Social Network Analysis of an Online Dating Network

Lin Chen
Computer Science Discipline
Queensland University of Technology
l33.chen@student.qut.edu.au

Richi Nayak
Computer Science Discipline
Queensland University of Technology
r.nayak@qut.edu.au

ABSTRACT
Online social networks can be found everywhere from chatting websites like MSN, blogs such as MySpace to social media such as YouTube and second life. Among them, there is one interesting type of online social networks, online dating network that is growing fast. This paper analyzes an online dating network from social network analysis point of view. Observations are made and results are obtained in order to suggest a better recommendation system for people-to-people networks.

Keywords
Online Dating Network, Social Network Analysis, User profile, User preference

INTRODUCTION
Social networking sites have skyrocketed in popularity with millions of users using social networking websites every day. Most commonly, online social networks serve as a virtual community where users make friends, keep in touch with old friends, share pictures and inform the world of major and minuscule life events. Facebook (www.facebook.com) and Twitter (http://twitter.com) are good examples of online social networking. Online social networks are referred to as any online community where users interact with one another and form relations [15]. A knowledge sharing website such as Wikipedia (www.wikipedia.com) can also be considered as a online social network where users interact through editing articles or it could be message sharing website like MySpace where relations are formed through sharing personal blogs, music, and videos.

Among all of online social networking sites, there is an interesting and demanding but not well studied site; online dating websites. Online dating websites are a type of social networking because users are related through various communication paths; including sending messages, emails or chat. Online dating services offer various advantages; such as enormous amount of choice for a user, no limitation from the physical distance, secure channel, less embarrassment when asking personal questions. Time has witnessed the great increase in demand of the online dating industry. It is reported that the online dating industry has gone from zero about 10 years ago to about Aus $90 million dollars in 2008 and about $100 million dollars in 2010[3]. The US online dating market is expected to grow to US $932 million dollars in 2011 [3].

This paper looks at the structure of an online dating network through Bow-Tie Analysis, Indegree & Outdegree, reachability analysis and attribute breakage. A popular Australian dating service was chosen for this study because of its rich data with a one year snapshot extracted. The results from this study will aid in building a better recommendation system for online dating networks.

C&T'11, 29 June – 2 July 2011, QUT, Brisbane, Australia.
ACM 978-1-4503-0824-3

LITERATURE REVIEW
Social Network Analysis (SNA) is the study of social entities (people in organisation) and their interactions and relationships [21]. The purpose of SNA is to understand the structure, behaviour and composition of social networks and thus improve the social network and social relations contained within. SNA has been applied in many situations from product marketing to search engines and organizational dynamics [6] [19]. SNA has been used to discover how a rumor spreads and what social structures exist among people [21]. By analysing a company's email, SNA helps the manager to find the hidden leader who may not necessarily have a high official position, or much responsibility, but plays an important role in the company [19] [5] [14] [9]. In the same context, studies have analysed "who knows what knowledge", "who think who knows who" and "who believes in what" [17]. The Google search engine provides a classic example of social network analysis. The famous PageRank algorithm is based on the concept that the linked pages have relevant relationships to a certain extent [16].

The characteristics and features of an online social network can be observed when the graph structure of a social network is modeled and analyzed. An in-depth understanding of the graph structure of online social networks is necessary in order to evaluate the current online social networks and to understand the impact online social networks have on the Internet [15]. The application of graph structure theory to online dating network might result in the detecting of the flaws of the existing methods and help propose suitable methods for online dating network; like how the study of the Web led to the discovery of algorithms for finding the sources of authority in the Web. The graph structure theory is able to identify the power law, small-world networks and scale-free networks in online social networks [15] [13] [1] and can help show the users' distribution in online dating network. Graph structure theory is also able to highlight SCC (Strongly Connected Component) and WCC (Weakly Connected Component) which gives an insight on how well the users in a social network are connected and how easy it is for a user to find ideal partner.

The relevant papers on online dating topic mainly come from psychology such as what contributes to successful marriage [12], assessing attractiveness in online dating profile [7] and many others. Another source is recommendation systems. Traditional recommendation algorithms, including user-user algorithms and item-item algorithms are proposed for online dating recommendation by [4]. Kazienko [10] claims that not only users' interests and demographic data need to be considered, but also their activities and some measures of relationships with other users should be considered. So far, however, no research has been done to study the online dating network from the SNA point of view.

AN ONLINE DATING NETWORK

Online dating networks are places where people upload their personal information and preferred partner information and allow them communicate with each other on the Internet in order to facilitate the user developing a personal romantic relationship.

Data required by a dating network for recommending potential partners can be divided into the following features: (1) Personal profile for each user which includes self details on demographic, fixed-choice responses on Physical, Identity, Lifestyle, Career and Education, Politics and Religion and other attributes, free-text responses to various interests such as sport, music etc, and optionally, one or more photographs; (2) Ideal partner profile for each user which includes information about what user prefers in Ideal partner, usually the multiple choices on the attributes discussed before; (3) User activities on the network such as viewing the profiles of other members, sending pre-typed messages to other users; sending emails or chat invitations; and (4) Measure of relationships with other users such as willingness to initialize relationships and responding to invitations, and frequency and intensity with which all relationships are maintained. A relationship can be called successful for the purpose of match making when a user initiates a pre-typed message as a token of interest and the target user sends back a positive message reply. Message is called positive if the user replies in positive manner, and vice versa for a negative reply.

The users' online dating activities can be represented as bipartite graph. Bipartite graph have two distinguishable groups: female and male. The data analysis shows that 97% of members in this network are looking for a straight relationship, therefore, any other types of relationships are not considered in this paper. It is assumed that direct links inside a group do not exist. Figure 1 shows an example of this bipartite graph with a male group and a female group. The links between two groups are messages which are directed. The link shows the existence of a relationship between two members essentially showing the user interest. As shown in Figure 1, a male user u_{m1} sends a message to a female user u_{f4} indicating the male user is interested in the female user. But the female user u_{f4} does not reply. So their relation is not established in this case. A user can choose to send many messages like u_{m3}. A user can choose not to send a message and rather wait for a message being sent to him or her; u_{f2} is such an example.

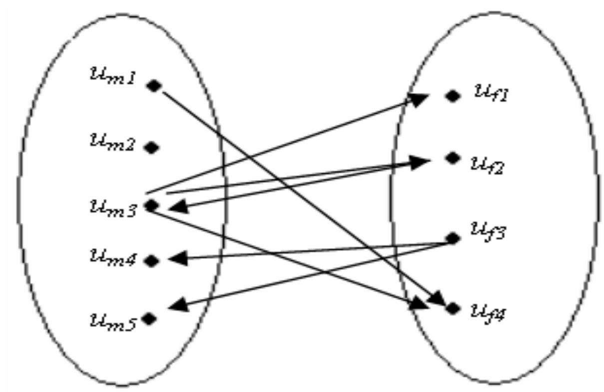

Figure1. Bipartite Graph.

The SOCIAL NETWORK ANALYSIS METHODs

There are three important concepts in social network analysis: *actors*, *relations*, and *ties*. Actors refer to people or organizations [21]. Relations are characterized by content, direction, and strength [8]. A tie connects a pair of actors by one or more relations.

An *actor* in the online dating network refers to the user who logs into the network as a user and view profiles, send/receive messages and emails. Actors can be described by the static profile that they create in the beginning of their registration and update over the period of membership. The static profile is defined by (1) users interest, that includes structured information such as Personality, Smoking Habits, Dietary Habits, Have Pets, Political Persuasion, etc and and free text information such as Headline, About Me, Music, Reading, Movie, Sports, etc, (2) demographic information in structured format, and (3) users matching preferences that includes structured text information such as Have Pets, Relationship Status, Have Children, Hair Color, etc and free text information such as Ideal Partner Description.

The *content* of *relation* in the network is the exchange of information by sending and/or receiving messages and/or emails. The *direction* of *relation* in the network can be bidirection, because, the studied portal allows a user to both send and receive the message/email from another user. The *strength* of *relation* in the network can be identified by the frequency of two users' exchange of the information. Each user can be given a social position based on strength of these relations. Two actors in the network may have multiple ties which are connected by relations of sending and receiving different messages.

Followings are some of the social network analysis methods that we have used in this paper to understand the underlying network.

Bow Tie Structure

The Bow Tie structure, shown in Figure 2, has been successfully used to explain the dynamic behaviour of the web and helps to understand its structure [2]. It has four distinct components: Core (SCC), In, Out and Others (Tendrils and Tubs). This research applies the Bow Tie structure to understand the behavior of users in online dating network. The core is made of the users who frequently participate by sending and receiving messages/emails. A large core indicates the presence of a community where many users interact, directly or indirectly. The "In" is made of users who predominantly receive messages/emails. The "Out" is made of the users who predominately send messages/emails. The "Tendrils" or "Tubes" attaches to either the "In" or the "Out" components or both. Tendrils are those users who only send messages/emails to "In" users. Tubes are those users who receive the messages/emails by "Out" users.

Figure 3 shows the result of conducting Bow Tie structure analysis for the studied dating network. The "Core" part of Bow Tie structure is calculated by using Tarjan's [20] strongly connected components algorithm. The Tarjan's algorithm is a graph theory algorithm for finding the strongly connected components of a graph. Under this theory a vertex or node A is strongly connected to a vertex or node B if there exist two paths, one from A to B and another from B to A. For the purpose of our analysis, the experiments randomly select the starting node. The strongly connected components are only those that are reachable from the start node and thus it is possible that not all strongly connected nodes will be visited. This can be overcome by start the algorithm several times from randomly chosen starting nodes. From the algorithm, "In" and "Out" can be calculated as well. "In" can be calculated by counting the edges which has arrows pointing outward to other users. "Out" can be calculated by counting the edges which has the arrows pointing forward to the user. The "Tendril", "Tubes" and "Disconnected" are considered as "Others" part. The "Others" part can be calculated by deducting the values of the "Core", "In", and "Out" from the total number of edges. The results are compared with the Bow Tie Structure analysis of web [2] and Yahoo! Answers [25] (a social network for question answering). The core of the network is relatively larger than the Web and Yahoo! Answers. This shows that 60% of active network users are involved in all kind of information exchange activities. The majority of users actively participate in the network by sending and receiving messages/emails. Only a few users (5%) just receive messages and 12.5% just send messages. This phenomenon proves that most users use the network with the right intensions of finding compatible matches.

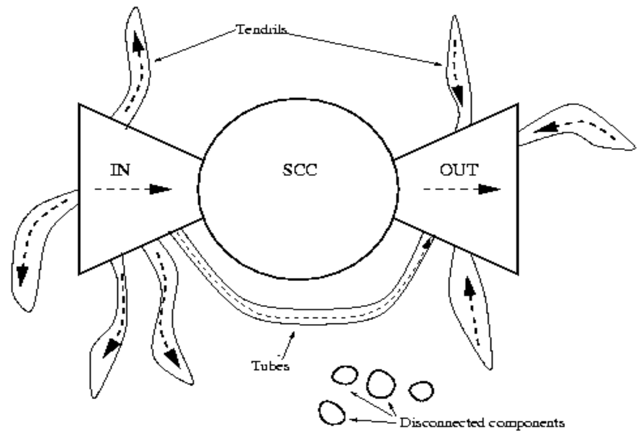

Figure 2. Bow Tie Structure.

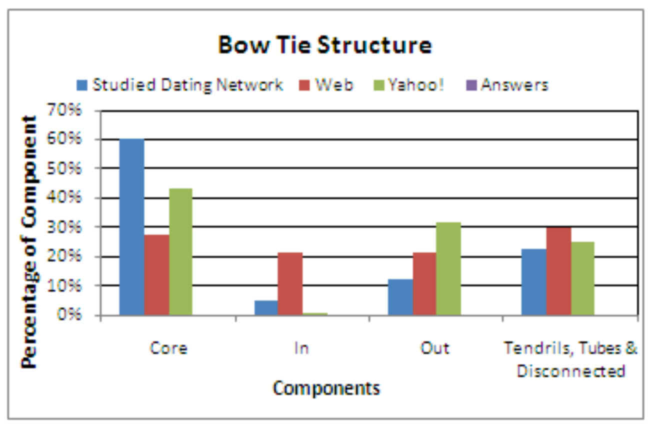

Figure 3. Bow Tie Structure of Networks

The Small World Phenomena

A social network exhibits the small world phenomena if any two individuals in the network are likely to be connected through a short sequence of intermediate acquaintance [11]. For example the Web and YouTube have small world properties [15]. Small world networks can be characterized with short average path length, small diameter and high clustering coefficient [15]. Average path length is simply the average path of all-pairs-shortest paths on social network. Eccentricity is the maximum shortest path distance between a node and any other node. The diameter is the maximum eccentricity across all vertices. Clustering coefficient is a measure of degree to which nodes in a graph tend to cluster together.

To test whether the underlying online dating social network is small world or not, 1000 random users who are active during the selected 6 months period are extracted from the network. Analysis of the network shows that the network diameter is 14 and the average path length is 4.923. The Web, on the other hand, has a diameter of 16.12 and an average path length of 905 [22]. Compared to the Web, the online dating social network has smaller diameter and shorter average path length. However, the clustering coefficient is 0 for these 1000 nodes. The reason can be explained by this social network structure. In online dating social network, 97% of links are between males and females. The number of links existing in the same group is small. So the neighborhood of a male user only has female users and female users are rarely directly linked, similarly, the neighborhood of a female user only have male users.

The top graph in Figure 4 is the graph of 10 out of 1000 random users and their links. The bottom graph is 100 random users and their links. It shows that some users link to many other users and some users only link to a few users.

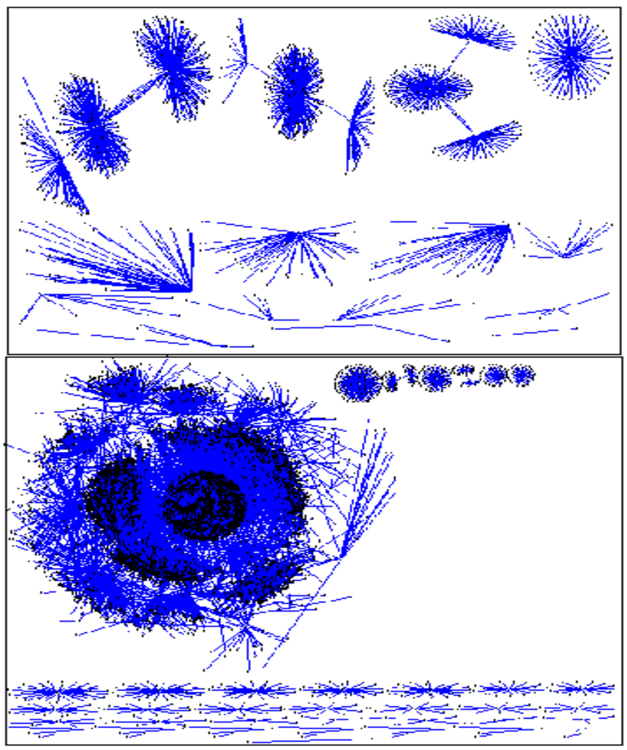

Figure 4. Selected Nodes Visualiation.

Indegree & Outdegree

Degree centrality identifies highly connected actors by measuring their activity and participation in the network. Degree centrality in the network can be used to measure the users' popularity or prestige based on: the number of messages and/or Emails they have sent/received, the stamps they have bought/used, the channels they have initiated and the profiles they have viewed or been viewed.

In the case of a relationship that considers the direction of the link, two indexes are defined: indegree and outdegree. Indegree is the number of links terminating at the node. In the dating network, it refers to the number of emails/messages/channels that a user has received. Outdegree is the number of links originating from the node. It refers to the number of emails/messages/channels that a user has sent. Due to the space constraint, only email degree centrality is displayed.

As the results show (Figure 5) the indegree and outdegree follow the power law which explains the phenomena where large events are rare, but small ones quite common. It can be seen that only a small number of users send/receive a large number of emails. Most of the users send/receive very few emails. For example only about 250 users received more than 1000 emails, whereas, about 400,000 users received less than 20 emails. Only about 400 users send more than 1000 emails, whereas, about 500,000 of users send less than 20 emails.

Indegree and outdegree indicates that only a small number of users can really reach high level of popularity (ie, exchanging a large amount of information), and large number of network users are at low level of popularity (ie, exchanging a small amount of information). Therefore, only small numbers of users are profitable customers for the studied dating network and majority of users are not that profitable in comparison. This observation should be noted while proposing a method that determines the popularity of users in the network. Thus using traditional methods, only a small numbers of users should receive high popularity score and large numbers of users should receive low popularity score.

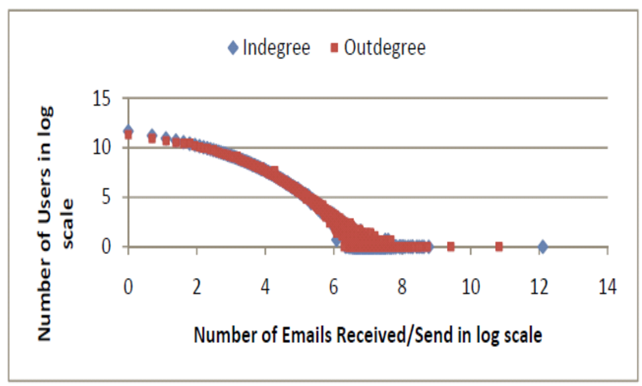

Figure 5. Degree Centrality (Email) in Studied Dating Network.

Reachability

Reachability is defined as the ability of a node to pass to another node in the network. If every node in the network can directly connect to the majority of all the other nodes, then the network is well-connected. Breadth-first-search (BFS) is implemented to test the reachability of the underlying online dating social network. BFS on a directed graph starts with a node *u* in the graph. It then counts the number of nodes reachable from *u* in a series of layers which are disjoint. The first layer has all nodes that are pointed to by links from *u*. A layer *k* consists of links which connected to nodes from layer *k-1* excluding those in any earlier layer. For the analysis purpose, we randomly selected 300 users who have logged in the dating network at any time during the defined six months as the starting nodes. Their communication records are observed for this experiment purpose and the direction of communication is the forward direction which means the layer k users are the initiators of the communications to layer k+1 users.

As a result, lots of nodes die out. These nodes are connected to few other nodes which also have few links or no links to other nodes. A small amount of nodes explode quickly after a few layers. Figure 6 shows the results for 2 to 5 layers. Noticed from Figure 6, 221 starting nodes out of 300 starting nodes are only linked to 1 other node or do not link to any other nodes where the second layer links have limited or no connections to other nodes. The remaining 79 nodes' reachability grows quickly. For example, the node with the maximum reachability in this test can reach 100 nodes at the second layer, 10^4 at third layer, $10^{5.2}$ at fourth layer and $10^{5.5}$ at fifth layer. This experiment shows that around 73% of nodes are linked to a few nodes, and only 26% of nodes are able to connect to lots of nodes (more than 10000). These experiments ascertain that the method, which would need to walk through the graph, needs to control the number of layers in the walk. Otherwise, it would become computationally untraceable to load the whole graph especially when a walk involves millions of nodes in this network

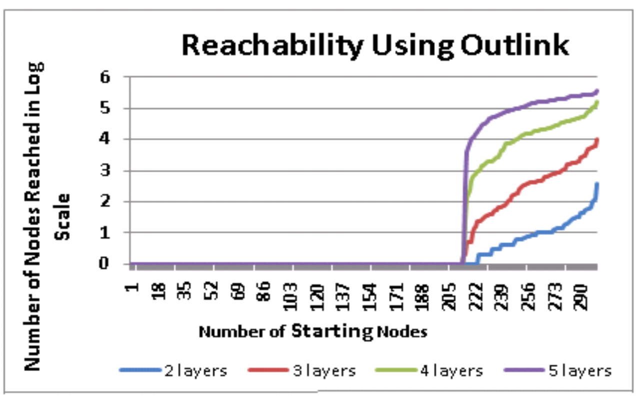

Figure 6. Reachability of the Dating Network.

Attributes Breakage

A number of web studies show that what users do online is actually different from what users say explicitly [23][24] An analysis is carried out to check if the users' online behavior is consistent with the description of the preferences that they write in their user profiles in an online dating network. It is critical to analyse the consistency of a user's actions and say, in order to provide more accurate and better dating services. In the dating networks, the inconsistency in online behavior and explicit information on user profile may be caused by misunderstanding of the users real needs. For example, Emma says she wants a solid, stable man who earns $100,000 plus but keeps clicking on profiles of muscle-bound bad boys. It may also be caused by the change of the taste over time, but the change is not incorporated in the described user profile.

An experiment was conducted by observing users' successful message activity and then comparing it with the preferences given in the user profiles. As discussed in online dating network section, a message is successful only when the message target / receiver sends back a positive message to the message initiator. In this experiment, a comparison is done by looking up the preference attributes of the message initiators and the message receivers profile attributes.

According to the experiment results, in 90% of the instances, either the message initiators preference does not match with the contacted message receivers' profile or the message receiver's preference does not match with the message initiator's profile. This means that the online dating system which adopts information filtering techniques could fail for 90% of the cases when recommending dating partners because of the inconsistency of what a user says and how they act.

More experiments are conducted to look closer at which attribute is most easily broken by users in contacting users online. Results in Table I shows that the message initiator is most likely to break the attribute "Occupation". Overall, 46.40% of times in successful messages, the initiator send message to a receiver whose "Occupation" value in the profile does not comply to the attribute value defined in initiator's preference. There are variations between male and female in terms of the order of attribute breakage. In 36.30% of times, the message target breaks the age attribute as show in Table II. This means that the message target sends back a positive message at 36.30% of times even though the message initiator's age is not in the message target's desired age band. Comparing the message target to message initiator, the difference of the percentage of "Age" breakage comes to around 20%.

Table III and Table IV shows the number of attribute breakage in successful messages. Notice that not every one defines all the attributes in profile or preference. Lots of times, users only define a few attributes such as "Age", "Build" in their ideal partner preferences. In 59.86% of successful messages, the attributes of message target in his/her profile does not contradict the attributes defined by the initiator in his/her preference. Only in 34.66% cases, as show in Table IV, the attributes of message initiator in his/her profile does not contradict the attributes of the message target in his/her ideal partner preference. Seen from Table III and IV, few users consider the candidates who have or would cause 6 or more attributes breakages.

Table I. Attribute Breakage by Initiator in Successful Messages (SM).

Attribute Breakage	Overall	Male	Female
Occupation	46.40%	46.13%	46.19%
Children	29.77%	29.92%	29.41%
Star Sign	25.82%	23.05%	33.86%
Level	25.39%	26.58%	23.81%
Education	24.61%	24.21%	25.27%
Personality	19.47%	19.34%	19.76%
Religion	18.92%	17.18%	22.34%
Eye Color	17.66%	16.03%	24.47%
Have Pets	16.04%	18.46%	11.02%
Age	16.04%	16.06%	24.47%
Politics	15.95%	13.90%	20.44%
Ethnic	14.24%	12.69%	18.36%
Hair Color	13.99%	13.06%	17.47%
Build	12.64%	11.62%	16.72%
Height	11.82%	10.94%	13.70%
Smoke	10.53%	10.90%	9.60%
Marital Status	4.94%	4.15%	6.75%
Diet	4.82%	5.17%	3.84%
Drink	4.38%	4.14%	4.91%
Sex	0	0	0

Table Vshows the attributes usage in profile or preference. The more frequently the attribute used in profile or preference, the more important the attributes should be for users. If that is the case, the most important attribute should be broken the least. According to Table I, it shows the "Occupation" and "Children" are not important attributes in an initiator ideal partner preference, due to them being the attributes broken the most. According to Table V, "Occupation" is the second least important attribute when specifying an ideal partner. However, "Children" is relatively important attribute. In both Table I & V, attribute "Sex" is the most important attribute. Observe Tables II and V, the most interesting attribute is "Age". According to Table II, "Age" is the second most likely attribute to be broken by the target (receiver) of a message. So it should be the second lest important attribute. According to Table V, "Age" is the second most important attribute when describing a user's ideal partner. This analysis shows that the message receiver is more tolerant to the sender even though the sender's age does not satisfy what the receiver described in their ideal partner preference. The message initiator is pickier in terms of "Age" when choosing the target and thus is more likely to send a message to those in the defined age band. This observation further proves the inconsistency of users.

Table II. Attribute Breakage by Target in SM.

Attribute Breakage	Overall	Male	Female
Occupation	50.40%	44.62%	51.27%
Age	36.30%	31.07%	37.69%
Star Sign	36.27%	28.99%	38.06%
Children	33.71%	31.90%	34.03%
Education	32.45%	25.33%	33.59%
Level	26.68%	26.32%	26.73%
Religion	23.94%	19.83%	24.60%
Eye Color	21.92%	18.70%	22.79%
Height	21.46%	15.48%	22.40%
Ethnic	21.09%	17.45%	21.77%
Person	19.56%	18.89%	19.68%
Politics	19.27%	14.87%	20.35%
Hair Color	17.81%	16.69%	18.04%
Build	16.99%	17.56%	16.83%
Smoke	13.36%	10.46%	14.05%
Have Pets	9.07%	17.26%	7.80%
Marital Status	7.70%	5.36%	8.16%
Drink	5.87%	5.81%	5.89%
Diet	4.56%	4.93%	4.46%
Sex	0	0	0

DISCUSSION

This paper utilized a number of social network analysis methods to understand the people behavior in an online dating network. Previous online dating research have been done focusing on existing scam issues, attractive facial features to find similarities between online dating partners for suggesting partners. A leading Australian online dating network was chosen for the study of online dating network because of its rich data.

The Bow Tie structure analysis was found useful in identifying the dynamic structure of the studied network. It was found that majority of users participate in all kinds of network activities including sending and receiving messages. However, there is still 22.5% of users either communicate to those users who only "predominately receive messages" or "predominately send messages" users, or do not have any communication at all. The existence of this portion of users does not bring in any profit to the online dating company. Many online dating companies require users to pay a fee based on their communication usage. This may lead to some users not spending any money on sending messages before they find a candidate worthy of sending messages. A recommendation engine should pay attention to these users while making recommendations of potential partners.

Table III. Number of Attribute Breakage by Initiator in SM.

Number of Attributes Broken	Overall	Male	Female
0	59.86%	62.50%	50.01%
1	25.53%	24.94%	27.75%
2	8.82%	7.87%	12.38%
3	3.36%	2.80%	5.48%
4	1.36%	1.08%	2.42%
5	0.57%	0.44%	1.07%
6	0.25%	0.20%	0.47%
7	0.11%	0.09%	0.20%
8	0.05%	0.04%	0.09%
9	0.028%	0.02%	0.051%
10	0.015%	0.01%	0.031%
11	0.008%	0.006%	0.019%
12	0.0046%	0.0028%	0.011%
13	0.0023%	0.0011%	0.007%
14	0.001%	3.985E-4%	0.003%
15	4.25E-4%	1.549E-4%	0.001%
16	1.16E-4%	8.11E-5%	2.48057E-4%
17	2.33E-5%	1.47E-5%	5.51239E-5%
18	0.0	0.0	0.0
19	0.0	0.0	0.0
20	0.0	0.0	0.0
21	0.0	0.0	0.0

Indegree & Outdegree illuminate the wide existing phenomenon that large events are rare and small ones are common. Due to time, effort, cost factors, the majority of users in online dating network selectively send messages to their interested candidates. Only small portion of users broadcast their messages to many users. As such, recommendation system should take both users point of view and making profit into consideration. Many recommendations for user can potentially bring more profit for online dating company. However, too many recommendations may cause information overloading for users. Balancing the number of recommendations which both users and online dating company are content with is necessary.

Reachability analysis gives us an insight of the network in terms of the computational complexity of the algorithm that should be deployed in the network to make a recommendation to users. This analysis shows that only a small number of nodes can reach many other nodes and it would be expensive to reach to all nodes if an algorithm such as random walk is employed.

To find out whether users "lie" in their profile, this paper compared the user defined ideal partner preferences with the profile of the contacted love interest. In 90% of cases, users' defined ideal partner preferences do not match with the profiles of users' contacted partners. In the majority of cases, there is one attribute breakage. From the result, it seems that users are tolerant with the inconsistency. Comparing the message initiators and message receivers, the receivers are more willingly to accept someone does not match with their ideal partner preference. Even some important attributes such as "Age", the receivers are more willing to break than the initiators. Finally, the finding of the attribute breakage means that users' personal profile and their contacted users' personal profiles should be taken into the consideration in combination during recommendation. The users' ideal partner preferences alone are not a good indicator when making a recommendation.

Table IV. Number of Attribute Breakage by Target in SM.

Number of Attributes Broken	Overall	Male	Female
0	34.66%	45.08%	31.88%
1	33.19%	32.31%	33.42%
2	18.08%	13.99%	19.17%
3	8.29%	5.37%	9.07%
4	3.47%	2.00%	3.86%
5	1.39%	0.75%	1.57%
6	0.55%	0.29%	0.62%
7	0.22%	0.12%	0.24%
8	0.09%	0.05%	0.09%
9	0.04%	0.02%	0.04%
10	0.02%	0.01%	0.02%
11	0.007%	0.006%	0.007%
12	0.003%	0.002%	0.003%
13	0.001%	9.389E-4%	0.001%
14	4.598E-4%	3.313E-4%	4.942E-4%
15	1.804E-4%	3.038E-4%	1.475E-4%
16	4.657E-5%	1.105E-4%	2.950E-5%
17	3.492E-5%	0.0	4.425E-5%
18	0.0	0.0	0.0
19	0.0	0.0	0.0
20	5.821E-6%	0.0	7.375E-6%
21	0.0	0.0	0.0

Table V. Attribute Usage in Personal User Profile.

Attribute	Usage in Profile	Usage in Preference
Gender	100.00%	Not available
Age	99.88%	93.81%
Marital Status	98.92%	30.91%
Height	95.21%	46.74%
Children	92.54%	29.24%
Build	91.37%	53.37%
Star Sign	88.90%	1.75%
Hair Color	88.87%	8.90%
Eye Color	87.66%	5.33%
Drink	86.87%	38.98%
Smoke	85.98%	43.77%
Personality	84.13%	36.76%
Have Pets	78.06%	14.69%
Education	77.87%	14.46%
Occupation	77.10%	3.99%
Level	71.67%	6.28%
Diet	69.24%	10.03%
Ethnic	48.55%	11.25%
Politics	43.92%	7.56%
Religion	41.89%	8.22%

ACKNOWLEDGEMENT

The content presented in this paper is part of an ongoing cooperative study between Queensland University of Technology (QUT) and an Industry Partner, with sponsorship from the Cooperative Research Centre for Smart Services (CRC-SS). The authors wish to acknowledge CRC-SS for partly funding this work and the Industry Partner for providing data for analysis. The views presented in this paper are of the authors and not necessarily the views of the organizations.

REFERENCES

1. Adamic, L. A., Buyukkokten, O., & Adar, E. (2003). A social network caught in the Web. *First Monday, 8*(6).
2. Borodin, A., & al., e. (2003). Graph structure in the web. Retrieved 18th June, 2008, from http://www.cis.upenn.edu/~mkearns/teaching/NetworkedLife/broder.pdf
3. Bowden,T., Ed., (2010). *Online dating revolution* (7:30 Report. Australian Broadcasting Corporation, 2010.
4. Brozovsky, L., & Petricek, V. (2007). Recommender System for Online Dating Service. Retrieved on December 10th from http://www.occamslab.com/petricek/papers/dating/brozovsky07recommender.pdf
5. Cai, D., et al. (2005). Mining Hidden Community in Heterogeneous Social Networks. Proceedings of LinkKDD'05, 58-65.
6. Domingos, P., & Richardson, M. (2001). Mining Network Value of Customer. Paper presented at KDD'01.
7. Fiore*,A.T., et al.*, (2008). "Assessing Attractiveness in Online Dating Profiles," in *26th Annual SIGCHI Conference on Human Factors in Computing Systems (CHI'08)*, Florence, Italy, 2008, pp. 797 - 806.
8. Garton, L. (1997). Studying Online Social Networks. Retrieved September 14th, 2007, from http://jcmc.indiana.edu/vol3/issue1/garton.html
9. Kautz, H., et al. (1997). Combining Social Networks and Collaborative Filtering. Communication of the ACM, 40 (3), 63-65.
10. Kazienko, P., & Musial, K. (2006). Recommendation Framework for Online Social Networks. Paper presented at the 4th Atlantic Web Intelligence Conference (AWIC'06).
11. Kleihberg. (1999). "The small world phenomenon: An algorithm perspective," Cornell 99-1776, 1999.
12. Kreider, R., & Fields, J. (2002). Number, timing and duration of marriages and divorces: 1996. Washington DC. Census Bureau.
13. Kumar, R., Raghavan, P., Rajagopalan, S., & Tomkins, A. (1999). Trawling the Web for Emerging Cyber-Communities. *Computer Networks, 31*, 1481-1493.
14. Matsuo, Y. et al. (2006). POLYPHONET: An Advanced Social Network Extraction System from the Web. Proceedings of WWW'06, 397-406.
15. Misolve*,A., et al.*, (2007). "Measurement and Analysis of Online Social Networks," presented at the Internet Measurement Conference, 2007.
16. Page, L., Brin, S. (1998). The PageRank citation ranking: brining order to the web. Tech. rep. Standford Digital Library Technologies Project.
17. Pathak, N., Mane, S., & Srivastava, J. (2007). Analysis of Cognitive Social and Knowledge Networks from Electronic Communication. International Jouranl of Semantic Computing. 1 (1), 87-118.
18. Preiss, B. R., (1998). Data structures and algorithms with object oriented design patterns in Java: John Wiley & Sons, 1998.
19. Song, X. D., et al. (2005). Modeling and Predicting Personal Information Dissemination Behavior. Proceedings of KDD'05, 479-488.
20. Tarjan, R. (1972). Depth-first search and linear graph algorithms. *SIAM Journal on Computing, 1*(2), 146-160.
21. Wasserman, S., & Faust, K. (1994). Social Network Analysis Method and Applications. Cambridge University Press.

22. Wilson, *C., et al.*, (2009). "User interaction in social networks and implications," presented at the EuroSys'09, Germany, 2009.

23. Online Dating: a sign of desperation? (2010) Retrieved 2nd March from http://www.mmegi.bw/index.php?sid=1&aid=932&dir=2011/February/Monday14

24. Gates, V. (2007). " Internet Dating 2.0" Retrieved 2nd March from http://www.time.com/time/business/article/0,8599,1580609,00.html

25. Chen, L., Nayak, R. "Expertise analysis in a question answer portal for author ranking" in International Conference on Web Intelligence pp.134-140, 2008.