



2014 Mathematical Contest in Modeling (MCM/ICM) Control Sheet

Please review this page before submitting your solution to ensure that all of the information is correct. Do not make changes by hand to the information on this control sheet. If you need to change any of the information on this sheet, login via the **Advisor Login** link on the MCM web site at <http://www.comap.com/undergraduate/contests/mcm>, make the changes online, and print a new sheet.

You may NOT photocopy this control sheet to give to a new team, nor may you assign any team a control number. Each team must have its own control number, obtained by registering via the MCM web site at <http://www.comap.com/undergraduate/contests/mcm>.

Advisor
Name: Yanhui Guan

Department: Department of Mathematics

Institution: Sun Yat-Sen University

Address: No. 135, Xingang Xi Road
School of Mathematics &
Computational Science, Sun Yat-sen University
Guangzhou, Guangdong 510275

Phone: 86-20-84113190

Fax: 86-20-84037978

Email: g_yanhui@163.com

Home
Phone: 86-20-84113190

**Your team's control number
is:**

24367

(Place this control number on
all pages of your solution paper
and on any support material.)

Problem Chosen:

C

The names of the team members will appear on your team's certificate **exactly** as they appear on this page, including all capitalization and punctuation, if any. Gender data is optional and will be used for statistical purposes only; it will not appear on the certificate.

Team Member

Gender

Hang Chen

M

Zemin Wang

M

Qing Pan

F

Each team member must sign the statement below:

(Failure to obtain signatures from each team member will result in disqualification of the entire team.)

Each of us hereby testifies that our team abided by all of the contest's rules and did not consult with anyone who was not on this team in developing the enclosed solution paper.

Mailing Address and Signature of Hang Chen

Mailing Address and Signature of Zemin Wang

Mailing Address and Signature of Qing Pan

This signed original must be stapled to the top of your team's solution paper.

	Team Control Number	
For office use only	24367	For office use only
T1 _____		F1 _____
T2 _____		F2 _____
T3 _____	Problem Chosen	F3 _____
T4 _____	C	F4 _____

P-Rank Network Model: A New Method To Measure Influence and Impact

Summary

A particularly interesting and challenging problem in network analysis is to measure influence and impact. In this article, we develop a P-Rank Network Model to figure out the influence of each node within research networks and other areas of society. Our model is based on PageRank algorithm and double-weighting factor p . To further improve the model, we also add Max-Degree algorithm to our model in order to discuss how to maximize the network-based influence when applying our model into practice within networks.

After carefully analyzing the problem, we divide our paper into five parts:
To complete the first task, after effective data extraction, we use a set of nodes and their links to build the co-author network of the Erdos1 authors with basic assumptions and then conduct the properties analysis.

As for the second and third tasks, we use the P-Rank Network Model (PRNM) to measure the influence and impact within research networks. With corresponding network influence data, we figure out the most influential Erdos1 author and determine the most influential paper in network science. In addition, we further discuss the model applicability to generic network influence-ranking problems.

To deal with the fourth question, we implement our model on network influence of movie actors. To achieve more reliable results, we restrict the network to only ten famous movie actors and then use the algorithm in PRNM to work out the most influential movie actor within this network. The effect of different value of p on our results is also discussed in this part.

Afterwards, we apply our model into practice within networks. In this case, we focus on how to maximize the network-based influence based on max-degree algorithm. To explain the application, we present two specific situations to maximize impact. We also discuss in detail on how to optimize our model by adding the new algorithm.

Finally, the sensitivity, strengths and weakness of the model are discussed, and the future work is pointed out.

Contents

1	Introduction	4
1.1	Background	4
1.2	Breaking Down the Problem	4
1.3	Our Work	4
2	Co-author Network of the Erdos1 Authors	5
2.1	Basic Assumptions	5
2.2	Definition	5
2.3	Establishment	5
2.4	Properties Analysis	6
3	P-Rank Network Model	6
3.1	Model Description	6
3.2	Modeling Methodology	7
3.3	Influence Measures and Algorithm	7
3.4	Figuring Out the Most Influential Author	8
3.5	Most Influential Paper in Network Science	9
3.6	Applicability to Generic Network Influence-Ranking Problems	10
4	Implementing the Model on Network Influence of Movie Actors	11
4.1	Introduction	11
4.2	Network of Ten Movie Actors	11
4.3	Influence Measures of Movie Actors	12
4.4	Simulation and Analysis	12
4.4.1	First results	12
4.4.2	Effect of different p value	13
4.4.3	Simulation: The promotion of the model	14
5	Application within Networks	14
5.1	Our Goal	14
5.2	Maximizing Influence	14
5.3	Model Optimizing	16
6	Conclusions	16
6.1	Sensitivity and Uncertainty Analysis	16
6.2	Strengths	16
6.3	Weaknesses	17
6.4	Future Work	17
	References	17

1 Introduction

1.1 Background

The study of networks has emerged in diverse disciplines as a means of analyzing complex relational data. The earliest known paper in this field is the famous Seven Bridges of Königsberg written by Leonhard Euler in 1736. Euler's mathematical description of vertices and edges was the foundation of graph theory, a branch of mathematics that studies the properties of pairwise relations in a network structure. The field of graph theory continued to develop and found applications in chemistry.[1]

With the development of network study, network science has been a hot research field in recent years. Network science is an interdisciplinary academic field which studies complex networks such as telecommunication networks, computer networks, biological networks, cognitive and semantic networks, and social networks. The field draws on theories and methods including graph theory from mathematics, statistical mechanics from physics, data mining and information visualization from computer science, inferential modeling from statistics, and social structure from sociology.

In this problem, we focus on using networks to measure influence and impact. In research networks, one of the techniques to determine influence of academic research is to build and measure properties of citation or co-author networks. Citation between papers and co-authoring between researchers are the most common relationship in research networks. Identifying influential papers or researchers in academic networks is very meaningful for people to quickly understand related academic fields. Our goal for ICM 2014 is to analyze influence and impact in research networks and other areas of society.

1.2 Breaking Down the Problem

After analyzing the problem, we conclude five main sub-problems to tackle in our paper:

- Build the co-author network of the Erdos1 authors and analyze the properties of this network.
- Develop influence measures to determine the most influential Erdos1 author
- Compare the relative significance of a research paper by analyzing related works
- Implement the algorithm on a different set of network influence data
- Discuss the science, understanding and utility of modeling influence and impact within networks

1.3 Our Work

To tackle the first problem, we use a set of parameters to build the co-author network of the Erdos1 authors with basic assumptions and then conduct the properties analysis. As for the second and third tasks, we want to use the Weighted Page Rank Network Influence Model (WPRNLM) to measure the influence and impact within research networks. With corresponding network influence data, we will figure out the most influential Erdos1 author and determine the most influential paper in network science. In addition, we will further discuss the model applicability to generic

network influence-ranking problems. To deal with the fourth question, we plan to implement our model on network influence of movie actors. To achieve better and more reliable results, we restrict the network to only ten famous Hong Kong and Taiwan movie actors. Then we use the algorithm in WPRNLM to work out the most influential movie actor within this network. Finally, we apply our model into practice within networks. In this case, we focus on how to maximize the network-based influence based on max-degree algorithm. To explain the application, we present two specific examples in life. We also discuss in detail on how to optimize our model by adding the new algorithm.

2 Co-author Network of the Erdos1 Authors

2.1 Basic Assumptions

- Only taking the total influence of the important articles into consideration when determining authors original influence. We use H-index to measure the original influence of one author. H-index connotes the quantity of important papers in academic research as a scientist has index h if h of his/her N_p papers have at least h citations each, and the other $(N_p - h)$ papers have no more than h citations each.
- To define the Erdos number, Erdos alone was assigned the Erdos number of 0 (for being himself), while his immediate collaborators could claim an Erdos number of 1, their collaborators have Erdos number at most 2, and so on.
- When calculating the attached influence, we only consider the relationship between Erdos1 authors in the method of weighting and ignore the influence of other Erdos2 authors.
- Use only cooperation times to describe the line between two authors regardless of the intensity and effect of each cooperation.
- Neglect the impact of interference factor outside the network.

2.2 Definition

- **Node:** Person who has co-author with Erdos1, every person is a rectangle with a number aside
- **Link:** Cooperation between Erdos1 authors

2.3 Establishment

We get data from website supported, the list of 511 coauthors of Paul Erdos, together with their coauthors listed beneath them, we only take relationship between the Erdos1 authors.

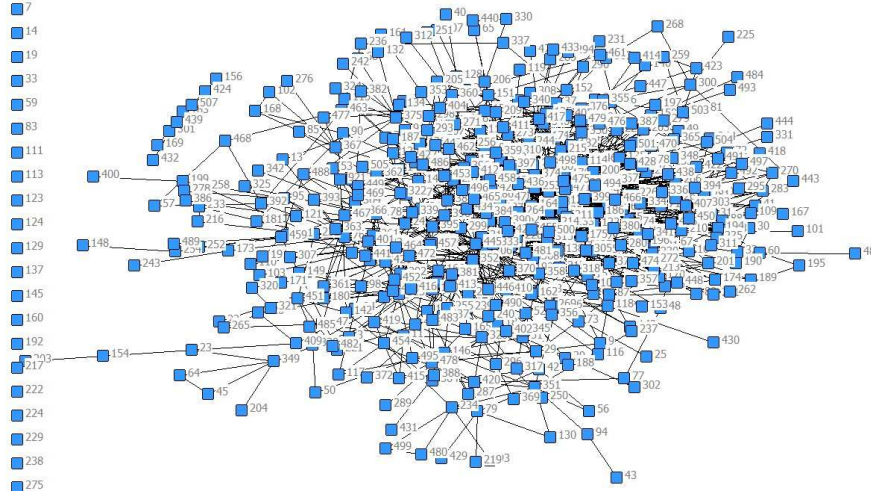


Figure 2.3.1: Co-author network

2.4 Properties Analysis

- **Size:** this network consists of 511 nodes and 1640 undirected edges .
- **Density:** density D of a network is defined as:

$$D = \frac{2E}{N(N-1)} \quad (2.4.1)$$

this network's density is 0.01258

- **Betweenness Centrality:** a measure of a node's centrality in a network equal to the number of shortest paths from all vertices to all others that pass through that node (Freeman et, al., 1979).

The betweenness centrality of node is defined as following:

$$C_B(i) = \sum_{s \neq i \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}} \quad (2.4.2)$$

Where σ_{st} is the total number of shortest paths from node s to node t and $\sigma_{st}(i)$ is the number of those paths that pass through node i .

- **Edge Betweenness:** the number of shortest paths between pairs of nodes that run along it (Girvan M. et, al., 2002). The edges that connect communities in the network have high edge betweenness.

3 P-Rank Network Model

3.1 Model Description

The inspiration about how to establish the model to figure out the influence of the network comes from Page Rank algorithm. We introduce double-weighting factor based on the Page Rank algorithm to make it more suitable to figure out the influence of the network.

In our model, every person's influence is determined by three part

- The sum of the votes on its connections with other important reseracher in Erdos1 network
- The number of person own publications
- Times he co-autherd with Erdos

Then we do some **Symbol definition** as follows:

Symbol	Meaning
S	the composite influence of a node
O	the origin influence of a node
A	attached influence of other nodes
p	double-weighting factor
S_i	S of the ith author
O_i	O of the ith author
A_i	A of the ith author
H_i	the h-index of the ith author
H	the sum of the h-index of 511 authors
m	the number of the author who cooperate with the ith author
S_{ij}	the S of author j in m authors cooperated with i
CT_{ij}	cooperation time between the author j and i

3.2 Modeling Methodology

- Each person's vote is proportional to the importance of its source co-author
- If person P with importance X has n co-authors, auther j have co-autherd with P CT_j times, auther j will get influence $\frac{XCT_j}{\sum_{0 < i < n} CT_i}$ from P
- If person P cooperate with Erdos, CT_p times, he will get influence $\frac{CT_p}{\sum_{0 < i < n} CT_i}$

3.3 Influence Measures and Algorithm

Matrix M has one row and one column for each author Suppose person j has n partners [3]

$$M_{ij} = \begin{cases} \frac{CT_{ji}}{\sum_{0 < k < n} CT_{jk}}, & \text{if } j \rightarrow i; \\ 0, & \text{others} \end{cases} \quad (3.3.1)$$

M is a **column stochastic matrix**, which columns sum to 1

s, o both are vectors with one entry per person

s_i is the influence score of person i, call it the rank vector, $|s| = 1$

o_i is the own influence of person i, $o_i = \frac{h_i}{H}$, $|o| = 1$

we use **Power Iteration method**:

Suppose there are N person

Initialize: $s_0 = [1/N, \dots, 1/N]^T$

Iterate: $s^{k+1} = Bs^k$

Stop : when $|s^{k+1} - s^k|_1 < \epsilon$, $|x|_1$ is the L_1 norm

$$\begin{aligned}
B_{ij} &= (1 - p)M_{ij} + po_i \\
s_i &= \sum_{1 \leq j \leq N} B_{ij}s_j \\
s_i &= \sum_{1 \leq j \leq N} [(1 - p)M_{ij} + po_i]s_j \\
&= (1 - p) \sum_{1 \leq j \leq N} M_{ij}s_j + po_i \sum_{1 \leq j \leq N} s_j \\
&= (1 - p) \sum_{1 \leq j \leq N} M_{ij}r_j + po_i, \text{ since } |s| = 1 \\
s &= (1 - p)Ms + po
\end{aligned}$$

3.4 Figuring Out the Most Influential Author

To figure out the most influential Erdos1 author in our network, we assume each node represents an Erdos1 author. The sum influence of each node equals to original influence (O_I) plus attached influence (A_I). As long as two authors have cooperated for at least one time, we think there is a link between them. [2] For author i in our network, we can calculate his sum influence using the following formula:

$$S_i = p * O_i + (1 - p) * A_i \quad (3.4.1)$$

$$O_i = \frac{h_i}{H} \quad (3.4.2)$$

$$A_i = \sum_{j=1}^m S_{ij} \frac{CT_{ij}}{\sum_{j=1}^m CT_{ij}} \quad (3.4.3)$$

We use PRNM model to figure out the rank of 511 authors when p changes from 0.1 to 0.9, and find out authors who always appear in top10 when p changes. And then we draw the figure about them as figure 3.4.1.

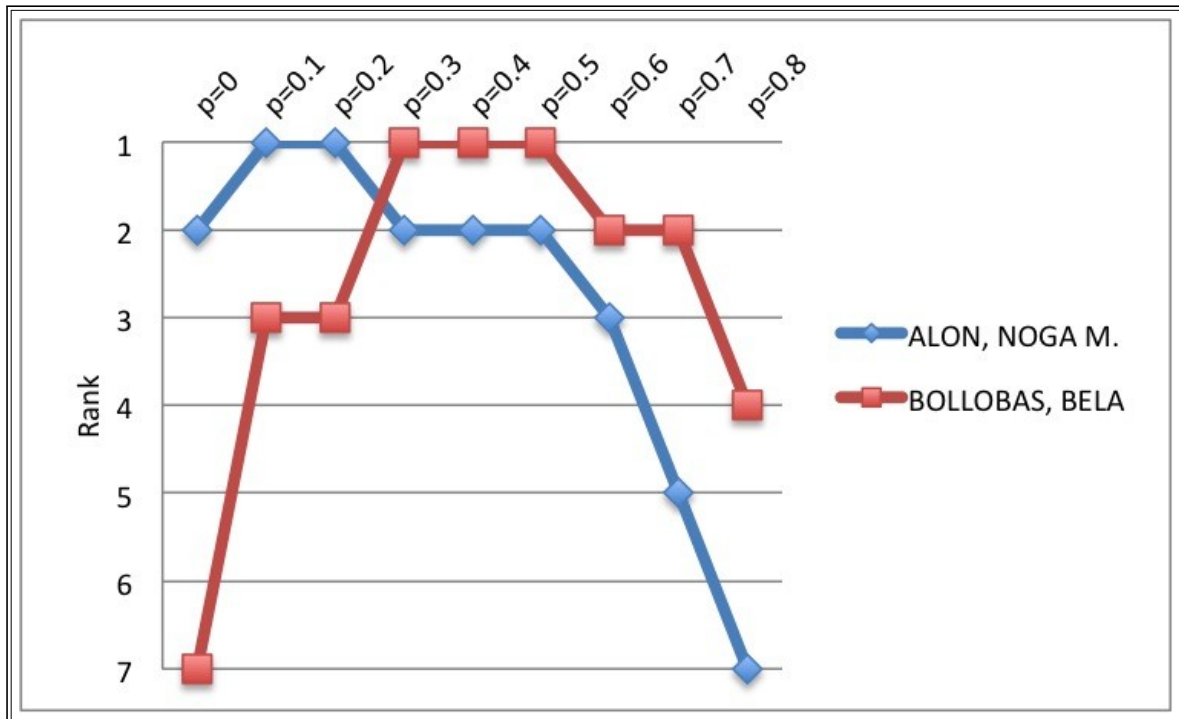


Figure 3.4.1: Influence ranking for 2 most influential Edors1 authors

During the process of p changes, ALON , NOGA M and BOLLOBAS, BELA are always appear in top10 ,so we can conclude that ALON , NOGA M and BOLLOBAS, BELA are the most influential authors.

3.5 Most Influential Paper in Network Science

In order to understand the problem better, we draw the network of 16 papers' citation relationship as figure(3.5.1):

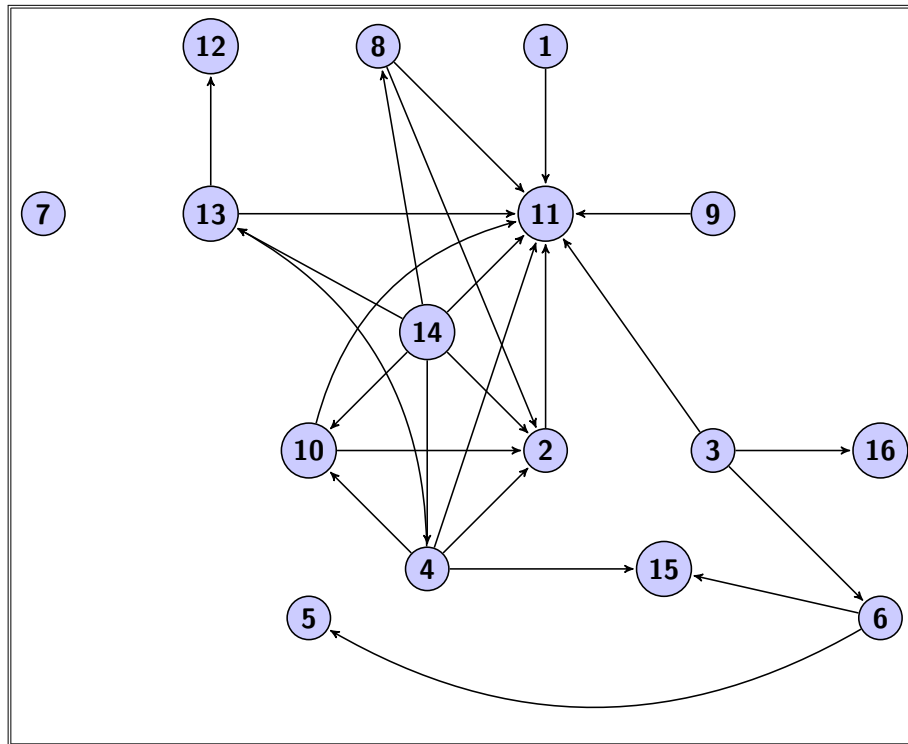


Figure 3.5.1: Topological order of paper network

To find out the most influential paper among the 16 given papers, we make some assumptions:

- Assuming that each paper has an initial relative influence number 1
- Assuming that paper A is cited by paper B, then we call B has one citation number of 1 from A, and the influence factor of the relative influence in network from B to A is $15/16$; and if B is cited by C, then C has a citation number of 2 from A, and the influence factor of the relative influence in network from C to A is $14/16$; and so on. The distribution of the influence factor is as follows:

$$\text{factor} = -\frac{\text{Citation number}}{16} + 1$$

- Calculating the minimum citation numbers influence caused by other papers when calculate the papers relative influence in network. For instance: paper A is cited by paper B and B both has citation number 1 and citation number 2, and when calculate the influence of A, only the citation number 1 of B is relative, which is the minimum citation number.

$$\text{Papers comparative influence} = 1 + \sum_{1 \leq i \leq m} n_i * f_i \quad (3.5.1)$$

where n_i is the independent paper number which is cited i times by this paper, f_i is the influence factor of citation number i , m is the maximum citation number. According to this method, the relative influence of these 16 papers is as figure (3.5.2):

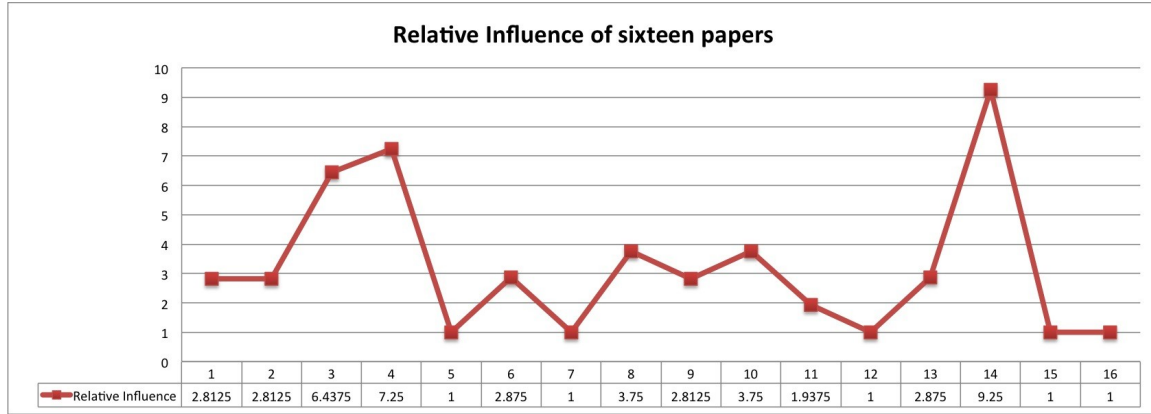


Figure 3.5.2

In this figure, the 14th paper has maximum relative influence among these 16 papers[4]

3.6 Applicability to Generic Network Influence-Ranking Problems

We make some changes in the measuring network influence model that we established to let it meet the demand of obtaining the relative influence of each paper in the paper network model. Similarly, we change some parameters and apply the model to other networks in the real world. For example: when calculate a scholars influence in a network, we use the equation the overall influence = original influence (1-p) + attached influence to get each scholars influence by iteration (the original influence is represented by each scholars research achievements, and the attached influence is obtained by the weight distribution of the overall influence of other scholars that have a collaboration relationship to this scholar). [5] As for obtaining the influence of a school, department or a journal in the network, it needs the parameters and the relationships mathematic description to be fixed, and we can use the model we established too. For instance, when we set up a network in which the school is regarded as the node to obtain the research influence of each school, and then the schools original influence is its own research achievements, and the amount of co-achievements with other schools is represented as the weighted index from other schools. And next, we use the weighted page rank algorithm to get each schools influence in the network. Hence, we only need to collect each schools independent research achievements and use some corresponding data to represent them, and then gather the amount of co-achievements with other schools, and finally the work of studying the network relationship figure among schools is finished.

4 Implementing the Model on Network Influence of Movie Actors

4.1 Introduction

To test the validity of our model and algorithm derived from academic research networks, we apply PRNM to a totally different movie actor collaboration network. To simplify the model and have an accurate result, we limit the scope of our movie actor collaboration network within ten famous Taiwan and Hong Kong actors, Shawn Yue, Simon Yam, Kot Eric, Eric Anthony, Perry Francis, NG Eric, Tsang Tommy, Tam Andy, Lau Daniel, Wu and Bowie Tsang. They have good fame among the movie collaboration network for they have tight relationship between each other and they are known well by a lot of movie fans. We find relevant actor collaboration data from a famous online film community on the Internet, Mtime, a website co-constructed by technical experts from overseas and some domestic senior media professionals. Compared to the famous IMDB website which provides the search service of the film database, Mtime has a comparatively complete collection of Chinese movies so it is more suitable for the Chinese film network research.[6]

4.2 Network of Ten Movie Actors

In the process of constructing network, a node represents an actor, and if two of these actors appeared in one film, we use one undirected weighted edge to connect the two corresponding nodes. The following figure is the network of these ten movie actors fig(4.2.1):

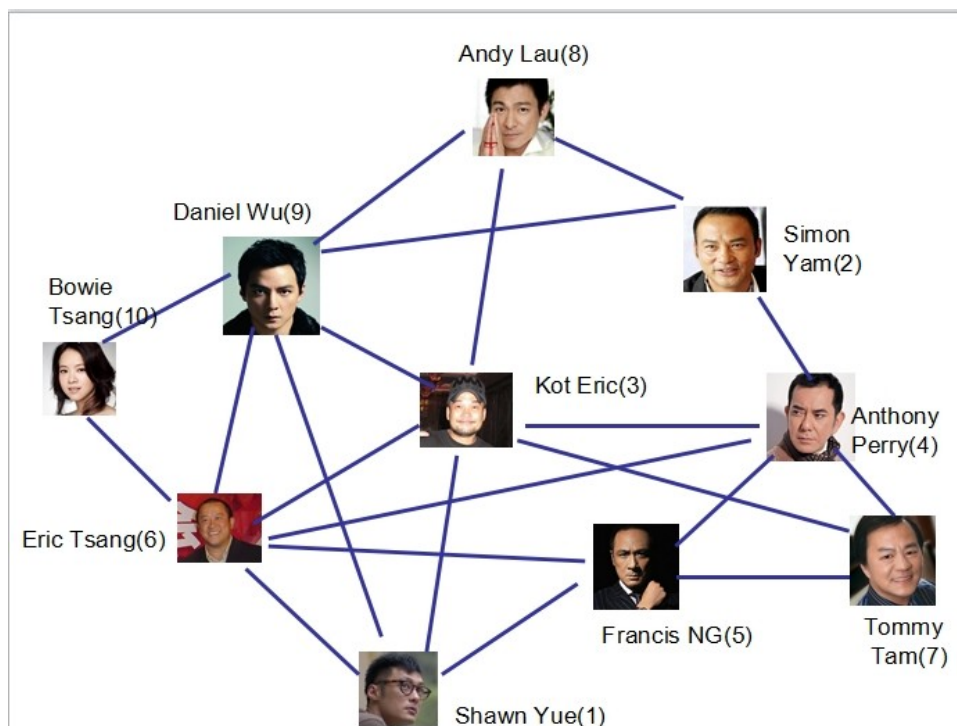


Figure 4.2.1: Actor Network

4.3 Influence Measures of Movie Actors

Influence Measures of Movie Actors According to our model, the sum influence of each actor in the actor collaboration network is given by the following formula: Sum influence equals to original influence plus attached influence (measured by collaboration times)

$$S_i = p * O_i + (1 - p) * A_i \quad (4.3.1)$$

The original influence of one actor means his own influence in the film circle, according to the data in Mtime, we can find a pie chart about the relative influence of these ten actors, which are measured by public voting, and then we turn the percentage numbers into specific decimals.

Pie chart data of actors original influence and corresponding table data are as follows:

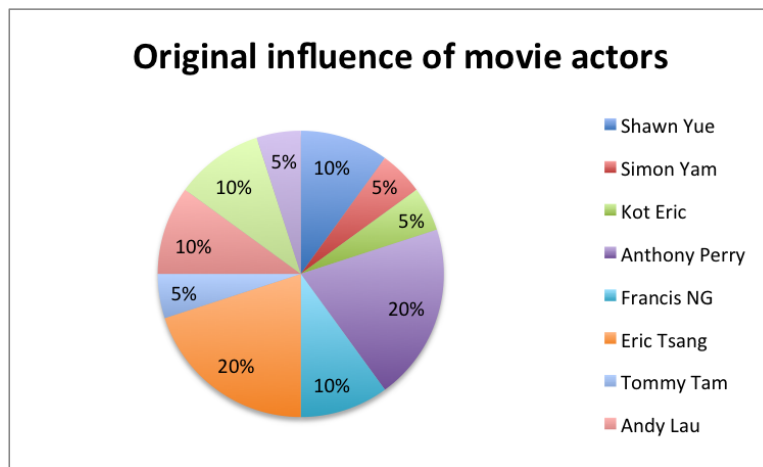


Figure 4.3.1: 1

Name	Origin influence
Shawn Yue	0.10
Simon Yam	0.05
Kot Eric	0.05
Anthony Perry	0.20
Francis NG	0.10
Eric Tsang	0.20
Tommy Tam	0.05
Andy Lau	0.10
Daniel Wu	0.10
Bowie Tsang	0.05

Figure 4.3.2: Actor's origin influence

4.4 Simulation and Analysis

4.4.1 First results

As we cannot decide the real value of p , we suppose p to be 0.5 which means that the original influence and the attached influence are equally weighted. According to our model and algorithm, we have found actor Anthony Perry is the most influential actor in the actor network whose sum influence reaches 0.154184, followed by Eric Tsang and Daniel Wu (0.154844 and 0.137581 respectively). The specific influence ranking when $p=0.5$ are as figure(4.4.1):

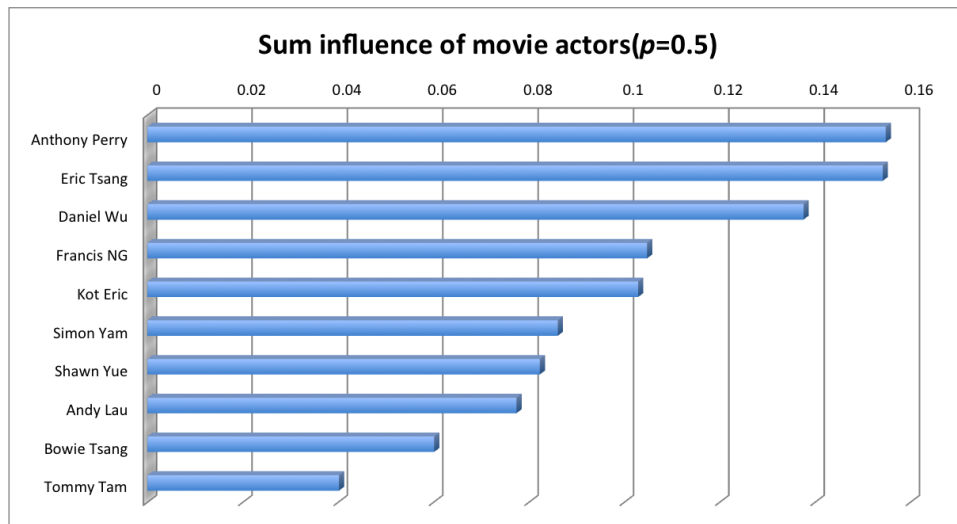


Figure 4.4.1: sum influence of movie actors(p=0.5)

4.4.2 Effect of different p value

To make our model more reliable and accurate, we need to discuss our results with different double-weighting factor p . In this case, we use the algorithm in our model to figure out the sum influence of movie actors with different p value. We adopt line graph to make our outcome more vivid. [7]

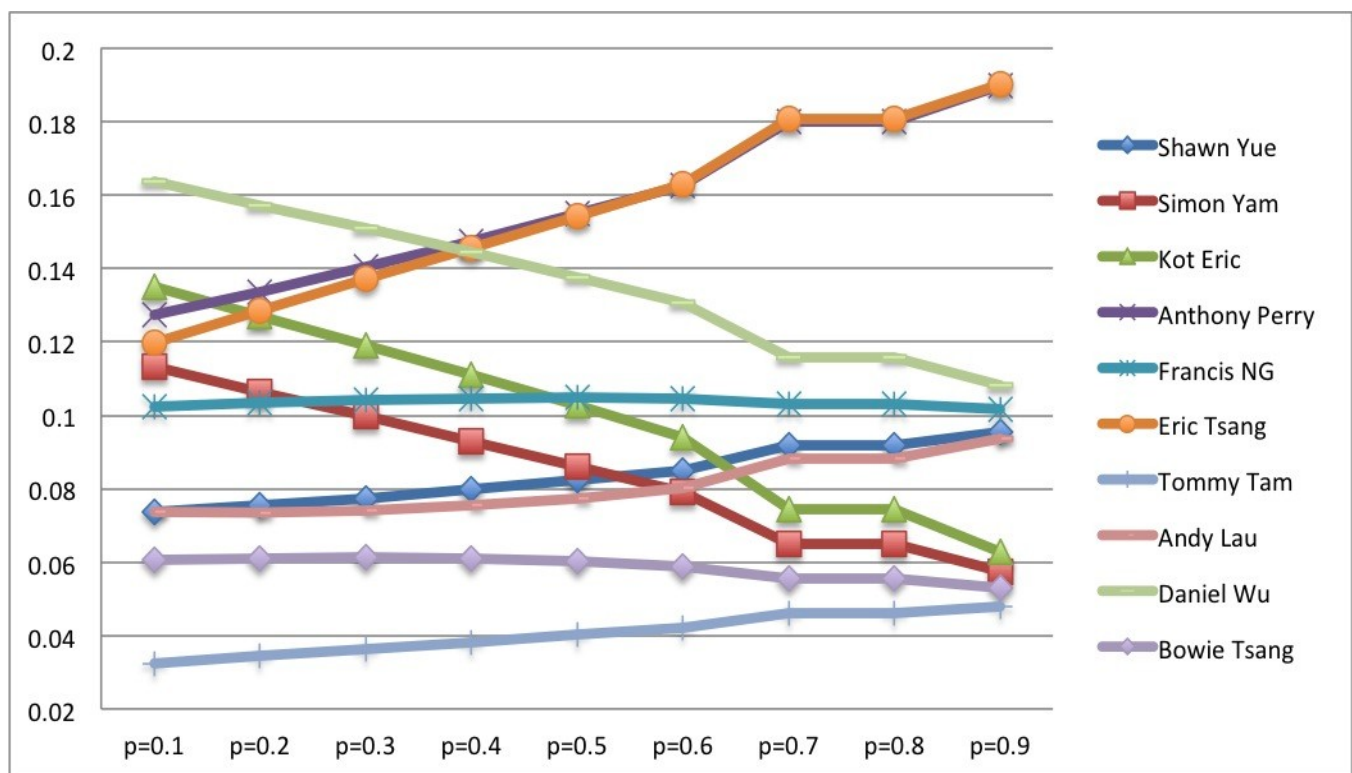


Figure 4.4.2: Sum influence of movie actors with different p value

In addition, according to our experiment data, we find that total sum influence in the network

decreases with increasing p value. This can be seen from the bar chart below. To explain the phenomenon, we think the reason is that the weight of attached influence shows a downward trend when value of p increases.

4.4.3 Simulation: The promotion of the model

In process of model building, we have used many network techniques to empower our model. To apply our model into generic network influence-ranking problems, we need to estimate accurate p value in specific networks. Hence, we state a general approach, by which can measure network influence as follows:

First, perform statistics of various network on the network analyzed to estimate p value of that network. Then, according to features of specified network, transform them into original influence and attached influence about nodes just like what we do in our model.

Finally, combine both of them together to figure out the sum influence. Thus, we can measure the influence and impact within general network based on our model and algorithm.

5 Application within Networks

5.1 Our Goal

- To calculate how the individuals choose collaboration partners in the network to make his influence get maximum promotion through the network influence model we have created.
- To make one given network systems transmission result maximum by using the network influence model we have created.

5.2 Maximizing Influence

- In a network that is in some certain area, we want to choose a number of nodes in it to develop relationships with ourselves in order to improve ourselves influence as high as possible. We can assume that there are m nodes in the network, we use the PRNM model to calculate the influence x_i of node i , then we denote the influence that node i shared with other nodes that are connected with node i by Y_i , $Y_i = X_i/n$, n is the number of the nodes that connect with node i , then we arrange y in order of descending.

If our target is to select k nodes in the network to develop relationships with ourselves to improve ourselves influence as high as possible, So it is more suitable for us to select the first k nodes. [8]

- We use the MaxDgree algorithm[9] (also be called the MDA algorithm) to calculate a networks maximum influence, which has on greatest advantage that has a small time complexity, and it has the ability to handle any kind of network.

Because it just needs to collect some properties of network such as degree centrality, number of degrees when selecting initial activation node. Then utilize these properties of network to directly choose initial activation node with the number of K .

Finally calculate their influences. The biggest disadvantage of the heuristic algorithm is the result of unstable solution. And its effect varies with the change of the internet. MaxDegree Algorithm is kind of simple and effective.

It requires to directly choose K biggest node degree as the initial activation node at the rank of the internet. Therefore, in some cases, it is feasible to solve the problem of MaxInfluence by adopting MaxDegree Algorithm. On the one hand, it has a small time complexity, on the other hand, it shows great effects. Hereinafter is the detail description of the MaxDegree Algorithm. As shown in table 5.2.1

Table 5.2.1: MaxDegree Algorithm

Input:	Social network G and k initial activation nodes with the number of K
Output:	k nodes of the max-degree and sum number of affected nodes

MAXDEGREE ALGORITHM

```

1  Social network G and k initial activation nodes with the number of K
2  readFile ( File*file ); // read network
3  buildModels (int model); //build IC model and LT model according to the parametres
4  simulateMoeels(int model,int num); //simulate spread model
5   $A \leftarrow \emptyset$ .
6  int sum=0;
7   $A \leftarrow \text{maxDegree}(k)$ 
8  for( $i = 1$  to  $M$ )
9    sum=sum+F(A,G)
10 end for
11 return A and sum/M
```

Obtain the function $\text{maxDegree}(k)$ that has k nodes of the maximum degree, and use the direct selection algorithm and the time complexity is $O(kn)$. And we prepare to obtain the node assemblage $F(A, G_r)$ in the assemblage A from the propagation diagram. . About the analysis of the reachability of each node, it can be accomplished in (s) that is the network diameter. The time complexity is mainly related to that algorithm should be done M times, and hence its time complexity is the $O((kn - k^2/2)M)$, n is the node number, k is the initial given node number. Although the maxDegree algorithm can have a quick selection for the initial maximum influence node, it has the flaw of penitential neighbor-overlap. In the figure ??, the social network which divided by dotted lines has two communities, and obviously, the node connection in the community is relatively close, while the connections among communities are comparatively sparse. There are two initial maximum influence nodes. According to the maxDegree algorithm, node A and F are selected as the initial nodes. A is next to F node, and therefore, the nodes chosen by the maxDegree algorithm is overlapped. To solve this problem, each community should select one node as the initial maximum influence node and then the initial nodes are spread to each community, which pause a better and faster information transmission.

5.3 Model Optimizing

The two models we mentioned in 5.2 is a kind of optimization based on the original model. In the first model in 5.2, we firstly run the PRNM model to figure out every nodes S , then we share the nodes S to n people who has cooperated with the node evenly. And if we want to improve ourselves influence mostly, we could choose the person who can share more influence to others. And in second model 5.2, we use Maxdegree algorithm to figure out the maximize influence of a network.

6 Conclusions

Although Mr. Gore has expressed concerns to some associates about the damage a brokered convention could cause, several associates said he was hopeful that one candidate would soon break through, sparing the party such an outcome. He told a close friend recently that his decision not to endorse feels like the right thing and that he remained optimistic the race is going to tip at some point, the friend said.

6.1 Sensitivity and Uncertainty Analysis

Sensitivity analysis is the study of how the uncertainty in the output of our model can be apportioned to different sources of uncertainty in its inputs. When applying our model to measure influence and impact in different networks, it is very difficult to guarantee the data that are completely clean, which means correct and complete. Fortunately, the sensitivity of PRNM model is very robust. PRNM not only considers the attached influence based on the relationship between nodes, but also analyzes nodes original influence using given data. The result is likely to take into consideration the sum influence measures of every node. So changing or missing a few connections would not have big influence on our model. Our model also considers some uncertain factors, such as the number of total nodes, specific measure methods of different relationship between nodes or other unexpected events. We have not included the interference from external environment, which may be too complex for us to study. In the future we may find a better method, and then we will take more factors into our model.

6.2 Strengths

We use P-Rank Network Model to study and measure the influence and impact within networks. Advantages are listed as follows:

- **Comprehensive:** we take both original influence and attached influence into consideration for network impact measures. Thus, the solution of our model pursues high credibility, while reducing the misjudgment rate.
- **Extendable:** the result of simulation shows that our model can be applied to other networks, not just research networks.
- **Reasonable:** the result of our model match well with the real situation, which proves the rationality and correctness of our model.
- **Flexible:** the sensitivity of our model is very robust. Since everything may be an accident, our model has its flexibility that allows the unexpected things to happen.

6.3 Weaknesses

- Our model is creative and robust, however, we do not have enough data to verify our model although we have discussed the data required. Hence, the correctness of our model remains to be verified.
- We use only one single effective parameter in our model: double-weighting factor p (that is, the relative weight of original influence), Sometimes this parameter are determined manually, which could cause uncertainty of our model solutions and the priority will change to some extent if different constant is set. Although we need to estimate the value of p in advance to validate the models stability, we still cant explain why the way we set the value of p can guarantee the accordance of our model results to the reality.
- When analyzing different networks in our models, we actually only focus on limited fields. So it is better if we have more detailed discuss about other network fields in addition to academic research or movie actors.

6.4 Future Work

- We will apply our model into more complex network influence analysis to optimize the PRNM algorithm. And our model will be enhanced to a larger extent.
- Find enough data to verify our model in order to improve both the sensitivity and robustness of our model.

References

- [1] Jin-qing Fang,Qiao Bi,Yong Li. Advances in theoretical models of network science[J]. Frontiers of Physics in China . 2007
- [2] Matthew Richardson, Pedro Domingos. The intelligent surfer :Probabilistic combination of link and content information in PageRank. Advances in Neurology . 2002
- [3] Narayanam, R,Narahari, Y.A shapley value-based approach to discover influential nodes in social networks. IEEE Transactions on Automation Science and Engineering . 2010
- [4] D. Garlaschelli,M. I. Loffredo.Generalized Bose-Fermi Statistics and Structural Correlations in Weighted Networks. Physical Review . 2009
- [5] Liu Z H,et al.Connective distribution and attack tolerance of general networks with both preferential and random attachments. Physics Letters . 2002
- [6] ZHANG Pei-pei,CHEN Kan,HE Yue,ZHOU Tao,SU Bei-bei,JIN Ying-di,CHANG Hui,ZHOU Yue-ping,SUN Li-cheng,WANG Bing-hong,HE Da-ren.Model and empirical study on some collaboration networks. Physica A Statistical Mechanics and its Applications . 2006
- [7] Newman MEJ.The structure and function of complex networks. SIAM Review . 2003
- [8] Zaixin Lu,Wei Zhang,Weili Wu,Joonmo Kim,Bin Fu.The complexity of influence maximization problem in the deterministic linear threshold model. The Journal of Combinatorics . 2011

- [9] W. Chen, Y. Wang, S. Yang. Scalable influence maximization for prevalent viral marketing in large scale social networks. Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining . 2010