

Beyond Personalization: Social Content Recommendation for Creator Equality and Consumer Satisfaction Wenyi Xiao*, Huan Zhao*, Haojie Pan*, Yangqiu Song*, Vincent W. Zheng⁺, Qiang Yang[†]* *Hong Kong University of Science and Technology, Hong Kong

⁺WeBank, China



INTRODUCTION

MCTS BASED SOCIAL EXPLORATION

RESULTS

Content Recommendation



The volume of news is overwhelming to users...

It's critical to help users select the interesting contents in real time!



The selection strategy is defined by:

		Steemit	-English			Steemit	-Spanish	
	AUC	F1	Gini	C&C	AUC	F1	Gini	C&C
NCF	52.83 ± 0.13	42.14 ± 0.21	$66.04 {\pm} 0.25$	$37.71 {\pm} 0.22$	$50.46 {\pm} 0.21$	$35.02 {\pm} 0.26$	58.13 ± 0.34	38.14 ± 0.2
SAMN	53.05 ± 0.35	42.28 ± 0.45	65.98 ± 0.21	$37.80 {\pm} 0.28$	51.10 ± 0.24	35.24 ± 0.31	58.29 ± 0.32	38.20 ± 0.3
LR	$52.89 {\pm} 0.07$	34.50 ± 0.11	62.89 ± 0.11	$35.86 {\pm} 0.11$	53.15 ± 0.06	$36.50 {\pm} 0.29$	$55.84 {\pm} 0.09$	39.97 ± 0.1
LibFM	50.01 ± 0.12	40.43 ± 0.22	66.42 ± 0.13	36.79 ± 0.16	$47.71 {\pm} 0.30$	22.37 ± 0.33	56.50 ± 0.21	29.55 ± 0.2
DKN	$62.71 {\pm} 0.22$	42.85 ± 0.45	62.29 ± 0.26	40.22 ± 0.33	57.02 ± 0.39	41.27 ± 0.45	$53.98{\pm}0.25$	43.52 ± 0.3
SEAN	$65.57{\pm}0.17$	$47.69{\pm}0.46$	$61.78{\pm}0.24$	$42.43{\pm}0.33$	$59.98{\pm}0.34$	$42.99{\pm}0.37$	53.99 ± 0.23	$44.46{\pm}0.5$

> Comparison of social exploration methods

Previous RS VS. Current RS.





• User = Consumer

- User = Consumer
- Editor = Creator
- Only consider consumer
- User = Creator • Consider creator & consumer
- Measure on Steemit





$a = \operatorname{argmax}_{v} \{ Q_t(v) + \lambda \cdot U_t(v) \},\$

 $Q_t(v)$ denotes the exploitation reward of node v at time t $U_t(v)$ is the utility to explore node v λ is used to balance the two terms



- Goal: Select the nodes who
- have less been explored
- $N_t(c_l)$: Times the current node c_1 has been selected
- $N_t(v)$: Times the neighbor node v has been selected

> Exploitation

Four strategy for $Q_t(v)$:

- Average F1 evaluated based on RS model
- PageRank value obtained from social network
- PageRank value obtained from dynamic activity network
- Payout that a user gains from platform

Variants	F1	Gini	C&C
w/o social	42.40±0.30	58.56±0.43	41.91±0.35
w/o social attention	44.79±0.17	62.22±0.36	41.98±0.23
one-hop friends	43.08±0.16	60.85±0.25	41.04 ± 0.20
w/o CNN	45.25±0.22	59.98±0.26	42.58±0.21
w/o GRU	45.07±0.31	60.47 ± 0.14	42.12 ± 0.27
w/o CNN & GRU	44.08±0.26	60.06±0.20	41.91±0.25
SEAN	47.69±0.46	61.78±0.24	42.43±0.33

> Comparison of different variants

Models	F1	Gini	C&C
Random Select	$42.48 {\pm} 0.38$	$59.13{\pm}0.22$	$41.09 {\pm} 0.28$
Random Walk	45.05 ± 0.39	$60.98 {\pm} 0.09$	$41.77 {\pm} 0.20$
SEAN-RS-F1	$47.69{\pm}0.46$	$61.78 {\pm} 0.24$	42.43 ± 0.33
SEAN-SPR	$45.99 {\pm} 0.35$	$60.90 {\pm} 0.32$	42.27 ± 0.33
SEAN-DPR	$45.96 {\pm} 0.44$	$60.98 {\pm} 0.22$	42.21 ± 0.29
SEAN-Payout	46.26 ± 0.36	$60.65 {\pm} 0.40$	$42.53{\pm}0.37$

Parameter Sensitivity



(b) Beam width B.

MODEL FRANMWORK



- The left side is a social exploration module which explores high-order friends.
- The right side is a hierarchical architecture from CNN layer to encode words to GRU layer to encode sentences in the document.



	${f Steemit-English}$	Steemit-Spanish
Duration (days)	370	126
# Consumers	$7,\!242$	$1,\!396$
# Creators	44,220	4,073
# Relations	$273,\!942$	$25,\!593$
# Documents	$177,\!134$	$14,\!843$
Avg. word per document	290	509
# Logs	$220,\!909$	$20,\!893$
# Samples	684,752	$92,\!236$



(c) Trade-off λ .

CONCLUSIONS

1. Consider both content creators and consumers Gini index to measure the inequality 2. Social explorative attention Network (SEAN) Incorporate users' high-order friends Social exploration mechanism based on MCTS 3. Construct two datasets from Steemit A Large dataset in English

A small dataset in Spanish





Creator Equality: Gini index

Both consumer and creator: harmonic mean





• email] wxiaoae@cse.ust.hk

• [codes]https://github.com/HKUST-KnowComp/Social-Explorative-Attention-Networks