Mechanism Design in Shared Infrastructures

Richard Weber

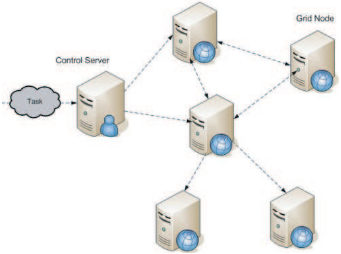
Statistical Laboratory, University of Cambridge

(work with Costas Courcoubetis, Department of Computer Science, Athens University of Economics and Business)

Workshop on Distributed Decisions via Games and Price Mechanisms
Lund Center for Control of Complex Engineering Systems
10–12 March, 2010

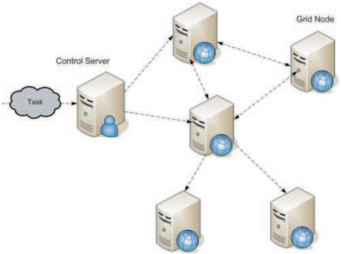
Motivation: managing shared infrastructures

Grid computing



Motivation: managing shared infrastructures

Grid computing



This is an example of a **virtual facility**, composed of shared resources, — such as computers, routers, and communication links —, which are used together to so that agents can perform tasks.

Issues for shared infrastructures

Computer Science/Engineering

Fast Internet, combined with ever increasing performance and reducing cost, has made Grid computing a reality.

Business/Economics

Adoption has been slowed because of perceived economic barriers.

Issues for shared infrastructures

Computer Science/Engineering

Fast Internet, combined with ever increasing performance and reducing cost, has made Grid computing a reality.

Amazon, IBM, and Sun are already providers of simple Grid services. However, the ultimate goal is infrastructures in which participants can be both providers and consumers.

Business/Economics

Adoption has been slowed because of perceived economic barriers.

Issues for shared infrastructures

Computer Science/Engineering

Fast Internet, combined with ever increasing performance and reducing cost, has made Grid computing a reality.

Amazon, IBM, and Sun are already providers of simple Grid services. However, the ultimate goal is infrastructures in which participants can be both providers and consumers.

Business/Economics

Adoption has been slowed because of perceived economic barriers.

GridEcon: a 'Sixth Framework Programme' of European Community, exploring the perceived economic barriers to the adoption of grid, or cloud, computing, 07/06–05/09.

Economic/business questions

- Participant asks, 'How will I benefit by participating in a shared infrastructure?'

Economic/business questions

- Participant asks, 'How will I benefit by participating in a shared infrastructure?'
- 'How are the costs and benefits going to be shared?'

Economic/business questions

- Participant asks, ‘How will I benefit by participating in a shared infrastructure?’
- ‘How are the costs and benefits going to be shared?’
- ‘How is a shared infrastructure best organized?’

Economic/business questions

- Participant asks, ‘How will I benefit by participating in a shared infrastructure?’
- ‘How are the costs and benefits going to be shared?’
- ‘How is a shared infrastructure best organized?’

Participants will be making *distributed decisions* (about their participation, contributions and usage).

Economic/business questions

- Participant asks, 'How will I benefit by participating in a shared infrastructure?'
- 'How are the costs and benefits going to be shared?'
- 'How is a shared infrastructure best organized?'

Participants will be making *distributed decisions* (about their participation, contributions and usage).

Can these decisions be coordinated and optimized through *price mechanisms* — or is something additional needed?

A ‘market solution’

The GridEcon project developed software that ‘intelligently matches requests for resource against resources which are available.’

A ‘market solution’

The GridEcon project developed software that ‘intelligently matches requests for resource against resources which are available.’

Providers and consumers of computing resources go to trade.

An organization might post that it needs 10 virtual machines of a certain type for 8 hours and state that the maximum price it is willing to pay is 100 euros.

This corresponds to a ‘bid’ in this market.

A ‘market solution’

The GridEcon project developed software that ‘intelligently matches requests for resource against resources which are available.’

Providers and consumers of computing resources go to trade.

An organization might post that it needs 10 virtual machines of a certain type for 8 hours and state that the maximum price it is willing to pay is 100 euros.

This corresponds to a ‘bid’ in this market.

Similarly, an organization can post computing resources with an ‘ask’ of the minimum price at which it is willing to sell.

A ‘market solution’

The GridEcon project developed software that ‘intelligently matches requests for resource against resources which are available.’

Providers and consumers of computing resources go to trade.

An organization might post that it needs 10 virtual machines of a certain type for 8 hours and state that the maximum price it is willing to pay is 100 euros.

This corresponds to a ‘bid’ in this market.

Similarly, an organization can post computing resources with an ‘ask’ of the minimum price at which it is willing to sell.

The market matches the asks and bids.

A ‘rule-based solution’

At each point in time, participants can either contribute resources or ask for resources.

A ‘rule-based solution’

At each point in time, participants can either contribute resources or ask for resources.

Resources are allocated amongst those who ‘ask’ according to some rule.

A ‘rule-based solution’

At each point in time, participants can either contribute resources or ask for resources.

Resources are allocated amongst those who ‘ask’ according to some rule.

Some rules are obviously bad: like ‘equal shares’, which leads to the familiar free-rider problem.

A ‘rule-based solution’

At each point in time, participants can either contribute resources or ask for resources.

Resources are allocated amongst those who ‘ask’ according to some rule.

Some rules are obviously bad: like ‘equal shares’, which leads to the familiar free-rider problem.

An alternative rule might be ‘proportional shares’.

Can we find good rules?

A ‘fully managed solution’

Let us model operation of an infrastructure.

- It is composed of resources (links, servers, buffers, etc).

A 'fully managed solution'

Let us model operation of an infrastructure.

- It is composed of resources (links, servers, buffers, etc).
- It an be operated in different ways, say $\omega \in \Omega$,
  (by scheduling, routing, bandwidth allocation, etc.)

A 'fully managed solution'

Let us model operation of an infrastructure.

- It is composed of resources (links, servers, buffers, etc).
- It an be operated in different ways, say $\omega \in \Omega$,
  (by scheduling, routing, bandwidth allocation, etc.)
- On a given day the subset of agents who wish to use the infrastructure is some $S \subseteq \{1, \ldots, n\}$.

A ‘fully managed solution’

Let us model operation of an infrastructure.

- It is composed of resources (links, servers, buffers, etc).
- It an be operated in different ways, say $\omega \in \Omega$,
  (by scheduling, routing, bandwidth allocation, etc.)
- On a given day the subset of agents who wish to use the infrastructure is some $S \subseteq \{1, \dots, n\}$.
- If operated in manner ω then agent i has benefit

$$\theta_i u_i(\omega)$$

A ‘fully managed solution’

Let us model operation of an infrastructure.

- It is composed of resources (links, servers, buffers, etc).
- It an be operated in different ways, say $\omega \in \Omega$,
  (by scheduling, routing, bandwidth allocation, etc.)
- On a given day the subset of agents who wish to use the infrastructure is some $S \subseteq \{1, \dots, n\}$.
- If operated in manner ω then agent i has benefit

$$\theta_i u_i(\omega)$$

- $u_i(\cdot)$ is pubic knowledge, but only agent i knows θ_i.

A 'fully managed solution'

Let us model operation of an infrastructure.

- It is composed of resources (links, servers, buffers, etc).
- It an be operated in different ways, say $\omega \in \Omega$,
  (by scheduling, routing, bandwidth allocation, etc.)
- On a given day the subset of agents who wish to use the infrastructure is some $S \subseteq \{1, \ldots, n\}$.
- If operated in manner ω then agent i has benefit

$$\theta_i u_i(\omega)$$

- $u_i(\cdot)$ is pubic knowledge, but only agent i knows θ_i.
- ω is to be chosen on the basis of S and declared θ_i.

Agents pay for operating cost

Fees are collected to cover some daily operating cost, c.

Agents pay for operating cost

Fees are collected to cover some daily operating cost, c.

In some situations we may take the fee as money.

Agents pay for operating cost

Fees are collected to cover some daily operating cost, c.

In some situations we may take the fee as money.

In others, we may wish to collect fees 'in kind', i.e., as contributions to the pool of resources which comprises the infrastructure.

Agents pay for operating cost

Fees are collected to cover some daily operating cost, c.

In some situations we may take the fee as money.

In others, we may wish to collect fees 'in kind', i.e., as contributions to the pool of resources which comprises the infrastructure.

Suppose the fees are money, and agent i is charged $p_i(S, \theta)$.

Agents pay for operating cost

Fees are collected to cover some daily operating cost, c.

In some situations we may take the fee as money.

In others, we may wish to collect fees 'in kind', i.e., as contributions to the pool of resources which comprises the infrastructure.

Suppose the fees are money, and agent i is charged $p_i(S, \theta)$.

The (ex-ante) budget constraint is

$$E_{S,\theta}\left[\, p_1(S, \theta) + \cdots + p_n(S, \theta)\right] \geq c$$

The incentive compatibility issue

Agent i wishes to maximize his (ex-ante) expected net benefit

$$nb_i(\theta_i) = E_{S,\theta_{-i}}\left[\theta_i u_i(\omega(S,\theta)) - p_i(S,\theta)\right]$$

The incentive compatibility issue

Agent i wishes to maximize his (ex-ante) expected net benefit

$$nb_i(\theta_i) = E_{S,\theta_{-i}}\left[\theta_i u_i(\omega(S,\theta)) - p_i(S,\theta)\right]$$

Unless $p_i(S,\theta)$ and $\omega(S,\theta)$ are chosen carefully, agent i may benefit by being untruthful in declaring θ_i.

A simple mathematical example

Consider 2 agents, both present on all days.

A simple mathematical example

Consider 2 agents, both present on all days.

Infrastructure is described by a single resource, parameterized by a number (such as computing cycles); so operating methods are:

$$\{\omega\} = \{(x_1, x_2) \, : \, x_1 + x_2 \leq 1\}$$

A simple mathematical example

Consider 2 agents, both present on all days.

Infrastructure is described by a single resource, parameterized by a number (such as computing cycles); so operating methods are:

$$\{\omega\} = \{(x_1, x_2) : x_1 + x_2 \leq 1\}$$

On day t, agent i has utility for resource of $\theta_{i,t} u(x)$, where $u(x) = x$, and $\theta_{i,t}$ is ex-ante known to be distributed $U[0, 1]$.

A simple mathematical example

Consider 2 agents, both present on all days.

Infrastructure is described by a single resource, parameterized by a number (such as computing cycles); so operating methods are:

$$\{\omega\} = \{(x_1, x_2) \,:\, x_1 + x_2 \leq 1\}$$

On day t, agent i has utility for resource of $\theta_{i,t} u(x)$, where $u(x) = x$, and $\theta_{i,t}$ is ex-ante known to be distributed $U[0,1]$.

Focus on one day; let $\theta_{i,t} = \theta_i$. Aim to maximize

$$E_{\theta_1,\theta_2}\left[\max_{x_1,x_2}\{\theta_1 u(x_1) + \theta_2 u(x_2)\} - c\right]$$

A simple mathematical example

Consider 2 agents, both present on all days.

Infrastructure is described by a single resource, parameterized by a number (such as computing cycles); so operating methods are:

$$\{\omega\} = \{(x_1, x_2) \,:\, x_1 + x_2 \leq 1\}$$

On day t, agent i has utility for resource of $\theta_{i,t} u(x)$, where $u(x) = x$, and $\theta_{i,t}$ is ex-ante known to be distributed $U[0,1]$.

Focus on one day; let $\theta_{i,t} = \theta_i$. Aim to maximize

$$E_{\theta_1,\theta_2}\left[\max_{x_1,x_2}\{\theta_1 u(x_1) + \theta_2 u(x_2)\} - c\right] = E\left[\max\{\theta_1, \theta_2\} - c\right]$$

A simple mathematical example

Consider 2 agents, both present on all days.

Infrastructure is described by a single resource, parameterized by a number (such as computing cycles); so operating methods are:

$$\{\omega\} = \{(x_1, x_2) \,:\, x_1 + x_2 \leq 1\}$$

On day t, agent i has utility for resource of $\theta_{i,t} u(x)$, where $u(x) = x$, and $\theta_{i,t}$ is ex-ante known to be distributed $U[0,1]$.

Focus on one day; let $\theta_{i,t} = \theta_i$. Aim to maximize

$$\begin{aligned} E_{\theta_1,\theta_2}\left[\max_{x_1,x_2}\{\theta_1 u(x_1) + \theta_2 u(x_2)\} - c\right] &= E\left[\max\{\theta_1, \theta_2\} - c\right] \\ &= \tfrac{2}{3} - c \end{aligned}$$

A simple mathematical example

Consider 2 agents, both present on all days.

Infrastructure is described by a single resource, parameterized by a number (such as computing cycles); so operating methods are:

$$\{\omega\} = \{(x_1, x_2) \,:\, x_1 + x_2 \le 1\}$$

On day t, agent i has utility for resource of $\theta_{i,t} u(x)$, where $u(x) = x$, and $\theta_{i,t}$ is ex-ante known to be distributed $U[0,1]$.

Focus on one day; let $\theta_{i,t} = \theta_i$. Aim to maximize

$$\begin{aligned} E_{\theta_1,\theta_2}\left[\max_{x_1,x_2}\{\theta_1 u(x_1) + \theta_2 u(x_2)\} - c\right] &= E\left[\max\{\theta_1, \theta_2\} - c\right] \\ &= \tfrac{2}{3} - c \end{aligned}$$

But as θ_1, θ_2 are unknown, this 'first best' cannot be achieved.

Comparison to auction design

Auction

Aim is to maximize seller's expected revenue:

$$E_{S,\theta}\left[\, p_1(S,\theta) + \cdots + p_n(S,\theta)\right]$$

Infrastructure optimization

Aim is to maximize expected welfare:

$$E_{S,\theta}\left[\theta_1 u_1(\omega(S,\theta)) + \cdots + \theta_n u_n(\omega(S,\theta))\right] - c$$

subject to

$$E_{S,\theta}\left[\, p_1(S,\theta) + \cdots + p_n(S,\theta)\right] \geq c$$

Both problems also have 'individual rationality' and 'incentive compatibility constraints'.

Second-best solution

In practice we do not know θ_1 and θ_2.

A 'second-best' mechanism can be constructed as follows.

If agent i declares θ_i then he is charged a fee

$$p(\theta_i) = \begin{cases} (1/2)(\theta_i^2 + \theta_0^2)\,, & \theta_i \geq \theta_0 \\ 0\,, & \theta_i < \theta_0 \end{cases}$$

He obtains $x_i = 1$ if $\theta_i = \max\{\theta_1, \theta_2\}$ and $\theta_i \geq \theta_0$.

Note that the resource is given wholly to one agent, and may be given to neither.

Properties of the 'solution'

- Agents are incentivized to be truthful.

Properties of the 'solution'

- Agents are incentivized to be truthful.
- Expected social welfare is decreasing in θ_0.

Properties of the 'solution'

- Agents are incentivized to be truthful.
- Expected social welfare is decreasing in θ_0.
- Sum of the expected payments is

$$E\Big[p(\theta_1) + p(\theta_2)\Big] = 1/3 + \theta_0^2 - (2/3)\theta_0^2\,,$$

increasing in θ_0.

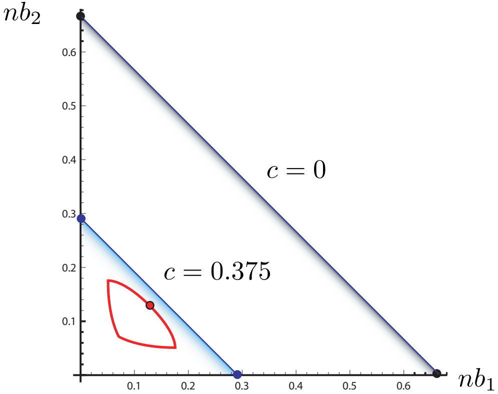
Properties of the 'solution'

- Agents are incentivized to be truthful.
- Expected social welfare is decreasing in θ_0.
- Sum of the expected payments is

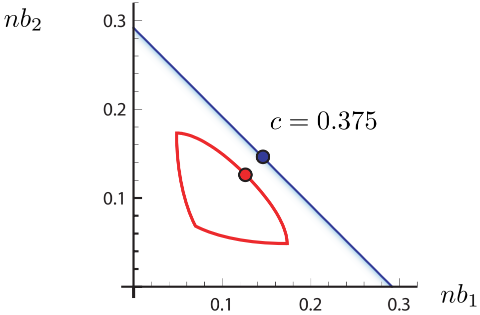
$$E\Big[p(\theta_1) + p(\theta_2)\Big] = 1/3 + \theta_0^2 - (2/3)\theta_0^2 \,,$$

  increasing in θ_0.
- Choosing θ_0 so that the above equals c, maximizes the social welfare subject to covering cost c.

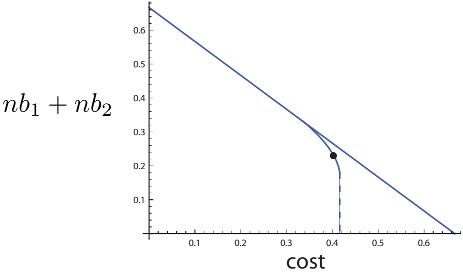
Second-best versus first-best



Second-best versus first-best



Second-best versus first-best



Expected social welfare as a function of c, compared to first-best.

For $c \in [0.333, 0.416]$ the second-best falls short of the first-best.

There is no way to cover a cost greater than $\frac{5}{12} = 0.416$.

Other second-best mechanisms

Other mechanisms can be designed that also work.

(a) There is a mechanism that has ex-post cost-covering, i.e., so that $p_1(\theta_1, \theta_2) + p_2(\theta_1, \theta_2) = c$.

Other second-best mechanisms

Other mechanisms can be designed that also work.

(a) There is a mechanism that has ex-post cost-covering, i.e., so that $p_1(\theta_1, \theta_2) + p_2(\theta_1, \theta_2) = c$.

$$p_1(\theta_1, \theta_2) = \frac{1}{2}c + \frac{1}{2}(\theta_1^2 + \theta_0^2)1_{\{\theta_1 > \theta_0\}} - \frac{1}{2}(\theta_2^2 + \theta_0^2)1_{\{\theta_2 > \theta_0\}}$$

Other second-best mechanisms

Other mechanisms can be designed that also work.

(a) There is a mechanism that has ex-post cost-covering, i.e., so that $p_1(\theta_1, \theta_2) + p_2(\theta_1, \theta_2) = c$.

$$p_1(\theta_1, \theta_2) = \tfrac{1}{2}c + \tfrac{1}{2}(\theta_1^2 + \theta_0^2)1_{\{\theta_1 > \theta_0\}} - \tfrac{1}{2}(\theta_2^2 + \theta_0^2)1_{\{\theta_2 > \theta_0\}}$$

(b) There is a mechanism that has ex-post incentive compatibility and rationality.

Other second-best mechanisms

Other mechanisms can be designed that also work.

(a) There is a mechanism that has ex-post cost-covering, i.e., so that $p_1(\theta_1, \theta_2) + p_2(\theta_1, \theta_2) = c$.

$$p_1(\theta_1, \theta_2) = \frac{1}{2}c + \frac{1}{2}(\theta_1^2 + \theta_0^2)1_{\{\theta_1 > \theta_0\}} - \frac{1}{2}(\theta_2^2 + \theta_0^2)1_{\{\theta_2 > \theta_0\}}$$

(b) There is a mechanism that has ex-post incentive compatibility and rationality.

$$p_1(\theta_1, \theta_2) = \max(\theta_0, \theta_2)1_{\{\theta_1 > \max(\theta_0, \theta_2)\}}$$

A model with true sharing

Suppose $u(x) = \sqrt{x}$. The resource is shared differently.

A model with true sharing

Suppose $u(x) = \sqrt{x}$. The resource is shared differently.
The optimal policy is found by solving a Lagrangian dual problem

$$\min_{\lambda \geq 0} \left\{ E_{\theta_1, \theta_2} \left[\max_{\substack{x_1,\, x_2 \geq 0 \\ x_1 + x_2 \leq 1}} \sum_{i=1}^{2} h_\lambda(\theta_i) u(x_i) \right] - (1+\lambda)c \right\} .$$

where $h(\theta_i) = (\theta_i + \lambda(2\theta_i - 1))$ and

$$x_i(\theta_1, \theta_2) = \frac{h_\lambda(\theta_i)^2}{\sum_{j=1}^{2} h_\lambda(\theta_j)^2}$$

A model with true sharing

Suppose $u(x) = \sqrt{x}$. The resource is shared differently.

The optimal policy is found by solving a Lagrangian dual problem

$$\min_{\lambda \geq 0} \left\{ E_{\theta_1, \theta_2} \left[\max_{\substack{x_1, x_2 \geq 0 \\ x_1 + x_2 \leq 1}} \sum_{i=1}^{2} h_\lambda(\theta_i) u(x_i) \right] - (1+\lambda)c \right\} .$$

where $h(\theta_i) = (\theta_i + \lambda(2\theta_i - 1))$ and

$$x_i(\theta_1, \theta_2) = \frac{h_\lambda(\theta_i)^2}{\sum_{j=1}^{2} h_\lambda(\theta_j)^2}$$

Fees increase with λ.

Social welfare decreases with λ, but is maximal subject to the constraint of covering the cost.

The role of the operating policy

Interestingly, the resource is not allocated in the 'most efficient' way.

That would be $x_i(\theta_1, \theta_2) = \theta_i^2/(\theta_1^2 + \theta_2^2)$.

The role of the operating policy

Interestingly, the resource is not allocated in the 'most efficient' way.

That would be $x_i(\theta_1, \theta_2) = \theta_i^2/(\theta_1^2 + \theta_2^2)$.

This is one of our most important lessons:

To optimally incentivize participation in shared infrastructures, and make the most of the resources available, one should appreciate that both (i) fee structure, and (ii) operating methods, must both play a part in providing the correct incentives to users.

Further research

- ‘Can *distributed decisions* be coordinated and optimized through *price mechanisms* — or is something additional needed?’

Further research

- ‘Can *distributed decisions* be coordinated and optimized through *price mechanisms* — or is something additional needed?’
- Contribution schemes: agent i who contributes q_i gets resource (when he requests it) in proportion to q_i^s.

Further research

- ‘Can *distributed decisions* be coordinated and optimized through *price mechanisms* — or is something additional needed?’
- Contribution schemes: agent i who contributes q_i gets resource (when he requests it) in proportion to q_i^s.
- As the number of participants becomes large the management simplifies. Allocation of the resource can be that which maximizes

$$\sum_{i \in S} \theta_i u(x_i)$$

Further research

- 'Can *distributed decisions* be coordinated and optimized through *price mechanisms* — or is something additional needed?'
- Contribution schemes: agent i who contributes q_i gets resource (when he requests it) in proportion to q_i^s.
- As the number of participants becomes large the management simplifies. Allocation of the resource can be that which maximizes

$$\sum_{i \in S} \theta_i u(x_i)$$

- Other models?

Building a facility from scratch

A different model: facility of size Q, costing $c(Q) = Q$ (per slot), is formed by initial contributions of agents.
These are incentivized to contribute because their contribution will affect the amount of resources they will get at run time.
Probably a good model for virtual Grid infrastructures.

Building a facility from scratch

A different model: facility of size Q, costing $c(Q) = Q$ (per slot), is formed by initial contributions of agents.
These are incentivized to contribute because their contribution will affect the amount of resources they will get at run time.
Probably a good model for virtual Grid infrastructures.

- $\theta_{i,t} = \theta_i$ for all t (private information).

Building a facility from scratch

A different model: facility of size Q, costing $c(Q) = Q$ (per slot), is formed by initial contributions of agents.
These are incentivized to contribute because their contribution will affect the amount of resources they will get at run time.
Probably a good model for virtual Grid infrastructures.

- $\theta_{i,t} = \theta_i$ for all t (private information).
- Agent i is 'on-off' w.p. α_i, $1 - \alpha_i$. (public information).

Building a facility from scratch

A different model: facility of size Q, costing $c(Q) = Q$ (per slot), is formed by initial contributions of agents.
These are incentivized to contribute because their contribution will affect the amount of resources they will get at run time.
Probably a good model for virtual Grid infrastructures.

- $\theta_{i,t} = \theta_i$ for all t (private information).
- Agent i is 'on-off' w.p. α_i, $1 - \alpha_i$. (public information).
- Sharing policy is $x_i(\boldsymbol{\theta}, S)$, $S=$ set of agents 'on'.

Building a facility from scratch

A different model: facility of size Q, costing $c(Q) = Q$ (per slot), is formed by initial contributions of agents.
These are incentivized to contribute because their contribution will affect the amount of resources they will get at run time.
Probably a good model for virtual Grid infrastructures.

- $\theta_{i,t} = \theta_i$ for all t (private information).
- Agent i is 'on-off' w.p. α_i, $1 - \alpha_i$. (public information).
- Sharing policy is $x_i(\boldsymbol{\theta}, S)$, $S=$ set of agents 'on'.
- System planner posts how he will compute agents' contributions and the $x_i(\boldsymbol{\theta}, S)$ as functions of the θ_is that they declare.

Building a facility from scratch

A different model: facility of size Q, costing $c(Q) = Q$ (per slot), is formed by initial contributions of agents.
These are incentivized to contribute because their contribution will affect the amount of resources they will get at run time.
Probably a good model for virtual Grid infrastructures.

- $\theta_{i,t} = \theta_i$ for all t (private information).
- Agent i is 'on-off' w.p. α_i, $1 - \alpha_i$. (public information).
- Sharing policy is $x_i(\boldsymbol{\theta}, S)$, $S=$ set of agents 'on'.
- System planner posts how he will compute agents' contributions and the $x_i(\boldsymbol{\theta}, S)$ as functions of the θ_is that they declare.
- Agents declare θ_is and system runs according to posted policy.

Analysis of proportional sharing

Consider simple case of 2 identical agents $\theta_1 = \theta_2 = 1$, $c(Q) = Q$.

Analysis of proportional sharing

Consider simple case of 2 identical agents $\theta_1 = \theta_2 = 1$, $c(Q) = Q$.

Agent i contributes q_i. Agent 1 has net benefit

$$\alpha_1(1-\alpha_2)u(x_1^{\{1\}}) + \alpha_1\alpha_2 u(x_1^{\{1,2\}}) - q_1\,.$$

Analysis of proportional sharing

Consider simple case of 2 identical agents $\theta_1 = \theta_2 = 1$, $c(Q) = Q$.

Agent i contributes q_i. Agent 1 has net benefit

$$\alpha_1(1 - \alpha_2)u(x_1^{\{1\}}) + \alpha_1\alpha_2 u(x_1^{\{1,2\}}) - q_1 \,.$$

Consider 4 possible sharing disciplines:

- Acting alone: $x_i^{\{i\}} = x_i^{\{1,2\}} = q_i$.

Analysis of proportional sharing

Consider simple case of 2 identical agents $\theta_1 = \theta_2 = 1$, $c(Q) = Q$.

Agent i contributes q_i. Agent 1 has net benefit

$$\alpha_1(1-\alpha_2)u(x_1^{\{1\}}) + \alpha_1\alpha_2 u(x_1^{\{1,2\}}) - q_1\,.$$

Consider 4 possible sharing disciplines:

- Acting alone: $x_i^{\{i\}} = x_i^{\{1,2\}} = q_i$.
- Equal sharing: $x_i^{\{i\}} = q_1 + q_2$ and $x_i^{\{1,2\}} = \frac{1}{2}(q_1 + q_2)$.

Analysis of proportional sharing

Consider simple case of 2 identical agents $\theta_1 = \theta_2 = 1$, $c(Q) = Q$.

Agent i contributes q_i. Agent 1 has net benefit

$$\alpha_1(1-\alpha_2)u(x_1^{\{1\}}) + \alpha_1\alpha_2 u(x_1^{\{1,2\}}) - q_1\,.$$

Consider 4 possible sharing disciplines:

- Acting alone: $x_i^{\{i\}} = x_i^{\{1,2\}} = q_i$.
- Equal sharing: $x_i^{\{i\}} = q_1 + q_2$ and $x_i^{\{1,2\}} = \frac{1}{2}(q_1 + q_2)$.
- Proportional sharing:

$$x_i^{\{i\}} = q_1 + q_2\,, \quad x_i^{\{1,2\}} = \frac{q_i}{q_1 + q_2}(q_1 + q_2)\,.$$

Analysis of proportional sharing

Consider simple case of 2 identical agents $\theta_1 = \theta_2 = 1$, $c(Q) = Q$.

Agent i contributes q_i. Agent 1 has net benefit

$$\alpha_1(1 - \alpha_2)u(x_1^{\{1\}}) + \alpha_1\alpha_2 u(x_1^{\{1,2\}}) - q_1 \,.$$

Consider 4 possible sharing disciplines:

- Acting alone: $x_i^{\{i\}} = x_i^{\{1,2\}} = q_i$.
- Equal sharing: $x_i^{\{i\}} = q_1 + q_2$ and $x_i^{\{1,2\}} = \frac{1}{2}(q_1 + q_2)$.
- Proportional sharing:

$$x_i^{\{i\}} = q_1 + q_2 \,, \quad x_i^{\{1,2\}} = \frac{q_i}{q_1 + q_2}(q_1 + q_2) \,.$$

- s-Proportional sharing:

$$x_i^{\{i\}} = q_1 + q_2 \,, \quad x_i^{\{1,2\}} = \frac{q_i^s}{q_1^s + q_2^s}(q_1 + q_2) \,.$$

Results for $\alpha_i = \alpha = 0.8$, $u(x) = 10 - 1/x$

scheme	social welfare	values of q_1, q_2
Acting alone	$r\alpha - 2\sqrt{\alpha}$	$\sqrt{\alpha}$
	6.21115	0.894427
Equal sharing	$r\alpha - \frac{3}{2}\sqrt{\alpha(1+\alpha)}$	$\frac{1}{2}\sqrt{\alpha(1+\alpha)}$
$s = 0$	6.2	0.6
Proportional sharing	$r\alpha - \frac{\sqrt{\alpha}\,(3+5\,\alpha)}{2\,\sqrt{1+3\,\alpha}}$	$\frac{1}{2}\sqrt{\alpha(1+3\,\alpha)}$
$s = 1$	6.30225	0.824621
Central planner	$r\alpha - \sqrt{2\alpha(1+\alpha)}$	$\sqrt{\alpha(1+\alpha)/2}$
$s = \frac{1}{2}(1 + 1/\alpha)$	6.30294	0.848528

Results for $\alpha_i = \alpha = 0.8$, $u(x) = 10 - 1/x$

scheme	social welfare	values of q_1, q_2
Acting alone	$r\alpha - 2\sqrt{\alpha}$	$\sqrt{\alpha}$
	6.21115	0.894427
Equal sharing	$r\alpha - \frac{3}{2}\sqrt{\alpha(1+\alpha)}$	$\frac{1}{2}\sqrt{\alpha(1+\alpha)}$
$s = 0$	6.2	0.6
Proportional sharing	$r\alpha - \frac{\sqrt{\alpha}\,(3+5\,\alpha)}{2\,\sqrt{1+3\,\alpha}}$	$\frac{1}{2}\sqrt{\alpha(1+3\,\alpha)}$
$s = 1$	6.30225	0.824621
Central planner	$r\alpha - \sqrt{2\alpha(1+\alpha)}$	$\sqrt{\alpha(1+\alpha)/2}$
$s = \frac{1}{2}(1 + 1/\alpha)$	6.30294	0.848528

How do these results generalize?

The exact solution

Define $g_i(\theta_i) = \theta_i - (1 - F_i(\theta_i))/f_i(\theta_i)$

E.g., $g(\theta_i) = 2\theta_i - 1$ when F_i is $U[0,1]$.

The exact solution

Define $g_i(\theta_i) = \theta_i - (1 - F_i(\theta_i))/f_i(\theta_i)$

E.g., $g(\theta_i) = 2\theta_i - 1$ when F_i is $U[0,1]$.

There is a $\lambda \geq 0$, such that for all S the optimal way to share resource amongst a set of active agents S is to maximize

$$\textstyle\sum_{i \in S}(\theta_i + \lambda g(\theta_i))u(x_i(\theta, S))\,, \tag{1}$$

over $\sum_i x_i(\theta, S) \leq Q(\theta)$.

The exact solution

Define $g_i(\theta_i) = \theta_i - (1 - F_i(\theta_i))/f_i(\theta_i)$

E.g., $g(\theta_i) = 2\theta_i - 1$ when F_i is $U[0,1]$.

There is a $\lambda \geq 0$, such that for all S the optimal way to share resource amongst a set of active agents S is to maximize

$$\sum_{i \in S} (\theta_i + \lambda g(\theta_i)) u(x_i(\theta, S)) \,, \tag{1}$$

over $\sum_i x_i(\theta, S) \leq Q(\theta)$.

Here λ is a Lagrange multiplier for a constraint

$$E\Big[\sum_i p_i(\theta)\Big] \geq E\big[c(Q(\theta))\big] \,.$$

The exact solution

Define $g_i(\theta_i) = \theta_i - (1 - F_i(\theta_i))/f_i(\theta_i)$

E.g., $g(\theta_i) = 2\theta_i - 1$ when F_i is $U[0, 1]$.

There is a $\lambda \geq 0$, such that for all S the optimal way to share resource amongst a set of active agents S is to maximize

$$\textstyle\sum_{i \in S} (\theta_i + \lambda g(\theta_i)) u(x_i(\theta, S)) \,, \tag{1}$$

over $\sum_i x_i(\theta, S) \leq Q(\theta)$.

Here λ is a Lagrange multiplier for a constraint

$$E\Big[\textstyle\sum_i p_i(\theta)\Big] \geq E\big[c(Q(\theta))\big] \,.$$

Note $g(\theta_i)$ is increasing in θ_i, but $E[g(\theta_i)] = 0$.

So an agent who declares a greater θ_i is receives more than a market allocation would give him when sharing the resource.

A large N solution

- Assume $c(Q) = Q$ and N very large.

A large N solution

- Assume $c(Q) = Q$ and N very large.
- Then S will always be near its typical value and $x_i(\boldsymbol{\theta}, S)$ become $x_i(\boldsymbol{\theta}) \approx E_S[x_i(\boldsymbol{\theta}, S)]$.

A large N solution

- Assume $c(Q) = Q$ and N very large.
- Then S will always be near its typical value and $x_i(\boldsymbol{\theta}, S)$ become $x_i(\boldsymbol{\theta}) \approx E_S[x_i(\boldsymbol{\theta}, S)]$.
- The allocations should satisfy $\sum_i \alpha_i x_i(\boldsymbol{\theta}) \leq Q$.

A large N solution

- Assume $c(Q) = Q$ and N very large.
- Then S will always be near its typical value and $x_i(\boldsymbol{\theta}, S)$ become $x_i(\boldsymbol{\theta}) \approx E_S[x_i(\boldsymbol{\theta}, S)]$.
- The allocations should satisfy $\sum_i \alpha_i x_i(\boldsymbol{\theta}) \leq Q$.

A large N solution

- Assume $c(Q) = Q$ and N very large.
- Then S will always be near its typical value and $x_i(\boldsymbol{\theta}, S)$ become $x_i(\boldsymbol{\theta}) \approx E_S[x_i(\boldsymbol{\theta}, S)]$.
- The allocations should satisfy $\sum_i \alpha_i x_i(\boldsymbol{\theta}) \leq Q$.

It turns out that the solution of the Mechanism Design problem implies a simple 'effective bandwidth' tariff for type i agents:

- System guarantees (with prob (1-ϵ)) resource y for a contribution of $\alpha_i y$ ($\alpha_i(1+\epsilon)y$).
- Agent i indirectly declares his θ_i by selecting y to maximize $\max_y\{\theta_i u(y) - \alpha_i y\}$.
- No information on F_i required!

Declaring activity frequencies

Now the α_i are private information, i.i.d. uniform on $[0, 1]$, and $\theta_{i,t} = \theta_i = 1$. Sensible if accounting of activity is costly.

Declaring activity frequencies

Now the α_i are private information, i.i.d. uniform on $[0, 1]$, and $\theta_{i,t} = \theta_i = 1$. Sensible if accounting of activity is costly.

The facility is built from agent contributions.
We wish to compute the set of optimal tariffs $q(\omega), x(\omega)$ parametrized by ω the 'type' of the agent, where an agent that contributes $q(\omega)$ gets $x(\omega)$ when he is 'on'.

Declaring activity frequencies

Now the α_i are private information, i.i.d. uniform on $[0,1]$, and $\theta_{i,t} = \theta_i = 1$. Sensible if accounting of activity is costly.

The facility is built from agent contributions.
We wish to compute the set of optimal tariffs $q(\omega), x(\omega)$ parametrized by ω the 'type' of the agent, where an agent that contributes $q(\omega)$ gets $x(\omega)$ when he is 'on'.

An agent maximizes his net benefit $f(\alpha)$, where

$$f(\alpha) = \max\left\{ \max_{\omega} \left[\alpha u(x(\omega)) - q(\omega) \right], 0 \right\}.$$

So need $d[\alpha u(x(\omega)) - q(\omega)]/d\omega|_{\omega=\alpha} = \alpha u'(\alpha) - g'(\alpha) = 0$.

So if an agent with α^* has net benefit 0 then

$$q(\alpha) = \alpha u(x(\alpha)) - \int_{\alpha^*}^{\alpha} u(x(\omega))\, d\omega\,.$$

giving

$$\int_{\alpha^*}^{1} q(\alpha)\, d\alpha = \int_{\alpha^*}^{1} (2\alpha - 1) u(x(\alpha))\, d\alpha\,.$$

So if an agent with α^* has net benefit 0 then

$$q(\alpha) = \alpha u(x(\alpha)) - \int_{\alpha^*}^{\alpha} u(x(\omega))\, d\omega\,.$$

giving

$$\int_{\alpha^*}^{1} q(\alpha)\, d\alpha = \int_{\alpha^*}^{1} (2\alpha - 1) u(x(\alpha))\, d\alpha\,.$$

The resource constraint is

$$\int_{0}^{1} \left[\alpha x(\alpha) - q(\alpha)\right]\, d\alpha \leq 0$$

So if an agent with α^* has net benefit 0 then

$$q(\alpha) = \alpha u(x(\alpha)) - \int_{\alpha^*}^{\alpha} u(x(\omega))\, d\omega\,.$$

giving

$$\int_{\alpha^*}^{1} q(\alpha)\, d\alpha = \int_{\alpha^*}^{1} (2\alpha - 1) u(x(\alpha))\, d\alpha\,.$$

The resource constraint is

$$\int_{0}^{1} \left[\alpha x(\alpha) - q(\alpha)\right]\, d\alpha \leq 0$$

So we seek to maximize a Lagrangian

$$L = \int_{\alpha^*}^{1} \Big[(\alpha + \lambda(2\alpha - 1))u(x(\alpha)) - (1+\lambda)\alpha x(\alpha)\Big]\, d\alpha\,,$$

For $u(x) = \sqrt{x}$, this gives

$$x(\omega) = \left(\frac{2\lambda + 1}{2(\lambda + 1)} - \frac{\lambda}{2(\lambda + 1)\omega}\right)^2$$

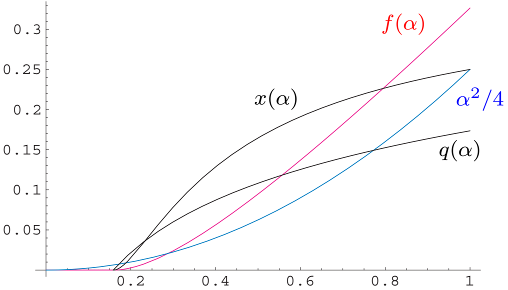
We find the correct λ by minimizing with respect to λ, giving $\lambda = 0.232206$. So for $\omega \geq 0.158566$,

$$q(\omega) = 0.173521 + 0.0942239 \log \omega$$
$$x(\omega) = \left(0.594224 - \frac{0.0942239}{\omega}\right)^2$$

and $q(\omega) = x(\omega) = 0$ for $\omega < 0.158566$ ($= \lambda/(1 + 2\lambda)$).

Note that agents with small α (less than $\alpha^* = 0.158566$) are prevented from participating.

The optimal solution for $u(x) = \sqrt{x}$



The black lines show $q(\alpha)$ and $x(\alpha)$, with $q(\alpha) < x(\alpha)$ when $\alpha > 0.2339$.
The red line is the net benefit $f(\alpha) = tx(\alpha) - q(\alpha)$.
The the blue line is $\alpha^2/4$, the net benefit obtained acting alone.

Note that some agents would prefer self-provisioning.

Conclusions

- In most realistic resource allocation problems there is private information to participants.

Conclusions

- In most realistic resource allocation problems there is private information to participants.
- Resource allocation policies need to take account of need to give right incentives. To encourage agents who value the resource more to say so, and so be willing to contribute more towards the cost, we need to reward them better than an internal market would do. But figuring out exactly how to do this is not a simple task!

Conclusions

- In most realistic resource allocation problems there is private information to participants.
- Resource allocation policies need to take account of need to give right incentives. To encourage agents who value the resource more to say so, and so be willing to contribute more towards the cost, we need to reward them better than an internal market would do. But figuring out exactly how to do this is not a simple task!
- Simple-minded sharing policies (like proportional sharing) may not to produce sufficient incentives for participants to contribute resources.

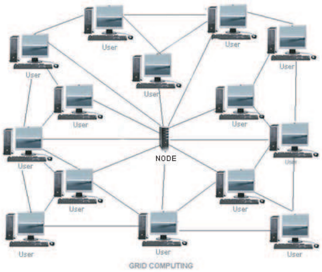
Conclusions

- In most realistic resource allocation problems there is private information to participants.
- Resource allocation policies need to take account of need to give right incentives. To encourage agents who value the resource more to say so, and so be willing to contribute more towards the cost, we need to reward them better than an internal market would do. But figuring out exactly how to do this is not a simple task!
- Simple-minded sharing policies (like proportional sharing) may not to produce sufficient incentives for participants to contribute resources.
- Many new interesting problems!!!

Motivation: Grid Computing

Grid Computing

A virtual computer composed of a cluster of networked, loosely coupled computers, acting in concert to perform very large tasks.



A simple example with full information

- Discrete time with slots $1, 2, \dots$.
- Facility of size Q (either given, or to be determined), costing c per slot to operate.
- In slot t agent i has utility $\theta_{i,t}\sqrt{x_i}$, where $\theta_{i,t}$ are i.i.d. $\sim F_i$.

A simple example with full information

- Discrete time with slots $1, 2, \ldots$.
- Facility of size Q (either given, or to be determined), costing c per slot to operate.
- In slot t agent i has utility $\theta_{i,t}\sqrt{x_i}$, where $\theta_{i,t}$ are i.i.d. $\sim F_i$.

Resource sharing problem: At each time t allocate resource to maximize sum of utilities, and obtain payments to cover the cost.

$$\underset{\{x_i\}}{\text{maximize}} \sum_{i=1}^N \theta_{i,t}\sqrt{x_i}\,, \quad \text{such that } \sum_{i=1}^N x_i \le Q.$$

A simple example with full information

- Discrete time with slots $1, 2, \ldots$.
- Facility of size Q (either given, or to be determined), costing c per slot to operate.
- In slot t agent i has utility $\theta_{i,t}\sqrt{x_i}$, where $\theta_{i,t}$ are i.i.d. $\sim F_i$.

Resource sharing problem: At each time t allocate resource to maximize sum of utilities, and obtain payments to cover the cost.

$$\underset{\{x_i\}}{\text{maximize}} \sum_{i=1}^N \theta_{i,t}\sqrt{x_i}\,, \quad \text{such that } \sum_{i=1}^N x_i \le Q.$$

Solution:

$$x_i = \frac{\theta_{i,t}^2}{\sum_{k=1}^N \theta_{k,t}^2} Q\,,\ V_{i,t} = \frac{\theta_{i,t}^2}{\sqrt{\sum_{k=1}^N \theta_{k,t}^2}}\sqrt{Q}\,,\ V_t = \sqrt{\sum_{k=1}^N \theta_{k,t}^2}\,\sqrt{Q}\,.$$

A simple example with full information

- Discrete time with slots $1, 2, \ldots$.
- Facility of size Q (either given, or to be determined), costing c per slot to operate.
- In slot t agent i has utility $\theta_{i,t}\sqrt{x_i}$, where $\theta_{i,t}$ are i.i.d. $\sim F_i$.

Resource sharing problem: At each time t allocate resource to maximize sum of utilities, and obtain payments to cover the cost.

$$\underset{\{x_i\}}{\text{maximize}} \sum_{i=1}^N \theta_{i,t}\sqrt{x_i}\,, \quad \text{such that } \sum_{i=1}^N x_i \le Q.$$

Solution:

$$x_i = \frac{\theta_{i,t}^2}{\sum_{k=1}^N \theta_{k,t}^2} Q\,,\ V_{i,t} = \frac{\theta_{i,t}^2}{\sqrt{\sum_{k=1}^N \theta_{k,t}^2}}\sqrt{Q}\,,\ V_t = \sqrt{\sum_{k=1}^N \theta_{k,t}^2}\,\sqrt{Q}\,.$$

If $E[V_t] > c$, we could ask agent i to make payment p_i such that $E[V_{i,t}] \ge p_i$ and $\sum_i p_i = c$.

Note that this is not the same as setting a price p and then letting agent i to buy x_i to maximize

$$\theta_{i,t}\sqrt{x_i} - px_i\,,$$

where we choose p so that $\sum_{i=1}^{N} x_i \leq Q$.

The problem is that $p\sum_k x_k$ does not necessarily cover cost c.

The general model

- N agents share common resource Q **over time** $t = 1, 2, \ldots$.

The general model

- N agents share common resource Q **over time** $t = 1, 2, \ldots$.
- At time t agent i has utility $\theta_{i,t} u(x_i)$, where
  - $u(\cdot)$ convex increasing,
  - x_i is the amount of resource allocated to agent i,
  - $\theta_{i,t}$ is its *personalization parameter*, i.i.d., realized from some distribution F_i (e.g., $U[0, \phi_i]$).

The general model

- N agents share common resource Q **over time** $t = 1, 2, \ldots$.
- At time t agent i has utility $\theta_{i,t} u(x_i)$, where
  - $u(\cdot)$ convex increasing,
  - x_i is the amount of resource allocated to agent i,
  - $\theta_{i,t}$ is its *personalization parameter*, i.i.d., realized from some distribution F_i (e.g., $U[0, \phi_i]$).
- System operates according to some 'rules of a game' G.

The general model

- N agents share common resource Q **over time** $t = 1, 2, \ldots$.
- At time t agent i has utility $\theta_{i,t} u(x_i)$, where
  - $u(\cdot)$ convex increasing,
  - x_i is the amount of resource allocated to agent i,
  - $\theta_{i,t}$ is its *personalization parameter*, i.i.d., realized from some distribution F_i (e.g., $U[0, \phi_i]$).
- System operates according to some 'rules of a game' G.
- **Aim:** design G so that at the Nash equilibrium the expected sum of net benefits of the agents is maximized while recovering a cost c per slot in the long run.

The general model

- N agents share common resource Q **over time** $t = 1, 2, \ldots$.
- At time t agent i has utility $\theta_{i,t} u(x_i)$, where
  - $u(\cdot)$ convex increasing,
  - x_i is the amount of resource allocated to agent i,
  - $\theta_{i,t}$ is its *personalization parameter*, i.i.d., realized from some distribution F_i (e.g., $U[0, \phi_i]$).
- System operates according to some 'rules of a game' G.
- **Aim:** design G so that at the Nash equilibrium the expected sum of net benefits of the agents is maximized while recovering a cost c per slot in the long run.

The size Q may be determined as part of the game, given $c(Q)$.

The general model

- N agents share common resource Q **over time** $t = 1, 2, \ldots$.
- At time t agent i has utility $\theta_{i,t} u(x_i)$, where
  - $u(\cdot)$ convex increasing,
  - x_i is the amount of resource allocated to agent i,
  - $\theta_{i,t}$ is its *personalization parameter*, i.i.d., realized from some distribution F_i (e.g., $U[0, \phi_i]$).
- System operates according to some 'rules of a game' G.
- **Aim:** design G so that at the Nash equilibrium the expected sum of net benefits of the agents is maximized while recovering a cost c per slot in the long run.

The size Q may be determined as part of the game, given $c(Q)$.

Each agent should be better of by participating in this system than by building his own facility.

The game G

System designer posts operating rules of the facility,

1. Participants chose contracts (*), i.e., they implicitly reveal something of their private information. They may make some initial payments depending on the contracts chosen.

The game G

System designer posts operating rules of the facility,

1. Participants chose contracts (*), i.e., they implicitly reveal something of their private information. They may make some initial payments depending on the contracts chosen.

2. The facility operates in discrete slots $t = 1, 2, \ldots$. At each t the agents derive value from the facility and may make further payments to cover the running cost. Agents may also be asked to disclose some further private information regarding their utility of the service at time t (**).

The game G

System designer posts operating rules of the facility,

1. Participants chose contracts (*), i.e., they implicitly reveal something of their private information. They may make some initial payments depending on the contracts chosen.

2. The facility operates in discrete slots $t = 1, 2, \ldots$. At each t the agents derive value from the facility and may make further payments to cover the running cost. Agents may also be asked to disclose some further private information regarding their utility of the service at time t (**).

3. The resource sharing and the payment policies take into account the information provided in (*) and (**).

Example: allocating a single item

We wish to share a single machine between 2 agents. On day t agent i has utility $\theta_{i,t}$, where $F_1 = U[0,1]$ and $F_2 = U[0,2]$ are distributions that are known to system operator. How do we allocate the machine and take payments to cover the cost c?

Example: allocating a single item

We wish to share a single machine between 2 agents. On day t agent i has utility $\theta_{i,t}$, where $F_1 = U[0,1]$ and $F_2 = U[0,2]$ are distributions that are known to system operator. How do we allocate the machine and take payments to cover the cost c?

Lets first consider a simple intuitive policy (A1):

Example: allocating a single item

We wish to share a single machine between 2 agents. On day t agent i has utility $\theta_{i,t}$, where $F_1 = U[0, 1]$ and $F_2 = U[0, 2]$ are distributions that are known to system operator. How do we allocate the machine and take payments to cover the cost c?

Lets first consider a simple intuitive policy (A1):

- Set a price p for using the machine.

Example: allocating a single item

We wish to share a single machine between 2 agents. On day t agent i has utility $\theta_{i,t}$, where $F_1 = U[0,1]$ and $F_2 = U[0,2]$ are distributions that are known to system operator. How do we allocate the machine and take payments to cover the cost c?

Lets first consider a simple intuitive policy (A1):

- Set a price p for using the machine.
- If just one agent wants to use it, he gets it and pays p
  E.g. agent 2 gets it if $\theta_{1,t} < p < \theta_{2,t}$

Example: allocating a single item

We wish to share a single machine between 2 agents. On day t agent i has utility $\theta_{i,t}$, where $F_1 = U[0,1]$ and $F_2 = U[0,2]$ are distributions that are known to system operator. How do we allocate the machine and take payments to cover the cost c?

Lets first consider a simple intuitive policy (A1):

- Set a price p for using the machine.
- If just one agent wants to use it, he gets it and pays p
  E.g. agent 2 gets it if $\theta_{1,t} < p < \theta_{2,t}$
- If both agents want to use it, agent 2 gets it and pays p.

Example: allocating a single item

We wish to share a single machine between 2 agents. On day t agent i has utility $\theta_{i,t}$, where $F_1 = U[0,1]$ and $F_2 = U[0,2]$ are distributions that are known to system operator. How do we allocate the machine and take payments to cover the cost c?

Lets first consider a simple intuitive policy (A1):

- Set a price p for using the machine.
- If just one agent wants to use it, he gets it and pays p
  E.g. agent 2 gets it if $\theta_{1,t} < p < \theta_{2,t}$
- If both agents want to use it, agent 2 gets it and pays p.
- Choose p so that expected payment per time slot is c.

Using an auction

Can we do any better? The theory of optimal auctions suggests the following scheme.

Using an auction

Can we do any better? The theory of optimal auctions suggests the following scheme.

- Pick a nonnegative real number b and construct the prices

$$p_1 = \max\{b, \theta_{2,t} - b\},\ p_2 = \max\{2b, \theta_{1,t} + b\}\,.$$

Using an auction

Can we do any better? The theory of optimal auctions suggests the following scheme.

- Pick a nonnegative real number b and construct the prices

$$p_1 = \max\{b, \theta_{2,t} - b\},\ p_2 = \max\{2b, \theta_{1,t} + b\}\,.$$

- Ask agents to reveal their $\theta_{i,t}$. If $\theta_{i,t} > p_i$, the machine is allocated to agent i and he pays p_i (This cannot hold for both 1 and 2.)

Using an auction

Can we do any better? The theory of optimal auctions suggests the following scheme.

- Pick a nonnegative real number b and construct the prices

$$p_1 = \max\{b, \theta_{2,t} - b\},\ p_2 = \max\{2b, \theta_{1,t} + b\}\,.$$

- Ask agents to reveal their $\theta_{i,t}$. If $\theta_{i,t} > p_i$, the machine is allocated to agent i and he pays p_i (This cannot hold for both 1 and 2.)

- Choose b so that expected payment per slot is c.

Using an auction

Can we do any better? The theory of optimal auctions suggests the following scheme.

- Pick a nonnegative real number b and construct the prices

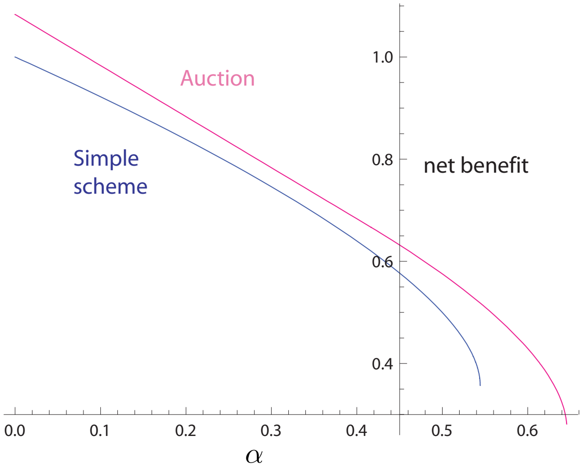
$$p_1 = \max\{b, \theta_{2,t} - b\},\ p_2 = \max\{2b, \theta_{1,t} + b\}\,.$$

- Ask agents to reveal their $\theta_{i,t}$. If $\theta_{i,t} > p_i$, the machine is allocated to agent i and he pays p_i (This cannot hold for both 1 and 2.)

- Choose b so that expected payment per slot is c.

This solution maximizes sum of expected agent utilities conditional on recovering c on the average, is incentive compatible. Note that agent 1 can win even if $\theta_{1,t} < \theta_{2,t}$.

Comparing the policies



What did we learn?

The cost coverage constraint requires we make inefficient allocations. I.e., we should not simply allocate the resource to the agent who declares the greatest θ_i.

What did we learn?

The cost coverage constraint requires we make inefficient allocations. I.e., we should not simply allocate the resource to the agent who declares the greatest θ_i.

But if there is **infinite repetition** over many slots, this simplifies the problem significantly. If we know the types of the agents and can police them we can easily achieve efficient allocation:

What did we learn?

The cost coverage constraint requires we make inefficient allocations. I.e., we should not simply allocate the resource to the agent who declares the greatest θ_i.

But if there is **infinite repetition** over many slots, this simplifies the problem significantly. If we know the types of the agents and can police them we can easily achieve efficient allocation:

1. Assume that the system will run with full information at each t; charge each agent a fixed payment/slot less than his average utility/slot, to cover cost.
2. Ask the value of the $\theta_{i,t}$ at each t and share the server efficiently; police the declarations according to the F_is.

What did we learn?

The cost coverage constraint requires we make inefficient allocations. I.e., we should not simply allocate the resource to the agent who declares the greatest θ_i.

But if there is **infinite repetition** over many slots, this simplifies the problem significantly. If we know the types of the agents and can police them we can easily achieve efficient allocation:

1. Assume that the system will run with full information at each t; charge each agent a fixed payment/slot less than his average utility/slot, to cover cost.
2. Ask the value of the $\theta_{i,t}$ at each t and share the server efficiently; police the declarations according to the F_is.

This scheme is incentive compatible (not immediate, requires a proof).

What did we learn?

The cost coverage constraint requires we make inefficient allocations. I.e., we should not simply allocate the resource to the agent who declares the greatest θ_i.

But if there is **infinite repetition** over many slots, this simplifies the problem significantly. If we know the types of the agents and can police them we can easily achieve efficient allocation:

1. Assume that the system will run with full information at each t; charge each agent a fixed payment/slot less than his average utility/slot, to cover cost.
2. Ask the value of the $\theta_{i,t}$ at each t and share the server efficiently; police the declarations according to the F_is.

This scheme is incentive compatible (not immediate, requires a proof).

Open problem: optimal scheme if we do not know the F_is?

Issues for this talk

Problem of private information. What are sensible resource allocation policies in shared infrastructures when participants have private information?

Issues for this talk

Problem of private information. What are sensible resource allocation policies in shared infrastructures when participants have private information?

Naive policies (like ‘internal market’, or ‘equal sharing’) may not be suitable.

Issues for this talk

Problem of private information. What are sensible resource allocation policies in shared infrastructures when participants have private information?

Naive policies (like 'internal market', or 'equal sharing') may not be suitable.

We wish to

- eliminate the free-rider problem;
- incentivize agents to truthfully reveal private information.

Issues for this talk

Problem of private information. What are sensible resource allocation policies in shared infrastructures when participants have private information?

Naive policies (like 'internal market', or 'equal sharing') may not be suitable.

We wish to

- eliminate the free-rider problem;
- incentivize agents to truthfully reveal private information.

Key observation: agents will adopt strategies that depend on how a system is operated.

Example: a Bridge

A bridge may or may not be built. There are 2 potential users.

Mathematical Bridge, Queens' College, Cambridge

Example: a Bridge

A bridge may or may not be built. There are 2 potential users.

Mathematical Bridge, Queens' College, Cambridge

If it is built (at cost $\$1$) then user i benefits by $\$\theta_i$.
Knowing θ_1 and θ_2, we should build the bridge if $\theta_1 + \theta_2 > 1$.

Example: a Bridge

A bridge may or may not be built. There are 2 potential users.

Mathematical Bridge, Queens' College, Cambridge

If it is built (at cost $\$1$) then user i benefits by $\$\theta_i$.
Knowing θ_1 and θ_2, we should build the bridge if $\theta_1 + \theta_2 > 1$.

If we build the bridge we must charge for the cost.
Suppose we decide to charge user i a fee of $\theta_i/(\theta_1 + \theta_2)$.

Problem: user i has incentive to under-report his true value of θ_i.

Example: a Bridge

A bridge may or may not be built. There are 2 potential users.

Mathematical Bridge, Queens' College, Cambridge

If it is built (at cost $\$1$) then user i benefits by $\$\theta_i$.
Knowing θ_1 and θ_2, we should build the bridge if $\theta_1 + \theta_2 > 1$.

If we build the bridge we must charge for the cost.
Suppose we decide to charge user i a fee of $\theta_i/(\theta_1 + \theta_2)$.

Problem: user i has incentive to under-report his true value of θ_i.

Fees should incentivize users to truthfully reveal θ_1, θ_2, with

$$p_1(\theta_1, \theta_2) + p_2(\theta_1, \theta_2) = 1 \text{ or } 0\,,\ \text{as bridge is built or not built}\,.$$

Example: a Bridge

In the example above the bridge is a **nonrivalrous good**.

Example: a Bridge

In the example above the bridge is a **nonrivalrous good**.

But what if its use is **rivalrous**?

E.g., suppose that on each given day only one person may use the bridge. On each day, users 1 and 2 wish to use the bridge, independently, with probabilities α_1 and α_2.

Example: a Bridge

In the example above the bridge is a **nonrivalrous good**.

But what if its use is **rivalrous**?

E.g., suppose that on each given day only one person may use the bridge. On each day, users 1 and 2 wish to use the bridge, independently, with probabilities α_1 and α_2.

Now we must decide (as functions of the initially declared θ_1, θ_2)

- whether or not the bridge is built;

Example: a Bridge

In the example above the bridge is a **nonrivalrous good**.

But what if its use is **rivalrous**?

E.g., suppose that on each given day only one person may use the bridge. On each day, users 1 and 2 wish to use the bridge, independently, with probabilities α_1 and α_2.

Now we must decide (as functions of the initially declared θ_1, θ_2)

- whether or not the bridge is built;
- what contributions the users should make towards its cost;

Example: a Bridge

In the example above the bridge is a **nonrivalrous good**.

But what if its use is **rivalrous**?

E.g., suppose that on each given day only one person may use the bridge. On each day, users 1 and 2 wish to use the bridge, independently, with probabilities α_1 and α_2.

Now we must decide (as functions of the initially declared θ_1, θ_2)

- whether or not the bridge is built;
- what contributions the users should make towards its cost;
- who gets to use the bridge on those days that both users say that they wish to do so.

Motivation

Similarly, in grid computing:

- how do we incentivize agents to participate and contribute computational resource?
- what size of computational resource will be installed?
- what contributions should agents make towards its cost — or what amounts of resource should they be willing to contribute?
- how should the resource be shared?

Motivation

Similarly, in grid computing:

- how do we incentivize agents to participate and contribute computational resource?
- what size of computational resource will be installed?
- what contributions should agents make towards its cost — or what amounts of resource should they be willing to contribute?
- how should the resource be shared?

Are auction and mechanism design theory appropriate? And under what assumptions on our model are these applicable?

What is fundamentally new in this problem?

Can we describe optimal policies?

Our infrastructure optimization problem

Our infrastructure optimization problem is as follows.

- Say how the infrastructure will be operated for possible subset of users S.
- Say what fees will be collected from users.

Our infrastructure optimization problem

Our infrastructure optimization problem is as follows.

- Say how the infrastructure will be operated for possible subset of users S.
- Say what fees will be collected from users.

Do the two things above, as function of declared θ_i, so that:

1. Users find it in their best interest to truthfully reveal their θ_i.
2. Users see positive expected net benefit from participation.
3. Expected total fees cover the daily running cost, say c.
4. Expected social welfare (total net benefit) is maximized

Assumptions

Two possibilities:

1. The facility already exists; it has fixed size Q and known operating cost, or
2. The facility does not exist; its size will be the sum of participants' contributions.

Assumptions

Two possibilities:

1. The facility already exists; it has fixed size Q and known operating cost, or
2. The facility does not exist; its size will be the sum of participants' contributions.

How to share resources and recover costs?

Assumptions

Two possibilities:

1. The facility already exists; it has fixed size Q and known operating cost, or
2. The facility does not exist; its size will be the sum of participants' contributions.

How to share resources and recover costs?

- Easy when we know utilities of participants.

Assumptions

Two possibilities:

1. The facility already exists; it has fixed size Q and known operating cost, or
2. The facility does not exist; its size will be the sum of participants' contributions.

How to share resources and recover costs?

- Easy when we know utilities of participants.
- In practice agents' utilities are private information.
  We must design the system to operate well, under the constraint that each agent will reveal information in a manner that is to his best advantage.

Example: scheduling a server

- Suppose N agents share a single server. Agent i generates a jobs as a Poisson process of rate λ_i, whose service times are exponentially distributed with parameter 1.

Example: scheduling a server

- Suppose N agents share a single server. Agent i generates a jobs as a Poisson process of rate λ_i, whose service times are exponentially distributed with parameter 1.

- Initially, agents contribute resource amounts $y_1, \ldots, y_N$. This results in a server of rate $\sum_k y_k$. Under FCFS scheduling all jobs have mean waiting time $1/(\sum_k y_k - \sum_k \lambda_k)$.

Example: scheduling a server

- Suppose N agents share a single server. Agent i generates a jobs as a Poisson process of rate λ_i, whose service times are exponentially distributed with parameter 1.

- Initially, agents contribute resource amounts $y_1, \ldots, y_N$. This results in a server of rate $\sum_k y_k$. Under FCFS scheduling all jobs have mean waiting time $1/(\sum_k y_k - \sum_k \lambda_k)$.

- Agent i suffers delay cost, so his net benefit is, say,

$$nb_i = \lambda_i r - \theta_i \lambda_i \frac{1}{\sum_k y_k - \sum_k \lambda_k} - y_i \,.$$

θ_i is private information of agent i, but it has an *a priori* distribution that is public information.

Optimal queue scheduling

Instead of declaring contributions they are willing to make, we can imagine that agents (equivalently) declare their θ_i.

Suppose $\theta_1 < \theta_2 < \cdots < \theta_n$.

As a function of these declarations we take contributions of the form $y(\theta_i)$ from some subset of agents $i = 1, \ldots, j$ (a set with smallest θ_i).

Optimal queue scheduling

Instead of declaring contributions they are willing to make, we can imagine that agents (equivalently) declare their θ_i.

Suppose $\theta_1 < \theta_2 < \cdots < \theta_n$.

As a function of these declarations we take contributions of the form $y(\theta_i)$ from some subset of agents $i = 1, \ldots, j$ (a set with smallest θ_i).

We employ a priority scheduling policy in which priority is always given to the current job belonging to the agent with greatest θ_i.

Optimal queue scheduling

Instead of declaring contributions they are willing to make, we can imagine that agents (equivalently) declare their θ_i.

Suppose $\theta_1 < \theta_2 < \cdots < \theta_n$.

As a function of these declarations we take contributions of the form $y(\theta_i)$ from some subset of agents $i = 1, \ldots, j$ (a set with smallest θ_i).

We employ a priority scheduling policy in which priority is always given to the current job belonging to the agent with greatest θ_i.

Under this scheme, an agent with too great a θ_i will find unprofitable to consider participating.

$y_i(\theta_i)$ is increasing in θ_i, and is determined by an incentive compatibility condition.