Exploration for opinion evolutionary mechanism on Internet Live Broadcast Platforms

Bin Hu, Yutao Chen,

Zeyu Zuo, Chang Li, Shengnan Zhang

School of management, Huazhong University of Science &Technology, Wuhan, China

Abstract:
This study attempts to utilize computer simulation to study the evolution of opinions in online live broadcast platforms. Based on the SJBO model, we model the opinion dynamics and add psychological factors like Negativity Bias & Positivity Offset into the model. The opinion expression mechanism of agents is also considered and designed in this study. Moreover, we plan to validate the simulation model with the sentiment scores of opinions from real Internet live broadcast platforms, to make it an effective reference for further management of these platforms.

Keywords: Opinion Dynamics, Negativity Bias & Positivity Offset, Sentiment Analysis, Multi-agent Simulation

(Notification: This is a working paper for the project *Computational Organization Theory-based Research on Evolution Mechanism of Public Opinions on Internet Live Broadcast Platforms*, not an official one.**)**

1.Introduction

With the development of social networks sites (SNS), people have more extensive and convenient ways to express and share their opinions. In traditional online social media such as Facebook, Twitter and Microblog, users interact with each other by leaving comments on these platforms. Large quantities of online user generated contents (UGC), especially those comments left by users, provide researchers with plentiful resources to study the opinions of users in online communities.

Recently, a new kind of online social networking platform, live broadcast platform, has become popular in Asian countries such as China, South Korea and Japan. Users can watch live broadcast content by entering a broadcast room in these platforms. The biggest difference between these platforms and traditional video websites is that in the process of watching live broadcast, one or more hosts interact with users as if talking to them (i.e., the audiences can watch the hosts while the hosts cannot see the audiences), and the content of live broadcast is usually determined and generated by the hosts. Taking China's online live broadcasting platforms as an example, the content of live broadcasting mainly includes e-sports commentaries, entertainment performances and outdoor activities. Moreover, the form of user comments on the platforms is also very special: users' comments will be in the form of Danmu (screen bullets) scrolling across the broadcasting room (see Figure 1), that is to say, viewers will see all comments sent at the same time scrolling on the screen (not to consider the problem of network delay and blocking of certain users). In this process, hosts can also see the screen bullets left by the audiences in the background and respond to them. Besides leaving comments, audiences can also interact with the hosts by sending virtual gifts to them. Virtual gift giving, viewed as a sort of consumption behavior, requires users to pay a certain fee to buy gifts, and usually is not displayed in the bullet screens unless the gift is expensive enough.

Figure 1. Snapshot for Online Live Broadcast

Since the mode of online live broadcast is relatively new, there are only few researches conducted on it among the literature we have consulted. In these limited resources, the core issues focus on the impact of bullet screens on people's experience when watching live broadcasts (e.g., Djamasbi et al, 2016 and Jia et al, 2018), and a few studied on

natural language processing (NLP) of bullet screens (e.g., Xu et al).
In this study, we will focus on exploring the evolution mechanism of opinions in live broadcast platforms, and identify the corresponding factors in different conditions that contribute significantly to the results of the opinion evolution (e.g. consensus, bipolarization, or dispersion), to shed light on potential approaches for management of online live broadcast platforms. Firstly, in Section 2, we review the relevant literature on the study of dynamics models for opinion evolution and the social psychological theories applied to our research. In Section 3, we elaborate on the opinion dynamic model we constructed in this study for the online broadcast platforms. Different from previous studies, we consider people's difference in awareness of positive and negative opinions in the modeling process, as well as the conditions when people would express their opinions in the live broadcast platform environment (unexpressed opinions would not affect other audiences). Then, in Section 4, we have collected the time series data of bullet screens from the live broadcast platforms, and plan to obtain the evolutionary trend of the sentiment value in those bullet screens by means of sentiment analysis, so as to test and modify the above opinion dynamic model.

2. Literature Review

2.1 Dynamic Models for Opinion Evolution

The study of the evolution of opinions in communities has always been an important and interesting topic in the field of dynamics. Researchers built different models of opinion evolution based on different backgrounds and theoretical bases. Among many different models, we highlight the NSCRL model developed from the epidemic model, discrete opinion dynamic models based on finite opinion selection set as well as continuous opinion dynamic models based on the concept of bounded confidence.

In the field of propagation dynamics, Daley and Kendall introduced the traditional dynamic model of epidemic (Kermack and McKendrick, 1927) into the study of rumor propagation in 1964 and translated the concepts of S (susceptible), I (infectious), and R (recovered) into the context of rumor propagation (Daley and Kendall, 1964). On this basis, more complex models (SIRS, SEIR, SEIRS) can be formed through further subdivision of population states and transition mechanism. Schramm and Harrison C extended this ideology in rumor propagation model, and proposed the NSCRL model for studying public attitudes evolution towards government proposals (Schramm and Harrison C, 2006). In their work, the crowd was divided into five types:

- N: *Neutrals* who hold no firm ideology or opinion and who may be susceptible to ideological influence;
- S: *Active Supporters* who actively spread pro-government ideas;
- R: *Latent Supporters* who support the government but do not actively spread their ideas;
- C: *Active Contrarians* who actively spread anti-government ideas;
- L: *Latent Contrarians* who oppose the government but do not actively spread their ideas.

Based on the above definitions, the authors established different but relevant dynamic models such as NSRCL model, NSRL model, NSC recirculation model and stochastic model to adapt to different propagation scenarios. In these models, the evolution of viewpoints over time is realized through the state transitions between people in different states (see Figure 2), which is usually expressed as ordinary differential equations (ODEs) of proportion of people in different states for time as below: (β_i denotes the transition coefficient)

$$\frac{dN}{dt} = -\beta_1 NS - \beta_2 NC$$
$$\frac{dS}{dt} = \beta_1 NS - \beta_3 S^2 - \beta_4 SR + \beta_5 SC$$
$$\frac{dR}{dt} = \beta_3 S^2 + \beta_4 SR$$
$$\frac{dC}{dt} = \beta_2 NC - \beta_5 SC - \beta_6 C^2 - \beta_7 CL$$
$$\frac{dL}{dt} = \beta_6 C^2 + \beta_7 CL$$

State	**N**	**S**	**R**	**C**	**L**
N		*N->S*		*N->C*	
S	*N->S*	*S->R, S->R*	*S->R*	*S->C, C->S*	
R		*S->R*			
C	*N->C*	*S->C, C->S*		*C->L, C->L*	*C->L*
L				*C->L*	

Figure 2. NSRCL Transition Matrix

NSRCL, with its derivative models gives a systematic interpretation of the evolution of viewpoints, which is suitable for measuring the distribution of binary opinions (e.g., yes or no, for or against) of a specific idea or subject among groups. However, the application of this method will be limited in the scenarios where the subjects of public opinions are relatively scattered and views are non-binary. Additionally, the epidemic based model also can hardly provide insights on how inner features and activities of individuals affect their opinion formations and transitions.

The following discrete and continuous opinion dynamics models attempts to study the evolution of opinions in communities from the perspective of individuals rather than the whole system. In the discrete dynamic model, each agent selects its own opinion at each moment from a set of two or more finite opinions (e.g., yes or no; big, small or medium). In different discrete models, agents choose their own opinions with different strategies. For example, in the Voter model (Holley and Liggett, 1975), which view the whole evolution as a stochastic process, an agent (voter) shifts its opinion at each moment based on whether specific opinions exist in some pre-defined subsets of all agents (e.g. the nearest neighbors of this agent) at the last moment as well as a pre-defined probability distribution on the weight of these subsets. In another interesting discrete model, the Sznajd model, the evolution of opinions among agents relies on the consistency in pair-wise neighboring agents. K. Sznajd-Weron and J. Sznajd proposed this model in 2000, stating that, based on social validation phenomena, when an adjacent pair of agents has the same opinion, they will spread that opinion to another neighbor agent, respectively. Though the discrete models are able to simulate the evolution of opinions from the aspect of individuals, they are still confined to the scenarios where the opinions of agents belong to a finite set.

Different from the two types of models mentioned above, in the continuous opinion dynamics models, the opinion of every agent at each moment can be any real value within a pre-defined range. The early continuous models characterized the changes of opinions of agents as linear formats, such as in classic model (De Groot, 1974; Lehrer, 1975). The updated opinion of an agent is formed by considering all the opinions in the community with different weights assigned. Friedkin and Johnsen modified this model in 1990 to involve the adherence to one's initial opinion while maintain its linearity. In these continuous models, the opinion transition mode of the whole group can be expressed as a transition matrix. When the matrix is fixed or irrelevant to the opinions themselves, which keeps the model linear, the evolution results are likely to be obtained through mathematical analysis (Hegselmann and Krause, 2002; JAVIER GÓMEZ-SERRANO et al, 2012).

The Bounded Confidence (BC) model (Krause, 1997 and 2000; Deffuant, 2000) breaks down the constraint of linear form by associating the opinion transition function with the opinion values themselves. In this model, agents' opinions influence each other only when the difference of their views is below a certain threshold. This threshold, that is, bounded confidence, can be interpreted as the uncertainty of the agent's opinion. When the difference from another opinion exceeds the scope of such uncertainty, the agent will not consider incorporating such view. The general form of the BC model can be expressed as follows:

$$x_i^{t+1} = x_i^t + f\left(x_i^t, x_j^t\right)\left(x_j^t - x_i^t\right)$$

$$f\left(x_i^t, x_j^t\right) = \begin{cases} \mu, & \left|x_j^t - x_i^t\right| \leq u \\ 0, & \left|x_j^t - x_i^t\right| > u \end{cases}$$

Where x_i^t represents the opinion value of agent *i* at time *t*, u represents the threshold (uncertainty) for adopting the opinion from others and μ denotes the degree of adoption. Noting that function of $f\left(x_i^t, x_j^t\right)$ plays a decisive role in the opinion update, we regard it as the kernel function of the BC model.

After the BC model was proposed, due to its flexibility and comprehensibility, it was widely used in the study of opinion dynamics. Researchers adjust the basic model according to different backgrounds and problems, mostly in the point of kernel function. We emphasize three of these adaptations: The Gaussian Bounded Confidence model (GBC) (Deffuant, 2006) change the kernel function of the opinion transition into a Gaussian form and considering dynamic changes of uncertainties (i.e. thresholds) for agents; The Relative Agreement (RA) model (Guillaume et al., 2002) introduces a new assumption that individuals with lower uncertainty are more influential (also harder to be influenced) than individuals with higher uncertainty during opinion interaction, by defining the relative agreement as a ratio between a segment of the opinion overlap from the interactors and their individual uncertainty, and replacing the original kernel function in BC model with one containing the expression of relative agreement; The Social Judgment Based Opinion (SJBO) dynamics model (Jager and Amblard, 2005) differs from the basic BC model by taking in the idea of social judgement theory (Sherif et al, 1961 and 1965), which states that when the opinions of two agents are close enough, their opinions tend to assimilate with each other; when the gap between two opinions is large enough, they tend to repulse to each other; otherwise, the two agents will not influence the opinion of each other. The introduction of repulsion effect, or Boomerang effect (Hovland et al.,1953 and 1957), makes this model more true-to-life than the original BC model and more consistent with people's behaviors in reality. Therefore, we chose the SJBO model as the framework in this study for our modeling and the general mathematical form is shown below:

$$x_i^{t+1} = x_i^t + f\left(x_i^t, x_j^t\right)\left(x_j^t - x_i^t\right)$$

$$f\left(x_i^t, x_j^t\right) = \begin{cases} \alpha, & \left|x_j^t - x_i^t\right| < d_1 \\ 0, & d_1 \le \left|x_j^t - x_i^t\right| \le d_2 \\ -\beta, & \left|x_j^t - x_i^t\right| > d_2 \end{cases}$$

Where x_i^t represents the opinion value of agent i at time t, and d_1 represents the threshold for assimilation while d_2 is that for repulsion. α and β refer to the degree of assimilation and repulsion, respectively.

Based on the SJBO model, researchers have studied the evolution of public opinion at different levels. Wong et al. (2014) made corrections to the original assumptions and identify how factors like the assimilation and repulsion threshold affect the opinion clustering; Fan and Pedrycz (2016) combined the SJBO model with a simplified discrete model to simulate the inner continuous opinion variation and the observable discrete choices of agents and considered the heterogeneity from informed and regular agents; Fan and Pedrycz (2017) also studied on the effect of network topologies to opinion evolutions by applying SJBO to both a broadcast (one-to-many) network and a BA scale-free social network (many-to-many); Wang and Fu (2016) applies SJBO to the analysis of media effect and antagonistic interactions in online-offline networks. However, these studies mostly rely on dynamics modeling and simulation analysis of opinion evolutions, and the validity of those models are not supported by data from real social network. Therefore, in our study, we work to build a simulation model consistent with the opinion evolution in realistic live broadcast platforms, using sample data from online live broadcast platforms for validation.

2.2 Negativity Bias and Positivity Offset

The concept of "Negativity Bias and Positivity Offset", which we integrated into the opinion dynamics model, reveals that people tend to respond differently according to whether the stimuli elicit a positive or a negative cognitive reaction. The idea of negative bias has been elaborated in the review of Rozin and Royzman (2001). They suggested that Negativity Bias is manifested in four aspects:

- Negativity Potency: Negative entities are stronger than the equivalent positive entities
- Steeper Negative Gradients: The negativity of negative events grows more rapidly with approach to them in space or time than does the positivity of positive events
- Negativity Dominance: Combinations of negative and positive entities yield evaluations that are more negative than the algebraic sum of individual subjective valences would predict
- Negative Differentiation: Negative entities are more varied, yield more complex conceptual representations, and engage a wider response repertoire.

We hereof highlight the second of these manifestations, which we choose to incorporate in this research. The measurement of the steeper negative gradients is originally proposed in Cacioppo's works in 1994 and 1997, where the researchers utilize a bivariate plane to quantitatively analyze the relationship between individuals' attitudes and their evaluation of positive and negative stimuli. The bivalent models of evaluative activation can be expressed as follow:

$$Attitude = \frac{W_p}{W_p + W_n} P_i - \frac{W_n}{W_p + W_n} N_j + I_{ij}$$

Where P_i (N_j) represents the level of positivity (negativity) activated by an attitude object; W_p (W_n) represents the relative attitudinal effect of variations in positivity (negativity); I_{ij} represents nonadditive effects. On the basis of the approach-avoidance experiment conducted by miller et al. (1959), Cacioppo deduced that the negative weight in the above model should be greater than the positive weight (Negativity bias), and proposed that when the stimulus intensity is low, the positive evaluation process (Positivity Offset) is stronger, which is expressed by adding a higher intercept for positive evaluation. The modified version of this computational model is presented below with a visualized example result (see Figure 3):

$$Attitude = \left(W_p P_i + c\right) - W_n N_j + I_{ij}$$

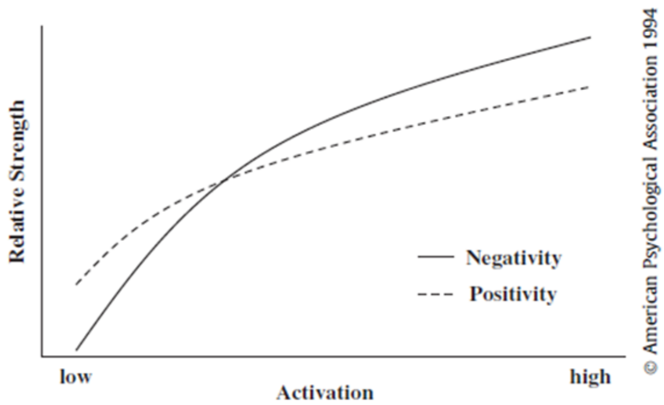

Figure 3. Activation function example result (1994)

The ideology of negative bias and positive offset is supported by many psychological and neuroscience findings as well (e.g., Smith et al., 2003; Carretié et al., 2001). Most of these studies conducted event-related brain potential (ERP) experiments to measure the amplitude of participants' response to positive or negative stimuli, so as to verify the existence of negative bias and positive offset in human reactions. In addition, some researchers (e.g., Ito and Cacioppo, 2005; Norris et al, 2011) have conducted a deeper extension of these cognitive biases: looking for individual differences existing in these cognitive biases due to different personalities. In view of the universality of both theoretical and experimental studies on negativity bias and positivity offset, and their desirable foundation in quantitative analysis, as well as their strong relevance with the evolution of opinions in online live broadcast platforms since the formation and transition of people's opinions in the live broadcast networks are influenced by a large number of positive and negative views, it is reasonable to include this psychological theory into the modeling process.

2.3 Opinion Expression Mechanism

In the online live broadcast platforms, the audiences will not express their opinions in Danmu all the time, so in the construction of the dynamic model, we need to consider under what conditions audiences will send screen bullets on the platform. In this study, our consideration for the mechanism of people's opinion expression is related to a well-known communication theory -- the spiral of silence (Noelle-Neumann, 1974, 1991). This theory generally stipulates that people are only willing to express their opinions when they are in line with the dominant opinion of the public. Otherwise, they will choose to remain silent due to the fear of being isolated or excluded. This inference about the spiral of silence rests on three core assumptions:

- People are in fear of isolation.
- People can assess the dominant opinions in their environment.
- The evaluation of the opinion climate can influence people's expression of their opinions.

In the online community, however, the validity of these assumptions, especially the first assumption, has been questioned through later studies: Yun et al (2016) showed that people with minority perceptions will not keep silence in fear of isolation but will increase their willingness for speaking because of the feeling of hostility; Chun and Lee (2017) proposed that mediation factors such as the sense of social support and control also play an important role in whether or not people express opinions, not just the fear of being isolated; Liu and Rui (2017) examined the role of self-presentational concern in the spiral of silence under online settings like Facebook.

In general, expression of opinions in the online social networks do not strictly follow the traditional theory of the spiral of silence, and many factors that have been studied or not been studied may play roles in it. In this study, we defined the hypothesis of fear of isolation in the traditional theory as a state of complete conformity, and individuals can be in a state between complete conformity and complete non-conformity, which suggests that factors other than fear of isolation might make an individual less conform. Individuals with different degrees of conformity have different disciplines of speaking after identifying the dominant opinion. The specific opinion expression mechanism design will be elaborated in Section 3.

3. The Model

In this section, we are going to illustrate the SJBO model-based opinion evolutionary dynamics model of live broadcasting platform and its setup philosophy. We firstly notify the assumptions and notations for our model, and then elaborate on how we extended the traditional SJBO model to accustom to the online live broadcast platforms and also how we incorporated the cognitive or communication theories like negativity bias, positivity offset and the modified version of the spiral of silence into our model.

3.1 Assumptions and Annotations

Assumptions:

- The live broadcast platform is assumed to be closed or in a dynamic equilibrium during the simulated evolution process, which means we preclude the effect of the variance in the scale of the social network.
- The networks among the audiences (see in Figure 4) and between the host (see in Figure 5) and audiences are considered to be fully connected, which means that we do not consider the policy to shield the messages sent by certain agents.
- The impact from the gift sending messages is precluded in this study, since the messages of gift sending are displayed in screen bullets after filtration on their expense and sometimes irrelevant to the current broadcast room. (In addition, we believe that the analysis of consumption behaviors like virtual gift sending requires different models from the analysis of pure opinion evolutions.)

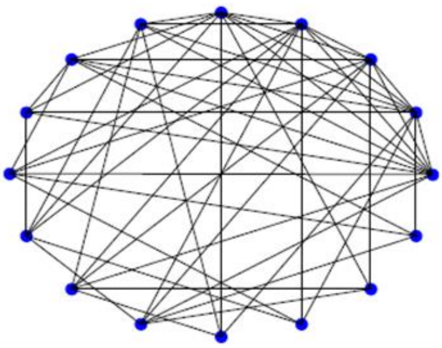

Figure 4. Audience-Audience Network

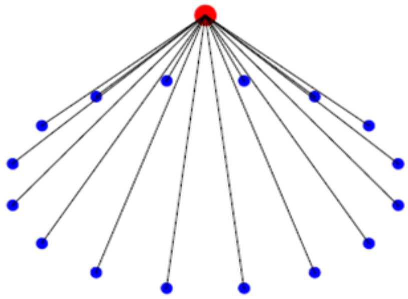

Figure 5. Host-Audience Network

Index	Definition	Range
x_i^t	*opinion value of agent i at time t*	$[-1,1]$
$f\left(x_i^t, x_j^t\right)$	*kernel function of opinion impact degree*	
α	*assimilation coefficient*	$[0,1]$
β	*repulsion coefficient*	$[0,1]$
d_1	*assimilation threshold*	$[0,1]$
d_2	*repulsion threshold*	$[1,2]$
X_i^t	*agent i 's integrated opinion value from all other agents expressing opinions at time t*	
ω_{ij}^t	*weight of opinion from agent j in the integrated opinion of agent i*	
S^t	*the set of agents expressing opinions at time t*	
p	*probability for agents to express when in majority*	$[0,1]$
τ_i^t	*opinion expression state for agent i at time t*	
x_H^t	*opinion/information value of host at time t*	$[-1,1]$
X_H^t	*host's integrated opinion value from all other agents expressing opinions at time t*	
m	*stabilization coefficient in opinion impact degree*	$[0,1]$
n	*stabilization coefficient in threshold*	$[1,2]$
b	*weight of opinion from host in the integrated opinion of agent i*	$[0.5,1]$
k_P	*the slope for positive opinion perception*	$(0,1]$
c	*the intercept for positive opinion perception*	$[0,1)$
k_N	*the slope for negative opinion perception*	$(0,1]$
θ	*the exponent for marginal diminishing effect*	$(0,1]$
$\widetilde{x_i^t}$	*perceived opinion of agent i by others at time t*	$[0,1]$

Figure 6. Annotation Table

3.2 Modeling Opinion Evolution

The original SJBO model was limited to the interaction between two agents, so we firstly extended the opinion dynamics model of SJBO from a dual scenario to a multiple-agent one by setting the intermediate variable X_i^t, suggesting each agent can be influenced by all revealed opinions as well as the information from the host. In addition, we modified the kernel function by binding a constraint factor (see Appendix 1) so that the opinion value will not exceed the predetermined range after every change. The extended model is now in the following form:

$$x_i^{t+1} = x_i^t + f(x_i^t, X_i^t)(X_i^t - x_i^t)$$

Where

$$f(x_i^t, X_i^t) = \begin{cases} \frac{\alpha(1 - |x_i^t|)}{2}, & |x_i^t - X_i^t| < d_1 \\ 0, & d_1 \le |x_i^t - X_i^t| \le d_2 \\ -\frac{\beta(1 - |x_i^t|)}{2}, & |x_i^t - X_i^t| > d_2 \end{cases}$$

Where X_i^t represents an integrated opinion from all agents other than agent *i*, including the host:

$$X_i^t = \sum_{j \in S^t \cap j \neq i} \omega_{ij}^t x_j^t$$

ω_{ij}^t represents the weight of the opinion of agent *j* in the integrated opinion of agent *i*. When there is no differentiation existing in evaluating all the expressed opinions, the weight is identical for each opinion, which means:

$$\omega_{ij}^t = \frac{1}{|S^t| - 1}$$

Next, we considered the influence posed by the host. Each audience in the broadcast channel will be affected by both the other audiences and the host. Usually, the impact from the host outweighs those from other audiences, and the host is always more difficult to be affected by the audiences' opinions because they have pre-planned interactive goals in broadcasting. With these distinct traits, we set the host to be a special agent in the whole online broadcast network, denoted by *H*. For *H*, we can obtain the following transition function for the host:

$$x_H^{t+1} = x_H^t + f(x_H^t, X_H^t)(X_H^t - x_H^t)$$

Where

$$f(x_H^t, X_H^t) = \begin{cases} \dfrac{m\alpha(1-|x_H^t|)}{2}, & |x_H^t - X_H^t| < \dfrac{d_1}{n} \\ 0, & \dfrac{d_1}{n} \leq |x_H^t - X_H^t| \leq nd_2 \\ -\dfrac{m\beta(1-|x_H^t|)}{2}, & |x_H^t - X_H^t| > nd_2 \end{cases}$$

m and *n* are stabilization coefficients that represent the difference in the degree and thresholds of being impacted between the host and ordinary audiences. Correspondingly, the ω_{ij}^t will be different from the original unbiased state. ω_{iH}^t needs to be larger than the weight of any other audience's opinion. So, we let ω_{iH}^t to be a constant over 0.5 under any circumstance. The new function for the ω_{ij}^t containing the bias to pay more attention to the host's message is shown below:

$$\omega_{ij}^t = \begin{cases} \dfrac{1-b}{|S^t|-2} & i \neq H, j \neq H \\ b & i \neq H, j = H \\ \dfrac{1}{|S^t|-1} & i = H \end{cases}$$

In the following part, we were going to incorporate the cognitive bias of positivity offset and negativity bias into the analysis of our model. Positivity offset refers to an additional impact posed by a positive stimulus which presents an advantage when the intensity is relatively slight, while negativity bias refers to a higher impact posed by a negative stimulus as the stimulus becomes more intense. These two effects can be expressed as an additional intercept for the positivity activation function and a larger slope for the negativity activation function in the mathematic form:

$$\widetilde{x_j^t} = \begin{cases} k_P x_j^t + c & x_j^t > 0 \\ k_N x_j^t & x_j^t \leq 0 \end{cases}$$

Where $\widetilde{x_j^t}$ represents a single perceived opinion of agent *j* by others at time *t,* after the cognitive distortion. In line with this, the process for the formation of X_i^t can be modified as below:

$$X_i^t = \sum_{j \in S^t \cap j \neq i} \omega_{ij}^t \widetilde{x_j^t}$$

To ensure the value of the $\widetilde{x_j^t}$ does not cross the border, we set the following constraints:

$$0 \leq c \leq k_N - k_P < 1 - k_P$$

We also noticed that people's perceptions of positive and negative things are not necessarily in a linear relationship towards the intensity of stimuli. Like what has been expressed in the prospect theory, the relationship could be in a marginal decreasing form. To hold this condition, we could modify part of the perceived opinion function as below:

$$\widetilde{x_j^t} = \begin{cases} k_P\left(x_j^t\right)^\theta + c & x_j^t > 0 \\ -k_N\left(-x_j^t\right)^\theta & x_j^t \le 0 \end{cases}$$

Where $\theta \in (0,1)$.

In the following part, we considered the individual expression mechanism in the live broadcasting platforms. The audiences will not express their opinions all the time, but the opinions they convey usually reflect the exact inner ideas. We set up the mechanism to distinguish whether an audience will express his or her opinion based on the audience's judgment of the dominant opinion at a certain moment and the audience's degree of conformity (i.e. fear of isolation).

When audience agents believe that their opinions are qualitatively consistent with the dominant opinion (i.e. same in positivity or negativity), they will judge themselves as belonging to the majority, otherwise they will consider themselves as belonging to the minority. In the judgment of dominant opinion, audiences will not actually calculate the number of positive and negative opinions to judge, but use the quasi-statistical sense: The dominant opinion is determined by the integration of all others' opinions, which is notified as X_i^t in our model. This process can be expressed as follows:

$$x_i^{t+1} X_i^t \ge 0, i - majority$$
$$x_i^{t+1} X_i^t < 0, i - minority$$

Noticing that when agents have formed their integrated opinion of others at time *t*, their opinions have actually transited to the state at time *t*+1. But since other audiences' expression states at time *t*+1 have not been formed, they can only rely on their integrated opinion at time *t*, not *t*+1, to make judgement. After the judgement, the expression pattern of the audiences according to their perceived majority or minority is demonstrated as follow:

$$\tau_i^{t+1} \sim \begin{cases} 1 & i = H \\ B(1, p) & i \ne H, i - majority \\ B(1, 1-p) & i \ne H, i - minority \end{cases}$$

Where τ_i^t is a random variable subject to a zero-one distribution with probability *p*, whose value depends on the degree of conformity of the agents as the following figure suggests:

(p value)	*Conformity*	……	*Neutrality*	……	*Inconformity*
Majority	1	……	0.5	……	0
Minority	0	……	0.5	……	1

Figure 7. Expression Probability Chart

In addition, the design of opinion expression mechanism also determines the evolution process of set *S*, which will be updated at every moment with the change of agents who express their opinions. The definition of set S could be interpreted mathematically as below:

$$S^t = \{i | \tau_i^t = 1\}$$

4. Model Validation (in progress)

In this section, we are going to test the validity of the model we proposed with the opinion data collected from Douyu, an internet live broadcast platform in China. We would firstly introduce the data types and information obtained from the platform, then elaborate on the processing method and analysis of the bullet screen data, and finally briefly explain the idea of model validity test that we might adopt.

4.1 Data Description

The data we collected and stored in SQL Server comes from *Douyu TV*, a leading live broadcasting enterprise in China. Each sample contains all messages generated in the process of live broadcasting in a live broadcasting room during a certain period of time. For each bullet screen data item in the sample, it includes information like the time when the bullet screen was sent, the nickname of the audience who sent it, the bullet screen content, the ID of the broadcast room and so on (see Figure 8). We have currently collected over 150000 messages from 6 different rooms.

Index	**Data-Type**	**Example/Range**
MessageID	Ordinal number	123
Time	Datetime	2018/10/12 19:48
Type	Category	Danmu or Gift
Nickname	Text	“小 v 爷 7788”
Content	Text and Emoji	“大气大气带起”
RoomID	Index	5706218

Figure 8. Metadata of Bullet Screen

4.2 Sentiment Analysis

In the validation of our model, the acquirement of sentiments score of the bullet screen data is the premise of the subsequent analysis. The sentiment score of a specific Danmu comes from the sentiment analysis, an active field in natural language processing (NLP), of the content of this Danmu. The result of sentiment analysis is usually expressed as the probability of an opinion to be positive (negative), and the final sentiment score is usually presented as the mean of the positivity (+1) and negativity (-1) of this Danmu. The calculation of a sentiment score is presented below:

$$\begin{aligned} Sentiment\ Score &= (+1) * P + (-1) * N \\ &= P - N \\ &= P + (1 - P) \\ &= 1 - 2P \end{aligned}$$

Where *P* represents the probability for the opinion to be positive, and *N* represents the probability for it to be negative. The results of sentiment scores of Danmu are viewed as equivalent to the opinion values (i.e. x_i^t) in the simulation model, which set stage for further comparison.

According to previous literature and practices, there are multiple ways to conduct sentiment analysis like keyword recognition, lexical association, and machine learning methods. Hereof, we highlight two of them that we are currently putting into practice in our research:

- Sentiment analysis based on Semi-supervised learning with API of *BosonNLP*. Boson conducted sentiment analysis on Chinese text based on semi-supervised learning models derived from millions of social network balanced corpus and thousands of news balanced corpus. Due to business confidentiality, we are not able to learn more details about the model, but we have learned from the authoritative statistics that the accuracy of its sentiment analysis is about 80% to 85%, which can be improved to 85% or 90% after industrial data labeling.
- Sentiment analysis based on Naïve Bayes model (see Appendix 2) defined by *snowNLP* as well as customized lexicons. SnowNLP is an open source package for Chinese natural language processing in python. The advantage of this package is that it allows users to customize input data as their training sets. Therefore, we added some Chinese sentiment lexicons in line with the context of online social media as a supplement to the original training set of *snowNLP*, which contribute to the improvement of accuracy in sentiment analysis.

Among the bullet screens, apart from general text messages, there is a special kind of data – emojis (see Figure 9) for us to deal with. For the sentiment analysis of emoticons, we refer to the work of Novak et al (2015) to gain emoji's sentiment scores through empirical study. Participants will be asked to fill out a questionnaire containing all the emojis in the Douyu platform, and they will make judgments between 1, 0, and -1 for each emoji, where 1 represents positive feelings, 0 stands for neutral or unclear, and -1 means negative feelings. Then, the final sentiment score for an emoji will be obtained with similar methodology to that of usual text. The only difference is that values for *P* and N are from statistical ratios, rather than probabilities predicted by machine learning methods.

index	content	picture
dy101	666	
dy102	发呆	
dy103	拜拜	
dy104	晕	

Figure 9. Metadata of Emoji

5. Further Work

At present, the project is in the phase of collecting the data for Emoji analysis. When the data acquisition is completed, we will conduct comparison analysis between the sentiment - time fluctuation curve derived from screen bullets and the opinion evolutionary curve of the simulation model. Then, by using the ideas of the mean value of opinions among agents at different times, the clustering of opinions in the evolution process and other statistical indicators, we can measure the correlation between the simulation model and the real scenario to complete the validation.

When the model validation is done, we will start the sensitivity analysis of parameters, such as the slopes and the intercept corresponding to negative bias and positive offset, expression probability in expression mechanism and so on. Through sensitivity analysis, we can identify the factors that play decisive roles in the process and results of opinion evolution. (Some test simulations are presented in Appendix 3)

After the identification, we will try to design some strategies, such as shielding the agents who are actively sending negative opinions, to manage the evolutionary trend of public opinion in the online broadcast platforms.

Selected Reference

Djamasbi S , Hall-Phillips A , Liu Z , et al. Social Viewing, Bullet Screen, and User Experience: A First Look[C]// Hawaii International Conference on System Sciences. IEEE, 2016.

Yang X , Binglu W , Junjie H , et al. Natural language processing in “Bullet Screen” application[C]// International Conference on Service Systems & Service Management. IEEE, 2017.

Jia A L, Shen S, Li D, et al. Predicting the implicit and the explicit video popularity in a User Generated Content site with enhanced social features[J]. Computer Networks, 2018.

Chau H F , Wong C Y , Chow F K , et al. Social judgment theory based model on opinion formation, polarization and evolution[J]. Physica A Statistical Mechanics & Its Applications, 2014, 415:133-140.

Fan K , Pedrycz W . Opinion evolution influenced by informed agents[J]. Physica A Statistical Mechanics & Its Applications, 2016, 462:431-441.

Wang S , Fu X . Opinion dynamics on online-offline interacting networks:media influence and antagonistic interaction[J]. 2016.

Fan K , Pedrycz W . Evolution of public opinions in closed societies influenced by broadcast media[J]. Physica A: Statistical Mechanics and its Applications, 2017, 472:53-66.

Deffuant, Guillaume (2006). 'Comparing Extremism Propagation Patterns in Continuous Opinion Models'. Journal of Artificial Societies and Social Simulation 9(3)8

Amblard F, Deffuant G, Weisbuch G. How Can Extremism Prevail? a Study Based on the Relative Agreement Interaction Model[J]. Journal of Artificial Societies & Social Simulation, 2002, 5(4):1.

JAVIER GÓMEZ-SERRANO, Graham, C. , Boudec, J. Y. L. , & Bellomo, N. . (2012). The bounded confidence model of opinion dynamics. Mathematical Models and Methods in Applied Sciences, 22(02), 1150007-1-1150007-46.

Degroot M H . Reaching a Consensus[J]. Journal of the American Statistical Association, 1974, 69(345):118-121.

Hegselmann, Rainer & Krause, Ulrich. (2002). Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation. Journal of Artificial Societies and Social Simulation. 5.

Cacioppo J T , Gardner W L , Berntson G G . Beyond Bipolar Conceptualizations and Measures: The Case of Attitudes and Evaluative Space[J]. Personality and Social Psychology Review, 1997, 1(1):3-25.

Cacioppo JT, Bernston CG (1994) Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. Psychol Bull 115 401–423. doi: 10.1037//0033-2909.115.3.401

Norris C J , Larsen J T , Crawford L E , et al. Better (or worse) for some than others: Individual differences in the positivity offset and negativity bias[J]. Journal of Research in Personality, 2011, 45(1):100-111.

Rozin P , Royzman E B . Negativity Bias, Negativity Dominance, and Contagion[J]. Personality & Social Psychology Review, 2001, 5(4):296-320.

Smith N K , Cacioppo J T , Larsen J T , et al. May I have your attention, please: Electrocortical responses to positive and negative stimuli[J]. Neuropsychologia, 2003, 41(2):171-183.

Chen A . Are we sensitive to valence differences in emotionally negative stimuli? Electrophysiological evidence from an ERP study[J]. Neuropsychologia, 2007, 45(12):2764-2771.

Luis Carretié, Mercado F , Tapia M , et al. Emotion, attention, and the 'negativity bias', studied through event-related potentials[J]. International Journal of Psychophysiology Official Journal of the International Organization of Psychophysiology, 2001, 41(1):75-85.

Liu Y , Rui J R , Cui X . Are People Willing to Share Their Political Opinions on Facebook? Exploring Roles of Self-Presentational Concern in Spiral of Silence[J]. Computers in Human Behavior, 2017, 76.

Morrison K R . A license to speak up: Outgroup minorities and opinion expression[J]. Journal of Experimental Social Psychology, 2011, 47(4):0-766.

Pang N , Ho S S , Zhang A M R , et al. Can spiral of silence and civility predict click speech on Facebook?[J]. Computers in Human Behavior, 2016, 64:898-905.

Yun G W , Park S Y , Lee S . Inside the Spiral: Hostile Media, Minority Perception, and Willingness to Speak Out on a Weblog[J]. Computers in Human Behavior, 2016, 62:236-243.

Chun J W , Lee M J , Chun J W , et al. When does individuals' willingness to speak out increase on Social Media? Perceived Social Support and Perceived Power/Control[J]. Computers in Human Behavior, 2017, 74.

Novak P K, Smailović J, Sluban B, et al. Sentiment of Emojis[J]. Plos One, 2015, 10(12).

Appendix

1. The mathematical proof for the factor $\frac{(1-|x_i^t|)}{2}$ in the kernel function:

The usage for this is to constrain the opinion value between -1 and 1. Given $x_i^t < 0$, the repulsion function will be used when X_i^t is positive and we need to ensure that $x_i^{t+1} > -1$.

$$\begin{aligned} x_i^{t+1} &= x_i^t + f(x_i^t, X_i^t)(X_i^t - x_i^t) \\ &= x_i^t - \frac{\beta(1-|x_i^t|)}{2}(X_i^t - x_i^t) \\ &= x_i^t - \frac{\beta(1+x_i^t)}{2}(X_i^t - x_i^t) \\ &\geq x_i^t - \frac{\beta}{2}[1-(x_i^t)^2] \end{aligned}$$

Given

$$\begin{aligned} &f(x) = x - \frac{\beta}{2}[1-x^2] \\ &If\ \beta \in [0,1],\ f(x)_{min} = f(-1) = -1 \\ &If\ \beta > 1, \quad f(x)_{min} = f\left(-\frac{1}{\beta}\right) = -\frac{1}{2}\left(\beta + \frac{1}{\beta}\right) \end{aligned}$$

To make $f\left(-\frac{1}{\beta}\right) \geq -1$, we obtain $\beta = 1$, posing a contradict. Therefore, β has to be in $[0,1]$ with the constrain factor to make sure the opinion value will not cross the boundary.

2.The sentiment identification of Naïve Bayes model:

$$P(c_1|w_1, \dots w_n) = \frac{P(w_1, \dots w_n|c_1) * P(c_1)}{P(w_1, \dots w_n|c_1) * P(c_1) + P(w_1, \dots w_n|c_2) * P(c_2)}$$

$$= \frac{1}{1 + \frac{P(w_1, \dots w_n|c_2) * P(c_2)}{P(w_1, \dots w_n|c_1) * P(c_1)}}$$

$$= \frac{1}{1 + exp\left[log\left(\frac{P(w_1, \dots w_n|c_2) * P(c_2)}{P(w_1, \dots w_n|c_1) * P(c_1)}\right)\right]}$$

$$= \frac{1}{1 + exp\left[log\left(P(w_1, \dots w_n|c_2) * P(c_2)\right) - log\left(P(w_1, \dots w_n|c_1) * P(c_1)\right)\right]}$$

Where c_1, c_2 represents the class labels of the object (in our study, positive or negative); and w_i presents the properties of the text after sentence segmentation.

3.Test simulation results:

One host and five hundred audiences with time span of 150 units.

Scenario 1:(standard)

$x_H^0 = 0.5, x_i^0 = 0, \tau_i^0 = 1, a = 0.5, \beta = 0.5, d_1 = 1, d_2 = 1, \mathrm{b} = 0.5$

$k_P = 0.5, k_N = 1, c = 0.25, \theta = 1, p = 0.5, m = 0.5, n = 1.5$

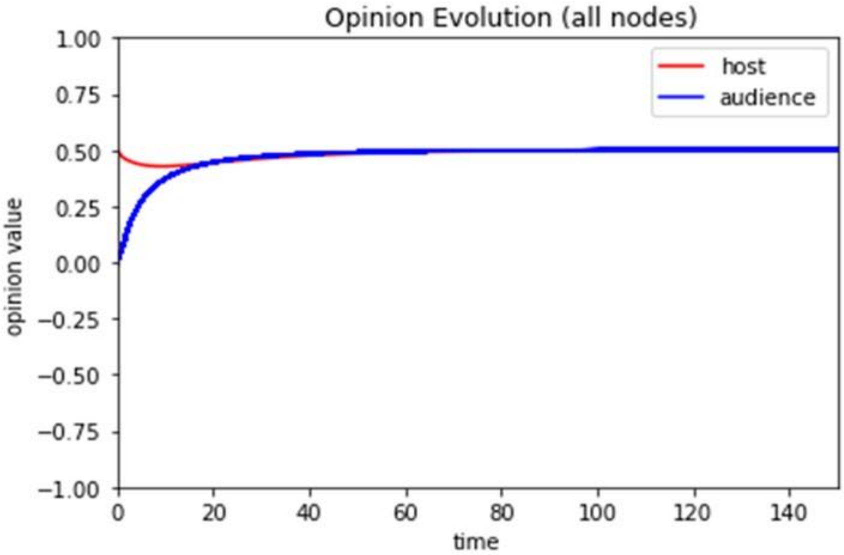

Figure 10. Simulation Result: Scenario 1

Scenario 2: (considering the difference of the host)

2.1: diminishing the difference

$x_H^0 = 0.5, x_i^0 = 0, \tau_i^0 = 1, a = 0.5, \beta = 0.5, d_1 = 1, d_2 = 1, \mathrm{b} = 0.5$

$k_P = 0.5, k_N = 1, c = 0.25, \theta = 1, p = 0.5, m = 1, n = 1$

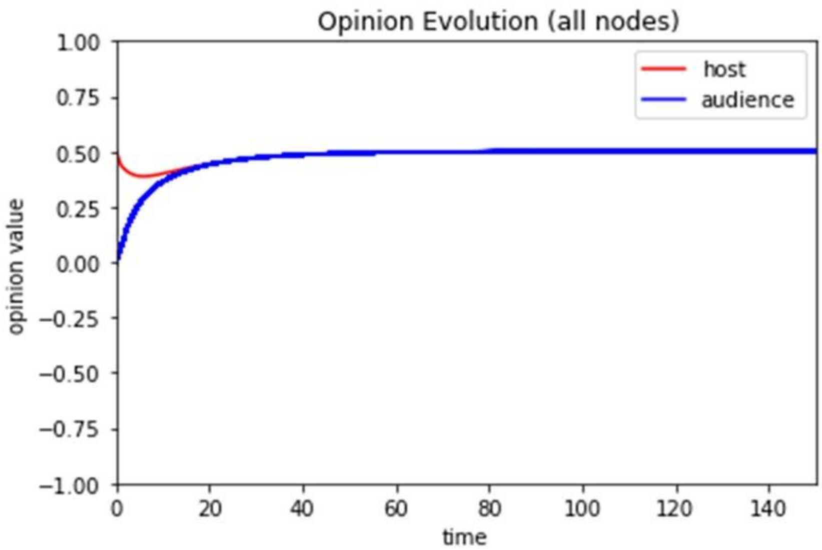

Figure 11. Simulation Result: Scenario 2.1

2.2: increasing the difference

$x_H^0 = 0.5, x_i^0 = 0, \tau_i^0 = 1, a = 0.5, \beta = 0.5, d_1 = 1, d_2 = 1, \mathrm{b} = 0.5$

$k_P = 0.5, k_N = 1, c = 0.25, \theta = 1, p = 0.5, m = 0.2, n = 1.8$

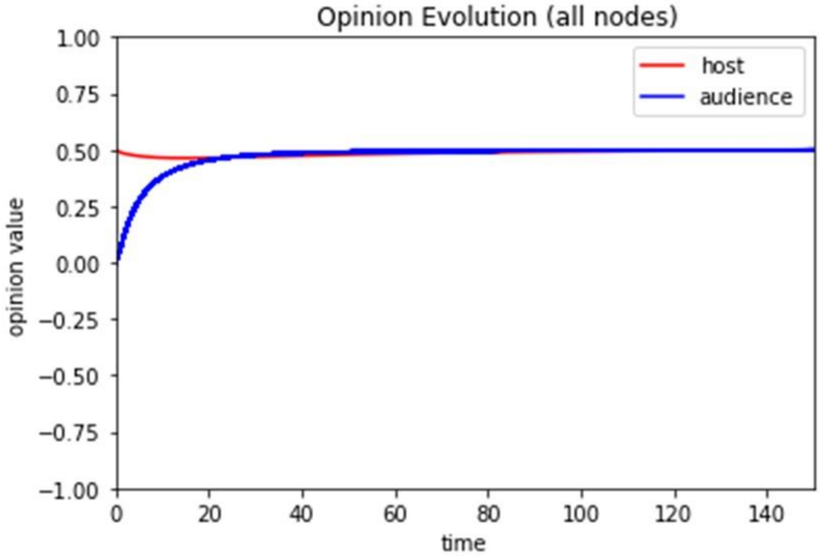

Figure 12. Simulation Result: Scenario 2.2

To test the impact of parameter m and n towards the host, we duplicate the simulations with different values of m and n. The results are displayed below:

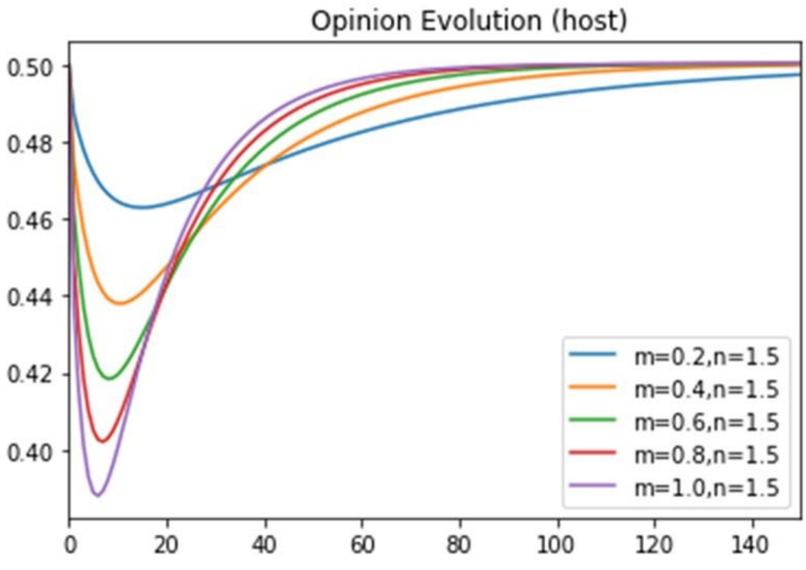

Figure 13. Simulation Result of the host for the variation in *n*

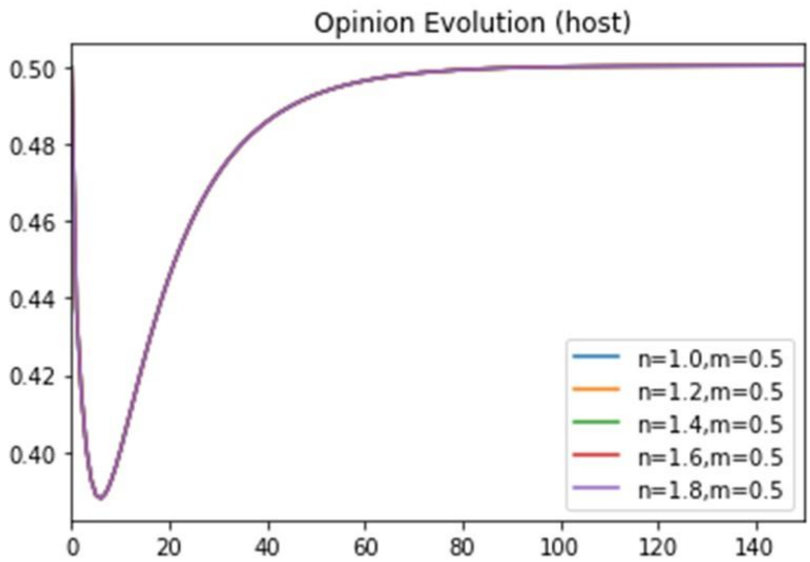

Figure 14. Simulation Result of the host for the variation in m

These two graphs indicate that when the audiences have no biases in opinions at the beginning, the stabilization coefficient m has the decisive impact on the opinion evolution of the host.

Scenario 3: (considering the differentiation of initial opinions)
$x_H^0 = 0.5, x_i^0 = uniform(-1,1), \tau_i^0 = 1, a = 0.5, \beta = 0.5, d_1 = 1, d_2 = 1$, b $= 0.5$
$k_P = 0.5, k_N = 1, c = 0.25, \theta = 1, p = 0.5, m = 1, n = 1$

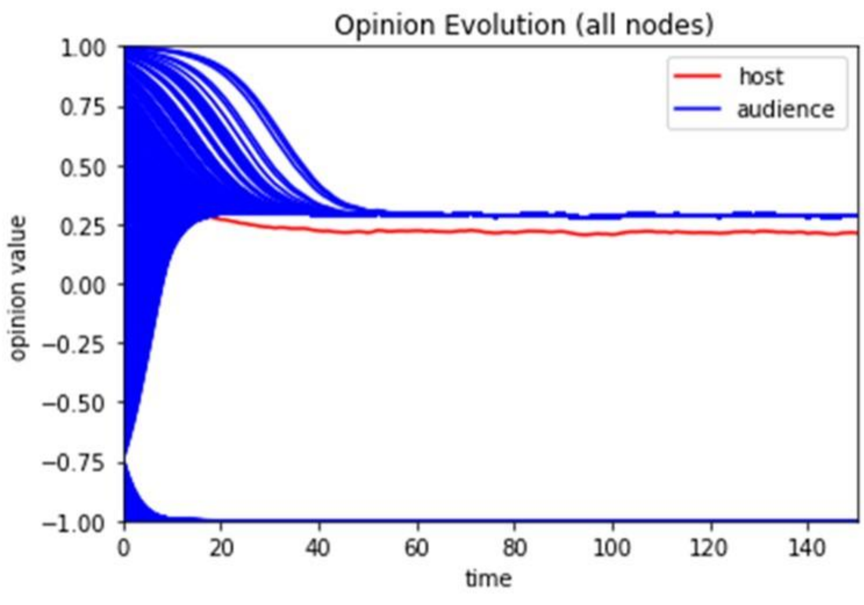

Figure 15. Simulation Result: Scenario 3