Team Control Number

For office use only

T1 _____

T2 _____

T3 _____

T4 _____

42931

Problem Chosen

D

For office use only

F1 _____

F2 _____

F3 _____

F4 _____

2016

Measuring the Evolution and Influence in Society's Information Networks

Summary Sheet

Information is spread quickly in today's tech-connected communications network. In order to measure the evolution and influence in society's information networks, we've determined to build four models.

Firstly, we've built an interpersonal information propagation model based on **Cellular Automata**. According to **the Law of Large Number**, we find that the average spread speed of information is of positive relevance with the strength of the information, while the instantaneous propagation velocity of information is of negative relevance with the online information sharing level. We've also found a critical value of 0.6(out of 1) existing in information diffusion–the information can be diffused only when **the inherent value of information** exceeds 0.6, otherwise it will be filtered out.

Secondly, through **fuzzy evaluation model**, we are able to judge the inherent value of news from different times. We've defined that only when the inherent value of the information exceeds the critical value 0.6, it will be qualified as news.

Thirdly, we've built **a gray prediction model**. Based on the existing data, we've made predictions to usage rates of different media, with the result of relative error for 4.5-6.5%. Moreover, we've predicted the situation in 2050 based on our model: by 2050, the usage rates of the several primary media will respectively reach: newspapers 12.65%, radio 16.65%, TV 36%, Internet >100%.

Finally, we've built **the fuzzy cellular automata model** to reflect people's attitudes (indicated as 0-9) towards information by fuzzy inferences. Then, through **Mamdani fuzzy reasoning algorithm**, we've come to the internal probability for different cellular to influence each other and the external probability from the influence of the topology of mass media. These two probabilities are independent, but they together affect the flow of information. Thus we can draw the following conclusions: during the process of information dissemination , people with similar views are likely to stay together; Mass media plays a guiding role in the process of information dissemination. People tend to believe what mass media report; Without mass media, people's opinions toward the news tend to be neutral, which is to say, extreme views will gradually disappear.

KEY WORDS: Fuzzy Cellular Automata, Fuzzy Inference, Gray Prediction, Mamdani Reasoning Algorithm, Law of Large Number

Contents

1 Introduction

Technological and social developments have increased the speed with which information can spread, as well as influenced its content. The effects of information dissemination includes organizational functions which consists of informing, expressing, explaining and guiding, and social functions which consists of politic, economic, education and culture.

Information is spread quickly in today's tech-connected communications network; sometimes it is due to the inherent value of the information itself, and other times it is due to the information finding its way to influential or central network nodes that accelerate its spread through social media. While content has varied, the prevailing premise is that this cultural characteristic to share information (both serious and trivial) has always been there.

By taking a historical perspective of flow of information relative to inherent value of information, the Institute of Communication Media (ICM) seeks to understand the evolution of the methodology, purpose, and functionality of societys networks. Our task is to analyze the relationship between speed/flow of information vs inherent value of information.

2 Task 1: Explore the Flow/Filter of Information with Models

2.1 Model A: Cellular Automata Information Dissemination Model

2.1.1 Assumptions

1) In the relationship network, a member receive information by the information dissemination process from members associated with him. And the process is influenced by trust between members, will to spread and internal value of information.

2) In the original stage of dissemination, only one member knows the information while the others do not know.

3) The process of dissemination is unidirectional, and member will not forget the information once he receives it.

4) The internal value of information remain unchanged during the dissemination process.

5) The member can only send information in the next step time after he received the information.

6) The network is sealed, so there is no information from outside and to outside.

7) The only way for members to receive information is the dissemination process, which means he can not know the information by self-propagation like learning.

2.1.2 Small-world Theory

Small-world Model is a model with large clustering coefficients[1] and short average path length.Most of the average path length of real network is much less than imagination, which can be called *Small-world Effect*.

The steps to analyze the small-world model are that: Firstly find out the topology of abstract of social network based on complex network theory. Then analyze the propagation mechanism according to certain rules. Finally analyze how to affect this process by certain measures.

2.1.3 Cellular Automata Model

Cellular Automata is a special methods available on computer simulation which designed according to the characteristics of complex system, its space, time and state of system is all discrete[2]. In fact, cellular automata is a discrete dynamical systems. Cellular Automata is abstract of objective world, which divides the continuous space into discontinuous points, and use a limited number of combiations of the discrete states to characterize state of space. Cellular automata can be viewed as a simple model of a spatially extended decentralized system made up of a number of individual components (cells)[3]. The communication between constituent cells is limited to local interaction. Each individual cell is in a specific state which changes over time depending on the states of its local neighbors. The overall structure can be viewed as a parallel processing device.

Interpersonal Information Communication is information exchange between the one and another, including direct face-to-face communication and indirect communication with by means of media. According to the assumption, in particular interpersonal networks, when a member receive information he will spread this information towards individuals associated with him. Though there are many factors that can influence the next member's successful reception of information, this paper mainly takes the credit between members and the will to spread into consideration. Additionally, while paths and speed of the information dissemination is different in different networks, this model is based on small-world networks.

In the network each node stands for a member. Fig.1 is a simulation of interpersonal network with 20 nodes. Each node is connected to the K nodes closest to it.

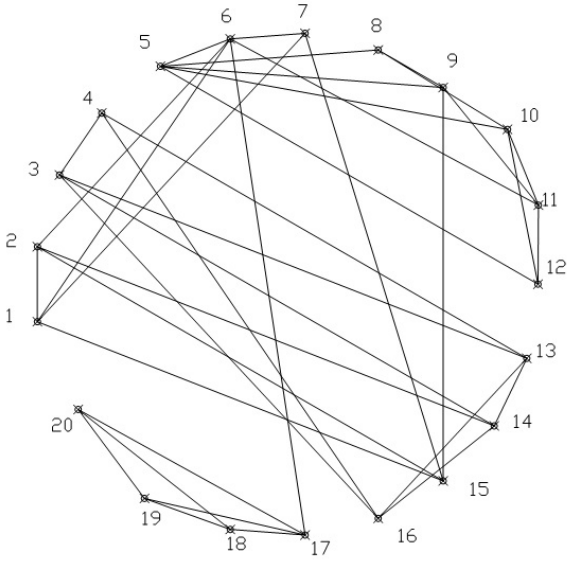

Figure 1: Interpersonal Network

Cellular Automata is a set of cells with discret states that distribute in a simulated space[4].

The credit(credit $\in [0, 1]$) is the parameter to measure how close relationship two members have. The higher the credit is, the closer the relationship will be. As the network is undirected, take member 1 and member 2 for example, credit(1,2)=credit(2,1). The will to spread indicates the intensity of one's subjective willingness to disseminate the information to other members in the network. Just the same as credit, the higher the value of will is, the greater probability he will disseminate the information to other members. In this paper, we mainly consider the influence of these two elements to the dissemination process by studying

the process under different distribution.

Under the structure of interpersonal network, consider two situations:

- While the credit is uniform distribution, the will is normal distribution, uniform distribution or binomial distribution
- While the will is uniform distribution, the credit is normal distribution, uniform distribution or binomial distribution

In the simulation, we use Matlab to get data meeting the required distribution at random. To evaluate the dissemination, we define:

1. **Spread speed** = all the nodes that have received the information/the step time of the dissemination.
2. **Average spread speed** = sum of spread speed of all groups/number of experiment.
3. **Instantaneous velocity** = (nodes that have received the information in time i - nodes that have received the information in time (i-1))/i $(1 < i \leq 20$,i is integer)

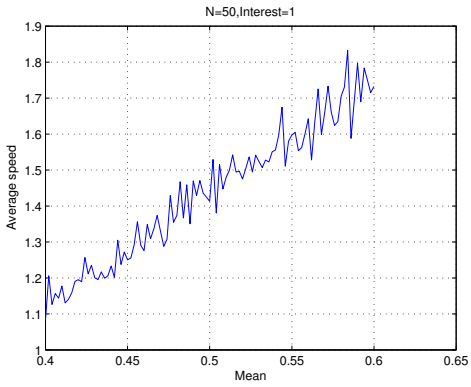

Figure 2: The influence of means to average velocity when credit is normal distributed(Variance = 0.02)

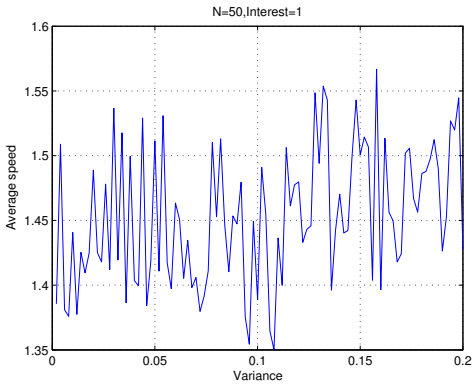

Figure 3: The influence of variance to average velocity when will is normal distributed(Mean = 0.5)

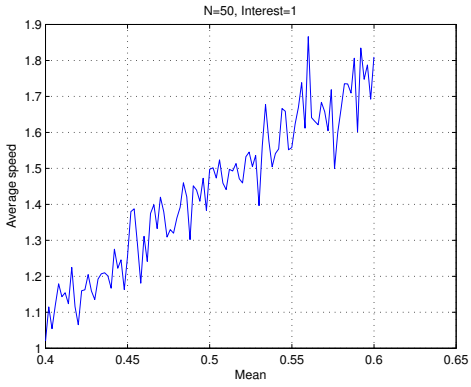

Figure 4: The influence of means to average velocity when will is normal distributed(Variance = 0.02)

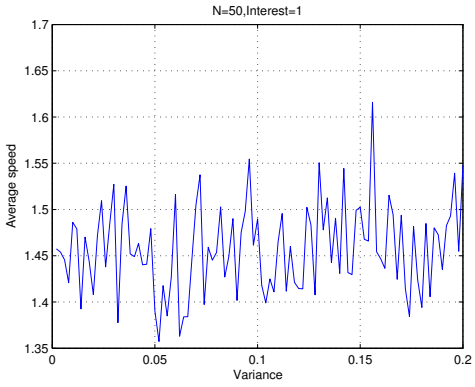

Figure 5: The influence of variance to average velocity when will is normal distributed(Mean = 0.5)

The result of the simulation suggests that:

In the case of a uniform distribution of trust, means of will to disseminate information to the average propagation velocity is approximately in proportional relationship, and the average speed of spread increases as the variance of trust increases.

In the case of a uniform distribution of will to disseminate, means of trust to average speed of spread is approximately in proportional relationship, and the average speed of spread decreases as the variance of trust increases.

The instantaneous velocity of information spread decreases as the information sharing level in the network increases. When the information sharing level becomes saturated, the instantaneous velocity is zero. The interpersonal spread speed is mainly influenced by interpersonal network.

2.1.4 Effect of Inherent Value

We use I to represent the attractiveness of information, $I \in [0, 1]$. It's obvious that the higher the I is, the larger probability individuals can receive the information. Another external factor that can influence the receiving of information is the effects of surrounding people. To quantify this influence of surrounding people, we take the number of people that have received this information into consideration. The more the nearby people have received the information, the larger the probability the individual will receive the information. In that influence of different people is the same, so the influence of an individual towards other people can be represented by P, $P \in [0, 1]$, which means the proportion of disseminators of information accounted for all the associated people. Personal preference is represented by ϕ, $\phi \in [0, 1]$. We get the factor value P from cellular automata model. Among the external factors that influence the receiving of information, the two factors I and P play the different roles. C is used to represent the proportion of I towards the whole influence, $C \in [0, 1]$.

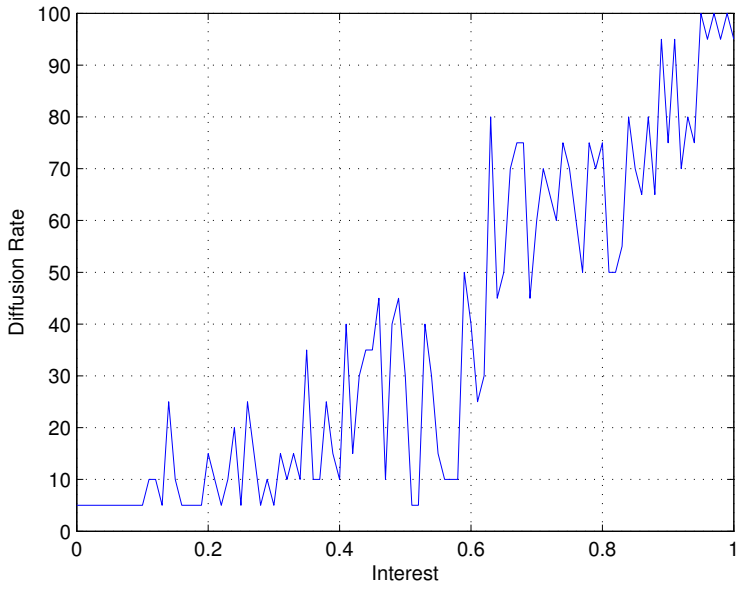

Figure 6: The influence of information attractiveness to diffusion rate

At the beginning, we choose an individual at random, and assume that he receive the information. $A^{(n)}$ is the set of people that receive the information at time n. Then at time n+1, the state of an individual that is not in $A^{(n)}$ is decided by state of people in time n who are associatied with him. If no one around him has received the information, he would not know the information. Otherwise whether he receives the information or not is decided

by the whole influence and his own preference ϕ. If the whole influence is larger than ϕ, he will receive the information and get into set $A^{(n+1)}$.

Fig.6 shows the influence of information attractiveness to diffusion rate. From it we can draw the conclusion that there is a threshold for information dissemination, which is 0.6. Only when the attractiveness is above 0.6 can the information spread out.

Fig.7 to Fig.10 below shows the influence of information attractiveness to spread speed:

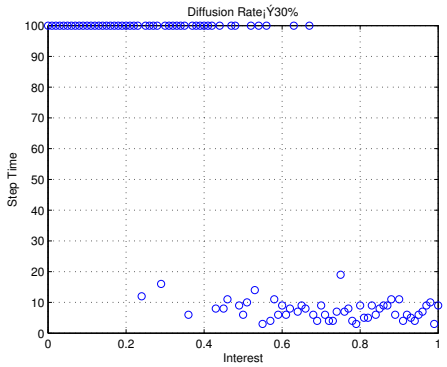

Figure 7: The influence of information attractiveness to spread speed(greater than or equal to 0.3)

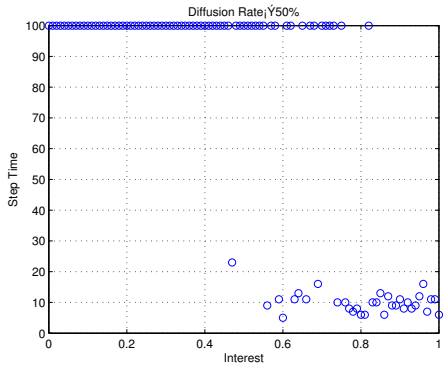

Figure 8: The influence of information attractiveness to spread speed(greater than or equal to 0.5)

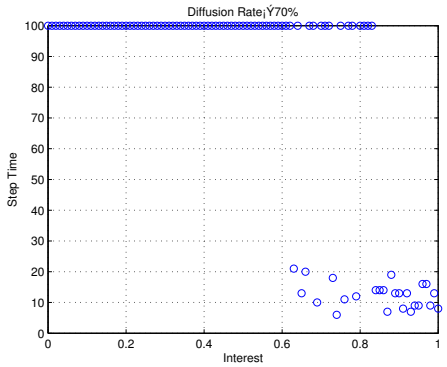

Figure 9: The influence of information attractiveness to spread speed(greater than or equal to 0.7)

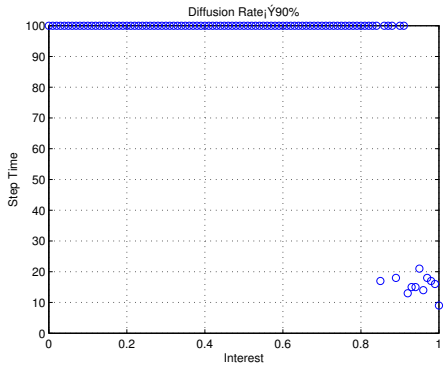

Figure 10: The influence of information attractiveness to spread speed(greater than or equal to 0.9)

From figures that show the influence of information attractiveness to spread speed and time to spread we can conclude:

1. The diffusivity increases as the information attractiveness increases.
2. When the information attractiveness is large enough, it has no effect on time of information dissemination.
3. When the information attractiveness is small enough(far less than 0.2), time to disseminate is not long. However, in such case, information diffusion area is very small, which there is nearly no diffusion of information.

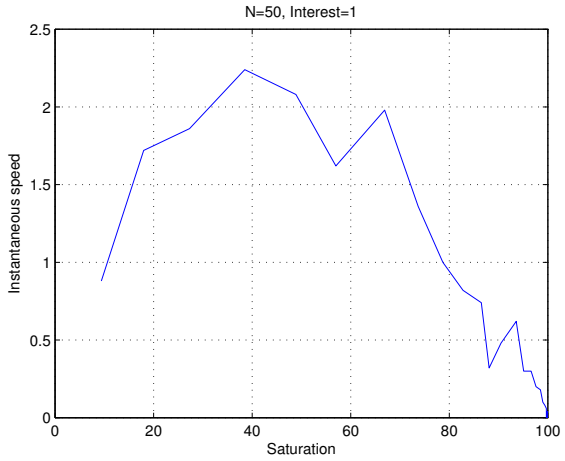

Figure 11: The influence of saturation to instantaneous velocity

2.2 Model B: Fuzzy Evaluation Model

2.2.1 Establishment of Evaluation System

News is packaged information about current events happening somewhere else.

To figure out what qualifies as news, we should firstly choose five properties of news: Importance, Significance, Authenticity, Newness[5], and Interest.

We use a set $U = \{U_1, U_2, U_3, U_4, U_5\}$ to represent the five elements we choose. U_1 to U_5 correspond to them respectively. As for the weights of the five elements, we use AHP(Analytic Hierarchy Process)[6] to decide. Considered that there are five media(television, newspaper, internet, radio and mobile), five coefficient matrix is shown as below:(the script t,n,i,r,m represent television, newspaper, internet, radio and mobile)

$$A_t = \begin{bmatrix} 1 & 1/2 & 4 & 3 & 3 \\ 2 & 1 & 7 & 5 & 5 \\ 1/4 & 1/7 & 1 & 1/2 & 1/3 \\ 1/3 & 1/5 & 2 & 1 & 1 \\ 1/3 & 1/5 & 3 & 1 & 1 \end{bmatrix} \quad A_n = \begin{bmatrix} 1 & 1/3 & 4 & 1 & 3 \\ 3 & 1 & 6 & 3 & 5 \\ 1/4 & 1/6 & 1 & 1/3 & 1/2 \\ 1 & 1/3 & 2 & 2 & 1 \\ 1/3 & 1/5 & 2 & 2 & 1 \end{bmatrix}$$

$$A_i = \begin{bmatrix} 1 & 1/2 & 3 & 4 & 3 \\ 2 & 1 & 5 & 7 & 5 \\ 1/3 & 1/5 & 1 & 2 & 1 \\ 1/4 & 1/7 & 1/2 & 1 & 1/2 \\ 1/3 & 1/5 & 1 & 2 & 1 \end{bmatrix} \quad A_r = \begin{bmatrix} 1 & 1/2 & 4 & 3 & 4 \\ 2 & 1 & 7 & 5 & 7 \\ 1/4 & 1/7 & 1 & 1/2 & 1 \\ 1/3 & 1/5 & 2 & 1 & 2 \\ 1/4 & 1/7 & 1 & 1/2 & 1 \end{bmatrix} \quad A_m = \begin{bmatrix} 1 & 2 & 4 & 6 & 4 \\ 1/2 & 1 & 2 & 3 & 2 \\ 1/4 & 1/2 & 1 & 2 & 1 \\ 1/6 & 1/3 & 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & 1 & 2 & 1 \end{bmatrix}$$

The weight vector can be got:

$$U_t = \begin{bmatrix} 0.2636 & 0.4758 & 0.0538 & 0.0981 & 0.1087 \end{bmatrix}^{\mathrm{T}} \tag{1}$$

Check the consistency: CI = 0.0180, CR = 0.0161.

$$U_n = \begin{bmatrix} 0.2093 & 0.4683 & 0.0558 & 0.1385 & 0.1282 \end{bmatrix}^{\mathrm{T}} \tag{2}$$

Check the consistency: CI = 0.0775, CR = 0.0692.

$$U_i = \begin{bmatrix} 0.2649 & 0.4787 & 0.0992 & 0.0581 & 0.0992 \end{bmatrix}^{\mathrm{T}} \tag{3}$$

Check the consistency: CI = 0.0067, CR = 0.0060.

$$U_r = \begin{bmatrix} 0.2699 & 0.4925 & 0.0637 & 0.1101 & 0.0637 \end{bmatrix}^{\mathrm{T}} \tag{4}$$

Check the consistency: CI = 0.0665, CR = 0.0058.

$$U_m = \begin{bmatrix} 0.4590 & 0.2295 & 0.1216 & 0.0684 & 0.1216 \end{bmatrix}^{\mathrm{T}} \tag{5}$$

Check the consistency: CI = 0.0033, CR = 0.0030. All the consistency of the matrix can be accepted.

2.2.2 Establishment of Evaluation Set

A evaluation set is a collection of the overall results of the evaluation of the target: $V = \{V_1, V_2, \cdots, V_n\}$(n is integer). We consider the situation n=5, which represents "high","relatively high","medium","relatively low","low".(V_1 corresponds to "high")

To determine the judgement matrix, judge element i of the target in the set U. Here consider a concept called degree of membership r_i, which means the degree that the judgement correspond to the element j in the evaluation set V. In such case, the result of element u_i can be represented by fuzzy set R.

Take the event that President Lincoln was assasinated for example. In that period, the main method of information diffusion is newspaper. Ten judgements is taken towards the newness of the information, one of which are V_1, two V_2, four V_3, three V_4, one V_5. So r_{41}=0.1,r_{42}=0.2,r_{43}=0.4,r_{44}=0.3,r_{45}=0.2. And matrix R:

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0.1 & 0.2 & 0.4 & 0.3 & 0.2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{6}$$

2.2.3 Overall Merit of Influence of News

In that the weight matrix U=$\begin{bmatrix} U_1 & U_2 & U_3 & U_4 & U_5 \end{bmatrix}$ = $\begin{bmatrix} 0.1362 & 0.3515 & 0.3649 & 0.0901 & 0.0476 \end{bmatrix}$, we can get the overall evaluation matrix:

$$Y = U \times R = \begin{bmatrix} 0.7472 & 0.0277 & 0.0554 & 0.0416 & 0.1558 \end{bmatrix} \tag{7}$$

After normalization:

$$Y^{'} = \begin{bmatrix} 0.7271 & 0.0270 & 0.0539 & 0.0404 & 0.1516 \end{bmatrix} \tag{8}$$

The matrix $Y^{'}$ means the degree of membership to high influence is 0.1375, the degree of membership to relatively high influence is 0.3549, and so on.

2.2.4 Defuzzification

1. **Max degree of membership**

   Choose the max value of the five number. From the example, as the max degree of membership is 0.3685. So the influence of the news can be seen as "medium".

2. **Quantify with percentage**

   Assume the five degrees V_1 to V_5 correspond to the percentage as the table below shows:

   So the overall evaluation in the example can be quantified as

$$Result = 95 \times 0.7271 + 85 \times 0.0270 + 75 \times 0.0539 + 65 \times 0.404 + 30 \times 0.1516 = 82.5854 \tag{9}$$

V_1	V_2	V_3	V_4	V_5
100-90	89-80	79-70	69-60	59-0

Table 1: Percent corresponding to the five degrees

Turn the hundred mark system into one point system. In this model, the highest result is 95 and the lowest result is 30, so the final result for television is

$$Result = \frac{82.5854 - 30}{95 - 30} = 0.81 \tag{10}$$

From the information propagation model, the information can only spread out when its attractiveness is above 0.6. So we use the value of 0.6 to be the threshold of determination of news. If the final result is above 0.6, we say this information is qualified as news.

2.3 Model C: Fuzzy Gray Prediction Model

A fuzzy gray prediction model is a statistical model that is usually used to predict future situation using past data. In that the data we can get has irregular intervals, we choose to establish irregular intervals sequence prediction model[7].

Assume the irregular intervals sequence is $X^{(0)}(t) = [x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)]$, the time sequence corresponding to it is $T^{(0)}(i) = [t_1, t_2, \cdots, t_n]$. All the data we use is chosen from existing data. Then the average time interval is

$$\Delta t_0 = \frac{1}{n-1}(t_n - t_1) \tag{11}$$

The unit period of difference coefficients of actual observation period and the average period is

$$\mu(t_i) = \frac{t_i - (i-1)\Delta t_0}{\Delta t_0} \tag{12}$$

The difference sequence of actual observation period:

$$\Delta x^{(0)}(t_i) = \mu(t_{i+1})[x^{(0)}(t_{i+1}) - x^{(0)}(t_i)] \tag{13}$$

Do a cumulative generation of difference sequence $\Delta X^{(0)}(t)$ and the observation sequence $X^{(0)}(t)$, we can get sequence $\Delta X^{(0)}(t)$ and $X^{(0)}(t)$. After equalization,

$$\Delta \bar{x}^{(1)}(t) = \frac{1}{2}[\Delta x^{(1)}(t) + \Delta x^{(1)}(t+1)] \tag{14}$$

$$\bar{x}^{(t)} = \frac{1}{2}[x^{(1)}(t) + x^{(1)}(t+1)] \tag{15}$$

From equation 4 and 5,the cumulative mean of the interval is :

$$\otimes \bar{x}^{(t)} = \bar{x}^{(t)} + \Delta \bar{x}^{(1)}(t) \tag{16}$$

Based on this can the differential equation be established:

$$\frac{d \otimes \bar{x}^{(1)}(t)}{dt} + a \otimes \bar{x}^{(1)}(t) = u \tag{17}$$

To get the parameter vector according to the principle of least squares method. Then we can get the prediction of $\hat{x}^{(0)}(t)$

$$\hat{x}^{(0)}(t) = \hat{x}^{(1)}(t) - \hat{x}^{(1)}(t - \Delta t_0) \tag{18}$$

2.4 Model D: Fuzzy Cellular Automata Model

2.4.1 Assumptions

1) The type of cellular automata is Moore type.

2) In order to derive the distribution of exposures a node receives over time as a function of time, we model the arrival of exposures as a binomial distribution[8].

3) We approximated the flux of exposures as constant in time such that each interval of time has an equal probability of an exposure arriving, so the sum of the events is a standard binomial random variable.

2.4.2 Influence of Internal Exposures by Infected Neighbours

In daily life, many vague concept exists in the way people think, rather than purely positive or negative attitude. So we can only use fuzzy set to describe this situation. Due to the characteristic the information have during dissemination, and taking cellular automata theory and fuzzy inference algorithm into consideration, we establish the fuzzy cellular automata model. Two input variable including environmental adaptability k and preference degree h are led into this model.

The fuzzy cellular automata model is based on a N*N dimension network, in which each cell represents an individual. Compared to the two states(received and not received) for each individual in model A, the value of cell state is continuous, $H \in [0, 9]$. On the other hand, there are three main attitudes of people, which are positive, neutral and negative. So we define that if $H < 3$, the attitude of the individual is negative, if $3 \leq H \leq 6$, the attitude is neutral, if $H > 6$, the attitude is positive.

Fuzzy inference is a process using known fuzzy conditions and rules to get a fuzzy conclusion[9]. As shown in Fig.12, two input variable k and h experience the process of fuzzy interface. Then the conclusion can be drawn from fuzzy rules and the input. Because in reality, the overall result should be a certain one, so we should clear the fuzzy conclusion to get a proper accurate result p. The general fuzzy rule is that: if x_1 is U and x_2 is V then y is W. U, V and W are subset of fuzzy sets of universal set X,Y,Z.

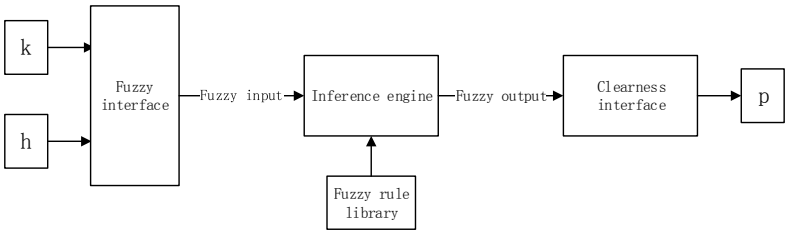

Figure 12: Process of fuzzy inference

In that different individuals have different knowledge background and personality, so they may have differnt attitudes towards the same thing, and the difference may be marked. Define preference H(i, j) as a fuzzy variable. Then we divide the preference H(i, j) into three fuzzy subsets: negative, neutral and positive. Fig.13 shows the fuzzy subset distribution of preference degree and membership.

In the actual spread process of the information, individuals' attitudes towards one message are decided not only by their own opinions, but also influenced by people surround them. To describe the degree one would change his attitude under the pressure of others and the society, we define environmental adaptation K(i, j) as a new input variable. And it

is also divided into three fuzzy subsets: conform, neutral and firm. Fig.14 shows the fuzzy subset distribution of environmental adaptability and membership.

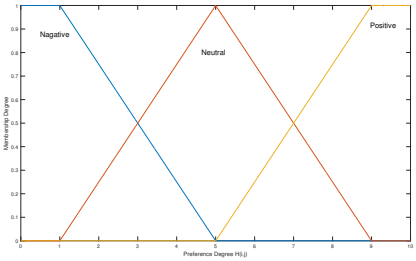

Figure 13: Fuzzy subset distribution of preference degree and membership

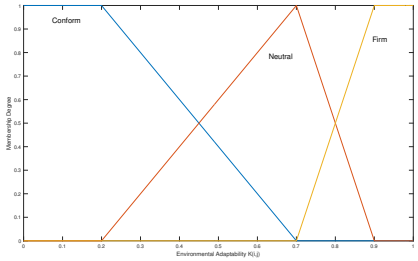

Figure 14: Fuzzy subset distribution of environmental adaptability and membership

On the other hand, the output variable P(i, j) should also be defined. P stands for the probability of the state of individual to transit in next time. Fig.15 shows the fuzzy subset distribution of environmental adaptability and membership.

Individual is influenced by surrounding individual, but the degree of influence is different, we define weight W to distinguish the difference of surrounding individual's influence. If a person's attitude is firm and have obviously tendentious, then W have large value.

$$W = \begin{cases} 4.5 - H, & 0 \leq H < 4.5 \\ H - 4.5, & 4.5 < H \leq 9 \end{cases} \tag{19}$$

The state of individual in next step time is

$$H_{ij}(t+1) = \frac{H_{sum}(t)}{W_{sum}(t)} \tag{20}$$

In the equation, $H_{sum}(t)$ means in time t, the product of preference H of individual and surrounding eight individuals and the weight W, $W_{sum}(t)$ means the sum of weight W of individual and surrounding eight individuals. Whether H transit is denpendent on the probability P.

There are nine fuzzy rules for the fuzzy inference as shown in Table 2:

2.4.3 The Influence of External Exposures by Mass Media

People usually cooperate with each other in daily life. The intensity of the influence lead to the consistency of the attitude of the whole system. Mass media serves as a guide in

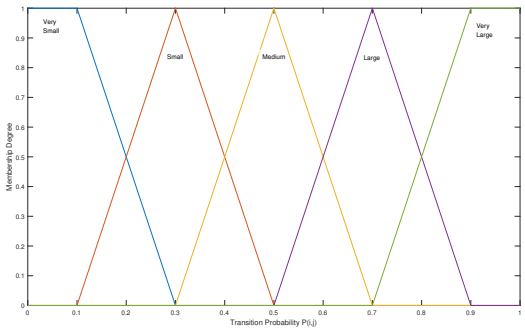

Figure 15: Fuzzy subset distribution of transition probability and membership

Table 2: Fuzzy rules

p \ k / h	conform	neutral	firm
negative	large	small	very small
neutral	very large	medium	small
positive	large	small	very small

the process of information diffusion, which makes the individual correspond to it. Affected by strong interaction, the society can face a phenomenon called classification phenomenon. There is a transformation point for the evolution result depending on the original proportion of approval. If the original density of approval is above 0.5, the system result tend to be approval, vice versa.

The influence of mass media is to add an external factor to the transition probability $P(t)$. So the transition probability $P(t) = P_{ext}(t) + P_{int}(t)$. $P_{ext}(t)$ is external influence, which is decided by mass media and the influence changes over time. $P_{int}(t)$ is internal influence, which is decided by neibouring cellular(this fuzzy inference has come to this conclusion). Fig.16 is the external influece of the model.

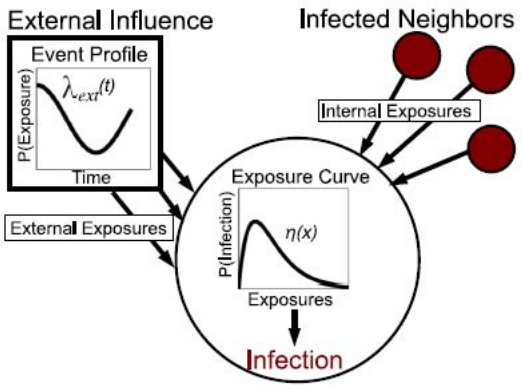

Figure 16: Information diffusion and external influence in networks[10]

3 Validity of Model

3.1 Task 2: Prediction for Situation In Contemporary

Apply model C to predict situation in contemporary. Based on the data of 1991-2006[11], we forcast the proportion of different media after several years. In that the lack of data on Internet, we ignore the prediction of it.

In Fig.17 solid lines represents the true data, dotted lines represents the data we predict using our model. From it we can see that our prediction has little deviation compared to true data today. We also calculate the relative deviation in 2012. The relative deviation of newspaper is 28.49%, the radio is 4.59%, and the TV is 6.54%.

The reason why the relative deviation of newspaper is larger than the others if that the market share of newspaper is small. From the result, our model can be proved valid. Additionally, this conclusion also correspond to the reality that people are more willing to get information from television compared to other media.

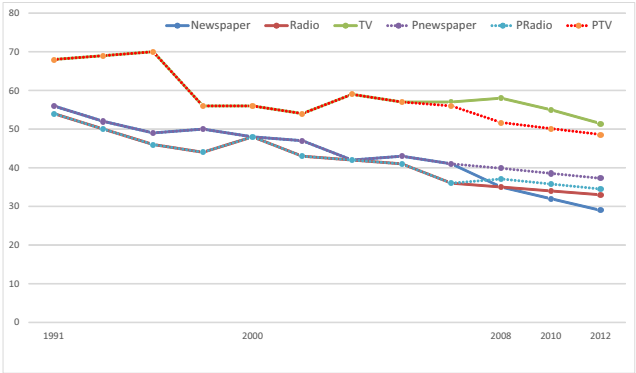

Figure 17: Prediction for situation in 2012

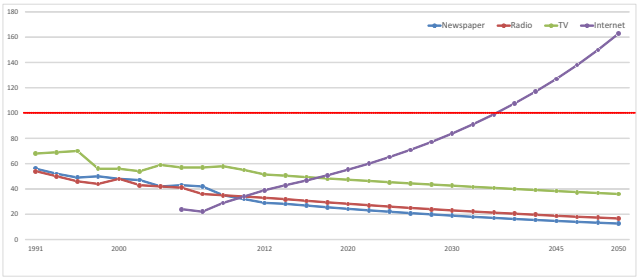

Figure 18: Prediction for situation in 2050

3.2 Task 3: Prediction for Situation In Future

We still use model C to predict the situation in 2050.From Fig.18 , its easy to learn that the usage of newspaper ,radio and TV ,which are traditional media, will decrease due to the impact of the internet. Nowadays , with the rapid development of information technology , new media has a much broader prospect than traditional media . We cannot live or work without the new media in this information era. However, the traditional media will not disappear ,just as shown on the figure traditional media still own a share. Besides, in our prediction curve, the usage of Internet will exceed 100% , the reasons might be: 1) One can own more than one Internet connector(e.g cell phone, personal computer, iPad, etc) 2) There might be a new medium in the future ,which may take share of Internet. But in our prediction , we ignore the possibility of the birth of new medium. This leads to the abnormal growth of Internet.

4 Influence of Information

4.1 Task 4: How Public Interest and Opinion Can Be Changed

4.1.1 Only Internal Factors Was Considered

Apply model D to analyze changes of people's opinions. Simulate the evolution based on model D. As Fig.19 shows, the white point stands for approval while the black one represents opposition. Under the circumstance that the original state of each cellular's state is random, the state of cellular change dramatically after a long period time of evolution. The state of neibouring cellular tend to have the same state.

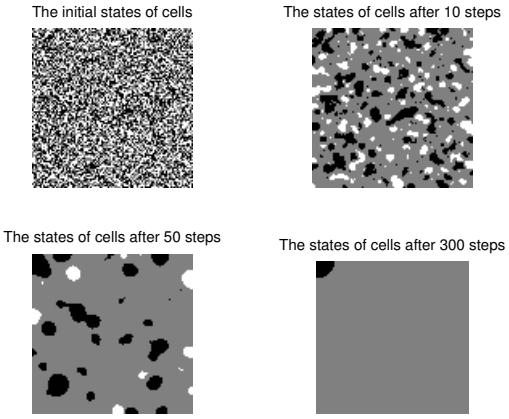

Figure 19: Simulation result of the cellular evolution

Fig.20 shows change of the proportion of each state during the evolution. We can come to a conclusion that at the beginning, the amount of people with three kinds of attitude is about the same. Because the original state is determined at random, so the evolution result is different after the same step time. But there exists a regular that after a long period of evolution, the amount of people who hold the neutral attitude increases, while the amount of other two kinds of people decreases and tend to be none. This phenomenon means people tend to get to conpromise in certain event.

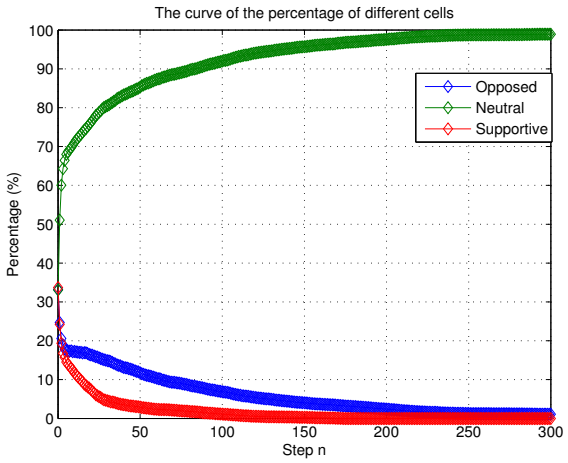

Figure 20: Change of the proportion of each state during the evolution

4.1.2 Influence From External Was Taken Into Consideration

When the factor of mass media is added to the model, do the simulation for another time. Fig.21 and Fig.22 show that when the original proportion of approval is above 0.5, the final state is stable, and almost of them become approval, only the rest little amount tend to be neutral. After a long time, people tend to believe what mass media lead.

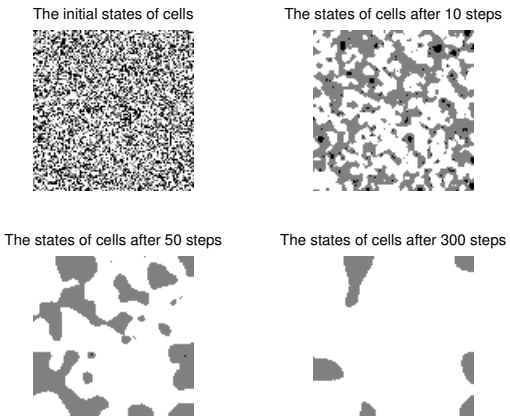

Figure 21: Simulation result of the cellular evolution(original proportion of approval is above 0.5)

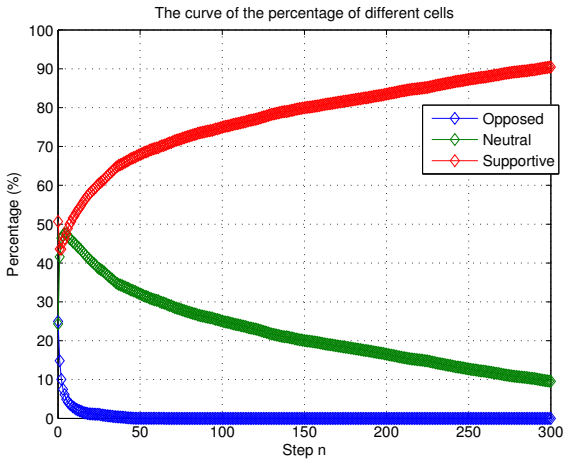

Figure 22: Change of the proportion of each state during the evolution(original proportion of approval is above 0.5)

4.2 Task 5: Other Factors That Influence Information Dissemination

4.2.1 Information Value

We have already discussed the influence of the internal value in model A.

4.2.2 People's Initial Opinion and Bias

We have discussed this influence in task 4 that it is influenced by internal and external factors.

4.2.3 Form of the Message or Its Source

During the research , weve found that information in the form of video can better be spread than picture, just as picture is more effective than information only in words. We also take the event that President Lincoln was assasinated for example. If this event took place in this era, we can easily conclude that the internal value of dissemination differs from different media ,by the fuzzy evaluation model we established above. So its necessary to analyze the influence of different media .Thus , we take secondary propagation theory into consideration . We introduce the multiplication model: Media Influence=number of audience(in 10,000) $\times$ average exposure time(hour per person) $\times$ credibility(%) $\times$ (1+rate of majority(%)).

Four indexes above are independent.From figure we know that in 2012 , TV has largest influence . However, according to the prediction model , in 2016, influence of internet will finally overcome that of TV.

4.2.4 The Topology or Strength of the Information Network in A Region, Country, or Worldwide

In our model A, the strength is measured by credit and transmission will, the bigger the mean number is, the faster the speed is.

In our model D, the strength is measured by k and h, and it also has positive correlation with the speed. So, the strength of the connection between nodes can affect how fast information spread. It is reflected in reality that best friends are more likely to share common beliefs and are more affected by each others opinions.

When considering medias influence in model D, we know that the more topology is, the more changing probability the cells have, which will accelerate the spread speed. And topology may play a more important role for information spread than the strength of the links.[12] If the scope is enlarged, the speed will decreases.

5 Sensitivity Analysis

1. **Model A:**

   Influence of different types of probability distribution of credit or will:

Table 3: Average speed of personal dissemination while dissemination will being different mean

Means of will to disseminate	Evenly distribution	Uniformed distribution	Binomial distribution
$\bar{b} = 0.57$	0.839	0.840	0.911
$\bar{b} = 0.53$	0.815	0.769	0.910
$\bar{b} = 0.49$	0.769	0.714	0.805
$\bar{b} = 0.45$	0.762	0.704	0.744

When the will of personal dissemination is uniform distributed, standard normal distributed and binominal distributed , the trend of average speed is consistent. But the average speed is faster when it comes to uniform and binominal distribution ,rather than standard normal distribution.

Table 4: Average speed of personal dissemination while credibility being different mean a

Means of trust	Evenly distribution	Uniformed distribution	Binomial distribution
$\bar{a} = 0.60$	1.336	1.373	1.493
$\bar{a} = 0.55$	1.184	1.242	1.402
$\bar{a} = 0.50$	1.101	1.142	1.363
$\bar{a} = 0.45$	0.933	0.971	1.199

When the credibility is uniform distributed, standard normal distributed and binominal distributed in interpersonal network , the trend of average speed is consistent. But the average speed is faster when it comes to standard normal and binominal distribution ,rather than uniform distribution.

The influence of mean value and variance has already been discussed.

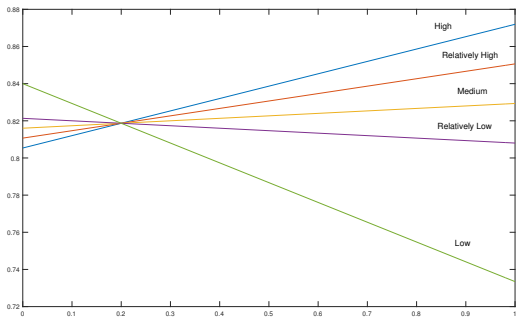

Figure 23: Sensitivity analysis of model B

2. **Model B:**

   During one given period(weight is constant), when we review one event, just like we discussed above, we can get one 5*5 review matrix. Assume that four of five elements are constant, which means theres only one variable(e.g. newness). Now we analyze five judgements towards the newness, from the figure, we find that high and low are more sensitive , which accords with our common cognition : the newer the event is , the more possible it will be reported.

3. **Model D:**

   We have already known the influence of preference H(I,j), now we are going to find out the influence of environmental adaptability K(I,j). When the average value of K increases by 10%, which stands for the increase of percentage of firm and neutral , the probability that crowd changes their opinions will decrease. This will increase the diffusion time towards a stable status.

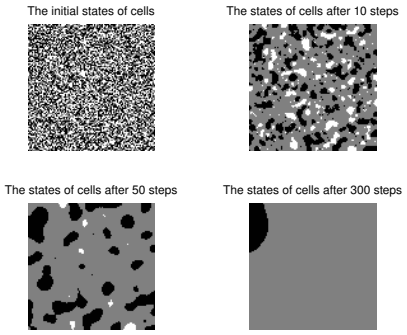

Figure 24: The increase of envoronmental adaptation

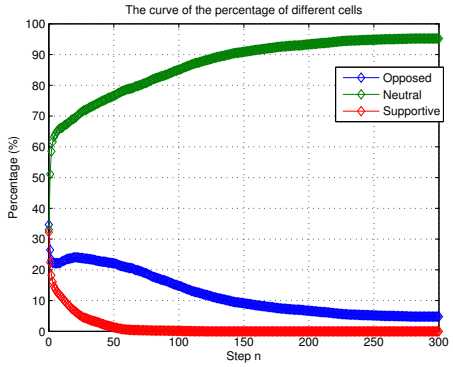

Figure 25: Change of cellular state with the increase of environmental adaptation

6 Error Analysis

1. **Model A:**

   According to model B, we can know that the inherent value of the information itself is not stable with time. So, if the valve change with time, we assume the value change as:$I(t+1) = I_0 \times e^{-\frac{I(t)}{t}}$, the result will change.

Table 5: Do 100 experiments to get the mean of T, shows as following, and record the step times.

Whether I is stable	Mean of step times	Variance of step times
NO	21.3	171.6429
YES	13.5	14.6633

Thus, if I is not stable, it will take more time to spread the information. But we just need to study the diffusion rate when the cells are in steady states. So the assumption of stable value of the information makes it easily for us to study other properties.

2. **Model C:**

Table 6: The error analysis of moder C

	newspaper	radio	TV
Predictive value	37.26	34.51	48.60
Actual value	29.00	33.00	52.00
The relative error(%)	28.49	4.59	6.54

For radio and TV, the relative errors are small, so we can accept the prediction. But when it comes to newspaper, it is a little out of reality.

3. **Model D:**

   Model D is similar to Model A, so we can get the same analysis.

7 Conclusion

We've built four models to measure the evolution and influence in social information networks.

With model A, we can simulate the process of information delivery. At the same time, we find a critical inherent value of information as 0.6 (out of 1), which means, only if the value of information itself over 0.6, can it be diffused.

Thus, we use model B to evaluate the value of information, and make the definition that when its value is over 0.6, it will qualify as news.

With model C, we can make some predictions and a comparison between the predicted data and the existing one. Moreover, we've brought in relative error to prove the correctness of our model, and have predicted the situation in 2050 based on our model: by 2050, the usage rates of the several primary media will respectively reach: newspapers 12.65%, radio 16.65%, TV 36%, Internet $> 100\%$.

Finally we create model D to reflect peoples attitudes towards information. In this way, we could reflect how public interest and opinion can be changed through information networks in todays connected world.

References

[1] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[2] S. Wolfram, "Cellular automata as models of complexity," *Nature*, vol. 311, no. 5985, pp. 419–424, 1984.

[3] N. Ganguly, B. K. Sikdar, A. Deutsch, G. Canright, and P. P. Chaudhuri, "A survey on cellular automata," 2003.

[4] P. Flocchini, F. Geurts, A. Mingarelli, and N. Santoro, "Convergence and aperiodicity in fuzzy cellular automata: revisiting rule 90," *Physica D: Nonlinear Phenomena*, vol. 142, no. 1, pp. 20–28, 2000.

[5] "News values," http://faculty.georgetown.edu/bassr/511/projects/letham/final/values.htm.

[6] T. L. Saaty, "Basic theory of the analytic hierarchy process: how to make a decision," *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales*, vol. 93, no. 4, pp. 395–423, 1999.

[7] Y.-P. Huang and C.-H. Huang, "A genetic-based fuzzy grey prediction model," in *Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century., IEEE International Conference on*, vol. 2. IEEE, 1995, pp. 1051–1056.

[8] M. Pradipta and P. P. Chaudhuri, "Fuzzy cellular automata for modeling pattern classifier," *IEICE Transactions on Information and Systems*, vol. 88, no. 4, pp. 691–702, 2005.

[9] Y. Chai, L. Jia, and Z. Zhang, "Mamdani model based adaptive neural fuzzy inference system and its application in traffic level of service evaluation," in *Fuzzy Systems and Knowledge Discovery, 2009. FSKD'09. Sixth International Conference on*, vol. 4. IEEE, 2009, pp. 555–559.

[10] S. A. Myers, C. Zhu, and J. Leskovec, "Information diffusion and external influence in networks," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 2012, pp. 33–41.

[11] "Pewresearchcenter," http://www.people-press.org/2012/09/27/section-1-watching-reading-and-listening-to-the-news-3/.

[12] L. Severini, M. El Chamie, and G. Neglia, "Topology versus link strength for information dissemination in networks," in *AlgoTel-15èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications*, 2013, pp. 1–4.