The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Analysis of Incentive Ratio in Top-Trading-Cycles Algorithms

Taiki Todo[*1*2]

[*1]Department of Informatics, Graduate School of ISEE, Kyushu University
[*2]Multi-Agent Optimization Team, RIKEN AIP

The main objective of this paper is to analyze some variants of the classical top-trading-cycles (TTC) algorithm for slightly modified models of the housing market. Extensions of TTC for such modified models are not necessarily strategy-proof, as pointed out by Fujita et al. (2015), and thus some alternative analysis of agents' selfish behavior is needed. In this paper, the incentive ratio, originally proposed by Chen et al. (2011), of the variants of TTC algorithm is analyzed in both (i) the multi-item exchange and (ii) an exchange model with a specific form of externalities.

1. Introduction

Exchange of indivisible items is a fundamental problem in the literature of economic theory, where each agent is endowed with a set of indivisible item and a preference over the items in the market, and monetary transfer is not allowed. The objective is the find an exchange rule that returns a socially-desirable redistribution (outcome) of items among agents. Most researches on exchange, especially those on the mechanism design perspective, have investigated agents' incentives, such as *strategy-proofness* that requires no agent can benefit by misreporting its preference to the mechanism.

The *housing market* [Shapley 74], is one of the well-studied model of exchange, where each agent's endowments are restricted to a single item. The *top-trading-cycles* (TTC) algorithm is a well-known exchange rule for the housing market, which satisfies strategy-proofness, and returns an efficient outcome in polynomial time. Furthermore, Ma showed that TTC is the only exchange rule that satisfies, under a natural condition, those two properties [Ma 94]. Recent years, there have been several extensions of TTC for various modified exchange problems, in the fields of economics and artificial intelligence [Pápai 00, Alcalde-Unzu 11, Aziz 12, Saban 13, Sonoda 14, Sun 15, Sikdar 17].

Strategy-proofness is, in general, a too demanding property, and is not compatible with optimal outcomes, e.g., Pareto efficient ones, in many realistic extensions, including exchange of multiple items [Sönmez 99]. Therefore, many researches in the literature of mechanism design have focused on developing sub-optimal mechanisms, in order to guarantee strategy-proofness [Todo 14].

However, it is still important to analyze, under mechanisms that are not strategy-proof, to what extent an agent can benefit by his selfish behavior. In particular, mechanism that are popular or easy-to-understand are, even if they are not strategy-proof, more likely to be used in practice, as the first-price auction is for selling items. One of such analysis is based on *incentive ratio* [Chen 11, Chen 17a, Chen 17b], which quantitatively evaluates the (ratio of) possible gain by a manipulation.

In this paper, we first algorithmically analyze the incentive ratio of some TTC variants, which indicates how much in the worst case an agent can gain by misreporting its preference under TTC. For the case of multi-item exchange, we show that (a) the incentive ratio is unbounded in general, and (b) it becomes 2 when the valuation is assumed to be lexicographic. For the case of service exchange, which is a special case of exchange with externalities, we show that (c) the incentive ratio is unbounded, even if the valuations have an upper bound.

Contact: Taiki Todo, Department of Informatics, Kyushu University, Motooka 744, Nishi-Ward, Fukuoka, Japan 819-0395, +81-92-802-3576, todo@inf.kyushu-u.ac.jp

2. General Model

In this section we introduce a general model of exchange. Let N be the set of n agents, and K be the set of indivisible items. Each agent $i \in N$ has an *endowment* $e_i \subseteq K$, satisfying $\bigcup_{i \in N} e_i = K$, $e_i \neq \emptyset$ for any $i \in N$, and $e_i \cap e_j = \emptyset$ for any pair $i, j \in N$. The profile $e := (e_i)_{i \in N}$ is called an *endowment profile*. An n-partition $(a_i)_{i \in N}$ of the set K, satisfying $\bigcup_{i \in N} a_i = K$, $a_i \neq \emptyset$ for any $i \in N$, and and $a_i \cap a_j = \emptyset$ for any pair $i, j \in N$, is called an *outcome*. Let A be the set of all outcomes. By definition, the endowment profile e is also an outcome, i.e., $e \in A$.

Each agent $i \in N$ also has a *valuation function* $v_i : A \to \mathbb{R}_{>0}$, which assigns a non-negative value for each outcome. The value $v_i(a)$ indicates the level of happiness for agent i when outcome a realizes. Let V be the set of all possible valuation functions. An *exchange rule* $f : A \times V \to A$ is a function that takes an endowment profile and a profile of valuation functions as an input and returns an outcome[*1].

In the literature of mechanism design, an incentive property called *strategy-proofness* has been extensively studied. An exchange rule is said to be strategy-proof if no agent can benefit by misreporting his valuation function, i.e., truth-telling is a dominant strategy of the game. Formally, it requires that $\forall N$, $\forall K$, $\forall e \in A$, $\forall i \in N$, $\forall v_{-i} \in V^{n-1}$,

*1 While we restrict our attention in this paper to deterministic exchange rules, most of the concepts defined in this paper can easily apply for *randomized exchange rules*, which returns a probability distribution over outcomes.

$\forall v_i \in V$, and $\forall v_i' \in V$,

$$v_i(f(e,(v_i,v_{-i}))) \geq v_i(f(e,(v_i',v_{-i}))).$$

As readers may wonder, however, the condition of strategy-proofness is quite demanding and actually very hard to satisfy; the inequality must hold for any profile of other agents' valuations and any misreport of the manipulator. Indeed, under very natural assumptions, strategy-proofness, and its variants such as false-name-proofness, are not achievable in various mechanism design problems [Gibbard 73, Satterthwaite 75, Yokoo 04, Todo 14].

Accordingly, Chen et al. [Chen 11] proposed an alternative measure, called *incentive ratio*, to evaluate the robustness of mechanisms/rules against agents' selfish behaviors.

Definition 1 (Incentive Ratio)**.** *For a given exchange rule f, the* incentive ratio in f *is* $\alpha \in \mathbb{R}_{\geq 1}$ *if* α *is the minimum real number such that* $\forall N$, $\forall K$, $\forall e \in A$, $\forall i \in N$, $\forall v_{-i}$, $\forall v_i$, $\forall v_i'$,

$$\alpha \cdot v_i(f(e,(v_i,v_{-i}))) \geq v_i(f(e,(v_i',v_{-i}))).$$

Obviously, when an exchange rule f is strategy-proof, the incentive ratio of f is one. The smaller, i.e., closer to one, the incentive ratio is, the more robust it is against manipulations (in our framework, against valuation misreports).

3. Housing Market and TTC

The classical *housing market*, originally proposed by Shapley and Scarf, is represented as a special case of our exchange problem[*2], by setting $|K| = n$ (and thus, automatically $|e_i| = |a_i| = 1$ for any $i \in N$ and $a \in A$) and $v_i(a) = v_i(b)$ if and only if $a_i = b_i$ for any $i \in N$ and any pair $a, b \in A$. Each outcome therefore corresponds to a permutation of endowments among agents. The following algorithm is called *Top-Trading-Cycles* (TTC in short), which is proposed to solve the housing market problem [Shapley 74].

Definition 2 (Top-Trading-Cycles)**.**

Step 1 *Construct a DAG with two types of vertices, agents and items, so that draw a directed edge from each agent vertex to his favorite item vertex, and a directed edge from each item vertex to its owner (agent) vertex. Assign to each agent, in each cycle, the item to which he is pointing and remove all such item vertices and agent vertices from the graph. Go to* **Step** 2*.*

Step $t(\geq 2)$ *The algorithm halts if no agent vertex remains; otherwise, each agent in the graph points to his favorite item among the remaining ones and each item points to its owner. Assign to each agent, in each cycle, the item to which he points and remove all such items and agents from the graph. Go to* **Step** $t+1$*.*

It was proven that the TTC algorithm is strategy-proof, and therefore, the incentive ratio of TTC in the housing market is 1.

*2 The original housing market is defined only with ordinal preferences, rather than with valuation functions. However, all the discussion on incentives in their paper can easily apply for the case with valuation functions.

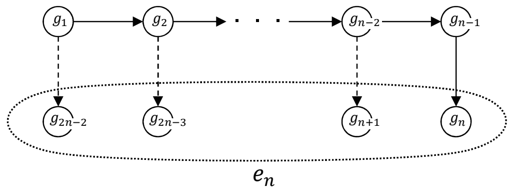

Figure 1: Example in Proof of Lemma 1

4. Multi-Item Exchange

Fujita et al. [Fujita 15] extended the classical housing market to *multi-item exchange* problem, in which agents' endowments are not restricted to a single item, i.e., remove the constraints of $|K| = n$ from the housing market.

They proposed a modification of TTC, so-called augmented TTC (ATTC), that still runs in polynomial times and always selects a core outcome. ATTC first splits each agent i into atomic players, so that each atomic player has the same valuation function with agent i only over single items, and owns exactly one item from e_i. ATTC then run the TTC algorithm for the market consisting of the atomic players. By definition, each agent finally obtains the same number of items as his original endowment, i.e., $|a_i| = |e_i|$ for any $a \in A$ and any $i \in N$. To simplify the notation, let us focus on the following *additive* valuation functions: there exists $u_i : K \to \mathbb{R}_{>0}$ such that $v_i(a_i) = \sum_{g \in a_i} u_i(g)$.

As they pointed out, ATTC is not strategy-proof. On the other hand, they also pointed out by Proposition 2 in [Fujita 18], even though ATTC is not strategy-proof, the best item that an agent receives under truth-telling cannot be improved by any misreport of valuation function.

Here, we provide a further observation on the manipulability of ATTC: there exists a problem instance under which an agent can improve all the other items, except for the best one, as much as possible.

Lemma 1. *For any N, there is a problem instance (K, e, v) such that an agent $i \in N$ with endowment $e_i \subset K$, who originally receives the best item and the worst $|e_i|-1$ items, receives the top $|e_i|$ item by misreporting valuation function.*

Proof. Consider $N = \{1, \ldots, n\}$, $K = \{g_1, g_2, \ldots, g_n, g_{n+1}, g_{n+2}, \ldots, g_{2n-2}\}$, $e = (e_i)_{i \in N} = (\{g_1\}, \{g_2\}, \ldots, \{g_n, g_{n+1}, g_{n+2}, \ldots, g_{2n-2}\})$, and the valuation functions $(v_i)_{i \in N}$ is given as follows:

$$\begin{aligned}
&u_1(g_2) > u_1(g_{2n-2}) > \cdots \\
&u_2(g_3) > u_2(g_{2n-3}) > \cdots \\
&\vdots \\
&u_{n-2}(g_{n-1}) > u_{n-2}(g_{n+1}) > \cdots \\
&u_{n-1}(g_n) > \cdots \\
&u_n(g_1) > u_n(g_2) > \cdots > u_n(g_n) > \cdots > u_n(g_{2n-2})
\end{aligned}$$

Figure 1 indicates the first and second best items for each agent except the manipulator $n \in N$.

When every agent truthfully reports their valuations, a cycle $g_1 \to g_2 \to \cdots \to g_n \to g_1$ is constructed at round

1 of ATTC. In the subsequent rounds only atomic players made for agent n remain. Agent n finally obtains $\{g_1, g_{n+1}, g_{n+2}, \dots, g_{2n-2}\}$, one of which is the best item and the others are the worst $n-2$ items.

Now consider the misreport v'_n by agent n, associated with u'_i given as follows:

$$u'_n(g_{n-1}) > u'_n(g_{n-2}) > \cdots > u'_n(g_1) > u'_n(g_n) > \cdots$$

Under this misreport, at each round r ($1 \le r \le n-1$) of ATTC, an atomic player, made for agent n, receives g_{n-r} for item g_{n+r-1}. Agent n thus finally obtains $\{g_1, g_2, \dots, g_{n-1}\}$, which are the best $n-1$ items. □

Obviously this is the best possible gain by an agent's valuation misreport, and thus the incentive ratio of ATTC is given as:

$$\frac{\sum_{k=1}^{n-1} u_n(g_k)}{u_n(g_1) + \sum_{l=n+1}^{2n-2} u_n(g_l)} \tag{1}$$

Based on this observation, we can now analyze the incentive ratio of ATTC in detail.

Theorem 1. *When the agents' valuation functions are additive, the ATTC for multi-item exchange has an unbounded incentive ratio.*

Proof. For given n, the maximum value of the incentive ratio, given in Eq. 1, approaches $n-1$, which realizes, for instance, when the first $n-1$ values, $u_n(g_1), \dots, u_n(g_{n-1})$, are close enough to one and all the lower values, $u_n(g_n), \dots, u_n(g_{2n-2})$, are almost zero. The ratio is therefore unbounded when $n \to \infty$. □

An additive valuation function is said to be *lexicographic* if its associated u_i satisfies the following:

$$\forall g \in K, u_i(g) \ge \sum_{h \in K \text{ s.t. } u_i(g) > u_i(h)} u_i(h).$$

Considering such a valuation function is natural when agents have extreme preferences so that their utility is determined almost solely by the best item they receives, and each of the other items he receives is considered as an extra. Under this assumption, the incentive ratio of ATTC is slightly improved.

Theorem 2. *When the agents' valuation functions are lexicographic, the ATTC for multi-item exchange has the incentive ratio of 2.*

Proof. When the valuation function u_n is given as

$$\forall g_k \in K, u_n(g_k) = 2^{|K|-k},$$

the incentive ratio becomes

$$\frac{2^{2n-3} + 2^{2n-4} + \cdots + 2^n}{2^{2n-3} + 2^{n-2} + \cdots + 2^0},$$

which converges to 2 for $n \to \infty$. □

5. Exchange with Externalities

Considering externalities in agents' utilities is a promising approach [Mumcu 07], as in most of the real-life market, a person's utility usually depends on other people's information/actions. However, in many mechanism design problems, including the exchange problem for housing market, such externality in agents' utilities causes a negative results, such as the non-existence of strategy-proof mechanisms.

A natural approach to avoid falling into such negative results is to focus on some specific structure of externalities. The *service exchange problem*, proposed by Lesca and Todo [Lesca 18] as a simple extension of the housing market, is one of such an approach. In their model, each agent considers both the item he receives *and* the agent who receives his endowment.

The service exchange problem is also a special case of exchange problem, which can be represented by setting $|K| = n$ and for any $i \in N$ and any pair $a, b (\neq a) \in A$, $v_i(a) = v_i(b)$ if and only if both $a_i = b_i$ and $\exists j \in N$ such that $a_j = b_j = e_i$. In words, an agent i is indifferent between two outcomes a and b if and only if (i) he receives the same item, and (ii) his endowment is taken by the same agent. To simplify the model, let us focus on the following valuation functions; there exists $p_i : V \to \mathbb{R}_{>0}$ and $q_i : N \to \mathbb{R}_{>0}$ such that $v_i(a) = p_i(a_i) + q_i(j)$, where $j \in N$ is the agent who takes e_i and $q_i(i) = 0$.

One naive way to implement the idea of TTC for the service exchange problem is to ignore the externality term q_i in agents' valuations and focus only on the item that each agent receives. By focusing on the receiving item, we can guarantee that the classical TTC runs without any modification.

The following theorem shows that the TTC for the problem has an unbounded incentive ratio.

Theorem 3. *The TTC for exchange with externalities has an unbounded incentive ratio, even if agents' externalities, $(q_i)_{i \in N}$, have an upper bound.*

Proof. Consider $N = \{1, 2, 3\}$, $K = \{g_1, g_2, g_3\}$, $e = (e_i)_{i \in N} = (\{g_1\}, \{g_2\}, \{g_3\})$, and the functions of manipulating agent 3, p_3 and q_3, are given as follows:

$$p_3(g_1) = p_3(g_2) + \epsilon$$
$$q_3(1) \ll q_3(2)$$

When both agents 1 and 2 most prefers item g_3, the TTC assigns g_1 to agent 3 and g_3 is assigned to agent 1, in which agent 3's valuation is $p_3(g_1) + q_3(1)$. By misreporting his valuation, agent 3 can receive g_2 and give his endowment g_3 to agent 2, in which his valuation is $p_3(g_2) + q_3(2)$. The ratio is therefore unbounded, since we can choose small enough $p_i(1)$. The ratio is still unbounded when both p_i and q_i have the same upper bound because, by adding more agents and items, we may find the above case in the very last step of TTC, where only a few items having small enough values p for agent 3 is left in the market, which still have large values q. □

6. Conclusions

Note that the analysis of incentive ratio still focuses on the worst case behavior of the market/algorithm, as the analysis of strategy-proofness does. It should be interesting to theoretically analyze the average case incentives [Kojima 09]. It is obvious that the TTC algorithms cannot always achieve the optimal outcome for extensions of , i.e., the one which maximizes *social welfare* which is defined as the sum of agents' valuations. Accordingly, analyzing their approximation factors is an open question, as several papers on both algorithms and mechanism design have considered for various problems. Developing different exchange rules that are still easy to understand, as well as having a better, i.e., smaller, incentive ratio, is another interesting direction.

Acknowledgement

This work is supported by JSPS KAKENHI Grant Numbers JP17H00761 and JP17H04695. The author thanks Etsushi Fujita, Julien Lesca, Akihisa Sonoda, and Makoto Yokoo for their helpful comments and discussion. All errors are my own.

References

[Alcalde-Unzu 11] Alcalde-Unzu, J. and Molis, E.: Exchange of indivisible goods and indifferences: The Top Trading Absorbing Sets mechanisms, *Games and Economic Behavior*, Vol. 73, No. 1, pp. 1 – 16 (2011)

[Aziz 12] Aziz, H. and Keijzer, de B.: Housing Markets with Indifferences: A Tale of Two Mechanisms, in *Proc. AAAI'12*, pp. 1249–1255 (2012)

[Chen 11] Chen, N., Deng, X., and Zhang, J.: How Profitable Are Strategic Behaviors in a Market?, in *Proc. the 19th Europ. Symp. Algorithms (ESA'11)*, pp. 106–118 (2011)

[Chen 17a] Chen, Z., Cheng, Y., Deng, X., Qi, Q., and Yan, X.: Agent Incentives of Strategic Behavior in Resource Exchange, in Proc. SAGT'17, pp. 227–239 (2017)

[Chen 17b] Chen, Z., Cheng, Y., Deng, X., Qi, Q., and Yan, X.: Limiting User's Sybil Attack in Resource Sharing, in *Proc. WINE'17*, pp. 103–119 (2017)

[Fujita 15] Fujita, E., Lesca, J., Sonoda, A., Todo, T., and Yokoo, M.: A Complexity Approach for Core-Selecting Exchange with Multiple Indivisible Goods under Lexicographic Preferences, in *Proc. AAAI'15*, pp. 907–913 (2015)

[Fujita 18] Fujita, E., Lesca, J., Sonoda, A., Todo, T., and Yokoo, M.: A Complexity Approach for Core-Selecting Exchange under Conditionally Lexicographic Preferences, *J. Artif. Intell. Res.*, Vol. 63, pp. 515–555 (2018)

[Gibbard 73] Gibbard, A.: Manipulation of voting schmes, *Econometrica*, Vol. 41, pp. 587–602 (1973)

[Kojima 09] Kojima, F. and Pathak, P. A.: Incentives and Stability in Large Two-Sided Matching Markets, *Amer. Econ. Rev.*, Vol. 99, No. 3, pp. 608–627 (2009)

[Lesca 18] Lesca, J. and Todo, T.: Service Exchange Problems, in *Proc. IJCAI-ECAI'18*, pp. 354–360 (2018)

[Ma 94] Ma, J.: Strategy-proofness and the strict core in a market with indivisibilities, *Intl. J. Game Theory*, Vol. 23, No. 1, pp. 75–83 (1994)

[Mumcu 07] Mumcu, A. and Saglam, I.: The core of a housing market with externalities, *Economics Bulletin*, Vol. 3, No. 57, pp. 1–5 (2007)

[Pápai 00] Pápai, S.: Strategyproof Assignment by Hierarchical Exchange, *Econometrica*, Vol. 68, No. 6, pp. 1403–1433 (2000)

[Saban 13] Saban, D. and Sethuraman, J.: House Allocation with Indifferences: A Generalization and a Unified View, in *Proc. ACM-EC'13*, pp. 803–820 (2013)

[Satterthwaite 75] Satterthwaite, M. A.: Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions, *J. Econ. Theory*, Vol. 10, pp. 187–217 (1975)

[Shapley 74] Shapley, L. and Scarf, H.: On cores and indivisibility, *J. Mathe. Econ.*, Vol. 1, No. 1, pp. 23–37 (1974)

[Sikdar 17] Sikdar, S., Adali, S., and Xia, L.: Mechanism Design for Multi-Type Housing Markets, in *Proc. AAAI'17*, pp. 684–690 (2017)

[Sönmez 99] Sönmez, T.: Strategy-proofness and Essentially Single-valued Cores, *Econometrica*, Vol. 67, No. 3, pp. 677–689 (1999)

[Sonoda 14] Sonoda, A., Fujita, E., Todo, T., and Yokoo, M.: Two Case Studies for Trading Multiple Indivisible Goods with Indifferences, in *Proc. AAAI'14*, pp. 791–797 (2014)

[Sun 15] Sun, Z., Hata, H., Todo, T., and Yokoo, M.: Exchange of Indivisible Objects with Asymmetry, in *Proc. IJCAI'15*, pp. 97–103 (2015)

[Todo 14] Todo, T., Sun, H., and Yokoo, M.: Strategyproof Exchange with Multiple Private Endowments, in *Proc. AAAI'14*, pp. 805–811 (2014)

[Yokoo 04] Yokoo, M., Sakurai, Y., and Matsubara, S.: The Effect of False-name Bids in Combinatorial Auctions: New Fraud in Internet Auctions, *Games and Economic Behavior*, Vol. 46, No. 1, pp. 174–188 (2004)