Mean Field Games in Societal Networks

Vijay G. Subramanian
EECS Department, University of Michigan
Joint work with J. Li, B. Xia, S. Paul, R. Bhattacharya, X. Geng, H. Ming, L. Xie and Srinivas Shakkottai, Texas A&M University

May 7th 2015
Mathematical and Algorithmic Sciences Lab
France Research Center, Huawei Technologies
Large Complex Networks In Modern Society

- **Telecommunications & Social Networks**
  
  Economic & social driver of change

  ~$1.5 trillion in 2010 (2.4% of world GDP)

- **Urban transportation systems**

  Congestion losses immense

  ~$67.5 billion productivity loss (0.7% of US GDP)

  Driver for intelligent transportation systems

- **Cyber physical systems**

  Smart grids, demand response & EV charging/storage

  DoE study: savings of $46-117 billion over 20 years
Societal networks as complex networks

“Properties” of complex networks

- Many agents, controllers, measurement devices, ...
- Complex connectivity & interaction structure
- Different agents possess different information
- Distributed, decentralized or competitive sub-systems
- Conflicts between agents’ preferred solutions & system’s preferred solutions

Extremely challenging to develop or analyze optimal solutions

Societal networks are complex networks

- Interconnected networks that are important to the functioning of society.
- Have a shared resource component, and participants have to periodically take decisions on when and how much to utilize such resources.
- Desire to incentivize good behavior that would benefit society as a whole.

Examples of real-world experiments on incentivizing good behavior in societal networks

- Merugu Prabhakar Rama’09, An incentive mechanism for decongesting the roads: A pilot program in Bangalore
- Prabhakar’13, Designing large-scale nudge engines.
Mean-field games paradigm

Mean-field game setting helps in many instances

- Models typical agent behavior with many other agents
- Naturally applies to distributed settings
- Yields simpler policies & simpler calculations

JovanovicRosenthal’88, GrahamMeleard’94, Lasry-Lions’07, HuangCainesMalhame’10
Mean-field analysis

- Rich history in statistical physics
  - Analysis of many body problems, Ising model, replica method, cavity method, etc.
- Asymptotic analysis of Markov processes
  - Kurtz's theorem for population density dependent Markov processes
  - Chemical Master Equation
  - Lots of recent work on analysis of multiple-access channel (802.11)
- Rich history in economics
  - Population games, Non-atomic players
  - Competitive equilibrium analysis, Analysis of wage distribution in labor markets
Mean-field games

- Large population of agents interact
  
  A. Number of agents each agent interacts with increases without bound
  
  B. Each agent interacts with a random pool of agents during finite lifetime with extremely small chance of interacting with any particular agent twice

- Agents use simpler strategies
  
  - Best-respond to population distribution of actions
    
    - Mean-field equilibrium (MFE): a self-consistent distribution of actions
    
    - Avoids complex belief structure & strategies of equilibrium analysis
  
  - For finite population setting, these strategies typically yields $\epsilon$-Nash equilibrium with $\epsilon$ decreasing to 0 with population size
Literature Overview

- Basic Literature
  - Model A - JovanovicRosenthal’88, Lasry-Lions’07, HuangCainesMalhame’10
  - Model A - existence and uniqueness under general conditions
    AdhlakaJohariWeintraub’12, BodohCreed’13
  - Model B - GrahamMeleard’94 developed approach using Sniztman’91 Propagation of Chaos paper

- Applications of Model B
  - Dynamic auctions with learning (second price auctions) - IyerJohariSundararajan’12
  - Scheduling in cellular systems - ManjrekarRamaswamyShakkottai’14

Specific strategic behavior is analyzed: no optimality considerations, no incentives

- Multi-armed bandit games - GummadiJohariYu’12

Some optimality considerations: nudging of equilibrium not considered
Outline of rest of talk

Discuss two problems with optimality considerations, incentives and nudging of equilibrium behavior

• **Real-time wireless streaming of video with device-to-device cooperation**
  - Cooperation leads to great benefits but needs to be incentivized as users can free-ride
  - User participation and truth-telling important considerations

• **Enabling demand-response to shift consumer peak energy usage via lottery schemes**
  - Load aggregator constraints indicated to users by coupling them using lottery schemes
  - Consumer actions can be observed so gains from demand-response passed on to the consumers via lottery schemes
Real-time wireless streaming of video with device-to-device cooperation
System Overview

Users want to view common real-time video stream
TV broadcast of sports or political event
Timing and Quality of Service

Tight timing constraints

- Create block $k$ at server
  - $N$ pieces per block
- Transmit to users over cellular link at time $k-2$
  - # of received pieces $e_i[k]$
  - i.i.d. with distribution $\zeta$
- Users exchange pieces of block at time $k-1$
  - Get $T(<N)$ broadcast transmission attempts
  - $a[k]$ vector of # of transmissions
- Users play out at time $k$ if all pieces available
  - Outcome $\chi_i(a[k], e_i[k])$
Timing and Quality of Service

Track deficit of unplayed packets

- $\eta$ - Target delivery ratio
  - View as arrival rate
- Successful play out as service
- Stable queue implies delivery ratio met
- Scheme of HouBorkarKumar’09

Cost of deficit

\[ c(d_i[k]) \]

Convex, monotone increasing function

\[ d_i[k] = \begin{cases} 
  d_i[k - 1] + \eta & \text{unsuccessful delivery} \\
  d_i[k - 1] - (1 - \eta) & \text{successful playout}
\end{cases} \]
Computation of $\chi_i(a[k], e_i[k])$

- Using network coding can make all broadcast transmissions useful to recipients
- # of useful transmissions for user $i$ is $g_i(a[k])$
  - Depends on other user transmissions
- Iff $e_i[k] + g_i(a[k]) = N$, play out possible

Can devise a greedy policy to stabilize deficit queues (Abedini, et al. '13)

- Maximizes one-step drift of Lyapunov function of deficits
- Dramatically reduces cellular usage even with 4 users
- Need users to collectively report $\theta[k] = (e[k], d[k-1])$
- Why would users report truthfully? Why would they transmit for others?
- User $i$ would like to free-ride, e.g., report high $d_i[k-1]$ and 0 for $e_i[k]$
Mean-field model

- Users extremely mobile
  - J clusters & M users per cluster
  - Users randomly permuted in each frame
- Users can quit at any time with new agent taking its place: regeneration w.p. \( 1-\beta \)
  - At regeneration deficit chosen independently with distribution \( \psi \)
- System Problem: minimize discounted sum of user costs
  - Users are assumed risk neutral
Mean-field calculations

Assume truth-telling + participation in all frames

- With J large, consider distributed setting where a[k] is chosen in each cluster
  - Information flow/overhead is minimized
- State of users in cluster chosen by distribution (φ × ζ)
  - Greedy policy optimal in each cluster
- Each non-regenerating user’s state is a Markov process

Users perceive everyone else via the mean-field!

- Denote stationary distribution of deficits (with regeneration) as \( \Pi_{φ × ζ} \)

Any self-consistent \( ρ = \Pi_{φ × ζ} \) defines an MFE

\[
d_i[k] = d_i[k-1] + \eta - \chi(a(θ_i, \hat{Θ}_{-i}), θ_i)
\]

\[
\hat{Θ}_{-i} \sim (ρ × ζ)_M^{M-1}
\]
MFE & Incentives

- MFE exists
  - Uses regenerative representation of $\prod_{0 \times \zeta}$
  - Owing to discrete nature of problem need to use stronger coupling than usual Skorohod coupling
  - Use Schauder fixed point theorem for existence
- Incentives based on unique features of problem
  - Devise transfers using mechanism design ideas for incentive compatibility & individual rationality
  - Values interdependent: different from traditional VCG setting
    - Groves mechanisms still work (Radner Williams’88)
  - Dynamic setting
    - Use dynamic mechanism design concepts (Bergemann Valimaki’10)
Details on transfers

System viewpoint and user viewpoint different

- Each user’s viewpoint
  - Self via Markov process (during life-time) and all others in mean-field
- System’s viewpoint
  - All users in mean-field

Need to align these viewpoints for correct incentives

- Transfers should make each user solve system problem from its viewpoint
  - System solves: system optimal allocation, optimal allocation from each user’s viewpoint
- Using Clarke pivot mechanism transfers positive with nice structure
  - System solves: system optimal allocation without each user
  - One part of transfer subsidizes user for participation: equals loss of gains from free-riding
  - Second part of transfer pays users for simpler nature of MFE: aligns viewpoints

Obtain per frame dominant strategy incentive compatibility & interim individual rationality
Test on Android phones

- Custom Android kernel for simultaneous 3G & WiFi use
- D2D allocations implemented via 802.11 back-off scheme
- 4 smartphones used

- At 800 kbps, 16 minute video costs $1 for cellular only download
- Hybrid scheme results in 40 cents of cellular usage
- System needs to pay users 36 cents
- User receives enough savings on cellular to use scheme
Enabling demand-response to shift consumer peak energy usage via lottery schemes
Problem setting

- Load Serving Entities (LSE) or Load Aggregator (LA) pays a variable price in the market.
- Peak prices in the day-ahead market are between 3pm-8pm with 5pm-6pm being maximum.
- Corresponds to increased demand.
- AC usage of about 3 kWh between 5pm-6pm in Texan homes.
- Nudge users to reduce this usage pattern by offering coupons to win weekly lotteries.
- Play lottery with M-1 other users: couples users under an LSE/LA.
- More discomfort, more coupons & better the chances to win.
- User’s surplus increases by w upon win and decreases by l upon loss.
- Users risk neutral & measure utility by a concave + increasing function of surplus.
Model details

Mean-field setting

- Users quit & regenerate
  - Models user churn
- Users play lottery in every frame with randomly chosen M-1 other users
- Users have finite action choices & MFE is a distribution on actions

Lottery schemes

- Viewed as choosing distribution over permutations
- Different actions yield different number of coupons
- More coupons should result in better chances of winning
- Use ranking models such a Plackett-Luce model to run lottery
  - Permutation given by increasing order of exponentials with parameter # of coupons
  - Easily generalizes: multiple rounds, complex but monotonic in coupons reward structure, etc.
MFE existence proof ideas

- An action distribution results in a surplus distribution
  - Surplus of each user is a Markov process driven by the lottery outcomes
- A surplus distribution results in a poset of action distributions
  - Actions are determined by solving discounted reward maximization problem
  - Threshold policy results with map between surplus & actions set-valued
  - Can focus on extreme pair of actions at any surplus value
  - Different choices of extreme actions results in poset of action distributions
  - Upper semicontinuous structure holds & we can take convex hull
- Use Kakutani fixed point theorem for existence

Logic follows Nash’s original proof of existence of (mixed) equilibria
Case study: Texan homes

- Home air conditioning is a major part of electricity usage in Texas.
- Typical 2500 sq. ft home consumes of the order of 30 kWh per day, and about 12 kWh in the peak period.
- Can we incentivize users to move some energy consumption from the peak period to off-peak?

Table 1: Parameters for a Residential AC Unit

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C$, Thermal Capacitance</td>
<td>10 kWh/°C</td>
</tr>
<tr>
<td>$R$, Thermal Resistance</td>
<td>2 °C/kW</td>
</tr>
<tr>
<td>$P_m$, Rated Electrical Power</td>
<td>6.8 kW</td>
</tr>
<tr>
<td>$\eta$, Coefficient of Performance</td>
<td>2.5</td>
</tr>
<tr>
<td>$\tau_r$, Temperature Setpoint</td>
<td>22.5 °C</td>
</tr>
<tr>
<td>$\Delta$, Temperature Deadband</td>
<td>0.3 °C</td>
</tr>
</tbody>
</table>
Action choices

- Change 5pm-6pm energy usage by adjusting set-points
- Compute hazard for LSE as a combination of mean & standard deviation of day-ahead price, and assign coupons to periods
- Pick 5 alternate actions: specific set-points
- Compute cost to user as combination of change in mean temperature & change in standard deviation of temperature
- Coupons for action vector based on total usage in all periods based on action, relative to base-line 0
- Coupon determination steers/nudges possible MFEs

Table 1. Actions, Costs and Energy Coupons

<table>
<thead>
<tr>
<th>Index</th>
<th>Action Vector</th>
<th>Cost</th>
<th>Coupons</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(22.5, 22.5, 22.5, 22.5, 22.5)</td>
<td>0</td>
<td>108.9</td>
</tr>
<tr>
<td>1</td>
<td>(22, 22.25, 24, 22, 21.75)</td>
<td>1.774</td>
<td>834.6</td>
</tr>
<tr>
<td>2</td>
<td>(21.75, 22.25, 24, 22.25, 22)</td>
<td>1.430</td>
<td>815.4</td>
</tr>
<tr>
<td>3</td>
<td>(21.75, 22, 24, 22.25, 22)</td>
<td>1.185</td>
<td>772.7</td>
</tr>
<tr>
<td>4</td>
<td>(21.75, 22, 24, 22, 22)</td>
<td>0.838</td>
<td>637.1</td>
</tr>
<tr>
<td>5</td>
<td>(22, 22, 24, 22.25, 22.25)</td>
<td>0.697</td>
<td>509.1</td>
</tr>
</tbody>
</table>
Lottery details & results

- Users win $40: perceive gain of $39 and loss of $1 per lottery
- Users stay in lottery for 50 weeks
- Almost linear utility
- Net reduction in hazard for LSE for 50 homes is $70 per week
Conclusion & Future Directions

Summary

- Suggested Mean-Field Games as a potential tool for analysis & design of large scale societal networks
- Explored incentives: either subsidies for cooperation or lottery rewards for tough actions
- Can be easily generalized from homogeneous set-up to a finite number of classes of heterogeneous agents with complex interactions

Future directions

- Consider systems with reputation effects & learning: crowdsourcing systems
- Consider systems that allow for reselling of resources: dynamic spectrum markets, peer-to-peer systems or cloud computing systems
- For a simpler setting understand how MFG could be steered/optimized
Acknowledgements

- Collaborators
- Funding from NSF

Thank you