Information Disclosure by Decentralized Coordination in Virtual Power Plants and District Energy Systems

Jörg Bremer[1] and Sebastian Lehnhoff[2]

Abstract: Grouping small, hardly predictable, and volatile energy resources to jointly operating virtual power plants with sufficient flexibility for coordination is widely seen as a key aspect of integrating renewable energy into the grid. For several reasons, self-organizing, agent-based systems are probably the best technology for coordination. A major drawback of many currently existing solutions is the necessity to communicate plain information for negotiation and optimization. Such information contains e.g. possible energy generation schemes or aggregated costs. Previous works have already shown that identification of anonymously sent information is possible. In this paper, we demonstrate the possibility of disaggregating cost structure information as an example of possible leakage of business information in the case of participation in virtual power plants or district energy systems. From this perspective, we derive measures to ensure privacy preservation in decentralized coordination algorithms.

Keywords: Virtual Power plant; Distributed Optimization; Self-Organization; Data Privacy

Addresses Sustainable Development Goal 7: Affordable and clean energy

1. Introduction

Integrating as much as possible renewable energy sources into the energy grid is one of the most crucial task today – not only for fighting global warming, but also for ensuring energy safety. One goal here is achieving energy efficiency by a higher penetration of renewable feed-in [Be13; Ka11]. Higher penetration of renewable energy needs modern concepts for integration into the power grid due to their volatile nature and low flexibility. Low voltage coupling points with the grid lead to partially inversed power flows and demand local optimization. One way to cope with volatility and small size is bundling of different energy resources and orchestrating them via communication and joint control. This concept is also known as virtual power plant (VPP) [NM12].

For many use cases, it is advantageous to bundle energy resources within a local region. Here, energetic neighborhoods come into play, especially regarding multi-modal energy systems in which for example a complex interplay of electricity and district heating grid are scrutinized for synergies [Th17]. The basic coordination problem and thus also the resulting problems are mostly the same as in VPP.

Multi-agent-based systems are widely considered to be a valid approach for coordinating

[1] University of Oldenburg, Energy Informatics, Uhlhornsweg 84, 26129 Oldenburg, Germany, jbr@offis.de
[2] University of Oldenburg, Energy Informatics, Uhlhornsweg 84, 26129 Oldenburg, Germany, lehnhoff@uol.de

a large number of distributed entities or sensing or operation equipment, and to solve distributed optimization and control problems in cyber-physical systems, especially for horizontal-distributed control tasks. In addition, hierarchical topologies supporting vertical- distributed control are already available. For optimizing industrial production and logistics processes, multi-agent systems have been on the research agenda and been used for many years. Use cases comprise for example supply chain management [CM06], or production planning [Tö02], and scheduling [OK07].

Solving a problem with the help of autonomously acting distributed entities – such as agents – naturally raises an issue with safety concerns; especially if a decentralized consensus solution has to be implemented in critical infrastructures or processes [Fa18]. Solving such problems within agent coalitions needs the exchange of information for negotiation. Exchanging information via messages is necessary to build up agents' beliefs for problem solving and inevitably allows insight into other agents' options.

As an example, in the predictive scheduling use case in VPPs frequently operational schedules are exchanged as proposal for the own choice of action [HS16]. Each schedule contains data of the possible portion of energy that may be generated (or consumed) during the same given time period. Today, the resolution is usually 15 minutes per time interval (often for a day). In future, finer grained resolutions can be expected. With each round during negotiation, a new possible operable schedule is sent to several other agents together with transient information on other agents' schedules. As the underlying problem is of a multi-objective nature, several performance indicators for evaluating a schedule according to different criteria often accompany the mere electro-technical information.

This is also known as gossiping principle that comes into play into many decentralized coordination algorithms [KV07]. Such information can be collected and aggregated by malicious agents. In the case of energetic neighborhoods, actors from a close vicinity are drawn closely together for long-term collaboration. This means that data from more than just a single coordination process could be collected to extract some meaningful information; [Da18] already showed that collected information can easily be assigned to specific energy devices, and even to corresponding businesses. Different machine learning methods have been tested to achieve this. Moreover, collected and aggregated schedules allow for deriving detailed information on internal processes - heating profiles, and thus working hours, machinery load factors in case of internal consumption optimization with batteries, or current capacity utilization [BL19; Da18].

In this contribution, we extend the concern and demonstrate that is also possible to derive cost and pricing information from individual tariffs even if contained only in aggregated form. The rest of the paper is organized as follows. We recap some previous work on data privacy in distributed algorithms and present a use case study that reveals information disclosure even for aggregated data in distributes optimization.

2. Related Work

Advances in information technologies have further increased long existing concerns of privacy. When it comes to autonomous agents acting on behalf of a business, several privacy and information leakage concerns can be raised. A good overview can be found in [SEG14].

Surprisingly low effort has so far been put into the question of data privacy when it comes to (decentralized) algorithm design. A method based on the alternating direction method of multipliers (ADMM) for solving decentralized optimization in an agent system with preserved privacy can be found in [ZAW19]. ADMM solves convex optimization problems by breaking them into smaller pieces that can for example be solved individually by agents [Bo11]. To incorporate privacy preservation, partially homomorphic cryptography has been integrated. Unfortunately, applications are limited to convex functions and the method cannot be applied to non-convex black-box optimization, what is often the case in decentralized agent coordination scenarios. An extension of the use of ADMM to distributed machine learning can be found in [Wa20].

Two frequently occuring tasks in multi-agent systems are distributed constraint satisfaction (DisCSP) and distributed constraint optimization (DCOP). For these two problem classes, algorithms have been developed that aim at preserving anonymity in multi-agent problem solving. The major concern in DisCSP and DCOP algorithms is that they usually leak information that can be exploited to infer private information of other agents [GPT06]. By integrating anonymity into specialized protocols, shared information cannot be linked to the corresponding agent. Examples can be found in [BM03; SGG07; SM04; YSH05]. Some of them still leak at least some information. All these approaches still communicate plain information and try just to disguise the sender of the information. On the other hand, [Da18] demonstrated for some coordination tasks in decentralized energy management that enough information is exchanged to still identify the origins of shared information.

A way more frequent use case in the energy sector is coordination of energy generation and consumption. In [BL19] a prototypical application was scrutinized that uses order preserving encryption [Ag04] to solve the predictive scheduling problem in virtual power plants. The possibility of direct integration as well as the performance which barely degraded by encoding offers some advantages over other encryption schemes. Simply the objective function that is minimized by the agents (locally as well as globally) has to be restated. Because no mathematical functions are supported, objective functions can only rely on the order of input values from different agents. An implementation for the sum of input values has been shown in [BL19]. On the other hand, this is already the biggest disadvantage of this method. Only a few number of special cases can be implemented with these methods.

Collected information can be analyzed with appropriate machine learning methods. In [Da18], an algorithm for decentralized, agent-based scheduling in virtual power plants has been scrutinized. It was found that schedules can be properly assigned to specific devices

(and thus businesses) with machine learning. In district energy systems, the concrete business behind an agent may even already be known due to public information on the other members in the energetic neighborhood. In the researched examples, the collected schedules allowed for deriving detailed information on internal processes and thus on heating profiles, working hours, machinery load factors in case of internal consumption optimization with batteries, current capacity utilization, etc. The same holds true for consumption patterns.

3. Data Disclosure in Self-Organization

As a case study for a possible leakage of information during self-organized coordination in VPP we consider the disclosure of individual cost or prices (depending on whether we spy out a generator or a consumer).

3.1 Problem Description

Predictive scheduling is the problem of finding an operation schedule (determining the individual course of generated or consumed power) for each energy resource within a VPP for a given future time horizon. Today, often planning is made for 96 time intervals of 15 minutes each for the next day.

This constitutes a distributed combinatorial nature of the optimization problem [18] for which several solution have been proposed [20], [35], [14]. Decentralized algorithms are seen as the most promising approach due to the distributed architecture and problem size. Additionally, in district energy systems of energy cooperatives of individual and self-dependent actors, centralized authorities that dictate the generated amount of power may spoil acceptance.

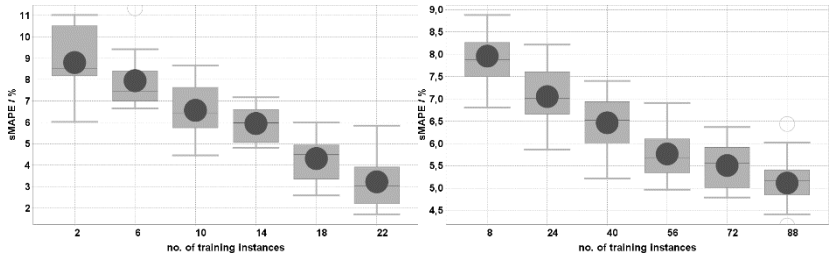
Solving distributed problems with agent-based, decentralized approaches leads to information exchange to build up the agents' beliefs for problem solving. This information could be collected and aggregated. In case of energetic neighborhoods, actors from a close vicinity are drawn closely together for long-term collaboration. In this case, schedules from more than just a single optimization process could be gathered. Collected schedules would allow for deriving detailed information on internal processes – heating profiles, and thus working hours, machinery load factors in case of internal consumption optimization with batteries, current capacity utilization, etc. The same holds true for consumption patterns. Thus, schedules and thereof derived phase spaces of device operations should actually not be publicly known.

3.2 Case Study

For our case study, we consider a small business that plans internal production schemes after some individually with the energy provider negotiated time of use tariff. In this way, for each time interval of the planning horizon, a different energy price has to be paid [So15]. W.l.o.g., we generated random prices $c \in [40, 80]$ cent for our experiments. We assume the following scenario. A group of distributed energy resources (as members of a VPP or a district energy system) is conducting a distributed (day-ahead) planning of energy consumption and generation with the goal of balancing as much as possible and to minimize overall energy cost. To incorporate individual cost in the decentralized balancing algorithm, different schedules are sent by the agents as proposal during negotiation. Each schedule s_i must be annotated with total individual energy cost

$$c_i = \sum_{j=1}^{d} s_i[j] \cdot c[j] \tag{1}$$

as these cannot be calculated by the other agents. Tariffs are not public. During coordination, a set of n different schedules could be collected by a fraudulent agent. If there is a number of schedules available equal to or greater than the schedule dimension, the system of equations is fully determined and the exact energy cost for each time interval can be derived. But, also if the system is undetermined, we can try to build an approximate model of the cost. We conducted some experiments with particle swarm optimization to fit the model.



(a) Scenario 1: 24 intervals, 1 day (b) Scenario 2: 96 15-minute intervals, 1 day

Fig. 1: Model fitting results of tariff estimation by with particle swarm optimization for two one day scenarios.

For the experiments, one agent was chosen to be the fraudulent agent during optimization. It was the task of this agent to collect all schedules of the other agents together with the total energy cost according to different individual time of use tariffs. From the collected unique schedules of a specific agent, tariff information can be calculated if the system is fully determined. This is the case as soon as the number of collected schedules is equal to or larger than the number of time intervals in a schedule. But even with less schedules an approximation can be calculated. We use particle swarm optimization to fit a model for the under determined system. For each model fitting a slightly different best guess for the tariff

approximation will be found. Thus, we repeated model fitting 20 times and took the mean tariff as the approximation.

Fig. 1 show the result for two different scenarios: 24 and 96 time intervals with different numbers of collected schedules that could be used for model fitting. As optimization protocol

dim.	no. of agents		
d	10	25	50
96	32.51 ± 15.65	36.23 ± 19.35	38.11 ± 14.23
24	10.53 ± 4.04	11.68 ± 1.74	15.39 ± 7.24
8	6.64 ± 3.05	7.25 ± 1.79	7.46 ± 2.92

Tab. 1: Number of mean unique schedules (per agent) of other agents that a fraudulent agent sees during a single distributed optimization process.

we used the one proposed in [HS17]. For measuring the quality of the model, we used the symmetric mean absolute percentage error (sMAPE):

$$sMAPE(f,a) = \frac{100}{n}\sum_{i=1}^{n}\frac{|f_i - a_i|}{(|a_i| + |f_i|) - 2} \tag{2}$$

The results show that a relative good fitting can be achieved with a rather small number of schedule information. Tab. 1 shows the result of another experiment. Each agent counted the number of unique schedules from other agents that were seen during the optimization process. This result shows that already during a single optimization process enough schedules are communicated to leak information also for indirect data that is contained only in aggregated form (like time interval individual pricing information).

From these results as well as from the findings of [Da18], we clearly see that is necessary to raise awareness for privacy interests and appropriate measures in distributed problem solving in (future) cyber-physical energy systems with a high share of autonomous functions. Research is still at the beginning, when it comes to suitable encryption schemes that could be used to secure the shared information in such systems. Some first examples for the centralized server can be found in [KLG19], but proper best practices and design schemes for systematically integrating these approaches into distributed problem solving are missing so far and are worth to be given more attention in the future.

4. Conclusion

With this contribution, we wanted to raise a general awareness of the information disclosure problem in distributed and self-organized systems. When local information on possible behavior is spread to other actors in a multi-agent system in order to achieve some consensus on coalition behavior, the chance of revealing private information is often unintentionally given. As today, obviously there is no technology, which could be used

out-of-the-box to tackle data masking in decentralized algorithms.

Thus, more research in the field of encrypted (distributed) optimization and self-organization is highly recommended.

Bibliography

[Ag04] Agrawal, R.; Kiernan, J.; Srikant, R.; Xu, Y.: Order preserving encryption for numeric data. In: Proceedings of the 2004 ACM SIGMOD international conference on Management of data. S. 563–574, 2004.

[Be13] Becker, T.; Boschert, S.; Hempel, L.; Höffken, S.; Obst, B.: Complex urban simulations and sustainable urban planning with spatial and social implicati- ons. In: Int Ann Photogrammetry, Remote Sensing Spat Inf Sci, ISPRS 8th 3DGeoInfo conference and WG II/2 workshop. Bd. 2, W1, 2013.

[BL19] Bremer, J.; Lehnhoff, S.: Encrypted Decentralized Optimization for Data Masking in Energy Scheduling. In: Proceedings of the 2019 3rd International Conference on Big Data Research. ICBDR 2019, Association for Computing Machinery, Cergy-Pontoise, France, S. 103–109, 2019, isbn: 9781450372015, url: https://doi.org/10.1145/3372454.3372487.

[BM03] Brito, I.; Meseguer, P.: Distributed forward checking. In: International Confe- rence on Principles and Practice of Constraint Programming. Springer, S. 801– 806, 2003.

[Bo11] Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. et al.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine learning 3/1, S. 1–122, 2011.

[CM06] Chaib-Draa, B.; Müller, J.: Multiagent based supply chain management. Springer Science & Business Media, 2006.

[Da18] Dabrock, K.: Privacy in der automatisierten prdiktiven Einsatzplanung von Energieanlagen im Smart Grid, Magisterarb., University of Oldenburg, Dept. of Energy Informatics, Germany, 2018.

[Fa18] Farris, I.; Taleb, T.; Khettab, Y.; Song, J.: A survey on emerging SDN and NFV security mechanisms for IoT systems. IEEE Communications Surveys & Tutorials 21/1, S. 812–837, 2018.

[GPT06] Greenstadt, R.; Pearce, J. P.; Tambe, M.: Analysis of privacy loss in distributed constraint optimization. In: AAAI. Bd. 6, S. 647–653, 2006.

[HS16] Hinrichs, C.; Sonnenschein, M.: Design, analysis and evaluation of control algorithms for applications in smart grids. In: Advances and New Trends in Environmental and Energy Informatics. Springer, S. 135–155, 2016.

[HS17] Hinrichs, C.; Sonnenschein, M.: A distributed combinatorial optimisation heuristic for the scheduling of energy resources represented by self-interested agents. International Journal of Bio-Inspired Computation 10/2, S. 69–78, 2017.

[Ka11] Karnouskos, S.: Demand side management via prosumer interactions in a smart city

energy marketplace. In: 2011 2nd IEEE PES international conference and exhibition on innovative smart grid technologies. IEEE, S. 1–7, 2011.

[KLG19] Kaiser, P.; Langer, A.; Gaedke, M.: MPCC: Generic Secure Multi-Party Com- putation in Centralized Cloud-based Environments. In: Proceedings of the 2019 3rd International Conference on Big Data Research. S. 60–66, 2019.

[KV07] Kermarrec, A.-M.; Van Steen, M.: Gossiping in distributed systems. ACM SIGOPS operating systems review 41/5, S. 2–7, 2007.

[NM12] Nikonowicz, L. B.; Milewski, J.: Virtual power plants-general review: structure, application and optimization. Journal of power technologies 92/3, S. 135, 2012.

[OK07] Opadiji, J. F.; Kaihara, T.: Distributed production scheduling using federated agent architecture. In: International Conference on Industrial Applications of Holonic and Multi-Agent Systems. Springer, S. 195–204, 2007.

[SEG14] Such, J. M.; Espinosa, A.; García-Fornes, A.: A survey of privacy in multi-agent systems. The Knowledge Engineering Review 29/3, S. 314–344, 2014.

[SGG07] Smith, M.; Grosz, B.; Greenstadt, R.: SSDPOP: Improving the Privacy of DCOP with Secret Sharing. Proceedings of Autonomous Agents and Multi- Agent Systems, AAMAS-2007/, 2007.

[SM04] Silaghi, M.-C.; Mitra, D.: Distributed constraint satisfaction and optimization with privacy enforcement. In: Proceedings. IEEE/WIC/ACM International Conference on Intelligent Agent Technology, 2004.(IAT 2004). IEEE, S. 531– 535, 2004.

[So15] Sonnenschein, M.; Lünsdorf, O.; Bremer, J.; Tröschel, M.: Decentralized control of units in smart grids for the support of renewable energy supply. Environmental Impact Assessment Review 52/, S. 40–52, 2015, issn: 0195-9255.

[Th17] Thiem, S.; Danov, V.; Metzger, M.; Schäfer, J.; Hamacher, T.: Project-level multi-modal energy system design-Novel approach for considering detailed component models and example case study for airports. Energy 133/, S. 691– 709, 2017.

[Tö02] Tönshoff, H.; Herzog, O.; Timm, I.; Woelk, P.: Integrated process planning and production control based on the application of intelligent agents. Proc. 3rd CIRP ICME/, S. 135–140, 2002.

[Wa20] Wang, X.; Ishii, H.; Du, L.; Cheng, P.; Chen, J.: Privacy-preserving distributed machine learning via local randomization and ADMM perturbation. IEEE Transactions on Signal Processing 68/, S. 4226–4241, 2020.

[YSH05] Yokoo, M.; Suzuki, K.; Hirayama, K.: Secure distributed constraint satisfac- tion: Reaching agreement without revealing private information. Artificial Intelligence 161/1-2, S. 229–245, 2005.

[ZAW19] Zhang, C.; Ahmad, M.; Wang, Y.: ADMM Based Privacy-Preserving Decentra- lized Optimization. IEEE Transactions on Information Forensics and Security 14/3, S. 565–580, 2019.