# SHORT TERM SCIENTIFIC MISSION (STSM) – SCIENTIFIC REPORT

**The STSM applicant submits this report for approval to the STSM coordinator**

**Action number: CA15210**
**STSM title: Stable Solutions of International Kidney Exchange Programs**
**STSM start and end date: 26/11/2017 to 08/12/2017**
**Grantee name: Aris Filos-Ratsikas**

## PURPOSE OF THE STSM/

(max.500 words)

The purpose of the STSM was to work on problems related to kidney exchange programs, specifically ones where the selfish/rational nature of the participants is taken into account. In particular, the aim was to appropriately model kidney exchange programs as cooperative games, where the participants are either different hospitals or even different countries (hence the term "International" in the STSM title).

The initial idea was motivated by the following scenario: Imagine a set of different countries (e.g. within the EU) participating in an international kidney exchange program that combines the donor-patient pools from all the participating hospitals in each country. Here, a "pool" is a set of patient-donor pairs, where the patient is incompatible with the donor (due to several possible reasons such as blood type incompatibility). Each such pair enrols into a hospital, and the hospital enlists the pair in a system involving multiple hospitals. All these pairs form a compatibility graph, where an edge between node u to node v signifies that the donor of the pair corresponding to node u is compatible with the patient of the pair corresponding to node v. Then the goal is to find cycles in this graph along which to facilitate the trade of kidneys, optimizing some objective – and while keeping the cycle length short for practical reasons.

Now imagine that in such a scenario, each country can willingly abstain from the program, electing to perform as many trades as possible internally, in some national exchange program. The reason for doing that is that perhaps in this manner, the country will have more patients receiving kidneys in the end, or the total cost for the country will be smaller (for example, having donors/patients fly from one country to another to be operated -since legally all operations in a cycle have to be performed simultaneously- can be quite costly).

A solution that guarantees that no country would like to leave the system and perform kidney exchanges locally is called *individually rational* and a very related problem was studied in [1] (see below) by Ioannis Caragiannis and co-authors. In cooperative game theory terms, such a solution is called an *imputation*.

The purpose is to see what happens if we take this a few steps further, and start looking for stability, i.e. a guarantee that no coalitions of countries will have incentives to leave the system and perform exchanges locally among their combined internal pool. Stability against coalitions of size 2 is the next step and stability against coalitions of any size is called *in the core*. Ultimately, the goal is to see if we can find solutions that are in the core.

[1]. Blum, Avrim, et al. "Opting into optimal matchings." *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*. Society for Industrial and Applied Mathematics, 2017.

**DESCRIPTION OF WORK CARRIED OUT DURING THE STSMS**

(max.500 words)

The first main task of the STSM visit was to properly formulate the problem and set the goals of the project clearly. This is a complicated process because the model often has to be modified, depending on how "interesting" the results that can be obtained are. After many fruitful discussions, we concluded that the model that the investigation should start from is the following.

There is a set of agents (e.g. the countries) and each one of those "represents" a set of nodes in a compatibility graph. For now, we will assume that the graph is undirected and an edge between two nodes signifies that the donor of the pair corresponding to the first node is compatible with the patient of the second and vice versa. The goal is to find a matching (exchanges of length 2) of maximum cardinality, which we will refer to as an *optimal matching* – such a matching is the one that makes sure as many patients as possible receive kidneys.

The *utility* or *payoff* of an agent is simply the number of matched patients that it represents and are included in the outputted matching. A matching is *individually rational* if no agent would have a higher utility by leaving the system and matching the nodes that it represents internally using an optimal matching for those. A matching is *stable (or in the core)* if no set of agents would have a higher combined utility (sum of matched patients over all the agents in the set) by leaving the system and matching the set of nodes that they represent internally, in a combined pool consisting only of those nodes. A matching is *stable against coalitions of size k* if it is stable when the size of the deviating coalitions is upper bounded by k.

The stable solutions mentioned above do not take into account *how* the solution looks like for the members of the deviating coalition. A weaker but quite meaningful requirement is to ask that no agent would want to be part of a coalition, if they can not improve their individual utility by doing that. A matching is *in a strong Nash equilibrium* if there is no coalition of agents and an optimal matching between their internal pool of nodes that makes each agent in that coalition better off. The definition for *strong Nash equilibrium against coalitions of size k* is similar.

During the visit, we worked mainly towards trying to answer the following questions: "Can we find the best solution which is at a strong Nash equilibrium?", "If we use an optimal matching, is it always in a strong pure Nash equilibrium? Is it in an *approximate* strong NE?", "Are there natural restrictions on the input graphs for which the optimal matching is in a strong NE?"

We also discussed possible extensions to the model where costs are associated with exchanges and the cost must also be shared among hospitals using some cost-sharing method.

**DESCRIPTION OF THE MAIN RESULTS OBTAINED**

(max. 500 words)

In [1], (see above), the authors proved that for any graph, if one constructs an optimal matching in a specific way, then the resulting matching is always individually rational (for pairwise exchanges, like in our model). The construction starts from an optimal internal matching for individual agents and then augments this matching maximally to obtain an optimal matching on the whole compatibility graph.

Our first result is that this is no longer possible, even if one is trying to obtain a matching which is stable against coalitions of size k>=2, as we come up with an instance where a deviating coalition can increase their total utility by matching nodes internally within their combined internal pool.

One observation that can be made based on the lower bound instance is that the increase in utility for each coalition is at most one node, i.e. one patient-pair. In other words, forming a coalition and matching the internal pool optimality would guarantee at most one more pair per coalition. This means that the example does not preclude the possibility of the optimal matching being a (strict) strong Nash equilibrium (but it does preclude the possibility of it being a (weak) strong NE, where at least one member of the deviating coalition is better off and all other are not worse off).

Then, we managed to extend this example to one where each member of the deviating coalition has one additional pair matched from the optimal matching within the combined internal pool of the coalition, which proves that even a (strict) strong Nash equilibrium is not always possible for the optimal matching.

However, we do believe that this result might be tight, in the sense that on any graph, the most incentive that any agent can have to be part of a deviating coalition can be at most one pair. If one imagines that each country might be responsible for hundreds of pairs, then it is conceivable that it wouldn't go through the effort of forming a different international program with other countries, just to ensure one more matched pair in the end. Formally, this is related to (additive) epsilon-equilibria or the notion of the epsilon-core in cooperative game theory.

The remainder of a visit was mostly dedicated to trying to prove this latter fact. We have ruled out several promising directions (e.g. starting from internal matchings and then augmenting in an "egalitarian" way, i.e. a way that makes sure the utility of the less satisfied agent is maximized) which unfortunately didn't work but we do have some intuition (related to some "locally" egalitarian solutions) about how to move forward in trying to prove this statement.

We also discussed the prospect of introducing "compensations" to the model which would eliminate the (small – assuming our intuition above is correct) incentives to deviate, but we haven't formulated this explicitly yet.

**FUTURE COLLABORATIONS (if applicable)**

(max.500 words)

We plan to continue working on this problem, as we think that it is very natural with clear real-life applications but also many interesting questions from a technical point of view and many possible extensions (costs and cost-sharing, compensations, graph restrictions or input distributions, experimental analysis e.t.c).

We are currently trying to organize a visit by Ioannis Caragiannis to the University of Oxford around March, when we will attempt to make further progress on the project. Meanwhile, we are collaborating on the project via e-mails and skype.

We are discussing the prospect of inviting further collaborators to join the project, but we will postpone the decision for after we have a more solid understanding of the problems mentioned above and presumably after we manage to prove (or disprove) our claims regarding the near-stable optimal matchings.