MULTI-RESOURCE ALLOCATION

FAIRNESS-EFFICIENCY TRADEOFFS IN A UNIFYING FRAMEWORK

Carlee Joe-Wong
Princeton University

Soumya Sen
Princeton University

Tian Lan
George Washington University

Mung Chiang
Princeton University

March 28, 2012

IEEE Infocom
What is Fairness?

- Politics, economics, sociology, engineering...

**How do you allocate a resource to different users?**

- Variance, Jain’s index, entropy (see TR for references)...

- Isoelastic or $\alpha$-fairness

- Unifying axiomatic theory of decomposable fairness measures

\[ \text{sgn}(1 - \beta) \left( \sum_{i=1}^{n} \left( \frac{x_i}{\sum_{j=1}^{n} x_j} \right)^{1-\beta} \right)^{\frac{1}{\beta}} \left( \sum_{i=1}^{n} x_i \right)^{\lambda} \]

Our Question

• Suppose you have multiple non-substitutable resources.
  • Memory
  • CPU
  • Bandwidth
• They combine to make something...
  • Jobs in a datacenter
• that multiple people want.
  • Different bundles of resource requirements
• But the resources are finite.

How do you allocate the resources to different people?
Two-Resource Example
Generalized Fairness on Jobs (GFJ)

\[ f_{\beta, \lambda}(x) = \text{sgn}(1 - \beta) \left( \sum_{i=1}^{n} \left( \frac{x_i}{\sum_{j=1}^{n} x_j} \right)^{1-\beta} \right)^{\frac{1}{\beta}} \left( \sum_{i=1}^{n} x_i \right)^{\lambda} \]

- Unique family of functions: \( \beta \) and \( \lambda \) parameters
  - \( \beta \): type of fairness
  - \( \lambda \): importance of efficiency
Defining “Fairness”

• An equal allocation?
  • 1 job for each user

• But not efficient

• Ranking the fairness of different allocations

\[
\text{sgn}(1 - \beta) \left( \sum_{i=1}^{n} \left( \frac{x_i}{\sum_{j=1}^{n} x_j} \right)^{1 - \beta} \right)^{\frac{1}{\beta}}
\]
Defining “Efficiency”

• Maximize the total number of jobs?
  • 0 jobs to user 1
  • 3 jobs to user 2
• But not that fair
• Ranking the efficiency of different allocations

\[
\left( \sum_{i=1}^{n} x_i \right)^\lambda
\]
Heterogeneous Users

- Different users need different mixes of resources...
- Is it fair to treat them the same way?
Visualizing Heterogeneity

\[ \sigma_1 = \frac{R_{21}}{R_{11}} \]
\[ \sigma_2 = \frac{R_{22}}{R_{12}} \]
\[ \sigma_3 = \frac{R_{23}}{R_{13}} \]

\[ x_1 R_{11} + x_2 R_{12} + x_3 R_{13} = C_1 \]

3 Users, 2 Resources
Dominant Shares

- Dominant shares $\mu_j x_j$ for each user

$$\mu_j = \max_i \left\{ \frac{R_{ij}}{C_i} \right\}$$

- Maximum share of any resource

Calculating Dominant Shares

- 2 of 6 apples and 3 of 4 oranges: \( \mu_1 = \max \left( \frac{1}{3}, \frac{3}{4} \right) \)

- 2 of 6 apples and 1 of 4 oranges: \( \mu_2 = \max \left( \frac{1}{3}, \frac{1}{4} \right) \)
Fairness on Dominant Shares (FDS)

- Use dominant shares instead of number of jobs
- If $\mu$ is larger, equal dominant shares for smaller number of jobs

$$\text{sgn}(1 - \beta) \left( \sum_{j=1}^{n} \left( \frac{\mu_j x_j}{\sum_{k=1}^{n} \mu_k x_k} \right)^{1-\beta} \right)^{\frac{1}{\beta}} \left( \sum_{j=1}^{n} \mu_j x_j \right)^{\lambda}$$
GFJ

Generalized Fairness on Jobs

FDS

Fairness on Dominant Shares
Resource Scalarization

Matrix

Scalarization
GFJ or FDS

Vector

Scalar

User Scalars

$\begin{aligned}
\text{FDS} \\
\text{DRF} \\
\text{GFJ}
\end{aligned}$

$\begin{aligned}
\Gamma_{\theta,\lambda}^{\text{FDS}} \\
\Gamma_{\theta,\lambda}^{\text{GFJ}}
\end{aligned}$

User $n$

User 2

User 1

Resource Requirements

Resource Requirements

Resource Requirements
Desirable Properties Of Fairness Functions
Property 1: Pareto-Efficiency

- \( f(x) > f(y) \) whenever the allocation \( x \) Pareto-dominates \( y \).
  - \( x_i \geq y_i \) for all entries \( i \), with strict inequality for some \( i \)
- Not an axiom: needs to be proven
- Does not hold for all parameter combinations
Parameter Conditions

- Necessary and sufficient conditions

\[ |\lambda| \geq \left| \frac{1-\beta}{\beta} \right| \quad \beta > 0 \]

- Holds for FDS and GFJ
  - Comes from the same conditions for single-resource fairness

- If \( \lambda = \frac{1-\beta}{\beta} \) and \( \beta > 0 \), fairness becomes \( \alpha \)-fairness with \( \alpha = \beta \).

Property 2: Sharing Incentive

- Each user receives at least a $\frac{1}{n}$ share of some resource.
  - Dominant share is over $\frac{1}{n}$
- Users don’t want to share the resources equally.
- Does it hold?
Parameter Conditions

• Sufficient conditions:

\[ \lambda = \frac{1-\beta}{\beta} \quad \beta > 1 \]

FDS

• Counterexamples exist:

\[ \lambda = \frac{1-\beta}{\beta} \quad \text{FDS} \quad 0 < \beta < 1 \quad \text{GFJ} \quad \beta > 0 \]
Property 3: Envy-Freeness

• A user can process more jobs with his own rather than another user’s resource allocation.

  • Users don’t want to switch allocations.

• Does it hold?
Parameter Conditions

- Sufficient conditions:
  \[ \lambda = \frac{1 - \beta}{\beta} \quad \beta > 1 \]

- Counterexamples exist:
  \[ \lambda = \frac{1 - \beta}{\beta} \quad \text{FDS} \quad 0 < \beta < 1 \quad \text{GFJ} \quad \beta > 0 \]
## Sufficient Conditions

<table>
<thead>
<tr>
<th>Fairness</th>
<th>Pareto-Efficiency</th>
<th>Sharing Incentive</th>
<th>Envy-Freeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDS</td>
<td>$</td>
<td>\lambda</td>
<td>\geq \left</td>
</tr>
<tr>
<td>GFJ</td>
<td>$</td>
<td>\lambda</td>
<td>\geq \left</td>
</tr>
</tbody>
</table>
Existence of a Counterexample

<table>
<thead>
<tr>
<th>Fairness</th>
<th>Sharing Incentive</th>
<th>Envy-Freeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDS</td>
<td>( \lambda = \frac{1-\beta}{\beta}, 0 &lt; \beta &lt; 1 )</td>
<td>( \lambda = \infty, \text{ any } \beta )</td>
</tr>
<tr>
<td>GFJ</td>
<td>( \lambda = \frac{1-\beta}{\beta}, \beta &gt; 0 )</td>
<td>( \lambda = \infty \text{ or } 0, \text{ any } \beta )</td>
</tr>
</tbody>
</table>

What about Efficiency?
Fair, Efficient, or Both?

User 1

User 2

3

2.29

0.57

0.76

1.71
Existence of a Tradeoff

- Nonlinear, non-separable, multidimensional, continuous state-space knapsack problem
  - Maximize fairness function subject to multiple linear capacity constraints
  - Allow fractional jobs

---

No Tradeoff

Tradeoff Exists

<table>
<thead>
<tr>
<th>Resource 1</th>
<th>Resource 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Resource 1</td>
<td>Resource 2</td>
</tr>
<tr>
<td>2</td>
<td>24</td>
</tr>
</tbody>
</table>
Equal Allocations at Maximum Efficiency

• Number of tight resource constraints = number of users

\[ \sum_{j=1}^{n} \gamma_{ij} x_j \leq 1 \quad \forall \ i. \]

\[ \sum_{j=1}^{n} \frac{\gamma_{ij}}{\mu_j} = \rho \]

\[ \sum_{j=1}^{n} \gamma_{ij} = r \]
Efficiency Operating Range

Optimal allocations for a range of $\beta$ and $\lambda$

Dominant Resource Fairness (DRF): max-min fairness on dominant shares
Job Allocation

\[ \lambda = \frac{1 - \beta}{\beta} \]

Optimal allocations for \( \alpha = \beta \)-fairness
Numerical Example

- Fairness: DRF-fairness divided by maximal DRF value
- Efficiency: Total jobs divided by maximum number of jobs
Psychological Perceptions

Parameter Values

- What parameters are compatible with the responses?
  - Do they satisfy Pareto-efficiency, etc.?

- Do people agree with each other?

- Online survey asking people to rank datacenter allocations

<table>
<thead>
<tr>
<th>Allocation Options</th>
<th>Allocated to Client A</th>
<th>Allocated to Client B</th>
<th>Total no. of Jobs Completed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CPU</td>
<td>TB</td>
<td>No. of Jobs Completed for Client A</td>
</tr>
<tr>
<td>Allocation 1</td>
<td>24</td>
<td>96</td>
<td>24</td>
</tr>
</tbody>
</table>
Are People Very Different?

Average Preferred Fairness Value

Average Preferred Efficiency
Actually They’re Pretty Similar

- Cluster 1 prefers efficiency to fairness
- Cluster 2 prefers fairness to efficiency
All Responses
Compatible Parameters (GFJ)

The darker the square, the more participant rankings were compatible.

Lines represent Pareto-efficient frontiers.
Back to the Motivation
Questions Answered

• How do we define fairness?
  • GFJ and FDS

• Are these properties satisfied?
  • Pareto-efficiency
  • Envy-freeness
  • Sharing incentive

• Does a fairness-efficiency tradeoff exist?

• What parameters are consistent with actual preferences?
  • Users fall into 2 clusters
Thank you!

Questions?