

# How People Affect Each Other on Social Networks?



Zhiping (Patricia) Xiao  
University of California, Los Angeles

## Background

- Motivation

- Paper List

## Problem Settings & Models

- Opinion Dynamics Models

- Data-Driven Analysis

The End

# Background

---



My previous works:

- ▶ Infer opinion from relation;
- ▶ Infer opinion from text content of posts / tweets.

More:

- ▶ Spread existing opinions;
- ▶ Influence some others' opinions.

Besides, nobody is satisfied with the social media environments nowadays. But we don't know what treatments can be applied yet.

### Opinion dynamics models:

- ▶ Predicting Opinion Dynamics via Sociologically-Informed Neural Networks (KDD'22)
- ▶ (\*) A model for the influence of media on the ideology of content in online social networks (Physical Review Research'20)

### Data-Driven Analysis:

- ▶ The effect of wording on message propagation: Topic- and author-controlled natural experiments on Twitter (ACL'14)
- ▶ (\*) Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement (Journal of Marketing Research'20)

### (\*) Some other related aspects:

- ▶ Integrating explanation and prediction in computational social science (Nature'21)

# Problem Settings & Models

---



Code: [https://github.com/mayaokawa/opinion\\_dynamics](https://github.com/mayaokawa/opinion_dynamics)

- ▶ Data unavailable, crawler script provided;
- ▶ Using Neural Networks to model opinion dynamics models.
- ▶ **NOT** considering network structure (e.g. follower-follower relations).

The idea:

- ▶ Use Neural ODE framework to learn the parameters of the social dynamical systems.
- ▶ Not using VAE for learning (i.e. NOT modeling a trajectory).

From step  $t$  to step  $t + 1$ , consider a single user  $x_u$ , given a opinion dynamics model  $f$  which predicts it as  $\tilde{x}_u$ , it says,

$$\tilde{x}_u(t + 1) = x_u(t) + \int_t^{t+1} f_{\theta}(\mathbf{X}(t))dt,$$

where  $\mathbf{X}$  can be all nodes' opinions in theory (but always select some in practice).  $f$  includes some learnable parameters.

Then, they use an MLP net  $g_{\phi}$  to model the dynamical system:  
code here

$$\hat{x}_u(t + 1) = g_{\phi}(\mathbf{X}(t))$$

- Note that: their code mentioned “attention” but there is no attention mechanism ([https://github.com/mayaokawa/opinion\\_dynamics/blob/main/modules.py#L35](https://github.com/mayaokawa/opinion_dynamics/blob/main/modules.py#L35))



Then the loss is computed as:

$$\ell_{all} = \ell_{data}(\hat{x}_u(t), x_u(t)) + \ell_{ode}\left(\frac{\partial g_\phi(\mathbf{X}(t))}{\partial t}, f_\theta(\mathbf{X}(t))\right) + \ell_{reg}(\theta),$$

where:

- ▶  $\ell_{data}$ : the difference (Cross Entropy Loss) between ground truth ( $x_u(t)$ ) and predicted result ( $\hat{x}_u(t)$ )
- ▶  $\ell_{ode}$ : the difference (MSE loss) between the SINN model's gradient ( $\frac{\partial g}{\partial t}$ , fetched via `torch.autograd.grad`), and the ODE gradient ( $f_\theta(\mathbf{X}(t))$ )
- ▶  $\ell_{reg}$ : regularization on the parameters of  $f$ , usually sum of their  $\ell_1$  norms (note that for different versions of opinion dynamics models,  $f$  are different, and there will be different ways of computing  $\ell_{reg}$ ).

Opinion dynamics: the study of how opinions emerge and evolve through exchanging opinions with others.

Problem Settings in General:

- ▶ Set of users  $\mathcal{U}$
- ▶ Each person  $u$  holds opinion  $x_u(t) \in [-1, 1]$  on a specific subject at time  $t$ .
- ▶ Users' opinions will affect each other. These models use math formula to model how opinions are updated from  $t$  to  $t + 1$ .
- ▶ There are many models with different **update rules**. (i.e. different versions of  $f$  as we mentioned in previous pages)

DeGroot: simple, basic

$$x_u(t+1) = x_u(t) + \sum_{\mathcal{U}/u} a_{uv} x_v(t),$$

where  $\mathcal{U}/u$  denotes all other users in the system and  $a_{uv}$  is the strength of the interactions between  $u$  and  $v$ .

It models *assimilation* (i.e. tendency of moving opinions towards others) well.

ODE version in SINN:

$$\frac{dx_u(t)}{dt} = \sum_{\mathcal{U}/u} a_{uv} x_v(t) = \sum_{\mathcal{U}/u} \mathbf{m}_u^T \mathbf{q}_v x_v(t)$$

Friedkin-Johnsen (FJ) model: allows for stubbornness

$$x_u(t+1) = s_u \sum_{v/u} x_v(t) + (1 - s_u)x_u(0),$$

where  $s_u \in [0, 1]$  denotes a user's susceptibility to persuasion. The  $s_u$ , the more open-minded a person is.

It models *susceptibilities* to persuasion (i.e., the tendency to defer to others' opinions) well.

ODE version in SINN:

$$\frac{dx_u(t)}{dt} = s_u \sum_{v/u} x_v(t) + (1 - s_u)x_u(0) - x_u(t)$$

Bounded confidence model: models *confirmation bias* (i.e., tendency to focus on information that confirms our preconceptions). A family of model. The most popular variant, the Hegselmann-Krause (HK) model:

$$x_u(t+1) = x_u(t) + \frac{1}{|N_u(t)|} \sum_{v \in N_u(t)} (x_v(t) - x_u(t)),$$

where  $N_u(t)$  denotes the set of users whose opinions fall within the bounded confidence interval of  $u$  at  $t$ :

$$N_u(t) = \{v \in \mathcal{U} \mid |x_u(t) - x_v(t)| \leq \delta\}$$

ODE version in SINN:

$$\frac{dx_u(t)}{dt} = \sum_{v \in \mathcal{U}} \sigma(\delta - |x_u(t) - x_v(t)|) (x_u(t) - x_v(t))$$

Stochastic Bounded confidence model: incorporating stochastic interactions based on opinion differences. Use  $p(z_{uv}^t = 1)$  to model the probability that  $u$  and  $v$  interact at time  $t$ .

$$p(z_{uv}^t = 1) = \frac{|x_u(t) - x_v(t)|^{-\rho}}{\sum_k |x_u(t) - x_k(t)|^{-\rho}},$$

ODE version in SINN:

$$\frac{dx_u(t)}{dt} = \sum_{v \in \mathcal{U}} \tilde{z}_{uv}^t (x_v(t) - x_u(t)),$$

where  $\tilde{z}_{uv}^t$  is computed from  $\mathbf{p}_u^t \in \mathbb{R}^U$ :

$$\tilde{z}_u^t = \text{Softmax}([\log(\mathbf{p}_u^t) + \mathbf{g}_u]/\tau),$$

with  $\mathbf{g}_u$  being random noise and  $\tau$  a temperature parameter. When  $\tau \rightarrow 0$ ,  $\tilde{z}_u^t$  approximates one-hot.

$$x_u(t+1) = \frac{1}{|I_u| + w} \left( wx_u(t) + \sum_{v \in \mathcal{M} + \mathcal{N}} \mathbf{A}_{uv} x_v(t) f(x_u(t), x_v(t)) \right),$$

where  $I_u$  is the set of accounts to which account  $u$  is receptive to,  $w$  is a pre-defined hyper-parameter,  $\mathbf{A}$  is the adjacency matrix of the network,  $f$  is a function that could be defined like “when [cond] then 1 else 0”.

From paper: *A Model for the Influence of Media on the Ideology of Content in Online Social Networks*

- ▶ Considered network structure;
- ▶ Somewhat more similar to GNN update rules.
- ▶ Note: this update rule is designed for modeling **media** influence.  $\mathcal{M}, \mathcal{N}$  are the sets of media and normal accounts respectively.

- ▶ Efforts were made to bridge the gap between opinion dynamics models and powerful computation tools (e.g. Neural Networks).
- ▶ More and more research works have considered graph structure in modeling social influence.
- ▶ Opinion dynamics models are having strong assumptions in general, bringing about a gap between theory and practice.



The effect of wording on message propagation: Topic- and author-controlled natural experiments on Twitter

- ▶ **Effect of Wording:** Investigate whether a different choice of words affects message propagation, *controlling speaker and the topic*.
- ▶ Measure propagation by #retweet

(\*) Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement

- ▶ **Image Content Engagement:** Investigate the influence of image content on social media engagement, empirically.
- ▶ Measure engagement by #retweet and #like

Observation: *it is unexpectedly common for the same user to post different tweets regarding the same URL.*<sup>1</sup>

Data:

- ▶ **TAC: Topic- and Author-Controlled pairs**
  - ▶ The previous one:  $t_1$ , the later one:  $t_2$ . Corresponding number of retweets:  $n_1, n_2$ .
  - ▶ How to control author: from the same account.
  - ▶ How to control topic: including the same URL.
- ▶ Examples:

author	tweets	#retweets
natlsecuritycnn	$t_1$ : FIRST ON CNN: After Petraeus scandal, Paula Broadwell looks to recapture 'normal life.' <a href="http://t.co/qy7GGuYW">http://t.co/qy7GGuYW</a>	$n_1 = 5$
	$t_2$ : First on CNN: Broadwell photos shared with Security Clearance as she and her family fight media portrayal of her [same URL]	$n_2 = 29$
ABC	$t_1$ : Workers, families take stand against Thanksgiving hours: <a href="http://t.co/J9mQHilEqv">http://t.co/J9mQHilEqv</a>	$n_1 = 46$
	$t_2$ : Staples, Medieval Times Workers Say Opening Thanksgiving Day Crosses the Line [same URL]	$n_2 = 27$
cactus.music	$t_1$ : I know at some point you've have been saved from hunger by our rolling food trucks friends. Let's help support them! <a href="http://t.co/zg9jwA5J">http://t.co/zg9jwA5J</a>	$n_1 = 2$
	$t_2$ : Food trucks are the epitome of small independently owned LOCAL businesses! Help keep them going! Sign the petition [same URL]	$n_2 = 13$

At: <https://chenhaot.com/pages/wording-for-propagation.html>

<sup>1</sup> Natural experiments and quasinalatural experiments (DiNardo, 2008)

Three versions of Features:

- ▶ Customize: Combining all 39 features (any feature can be used independently), including “ask people to share (explicitly)”, “1st person singular”, “positive/negative (sentiment)”, “informative” etc. These features are designed according to a lot of previous works.
- ▶ Also consider tagged bag-of-words (“BOW”) features, which includes all the unigram (word:POS pair) and bigram features.

Classifier: L2-regularized logistic regression, SVM

- ▶ A strong baseline: same classifier structure, including more features — time (day and hour) and follower-count, but not using TAC for training. (called  $\neg$ TAC+ff+time)

1. Do the wording effects exist?
  - ▶ Ask 106 humans to predict which version gets more widely spread (via Amazon Mechanical Turk experiment), and achieved an average accuracy of 61.3%.
  - ▶ It is somewhat possible to predict greater message spread from wording.
2. How to determine time-lag ( $|t_1 - t_2|$ ) and follower thresholds?

$$D = \sum_{0 \leq n_1 < 10} |\hat{E}(n_2|n_1) - n_1|$$

By examining  $D$  value's when other conditions are different. Here,  $\hat{E}(n_2|n_1)$  is the average value of  $n_2$  over pairs where  $t_1$  are retweeted  $n_1$  times.

Focus on analyzing the following aspects:

- ▶ Effectiveness: measured by attracting more retweets
- ▶ Author Prefer: measured by how often the authors have higher tendency of such feature in  $t_2$  than in  $t_1$
- ▶ Feature coefficients: measured by how well the model performs using that feature set.

Prediction performance:

- ▶ Human: 61.3%
- ▶  $\neg$ TAC+ff+time: 55.3%
- ▶ Using TAC: 65.6%

Some Findings:

- ▶ @-mentions and 2nd person pronouns are ineffective in promote retweeting, but these features are preferred by authors.

### Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement

- ▶ Similar Data Source and Ground Truth: from Twitter and Instagram, use likes and retweets counts as ground-truth engagements.
- ▶ Scope of Data narrowed to mostly commercial posts (sale, airline, etc.).
- ▶ Different Ways of Finding Pairs: Using propensity score matching approach to create a pseudo “treatment” (expose to image or not) group and a “control” group (1:1) on the basis of post and account characteristics.
- ▶ Models: logistic regression, multinomial naive Bayes, **linear support vector machine**, and random forest.

- ▶ Labeling is hard even to human beings;
  - ▶ e.g. *the Effect of Wording* work find human average accuracy of judging which message is more widespread 61.3%
- ▶ Hard to observe counter-factual pairs;
- ▶ Lack of ground-truth knowledge of the offline world;
- ▶ Most problems are not well-defined. It can be hard to convince your audience what you are studying in the first place.

(\*) *Integrating explanation and prediction in computational social science*

- ▶ Social Scientists

- ▶ Pros: Interpretable, Explainable, often invoking causal mechanisms.
- ▶ Cons: Fail to predict outcomes of interest, fail to offer solutions to real-world problems, fail to replicate results.

- ▶ Computer Science:

- ▶ Pros: Good at designing accurate predictive models.
- ▶ Cons: Neglecting causal mechanism, doesn't care whether or not the models are interpretable, easily biased.



Thank You All! 😊

Please feel free to discuss with me afterwards.