Popularity Prediction on Social Platforms with Coupled Graph Neural Networks

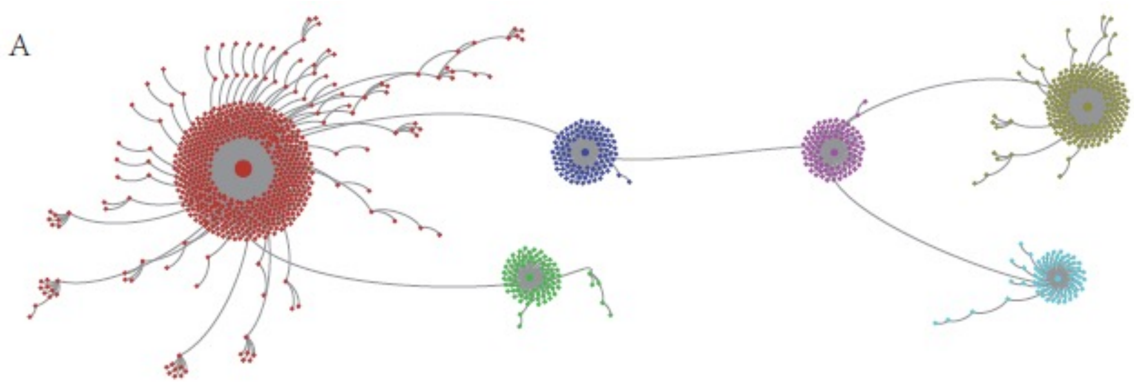
Qi Cao, Huawei Shen, Jinhua Gao, Bingzheng Wei, Xueqi Cheng

CAS Key Laboratory of Network Data Science and Technology,
Institute of Computing Technology,
Chinese Academy of Sciences, Beijing, China

Background

- Online social platforms provide convenience for the production and delivery of information

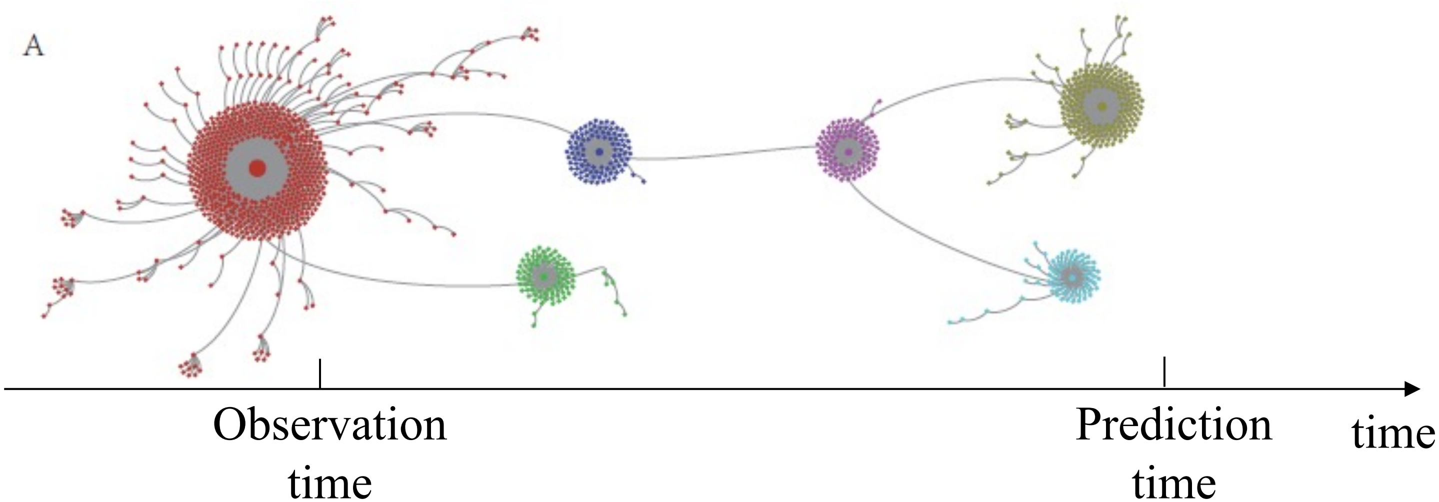
- Information cascades



Background

- Online social platforms provide convenience for the production and delivery of information

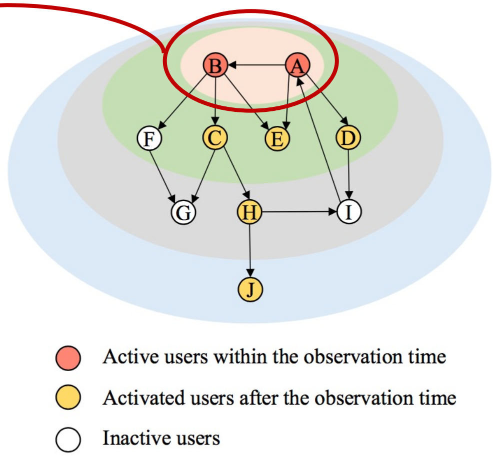
Predicting future popularity of information cascades on Social Platforms!



Outline

Motivation

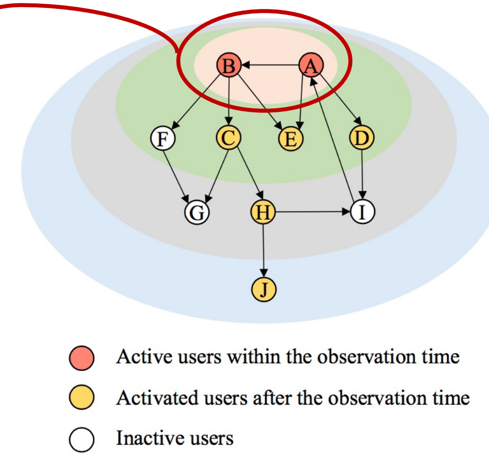
- Existing methods for popularity prediction mainly fall into three categories
  - ◆ Feature-based methods
    - Extract demographics, temporal and structural features of early adopters



Methods	consider early adopters	Approximate by statistics	Precisely capture cascading effect
Feature-based	✓	✗	✗
Deep learning-based			
Point process			

Motivation

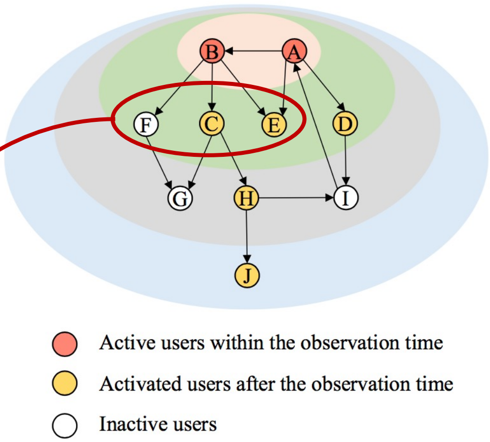
- Existing methods for popularity prediction mainly fall into three categories
  - Deep learning-based methods
    - Represent the subgraph of early adopters by deep learning



Methods	consider early adopters	Approximate by statistics	Precisely capture cascading effect
Feature-based	✓	✗	✗
Deep learning-based	✓	✗	✗
Point process			

Motivation

- Existing methods for popularity prediction mainly fall into three categories
  - Point process methods
    - Use average number of fans to approximate the impact of cascading effect in each generation



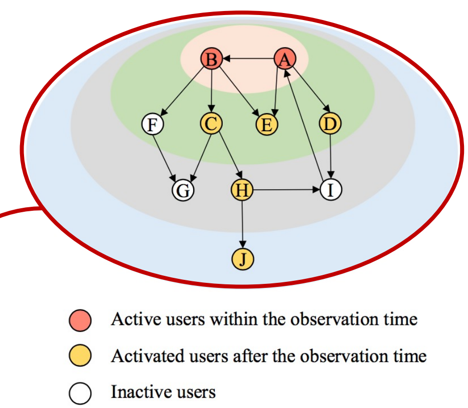
Methods	consider early adopters	Approximate by statistics	Precisely capture cascading effect
Feature-based	✓	✗	✗
Deep learning-based	✓	✗	✗
Point process	✗	✓	✗

Motivation

- Existing methods for popularity prediction mainly fall into three categories

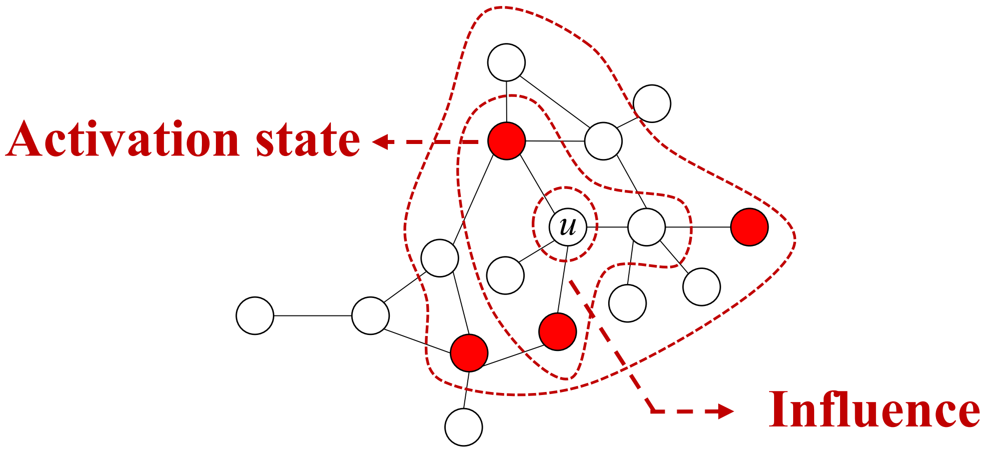
Methods	Only consider early adopters	Approximate by statistics	Precisely capture cascading effect
Feature-based	✓	✓	✗
Deep learning-based	✓	✗	✗
Point process	✗	✓	✗

- There still lack an effective method to **precisely capture the cascading effect** to predict the future popularity more accurately



Methods

- Cascading effect
  - The activation of a target user is **iteratively influenced by neighbors**.
  - The cascading effect is intrinsically the iterative interplay between **two key components**, i.e., **the state of neighbors** and **the spread of influence** over social networks.



Methods

- Framework of CoupledGNN
  - State Graph Neural Network
  - Influence Graph Neural Network

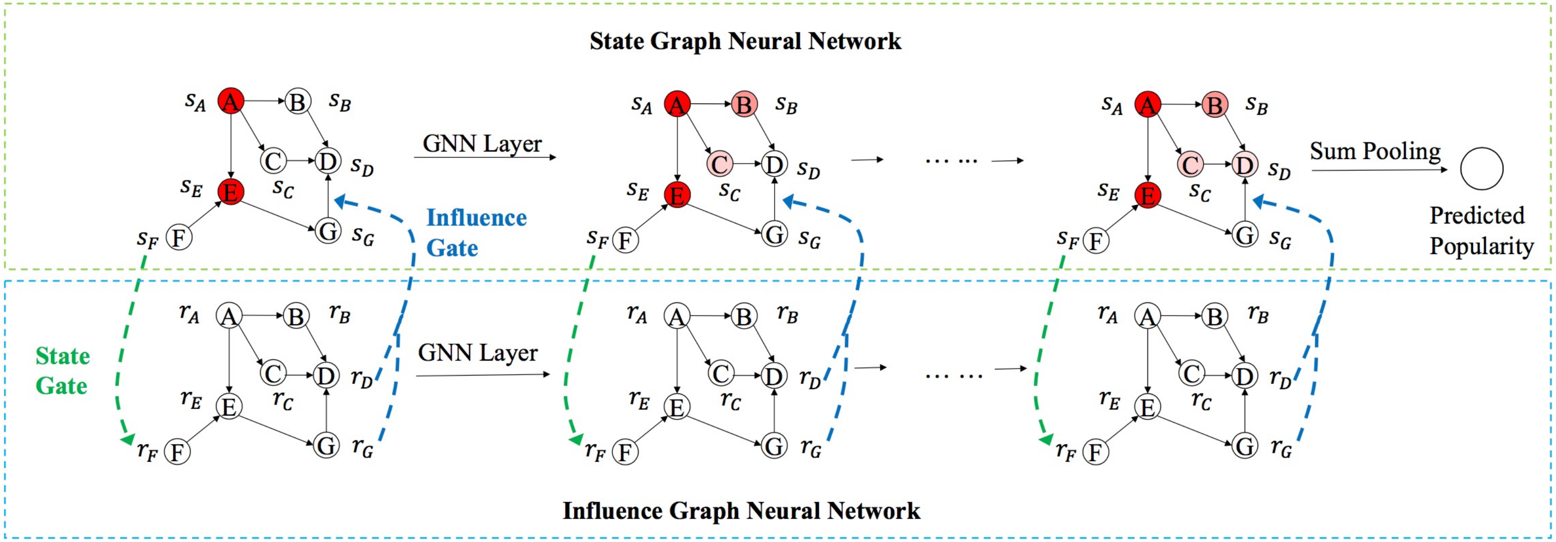


Figure 2: The framework of coupled graph neural networks for popularity prediction. s_* and r_* are the activation state and influence representation of user $*$ respectively.

Methods

- State Graph Neural Network
  - Model the activation state of user over network in a successive manner
    - Initial activation state

$$s_v^{(0)} = \begin{cases} 1, & v \in \boxed{C_T^m} \\ 0, & v \notin C_T^m \end{cases}$$

Active users within observation time

- Influence gating

$$\text{InfluGate}\left(\boldsymbol{r}_u^{(k)}, \boldsymbol{r}_v^{(k)}\right) = \boldsymbol{\beta}^{(k)}[\boldsymbol{W}^{(k)}\boldsymbol{r}_u^{(k)} \parallel \boldsymbol{W}^{(k)}\boldsymbol{r}_v^{(k)}],$$

Heterogeneous influence

- Aggregate

$$a_v^{(k)} = \boxed{\sum_{u \in \mathcal{N}(v)}} \text{InfluGate}\left(\boldsymbol{r}_u^{(k)}, \boldsymbol{r}_v^{(k)}\right) \boxed{s_u^{(k)}} + p_v,$$

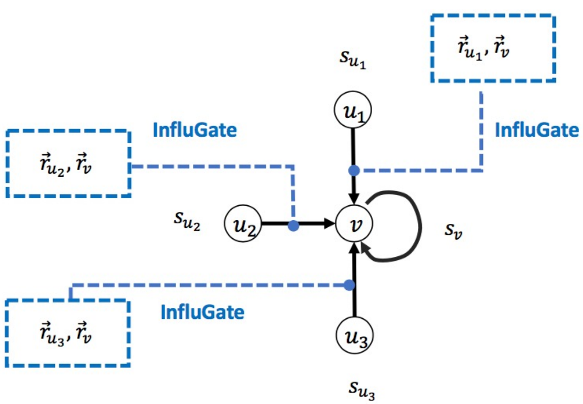
activation state

Expected activation state from neighborhood consider influence

- Combine

Update the activation state of user

$$s_v^{(k+1)} = \begin{cases} 1, & v \in C_T^m \\ \sigma\left(\mu_s^{(k)} s_v^{(k)} + \mu_a^{(k)} a_v^{(k)}\right), & v \notin C_T^m \end{cases}$$



Methods

- Influence Graph Neural Network
  - ◆ Model the diffusion of interpersonal influence in the social network
    - Initial influence representation $\boldsymbol{r}_v^{(0)}$
      - – node embedding & node features
    - Aggregate

$$\boldsymbol{b}_v^{(k)} = \boxed{\sum_{u \in \mathcal{N}(v)}} \text{StateGate}\left(s_u^{(k)}\right) \alpha_{uv}^{(k)} \boldsymbol{W}^{(k)} \boxed{\boldsymbol{r}_u^{(k)}},$$

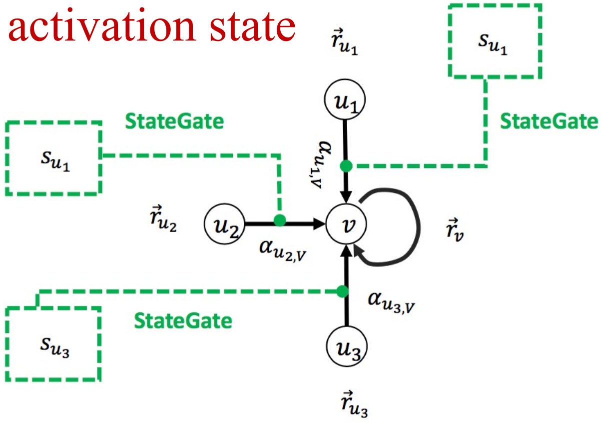
Influence representation

Expected influence representation from neighborhood gated by activation state

- Combine

$$\boldsymbol{r}_v^{(k+1)} = \boxed{\sigma\left(\zeta_r^{(k)} \boldsymbol{W}^{(k)} \boldsymbol{r}_v^{(k)} + \zeta_b^{(k)} \boldsymbol{b}_v^{(k)}\right)}$$

Update the influence representation of user



Methods

- Popularity prediction
  - Sum pooling mechanism

$$\hat{n}_{\infty}^{m} = \sum_{u \in \mathcal{V}} s_u^{(K)}$$

  - Loss function

$$L = L_{\mathrm{MRSE}} + L_{\mathrm{Reg}}$$

$$L_{\mathrm{MRSE}} = \frac{1}{M} \sum_{m=1}^{M} \left(\frac{\hat{n}_{\infty}^{m} - n_{\infty}^{m}}{n_{\infty}^{m}} \right)^2$$

$$L_{\mathrm{Reg}} = \eta \sum_{p \in \mathcal{P}} \| p \|_2 + \lambda L_{\mathrm{user}}$$

Experiments

- Datasets
  - Synthetic Data Set & Sina Weibo Data Set.
- Baselines
  - Feature-based (Cheng et al., WWW'14, Shulman et al., ICWSM'16, Gao et al., TKDD'19)
    - Features of early adopters
  - DeepCas (Li et al., WWW'17)
    - Representative deep learning based method
  - SEISMIC (Zhao et al., KDD'15)
    - Representative point process method
- Evaluation metrics
  - Mean/Median Relative Square Error (MRSE)
  - Mean Absolute Percentage Error (MAPE)
  - Wrong Percentage Error (WroPerc)

$$MRSE = \frac{1}{M}\sum_{m=1}^{M}\left(\frac{\hat{n}_\infty^m - n_\infty^m}{n_\infty^m}\right)^2.$$

$$MAPE = \frac{1}{M}\sum_{m=1}^{M}\frac{|\hat{n}_\infty^m - n_\infty^m|}{n_\infty^m}.$$

$$WroPerc = \frac{1}{M}\sum_{m=1}^{M}\mathbb{I}\left[\frac{|\hat{n}_\infty^m - n_\infty^m|}{n_\infty^m} \geq \epsilon\right].$$

$\hat{n}_\infty^m$: predicted popularity

n_∞^m : true popularity

M : number of information

Evaluation metrics

Experiments

- Q1: How does CoupledGNN performs on popularity prediction?
  - CoupledGNN **outperforms all the baselines with different observation time**.
  - The **longer** the **observation time** is, the **smaller** the **prediction errors** are.

Table 1: Popularity prediction in Sina Weibo

Observation Time	1 hour				2 hours				3 hours			
Evaluation Metric	MRSE	mRSE	MAPE	WroPerc	MRSE	mRSE	MAPE	WroPerc	MRSE	mRSE	MAPE	WroPerc
SEISMIC	-	0.2112	-	48.63%	-	0.1347	-	34.59%	-	0.0823	-	27.15%
Feature-based	0.2106	0.1254	0.3749	35.17%	0.1796	0.1041	0.3557	28.86%	0.1581	0.0804	0.3147	18.97%
DeepCas	0.2077	**0.0930**	0.3633	30.00%	0.1650	0.0670	0.3134	20.55%	0.1365	0.0361	0.2813	17.24%
CoupledGNN	**0.1816**	0.0946	**0.3515**	**25.68%**	**0.1397**	**0.0519**	**0.2989**	**17.81%**	**0.1120**	**0.0333**	**0.2611**	**13.01%**

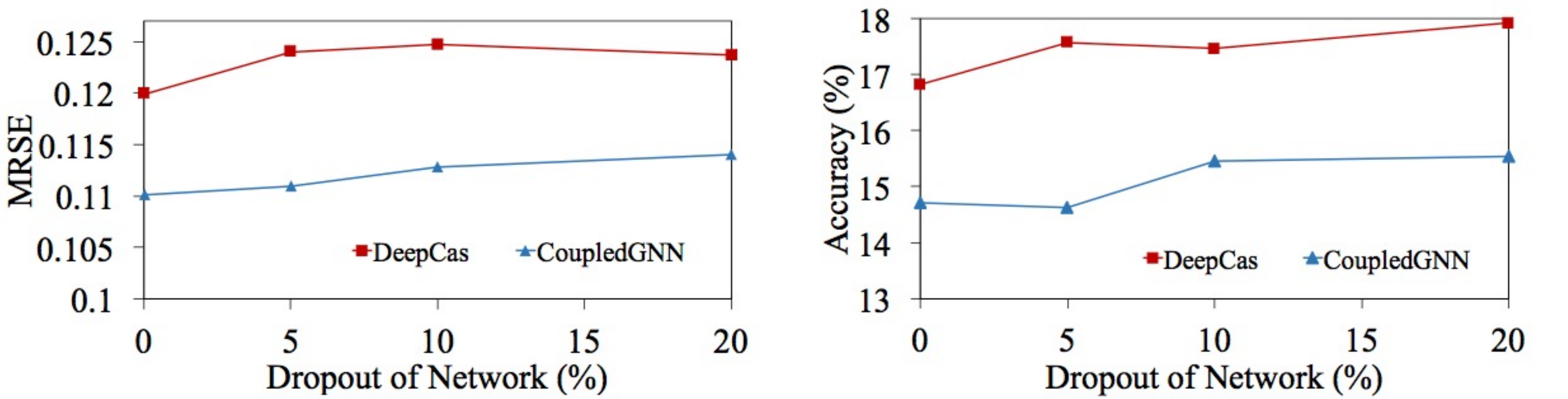
Experiments

- Q2: Is the coupled structure of our model effective?
  - **The coupled structure** of activation state and influence representation (CoupledGNN) **significantly improves the prediction performance** compared with the Single-GNN methods

Observation Time	1 hour			2 hour			3 hour		
Evaluation Metric	MRSE	MAPE	WroPerc	MRSE	MAPE	WroPerc	MRSE	MAPE	WroPerc
Single-GCN	0.1964	0.3707	29.11%	0.1595	0.3201	22.26%	0.1230	0.2653	16.10%
Single-GAT	0.1999	0.3754	30.82%	0.1569	0.3199	20.55%	0.1222	0.2655	16.10%
CoupledGNN	**0.1816**	**0.3515**	**25.68%**	**0.1397**	**0.2989**	**17.81%**	**0.1120**	**0.2611**	**13.01%**

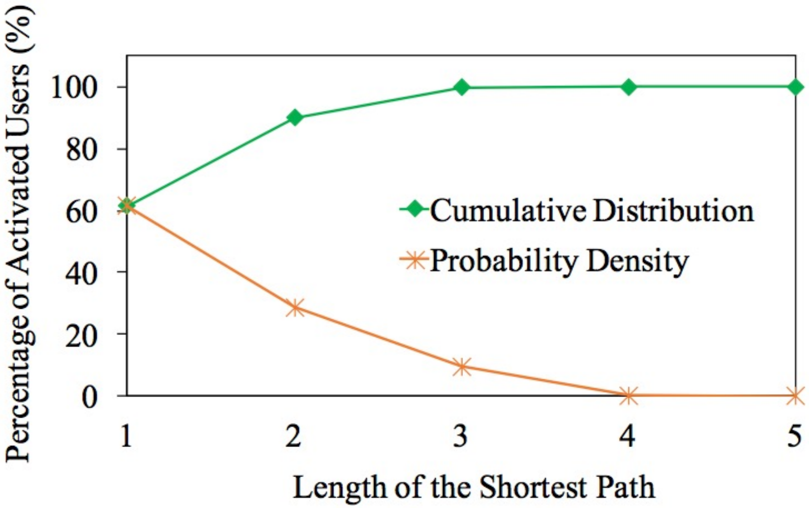
Experiments

- Q3: To what extent will a partial lack of network influences the prediction performance?
  - The **prediction performance** of both methods is **slightly degraded** by the dropout of the network
  - **CoupledGNN always significantly performs better** than DeepCas



Experiments

- Q4: Whether the learned CoupledGNN model can cover the scope of the cascading effect?
  - The **optimal number of layers** in CoupledGNN $K = 3$
  - **Almost all the activated users by cascading effect**, i.e., 99.76%, can be **covered within three-hops** in the neighborhood of early adopters.

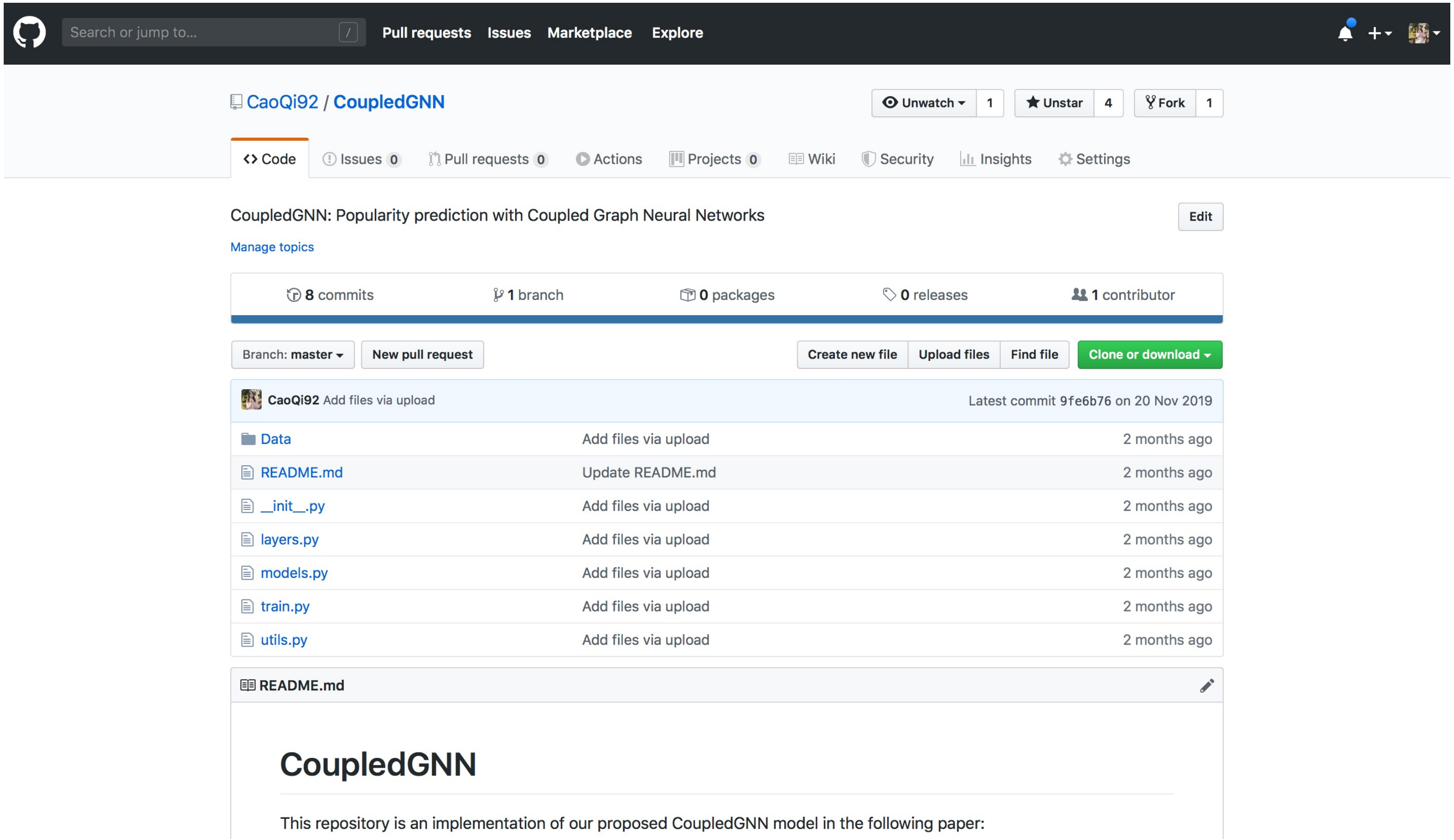


Conclusion

- We propose **CoupledGNN** model to precisely capture the cascading effect, improving the prediction performance of future popularity
  - View the cascading effect as the **iterative interplay** between two key components, i.e., **the state of neighbors** and **the spread of influence** over social networks.
  - Model the iterative interplay between node states and the spread of influence by **two graph neural networks** which **coupled** by gating mechanism

Public of Source Code

- We have made the source code of CoupledGNN public on github.



Contact Information

Thank you!
Welcome to contact us!

Qi Cao
caoqi@ict.ac.cn

Huawei Shen
shenhuawei@ict.ac.cn