Stability of the bipartite matching model

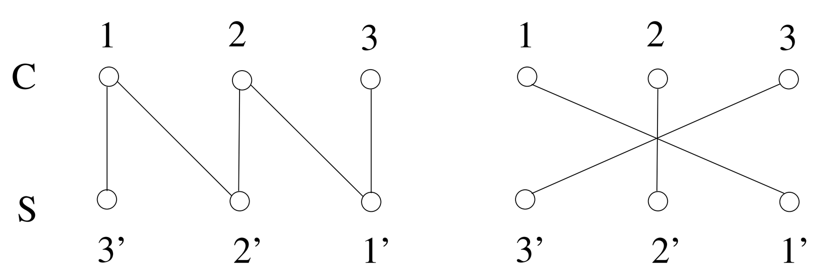
Ana BUŠIĆ (INRIA/ENS) ana.busic@ens.fr
Varun GUPTA (Carnegie Mellon University) varun@cs.cmu.edu
Jean MAIRESSE (CNRS/Université Paris 7) mairesse@liafa.jussieu.fr

The bipartite matching model

A multiclass queueing system with customers and servers playing symmetrical roles.

Def. A *bipartite matching structure* is a quadruple (C, S, E, F) where:

- C (resp. S) finite set of customer (resp. server) types;
- $E \subset C \times S$ is the set of possible matchings;
- $F \subset C \times S$ is the set of possible arrivals.



matching graph (C, S, E) The NN graph   arrival graph (C, S, F)

Evolution of the system

- Discrete time i.i.d. arrivals (of pairs customer/server) according to a joint probability measure μ on $F \subset C \times S$, independently of the past.
- Instantaneous matchings according to matching graph (C, S, E) and an *admissible* matching policy POL.
- Customers/servers that cannot be matched are stored in a buffer.

For a matching graph (C, S, E) we denote:

$$C(s) = \{c \in C \ : \ (c, s) \in E\}, \qquad S(c) = \{s \in S \ : \ (c, s) \in E\} .$$

A matching policy is *admissible* if:

- Only the current state of the buffer is taken into account;
- Buffer-first assumption: if the new arrival is $(c, s) \in E$, then c and s are matched together iff there are no servers from $S(c)$ and no customers from $C(s)$ in the buffer.

$\Longrightarrow$ Discrete time Markov chain on commutative (Match the Longest, Match the Shortest, Random, Priorities), or non-commutative state space (FIFO, LIFO).

Def. A *bipartite matching model* is a triple $[(C, S, E, F), \mu, \text{POL}]$, such that $\text{supp}(\mu) = F$ and the marginals of μ satisfy: $\text{supp}(\mu_C) = C$, $\text{supp}(\mu_S) = S$.

First introduced by Caldentey, Kaplan, and Weiss [1] (FIFO and $\mu = \mu_C \times \mu_S$).

Necessary conditions for stability

Def. The model is said to be *stable* if the Markov chain has a unique and attractive stationary probability measure (i.e. measure π such that $\pi P = \pi$ and for any initial measure ν, the sequence of Cesaro averages of νP^n converges weakly to π).

Prop. If the model is stable then the marginals of μ satisfy:

$$\text{NCOND}: \quad \begin{cases} \mu_C(U) < \mu_S(S(U)), & \forall U \subsetneq C \\ \mu_S(V) < \mu_C(C(V)), & \forall V \subsetneq S \end{cases}$$

Verifying NCOND

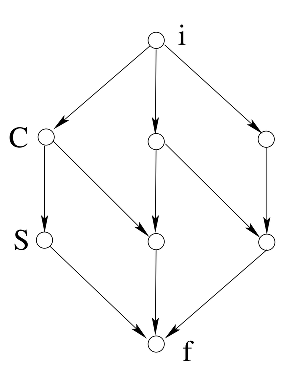
Prop. Given $[(C, S, E), \mu]$, there exists an algorithm of time complexity $O((|C| + |S|)^3)$ to decide if NCOND is satisfied.

Proof using network flow arguments:

$$\mathcal{N} = \big(C \cup S \cup \{i, f\}, E \cup \{(i, c), c \in C\} \cup \{(s, f), s \in S\}\big) .$$

Lemma.

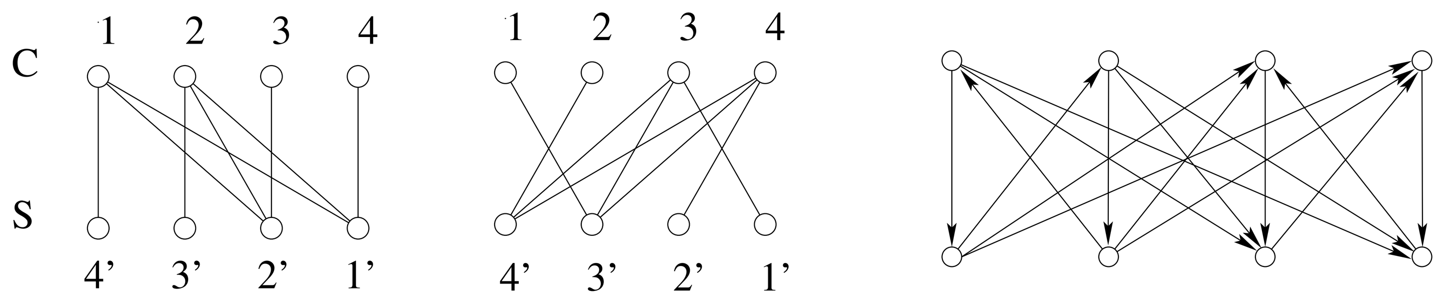
1. There exists a flow of value 1 in $\mathcal{N}$ iff (μ_C, μ_S) satisfies $\text{NCOND}_{\leq}$ ($<$ replaced by $\leq$ in NCOND).
2. There exists a flow T of value 1 such that $T(c, s) > 0$ for all $(c, s) \in E$ iff (μ_C, μ_S) satisfies NCOND.



Connectivity properties of the Markov chain

Consider a bipartite matching structure (C, S, E, F). Associated **directed** graph: the nodes are $C \cup S$ and the arcs are

$$c \longrightarrow s, \quad \text{if } (c, s) \in E, \qquad s \longrightarrow c, \quad \text{if } (c, s) \in F .$$



matching graph (C, S, E) arrival graph (C, S, F) associated directed graph

Thm. For a bipartite matching structure (C, S, E, F) the following properties are equivalent:

1. There exists μ such that $\text{supp}(\mu) = F$, $\text{supp}(\mu_C) = C$, $\text{supp}(\mu_S) = S$ and μ satisfies NCOND.
2. The associated directed graph is strongly connected.

Property for the transition graph of the Markov chain:

UTC : a unique (terminal) strictly connected component with all states leading to it.

Thm. If $(C \cup S, E \cup \widetilde{F})$ is strongly connected, then any bipartite matching model $[(C, S, E, F), \mu, \text{POL}]$ satisfies the property UTC.

Models that are stable for all admissible policies

The state space can be decomposed into facets, defined only by the non-zero classes.

Def. A *facet* is an ordered pair (U, V) such that: $U \subset C, V \subset S$ and $U \times V \subset (C \times S - E)$. The *zero-facet* is the facet $(\emptyset, \emptyset)$, we denote it shortly by $\emptyset$.

For a facet $\mathcal{F} = (U, V)$, define:

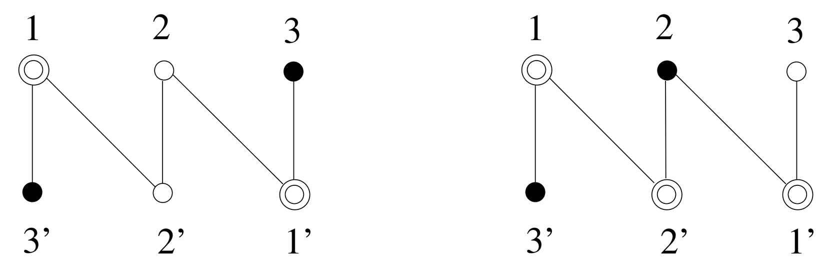
$$C_\bullet(\mathcal{F}) = U, \quad C_\circledcirc(\mathcal{F}) = C(V), \quad C_\circ(\mathcal{F}) = C - (C_\bullet(\mathcal{F}) \cup C_\circledcirc(\mathcal{F}))$$
$$S_\bullet(\mathcal{F}) = V, \quad S_\circledcirc(\mathcal{F}) = S(U), \quad S_\circ(\mathcal{F}) = S - (S_\bullet(\mathcal{F}) \cup S_\circledcirc(\mathcal{F})).$$

Denote by $\mathfrak{F}$ the set of facets. Define the following conditions on μ:

$$\text{SCOND}: \quad \mu_C(C_\circledcirc(\mathcal{F})) + \mu_S(S_\circledcirc(\mathcal{F})) \ > \ 1 - \mu(E \cap C_\circ(\mathcal{F}) \times S_\circ(\mathcal{F})), \quad \forall \mathcal{F} \in \mathfrak{F} - \{\emptyset\}$$

Def. A facet $\mathcal{F}$ is called saturated if $C_\circ(\mathcal{F}) = \emptyset$ or $S_\circ(\mathcal{F}) = \emptyset$.

SCOND $\Longrightarrow$ NCOND (considering only the saturated facets).



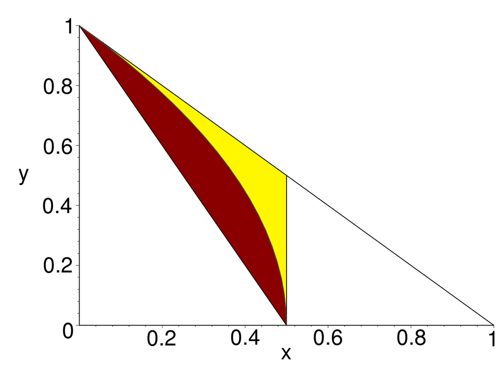
Prop. (Sufficient conditions) A bipartite model with probability μ satisfying SCOND is stable under any admissible matching policy.

Cor. Consider a bipartite graph in which any non-zero facet is saturated. For any admissible matching policy, the stability region is maximal.

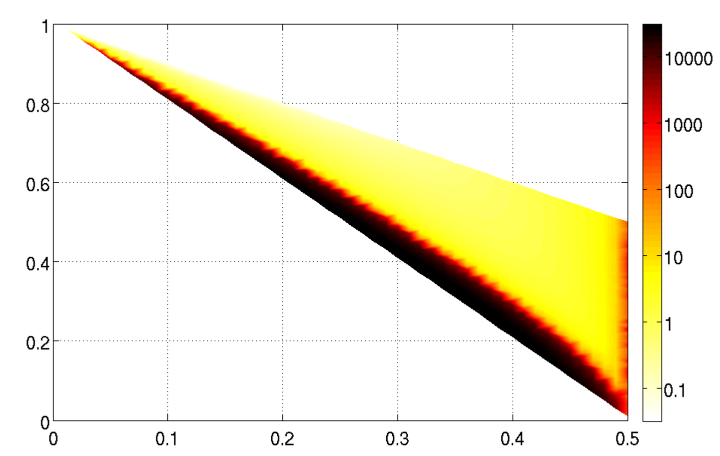
For the NN graph:
$\text{SCOND} = \{\text{NCOND} \cap (\mu_C(1) + \mu_S(1') > 1 - \mu(2, 2'))\}$.

For $\mu = \mu_C \times \mu_S$ and $\mu_C = \mu_S = (x, y, 1 - x - y)$:

$$\text{NCOND}: \ \begin{cases} x < 0.5 \\ 2x + y > 1 \end{cases} \qquad \text{SCOND}: \ \begin{cases} \text{NCOND} \\ 2x + y^2 > 1 \end{cases}$$

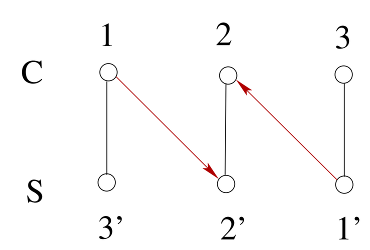


Priorities and Match the Shortest are not always stable



Simulation of the average buffer size up to time $n = 1000000$ for the NN-graph with $\mu = \mu_C \times \mu_S$, $\mu_C = \mu_S$, and MS policy.

Prop. NN model with either the MS policy or the PR (priority) policy such that customers of class 1 (resp. servers of class 1') give priority to servers of class 2' (resp. to customers of class 2):



For both policies, the stability region is not maximal.

Proof. Consider $\mu_C = (1/3, 2/5, 4/15)$, $\mu_S = \mu_C$, and $\mu = \mu_C \times \mu_S$. The conditions NCOND are satisfied, but the Markov chain is transient (for MS or PR defined as above).

Match the Longest is always stable

Thm. For any bipartite graph, the ML policy has a maximal stability region.

Open questions

- Is stability region is always maximal for the FIFO and RANDOM policies?
- For the MS and priority policies, how to compute the stability region?
- Better sufficient conditions for stability, valid for all admissible policies ?

Full paper: http://arxiv.org/abs/1003.3477

References

1. R. Caldentey, E.H. Kaplan, and G. Weiss. FCFS infinite bipartite matching of servers and customers. *Adv. Appl. Probab*, 41(3):695–730, 2009.