Social-judgment-based Modeling of Opinion Polarization in Chinese Live Streaming Platforms

2021 SBP-BRIMS Conference

Yutao Chen
cytwill@uchicago.edu
University of Chicago

Agenda

- Background: Live Streaming In China
- Research Goals:

  1. Modeling Opinion Evolution: Theories and Mathematic Formulation

  2. Scenario 1: no extremists, studying the effect of different parameters of the SJBO model

  3. Scenario 2: studying the effect of extremists in the initial state
- Indications and Future work

Live Streaming in China： ecosystem

- **WHEN**: the industry get boosted since the middle of 2010s, starting from small platforms to large IT service companies. In 2016, almost every 3 day a new live streaming platform came up.

- **WHAT**: almost everything! the content of live streaming ranges from art performance, eating (ASMR), video games, e-commerce and even everyday life.

- **WHO**: young people and celebrities! Low requirements with promising potential earnings. Online celebrity training programs for novices.

Live Streaming in China： traits

- **Synchronized Interaction**: audiences can interact with the streamers by sending comments (screen-bullet 弹幕) during the live streaming.
- **Complete Connectivity**: each audiences can see all the comments from other audiences (no blocking considered).
- **Information Overload**: too many comments scrolling on the screen, distracting the audiences from the streaming content.

Live Streaming in China：issues

- **From Content**: violence, pornography and other unsuitable content in small platforms. Much better in recent years after relevant legislation and reinforced supervision.
- **From Behaviors**: internet addiction (gaming live streaming), spending too much money on membership or virtual gift sending (teenagers).
- **From Comments**: online violence when the comments are dominated by anti-fans or extremists, threatening the mental health of streamers and giving bad examples of language use in online communities.

Modeling: Social-judgment-based Opinion Model (SJBO)

Social Judgment Theory (Sherif & Hovland 1961; Sherif et al. 1965): when people are exposed to new information, they will use their existing attitudes to perceive and evaluate it, and decide whether to approve the idea based on his existing position.

Bounded Confidence (BC) Model

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

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

$$\mu \in (0,1), u \in (0,1)$$

\+ Boomerang effect

Social-judgment-based (SJBO) Model

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

$$f\left(x_i^t, x_j^t\right) = \begin{cases} \dfrac{\alpha\left(1 - \left|x_i^t\right|\right)}{2}, & \left|x_j^t - x_i^t\right| < \varepsilon \\ 0, & \varepsilon \le \left|x_j^t - x_i^t\right| \le r \\ \dfrac{-\beta\left(1 - \left|x_i^t\right|\right)}{2}, & \left|x_j^t - x_i^t\right| > r \end{cases}$$

$$\varepsilon \in (0,1),\ r \in (1,2),\ \alpha, \beta \in (0,1)$$

Modeling: Social-judgment-based Opinion Model (SJBO)

Social Judgment Theory (Sherif & Hovland 1961; Sherif et al. 1965): when people are exposed to new information, they will use their existing attitudes to perceive and evaluate it, and decide whether to approve the idea based on his existing position.

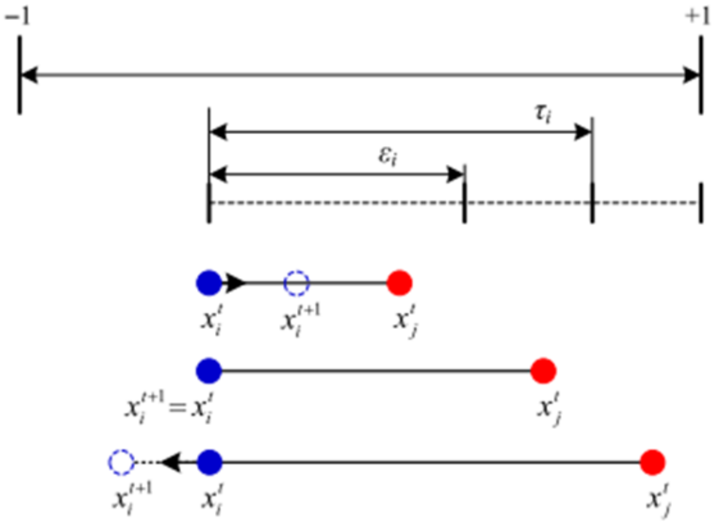


Fig. 2. Schematic diagram of SJBO model.

Social-judgment-based (SJBO) Model

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

$$f(x_i^t, x_j^t) = \begin{cases} \dfrac{\alpha(1-|x_i^t|)}{2}, & |x_j^t - x_i^t| < \varepsilon \\ 0, & \varepsilon \le |x_j^t - x_i^t| \le r \\ \dfrac{-\beta(1-|x_i^t|)}{2}, & |x_j^t - x_i^t| > r \end{cases}$$

$\varepsilon \in (0,1),\ r \in (1,2),\ \alpha, \beta \in (0,1)$

Modeling: Information Overload vs Host Impact

- The content of live streaming is quantified by a triangular distribution and is assumed not to be impacted by the audience's opinions as most hosts would prepare their contents before and make sure the live streaming progresses well.

- The audiences could interact with more than one opinions each time. However, when too many comments scroll on the screen, it is unrealistic for audiences to capture all of them. Thus, I define that for each time t, agents can only interact with $\min(k, |S^t| - 1)$ randomly picked comments.

- Live streaming content fully impacts the audiences' opinions when no comments are sent. As the comments increase, audiences' attention would be distracted by them, and the streaming content would have less impact as a result of information overload.

Modeling：Information Overload vs Host Impact

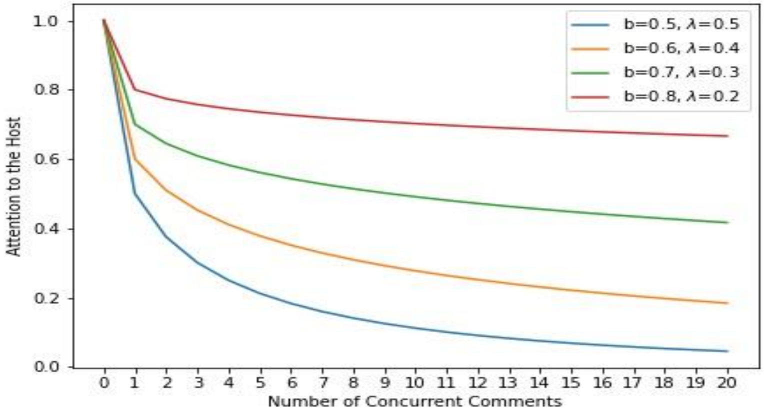
$$x_i^{t+1} = \begin{cases} x_i^t + f(x_i^t, X_i^t)(X_i^t - x_i^t) & i \neq host \\ Triang(l, m, 1) & i = host \end{cases}$$

$$f(x_i^t, x_j^t) = \begin{cases} \dfrac{\alpha(1 - |x_i^t|)}{2}, & |x_i^t - X_i^t| < \varepsilon \\ 0, & \varepsilon \leq |x_i^t - X_i^t| \leq r \\ \dfrac{-\beta(1 - |x_i^t|)}{2}, & |x_i^t - X_i^t| > r \end{cases}$$

$$X_i^t = \sum_{j \in S^t} w_{ij}^t x_j^t$$

$$w_{ij}^t = \begin{cases} b^{(|S^t|-1)^{\lambda}}, & j = host \\ \dfrac{1 - b^{(|S^t|-1)^{\lambda}}}{\min(k, |S^t| - 1)}, & j \neq host \end{cases}$$

b: the base attention for the host defined in [0,1]
λ: the distraction coefficient defined in [0,1]
Bigger b and smaller λ indicate less impact from the comments



Modeling: Opinion Expression (Entertainment Motivation)

- Audiences usually do not always send out their comments, so we need to design a mechanism to make the agents express their opinions sometimes rather than constantly.
- Entertainment motivation (Jarvinen et al. 2002): online behaviors like leaving comments or posting tweets can bring them relaxation and pleasure, which drive them to engage in online activities.

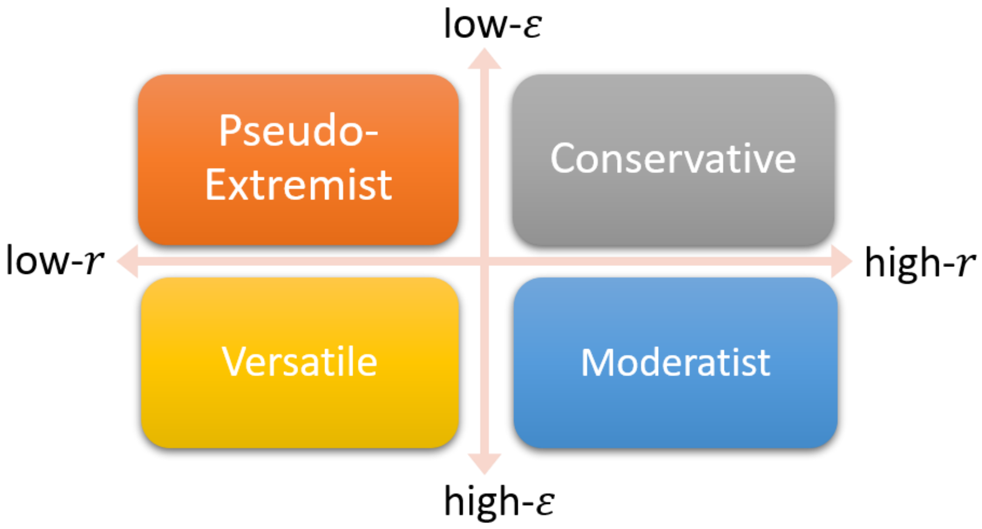
$$\tau_i^t = Bernoulli(\delta)$$

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

Scenario 1 (no extremists): Different Types of Audiences

- Setting: Population $N = 500$, $T = 600$, $m = 0.75$, $l = 0.5$, $\alpha, \beta = 0.5$, initial opinions uniformly drawn from [-0.8, 0.8].

- Audience type and Evolution tendency



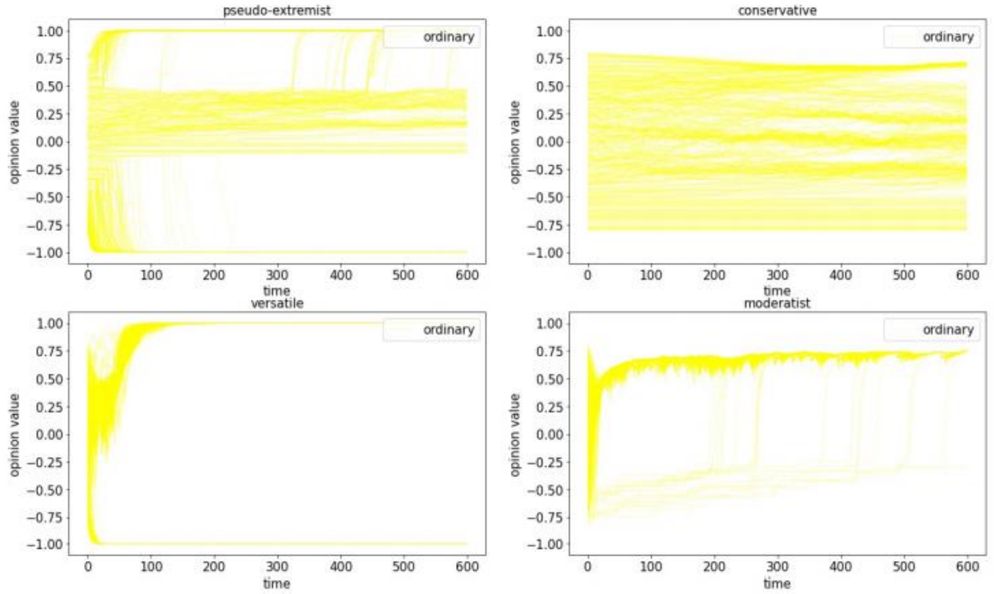


Fig. 2. Polarization tendency of different agent states (k=1, δ=0.02, b=0.6, λ=0.4): pseudo-extremist: ϵ=0.1, r=1.1, conservative: ϵ=0.1, r=1.9, versatile: ϵ=0.9, r=1.1, moderatist: ϵ=0.9, r=1.9

Scenario 1 (no extremists): impact from parameters.

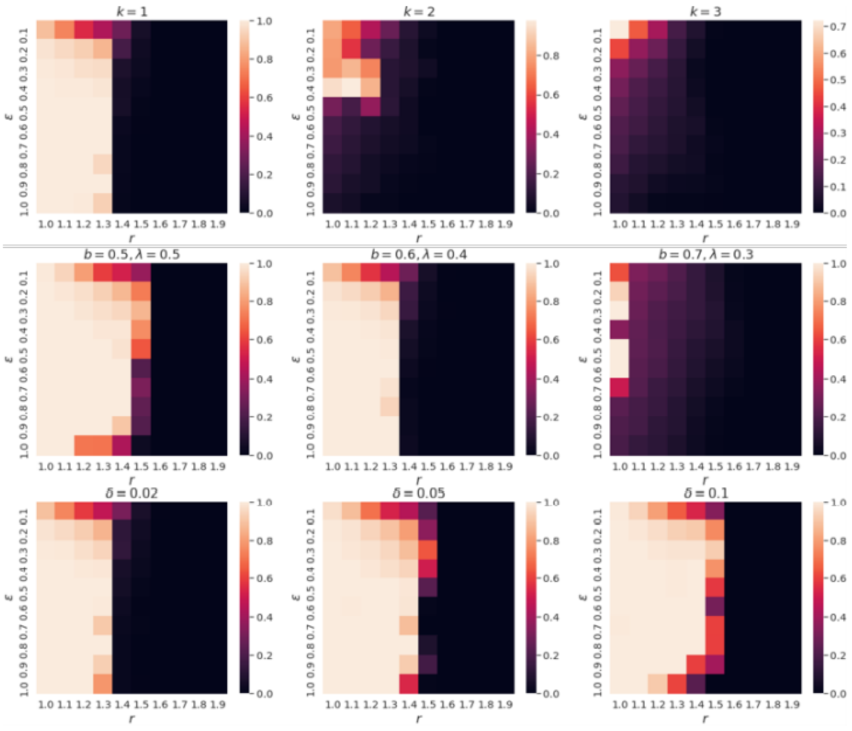


Fig. 3. Top: Polarization rate with different values of k (δ=0.02, b=0.6, λ=0.4); Middle: Polarization rate with different values of b, λ (k=1, δ=0.02); Down: Polarization rate with different values of δ (k=1, b=0.6, λ=0.4)

- Parameter k (number of comments to interact): reduce the likelihood to grow extreme opinions as parameter k increases. (people interacting with different opinions tend to develop more balanced opinions rather than going extreme)

Scenario 1 (no extremists): impact from parameters.

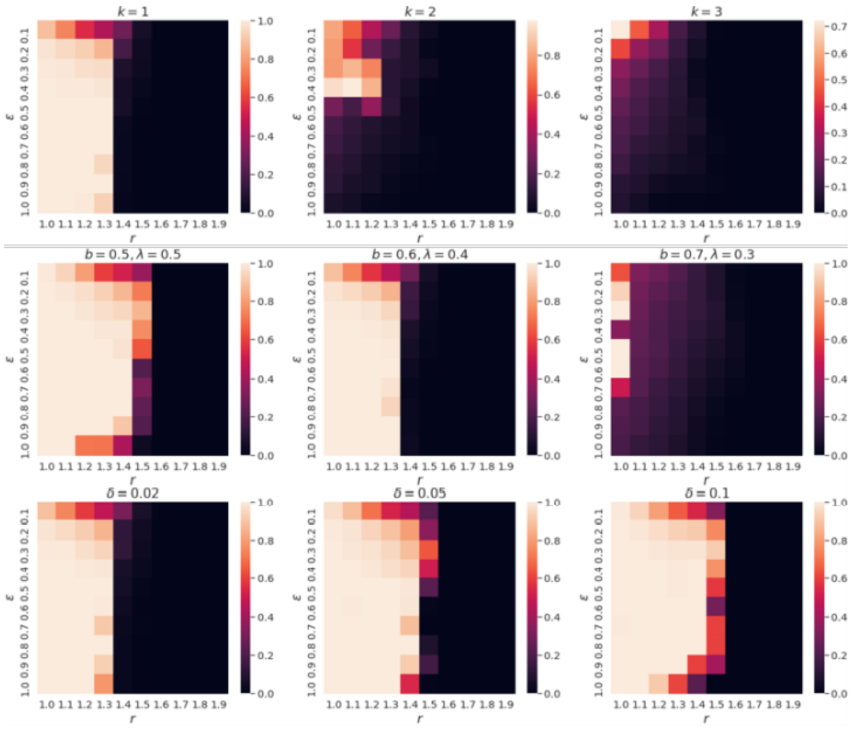


Fig. 3. Top: Polarization rate with different values of k (δ=0.02, b=0.6, λ=0.4); Middle: Polarization rate with different values of b, λ (k=1, δ=0.02); Down: Polarization rate with different values of δ (k=1, b=0.6, λ=0.4)

❖ Parameter b, λ, δ: as the attention for the streamer increases (larger b, smaller λ, smaller δ), the audiences need lower repulsion thresholds to develop extreme opinions. This would also make the audiences' opinions converge to the streaming content value.

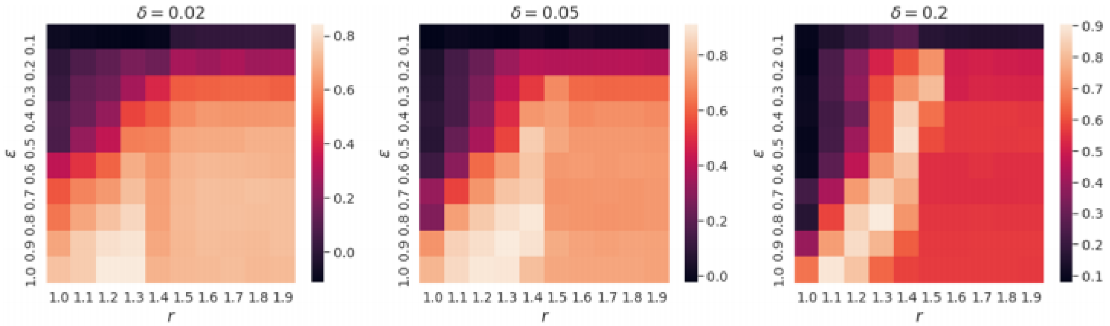


Fig. 5. Average final opinion value with different values of δ (k=1, b=0.6, λ=0.4)

Scenario 2 (with extremists)： defining extremists

Definitions and Quantification of extremists (Martins 2020):

- ✓ **Having more extreme opinions than ordinary audiences**: 0.8 is set to be the threshold of the absolute values of extreme opinions, the absolute values of extremists are in [0.8, 1]
- ✓ **Being more difficult to change their opinions to a milder direction**: extremists have lower assimilation threshold ($\epsilon_e < \epsilon$) and lower repulsion threshold ($r_e < r$)
- ✓ **being easier to take actions to advocate their opinions:** extremists have higher entertainment motivation ($\delta_e > \delta$)

Scenario 2 (with extremists): impact from initial extremists

- Setting: ordinary audiences are the same as in Scenario 1 except for $\epsilon = 0.5$, $r = 1.5$, $\delta = 0.02$. The initial opinions of positive extremists are uniformly drawn from [0.8, 1], and negative extremists' opinions are drawn from [−1, −0.8]. $\epsilon_e = 0.1$ and $r_e = 1.1$ is set for all extremists.
- Impact measured by positive polarization rate (the proportion of ordinary audiences holding positive extremist opinions in the end)
- More ordinary agents have extreme positive opinions when more positive extremists exist and fewer negative extremists exist
- Explain why buying pseudo-fans to make extremely positive comments and the effectiveness of controlling anti-fans or internet trolls to reduce cyberbullying

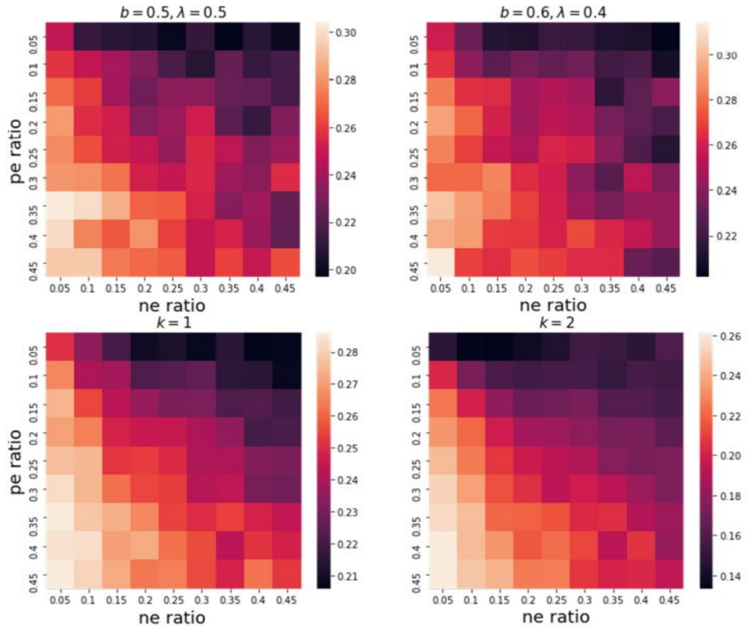


Fig. 6. Top: Positive polarization rate with different values of b, λ (k=1, δ=0.02); Down: Positive polarization rate with different values of k (b=0.6, λ=0.4, δ=0.02), *pe* and *ne* in the axis labels mean positive and negative extremists

Scenario 2 (with extremists)：impact from k

When one type of extremists dominate, increasing k could increase the interactions with that type of extreme opinions, and thus increases the corresponding polarization.

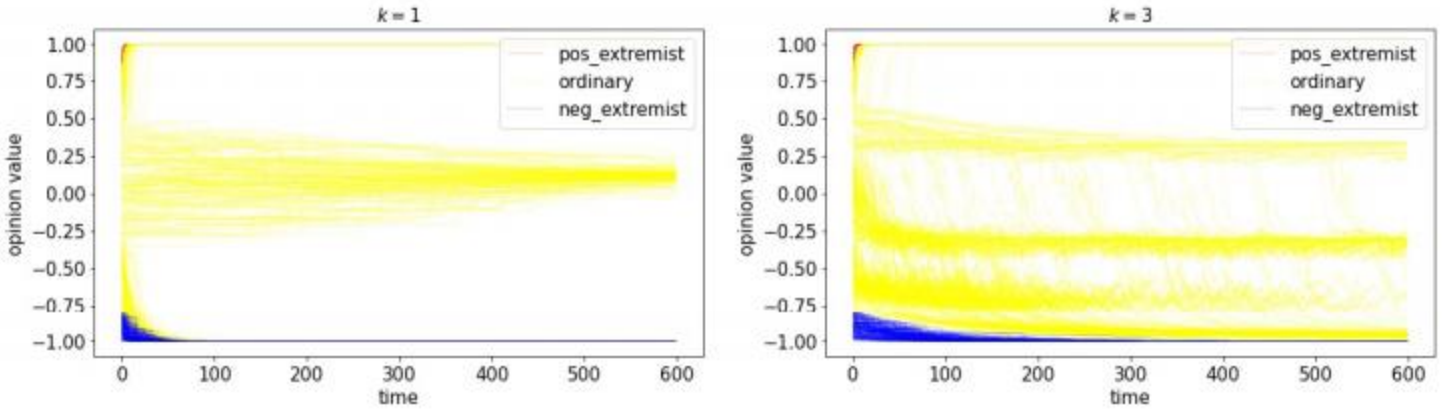


Fig. 7. Polarization tendency comparison between k=1 and k=3 with a negative extremist dominance (b=0.6, λ=0.4, δ=0.02, δ_e=0.8)

Indications

- The importance to build opinion dynamic models suiting to specific platforms or scenarios.
- **Internal Perspective**:

Extreme opinions can occur without the introduction of extremists when audiences are easy to repulse other opinions.

Audiences' attentions to the comments play a role in polarizing their opinions, and the role is varied according to the balance between the positive and negative extremists.

- **External Perspective**:

Controlling the comments available to the audiences might help to draw more attention to the hosts and reduce unwanted polarization.

Trying to identify and control the number of extremists (especially negative ones) are helpful to reduce unwanted polarization in live streaming.

Future Work

- Dynamics of the audiences network
- Different comment blocking strategies
- Other internal or external factors influencing the process
- Opinions to Consumptions

Thank you for watching!