Game-theoretic Paradoxes in Load Balancing for Distributed Computer Systems

H. Kameda, T. Kozawa, E. Altman

INRIA Sophia-Antipolis
2004 route des Lucioles, BP 93
06902 Sophia-Antipolis Cedex

E-mail: altman@sophia.inria.fr

USA, 1999

OUTLINE

I. Networking Games

II. Inefficiencies of equilibria

III. Distributed Computing

IV. Paradoxes

V. General model

VI. Uniqueness of Solutions

Networking games

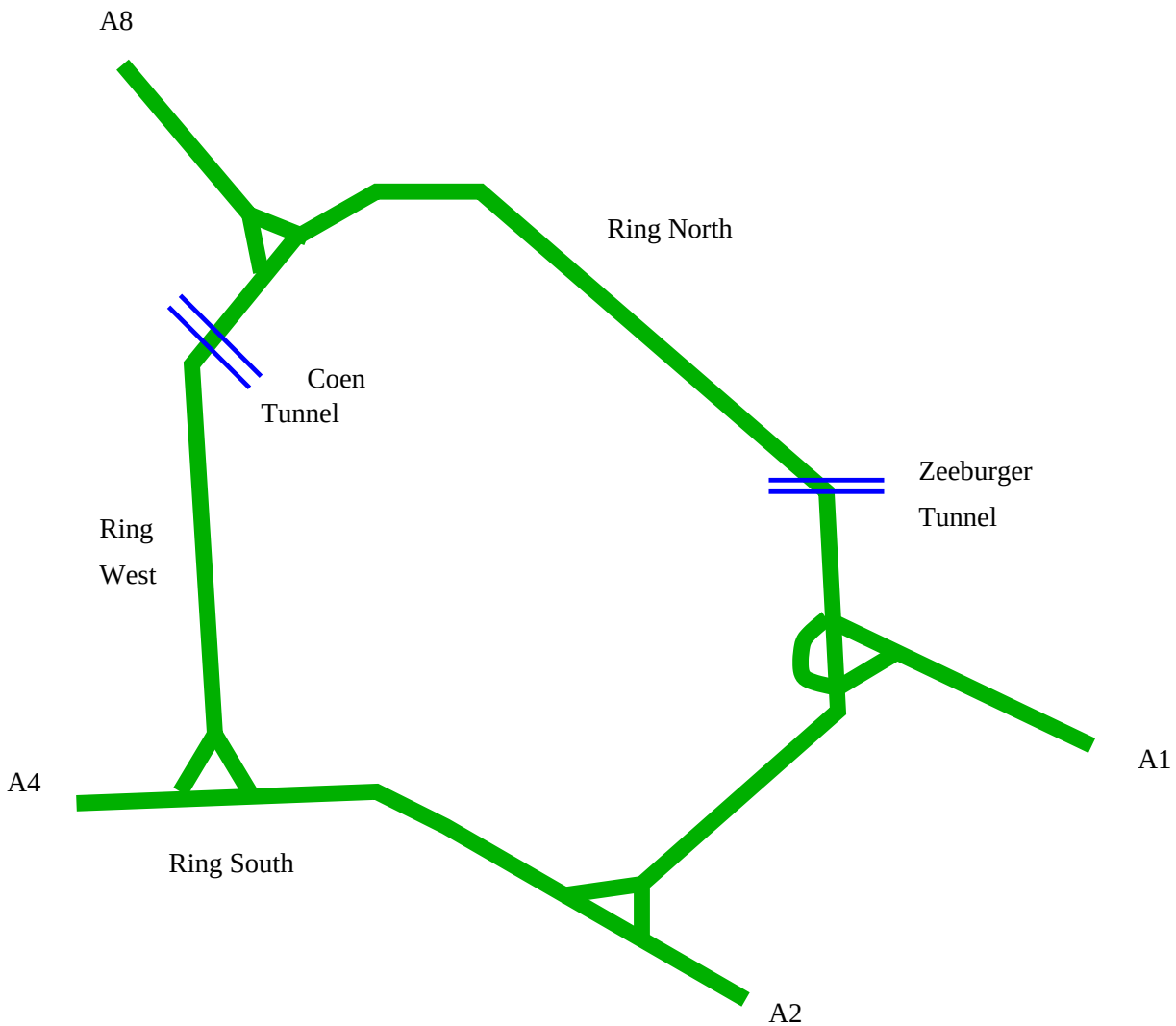
Road traffic

Individual choice of routes

- A large number of vehicles choose their route every day
- Each vehicle has a source and a destination.
- Typical objective of a vehicle: minimize delay
- Optimality concept: Wardrop Equilibrium

Example: The highway around Amsterdam

European project DRIVE II:
The vehicles have to decide the direction on the ring.
Objective: signalisation



Grouped choice of routes

- Example: transportation companies
- The influence of the choice of routs for a a large number of vehicles of a company has a nonnegligible influence on delays of other vehicles.
- Optimality concept: Nash Equilibrium.

Telecommunication networks

Individual choice of routes

- Ex: a large number of packets are transmitted over the network, and each packet makes its own routing choice.
- Ex: A large number of sessions share the network.
- A typical objective would be to minimise delays.

Grouped choice of routes

- Ex: A WEB application opens simultaneously several sessions (e.g. several images).

- Ex: Sessions cannot choses their routes; the choice is done by the service provider for all its subscribers.

Nash Equilibrium

- N joueurs
- Player i has X_i possible stragegies,
- $x = (x_1, ..., x_N)$ is a multi-strategy, $X \stackrel{\text{def}}{=} \Pi_{i=1}^{N} X_i$.
- $\psi_i(x)$:= The cost for player i when x is used
- Given a multi-strategy x, define (x^{-i}, y_i):= the multi-strategy obtained when player i switches to a strategy y_i.
- x is a Nash Equilibrium if for all i and $y \in X_i$,

$$\psi_i(x) \leq \psi_i((x^{[-i]}, y)).$$

The routing game

Network: a graph $G = (V, \mathcal{L})$ where

- V is a set of nodes
- L a class of directed links

Strategies of player i:

- player i sends a flow of Λ^i between the source $s(i)$ and the destination $d(i)$.
- The strategies are: $X_i = \{x_l^i\}, \quad x_l^i \in R_+$:
- Flow constraints:

$$r_v^i + \sum_{l \in In(v)} x_l^i = \sum_{l \in Out(v)} x_l^i,$$

$$\text{where} \qquad r_v^i = \begin{cases} \Lambda^i & \text{for } v = s(i), \\ -\Lambda^i & \text{for } v = d(i), \\ 0 & \text{otherwize} \end{cases}$$

Example of a network

- N players, all with the same source and destination.
- M links.
- $\lambda_l^{(k)}$:= the flow of player k over link l.
- The cost functions:

$$\psi_i(\lambda) = \sum_{m=1}^{M} \lambda_m^i f_m(\lambda_m)$$

where $f_m(\lambda_m)$ is the delay (or cost) of link m, and

$$\lambda_m = \sum_{i=1}^{N} \lambda_m^i$$

Wardrop Equilibrium

Definition: x is a Wardrop Equilibrium if

$$\sum_{l \in k^*} f_l(x) \leq \sum_{l \in k} f_l(x)$$

où $k^* \in P_i^*$ et $k \in P_i$,

– P_i are all the routes between $s(i)$ and $d(i)$,

– P_i^* are all the routes between $s(i)$ and $d(i)$ over which the flow is strictly positive.

Ref: J.G. Wardrop,
"Some theoretical aspects of road traffic research", *Proc. Inst. Civ. Eng.* Part 2, 1, 325–378, 1952.

- First Principle of Wardrop: The durations of a trip on any road actually used are equal, and are greater than the duration of a rtip of a single vehicle on a road which is not used.

- Second Principle of Wardrop:

The mean duration of the trip is minimal.

These principles have in fact appeared already in
A. C. Pigou, *The Economics of Welfare*, McMilan & Co., London, 1920.

Due to the fact that there are several individual decision makers, there are inefficiencies:

- The equilibrium is typically not socially optimal
- It has been observed that the increased capacity of a part of a system may lead to awkward behavior in terms of a system wide measure:

H. Kameda, J. Li, C. Kim and Y. Zhang, *Optimal Load Balancing in Distributed Computer Systems*, Springer, 1996.

- Braess paradox, in which increased capacity of a part of the system leads to the degradation in the benefits of *all* users

D. Braess, "Uber ein paradoxen der werkehrsplannung", *Unternehmenforschung*, **12**, 256-268, 1968.

J. E. Cohen and F. P. Kelly, "A Paradox of Congestion in a Queuing Network," *J. Appl. Prob. 27*, 730-734, 1990.

J. E. Cohen and C. Jeffries, "Congestion Resulting from Increased Capacity in Single-Server Queueing Networks," *IEEE/ACM Trans. on Networking 5*, 2, 1220-1225, 1997.

B. Calvert, W. Solomon and I. Ziedins, "Braess's Paradox in a Queueing Network with State-Dependent Routing," *J. Appl. Prob. 34*, 134-154, 1997.

Y. A. Korilis, A. A. Lazar and A. Orda, "Avoiding the Braess Paradox in Noncooperative Networks", *Proc. IEEE Conference on Decision & Control (San Diego)*, 864-878, 1997.

Distributed computing

- Two nodes: 1, 2.
- Each node: a single exponential server, param μ_i.
- Node i has the external Poisson arrival with rate ϕ_i
- x_{ii} are processed at node i,
- x_{ij} $(i \neq j)$ are forwarded through the communication means to the other node j. The results of those jobs are returned back through the communication means to node i.

$$x_{ii} + x_{ij} = \phi_i, \qquad (i \neq j),$$

$$x_{ij}, \geq 0, \quad i, j = 1, 2.$$

$\boldsymbol{x}$:=$(x_{11}, x_{12}, x_{21}, x_{22})$.

- $\mathbb{C}$:= the $\boldsymbol{x}$'s that satisfy the constraints,
- $\Phi = \phi_1 + \phi_2$.
- Goal: optimize x_{ij} $(i, j = 1, 2)$
- The load on node i is $\beta_i := x_{ii} + x_{ij}$ $(i \neq j)$
- The expected processing (including queueing) time of a job that is processed at node i, is

  $$1/(\mu_i - \beta_i)$$

  for $\beta_i < \mu_i$ (otherwize it is infinite).

Communication means:

Two alternatives.

(A) a single-channel communication line

$$G(\lambda) = \frac{t}{1 - (x_{12} + x_{21})t}$$

if $x_{12} + x_{21} < 1$, and is otherwize inifinite, where $\lambda = x_{12} + x_{21}$ is the network traffic.

The communication channel is modelled by a processor sharing server with service rate $1/t$; *i.e.*, the mean communication (without queueing) time is t, and, thus, the capacity of the communication line is t.

(B) Two unidirection communication lines 1 and 2.

The expected communication (with queueing) time of a job arriving at node i and being processed at node j ($\neq i$):

$$G_i(x_{ij}) = \frac{t}{1 - x_{ij}t}$$

if $x_{ij}t < 1$ (and is otherwize infinite).

For case (A), $(1/t) - x_{ij}$ $(i \neq j)$ is the amount of the line capacity left to be used by node i in forwarding the jobs arriving at node i.

The communication capacity for case (B) is greater than or equal to that for case (A).

The expected response time of a job that arrives at node i is

$$T_i(x) = \frac{1}{\phi_i} \sum_k x_{ik} T_{ik}(x),$$

where

$$T_{ii}(x) = \frac{1}{\mu_i - \beta_i}$$

if $\beta_i < \mu_i$ (and it is otherwize inifinite), and for $j \neq i$,

$$\begin{aligned} T_{ij}(x) &= \frac{1}{\mu_j - \beta_j} + \frac{t}{1 - (x_{ij} + x_{ji})t}, \qquad \text{for case (A)}, \\ &= \frac{1}{\mu_j - \beta_j} + \frac{t}{1 - x_{ij}t}, \qquad \text{for case (B)}. \end{aligned}$$

The overall expected response time of a job that arrives at the system is

$$T(x) = \frac{1}{\Phi} \sum_i \phi_i T_i.$$

The overall optimum: find $\bar{x}$ that achieves

$$T(\bar{x}) = \min T(x) \quad \text{s.t.} \quad x \in \mathbb{C}.$$

The node optimum (Nash E): find $\tilde{x}$ s.t. for all i

$$T_i(\tilde{x}) = \min T_i(x), \quad \text{s.t. } x \in \mathbb{C}.$$

The individual optimum (Wardrop E): find $\hat{x}$ s.t. for all i

$$T_i(\hat{x}) = \min_{i \neq j} \{T_{ii}(\hat{x}), T_{ij}(\hat{x})\} \quad \text{s.t.} \quad \hat{x} \in \mathbb{C}$$

Paradoxes

Mutual forwarding

There should be no mutual forwarding in overall and individual optima:

In overall and individual optima, either one of x_{ij} $(i \neq j)$ must be zero [Kameda et al].

Nash equilibrium: mutual forwarding paradox.

Monotonicity

The overall response time may be decreasing in the capacity, both in the individual and the node optimal cases (Fig. 3, Fig 4).

Braess Type Pardaox

Three Braess-like paradoxes (Fig 5-6):

- For case (A) with the decrease of t
- For case (B) with the decrease of t.
- By comparing cases (A) to case (B) where a single link has the same capacity in both cases.
- No paradox for the Wardrop E case (Fig 7).

INRIA

General model

- K sets of classes of jobs: $\mathcal{K}_i$ $i = 1, ..., K$ where $\mathcal{K}_i$ contains k_i classes of jobs.

$\mathcal{K}_i$ send jobs originally to node i.

$\mathcal{K} = \cup_{i=1}^{K} \mathcal{K}_i$.

- **Notation:**

$\phi^{(r)}$ input arrival rate of class r.

$x_{ii}^{(r)}$ flow originating at node i by class r that is processed at node i.

$x_{ij}^{(r)}$ the flow originating at node i by class r that is processed at node j.

$\mathbf{x} = (x_{ij}, i = 1, ..., K, j = 1, ..., K)$ set of all flows.

β_i total load on node i: $\beta_i = \Sigma_{j=1}^{K} \Sigma_{r \in \mathcal{K}_j} x_{ji}^{(r)}$.

λ_{ij} total flow forwarded from node i to node j.

λ the total network flow: $\lambda = \Sigma_{i=1}^{K} \Sigma_{j \neq i} \lambda_{ij}$.

Assumptions:

Case (A) – a single common bus.

Case (B) – 2 unidirectional links between any two processors.

System Elements:

Case A: $\mathcal{L}_A =$ set of nodes as well as the communication bus $\mathcal{L}_A = \{1, ..., K, c\}$.

Case B: set of nodes as well as the set of links $\mathcal{L}_B = \{1, ..., K, (ij), i = 1, ..., K, j \neq i\}$.

Costs:

$J_l^{(r)}(\mathbf{x})$:= Cost of element l for class r

Assumptions on the costs:

- **G1** $J^{(r)}$ is the sum of the local processing cost and the communication cost for class r where the latter are only functions of their local flow rate; i.e. in case (A) we have

$$J^{(r)}(\mathbf{x}) = \sum_{i=1}^{K} J_i^{(r)}(\mathbf{x}) + J_c^{(r)}(\mathbf{x})$$

and for case (B) we have for $r \in \mathcal{K}_i$

$$J^{(r)}(\mathbf{x}) = \sum_{j=1}^{K} \left(J_j^{(r)}(\mathbf{x}) + \sum_{i \neq j} J_{ji}^{(r)}(x_{ji}^{(r)}) \right).$$

- **G2** $J_l^{(r)}$ are continuous functions whose range is the nonnegative quadrant and their image is $[0, \infty]$.

- **G3** $J_l^{(r)}$ are convex functions in the rate sent by class r over the system element l. For example, if l is a processor and $r \in \mathcal{K}_l$ then $J_l^{(r)}$ is convex in $x_{ll}^{(r)}$.

- **G4** Whenever finite, $J_l^{(r)}$ is continuously differentiable in the flow sent by user r to system element l.

$K_l^{(r)}(\mathbf{x})$:= the partial derivative of $J_l^{(r)}(\mathbf{x})$ w.r.t. the flow sent by user r to system element l.

- **G5** If not all classes have finite cost and one of the classes has infinite cost then it can change its own flow to make this cost finite.

G5 ensures that any equilibrium has finite costs for all players.

Lemma: There exists an equilibrium.

Proof: Follows from Theorem 1 in Rosen (1965), see also Orda et al (1993).

INRIA

Uniqueness of Solutions

- Overall and Individual optimization: uniqueness is obtained by considering the equivalent routing game for which uniqueness holds for general topology.

[Ref:] H. Kameda and Y. Zhang, "Uniqueness of the Solution for Optimal Static Routing in Open BCMP Queueing Networks," *Mathematical and Computer Modelling 22*, 10-12, pp.119-130, 1995.

Nash Equilibrium

Assume

- (Π_1) $K_l^{(r)}$ is a function of two arguments:

(i) the total flow on the system element l,

(ii) the flow that class r sends to node element l. $K_l^{(r)}$ is strictly increasing in each of its two arguments.

Case (B)

Assume

(Π_2) : A class i that decides to ship some flow to a node $j \neq i$ should do it using a single hop (a single communication link); it is not possible for it to use two hops (i to k and then k to j), if there is a link that connect nodes i and j directly.

Equivalent routing model:

- elliminate the communicating costs between the processors;

this transforms the problem into a standard routing problem between parallel links.

- In order to compensate for the ij communicating cost that was elliminated we now add it to the cost of class i at processor j.

- Since for any i and $j \neq i$, only class i used the original communication link ij, the cost of class i in link j in the new problem depends only on the total rate at that link, as well as the rate at which jobs of class i are sent to link j.

This is an assumption of type (Π_1) .

- Hence the results from Orda et al (1993) are applicable.

Case (A), $K = 2$

For a fixed assignment of the other classes, class r is faced with a constrained minimization problem.

Its associated Lagrangian:

$$\Lambda^{(r)}(\mathbf{x}) = J^{(r)}(\mathbf{x}) - \alpha^{(r)}(\sum_{j=1}^{K} x_{ij}^{(r)} - \phi^{(r)}).$$

$\mathbf{x}^*$ is an equilibrium iff it satisfies the KT conditions:

$\exists\ \alpha^{(r)}, r \in \mathcal{K}$ s.t. for $r \in \mathcal{K}_i$, $i = 1, 2$, $j \neq i$:

$$K_i^{(r)}(\mathbf{x}^*) \geq \alpha^{(r)} \quad \text{and} \quad K_i^{(r)}(\mathbf{x}^*) = \alpha^{(r)} \text{ if } x_{ii}^{(r)} > 0, \tag{1}$$

$$\begin{aligned} K_c^{(r)}(\mathbf{x}^*) + K_j^{(r)}(\mathbf{x}^*) &\geq \alpha^{(r)} \\ K_c^{(r)}(\mathbf{x}^*) + K_j^{(r)}(\mathbf{x}^*) &= \alpha^{(r)} \text{ if } x_{ij}^{(r)} > 0; \end{aligned} \tag{2}$$

$$x_{ii}^{(r)}, x_{ij}^{(r)} \geq 0, \qquad \sum_{j=1}^{K} x_{ij}^{(r)} = \phi^{(r)}. \tag{3}$$

Let $\hat{\mathbf{x}}$ and $\mathbf{x}$ be two equilibria such that

$$\hat{\lambda} \geq \lambda. \tag{4}$$

Let i be such that $\hat{\beta}_i \geq \beta_i$. We shall show that for all $r \in \mathcal{K}_j$, $j \neq i$,

$$\hat{x}_{jj}^{(r)} \geq x_{jj}^{(r)}, \quad \text{or equivalently, } \hat{x}_{ji}^{(r)} \leq x_{ji}^{(r)} \tag{5}$$

(the equivalence follows from the constraint (3) on the sum of the flows.)

(5) holds trivially if $x_{jj}^{(r)} = 0$, so we have to check only the case $x_{jj}^{(r)} > 0$.

Fix $r \in \mathcal{K}_j$ and consider two subcases. Assume that

(a) $\hat{\alpha}^{(r)} \geq \alpha^{(r)}$. Note that $\hat{\beta}_i \geq \beta_i$ is equivalent to $\hat{\beta}_j \leq \beta_j$ (the equivalence follows from the constraint (3)). Hence

$$\begin{aligned} K_j^{(r)}(\hat{x}_{jj}^{(r)}, \hat{\beta}_j) &\geq \hat{\alpha}^{(r)} \geq \alpha^{(r)} \\ &= K_j^{(r)}(x_{jj}^{(r)}, \beta_j) \geq K_j^{(r)}(x_{jj}^{(r)}, \hat{\beta}_j). \end{aligned} \tag{6}$$

The equality as well as the first inequality follow from the KT conditions,

the last inequality follows from the monotonicity assumption (Π_1) .

Using the monotonicity assumption (Π_1) , for the first argument, we conclude from the fact $K_j^{(r)}(\hat{x}_{jj}^{(r)}, \hat{\beta}_j) \geq K_j^{(r)}(x_{jj}^{(r)}, \hat{\beta}_j)$ that (5) holds.

Thus we try instead of (a):

(b) $\hat{\alpha}^{(r)} \leq \alpha^{(r)}$. (5) holds trivially if $\hat{x}_{ji}^{(r)} = 0$. So it remains to check the case $\hat{x}_{ji}^{(r)} > 0$. We then have for $r \in \mathcal{K}_j$ $(j \neq i)$

$$\begin{aligned} & K_c^{(r)}(x_{ji}^{(r)}, \lambda) + K_i^{(r)}(x_{ji}^{(r)}, \beta_i) \\ & \geq \alpha^{(r)} \geq \hat{\alpha}^{(r)} = K_c^{(r)}(\hat{x}_{ji}^{(r)}, \hat{\lambda}) + K_i^{(r)}(\hat{x}_{ji}^{(r)}, \hat{\beta}_i) \\ & \geq K_c^{(r)}(\hat{x}_{ji}^{(r)}, \lambda) + K_i^{(r)}(\hat{x}_{ji}^{(r)}, \beta_i). \end{aligned}$$

The first inequality and the equality follow from the KT conditions,

the last inequality follows from the monotonicity of $K_l^{(r)}$ (property (Π_1)).

Using again the monotonicity, we conclude that (5) holds in case (b) as well.

We conclude that

$$\sum_{r \in \mathcal{K}_j} \hat{x}_{ji}^{(r)} \leq \sum_{r \in \mathcal{K}_j} x_{ji}^{(r)} \quad (j \neq i). \tag{7}$$

Combining this with $\hat{\beta}_i \geq \beta_i$, we conclude that

$$\sum_{r \in \mathcal{K}_i} \hat{x}_{ii}^{(r)} \geq \sum_{r \in \mathcal{K}_i} x_{ii}^{(r)}.$$

However, since for $i \in \mathcal{K}_i$, $\hat{x}_{ii}^{(r)} + \hat{x}_{ij}^{(r)} = \phi^{(r)}$, it follows that

$$\sum_{r \in \mathcal{K}_i} \hat{x}_{ij}^{(r)} \leq \sum_{r \in \mathcal{K}_i} x_{ij}^{(r)}. \tag{8}$$

Combining this with (7) we conclude that $\hat{\lambda} \leq \lambda$. This contradicts our assumption (4), unless we have equality in (4).

This implies in particular that (5) holds (as we derived above).

If for some $r \in \mathcal{K}_j$ (5) holds with strict inequality then (7) and (8) would hold with strict inequality, which would imply that $\hat{\lambda} < \lambda$.

But since we established that (4) holds with equality, we conclude that (5) holds with equality.

By a symmetric argument we establish (5) also for $r \in \mathcal{K}_i$. We conclude that $\mathbf{x} = \hat{\mathbf{x}}$.