Nash Stable Partitioning of Graphs and Community Detection in Social Networks

December 15, 2010

Y. NARAHARI

http://lcm.csa.iisc.ernet.in/hari

Computer Science and Automation
Indian Institute of Science, Bangalore
 PART 1: Social Network Analysis

 PART 2: Game Theoretic Models for Social Network Analysis

 PART 3: Community Detection and Nash Stable Partitions

 PART 4: SCoDA: Stable Community Detection Algorithm
Today’s Talk is a Tribute to

John von Neumann

The Genius who created two intellectual currents in the 1930s, 1940s

 Founded Game Theory with Oskar Morgenstern (1928-44)

Pioneered the Concept of a Digital Computer and Algorithms (1930s and 40s)
CENTRAL IDEA

Game Theoretic Models are very natural for modeling social networks
--------------------------------------
Social network nodes are rational, intelligent
--------------------------------------
Social networks form in a decentralized way
--------------------------------------
Strategic interactions among social network nodes

It would be interesting to explore Game Theoretic Models for analyzing social networks
--------------------------------------
Example 1: Discovering Communities
--------------------------------------
Example 2: Network Formation
--------------------------------------
Example 3: Finding Influential Nodes

Ramasuri Narayananam. Game Theoretic Models for Social Network Analysis, Ph.D. Dissertation, CSA, IISc, November 2010
Why Social Network Analysis?

- Diffusion of Information and Innovations
- To understand spread of diseases (Epidemiology)
- E-Commerce and E-Business (selling patterns, marketing)
- Job Finding (through referrals)
- Determine Influential Players (scientists, innovators, employees, customers, companies, genes, etc.)
- Build effective social and political campaigns
- Predict future events
- Crack terrorist/criminal networks
- Track alumni, etc…
Tools and Techniques for SNA

- Random Graphs
- Simulation
- Probabilistic Models
- Multi-Level Models
- Optimization Models
- Game Theoretic Models
Game Theory

Mathematical framework for rigorous study of conflict and cooperation among rational, intelligent agents

In the Internet era, Game Theory has become a valuable tool for analysis and design
Strategic Form Games (Normal Form Games)

\[ N = \{1, \ldots, n\} \]

\[ S_1, \ldots, S_n \]

Strategy Sets

\[ S = S_1 \times \ldots \times S_n \]

Players

\[ U_1 : S \rightarrow \mathbb{R} \]

\[ U_n : S \rightarrow \mathbb{R} \]

Payoff functions (Utility functions)
Example 1: Coordination Game

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>IIITB</th>
<th>MG Road</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>IIITB</td>
<td>100,100</td>
<td>0,0</td>
</tr>
<tr>
<td></td>
<td>MG Road</td>
<td>0,0</td>
<td>10,10</td>
</tr>
</tbody>
</table>

Models the strategic conflict when two players have to choose their priorities.
Example 2: Prisoner’s Dilemma

<table>
<thead>
<tr>
<th></th>
<th>No Confess NC</th>
<th>Confess C</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Confess NC</td>
<td>-2, -2</td>
<td>-10, -1</td>
</tr>
<tr>
<td>Confess C</td>
<td>-1, -10</td>
<td>-5, -5</td>
</tr>
</tbody>
</table>
Nash Equilibrium

A profile of strategies \( (s_1^*, s_2^*, \ldots, s_n^*) \) is said to be a pure strategy Nash Equilibrium if \( s_i^* \) is a best response strategy against \( s_{-i}^* \) for all \( i = 1, 2, \ldots, n \).

A Nash equilibrium profile is robust to unilateral deviations and captures a stable, self-enforcing agreement among the players.
Nash Equilibria in Coordination Game

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>IIITB</th>
<th>MG Road</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>IIITB</td>
<td>100,100</td>
<td>0,0</td>
</tr>
<tr>
<td></td>
<td>MG Road</td>
<td>0,0</td>
<td>10,10</td>
</tr>
</tbody>
</table>

Two pure strategy Nash equilibria: (IIITB, IIITB) and (MG Road, MG Road); one mixed strategy Nash equilibrium.
Nash Equilibrium in Prisoner’s Dilemma

<table>
<thead>
<tr>
<th></th>
<th>No Confess NC</th>
<th>Confess C</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Confess NC</td>
<td>-2, -2</td>
<td>-10, -1</td>
</tr>
<tr>
<td>Confess C</td>
<td>-1, -10</td>
<td>-5, -5</td>
</tr>
</tbody>
</table>

(C,C) is a Nash equilibrium
Nash’s Theorem

Every finite strategic form game has at least one mixed strategy Nash equilibrium

Mixed strategy of a player ‘i’ is a probability distribution on $S_i$.

$(\sigma_1^*, \sigma_2^*, \ldots, \sigma_n^*)$ is a mixed strategy Nash equilibrium if

$\sigma_i^*$ is a best response against $\sigma_{-i}^*$, $\forall i = 1, 2, \ldots, n$
Relevance of Nash Equilibrium

Nash equilibrium has several possible implications:

(a) Players are happy the way they are; do not want to deviate unilaterally

(b) Stable, self-enforcing, self-sustaining agreement

(c) Provides a principled way of predicting a steady-state outcome of a dynamic adjustment process
Example: Nash Equilibrium in Social Network Formation

\[ N = \{ 1, 2, \ldots, n \} \]  \text{Nodes}

\[ S_i = \text{Subsets of } N - \{ i \} \]

\[ u_i : S_1 \times S_2 \times \ldots \times S_n \rightarrow R \]

Each strategy profile results in a particular network

\textbf{Standard Assumptions:}

- A link is formed under mutual consent
- A link may be deleted unilaterally
\( N = \{1,2,3,4\} \)
\( s_1 = \{2,3,4\} \)
\( s_2 = \{1,3,4\} \)
\( s_3 = \{1,2,4\} \)
\( s_4 = \{1,2,3\} \)
Pairwise Nash Stability

1) No agent can increase its payoff by deleting a link. That is, \( \forall i, j \in N, \)
\[
  u_i(g) \geq u_i(g - ij) \\
  u_j(g) \geq u_j(g - ij)
\]

2) No two agents can both benefit, at least one of them strictly, by adding a link between themselves. That is, \( \forall i, j \in N, \) the following is not possible.
\[
  u_i(g) \leq u_i(g + ij) \\
  u_j(g) \leq u_j(g + ij)
\]

with at least one inequality strict.
Community Detection Problem

• Discover natural components such that connections within a component are dense and across components are sparse

• Important for social campaigns, viral marketing, search, and a variety of applications

• Extensively investigated problem

• Communities could be overlapping or non-overlapping. We are interested in non-overlapping communities.
## Community Detection: Relevant Work

### Optimization based approaches using global objective based on centrality based measures

### Spectral methods, Eigen vector based methods

### Multi-level Approaches

### State-of-the-Art Review
J. Leskovec et al. Empirical comparison of algorithms for community detection. WWW 2010
Existing Algorithms for Community Detection: A Few Issues

Most of these algorithms work with a global objective such as modularity, conductance, etc.

Do not take into account the strategic nature of the players and their associations

Most algorithms require the number of communities to be provided as an input to the algorithm
Our Approach

We use a strategic form game to model the formation of communities.

We view detection of non-overlapping communities as a graph partitioning problem and set up a graph partitioning game.

Only relevant existing work
Community Detection and Graph Partitioning

Non-overlapping community detection can be viewed as a graph partitioning problem
Graph Partitioning: Applications

1. VLSI circuit design
2. Resource allocation in parallel computing
3. Graph visualization and summarization
4. Epidemiology
5. Social Network Analysis
Email Network – Visualization and Summarization
Graph Partitioning Game

- Nodes in the network are the players
- Strategy of a node is to choose its community
- Utilities to be defined to reflect the network structure and the problem setting; preferably should use only local information
Proposed Utility Function

$U_i(S)$ is the sum of number of neighbours of node $i$ in the community plus a normalized value of the neighbours who are themselves connected.

The proposed utility function captures the Degree of connectivity of the node and also the density of its neighbourhood.

A Nash Stable Partition is one in which no node has incentive to defect to any other community.
Nash Stable Partition: An Example

\[ u_1(S_1) = 6 \quad u_1(S_2) = 0 \quad u_7(S_1) = 6 \quad u_7(S_2) = 1 \]
\[ u_2(S_1) = 9.33 \quad u_2(S_2) = 0 \quad u_8(S_1) = 9 \quad u_8(S_2) = 0 \]
\[ u_3(S_1) = 9.33 \quad u_3(S_2) = 0 \quad u_9(S_1) = 9.33 \quad u_9(S_2) = 0 \]
\[ u_4(S_1) = 9 \quad u_4(S_2) = 0 \quad u_{10}(S_1) = 9.33 \quad u_{10}(S_2) = 0 \]
\[ u_5(S_1) = 6 \quad u_5(S_2) = 1 \quad u_{11}(S_1) = 6 \quad u_{11}(S_2) = 0 \]
\[ u_6(S_1) = 1 \quad u_6(S_2) = 1 \]
SCoDA: Stable Community Detection Algorithm

Start with an initial partition where each community has a small number of nodes

Choose nodes in a non-decreasing order of degrees and investigate if it is better to defect to a neighbouring community

The algorithm terminates in a Nash stable partition
## Comparison of SCoDA with other Algorithms

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Girvan and Newman</td>
<td>M. Girvan and MEJ Newman. PNAS 2002</td>
</tr>
<tr>
<td>Spectral Algorithm</td>
<td>MEJ Newman. PNAS 2006</td>
</tr>
<tr>
<td>RGT Algorithm</td>
<td>W. Chen et al. DMKD, 2010</td>
</tr>
</tbody>
</table>
Performace Metrics

**COVERAGE**
Fraction of edges which are of intra-community type

**MODULARITY**
Normalized fraction of difference of intra-community edges
In the given graph and a random graph
## DATASETS

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Nodes</th>
<th>Edges</th>
<th>Triangles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karate</td>
<td>34</td>
<td>78</td>
<td>45</td>
</tr>
<tr>
<td>Dolphins</td>
<td>62</td>
<td>318</td>
<td>95</td>
</tr>
<tr>
<td>Les Miserables</td>
<td>77</td>
<td>508</td>
<td>467</td>
</tr>
<tr>
<td>Political Books</td>
<td>105</td>
<td>882</td>
<td>560</td>
</tr>
<tr>
<td>Football</td>
<td>115</td>
<td>1226</td>
<td>810</td>
</tr>
<tr>
<td>Jazz Musicians</td>
<td>198</td>
<td>274</td>
<td>17899</td>
</tr>
<tr>
<td>Email</td>
<td>1133</td>
<td>5451</td>
<td>10687</td>
</tr>
<tr>
<td>Yeast</td>
<td>2361</td>
<td>6913</td>
<td>5999</td>
</tr>
</tbody>
</table>
Karate Club Dataset with 3 Communities
Les Miserables Dataset with 7 Communities
FootBall Dataset with 12 Communities
Political Books Dataset with 5 Communities
<table>
<thead>
<tr>
<th>Data Set</th>
<th>SCoDA</th>
<th>GN</th>
<th>Greedy</th>
<th>Spectral</th>
<th>RGT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mod</td>
<td>Cov</td>
<td>Mod</td>
<td>Cov</td>
<td>Mod</td>
</tr>
<tr>
<td>Karate</td>
<td>0.4</td>
<td>82.05</td>
<td>0.4</td>
<td>75.64</td>
<td>0.38</td>
</tr>
<tr>
<td>Dolphins</td>
<td>0.525</td>
<td>80.50</td>
<td>0.519</td>
<td>82.38</td>
<td>0.495</td>
</tr>
<tr>
<td>Les Miserables</td>
<td>0.545</td>
<td>75.19</td>
<td>0.538</td>
<td>73.22</td>
<td>0.5</td>
</tr>
<tr>
<td>Political Books</td>
<td>0.524</td>
<td>89.11</td>
<td>0.496</td>
<td>91.83</td>
<td>0.509</td>
</tr>
<tr>
<td>Football</td>
<td>0.60</td>
<td>69.00</td>
<td>0.598</td>
<td>74.22</td>
<td>0.566</td>
</tr>
<tr>
<td>Jazz</td>
<td>0.439</td>
<td>79.83</td>
<td>0.403</td>
<td>77.93</td>
<td>0.438</td>
</tr>
<tr>
<td>Email</td>
<td>0.51</td>
<td>73.18</td>
<td>0.51</td>
<td>75.67</td>
<td>0.494</td>
</tr>
<tr>
<td>Yeast</td>
<td>0.571</td>
<td>67.85</td>
<td>0.568</td>
<td>74.55</td>
<td>0.571</td>
</tr>
</tbody>
</table>
SCoDA has comparable computational complexity and running time.

SCoDA maintains a good balance between Coverage and modularity.

SCoDA uses only local information.

Game theory helps solve certain KDD problems with Incomplete information.
POSSIBLE EXTENSIONS

Extend to weighted graphs, directed graphs, overlapping communities

There could be multiple Nash stable Partitions – choosing the best one is highly non-trivial
SOME MYTHS

Game theory is a panacea for solving SNA problems

Game theory makes all SNA algorithms much more efficient

Game Theory provides a complete alternative to SNA problem solving
SOME CHALLENGES

Game theory computations are among the hardest; For example, computing NE of even 2 player games is not even NP-hard!

Deciding when to use a game theoretic approach and mapping the given SNA problem into a suitable game could be non-trivial.
Some Promising Directions

- Designing scalable approximation algorithms with worst case guarantees
- Explore numerous solution concepts available in the ocean of game theory literature
- Exploit games with special structure such as convex games, potential games, matrix games, etc.
- Problems such as incentive compatible learning and social network monetization are at the cutting edge
SUMMARY

Game Theory imparts more power, more efficiency, more naturalness, and more glamour to social network analysis.

Sensational new algorithms for SNA problems? Still a long way to go but the potential is good. Calls for a much deeper study.
Questions and Answers …

Thank You …