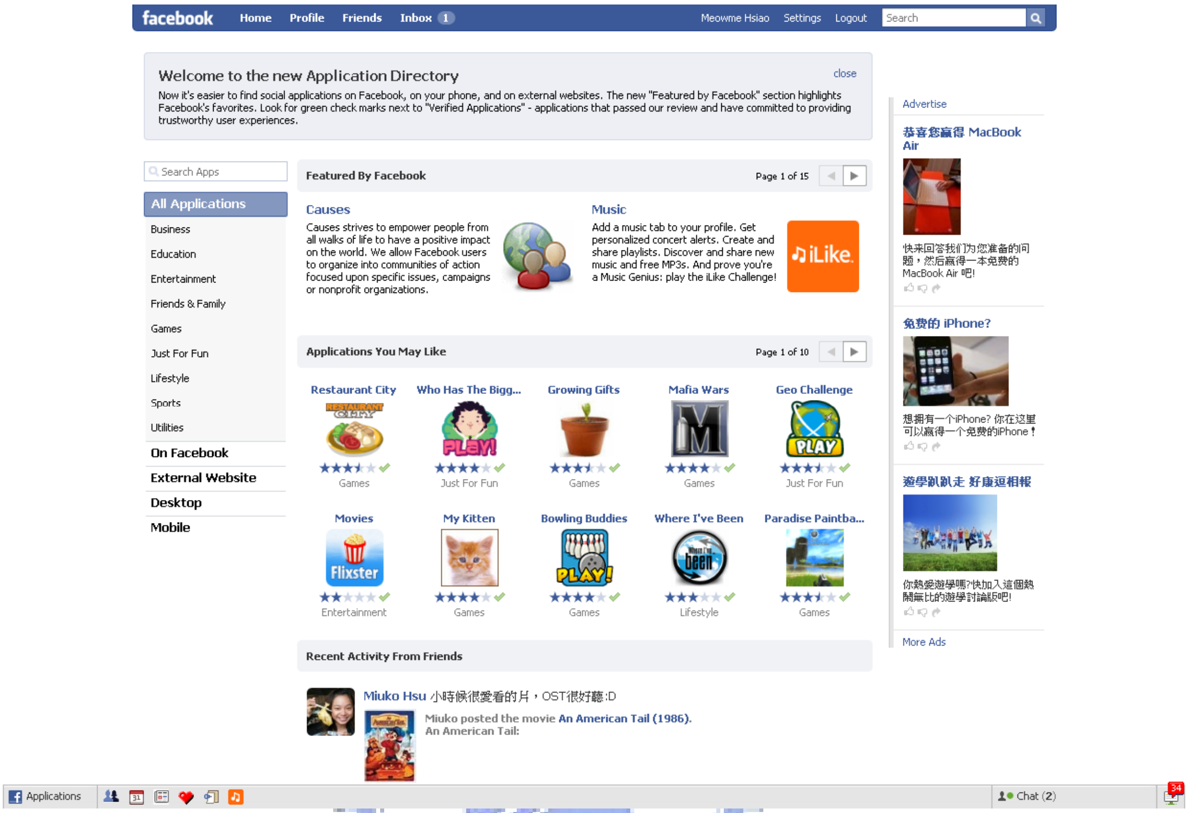
Appendix B

An example of social network based application on Facebook



Appendix C

The Application Directory of Facebook



Appendix D

The web based questionnaires for surveying popularity and reputation weighting