Dynamics of Small World Effect and Random Fuzzy Graph

Yijin Zhang (zhangyj@cqupt.edu.cn)
Chongqing University of Posts and Telecommunications
Jie Huang
Chongqing University of Posts and Telecommunications
Zongbing Lin
Panzhihua University

Research Article

Keywords: Small World Effect, Random Fuzzy Graph, Social Networks

Posted Date: July 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-666633/v2

Dynamics of Small World Effect and Random Fuzzy Graph

Yijin Zhang[1*], Jie Huang[1], Zongbing Lin[2]

[1]Key Lab of Intelligent Analysis and Decision on Complex Systems, Chongqing Municipal,
Key Laboratory of Industrial Internet of Things & Networked Control, Ministry of Education,
Chongqing University of Posts and Telecommunications, Chongqing 400065, P.R. China
[2]School of Mathematics and Computer Science, Panzhihua University, Panzhihua 617000, P.R. China
*Corresponding author. Email: zhangyj@cqupt.edu.cn

Abstract: Small World Effect (or Six Degrees of Separation theory) has generated significant influence in the world. Many researchers and institutes have done lots of work on the study of it. We model Small World Effect by random process and Graph, calculate the probability that any two people i and j in the world can contact each other after n steps (forwarding messages by intermediaries) based on different R (average number of acquaintances everyone has in the world).When R=50 or 80, if $n \geq 5$, the probability is 0.848 and the search is very likely to happen. When R=150, n=4, this probability is 0.99994, that is, after 4 steps, any two people in the world will establish connection almost surely. In the sense of Dunbar's number (R=150), six degrees of separation becomes four degrees of separation. We propose the concepts of (directed) Random Fuzzy Graph for the very first time which can describe the fact of recognition among people best, because the relation among persons is random and fuzzy.

Keywords: Small World Effect; Random Fuzzy Graph; Social Networks

1. Introduction

Small World Effect (or Six Degrees of Separation theory) is now accepted by more and more people and widely used in power grids, genetic networks, collaboration graph of film actors, social networks, World Wide Web, many real-life networks, etc. It has generated significant influence in the world.

Small World Effect was first proposed in 1929 by Hungarian writer Frigyes Karinthy in his short story called "Chains". Any two people on earth, he argued, can be connected on average through a chain of five contacts[1]. In 1967, sociologist Stanley Milgram came up with the famous concept of six degrees of separation: the notion that every American is connected to every other American by a chain of some six people. The concept of the "small world" was born[2]. The "six degrees of separation" phrase stuck after John Guare's 1990 eponymous play[3].

In 1998, Watts and Strogatz introduced the "small-world" model of networks, which describes the clustering and short separations of nodes found in many real-life networks[4]. The model has been cited by 18,682 papers (indicated on Web of Science and 37,639 on Microsoft Academic), which is now considered to be one of the benchmark network topologies. Duncan J. Watts et al report on a global social-search experiment in which more than 60,000 e-mail users attempted to reach one of 18 target persons in 13 countries by forwarding messages to acquaintances. They can reach their targets in a median of five to seven steps (*Science*, 2003)[5]. This paper has been cited 307 (indicated on Web of Science and 1,005 on Microsoft Academic). Microsoft researchers Jure Leskovec and Eric Horvitz, who have filtered MSN text messages for a single month in 2006, compared 30 billion of communications among 240 million users. The results showed that any user can be related to the 180 billion group pairing of the whole database by an average of 6.6 people. 48% of users could have a connection up to six times, and 78% within seven times[6]. Facebook and University of Milan have jointly announced their new research on six-degree separation theory: they have determined that the average number of people between any two independent individuals in the world is 4.74. Six degrees is slightly larger to describe the actual connection between two individuals. In fact, on Facebook, the probability of five degrees between any two users is 0.996 and 0.92 for only four degrees[6,7].

All the above results are obtained through experiments, involving only a few people in the world (compared with the total population of the world), are local. In this paper, we applies random process and graph theory to calculate the corresponding probability from global perspective (to explore total World population).

2. Modeling

We assume that everyone in the world is keeping in touch with other people. As long as two persons can contact each other and exchange information between them, we say they "know" (or " recognize", or "acquaint").

Of course, some people know more persons and others less. Since we study the connection between any two people from a global viewpoint, we suppose that the average number of acquaintances everyone has in the world is R. The total population of the world is N. $0 \leq R \leq N-1$. In this way, if any two persons i and j know each other, i (or j) must be one of the R acquaintances of another person, the probability is $\frac{R}{N-1}$. If i and j do not know each other, then i (or j) must be one of the (N-1-R) strangers of another person, the probability is $\frac{N-1-R}{N-1}$. That is to say, the probability of any two people i and j in the world can reach each other through 0 step (needn't being introduced by anyone else) is $\frac{R}{N-1}$.

Next, we consider the probability that any two people i and j in the world can contact each other after n steps (by forwarding messages). n is a natural number and may take 0,1,2,..., N-2. (Note: n= 0 i.e. 0 step, which is the case mentioned above; n=N-2 means that all the remaining N-2 people in the world have forwarded messages, one step one person, and i, j get reached after N-2 steps.) Imagine (Fig.1.). Step 1. i sends messages to his (or her) R acquaintances, trying to contact j (the same results for contacting i from j), that is, R participators have been involved at the first step. Step 2. The R participators give messages to their own R-1 acquaintances (besides i) respectively, attempting to get in touch with j. There are R*(R-1) participators at the second step. Because we assume that everyone in the world know R persons only, these R*(R-1) persons must not know i. Each subsequent step is the same. Then, and so on,..., the nth step covers $R*(R-1)^{n-1}$ senders. As long as there is one person (up to R) of this $R*(R-1)^{n-1}$ senders who can contact j, then i and j are connected (Fig. 1.). Among these, i has at most $R^{n}*(R-1)^{\frac{n-1}{2}}$ chains to link j. So long as one chain works, the purpose is achieved.

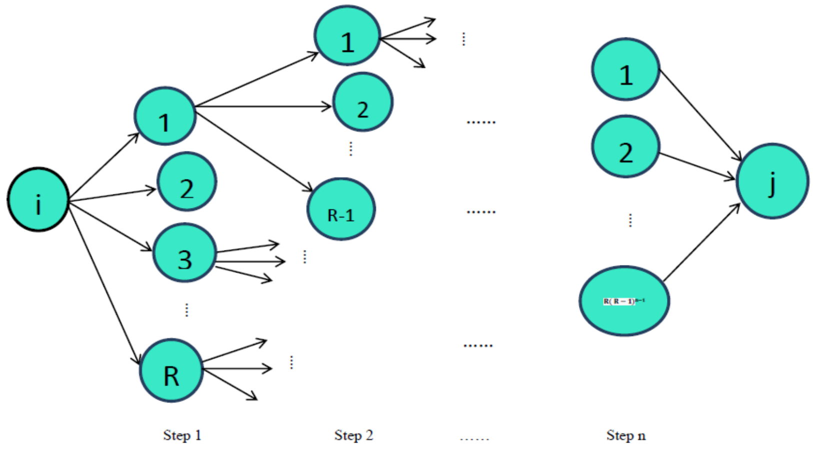

Fig. 1 Process of i trying to contact j. i sends messages to his (or her) R acquaintances. These R participators give messages to their own R-1 acquaintances (besides i) respectively. Then, and so on,..., the nth step covers $R*(R-1)^{n-1}$ senders. As long as there is one person (up to R) of these $R*(R-1)^{n-1}$ senders who can contact j, i and j are connected.

For convenience, we consider the probability that i can't contact j after n steps. Based on the above idea, because we assume that the average number of acquaintances everyone has in the world is R, the n steps can be completed smoothly in turn. Then why can't i contact j ? The reason must be that all of the $\mathbf{R*(R-1)^{n-1}}$ senders of the nth step can't connect j. They all belong to the group that j does not recognize. The probability is $\left(\frac{N-1-R}{N-1}\right)^{R*(R-1)^{n-1}}$. Therefore, the probability that any two people i and j in the world can contact each other through n steps (forwarding messages by intermediaries) is

$$\boldsymbol{P_{ij}^{(n)}} = \mathbf{1} - \left(\mathbf{1} - \frac{R}{N-1}\right)^{R(R-1)^{n-1}} \qquad n=1,2,... \tag{1}$$

After n steps, the number of all people included is $\frac{R*(R-1)^{n}-2}{R-2}$. If

$$\frac{R*(R-1)^{n}-2}{R-2} \geq N-1, \tag{2}$$

everybody in the world is reached. So once we obtain the minimum n satisfying Eq.2, the process terminates, and i contacts j.

3. Numeral calculating

Based on UN World Prospects, 2019 rev, World population is now 7,648,137,303[8]. In 2020, it was estimated to have reached 7.6 billion[9]. Here we select N =7,648,000,000.

There is a Chinese proverb that "News is always spread from one person to 10, then 10 to 100." One of its meanings is, news can be easily spread from one person to several or ten, even more persons, then to many many people rapidly. According to the research by Robin Dunbar, a professor of evolutionary psychology at Oxford University, the core circle of human beings may consist of three or five people, the closest ones of friends. Then there are 12 to 15 close friends (including confidants)... then 50 people. And the maximum number of close relationships one can maintain is 150 (Dunbar's number, also called Rule of 150), with an upper limit of 50 general friends. However, most people are far from reaching the upper limit [10-12]. So in this paper, R is taken as 10, 50, 80, 150 respectively. Note that N and R could be slightly larger or smaller, which will not have much effect on our results.

There are two special cases. If R=0, everyone knows 0 people, that is, any two people in the world don't know each other, and message can't be sent out, then $P_{ij}^{(n)} = 0$. If R=N-1, everyone knows all the other people, any two people in the world can contact each other, then $P_{ij}^{(0)} = 1$. These are consistent with Eq.1.

We substitute the values of N and R into Eq.1 and calculate with MATLAB\R2010b (4 digits after the decimal point. see Table 1).

Table 1. Values of $P_{ij}^{(n)}$. (End) indicates that N and R satisfy Eq. 2, the process terminates, and i is linked to j. If n≥11, $P_{ij}^{(n)} \equiv 1$, which are no longer listed.

Step n	R=10	R=30	R=50	R=80	R=150
0	1.3075e-009	3.9226e-009	6.5377e-009	1.0460e-008	1.9613e-008
1	1.3075e-008	1.1768e-007	3.2688e-007	8.3682e-007	2.9419e-006
2	1.1768e-007	3.4127e-006	1.6017e-005	6.6107e-005	4.3825e-004
3	1.0591e-006	9.8962e-005	7.8454e-004	5.2090e-003	6.3227e-002
4	9.5319e-006	2.8659e-003	3.7727e-002	3.3806e-001	9.9994e-001
5	8.5783e-005	7.9862e-002	8.4808e-001	1.0000e+000	1(End)
6	7.7179e-004	9.1052e-001	1(End)	1(End)	1(End)
7	6.9247e-003	1(End)	1(End)	1(End)	1(End)
8	6.0623e-002	1(End)	1(End)	1(End)	1(End)
9	4.3042e-001	1(End)	1(End)	1(End)	1(End)
10	9.9369e-001	1(End)	1(End)	1(End)	1(End)

In addition, there is an interesting case. Every country has administrative divisions. For example, the United States of American composed of states, Special District, and self-governing territories.They are divided into counties, cities, and town or township, then village or community. Well, any resident can contact the head of village or community, the head can contact the mayor of town, the mayor contacts the leader of the country or city, the leader contacts the governor, and the governor can certainly contact the president. There are 5 people participating. Then any other resident in the world can follow the similar path. Of course, the presidents of the two countries are connected. In this way, any two people in the world can be contacted through 10 people, and even in countries with complex administrative divisions, the figure is not much larger than 10.Therefore, $P_{ij}^{(n)} = 1 (n \geq 10)$. This case is quite special, and is artificial, unnatural, not very feasible.

4. Conclusions

Table 1 shows that when R=10, n=8, $P_{ij}^{(8)} \approx 0.061$, the value is greater than the probability of small probability event (0.05), that is, it can occur in a limited number of trials. When R=10, n=9, $P_{ij}^{(9)} \approx 0.43$, this is a relatively large probability, indicating that it is likely to occur at this time. When R=10, n=10, $P_{ij}^{(10)} > 0.99$. Furthermore, after 8 steps, the number of people covered is 53,808,401 and 4.3585e+009 after 10 steps, which is about 57% of the total population of the world. This embodies the speed of transmission, with same meaning as that Chinese proverb. When R = 30, n=5, $P_{ij}^{(5)} \approx 0.08$, the value is also greater than 0.05. When R=30, $n \geq 6$, $P_{ij}^{(n)} > 0.91$, the probability is very big. When R=50 or 80, if $n \geq 5$, the probability is bigger than 0.848 and the search is very likely to happen. When R=150, n=4, this probability is 0.99994, that is, after 4 steps, the connection will be accomplished almost surely. Six degrees of separation becomes four degrees of separation.

Yet, 150 is an upper limit, and most people are far from reaching it. Clustering coefficient, defined by D.J.W. and S.H.S. is the probability that two acquaintances of a randomly chosen person are themselves acquainted. They indicated that this clustering coefficient took values anywhere from a few percent to 40 or 50% for a variety of networks, and other researchers have also shown similar results for other networks[13,14]. Repeated calculations are possible, for example, person i_1 sends messages to person i_2 and i_3, but i_2 is likely to send message to i_3, or i_3 to i_2. What's more, not every acquaintance is willing to send messages. As a result, we think $R = 50$ or 80 is more realistic. Our results are in accordance with ones cited in Section 1.

Certainly, what we found are probabilities. Even though it is a large probability, there may exist some people in the world who will get touched in need of many steps (far more than the n we give) , even won't.

5. Definitions and examples

Random fuzzy graph

In the 1960s, Erds and Renyi, two great Hungarian mathematicians, established the random graph theory and initiated the systematic study of it. Random graph theory lies in the intersection of graph theory and probability theory, has always been the basic theory for studying complex networks[15,16]. In 1975, Rosenfeld introduced the concept of fuzzy graph[17], and fuzzy graph was widely studied. Now, We give the definitions of Random Fuzzy Graph. We regard each person as a point, and when two people know each other, we say that there is an edge connecting the two points. Besides, to everyone, some of his acquaintances are very familiar, and others are not, i.e. each acquaintance is fuzzy. So we introduce a membership function f to describe the degree of recognition which takes values on [0,1]. If two people are very familiar and all information can be exchanged between them, we take $f = 1$. If two people are familiar and can exchange partial information, we let $f=\frac{1}{2}$. If two people aren't acquainted, we have $f=0$. In other cases, according to the degree of recognition, we select f on $(0, \frac{1}{2})$ or $(\frac{1}{2}, 1)$.

Given a graph $G = (V, E)$, where V is the nonempty set of N vertices and E is the set of edges. Thus, we define a Random Fuzzy Graph $\tilde{G} = (V, E, P, f)$ on graph G, probability space $(\Omega, \mathcal{F}, P)$ and membership function f. Thus, in Small World Effect, we actually assume that the degree of recognition of any two people in the world is $f \geq \frac{1}{2}$. We denote by $f_{ij}^{(n)}$ the degree of recognition for people i and j after n steps in the world. $f_{ij}^{(n)}$ is the minimum one of the n steps. $P_{ij}^{(n)} = 0$ if and only if $f_{ij}^{(n)} = 0$.

As a supplement, there is such a situation in real life that one person i knows another person j well, and can even send messages to j, but j does not know i at all. In this case, we claim that degree of recognition for i to j through 0 step is $\frac{3}{4}$, and j to i through 0 step is 0, denoted by $f_{ij}^{(0)} = \frac{3}{4}$, $f_{ji}^{(0)} = 0$. At the same time, the probability that they know each other through 0 step is different, for example, the probability that i knows j is $P_{ij}^{(0)} = \frac{R}{N-1}$, and j knows i is $P_{ji}^{(0)} = 0$. So there exists a directed Random Fuzzy Graph.

Another definition

We now give another definition of Random Fuzzy Graph.

A Random Fuzzy Graph (RFG) or Fuzzy Random Graph (FRG) on a probability space $(\Omega, \mathcal{F}, P)$ and a universe X over a Graph G = (V, E), is a mapping:

$$\varphi\colon \mathrm{V} \times \mathrm{V} \to [0,1]_P \times [0,1]_f$$

$$(\mathrm{u},\mathrm{v}) \mapsto \varphi(\mathrm{u},\mathrm{v}) = (P(\mathrm{u},\mathrm{v}),\, f_A(\mathrm{u},\mathrm{v}))$$

for all $\mathrm{u},\mathrm{v} \in \mathrm{V}$, A is a fuzzy set in X, with the following properties:

$$\varphi(\mathrm{v},\mathrm{v}) = (1,1),$$

$$P(\mathrm{u},\mathrm{v}) = 0 \Leftrightarrow f_A(\mathrm{u},\mathrm{v}) = 0,$$

where (u,v) denotes the edge between u and v (can use the shorter notation uv), $(\mathrm{u},\mathrm{v}) \in \mathrm{E}$. $P(\mathrm{u},\mathrm{v})$ is the probability of the edge (u,v) and $P(\mathrm{v})$ is the probability of the vertex v. $f_A(\mathrm{u},\mathrm{v})$ is a membership function for the edge (u,v) belonging to the fuzzy set A in X and $f_A(\mathrm{u})$ for u. φ is called an undirected RFG if $f_A(\mathrm{u},\mathrm{v}) \equiv f_A(\mathrm{v},\mathrm{u})$ for any $\mathrm{u},\mathrm{v} \in \mathrm{V}$ and directed otherwise. If $P(\mathrm{u},\mathrm{v}) \equiv 1$, RFG becomes a Fuzzy Graph. If $f_A(\mathrm{u},\mathrm{v}) \equiv 1$, it is a Random Graph. φ is called a finite RFG if V is a finite set and infinite if V is an infinite set.

Another example

Another example is the spread of virus in a crowd. Every person is regarded as a point. Suppose that there are any two persons i and j in the crowd. Degree of their infection can be denoted by membership functions f_{i}, f_{j} respectively, which take values on [0,1]. 0 means no infection, and 1 infection; other cases, they take values in (0,1) (The following membership functions are the same). The probability that i infects j is $P_{ij}^{(0)}$, and j infects i is $P_{ji}^{(0)}$(through nobody, i.e., nobody between i and j) here $P_{ij}^{(0)} \neq P_{ji}^{(0)}$. The degree of i infecting j can be represented by membership function $f_{ij}^{(0)}$ and j infecting i is $f_{ji}^{(0)}$, (through nobody, i.e., nobody between i and j), $f_{ij}^{(0)} \neq f_{ji}^{(0)}$. Afterwards, i and j will also meet the third person, the fourth person,..., and so on. The probability that i infects j through n persons is denoted by $P_{ij}^{(\mathrm{n})}$, and j infects i through n persons is denoted by $P_{ji}^{(\mathrm{n})}$. The degree of i infecting j through n persons can be represented by membership function $f_{ij}^{(\mathrm{n})}$ and j infecting i through n persons is $f_{ji}^{(\mathrm{n})}$. Naturally, this is a Random Fuzzy Graph.

Third example

The third example is about computer networks. Each computer is regarded as a point. Any computer i receives or sends information more or less, represented by a membership function $f_i^{(0)}$. Two computers are said to be connected if they could exchange information. So any two computers i, j in the networks are likely to be connected or not. The probability that i connects j through n computers is $P_{ij}^{(\mathrm{n})}$, and j connects i by n computers is $P_{ji}^{(\mathrm{n})}$, here $P_{ij}^{(\mathrm{n})} = P_{ji}^{(\mathrm{n})}$. How much information is transmitted between two connected computers i, j is expressed by another membership function, for instance, $f_{ij}^{(\mathrm{n})}$ indicates how much information i sends to j through n computers and $f_{ji}^{(\mathrm{n})}$ j to i. It also becomes a Random Fuzzy Graph.

So, there are many cases in real world which can be described by Random Fuzzy Graph.

Finally, we point out that, although all our results come from small world network (one of social networks), we believe that our theory could be generalized to other kinds of networks. We will carry out further research on Random Fuzzy Graph and its application to other fields. And we hope that researchers studying other types of networks will also find our theory useful.

Acknowledgements

This work was supported by Natural Science Foundation Project of Chongqing(Grant No.cstc2019 jcyj-msxmX0716, cstc2020jcyj-msxm1397) and Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No.KJQN202000601).

Competing interests

The authors declare that they have no competing interests.

Data availability statement

The data that support the findings of this study are available from REFERENCES 8-12. In fact, there are several data used to support our work only, that is, total population of the world and average number of acquaintances everyone has in the world.

REFERENCES

[1] F. Karinthy. "Chain-Links", in *Everything is Different*. [Translated from Hungarian and annotated by Adam Makkai. Edited by Enikő Jankó.] (1929).

[2] J. Travers, S. Milgram, *Sociometry* 32, 425(1969).

[3] Six Degrees of Separation: A Play. Nov 21st(1990).

[4] Watts, D.J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* 393, 440–442 (1998).

[5] Peter Sheridan Dodds, Roby Muhamad, Duncan J. Watts. An Experimental Study of Search in Global Social Networks Science 301, 827; DOI: 10.1126/Science.1081058 (2003).

[6] Facebook and University of Milan have jointly announced new research on the six-degree separation theory [DB/OL], https://www.aliyun.com/zixun/content/2_6_11732.html, 2014-09-02.

[7] Backstrom L, Boldi P, Rosa M, et al. Four degrees of separation[J]. 2(9):1222-1227(2013).

[8] World Statistics [DB/OL], http://worldstatistics.org 2018-9-6.

[9] World mapper [DB/OL], https://worldmapper.org/maps/population-year-2018/

[10] Dunbar, R. M. Neocortex size as a constraint on group size in primates [J]. Journal of Human Evolution, 22(6):469-493(1992).

[11] Dunbar, Grooming, Gossip, and the Evolution of Language[M]. Cambridge MA: Harvard University Press(1996).

[12] Dunbar R. How many friends does one person need [J]. Evolutionary Psychology, (4):500-504(2010).

[13] Newman, M. E. J. *Proc. Natl. Acad. Sci.* 98,404–409(2001).

[14] Newman, M. E. J., Strogatz, S. H. &Watts, D. J. *Phys. Rev. E Stat. Phys.Plasmas Fluids Relat. Interdiscip.Top.*, 64, 026118(2001).

[15] Erdős, P. Rényi, A "On Random Graphs I" in Publ. Math. Debrecen 6, p. 290–297(1959).

[16] Béla Bollobás, *Random Graphs*, 2nd Edition, Cambridge University Press(2001).

[17] A. Rosenfeld, A. "Fuzzy graphs".In: Zadeh, L.A., Fu, K.S., Tanaka, K.,Shimura,M.(eds.), *Fuzzy Sets and their Applications to Cognitive and Decision Processes*, Academic Press, New York, 77–95(1975).

Figures

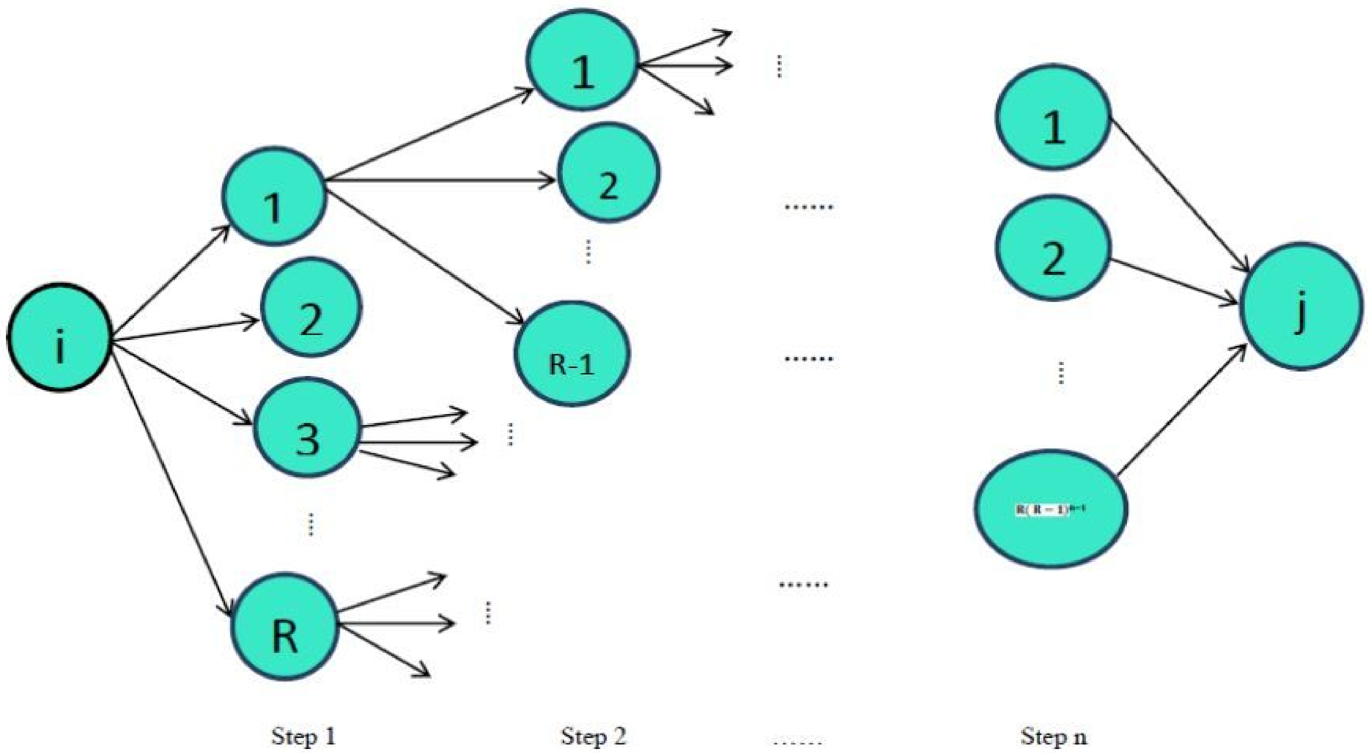

Figure 1

Process of i trying to contact j. i sends messages to his (or her) R acquaintances. These R participators give messages to their own R-1 acquaintances (besides i) respectively. Then, and so on,..., the nth step covers R * (R-1)n-1 senders. As long as there is one person (up to R) of these R * (R-1)n-1 senders who can contact j, i and j are connected.