Social Networks

Mayur Mohite[1]

E-Commerce Lab
Department of Computer Science and Automation
Indian Institute of Science

Lab Talk

[1] The author is fully supported by MHRD scholarship and is sincerely thankful to Mr. Arjun Singh for increasing MHRD Scholarship from INR 5000 to 8000 last year

Outline

Statistical Properties of Community Structure in Large Social and Information Networks
By Leskovec et al appeared in WWW08

Network Community Profile Plot

- Conductance of some $S \subseteq V$ is given by-

$$\phi = s/v$$

Where, v is the sum of the degrees of nodes in S and s is the number of edges that cross the cut S in the graph

- Now for a given size k the best conductance value for this size in the entire network is given by-

$$\Phi(k) = \min_{S \subset V, |S|=k} \phi(S)$$

- *NCP* plot measures the quality of best possible community as a function of size of the community

Network Community Profile Plot

- Conductance of some $S \subseteq V$ is given by-

$$\phi = s/v$$

  Where, v is the sum of the degrees of nodes in S and s is the number of edges that cross the cut S in the graph

- Now for a given size k the best conductance value for this size in the entire network is given by-

$$\Phi(k) = \min_{S \subset V, |S|=k} \phi(S)$$

- *NCP* plot measures the quality of best possible community as a function of size of the community

Network Community Profile Plot

- Conductance of some $S \subseteq V$ is given by-

$$\phi = s/v$$

  Where, v is the sum of the degrees of nodes in S and s is the number of edges that cross the cut S in the graph

- Now for a given size k the best conductance value for this size in the entire network is given by-

$$\Phi(k) = \min_{S \subset V, |S| = k} \phi(S)$$

- *NCP* plot measures the quality of best possible community as a function of size of the community

Network Community Profile Plot

- Conductance of some $S \subseteq V$ is given by-

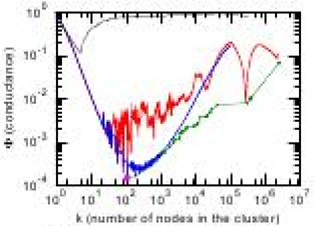
$$\phi = s/v$$

  Where, v is the sum of the degrees of nodes in S and s is the number of edges that cross the cut S in the graph
- Now for a given size k the best conductance value for this size in the entire network is given by-

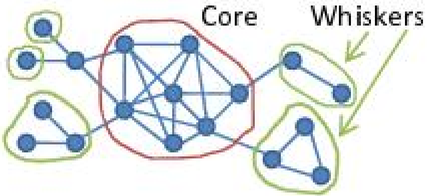
$$\Phi(k) = \min_{S \subset V, |S|=k} \phi(S)$$

- *NCP* plot measures the quality of best possible community as a function of size of the community

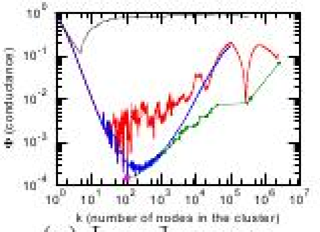
An Example NCP and General Graph Structure for large Social Networks



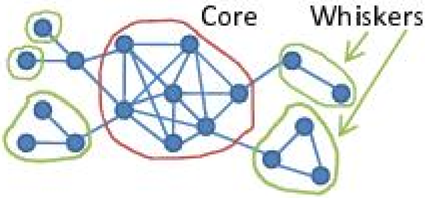
(a) LIVEJOURNAL



An Example NCP and General Graph Structure for large Social Networks



(a) LiveJournal



Observations

- Upto size of around 100 nodes the slope of NCP plot is usually sloping downward indicating good quality of communities
- Beyond the size of 100 nodes the NCP plot slopes upward indicating that quality of communities worsen with increasing size
- Does large sized communities really exist?
- Current random network models do not exhibit this behaviour except for the forest fire[2] model

[2] *Graph Evolution: Densification and shrinking diameters* by Leskovec et al in TKDD 07

Observations

- Upto size of around 100 nodes the slope of NCP plot is usually sloping downward indicating good quality of communities
- Beyond the size of 100 nodes the NCP plot slopes upward indicating that quality of communities worsen with increasing size
- Does large sized communities really exist?
- Current random network models do not exhibit this behaviour except for the forest fire[2] model

[2] *Graph Evolution: Densification and shrinking diameters* by Leskovec et al in TKDD 07

Observations

- Upto size of around 100 nodes the slope of NCP plot is usually sloping downward indicating good quality of communities
- Beyond the size of 100 nodes the NCP plot slopes upward indicating that quality of communities worsen with increasing size
- Does large sized communities really exist?
- Current random network models do not exhibit this behaviour except for the forest fire[2] model

[2] *Graph Evolution: Densification and shrinking diameters* by Leskovec et al in TKDD 07

Observations

- Upto size of around 100 nodes the slope of NCP plot is usually sloping downward indicating good quality of communities
- Beyond the size of 100 nodes the NCP plot slopes upward indicating that quality of communities worsen with increasing size
- Does large sized communities really exist?
- Current random network models do not exhibit this behaviour except for the forest fire[2] model

[2] *Graph Evolution: Densification and shrinking diameters* by Leskovec et al in TKDD 07

Observations

- Upto size of around 100 nodes the slope of NCP plot is usually sloping downward indicating good quality of communities
- Beyond the size of 100 nodes the NCP plot slopes upward indicating that quality of communities worsen with increasing size
- Does large sized communities really exist?
- Current random network models do not exhibit this behaviour except for the forest fire[2] model

[2] *Graph Evolution: Densification and shrinking diameters* by Leskovec et al in TKDD 07

Group Formation in Large Social Networks: Membership, Growth , Evolution
By Backstrom et al appeared in KDD06

Community membership

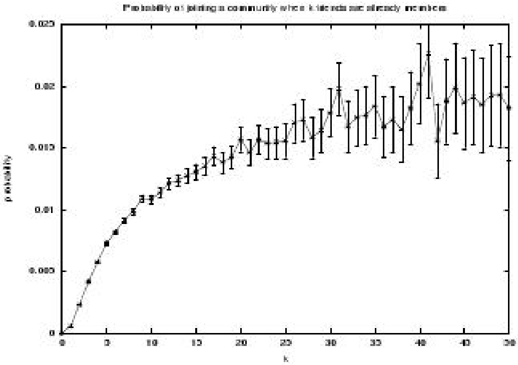


Figure 1: The probability p of joining a LiveJournal community as a function of the number of friends k already in the community. Error bars represent two standard errors.

Community membership

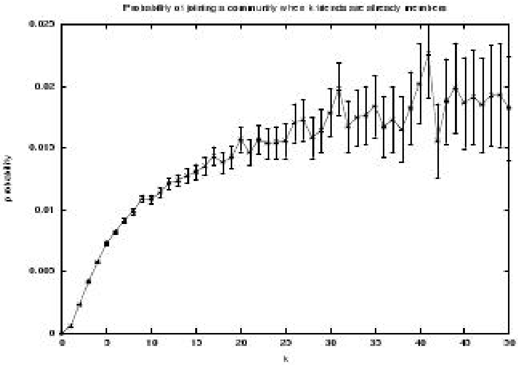


Figure 1: The probability p of joining a LiveJournal community as a function of the number of friends k already in the community. Error bars represent two standard errors.

Community membership...

Features Used	ROCA	APR	CXE
Number of Friends	0.69244	0.00301	0.00934
Post Activity	0.73421	0.00316	0.00934
All	0.75642	0.00380	0.00923

Community membership...

Features Used	ROCA	APR	CXE
Number of Friends	0.69244	0.00301	0.00934
Post Activity	0.73421	0.00316	0.00934
All	0.75642	0.00380	0.00923

Observations

- Internal connectedness of friends inside the community: Probability of joining community proportional to this feature
- Relation to Models of diffusion: Asynchronous Behaviour

Observations

- Internal connectedness of friends inside the community: Probability of joining community proportional to this feature
- Relation to Models of diffusion: Asynchronous Behaviour

Observations

- Internal connectedness of friends inside the community: Probability of joining community proportional to this feature
- Relation to Models of diffusion: Asynchronous Behaviour

Future Directions

- Asynchronous models of diffusion
- Incorporate other features like internal connectedness of the neighbours along with number of neighbours in models of diffusion

Future Directions

- Asynchronous models of diffusion
- Incorporate other features like internal connectedness of the neighbours along with number of neighbours in models of diffusion

Future Directions

- Asynchronous models of diffusion
- Incorporate other features like internal connectedness of the neighbours along with number of neighbours in models of diffusion

Maximizing the Spread of influence through Social Networks
By Kempe et al appeared in KDD03

Models of Diffusion

Linear Threshold Model

A node v is influenced by each neighbor w according to weights $b_{v,w}$ and initially each node chooses a random threshold(θ_v) and some initial set of nodes are activated/influenced initially. Now in each discrete time step activate the nodes whose total weight of its active neighbours crosses threshold i.e.

$$\sum_{w \text{ active neighbor of } v} b_{v,w} \geq \theta_v$$

Models of Diffusion

Linear Threshold Model

A node v is influenced by each neighbor w according to weights $b_{v,w}$ and initially each node chooses a random threshold(θ_v) and some initial set of nodes are activated/influenced initially. Now in each discrete time step activate the nodes whose total weight of its active neighbours crosses threshold i.e.

$$\sum_{w \text{ active neighbor of } v} b_{v,w} \geq \theta_v$$

Models of Diffusion

Independent Cascade Model

Initially some set of nodes are activated, then at the time step when a node v becomes active it is given *single* chance to activate each currently inactive neighbor w, it succeeds with probability say $p_{v,w}$, the process stops when no more activations are possible.

Models of Diffusion

Independent Cascade Model

Initially some set of nodes are activated, then at the time step when a node v becomes active it is given *single* chance to activate each currently inactive neighbor w, it succeeds with probability say $p_{v,w}$, the process stops when no more activations are possible.

Influence Maximization Problem

Theorem

1. Influence maximization problem is $NP-hard$ for both Linear Threshold and Independent cascade model.
2. Influence function is monotone and submodular for both the models and $1-1/e$ approximation algorithm exists.

Influence Maximization Problem

Theorem

1. Influence maximization problem is $NP-hard$ for both Linear Threshold and Independent cascade model.
2. Influence function is monotone and submodular for both the models and $1-1/e$ approximation algorithm exists.

Optimal Marketing Strategies Over Social Networks
By Mukund Sundararajan et al. appeared in *WWW 08*

The Model

- A seller and set V of potential buyers, seller approaches each buyer in sequence and offers a price
- Valuation of the buyer is $v_i : 2^V \rightarrow \mathbb{R}^+$capturing the influence that other buyers that already own the item have on each other
- Strategy for the seller consist of two components: Sequence of buyers and the price to offer

The Model

- A seller and set V of potential buyers, seller approaches each buyer in sequence and offers a price
- Valuation of the buyer is $v_i : 2^V \rightarrow \mathbb{R}^+$capturing the influence that other buyers that already own the item have on each other
- Strategy for the seller consist of two components: Sequence of buyers and the price to offer

The Model

- A seller and set V of potential buyers, seller approaches each buyer in sequence and offers a price
- Valuation of the buyer is $v_i : 2^V \rightarrow \mathbb{R}^+$capturing the influence that other buyers that already own the item have on each other
- Strategy for the seller consist of two components: Sequence of buyers and the price to offer

The Model

- A seller and set V of potential buyers, seller approaches each buyer in sequence and offers a price
- Valuation of the buyer is $v_i : 2^V \rightarrow \mathbb{R}^+$capturing the influence that other buyers that already own the item have on each other
- Strategy for the seller consist of two components: Sequence of buyers and the price to offer

Optimal Marketing Strategies

- For symmetric buyers i.e when their valuations are identically distributed then optimal marketing strategy can be computed in polynomial time.
- For assymmetric setting finding optimal marketing strategy is $NP-hard$
- **I**nfluence: and **E**xploit strategies for constant factor approximation algorithms

Optimal Marketing Strategies

- For symmetric buyers i.e when their valuations are identically distributed then optimal marketing strategy can be computed in polynomial time.
- For assymmetric setting finding optimal marketing strategy is $NP - hard$
- Influence: and Exploit strategies for constant factor approximation algorithms

Optimal Marketing Strategies

- For symmetric buyers i.e when their valuations are identically distributed then optimal marketing strategy can be computed in polynomial time.
- For assymmetric setting finding optimal marketing strategy is $NP - hard$
- Influence: and Exploit strategies for constant factor approximation algorithms

Optimal Marketing Strategies

- For symmetric buyers i.e when their valuations are identically distributed then optimal marketing strategy can be computed in polynomial time.
- For assymmetric setting finding optimal marketing strategy is $NP - hard$
- **I**nfluence: and **E**xploit strategies for constant factor approximation algorithms

Open Problems

- Finding another family of strategies which intelligently constructs the sequence of buyers to offer product to, to improve approximation ratio
- Pricing strategies for viral marketing[3]
- Strategic buyers with costs involved in maintaining the links

[3] Pricing strategies for viral marketing on social networks Rajeev Motwani et al. Under Submission 2009

Open Problems

- Finding another family of strategies which intelligently constructs the sequence of buyers to offer product to, to improve approximation ratio
- Pricing strategies for viral marketing[3]
- Strategic buyers with costs involved in maintaining the links

[3] Pricing strategies for viral marketing on social networks Rajeev Motwani et al. Under Submission 2009

Open Problems

- Finding another family of strategies which intelligently constructs the sequence of buyers to offer product to, to improve approximation ratio
- Pricing strategies for viral marketing[3]
- Strategic buyers with costs involved in maintaining the links

[3] Pricing strategies for viral marketing on social networks Rajeev Motwani et al. Under Submission 2009

Open Problems

- Finding another family of strategies which intelligently constructs the sequence of buyers to offer product to, to improve approximation ratio
- Pricing strategies for viral marketing[3]
- Strategic buyers with costs involved in maintaining the links

[3] Pricing strategies for viral marketing on social networks Rajeev Motwani et al. Under Submission 2009

Inoculation Strategies for the victims of viruses and sum of squares partition problem, James Aspnes et al. Appeared in SODA05

The Game

- A sets of n players each choose a strategy a_i which is 0 or 1, this gives strategy profile $\vec{a} \in [0,1]^n$
- Individual costs/utilities of nodes given a strategy profile $\vec{a} \in [0,1]^n$

$$cost_i(\overrightarrow{a}) = a_i C + (1 - a_i) L \cdot p_i(\overrightarrow{a})$$

- Social cost-

$$cost(\overrightarrow{a}) = \sum_{j=0}^{n-1} cost_j(\overrightarrow{a})$$

The Game

- A sets of n players each choose a strategy a_i which is 0 or 1, this gives strategy profile $\vec{a} \in [0,1]^n$
- Individual costs/utilities of nodes given a strategy profile $\vec{a} \in [0,1]^n$

$$cost_i(\overrightarrow{a}) = a_i C + (1 - a_i) L \cdot p_i(\overrightarrow{a})$$

- Social cost-

$$cost(\overrightarrow{a}) = \sum_{j=0}^{n-1} cost_j(\overrightarrow{a})$$

The Game

- A sets of n players each choose a strategy a_i which is 0 or 1, this gives strategy profile $\vec{a} \in [0,1]^n$
- Individual costs/utilities of nodes given a strategy profile $\vec{a} \in [0,1]^n$

$$cost_i(\overrightarrow{a}) = a_i C + (1 - a_i) L \cdot p_i(\overrightarrow{a})$$

- Social cost-

$$cost(\overrightarrow{a}) = \sum_{j=0}^{n-1} cost_j(\overrightarrow{a})$$

The Game

- A sets of n players each choose a strategy a_i which is 0 or 1, this gives strategy profile $\vec{a} \in [0,1]^n$
- Individual costs/utilities of nodes given a strategy profile $\vec{a} \in [0,1]^n$

$$cost_i(\overrightarrow{a}) = a_i C + (1 - a_i) L \cdot p_i(\overrightarrow{a})$$

- Social cost-

$$cost(\overrightarrow{a}) = \sum_{j=0}^{n-1} cost_j(\overrightarrow{a})$$

Attack Graph

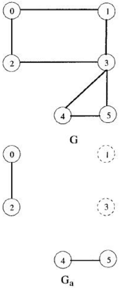


Figure 1: Sample graph G and its attack graph $G_{\vec{a}}$ for $\vec{a} = 010100$.

$$p_i(\vec{a}) = k_i/n$$

Attack Graph

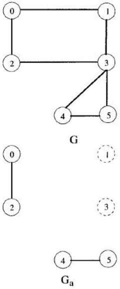


Figure 1: Sample graph G and its attack graph $G_{\vec{a}}$ for $\vec{a} = 010100$.

$$p_i(\vec{a}) = k_i / n$$

Results

- Nash Euilibrium of this game always exists and *some* Nash Equilibrium can be computed in $O(n^3)$ time starting from $\vec{a} = 1^n$
- Computing worst (w.r.t Social cost) and Best Nash Equilibrium is $NP-hard$ and Price of Anarchy is $\Theta(n)$
- Computing Social optimum is $NP-hard$

Results

- Nash Euilibrium of this game always exists and *some* Nash Equilibrium can be computed in $O(n^3)$ time starting from $\vec{a} = 1^n$
- Computing worst (w.r.t Social cost) and Best Nash Equilibrium is $NP-hard$ and Price of Anarchy is $\Theta(n)$
- Computing Social optimum is $NP-hard$

Results

- Nash Euilibrium of this game always exists and *some* Nash Equilibrium can be computed in $O(n^3)$ time starting from $\vec{a} = 1^n$
- Computing worst (w.r.t Social cost) and Best Nash Equilibrium is $NP-hard$ and Price of Anarchy is $\Theta(n)$
- Computing Social optimum is $NP-hard$

Results

- Nash Euilibrium of this game always exists and *some* Nash Equilibrium can be computed in $O(n^3)$ time starting from $\vec{a} = 1^n$
- Computing worst (w.r.t Social cost) and Best Nash Equilibrium is $NP - hard$ and Price of Anarchy is $\Theta(n)$
- Computing Social optimum is $NP - hard$

Open Problems

- Better Approximation algorithm for social optimum and/or inapproximability result
- Model complexity:The Virus need not infect all the neighbors of the infected node, the loss due to virus need not be known deterministically, the incomplete information about strategies

Open Problems

- Better Approximation algorithm for social optimum and/or inapproximability result
- Model complexity:The Virus need not infect all the neighbors of the infected node, the loss due to virus need not be known deterministically, the incomplete information about strategies

Open Problems

- Better Approximation algorithm for social optimum and/or inapproximability result
- Model complexity:The Virus need not infect all the neighbors of the infected node, the loss due to virus need not be known deterministically, the incomplete information about strategies

On the Windfall of Friendship: Virus Inoculation Strategies on Social Networks
by Dominic Meier et al. appeared in EC08

An Excerpt from a watercooler conversation

- Prof. Jayant: How do you do Prof. Narahari?
- Prof. Narahari: Things are great on my end, I think game theory is the most interesting research area modeling real world phenomenon perfectly
- Prof. Jayant: Now how is that? People in real world are not purely selfish and intelligent for example, my lab is filled with students who are light years away from being termed as intelligent, *I bet your's is too*.
- Prof. Narahari: Regarding lab situation, I totally agree, but I beg to differ with you on the former argument prof. Jayant, people in real world only have *strange* utility functions.

And prof. Narahari wins another intellectual argument against prof. Jayant

An Excerpt from a watercooler conversation

- Prof. Jayant: How do you do Prof. Narahari?
- Prof. Narahari: Things are great on my end, I think game theory is the most interesting research area modeling real world phenomenon perfectly
- Prof. Jayant: Now how is that? People in real world are not purely selfish and intelligent for example, my lab is filled with students who are light years away from being termed as intelligent, *I bet your's is too.*
- Prof. Narahari: Regarding lab situation, I totally agree, but I beg to differ with you on the former argument prof. Jayant, people in real world only have *strange* utility functions.

And prof. Narahari wins another intellectual argument against prof. Jayant

An Excerpt from a watercooler conversation

- Prof. Jayant: How do you do Prof. Narahari?
- Prof. Narahari: Things are great on my end, I think game theory is the most interesting research area modeling real world phenomenon perfectly
- Prof. Jayant: Now how is that? People in real world are not purely selfish and intelligent for example, my lab is filled with students who are light years away from being termed as intelligent, *I bet your's is too.*
- Prof. Narahari: Regarding lab situation, I totally agree, but I beg to differ with you on the former argument prof. Jayant, people in real world only have *strange* utility functions.

And prof. Narahari wins another intellectual argument against prof. Jayant

An Excerpt from a watercooler conversation

- Prof. Jayant: How do you do Prof. Narahari?
- Prof. Narahari: Things are great on my end, I think game theory is the most interesting research area modeling real world phenomenon perfectly
- Prof. Jayant: Now how is that? People in real world are not purely selfish and intelligent for example, my lab is filled with students who are light years away from being termed as intelligent, *I bet your's is too*.
- Prof. Narahari: Regarding lab situation, I totally agree, but I beg to differ with you on the former argument prof. Jayant, people in real world only have *strange* utility functions.

And prof. Narahari wins another intellectual argument against prof. Jayant

An Excerpt from a watercooler conversation

- Prof. Jayant: How do you do Prof. Narahari?
- Prof. Narahari: Things are great on my end, I think game theory is the most interesting research area modeling real world phenomenon perfectly
- Prof. Jayant: Now how is that? People in real world are not purely selfish and intelligent for example, my lab is filled with students who are light years away from being termed as intelligent, *I bet your's is too*.
- Prof. Narahari: Regarding lab situation, I totally agree, but I beg to differ with you on the former argument prof. Jayant, people in real world only have *strange* utility functions.

And prof. Narahari wins another intellectual argument against prof. Jayant

An Excerpt from a watercooler conversation

- Prof. Jayant: How do you do Prof. Narahari?
- Prof. Narahari: Things are great on my end, I think game theory is the most interesting research area modeling real world phenomenon perfectly
- Prof. Jayant: Now how is that? People in real world are not purely selfish and intelligent for example, my lab is filled with students who are light years away from being termed as intelligent, *I bet your's is too*.
- Prof. Narahari: Regarding lab situation, I totally agree, but I beg to differ with you on the former argument prof. Jayant, people in real world only have *strange* utility functions.

And prof. Narahari wins another intellectual argument against prof. Jayant

The Game and Windfall of Friendship

- The *Perceived* cost/utility of an Individual-

$$c_p(i,\vec{a}) = c_a(i,\vec{a}) + F \cdot \sum_{p_j \in \Gamma(p_i)} c_a(j,\vec{a})$$

where, $F \in [0,1]$ is the Friendship Factor and $\Gamma(p_i)$ are the neighbours of node p_i

Definition

The Windfall of Friendship (WoF) $\gamma(F,I)$ is defined as the ratio of worst Nash equilibrium to the worst Friendship Nash Equilibrium(FNE)

The Game and Windfall of Friendship

- The *Perceived* cost/utility of an Individual-

$$c_p(i,\vec{a}) = c_a(i,\vec{a}) + F \cdot \sum_{p_j \in \Gamma(p_i)} c_a(j,\vec{a})$$

where, $F \in [0,1]$ is the Friendship Factor and $\Gamma(p_i)$ are the neighbours of node p_i

Definition

The Windfall of Friendship (WoF) $\gamma(F,I)$ is defined as the ratio of worst Nash equilibrium to the worst Friendship Nash Equilibrium(FNE)

The Game and Windfall of Friendship

- The *Perceived* cost/utility of an Individual-

$$c_p(i,\vec{a}) = c_a(i,\vec{a}) + F \cdot \sum_{p_j \in \Gamma(p_i)} c_a(j,\vec{a})$$

where, $F \in [0,1]$ is the Friendship Factor and $\Gamma(p_i)$ are the neighbours of node p_i

Definition

The Windfall of Friendship (WoF) $\gamma(F, I)$ is defined as the ratio of worst Nash equilibrium to the worst Friendship Nash Equilibrium(FNE)

Results

- For any instance of the game $1 \leq \gamma(F, I) \leq PoA(I)$
- It is $NP - complete$ problem to compute the best and the worst Friendship Nash Equilibrium
- For Complete Graph and Star Graph there always exists a FNE, for K_n $\gamma(F, I) \leq 4/3$, for star graph under certain conditions it can be as high as $O(n)$.

Results

- For any instance of the game $1 \leq \gamma(F, I) \leq PoA(I)$
- It is $NP - complete$ problem to compute the best and the worst Friendship Nash Equilibrium
- For Complete Graph and Star Graph there always exists a FNE, for K_n $\gamma(F, I) \leq 4/3$, for star graph under certain conditions it can be as high as $O(n)$.

Results

- For any instance of the game $1 \leq \gamma(F, I) \leq PoA(I)$
- It is $NP - complete$ problem to compute the best and the worst Friendship Nash Equilibrium
- For Complete Graph and Star Graph there always exists a FNE, for K_n $\gamma(F, I) \leq 4/3$, for star graph under certain conditions it can be as high as $O(n)$.

Results

- For any instance of the game $1 \leq \gamma(F, I) \leq PoA(I)$
- It is $NP-complete$ problem to compute the best and the worst Friendship Nash Equilibrium
- For Complete Graph and Star Graph there always exists a FNE, for K_n $\gamma(F, I) \leq 4/3$, for star graph under certain conditions it can be as high as $O(n)$.

Open Problems

- Existence of Nash Equilibrium for any graph
- Introduction of Malicious Players[4] in the setting
- There could be number of viruses on each node and each node could be resistant to subset of those viruses

[4] *When Selfish meets Evil: Byzantine Players in a Virus Inoculation Game* by Thomas Moscibroda et al. Appeared in PODC06

Open Problems

- Existence of Nash Equilibrium for any graph
- Introduction of Malicious Players[4] in the setting
- There could be number of viruses on each node and each node could be resistant to subset of those viruses

[4] *When Selfish meets Evil: Byzantine Players in a Virus Inoculation Game* by Thomas Moscibroda et al. Appeared in PODC06

Open Problems

- Existence of Nash Equilibrium for any graph
- Introduction of Malicious Players[4] in the setting
- There could be number of viruses on each node and each node could be resistant to subset of those viruses

[4] *When Selfish meets Evil: Byzantine Players in a Virus Inoculation Game* by Thomas Moscibroda et al. Appeared in PODC06

Open Problems

- Existence of Nash Equilibrium for any graph
- Introduction of Malicious Players[4] in the setting
- There could be number of viruses on each node and each node could be resistant to subset of those viruses

[4] *When Selfish meets Evil: Byzantine Players in a Virus Inoculation Game* by Thomas Moscibroda et al. Appeared in PODC06

A Framework For Analysis of Dynamic Social Networks
by Tanya Berger-Wolf et al. appeared in KDD06

The Model

- Consider a population, $X = \{x_1, ..., x_n\}$, and $g \subseteq X$ and input temporal sequence of partitions of X, $P_1, P_2, ..., P_T$ where each partition is a *disjoint* set of groups and let $P(g)$ denote the index of the partition to which g belongs.

Definitions

Given Temporal sequence of partitions, and a similarity measure $sim(,)$, a turnover threshold β and a function $\alpha(T)$, then a MetaGroup (MG) is a sequence of groups $MG = \{g_1, ..., g_l\}$, $\alpha(T) \leq l \leq T$ such that, no two groups in MG are in the same partition and the groups are ordered by the partition time steps i.e. $\forall i, j \; 1 \leq i < j \leq l, \; P(g_i) < P(g_j)$ and the consecutive groups in MG are similar i.e $\forall 1 \leq i < l \; sim(g_i, g_{i+1}) \geq \beta$

The Model

- Consider a population, $X = \{x_1, ..., x_n\}$, and $g \subseteq X$ and input temporal sequence of partitions of X, $P_1, P_2, ..., P_T$ where each partition is a *disjoint* set of groups and let $P(g)$ denote the index of the partition to which g belongs.

Definitions

Given Temporal sequence of partitions, and a similarity measure $sim(,)$, a turnover threshold β and a function $\alpha(T)$, then a MetaGroup (MG) is a sequence of groups $MG = \{g_1, ..., g_l\}$, $\alpha(T) \leq l \leq T$ such that, no two groups in MG are in the same partition and the groups are ordered by the partition time steps i.e. $\forall i, j \; 1 \leq i < j \leq l, \; P(g_i) < P(g_j)$ and the consecutive groups in MG are similar i.e $\forall 1 \leq i < l \; sim(g_i, g_{i+1}) \geq \beta$

The Model

- Consider a population, $X = \{x_1, ..., x_n\}$, and $g \subseteq X$ and input temporal sequence of partitions of X, $P_1, P_2, ..., P_T$ where each partition is a *disjoint* set of groups and let $P(g)$ denote the index of the partition to which g belongs.

Definitions

Given Temporal sequence of partitions, and a similarity measure $sim(,)$, a turnover threshold β and a function $\alpha(T)$, then a MetaGroup (MG) is a sequence of groups $MG = \{g_1, ..., g_l\}$, $\alpha(T) \leq l \leq T$ such that, no two groups in MG are in the same partition and the groups are ordered by the partition time steps i.e. $\forall i, j\, 1 \leq i < j \leq l,\ P(g_i) < P(g_j)$ and the consecutive groups in MG are similar i.e $\forall 1 \leq i < l\ sim(g_i, g_{i+1}) \geq \beta$

Graph Theoretic Formulation

Graph Model

Consider a multipartite weighted Graph $G = (V_1, ..., V_T, E)$ where V_i is the set of groups in partition P_i and $(g_i, g_j) \in E$ if $P(g_i) < P(g_j)$ and $sim(g_i, g_j) \geq \beta$ and $w(g_i, g_j) = sim(g_i, g_j)$, so metagroup in this graph will be path of length atleast $\alpha(T)$.

Graph Theoretic Formulation

Graph Model

Consider a multipartite weighted Graph $G = (V_1, ..., V_T, E)$ where V_i is the set of groups in partition P_i and $(g_{i,}g_j) \in E$ if $P(g_i) < P(g_j)$ and $sim(g_i, g_j) \geq \beta$ and $w(g_i, g_j) = sim(g_i, g_j)$, so metagroup in this graph will be path of length atleast $\alpha(T)$.

Algorithms for MG Statistics

- Number of Metagroups, average metagroup length
- Extremal problems: Most persistent MG, Most stable MG
- Group Connectivity : Given a set of groups $g_1, ..., g_l$ in separate partitions ordered by their partition indices then is there a MG containing all these groups, among them find the most persistent,stable MG etc.

Algorithms for MG Statistics

- Number of Metagroups, average metagroup length
- Extremal problems: Most persistent MG, Most stable MG
- Group Connectivity : Given a set of groups $g_1, \ldots, g_l$ in separate partitions ordered by their partition indices then is there a MG containing all these groups, among them find the most persistent,stable MG etc.

Algorithms for MG Statistics

- Number of Metagroups, average metagroup length
- Extremal problems: Most persistent MG, Most stable MG
- Group Connectivity : Given a set of groups $g_1, \ldots, g_l$ in separate partitions ordered by their partition indices then is there a MG containing all these groups, among them find the most persistent,stable MG etc.

Algorithms for MG Statistics

- Number of Metagroups, average metagroup length
- Extremal problems: Most persistent MG, Most stable MG
- Group Connectivity : Given a set of groups $g_1, \ldots, g_l$ in separate partitions ordered by their partition indices then is there a MG containing all these groups, among them find the most persistent,stable MG etc.

Strategic Network Formation with Structural Holes
by Jon Kleinberg et al. appeared in EC08

The Model

Network Formation Game

Each node u links to a set of nodes $L(u)$

The constructed links by all the nodes form a undirected graph in which let $N(u)$ be the set of neighbors of u, let r_{vw} be the number of length two paths between v, w then payoff to a node u is given by-

$$\alpha_0 |N(u)| + \sum_{v,w \in N(u)} \beta(r_{vw}) - \sum_{v \in L(u)} c_{uv}$$

Where c_{uv} is the edge maintainance cost and β is some decreasing function

The Model

Network Formation Game

Each node u links to a set of nodes $L(u)$

The constructed links by all the nodes form a undirected graph in which let $N(u)$ be the set of neighbors of u, let r_{vw} be the number of length two paths between v, w then payoff to a node u is given by-

$$\alpha_0 \left|N(u)\right| + \sum_{v,w \in N(u)} \beta(r_{vw}) - \sum_{v \in L(u)} c_{uv}$$

Where c_{uv} is the edge maintainance cost and β is some decreasing function

Results and Analysis

- The best response dynamics of a node can be computed in polynomial time and for arbitrary cost function the best response dynamics can cycle
- In uniform metric i.e when all edge maintenance costs $c_{uv} = 1$ then a multipartite graph $G_{n,k}$ whose vertex set $V = V_1 \cup .... \cup V_q \cup V_{q+1}$ where $V_i \cap V_j = \emptyset$ for $i \neq j$ and $|V_1| = ... = |V_q| = k$ and $|V_{q+1}| = n \, mod \, k$, for each $u \in V_i$ $L(u) = \cup_{j=1}^{i-1} V_j$ is a Nash equilibrium of the game for some choice of k.
- They characterize the structure of equilibrium graph by showing that every equilibrium graph will have $\Omega(n^2)$ edges
- The question of existence of nash equilibrium for general cost matrix remains open.

Results and Analysis

- The best response dynamics of a node can be computed in polynomial time and for arbitrary cost function the best response dynamics can cycle
- In uniform metric i.e when all edge maintenance costs $c_{uv} = 1$ then a multipartite graph $G_{n,k}$ whose vertex set $V = V_1 \cup .... \cup V_q \cup V_{q+1}$ where $V_i \cap V_j = \emptyset$ for $i \neq j$ and $|V_1| = ... = |V_q| = k$ and $|V_{q+1}| = n \, mod \, k$, for each $u \in V_i$ $L(u) = \cup_{j=1}^{i-1} V_j$ is a Nash equilibrium of the game for some choice of k.
- They characterize the structure of equilibrium graph by showing that every equilibrium graph will have $\Omega(n^2)$ edges
- The question of existence of nash equilibrium for general cost matrix remains open.

Results and Analysis

- The best response dynamics of a node can be computed in polynomial time and for arbitrary cost function the best response dynamics can cycle
- In uniform metric i.e when all edge maintenance costs $c_{uv} = 1$ then a multipartite graph $G_{n,k}$ whose vertex set $V = V_1 \cup \ldots. \cup V_q \cup V_{q+1}$ where $V_i \cap V_j = \emptyset$ for $i \neq j$ and $|V_1| = \ldots = |V_q| = k$ and $|V_{q+1}| = n \, mod \, k$, for each $u \in V_i$ $L(u) = \cup_{j=1}^{i-1} V_j$ is a Nash equilibrium of the game for some choice of k.
- They characterize the structure of equilibrium graph by showing that every equilibrium graph will have $\Omega(n^2)$ edges
- The question of existence of nash equilibrium for general cost matrix remains open.

Results and Analysis

- The best response dynamics of a node can be computed in polynomial time and for arbitrary cost function the best response dynamics can cycle
- In uniform metric i.e when all edge maintenance costs $c_{uv} = 1$ then a multipartite graph $G_{n,k}$ whose vertex set $V = V_1 \cup .... \cup V_q \cup V_{q+1}$ where $V_i \cap V_j = \emptyset$ for $i \neq j$ and $|V_1| = ... = |V_q| = k$ and $|V_{q+1}| = n \, mod \, k$, for each $u \in V_i$ $L(u) = \cup_{j=1}^{i-1} V_j$ is a Nash equilibrium of the game for some choice of k.
- They characterize the structure of equilibrium graph by showing that every equilibrium graph will have $\Omega(n^2)$ edges
- The question of existence of nash equilibrium for general cost matrix remains open.

Results and Analysis

- The best response dynamics of a node can be computed in polynomial time and for arbitrary cost function the best response dynamics can cycle
- In uniform metric i.e when all edge maintenance costs $c_{uv} = 1$ then a multipartite graph $G_{n,k}$ whose vertex set $V = V_1 \cup .... \cup V_q \cup V_{q+1}$ where $V_i \cap V_j = \emptyset$ for $i \neq j$ and $|V_1| = ... = |V_q| = k$ and $|V_{q+1}| = n \, mod \, k$, for each $u \in V_i$ $L(u) = \cup_{j=1}^{i-1} V_j$ is a Nash equilibrium of the game for some choice of k.
- They characterize the structure of equilibrium graph by showing that every equilibrium graph will have $\Omega(n^2)$ edges
- The question of existence of nash equilibrium for general cost matrix remains open.

Thank You!

Have
a
Rocking
Weekend!!!

References

- M.E.J Newman et al. *Community Structure in Social and Biological Network. PNAS*, 2002.
- M.E.J Newman *Modularity and Community Structure in Networks.* PNAS. 2006.
- Tanya Berger-Wolf et al. *Constant Factor Approximation Algorithm for Identifying Dynamic Communities.* KDD 2009.
- Jon Kleinberg et al. *Structure of Information Pathways in Social Communication Network.* KDD. 2008.
- Moshe Babaioff et al *Congestion Games with Malicious Players.* EC. 2007.
- Jon Kleinberg et al. *Feedback Effects Between Similarity and Social Influence.* KDD. 2008.