Progress in Artificial Intelligence manuscript No.
(will be inserted by the editor)

Co-utility

Self-enforcing collaborative protocols with mutual help

Josep Domingo-Ferrer · David Sánchez · Jordi Soria-Comas

Received: date / Accepted: date

Abstract Peer-to-peer scenarios are increasingly important in globalized information societies. To guide the interaction between (rational) agents in these P2P settings, self-enforcing collaborative protocols are desirable, so that agents have no motivation to deviate from them. Moreover, to increase social welfare and ensure the success of protocols, it would be desirable to make the collaboration between agents mutually beneficial. In this paper, we present the notion of *co-utility*, which refers to protocols in which mutual help is the best rational option to take, even for purely selfish agents. This concept is illustrated through several P2P scenarios, framed both in the information society and in the physical world, for which we present and discuss co-utile self-enforcing protocols.

Keywords Protocols · Game theory · Multi-agent systems · Self-enforcement · Co-utility

1 Introduction

Digital information and communication technologies drive globalized information societies and set the stage for interactions among human or automatic agents. These interactions are governed by protocols that specify the expected behavior and the communication patterns available to the agents. Well-designed protocols should be such that the participating agents are willing to follow them, that is, they ought to be self-enforcing. This is crucially important in distributed environments without common legal frameworks, which is the case of most Internet interactions (e.g., peer-to-peer interactions).

Self-enforcement alone does not necessarily increase the welfare of the agents participating in a protocol. Here we are specifically interested in self-enforcing P2P protocols that also promote mutually beneficial collaboration between agents, that is, a collaboration that improves the utilities of the involved agents with respect to a non-collaborative setting. If such protocols can be designed, they would not only enable seamless completion of distributed tasks, but would also promote social welfare. To capture the idea of mutually beneficial self-enforcing collaboration, we propose the notion of *co-utility* [7]. Specifically, a protocol is *co-utile* if helping the other peers that participate in the protocol to increase their utilities is the way to increase one's utility.

Self-enforcing protocols are usually designed with the most usual type of agent in mind: *rational agents*. We consider an agent to be rational if her actions in a given protocol are aimed at maximizing her utility, that is, the outcome she obtains from the execution of the protocol (e.g., money, functionality, privacy, etc.). In this sense, a rational behavior is a selfish one. Therefore, self-enforcing protocols should be designed so that rational agents do not find any incentive (i.e., increase of payoffs) to deviate from them. Moreover, for a protocol to be also co-utile, adhering to it should lead to mutual help between agents.

In this paper, we first formalize co-utility in game-theoretic terms (Section 2). Then (Section 3), we describe several P2P application scenarios where co-utility can be enforced to design mutually beneficial collaboration and, thus, unleash self-enforcing win-win interac-

UNESCO Chair in Data Privacy, Department of Computer Engineering and Mathematics
Universitat Rovira i Virgili, Av. Països Catalans 26, E-43007 Tarragona, Catalonia
Tel.: +34 977 558270
Fax: +34 977 559710
E-mail: {josep.domingo, david.sanchez, jordi.soria}@urv.cat

tions that boost social welfare. Finally, some concluding remarks are given (Section 4).

2 Co-utility

Game theory is a natural framework to deal with rational agents that are utility maximizers. The kind of scenarios to which we apply co-utility can be represented as perfect-information games, that are games in which each agent making a decision knows the payoffs of all agents under the various possible actions (or sequences of actions), plus any previously made decisions [10]. This is coherent with the usual sequential nature of protocols, in which the current state is known. Specifically, we represent these games in the so-called extensive form, which is a tree where: (i) nodes are the points where decisions are made, (ii) each node is labeled with the name of the agent making the decision, (iii) outgoing edges in a node represent the available choices (actions) at that node, and (iv) each leaf node is labeled with the tuple of payoffs that agents obtain when the node is reached.

Fig. 1 shows a game tree for the interaction between two agents A and B; each agent has two possible actions C (cooperate) and D (defect).

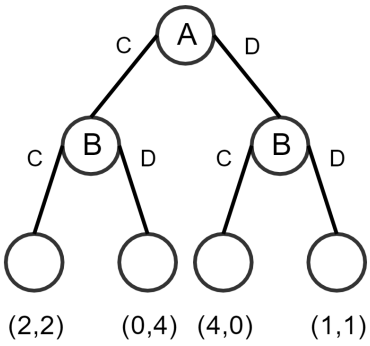


Fig. 1 Perfect-information game in extensive form. Agent A chooses an action first and then agent B reacts to the action selected by A. The payoffs are shown under each leaf node: the first value is A's payoff and the second B's payoff.

By using this extensive form, we can view a protocol execution (i.e., the actions needed for the completion of a task) as a path that traverses the tree representing the game. If the protocol allows random choices between actions, we identify the protocol with a subtree rather than a path. For the game tree in Fig. 1 there are four possible protocols (C, C), (C, D), (D, C) and (D, D), each of them representing a path from the root of the tree to a leaf node.

We will focus on self-enforcing protocols, which are those from which agents have no rational incentive to deviate. That is, no agent can increase her utility by deviating from the protocol, provided that the other agents stick to it. In game-theoretic terms, this means that, at each successive node of the protocol path, sticking to the next action prescribed by the protocol (taking the next edge in the path) is an *equilibrium* of the remaining subgame of the game (the subtree rooted at the current node), that is, a *subgame perfect equilibrium* of the game.

Finally, we say that a self-enforcing protocol is *co-utile* if it results in mutually beneficial collaboration between the participating agents. More specifically, a protocol P is co-utile if and only if *the three* following conditions hold:

1. P is self-enforcing;
2. The utility derived by each agent participating in P is strictly greater than the utility the agent would derive from not participating;
3. There is no alternative protocol P' giving greater utilities to all agents and a strictly greater utility to at least one agent.

The first condition ensures that, if participants engage in the protocol, they will not deviate. The second condition is needed to ensure that engaging in the protocol is attractive for everyone. The third condition can be rephrased in game-theoretic terms by saying that the protocol is a Pareto-optimal solution of the underlying game. The game tree in Fig. 1 does not allow for co-utile protocols: (D, D) is the only self-enforcing protocol (thus the only potentially co-utile protocol) but it is Pareto-dominated by (C, C). To illustrate co-utility, let us look at Fig. 2. Three different protocols are possible: (F, S), (F, R) and $(S, \emptyset)$. Among these (F, S) is self-enforcing and Pareto-optimal; therefore, assuming that the utility of the agents is zero when they do not participate in any protocol, (F, S) is co-utile.

In some scenarios, we can even reach a stronger case of mutual collaboration in which each agent achieves her *maximum* utility in the game. In this case, the protocol does not only lead to a Pareto-optimal payoff assignment to players, but it gives maximum payoff to all players. We call this *strict co-utility*. Protocol (F, S) in Fig. 2 is, in particular, strictly co-utile.

In general, given any game, there is no guarantee that the selfish behavior of the players will lead to co-utility, let alone strict co-utility. However, for a perfect-information game where there is one and only one leaf node that strictly maximizes the utility of all agents, the selfish behavior of the agents will cause them to follow a protocol that is strictly co-utile.

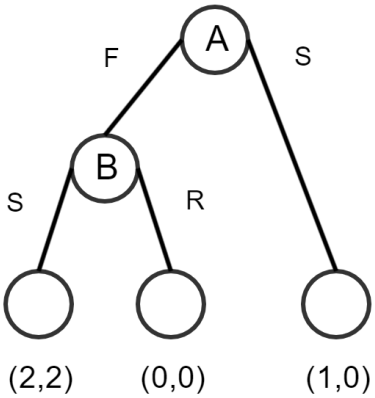


Fig. 2 Perfect-information game in extensive form. Agent A chooses first between actions F and S. When F is selected, agent B is given the chance to choose between S and R. Protocol (F, S) is co-utile.

3 Application cases

In order to illustrate the concepts introduced in Section 2, we discuss here several P2P scenarios in which co-utile protocols naturally arise.

3.1 Anonymous query submission

The first application scenario we consider involves users submitting queries to a Web search engine (WSE), a database or a Web service. In all cases, the service provider may learn some of the user's personal information, such as her census features, interests, etc., if her query refers to some of these topics [5,2]. As a result, service providers can build profiles of their users that could be used for marketing or even discriminatory purposes [15]. Because of this privacy threat, in addition to getting their queries answered, agents may also be interested in maintaining their privacy w.r.t. the service provider. In this scenario, these two utility components, functionality and privacy, are not easily reconciled, as an agent above cannot optimize both of them simultaneously.

The users' defense against being accurately profiled by the service provider consists in making their profile features flat (diverse) enough. To achieve this, users must hide their real queries about their own interests in a set of (ideally real) queries about diverse topics they are not interested in.

To implement this defense, we propose the following co-utile P2P query submission protocol: instead of submitting queries on their own, users may request other peers to submit their queries (and return the results). This may be not only good for the requester's privacy, but also for the submitter's privacy, because the requester's query can be viewed as noise concealing the actual features of the submitter to the service provider [5]; thus, if the user receiving the query request believes that the query contents can flatten his profile w.r.t. the service provider, he will be rationally interested to submit the query. Since all the users obtain a positive payoff as a result of their collaboration (i.e., functionality and privacy for the query originator and privacy for the submitter), a rational cooperation between peer agents emerges; thus, the protocol is co-utile. Even more, since the payoffs obtained by users are actually the maximum payoffs they can expect from the input query (i.e., for the query originator, having the query answered with no loss of privacy, and for the submitter, whatever privacy gain the query may bring to him), the protocol is strictly co-utile.

An example of this protocol is depicted in Figure 3, where agent profiles towards the service provider are represented as distributions of topics of interest.

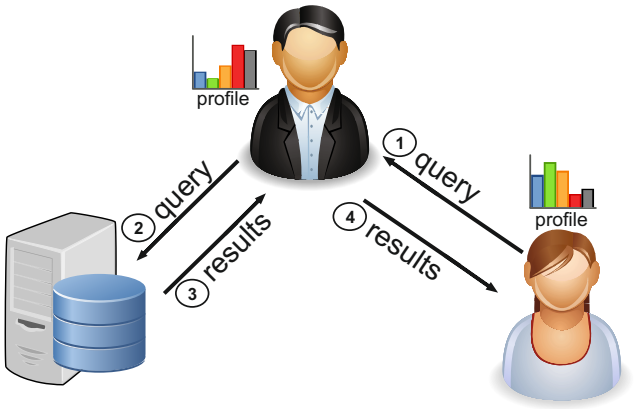


Fig. 3 Anonymous query submission protocol

3.2 P2P controlled distribution of copyrighted contents

The second co-utility application scenario we sketch refers to distribution of multimedia contents through P2P channels. Here, copyright infringement is a common practice that has become the most important threat to the content industry. To prevent this illegal content re-distribution, anonymous fingerprinting offers a technological solution based on embedding an imperceptible mark (the fingerprint) in the distributed content that is different for each content buyer. In case of illegal re-distribution, the fingerprint allows identifying the re-distributor through a tracing system, which facilitates taking legal actions against him. This should be enforced so that the identity of the buyers remains

anonymous and is only revealed in case of illegal redistribution.

Most fingerprinting schemes in the literature are centralized, so that the distribution of fingerprinted contents is basically unicast [3]; thus, they cannot be applied to the extremely popular and much more efficient P2P content distribution. With a P2P network, the merchant only needs to establish direct connections with a few seed buyers, and thereafter the content is further spread by those seeds. The challenge is how to ensure that the P2P spread content is properly fingerprinted so that it is still traceable in case of redistribution. Moreover, we would also like to ensure that the P2P fingerprinting scheme is self-enforcing and mutually beneficial (i.e., co-utile) for honest agents.

The co-utile fingerprinting protocol we propose is inspired in DNA propagation [11]. The data owner creates several copies of the content and embeds different fingerprints in each copy that identify the initial seed buyers. Then, further buyers can download the content in a P2P way by asking several seeds to provide disjoint fragments of the content (that also contain partial and different fingerprints). As a result, the new buyer will receive the content with an embedded unique fingerprint that has been created by combining fingerprint pieces of the different sources (similar to a DNA sequence). Because the resulting fingerprint is unique, from this moment onwards, it identifies the new buyer in case of illegal re-distribution. In turn, any buyer having downloaded the content in the previous way can now become a source of content for other buyers. The process is depicted in Figure 4: several sources (parents in the biological analogy) provide fragments of the content, and thus pieces of their fingerprints, to a destination (child or offspring in the biological analogy).

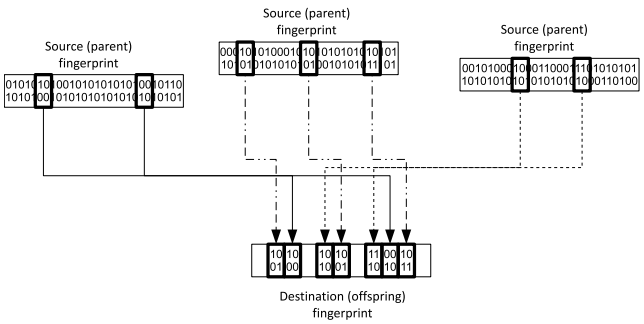


Fig. 4 P2P fingerprint construction

The described protocol is co-utile for honest agents because:

- No parent (source) is interested in sending her *entire* copy of the content (that carries the unique fingerprint identifying her) to any child (destination); otherwise, the parent could be held guilty for any misuse or undue redistribution performed by the child.
- No child is interested in receiving her copy from a single parent because, otherwise, she could be held guilty for any misuse by the parent.
- By following the P2P protocol, all the involved agents efficiently obtain the content they desire.

In this case, co-utility ensures automatic unique fingerprinting when the utility function of agents includes security (no one wants to be held guilty for other agents' misdeeds) and privacy (no one wants to lose her anonymity as a result of being held guilty).

3.3 Collaborative anonymization

The third co-utility scenario we discuss is collection of personal data from a set of privacy-conscious respondents, for example in a survey. Sets of individual data (e.g., census or health-related), also called microdata sets, are of great interest for many research initiatives [4]. However, releasing the original microdata is not feasible because it would lead to a violation of the privacy of the respondents. Statistical disclosure control (SDC) for microdata seeks to produce an anonymized version of the microdata file such that it enables valid statistical analyses but thwarts inference of confidential information about any specific individual [8].

The mainstream literature on SDC for microdata focuses on centralized anonymization, which features a trusted data collector. Local anonymization is an alternative disclosure limitation paradigm suitable for scenarios where the respondents do not trust (or only partially trust) the data collector. Each respondent anonymizes her own data before handing them to the data collector. However, the lack of information about the overall data set in local anonymization may lead to overprotection.

Collaborative anonymization is a distributed and collaborative process to generate an anonymized data set that satisfies: (i) it incurs no more information loss than the centralized paradigm for the same privacy level, and (ii) neither the respondents nor the data collector gain more knowledge about the confidential/sensitive attributes of any specific respondent than the knowledge contained in the final anonymized data set.

Several approaches for collaborative anonymization have been proposed in [13], including collaborative k-anonymity. k-Anonymity [12] is a well-known privacy model, which seeks to prevent record re-identification based on a set of attributes that are known as quasi-identifiers (these are attributes that may also be present

in external data sets containing identifiers, like for example geographic or census attributes). To prevent re-identification, k-anonymity requires each combination of quasi-identifier values present in the data set to be shared by at least k records. Attaining k-anonymity is usually a matter of running an algorithm on an already available data set. In collaborative k-anonymity, the agents giving away their data collaborate in a way that the collected data set is already k-anonymous, thus, preventing anyone (including the data collector) from seeing the non-anonymized data set.

Since the privacy protection offered by k-anonymity is based on hiding an individual within a group of at least k respondents, co-utility is natural in the collaborative approach. For each respondent, helping others to anonymize their data by following the specified protocol is the best option to keep her own data anonymous. Specifically, none of the respondents in a k-anonymous group is interested in making any other respondent in the group re-identifiable, because that renders her own record more easily re-identifiable. To illustrate, consider a respondent A whose record is hidden within a group of records G; if the record of another respondent in G is re-identified, the record of A becomes hidden within a group of only $|G| - 1$ records, and thus loses some protection. Hence A is interested in no one in G being re-identified.

3.4 Physical world applications

Although the information society is attractive to deploy co-utility because of the formalized and controlled nature of protocols, co-utility can also be applied in the physical world.

For example, sharing one's car with one or several passengers can result in mutual benefits: passengers are able to reach their destination faster and the car driver needs to support less expense for the trip (because toll or gas fees are split among the passengers or because the driver can use facilities reserved for high occupancy vehicles or because of both reasons) [1].

If only travel time and expenses are considered, it is in the best interest of the individuals to share the ride, that is, car sharing is co-utile.

However, several real-life circumstances may hinder co-utility in car sharing: for example, if origin and destinations of drivers and passengers are not exactly the same, or if they are reluctant to travel with strangers (for safety reasons, for fear of an unpleasant company, etc.). To neutralize these negative payoffs and facilitate co-utility, additional utilities/incentives can be added. Specifically, the additional expenses incurred by the car owner when origins and/or destinations are different can be compensated by requiring passengers to pay a reasonable fee that still makes sharing the ride attractive to everyone. Likewise, reluctance to each other's company can be mitigated by introducing a reputation mechanism, whereby all parties get reciprocally rated after sharing a trip. A party's reputation can be used as a subsequent reference. For reputation to be useful to facilitate co-utility, the reputation mechanism needs to be not only distributed (like [9]), but also co-utile itself: distributedness is necessary to avoid having to rely on a trusted third party, and reputation co-utility is necessary to ensure that all users will collaborate to keep reputations updated and reliable [6].

Co-utility can also be applied to foster the collaborative economy, and specifically crowdsourcing. Crowdsourcing means outsourcing a task to an anonymous group of self-interested individuals by means of an open call to the crowd offering rewards for work. Here, the crowdsourcing marketplace acts as the link between the crowd of workers and the crowdsourcers of tasks. The marketplace is used by crowdsourcers to publish tasks (including the associated requirements and the pay offered) and by workers to contact crowdsourcers that have published suitable tasks. This gives rise to co-utile interactions. In particular, the interaction is self-enforcing because agents act selfishly according to their interests. Moreover, it is *strictly co-utile* because each party selects, among the available options (i.e., work offers and matching workers), the one that maximizes her utility [14].

4 Conclusions

We have discussed and exemplified the co-utility notion. Co-utile protocols are not only self-enforcing, but they also provide Pareto-optimal utilities to the participants. If these Pareto-optimal utilities are also maximal for *all* participants, we speak of strict co-utility.

As we have shown through the application cases of Section 3, there are many P2P interactions for which co-utile protocols naturally arise. However, when this is not the case or when negative payoffs (e.g., costs, reluctance, fear, etc.) render mutual help not rational anymore, we need to implement mechanisms to help co-utility happen. This mechanism design may involve either changing/extending the set of actions available to agents to fulfill a task (i.e., changing the game) or adding artificial utilities/incentives (e.g., reputation-based utilities) that can compensate the effect of negative payoffs.

Acknowledgements Funding by the Templeton World Charity Foundation (grant TWCF0095/AB60 "CO-UTILITY") is gratefully acknowledged. Also, partial support to this work has been received from the Government of Catalonia (ICREA Acadèmia Prize to J. Domingo-Ferrer and grant 2014 SGR 537), the Spanish Government (projects TIN2011-27076-C03-01 "CO-PRIVACY" and TIN2014-57364-C2-1-R "SmartGlacis") and the European Commission (project H2020-644024 "CLARUS"). The authors are with the UNESCO Chair in Data Privacy, but the views in this paper are the authors' own and are not necessarily shared by UNESCO.

References

1. BlaBlaCar. http://www.blablacar.com/ (2015). Accessed 11 November 2015
2. Carminati, B., Ferrari, E., Tran, N.H.: A privacy-preserving framework for constrained choreographed service composition. In: 2015 IEEE International Conference on Web Services, pp. 297-304. IEEE (2015)
3. Chang, C.-C., Tsai, H.-C., Hsieh Y.-P.: An efficient and fair buyer-seller fingerprinting scheme for large scale networks. Computers & Security 29(2), 269–277 (2010)
4. Daries, J.P., Reich, J., Waldo, J., Young, E.M., Whittinghill, J., Seaton, D.T., Ho, A.D., Chuang, I.: Privacy, anonymity and big data in the social sciences. Communications of the ACM 57(9), 56-63 (2014)
5. Domingo-Ferrer, J., Ú. González-Nicolás, U.: Rational behavior in peer-to-peer profile obfuscation for anonymous keyword search. Information Sciences 185, 191-204 (2012)
6. Domingo-Ferrer, J., Farràs, O., Martínez, S., Sánchez, D., Soria-Comas, J.: Self-enforcing protocols via co-utile reputation management. Submitted manuscript (2015)
7. Domingo-Ferrer, J., Soria-Comas, J., Ciobotaru, O.: Co-utility: self-enforcing protocols without coordination mechanisms. In: Proc. of the 5th International Conference on Industrial Engineering and Operations Management, pp. 1-7. IEEE (2015)
8. Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Schulte Nordholt, E., Spicer, K., de Wolf, P.-P.: Statistical Disclosure Control. Wiley (2012)
9. Kamvar, S.D., Schlosser, M.T., Garcia-Molina, H.: The EigenTrust algorithm for reputation management in P2P networks. In: Proc. of the 12th International Conference on World Wide Web, pp. 640-651. ACM (2003)
10. Leyton-Brown, K., Shoham, Y.: Essentials of Game Theory: A Concise, Multidisciplinary Introduction. Morgan & Claypool (2008)
11. Megías, D., Domingo-Ferrer, J.: Privacy-aware peer-to-peer content distribution using automatically recombined ringerprints. Multimedia Systems 20(2), 105-125 (2014)
12. Samarati, P., Sweeney, L.: Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. Technical report, SRI International (1998)
13. Soria-Comas, J., Domingo-Ferrer, J.: Co-utile collaborative anonymization of microdata. Lecture Notes in Computer Science, Vol. 9321, pp. 192-206 (2015)
14. Turi, A.N., Domingo-Ferrer, J., Sánchez, D., Osmani, D.: Co-utility: conciliating individual freedom and common good for the crowd-based business model. In: Proc. of the 12th International Conference on e-Business Engineering (2015)
15. Viejo, A., Sánchez: Profiling social networks to provide useful and privacy-preserving web search. Journal of the Association for Information Science and Technology 65(12), 2444-2458 (2014).