From Local to Global Norm Emergence

——Dissolving Self-reinforcing Substructures with Incremental Social Instruments

Outline

1. Multi-agent system and norm
2. Social learning and SRS
3. BA-ratio
4. Experiments

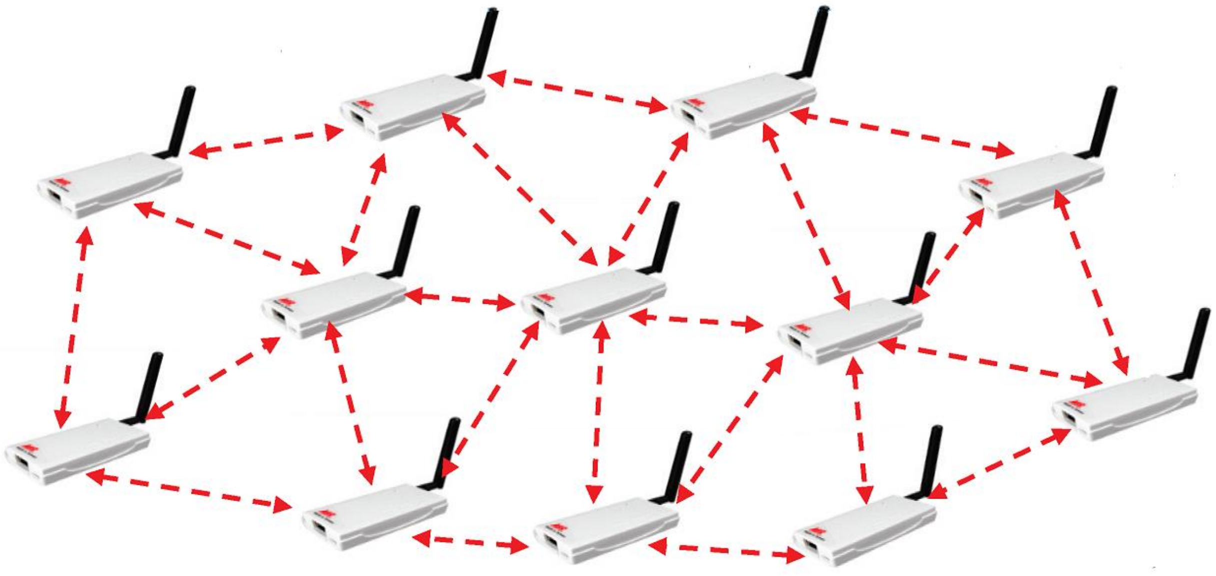
1.1 Multi-Agent System

- **Agent:** any entity that can autonomously act in its environment is called an agent.

1.1 Multi-Agent System

- **Multi-Agent System:** A multi-agent system is a loosely coupled network of agents that work together to solve problems that are beyond the capabilities or knowledge of individual agents.

- **Example:** Wireless Sensor Networks(WSN)



Sensor networks

1.1 Convention

- Pure coordination game is described as the problem where all agents need to select the same action to avoid conflict.

Table 1: Payoff matrix of the Two-lane road game

	left	right
left	(1,1)	(0,0)
right	(0,0)	(1,1)

- A **convention** in a MAS is a behavior that is common among agents, e.g. driving either on the right side or the left side of the road.

Remark. If all agents have learned to select the same action at every step in repeated pure coordination games, they belong to a convention.

1.2 Norm emergence

- Norm emergence is a kind of convention in social network. Social norms arise naturally through the interactions between individuals.
- Norm emergence problem can be modeled as a pure coordination game confined to a social topology $G = (N, E)$. Each pair of entities in E paly a 2-player-m-action coordination game.

Table 2: 2-player-m-action coordination game

	a_1	a_2	⋯	a_m
a_1	(1,1)	(-1,-1)	⋯	(-1,-1)
a_2	(-1,-1)	(1,1)	⋯	(-1,-1)
⋯	⋯	⋯	⋯	⋯
a_m	(-1,-1)	(-1,-1)	⋯	(1,1)

2.1 Social learning

- Sen and Airiau proposed the social learning mechanism. This is a simple yet powerful paradigm for norm emergence.

- A number of algorithms have been used to implement the social learning paradigm. e.g., Q-Learning, WoLF-PHC, and Fictitious play (FP). Among them, Q-learning is the most widely used algorithm due to its efficiency and effectiveness.

2.1 Q-learning

- The Q-learning algorithm estimates the quality of agent's actions in each state in order to derive an optimal policy. The quality (also called **Q-value**) of an action in a given state indicates how good (or bad) the action is in that particular state.

- Given a graph $G = (V, E)$, the Q-value will be updated by a pairwise interactions $\{v_i, v_j\} \in E$. The updated equation of action a_j for agent v_i is:

$$Q_i(a_j) := (1-\lambda)(Q_i(a_j)) + \lambda r$$

where r is the reward in a 2-player m-action coordination game.

2.1 Q-learning

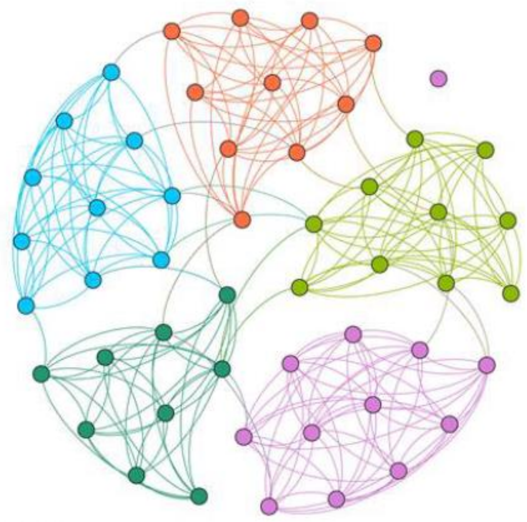
- The belief state b_i of v_i is the action a_j that has the maximum Q-value, i.e.,

$$b_i = \arg\max_{a_j} Q_i(a_j)$$

- If all agents have the same belief a_j in the subsequent pure coordination games, we say that they form a global norm.

2.2 SRS

- Despite the general belief that social learning prepares the way for norm emergence, there are situations where a norm fails to emerge.
- Sub-conventions or local norms are obstacles to global norm. Villatoro et al. refers to regions of a network that can maintain a sub-convention as Self Reinforcing Substructures (SRS).



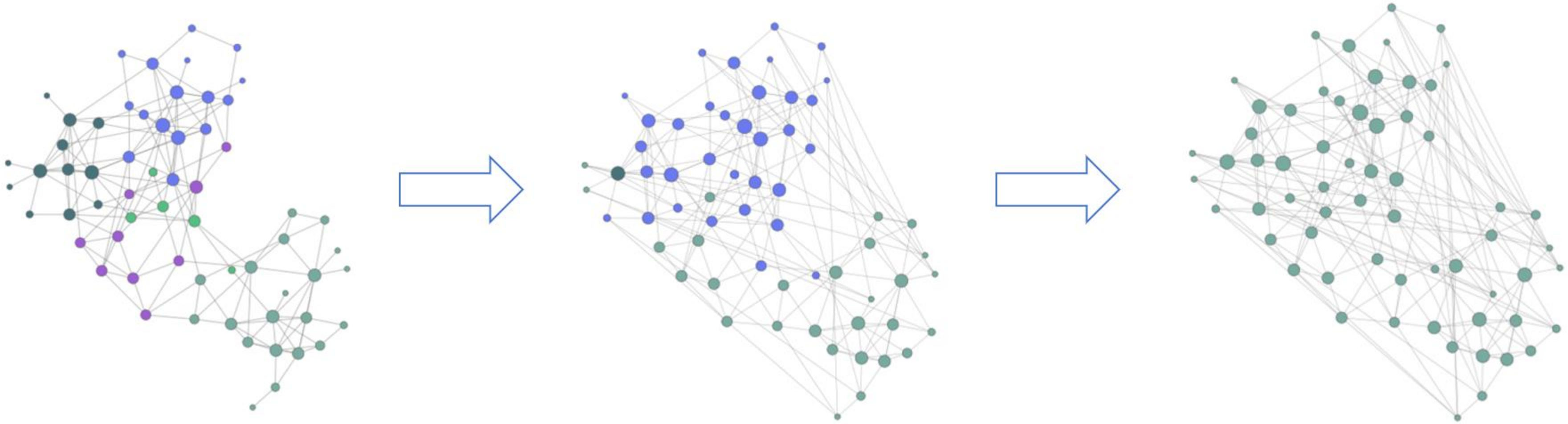
2.2 SRS

- To digest these failures of norm emergence, most of the literature assume that a norm is reached when at least 90% of the population select the same action. However,

  a) Villatoro et al argue that a threshold of 90% is not sufficient to say that agents' behavior has converged;
  b) even 90% relaxation might be too strong when the network maintains a sub-convention as SRS.

2.3 Problem statement

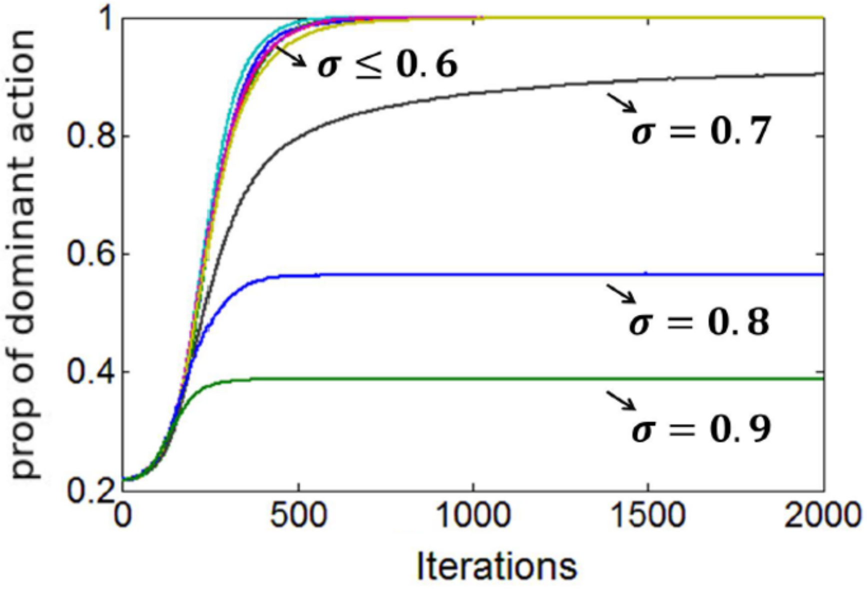
- Research has corroborated that network topology or structure has a significant impact on norm emergence.



2.3 Problem statement

Interestingly, a fully connected network does not exhibit any SRSs and global norm has been shown to always emerge.

- **Problem:** How the designer (central organization) makes extra connections (incremental social instrument) on the existing network to dissolve SRS?



3.1. A-entropy

Agent interaction network: $G = (N, E)$
$N = \{v_1, v_2, \cdots, v_n\}$
$E = \{\{v_i, v_j\} | when\ there\ is\ a\ social\ link\ between\ v_i\ and v_j\}$

- The probability of agent v_i participating in the interaction is: $\frac{d_i}{2|E|}$.
- The information of agent in their interaction is defined as A-entropy :

$$\mathcal{H}_a(G) = -\sum_{v_i \in N} \frac{d_i}{2|E|} \log_2 \frac{d_i}{2|E|}$$

Pairwise agent interaction sequence
$\{x_t, y_t\} \in E$

$\{\boldsymbol{x_1}, \boldsymbol{y_1}\}$
$\{\boldsymbol{x_2}, \boldsymbol{y_2}\}$
$\{\boldsymbol{x_3}, \boldsymbol{y_3}\}$
…
…
$\{\boldsymbol{x_t}, \boldsymbol{y_t}\}$
…

3.1. B-entropy

- The probability of belief state a_j participating in the interaction is:

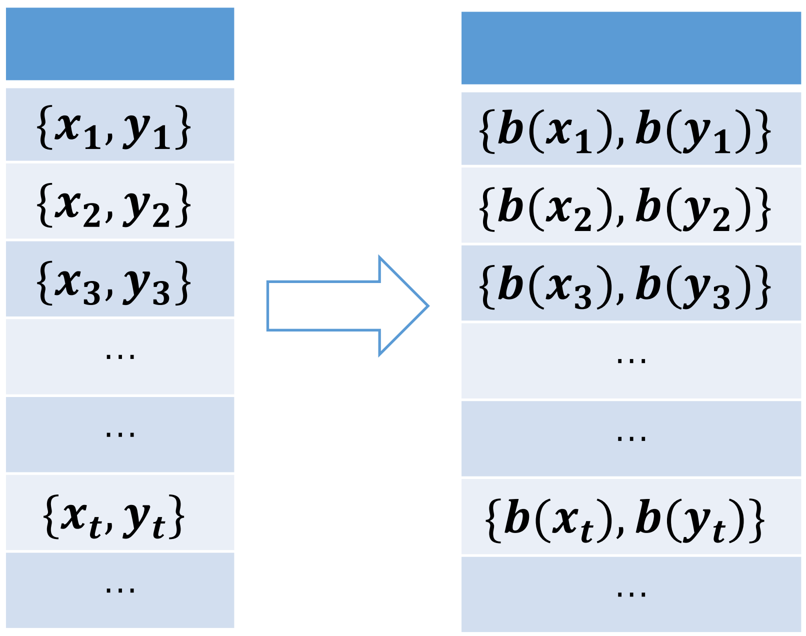
$$\Pr(a_j) = \sum_{b_i = a_j} \frac{d_i}{2|E|} = \sum_{v_i \in N_j} \frac{d_i}{2|E|} = \frac{\nu_j}{2|E|}$$

- The average information content of agents' beliefs in their interaction is defined as B-entropy :

$$\mathcal{H}_b^{\mathcal{P}}(G) = -\sum_{a_j \in A} \frac{\nu_j}{2|E|} \log_2 \frac{\nu_j}{2|E|}$$

- The distribution of belief state after social learning also reflects the structural information of the network.

Pairwise belief interaction sequence



$b(x_t)$ is the belief state of agent x_t.

3.2 BA-ratio

BA-ratio is defined as:

$$\rho_G(\mathcal{P}) := \mathcal{H}_b^{\mathcal{P}}(G)/\mathcal{H}_a(G)$$

- The BA-ratio reflects the level of diversity of the agents preferred actions
- If $\rho_G(\mathcal{P}) = 1$, then each agent holds its own belief; also, if $\rho_G(\mathcal{P}) = 0$ then all agents reach a consensus action.

3.3 Adaptive ISI

One way to dissolve SRS---minimizing BA-ratio

- Reduce the diversity of the agents preferred actions
- Reduce the amount of structural information of the network
- Increase the interaction between agents of different belief states

Theorem 2 (Small-Degree Principle). *Consider two NIAI pairs* $e = \{v_i, v_k\}, e' = \{v_{i'}, v_{k'}\} \in N_j \otimes N_\ell$.

(1) *If* $min\{d_i, d_k\} \leq min\{d_{i'}, d_{k'}\}$ *and* $max\{d_i, d_k\} \leq max\{d_{i'}, d_{k'}\}$, *then* $\rho_{G \oplus \{e\}}(\mathcal{P}) \leq \rho_{G \oplus \{e'\}}(\mathcal{P})$.

(2) *If* $e = \{v_i, v_k\}$ *is BA-ratio minimizing, then* v_i *has the smallest degree among those nodes in* N_j *that are not interacting with* v_k.

3.3 Adaptive ISI

Algorithm 1 Incremental Social Instrument (ISI)

Input: $G = (V, E), N = (N_1, N_2, ..., N_m)$
Output: a NIAI pair $\{u, v\}$

$\rho_{min} = 1$;
for $(j, \ell) \in \{1, \ldots, m\}^2$ where $\ell > j$ **do**
  $L_{j\ell} \leftarrow$ sort $N_j \cup N_\ell$ into non-decreasing degree order, $s \leftarrow 1, t \leftarrow |L_{j\ell}|$;
  while $s < t$ **do**
    for $k = s + 1 \rightarrow t$ **do**
      $e \leftarrow \{L_{j\ell}[s], L_{j\ell}[k]\}$;
      if $\delta(L_{j\ell}[s], L_{j\ell}[t]) = 0$ & $\rho_{G \oplus e} < \rho_{min}$ **then**
        $u \leftarrow L_{j\ell}[s], v \leftarrow L_{j\ell}[k]$;
        $\rho_{min} = \rho_{G \oplus e}, t = k - 1$; **Break**;
      end if
    end for
    $s \leftarrow s + 1$;
  end while
end for
RETURN $\{u, v\}$

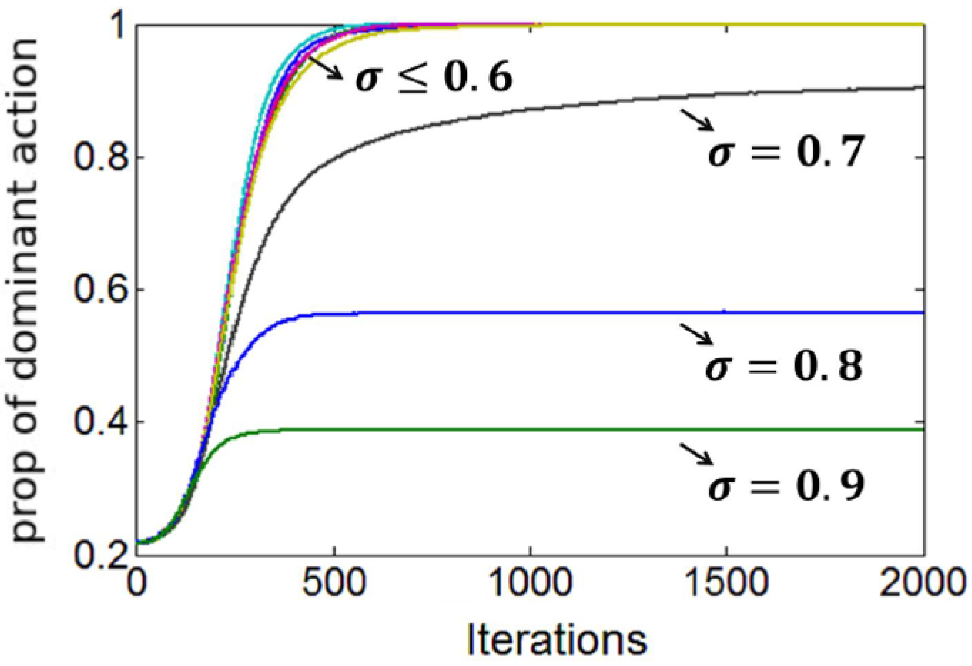
Algorithm 3 Adaptive Incremental Social Instrument (AISI)

Input: Social topology $G = (N, E)$, local norm partition $\mathcal{P} = \{N_1, \ldots, N_m\}$, budget $h \in \mathbb{N}$
Output: A sequence of h non-interacting pairs $e_1, e_2, \ldots$

while there exists v_i, v_k such that $b_i \neq b_k$ and $j = 1 \rightarrow h$ **do**
  Find an edge e by Alg.1 with respect to the current social topology G and partition $\mathcal{P}$;
  Update $G = G \oplus e$ and output $e_j = e$;
  Run the social learning process until the agents stabilize and record the current partition as the new $\mathcal{P}$;
end while

4.1 Experiment setup

- We set the number of iterations of social learning to be 5000 .



GRP setting: g = 50
l = 20
k = 20
m=10

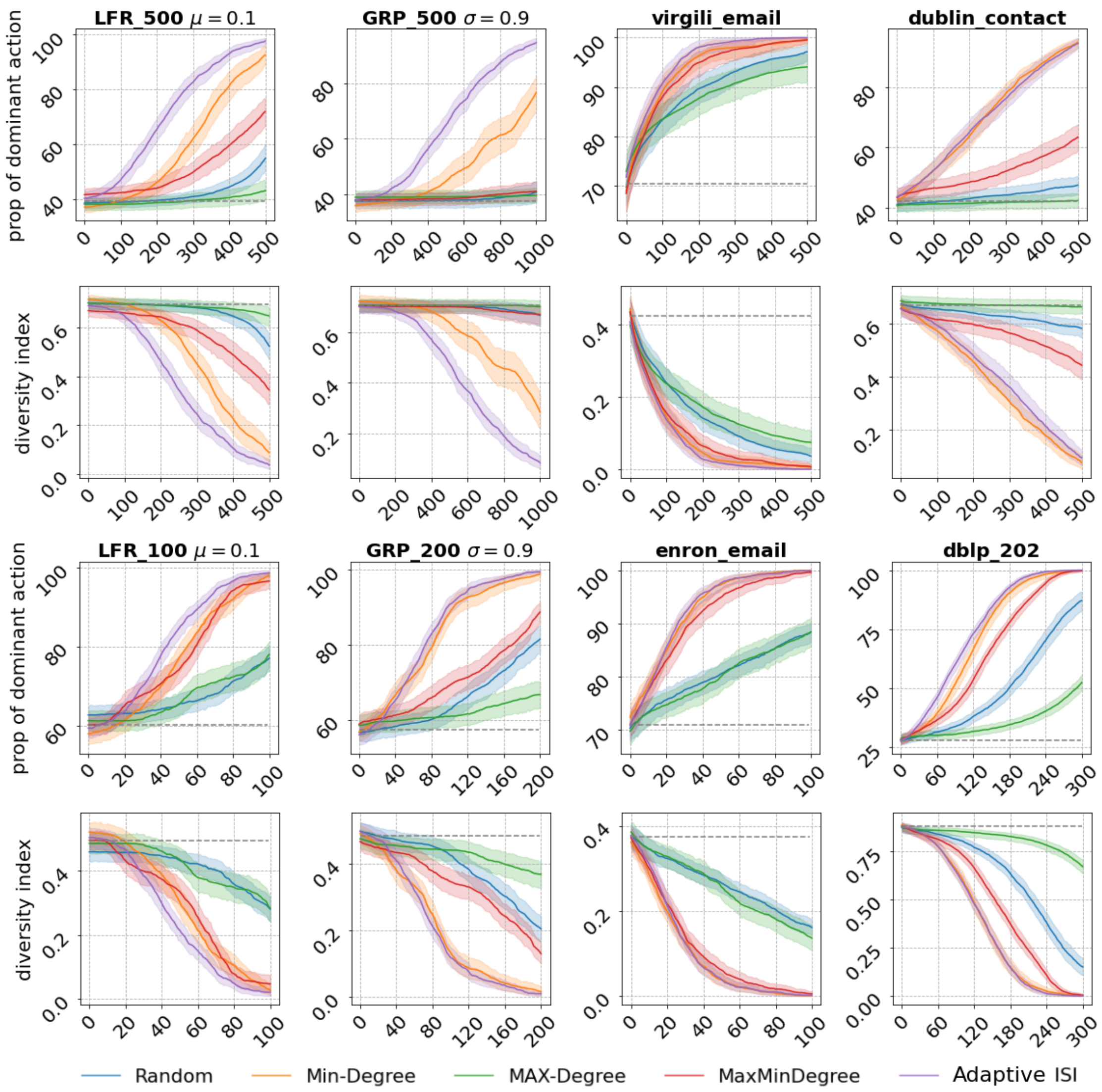
4.1 Experiment setup

Evaluation

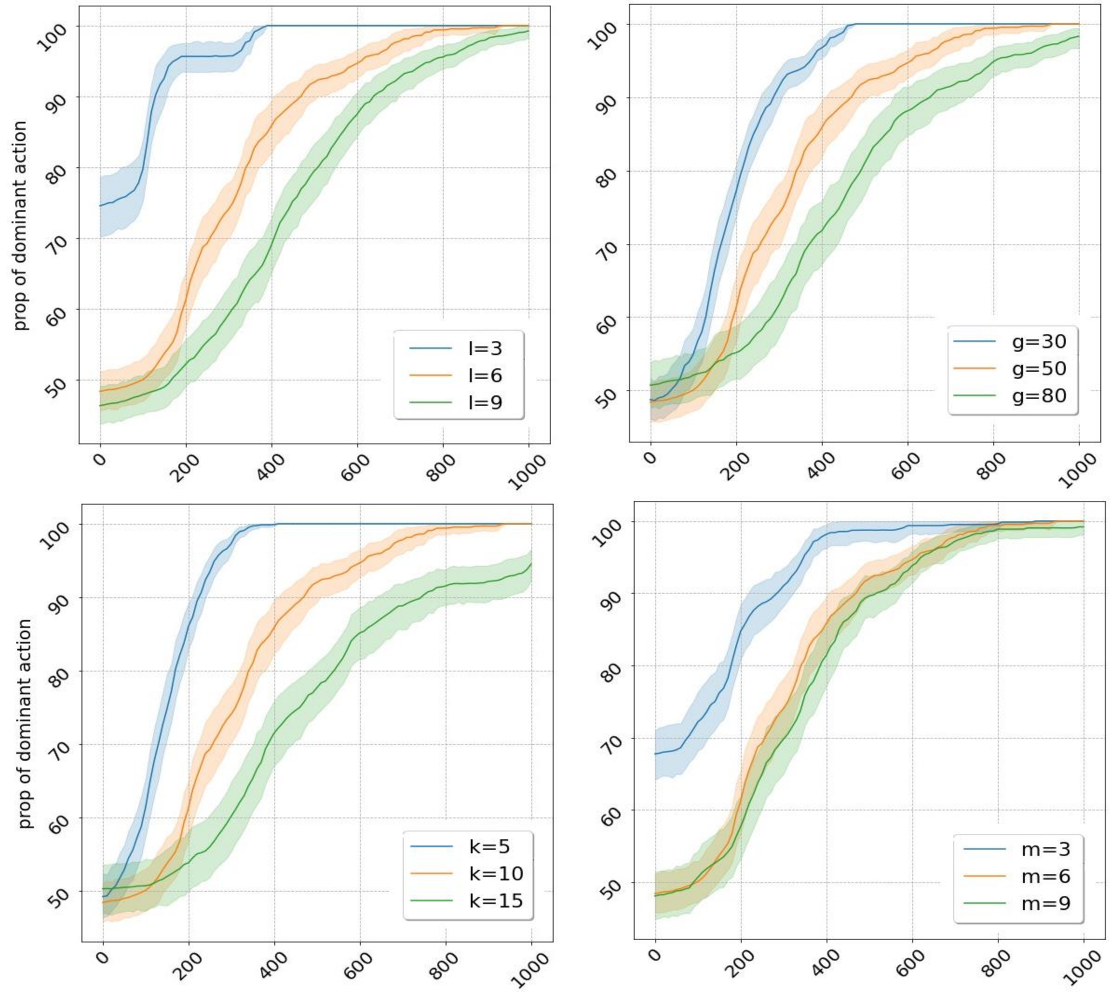
Input: $G = (V, E)$, An edge-created strategy $\mathcal{S}$;
Output: The proportion of the dominant action $\mathcal{R}$;
1. $\mathcal{R} \leftarrow$ Applying Q-learning on $G = (V, E)$;
2. While $\mathcal{R} < 100\%$
    $G \leftarrow$ Applying the strategy $\mathcal{S}$ on G;
    $\mathcal{R} \leftarrow$ Applying Q-learning on G;
    Output $\mathcal{R}$;
  End

Type	Graph	n	$\|E\|$	m	Modularity
GRP	grp_200_0.9	200	468	5	0.726
	grp_500_0.9	500	2408	8	0.789
LFR	lfr_100_0.1	100	302	5	0.638
	lfr_500_0.1	500	1411	8	0.788
RN	enron_email	143	623	7	0.568
	virgili_email	1133	5451	10	0.572
RN	dblp_202	202	387	5	0.510
	dublin_contact	410	2765	7	0.711

4.2 Experimental results



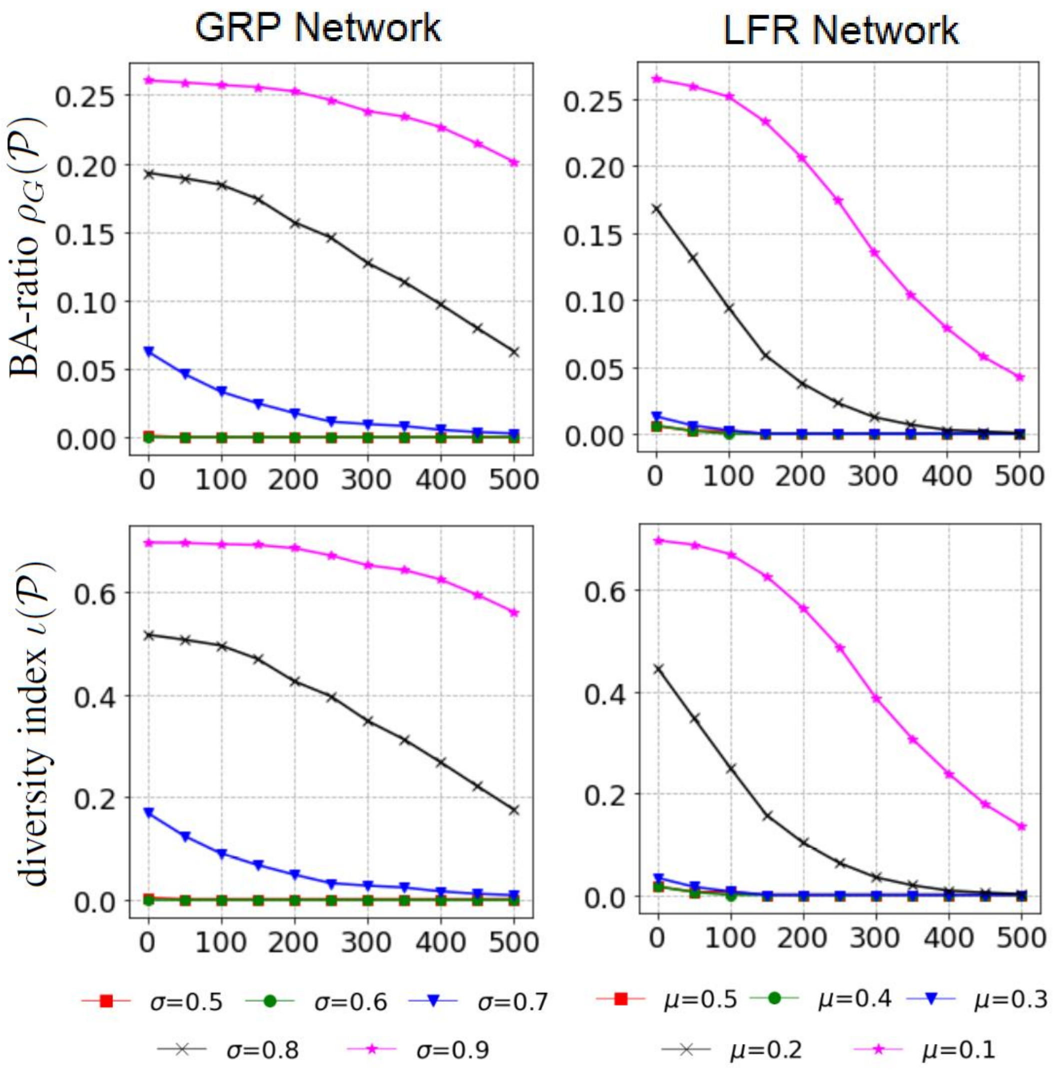
4.2 Experimental results



The reference network are generated by set l=6; g=50; k=10; m=6.

l: the number of communities
g: community size
k: average degree
m: the number of available action

4.2 Experimental results



Diversity index:

$$\iota(\mathcal{P}) := \frac{1}{\log_2 n} \sum_{1 \le j \le m} -\frac{|N_j|}{n} \log_2 \frac{|N_j|}{n}$$

Type	Parameter	n	$\lvert E \rvert$	m	Modularity
GRP	$\sigma = 0.9$	500	2408	10	0.789
	$\sigma = 0.8$	500	2465	10	0.692
	$\sigma = 0.7$	500	2472	10	0.596
	$\sigma = 0.6$	500	2565	10	0.495
	$\sigma = 0.5$	500	2524	10	0.398
LFR	$\mu = 0.1$	500	1411	9	0.789
	$\mu = 0.2$	500	1378	11	0.670
	$\mu = 0.3$	500	1571	8	0.465
	$\mu = 0.4$	500	1491	7	0.331
	$\mu = 0.5$	500	1362	9	0.300

5 Conclusion

- In this paper, we define the concept of BA-ratio. This is the first work that aims to bring information-theoretic argument on the bottleneck of norm emergence.

- We further prove the small-degree principle in minimizing BA-ratio at the process of creating links. Based on this principle, we design an adaptive ISI to dissolve SRS. Experiments verify the effectiveness of the algorithm.