Random Graph Model of Multi-hospital Kidney Exchanges

Panos Toulis
ptoulis@fas.harvard.edu
Harvard University, Dept. of Statistics

Joint work with David C. Parkes (SEAS, Harvard)

CIREQ Matching Conference, Montreal, 2014

Assume m hospitals indexed by h, each with n patient/donor pairs embedded in a directed *compatibility graph* G_h. We seek to design a *kidney exchange mechanism* $\mathcal{M}$, that selects a set of *cycles* out of the combined set of *reported graphs* $\bigcup_h G'_h$.

$$\text{Find } \mathcal{M}, \text{ computing } \mathcal{M}(\bigcup_h G'_h) = \{cycles\}$$
$$\text{subject to, efficiency}(\{\text{cycles}\}) \text{ \& incentives}(\mathcal{M})$$

- Focusing on a single graph G_h, we derive a formula for the expected #matches in a maximum 2-cycle matching and show this is given by a formula $\gamma n - \beta\sqrt{n} - 2$.
- The "benefit from pooling" m hospitals of size n is thus shown to be $\propto (m - \sqrt{m})\sqrt{n}$.
- We design a mechanism that is efficient and EPIC with hospitals revealing all pairs ($G'_h = G_h$) for moderately-sized hospitals.
- We perform extensive simulation studies using the `xCM` and the `Bonus` mechanisms (Ashlagi & Roth, 2013), and demonstrate significant advantages in efficiency and incentives of `xCM`.
- We develop and provide our source code written in `R` through `Github`; experimental results are fully reproducible.

Single-hospital setting

Notation

- We write (P, D) to denote a pair where patient has blood-type P and donor has blood-type D e.g., (A, B), $(\mathrm{O}, \mathrm{AB})$ and so on. We assume only 4 blood-types, namely O, A, B, AB.
- We write $(\mathrm{P}_1, \mathrm{D}_1) \rightarrow (\mathrm{P}_2, \mathrm{D}_2)$ to denote that the donor of pair #1 can donate to the patient of pair #2.
- The patient-donor list of a single hospital h can be represented as a *compatilibity graph* $G(V, E)$ where
  - $V = \{(P_h, D_h)\}$ = patient-donor pairs set.
  - $E = (e_{ij})$ = compatilibity relationships among pair such that, $e_{ij} = 1 \Leftrightarrow (P_i, D_i) \rightarrow (P_j, D_j)$
- For the most part of this talk, we focus on *2-way exchanges* (also called matchings). We write, $(\mathrm{P}_1, \mathrm{D}_1) \leftrightarrow (\mathrm{P}_2, \mathrm{D}_2)$ to denote that the two pairs can *exchange* kidney transplants.
- We will extend to 3-way exchanges at the end of this talk.

Blood-types in pairs are random; so are the compatibilities (edges). Let $\widetilde{G_n}$ denote the random compatibility graph; we assume the following generative process:

- **Start** with an empty set.
- **Sample** one pair donor/patient and their blood types.[1]
- **Add** pair to the collection, if
  - Donor/patient are blood-type incompatible (deterministic test).
    OR
  - Donor/patient are blood-type compatible but tissue-type incompatible (random test with success probability $1 - p_c$). [2]
- **REPEAT** until n pairs have been collected.

[1] e.g. 50% O, 30% A, 15% B, 5% AB (tunable parameters)

[2] p_c = "cross-match" probability that patient rejects the transplant from a random donor (e.g. 20%); we also test on a "non-uniform" model where p_c describes the patient's *sensitivity*.

Single-hospital setting

Structural properties of $\widetilde{G_n}$

- Four different "types" of pairs in the compatibility graph (Ünver, 2010).
- *Over-demanded* pairs (OD) are more central in the graph, but less frequent ($\sim 10\%$).
- *Under-demanded* pairs (UD) are the most frequent ($> 50\%$); can only donate to OD pairs; hardest to match.
- *Self-demanded* pairs (S) form disconnected components. Internal 2-way matches within the components are possible.
- *Reciprocal* pairs (R) form a bipartite graph. Strategic issues in kidney exchanges are mainly due to such pairs, since a large number ($\propto \sqrt{n}$) remains (internally) unmatched.

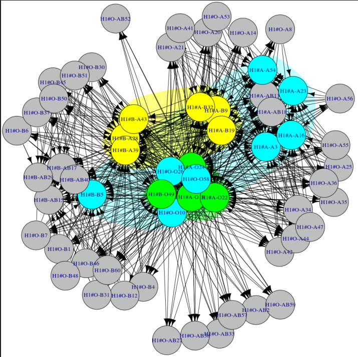


Figure : Sample of a compatilbility graph w/ 60 pairs.

Multi-hospital setting

Benefit from pooling; the "square-root law"

- A *regular matching* contains only (2-way) matches of the form (OD $\leftrightarrow$ UD), (S $\leftrightarrow$ S) and (R $\leftrightarrow$ R). When it exists, it is *maximum* (Roth et. al., 2004).
- We can show that a regular matching exists with high probability; then

$$\#\text{unmatched} = \underbrace{(\#\text{OD-UD})}_{\propto n} + \underbrace{(\#\text{surplus R})}_{\propto\sqrt{n}} + \underbrace{(\#\text{surplus S})}_{\propto O(1)}$$

- Thus, if $\mu(n) = \#$expected matches,

$$\mu(n) = \gamma n - \beta\sqrt{n} - 2$$

- **Square-root law**. Assume m hospitals of the same size n:

$$\text{Pooling benefit} = \mu(mn) - m\mu(n) \propto (m - \sqrt{m})\sqrt{n}$$

Multiple-hospital setting

Incentives: overview of assumptions

- To study incentives we assume an economy of m hospitals, with the **same** (moderate) patient/donor list size n.
- A hospital manipulates by **hiding pairs only**. Edges cannot be misreported (compatibility can be verified through medical tests).
- For our analysis, we assume that the combined graph of $(m-1)$ hospital graphs satisfy certain *perfect-matching* assumptions:
  - **R-perfect**: In balanced subgraph of an unbalanced R-subgraph, can be perfectly matched.
  - **S-perfect**: Any component in the S-subgraph (e.g. (A, A) pairs) can match all but one pairs.
- We also assume that every hospital is
  - **OD/UD-perfect**: All OD pairs can be matched with UD pairs.
- Analysis conditioned on aforementioned properties. We study EPIC, not DSIC.

Mechanism design: The xCM mechanism

2-way exchanges

The xCM mechanism works as follows:

Step 1. (Match S pairs) Match pooled S-pairs internally such that each hospital *h* matches *at least* as many as it can match internally.

Step 2. (Match R pairs) Match the *short side* of the pooled R-subgraph to the long-side under the *probabilistic uniform rule* (Ehlers & Klaus, 2003).[3] Each hospital matches *at least* as many as it can match internally.

Step 3. (Match remaining pairs locally) Enforce an *almost-regular* matching internally for each hospital, with all pairs that remain.

Step 4. (Match remaining pairs globally) Compute a random, almost-regular matching on the combined graph formed from all remaining unmatched pairs in the pool.

[3] The rule allocates +1 to all agents as long as the supply is larger than the #agents "in-demand"; it then allocates the remainder uniformly at-random.

Mechanism design: The xCM mechanism

Incentives and efficiency

Theorem 1. *The xCM mechanism is EPIC and 2-way efficient*[4] *for properties (i) S-perfect and R-perfect on compatibility graphs the size of every marginal economy and larger, and (ii) OD/UD-perfect on every hospital's compatibility graph.*

- Efficiency follows because xCM computes an overall matching that is regular, and thus maximum on the combined S- and R-subgraphs, and matches every OD pair to a UD pair.
- The OD/UD-perfect property holds for virtually any graph at 2% error (i.e., 2% OD pairs remain unmatched on average in regular matchings).
- The R-,S-perfect assumptions for the marginal economies hold at 2% for $m = 4$ hospitals, at $n \geq 25$.

[4] i.e., efficient allowing only 2-way matches

Mechanism design: xCM and Bonus

The Bonus mechanism (Ashlagi & Roth, 2013) focuses on the OD-UD subgraph and employs a lottery to allocate the "OD supply".

operation	xCM	Bonus
match S pairs	maximum matching [5] under IR constraints	maximum matching
match R pairs	uniform probabilistic rule under IR constraints	maximum matching under IR constraints
match OD/UD pairs	almost-regular matching	OD/UD lottery

[5] Both mechanisms make extensive use of *uniformly-random* maximum matchings.

GAME OF NODES

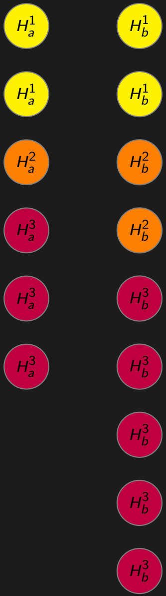
Game of Nodes

Matching on R subgraph (bipartite case)

- Assume 3 hospitals (different colors in figure) sharing only R pairs. Their goal is to maximize individual #matches.
- Assume *any* pair from one side can be matched to any other pair from the other side (stronger assumption than "R-perfect").
- If we just pick a random maximum then, by symmetry, H_3 will match

$$\underbrace{3}_{\text{all short-side pairs}} + \underbrace{6}_{\text{short-side supply}} \times \underbrace{5/(5+4)}_{\text{relative proportion in long-side}}$$

- By perfect-matching, a strategy can be represented by (x, y) where x, y are the #pairs reported in each side.



Game of Nodes

Matching on R subgraph: Numerical example

- Assume three hospitals with reports, $H_1(25, 10)$, $H_2(20, 50)$ and $H_3(15, 30)$
  - each node in figure = 5 pairs.
- Consider 3 mechanisms:
  - random max matching (`rCM`)
  - match internally → random max matching (IR+`rCM`)
  - uniform probabilistic rule (`uniform`)
- Also consider 3 strategies
  - truthful : share all nodes from each side
  - canonical : match internally, report remainder
  - "long-R": report the entire long side only

Game of Nodes

Matching on R subgraph: Numerical example

- $H_1(25, 10)$, $H_2(20, 50)$ and $H_3(15, 30)$
- Assume H_1, H_2 are truthful and H_3 is being strategic. Then the expected utilities for H_3 are given by the table below

strategy	mechanism		
	rCM	IR+rCM	uniform
truthful	35	35	37.5
canonical	39	35	37.5
long-R	45	37.5	37.5

Game of Nodes

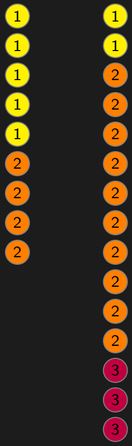
Matching on R subgraph: Numerical example

- $H_1(25, 10)$, $H_2(20, 50)$ and $H_3(15, 30)$
- Assume H_1, H_2 are truthful and H_3 is being strategic. Then the expected utilities are given by the table below

strategy	mechanism		
	`rCM`	IR+`rCM`	`uniform`
truthful	35	35	37.5
canonical	39	35	37.5
long-R	45	37.5	37.5

- For example, (`rCM` + canonical):

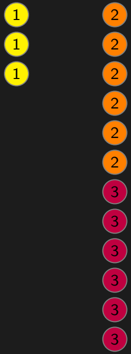
$$\underbrace{2 \cdot 15}_{\text{internal matches}} + \underbrace{\frac{15}{10 + 50 + 15}}_{\text{relative prop. in long-side}} \times \underbrace{(20 + 25)}_{\text{short side}} = 39$$



Game of Nodes

Matching on R subgraph: Numerical example

- $H_1(25, 10)$, $H_2(20, 50)$ and $H_3(15, 30)$
- Assume H_1, H_2 are truthful and H_3 is being strategic. Then the expected utilities are given by the table below



strategy	mechanism		
	rCM	IR+rCM	uniform
truthful	35	35	37.5
canonical	39	35	37.5
long-R	45	37.5	37.5

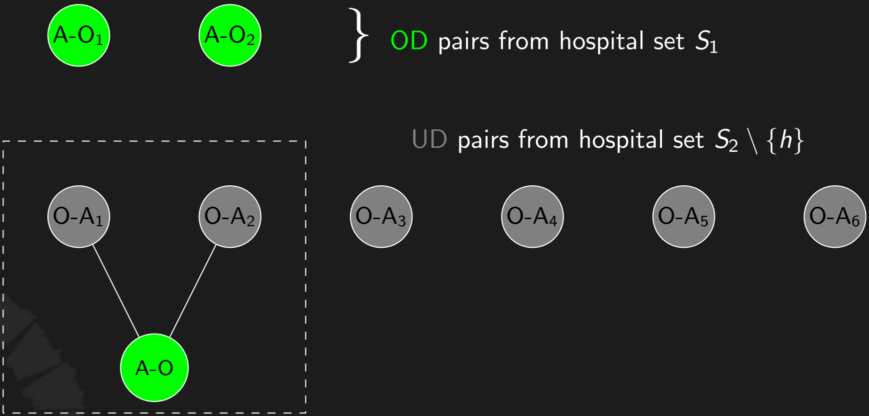
- For example, (IR+rCM and "long-R") equiv. to $H_1(15, 0)$, $H_2(0, 30)$ and $H_3(0, 30)$:

$$\underbrace{\frac{30}{30+30}}_{\text{relative prop. in long-side}} \times \underbrace{15}_{\text{short side}} + 2\underbrace{\min\{15, 30 - \text{matched}\}}_{\text{internal}} = 37.5$$

Game of Nodes

Matching on OD/UD subgraph: Numerical example

- Bonus splits hospitals in two groups, S_1 and S_2. Hospital h (in S_2) has 1 OD and 2 UD pairs. Same-side hospitals have 4 UD pairs. The 6 UD pairs in S_2 are to be matched with 2 OD pairs from S_1.
- Goal: Allocate the 2 OD pairs to the 6 UD pairs.



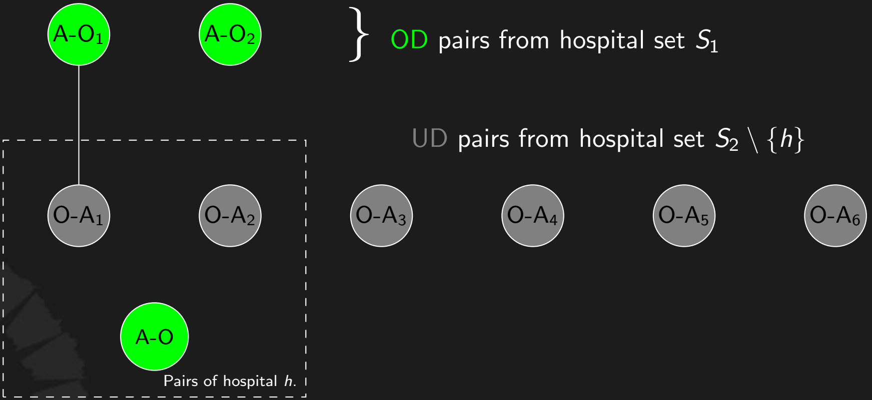
Pairs of hospital h.

Game of Nodes

Matching on OD/UD subgraph: Numerical example

- Case 1 (h truthful). Bonus preallocates a match. Probability of matching the other is **2/6** (not 1/5), so overall utility

$$\underbrace{1}_{\text{OD ``easy-to-match''}} + \underbrace{1}_{\text{UD preallocated}} + \underbrace{2/6}_{\text{lottery}} = 35/15.$$

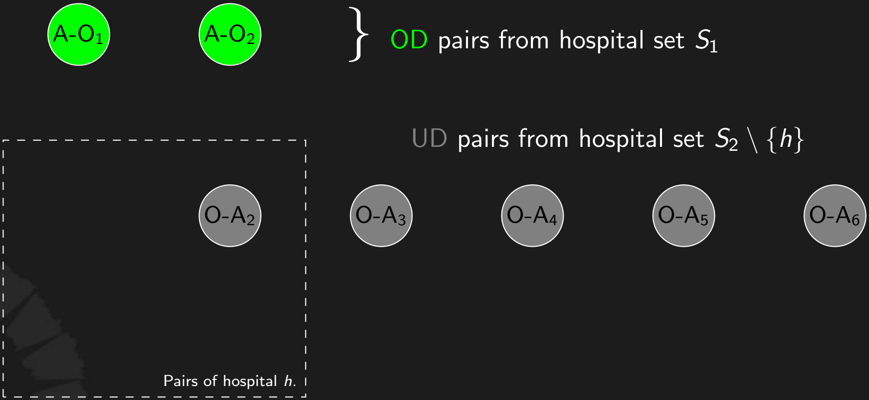


Game of Nodes

Matching on OD/UD subgraph: Numerical example

- Case 2 (h deviates). Probability of matching the other is 2/5, so overall utility is

$$\underbrace{2}_{\text{internal}} + \underbrace{2/5}_{\text{lottery}} = 36/15 > \underbrace{35/15}_{\text{truthful}}.$$



Extension to 3-way exchanges: 3-xCM

- One key idea is to define virtual-R pairs as
  $(A, O) \rightarrow (O, B) \equiv$ "virtual A-B" pair
  $(B, O) \rightarrow (O, A) \equiv$ "virtual B-A" pair
- The virtual-R pairs are important in clearing out the imbalance in the R subgraph and thus in achieving full efficiency (asymptotically) in 3-way exchanges.
- Another 3-cycle of interest from a welfare perspective involves 1 (OD) and 2 UD pairs. Such cycles are explicitly explored by 3-xCM but they are, generally, less frequent in practice.

Experiments

Incentives: 2-way exchanges

mech.	profile	strategy	avg.utility	#OD	#R	#S	#UD
rCM	tttttt	t	5.90 (0.03)	1629	2088	1778	1584
	tttttc	c	6.68 (0.06)	1604	2267	1802	2347
	tccccc	t	4.54 (0.05)	1492	1656	1476	371
	cccccc	c	5.57 (0.03)	1462	1878	1580	1201
xCM	tttttt	t	5.85 (0.03)	1618	2125	1811	1458
	tttttc	c	5.81 (0.06)	1600	2137	1765	1468
	tccccc	t	5.59 (0.07)	1485	1840	1582	1246
	cccccc	c	5.57 (0.03)	1457	1879	1583	1201
Bonus	tttttt	t	5.67 (0.03)	1537	2081	1778	1410
	tttttc	c	6.19 (0.06)	1571	1994	1784	2075
	tccccc	t	4.75 (0.06)	1405	1889	1511	424
	cccccc	c	5.50 (0.03)	1428	1870	1574	1179

Table : 2-cycles, with $m = 6$, and $n = 12$.

- Random maximum-matching (rCM) is not BNIC. Canonical deviation (hiding all possible internal matches) yields 6.68 matches vs. 5.9 when truthful.
- The canonical deviation is not useful in xCM. This is consistent with the EPIC theoretical property.
- The Bonus mechanism is not BNIC. Canonical deviation yields 6.19 vs 5.67 when truthful. The lottery seems to hurt matches of #R and UD pairs, as theoretically expected.

Experiments

Incentives: summary

mechanism	(m, n) =(#hospitals, #size each.) (4, 18)	(6, 12)	(12, 6)
rCM	1.148 (0.007)	1.133 (0.008)	1.141 (0.014)
3-rCM	1.124 (0.008)	1.113 (0.011)	1.090 (0.013)
xCM	0.995 (0.008)	0.994 (0.009)	1.021 (0.015)
3-xCM	1.022 (0.010)	1.018 (0.013)	0.997 (0.015)
Bonus	1.116 (0.008)	1.091 (0.009)	1.091 (0.015)
3-Bonus	1.158 (0.011)	1.131 (0.014)	1.058 (0.016)

Table : The average ratio of the utility to a hospital from deviating to the canonical strategy compared to the utility for truthful reporting.

- Neither `rCM` nor `Bonus` are BNIC for (m, n)-combination (hospital, size).
- The canonical deviation is not useful in `xCM` but may be marginally useful in 3-`xCM`. The incentives of `xCM`-* mechanisms are better even in environments where the perfect-matching properties do not hold precisely.

Experiments

Welfare: 3-way exchanges

mechanism	welfare	OO?	O[RS]	ORU	OSU	OUU	RRS	SSS
no pooling	29.73 (0.24)	153	572	1205	1479	81	573	369
max matching	39.51 (0.23)	1	50	1694	93	428	1686	1073
3-rCM-all-c	35.40 (0.20)	164	480	1203	1623	83	1055	505
		0	3.54	2.74	0.92	0	46.16	27.52
3-xCM-all-t	37.56 (0.22)	299	20	2408	200	170	0	1263
		100	100	100	100	100	-	100
3-Bonus-all-c	34.62 (0.20)	175	478	1167	1586	82	568	515
		1.14	0	0	0	0	0	28.93

- There is significant benefit from pooling (compare "no-pooling" with "max-matching").
- Main inefficiency of 3-xCM relative to max-matching, because of fewer UD and S pairs matched.
- Inefficiency of Bonus due to insufficient use of R pairs (for example, compare ORU matches in 3-xCM and Bonus).

Source code

- Source code and experiments fully available online: https://github.com/ptoulis/kidney-exchange
- IP solver powered by commercial Gurobi which is available for free with an academic license[6]
- Written in R, easy-to-use/extend, statistical tools readily available

```
> pool = rrke.pool(m=4, n=60, uniform.pra=T)
> kpd = kpd.create(pool, strategy.str="ttc", include.3way=T)
> out = Run.Mechanism(kpd, "xCM", include.3way=T)
> get.matching.utility(out)
[1] 58
> out$information
 info2.OO  info2.OR  info2.OS  info2.OU  info2.RR  info2.RS  info2.RU  info2.SS
        0         0         0         9         3         0         0         5
 info2.SU  info2.UU info3.OOO info3.OOR info3.OOS info3.OOU info3.ORR info3.ORS
        0         0         0         0         0         0         0         0
info3.ORU info3.OSS info3.OSU info3.OUU info3.RRR info3.RRS info3.RRU info3.RSS
        3         0         0         3         0         0         0         0
info3.RSU info3.RUU info3.SSS info3.SSU info3.SUU info3.UUU     n2way     n3way
        0         0         2         0         0         0        17         8
```

[6] http://www.gurobi.com/products/licensing-and-pricing/academic-licensing

Conclusions

- We apply random graph theory to quantify the statistical properties of kidney exchange graphs; pooling benefits sublinear to #hospitals ($m - \sqrt{m}$) and proportional to square-root of hospital size ($\propto \sqrt{n}$).
- We design a mechanism, namely `xCM`, to address incentives issues in multi-hospital kidney exchanges (e.g. hospital hiding pairs) in a static context.
- Our mechanism is efficient and EPIC under perfect-matching assumptions that are validated for moderately-sized hospitals (~ 30 pairs/hospital) and a "uniform-PRA" model (uniform crossmatch probability), and several blood-type distributions.
- In particular, our mechanism fares better compared to the `Bonus` mechanism (Ashlagi & Roth, 2013), which is shown to be vulnerable to deviations.
- We publicly release a code-base which can be used for reproducibility and further research on the domain.

THANK YOU