Environ Syst Decis (2017) 37:13–21
DOI 10.1007/s10669-016-9620-7

Privacy-preserving aggregation in life cycle assessment

Brandon Kuczenski[1] · **Cetin Sahin**[2] · **Amr El Abbadi**[2]

Published online: 1 December 2016

Abstract Life cycle assessment (LCA) is the standard technique used to make a quantitative evaluation about the ecological sustainability of a product or service. The life cycle inventory (LCI) data sets that provide input to LCA computations can express essential information about the operation of a process or production step. As a consequence, LCI data are often regarded as confidential and are typically concealed through aggregation with other data sets. Despite the importance of privacy protection in publishing LCA studies, the community lacks a formal framework for managing private data, and no techniques exist for performing aggregation of LCI data sets that preserve the privacy of input data. However, emerging computational techniques known as "secure multiparty computation" enable data contributors to jointly compute numerical results without enabling any party to determine another party's private data. In the proposed approach, parties who agree on a shared computation model, but do not trust one another and also do not trust a common third party, can collaboratively compute a weighted average of an LCA metric without sharing their private data with any other party. First, we formulate the LCA aggregation problem as an inner product over a foreground inventory model. Then, we show how LCA aggregations can be computed as the ratio of two secure sums. The protocol is useful when preparing LCA studies involving mutually competitive firms.

Keywords Life cycle assessment · Secure multiparty computation · Aggregation · Privacy · Confidentiality

1 Introduction

1.1 Confidentiality of process inventory data

Industrial Ecology (IE) comprises a collection of research methodologies for quantifying the flows of materials, products, and services through the industrial economy and for estimating the potential social and ecological (collectively environmental) implications of those flows. The objects of IE research are industrial activities, conducted either for subsistence or (more commonly) for profit. Because industrial processes are typically undertaken in a competitive economic context, the operators of these processes would like to prevent potential competitors from learning sensitive information about their activities. Information that may be valuable to a competitor is often termed confidential business information and therefore is often not freely available.

At the same time, many different kinds of organizations are motivated to make public disclosures about their environmental performance. These motivations may be inspired by regulatory requirements, marketing initiatives, or as part of a broader project of corporate sustainability. A central technique for evaluating sustainability performance is life cycle assessment (LCA), an analytic methodology for estimating the cumulative environmental impacts associated with delivering a particular product or service to

Originally Accepted in the Proceedings of the 2016 International Symposium on Sustainable Systems and Technology (ISSST 2016).

✉ Brandon Kuczenski
bkuczenski@ucsb.edu

[1] Institute for Social, Behavioral and Economic Research, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

[2] Department of Computer Science, University of California, Santa Barbara, Santa Barbara, CA 93106, USA

a consumer (ISO 2006). LCA considers the life cycle of a product from "cradle-to-grave," i.e., including impacts all the way upstream to the extraction of raw materials and all the way downstream to the final disposal of all products (Curran 1996; Finnveden et al. 2009).

There is an increasing pressure from consumers and policy makers to ensure the availability of inventory information to support LCA studies. Standards and specifications are emerging for a variety of environmental product declarations and product environmental footprints (Hunsager et al. 2014). In many markets, there are increasing requirements for disclosure of environmental footprints or availability of environmental data (Bateman et al. 2017), some of which can be satisfied through EPDs or other applications of LCA. Moreover, the International Forum on LCA Cooperation, coorganized by the United Nations Environment Programme, is supporting an initiative to develop a "global LCA data access network" that would provide inventory data in a manner "that allows defining fitness for purpose by any user" (UNEP 2016).

Because of the deeply interconnected nature of the industrial economy, the operation of a given process may generate impacts far afield from where the process of interest is located. Preparing an LCA study thus requires information about a wide range of industrial processes sometimes unrelated to the product system under direct investigation. This information is often provided in the form of a life cycle inventory (LCI) database, which is a comprehensive and self-consistent model of select processes in the global economy. Preparing an accurate and comprehensive LCI database is a tremendous task, and the development and maintenance of these resources are an ongoing challenge (UNEP/SETAC 2011).

The counterpart to the LCI database is the foreground data which directly describe the product system of interest (Kuczenski 2015). The foreground is made up of unit processes that together produce the product. A unit process is defined by its inventory, a list of economic or environmental flows into and out of an industrial process or network of processes. The magnitudes of the input and output flows can be developed through direct observation or engineering or economic modeling, and can express sensitive information about the operation of a production step. By implication, the preparation of LCI resources for general use, as well as the publication of LCA studies, must be done in a way that conceals proprietary information, while still establishing the veracity of the results to an independent observer.

As a consequence of the wide breadth of technical information required to prepare a process inventory, preserving the confidentiality of process information has been a principal concern since the very beginning of LCA (Hunt and Franklin 1996; Frischknecht 2004). Engagement with stakeholders and supply chain partners is often required for effective consideration of life cycle environmental sustainability, which accentuates confidentiality concerns (Kaenzig et al. 2010). Even when dealing with direct supply chain partners, it is not always possible to negotiate access to proprietary data (Nakano and Hirao 2011; Solér et al. 2010). While the secrecy of private data is often mentioned in publications of LCA results, there has been no significant development of techniques for managing private data. Trusted data providers often use aggregation techniques to publish simpler and more representative inventory models without violating nondisclosure requirements (Koffler 2016). Aggregation of results is often assumed to protect confidentiality to a suitable level (UNEP/SETAC 2011; p. 72; see also Section 1.2), but there is no established means for evaluating whether publications are effective at preserving secrecy, nor is there a means for validating the correctness of aggregations without revealing all the proprietary data to a third party. Finally, aggregation significantly reduces the transparency of the model, which limits the usefulness of the results to independent interpretation and reuse.

No current technique permits LCA to be used to support a comparison of the performance of mutually competitive firms, without requiring all the participants to share their data with a common third party to perform the aggregation. In this paper, we present a model for a cryptographic application that would enable a computation to be performed that preserves the privacy of all input data, without requiring disclosure to any party. All participants in the computation would together learn the result, but no member (nor any third party) would learn anything else. The parties could then maintain the secrecy of the result and use it as a benchmark for their own operations. Alternatively, they could choose to prepare a publication that would disclose aggregate results that could be used independently for LCA studies. Either course could be accomplished without relying on a third party to have direct knowledge of all the confidential inputs.

1.2 Aggregation in LCA

Data that are regarded as confidential by the owners can be concealed through aggregation with other data sets (see UNEP/SETAC 2011; ch. 3). There are several ways of aggregating data:

- Horizontal aggregation or horizontal averaging is used to combine the reports of several data providers who are all operating generally the same industrial process, usually through a volume-weighted average of data values.
- Vertical or gate-to-gate aggregation refers to combining several sequential production steps into a single data set, so that the contributions of the individual steps, as well as the identities of the data providers, are hidden.

Vertically aggregated processes are combined by summing individual inputs weighted by the relative activity levels of the processes involved in the aggregation.

- Cradle-to-gate or cradle-to-grave aggregation involves combining the direct and upstream inputs of some or all process inputs, including the resolution of loops or cyclical dependencies in the chain of upstream suppliers. Cradle-to-gate aggregation is usually performed through matrix inversion or via iterative techniques that approximate it.

Study authors and database publishers can also design different combinations of these methods. Interested readers are directed to the UNEP/SETAC report cited above. All aggregation methods currently require the data providers to agree to provide their data to the study author, who aggregates them and publishes the results. Often this role is played by an industry association or trade group (World Steel Association 2011; Franklin Associates 2007). In consideration of contemporary computer science, this technique is not regarded to preserve the privacy of the data contributors, because they may not trust the aggregator to protect the confidentiality of their information.

The outcome of an aggregation is a computation of a figure of merit, such as an environmental impact score or a resource demand that represents the aggregated system. In the first two types of aggregation, horizontal and vertical, the aggregation result is obtained from a fixed set of data contributors, each representing distinct processes, whose operations are proximate to one another in a given product system. In the horizontal case, the different processes are operating in parallel, whereas in the vertical case the processes are linked together sequentially. In both these cases, the processes involved in the computation are all in the foreground of the study. The computational result of interest may be represented as a weighted sum. In Sect. 2, we will show how such a quantity can be computed in a privacy-preserving manner.

In contrast, cradle-to-gate and cradle-to-grave aggregation each involve combining part or all of the study foreground with one or many background processes, including the complete supply chain. Performing this type of aggregation requires access to a complete LCI database or to a collection of cradle-to-gate inventories derived from such a database. This computation is not considered in the model presented here.

1.3 Privacy, secrecy, and anonymity

In computer science, the concept of "confidentiality" is represented in a variety of forms. Much of cryptography is concerned with the ability of one or more parties to perform a computational task in the presence of an adversary, who wishes to obtain knowledge of the computation. This is familiar in the concept of a "shared secret," which is a piece of information exchanged between parties in a secure communication (Menezes et al. 1996). This model reflects how LCI data providers ideally interact with data aggregators, including study authors and LCI database maintainers in current practice. The secret is known to both the provider and the aggregator but not to the public, and the aggregator can then perform an LCA computation to determine whatever result is desired.

In many cases, the aggregator (which may be working on the data provider's behalf) may wish to publish the results of the computation with an audience that may include the general public at large. In this case, the data provider will be concerned about the possibility that the secret can be deduced from the publication. The publication is said to be "privacy-preserving" if a reader of the data, who is potentially an adversary, cannot discern any information that the provider regards as private (Fung et al. 2010). "Privacy" may include any number of things, including the identity of the data provider, the form of the data provider's contribution, and the values of any data points. In practice, the meaning of privacy is quite vague, because it is impossible to account for an adversary's possible background knowledge about the individual (Dwork 2006). Privacy can be described in terms of the concepts of anonymity and secrecy. "Anonymity" indicates that an adversary cannot link a particular individual to a particular publication or to some aspect of a publication. On the other hand, "secrecy" indicates that an attacker cannot know the value of a variable in the computation.

In LCA, a unit process inventory dataset, a life cycle inventory, and a life cycle impact assessment result all constitute results derived from confidential data, and therefore, the publication of any of these elements should be considered from the perspective of privacy preservation. Both anonymity and secrecy are obtained through aggregation. The "background knowledge" held by an adversary may include first the identities of the various firms involved in a particular product system, industry group, or region. The background knowledge of a competitor likely extends to detailed information about process requirements on a generic basis. Often in LCA, anonymity is ensured by aggregating the results of multiple data providers whose identities are not disclosed (Finnveden et al. 2009; sec. 6). Secrecy is often provided in the same way by mandating that every data point represents "at least three" contributors (e.g., Weidema et al. 2013). The argument, often unstated, is that if an average result includes only two contributors, then each member would be able to deduce

the other's contribution by subtracting his own from the total. With three contributors, this is regarded as infeasible.

1.4 Secure multiparty computation

One of the classic problems of cryptography is secure multiparty computation (SMC), in which number of parties wish to compute a function over a set of inputs, where one input is held by each party, and no party wishes to reveal their input to anyone else (Lindell and Pinkas 2009). The original formulation, known as the "millionaire's problem," concerns two wealthy people who wish to determine which one is wealthier without either one revealing her net worth (Yao 1982). The security of an SMC protocol can be defined in terms of a number of different boundary conditions, including whether the parties share trust in a third party, whether an adversary is active or passive, and how many malicious parties must collude in order to violate the privacy of an honest party. A protocol involving k parties is strongly secure only if privacy is still assured even when $k - 1$ parties collude. Many SMC problems and solutions have been developed, but only recently has computing technology made such approaches feasible for practical use cases (Pinkas et al. 2009).

One of the simplest distributed SMC problems is the "secure sum," in which parties wish to compute the sum of their inputs without revealing them (Kantarcioglu 2008). A simple system can be implemented using homomorphic encryption (e.g., Paillier 1999) in which a third party performs the aggregation without being able to learn any of the inputs. The third party can be implemented with a secure coprocessor, which is a piece of hardware manufactured by a trusted entity and operated in isolation to support the protocol (Katz 2007). The use of homomorphic encryption for the computation of sustainability benchmarks was first proposed in Kerschbaum et al. (2011).

Below, we present a formulation of the LCA aggregation problem and discuss it in the context of current research on privacy-preserving computation. We show how an LCA computation can be performed as a secure multiparty computation that preserves the privacy of all inputs.

2 Formulating the aggregation problem

The LCA database computation is commonly presented as a sequence of large matrix multiplications (Heijungs and Suh 2002):

$$s = E \cdot B \cdot A^{-1} \cdot y \tag{1}$$

where the technology matrix A relates products (in rows) to processes (in columns); the environment matrix B relates processes (in columns) to environmental flows (in rows); the characterization matrix E relates environmental flows (in columns) to impact categories (in rows). The input vector y is a column of final demand representing the functional unit of the study, and the result of the multiplication s is a vector of impact scores. Confidential information is often needed to construct the A and B matrices. Computing the matrix product $B_x = B \cdot A^{-1}$ yields an aggregated background database and is considered to be irreversible (as long as the number of environmental flows is much smaller than the number of processes).

Practitioners authoring a study are not required to construct a technology matrix nor perform large matrix multiplication and inversion, instead making use of LCI databases provided in LCA software. The study-specific inventory model, called the foreground, makes reference to processes in the background, but the background does not make reference to the foreground. This allows the inventory model to be written as a block triangular matrix as shown in Fig. 1, where the foreground and background computations can be separated from each other (Kuczenski 2015). In this case, the technology matrix can be broken into three parts: a foreground matrix A_f, a dependency matrix A_d which shows the relationship of the foreground to the background, and an identity matrix to stand in for the background technology matrix, which is included in aggregated form in B_x. The B_x matrix is augmented by B_f to represent direct emissions by foreground processes. In many studies, $B_\mathrm{f} = 0$.

The A_d and B_f matrices make up the private data in the study because they describe which background processes are required by the foreground, how much of each background process is required, and any environmental exchanges associated with the foreground. A publication of results can be regarded as privacy-preserving only if it does not permit an adversary to learn anything about these private matrices.

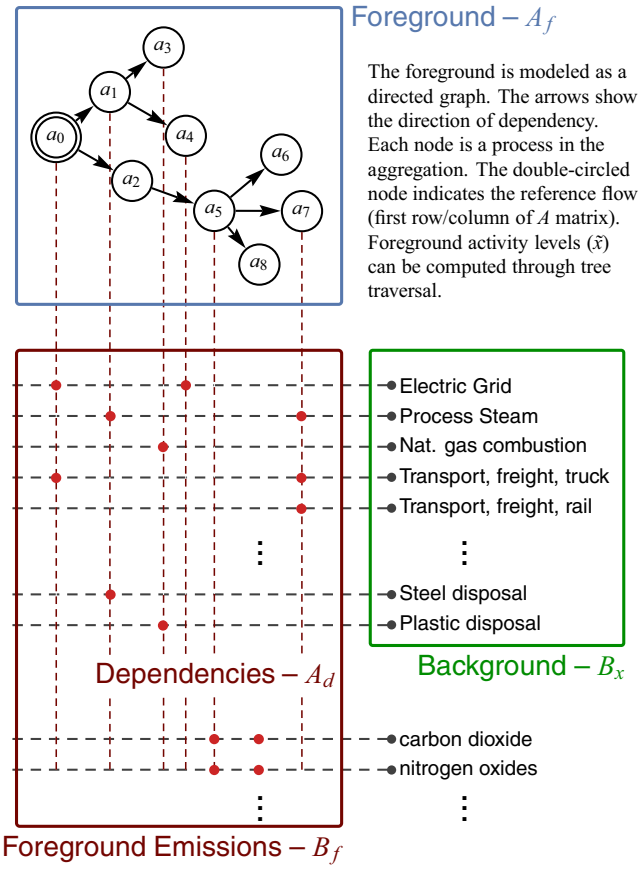
Solving a foreground LCA problem is reduced to inverting the small A_f matrix and using it to determine the activity levels of the foreground processes. The symbol $\tilde{y}$ represents the final demand of the foreground nodes only, and $\tilde{x}$ represents the activity levels of the foreground nodes. Often (if the foreground contains no loops) $\tilde{x}$ can be computed via tree traversal, without requiring matrix inversion. The LCA computation can then be written in terms of $\tilde{x}$:

$$a_\mathrm{d} = A_\mathrm{d} \cdot \tilde{x} \tag{2}$$

$$b = B_\mathrm{f} \cdot \tilde{x} + B_x \cdot a_\mathrm{d} \tag{3}$$

$$s = E \cdot b \tag{4}$$

The results shown above are termed LCA aggregation results. Computing the intermediate result a_d, called a unit



Unit Process aggregation:

$$x = A^{-1} \cdot y$$

$$x = \begin{bmatrix} A_f & 0 \\ A_d & I \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

$$a_d = A_d \cdot \tilde{x}$$

Inventory aggregation:

$$b = \begin{bmatrix} B_f & B_x \end{bmatrix} \cdot x$$

$$b = B_f \cdot \tilde{x} + B_x \cdot a_d$$

($B_x = B \cdot A^{-1}$ is a background LCI database)

Impact aggregation:

$$s = E \cdot b = E \cdot B \cdot x$$

Fig. 1 LCA aggregation results. An LCA foreground study is modeled as a directed graph (*top left*). The *arrows* show the direction of dependency. Each node is a process in the aggregation. The *double-circled node* indicates the reference flow. Foreground activity levels ($\tilde{x}$) can be computed through tree traversal

process aggregation, reports the background activity levels associated with the foreground. Similarly, computing b is called an inventory aggregation and computing s is an impact aggregation. Publishing an aggregation result can be evaluated with regard to whether it preserves the privacy of the input data.

3 A privacy-preserving multiparty LCA aggregation

3.1 Computational model

The foreground study formulation is useful because the LCA aggregations (at least, when $B_f = 0$) can be represented as inner products. In this section, we introduce a simple protocol to perform an LCA aggregation using secure sum operations, which would permit the computations to be performed in a privacy-preserving manner. In our approach, a secure coprocessor is used to generate a homomorphic encryption key and perform decryption, and a separate secure aggregator is used to perform the sum using a homomorphic encryption scheme with an addition operator. Under this scheme, all the participating parties share a common encryption key, but none can decrypt any encrypted values without the help of the coprocessor. The secure sum is accomplished as follows:

1. Each party i prepares its input by encrypting it with the shared key.
2. Each party sends its encrypted input to the aggregator.
3. The aggregator performs the addition on the encrypted data.
4. The aggregator returns the encrypted sum to all parties.
5. Any party is able to ask the coprocessor to decrypt the sum.

The setup is shown in Fig. 2. The result being computed could be at any level of aggregation: process level, inventory, or impact. The parties must agree in advance what metric is being evaluated and how to calculate it. They individually compute their inputs: a market volume or activity level for each entity (x_i), and a representative value of the metric for aggregation (a_i). The market shares could either be agreed upon in advance or selected privately. Parties are assumed to be passive (honest-but-curious) adversaries who do not collude.

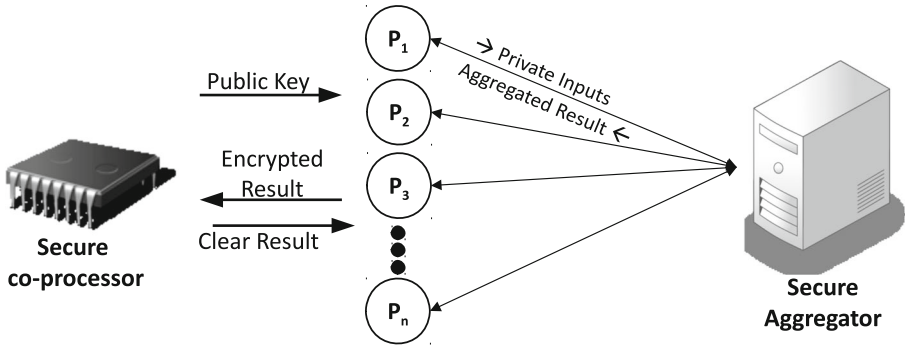


Fig. 2 Schematic diagram of the secure sum protocol. Each party has direct communication with a secure processor, which performs the homomorphic decryption, and a secure aggregator, which computes the sum of encrypted inputs

The foreground matrix for the aggregation is shown below using the convention that positive values are outputs and negative values are inputs:

$$A_{\mathrm{f}} = \begin{bmatrix} \Sigma_1 & 0 & 0 & \cdot & 0 \\ -x_1 & 1 & 0 & \cdot & 0 \\ -x_2 & 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -x_k & 0 & 0 & \cdot & 1 \end{bmatrix} \tag{5}$$

where Σ_1 represents the sum of all market shares.

Then, $\tilde{x} = [\ 1 \quad x_1 \quad x_2 \quad \cdot \quad x_k\ ]^T$ solves for the final demand:

$$A_{\mathrm{f}} \cdot \tilde{x} = [\ \Sigma_1 \quad 0 \quad 0 \quad 0 \quad \cdots\ ]^T \tag{6}$$

The parties then compute two secure sums: to the first, each party sends x_i (which is aggregated to Σ_1), and to the second, each party sends $a_i x_i$ (which is aggregated to Σ_2). The ratio Σ_2/Σ_1 is the market-weighted average of the environmental parameter. Each party can gain insight by comparing their private value to the average. No participant, nor any third party, has gained any private knowledge. This computation is illustrated schematically in Fig. 3.

3.2 Example

Consider the production of acetic acid via the catalytic carbonylation of methanol (Franklin Associates 2007). Acetic acid is a widely produced and inexpensive chemical intermediate, but there exist a variety of methods for producing it. A simplified production model for acetic acid is shown in Fig. 4. The two material inputs are methanol (CH_3OH, often abbreviated MeOH) and carbon monoxide (CO). Both inputs are typically produced from natural gas. The process also requires thermal energy from natural gas combustion, electricity, and the use of a catalyst. In the following example, the numbers are made up.

Facility operators may be interested to compare their efficiency or environmental performance with those of their peers or competitors without revealing details about their own processes. Suppose an environmental analyst had assembled a number of representatives of acetic acid manufacturing plants in order to assess their environmental performance as a group. These representatives are to be the parties to a secure computation. To protect the anonymity of the participants, each facility could be represented by an intermediary.

In order for the computation to take place, the following requirements must be met (see Fig. 4):

a. *Agree on a set of participants* The representatives must decide who among them wishes to participate.
b. *Agree on a production model* The parties must share a common model of the "black box" enclosing their processes. Although each specific plant will have a complex design, the generalized schematic shown in Fig. 4 can be applied to nearly any plant.
c. *Agree on a system boundary* The parties must include and exclude the same elements from their computations. For instance, the parties must agree whether or

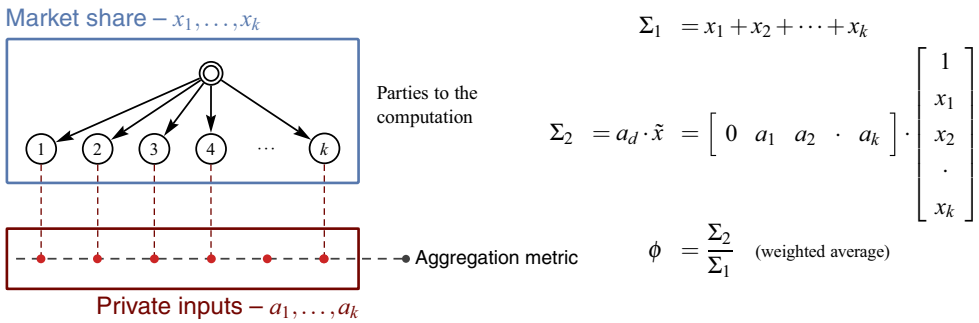


Fig. 3 Privacy-preserving market-weighted average. The foreground model for a horizontal average is a tree of height 1. The sum of activity levels Σ_1 and the sum of activity-level-weighted aggregation values Σ_2 are computed via distinct secure sums

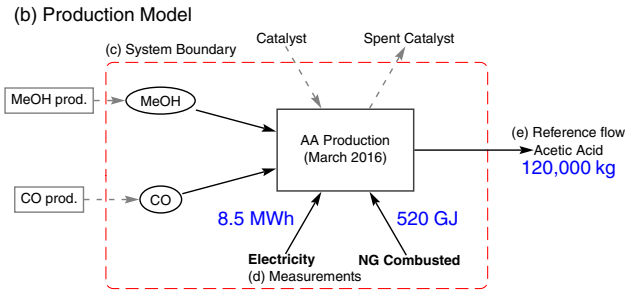


Fig. 4 Schematic representation of a black-box model for production of acetic acid from methanol and carbon monoxide. The figure makes reference to the list of requirements for privacy-preserving LCI computation in Sect. 3.2. One contributor's fictitious private inputs are shown alongside requirements (*d*) and (*e*)

not the production of the intermediates (methanol and carbon monoxide) is included in the boundary. For the schematic in Fig. 4, the production of intermediates is assumed to happen *outside* the boundary.

d. *Agree on measurements* The parties agree to measure electricity and thermal energy use, and to weight their inputs by total production. Catalyst use is excluded from the measurement.
e. *Agree on a reference flow* In order to make the comparison more meaningful, the parties agree to report on their facilities' total output during a common time period, in this example March of 2016.

Following the agreement, each party prepares its private inputs: total production, total electricity use, and total thermal energy use during the time period. In Eqs. (5) and (6), the total production corresponds to x_i, total electricity use corresponds to $a_1 x_i$, and total thermal energy use corresponds to $a_2 x_i$. In Fig. 4, one fictitious party's private inputs are shown in blue.

The parties then compute the secure sums and simultaneously learn the following results (also fictitious):

$$\Sigma_1 = x_1 + x_2 + x_3 \cdots = 5,680,000 \text{ kg} \quad (7)$$

$$\Sigma_2 = a_1 x_1 + a_2 x_2 + a_3 x_3 \cdots = 241 \text{ MWh} \quad (8)$$

$$\Sigma_3 = a_1 x_1 + a_2 x_2 + a_3 x_3 \cdots = 43,721 \text{ GJ} \quad (9)$$

The ratios Σ_2/Σ_1 and Σ_3/Σ_1 report the production-weighted average values for a_1 and a_2, respectively, which in this case are 0.071 kWh electricity and 4.3 MJ of thermal energy per kg of product. Each party can privately compare its performance to the average.

4 Discussion

When an LCA model is formulated as a foreground study, it is possible to give a precise definition to an LCA aggregation as an inner product of activity levels with some process-level environmental metric. We have shown how such a computation can be performed in a privacy-preserving way using a secure sum protocol. At the end of the computation, each party will have learned the average value of the environmental parameter, which can be compared against the private value. Our scheme uses a secure coprocessor implementing a homomorphic addition, but more robustly secure methods could also be applied (e.g., Goryczka et al. 2013). In practice, the demands for more technical security would not be required until the parties involved have developed a more sophisticated or routine usage of the technique presented; in the meantime, non-technical challenges associated with the encryption technology would outweigh the likely benefits to security that came from the enhanced techniques.

4.1 Limitations to SMC

The secure sum approach provides stronger privacy protection than current practice because the parties can compute the results without sharing their data with a common third party in clear text. However, it is also subject to a number of limitations. Foremost, the validity of the computation requires all parties to be honest. While computational methods can ensure that each party follows the protocol correctly, it is impossible to prevent a party from simply reporting a false number unless a third-party audit of the private inputs is permitted. Schemes have been developed that would permit a public audit of inputs while still maintaining privacy (e.g., Baum et al. 2014).

Second, though the parties may not trust each other or a third party with their private data, *technical trust* in the algorithm is still required. This is similar to any other cryptographic application: the user must trust that the software is well made, correct, and free of vulnerabilities. Third, the parties must trust each other to maintain the secrecy of the results: since all parties learn the same result, any one party can reveal it. Fourth, the privacy-preserving characteristics are weaker when parties are allowed to collude. In case of p corrupted parties, the aggregate results of the remaining $k - p$ parties can easily be found if the parties collude.

Finally, secure aggregation via homomorphic encryption requires fixed-point arithmetic. While most foreground aggregations can be performed with limited precision without significant information loss, the fixed-point requirement severely limits the capacity for privacy-preserving life cycle impact assessment (LCIA) owing to the wide dynamic range of emission factors and characterization factors.

4.2 Applications and utility

The model presented in this paper is suitable for performing any computation that can be represented as a set of

weighted sums, which includes foreground aggregation problems such as horizontal averaging and vertical aggregation. These operations are typically associated with the preparation of inventory data for publication, such as preparing industry-averaged models for inclusion in LCI databases. Current techniques for managing this style of data may appear to provide adequate mechanisms for preserving privacy. However, they suffer from the drawback that the aggregation process must be performed by a trusted party. This can constrain participation in the data collection effort and may provide an insurmountable barrier to entry for more secrecy-minded industries. The use of cryptographic techniques, such as the one presented here, may expand the reach of LCA into industries that are not presently well represented in the existing inventory databases, such as many classes of chemicals and pharmaceuticals, detergents, oil refineries, textile and dye producers, and others. Coupled with an auditing mechanism, privacy-preserving techniques may improve the accuracy of inventory results by broadening the base of participation and introducing the capacity to automate and self-manage inventory data gathering.

While the technique presented here focuses on horizontal averaging, the situation is similar for aggregation along a supply chain. The same challenges identified in Sect. 3.2 apply to the vertically aggregated case, but some aspects are much more complex. In particular, the coordination of multiple parties along a supply chain becomes much more difficult when the array of processes being considered is more diverse.

The privacy-preserving approach presented here also targets a new domain of information gathering that is not well addressed by conventional LCA studies—that of cross-firm collaboration for performance improvement. Rather than being motivated by a desire to publish environmental product declarations, firms who are confident in the privacy of their data may pursue benchmarking activities whose intended audience is strictly internal. At present, maintaining and updating a dataset managed by a third party require a renewal of effort on a periodic basis to elicit contributions from participants and a burdensome manual data collection; privacy-preserving techniques would allow interested parties to pursue information independently of a third party (Kerschbaum et al. 2011). Automated benchmarking, enabled by the use of privacy-preserving methods, could reduce the cost and effort required, after initial investments to establish the participant group and the technical infrastructure. This would enable participants to obtain routine and regularly updated benchmarking information to pursue broad-based improvement in environmental performance across an industry group.

The potential benefits would be large in the context of environmental product declarations and regulatory reporting requirements, enabling firms to collaborate to produce mutually privacy-protecting publications of a product. Very large groups of firms could conceivably participate in a mega-scale privacy-preserving computation to estimate national or global normalization values. More research is needed to demonstrate the technique in application.

5 Conclusion

When the scope of an LCA study includes firms that are in competition with one another, it can be challenging to gain much more cumbersome to obtain confidentiality. A single entity must be trusted by all contributors to view private information. The approach presented here may be useful when mutually competitive firms wish to gain private knowledge about their environmental performance by benchmarking against a cohort of similar firms.

The same model can be applied to vertical aggregation problems, such as those involving sequential steps in a product model. Using SMC techniques, firms that are supply chain partners could publish a validated report of their combined environmental footprint without exposing information to one another.

This innovation can increase the scope of participation in multistakeholder LCA projects, such as inventory database development and aggregated inventory publishing, by providing improved protection of confidential information. The aggregation model can also improve transparency in critical review by mechanizing key aspects of model structure, data validation, and computation.

Acknowledgements This work was supported by the National Science Foundation (CCF-1442966). We thank Omer Egecioglu (UCSB) for contributing to the development of this research.

References

Bateman AH, Blanco EE, Sheffi Y (2017) Disclosing and reporting environmental sustainability of supply chains. In: Bouchery Y, Corbett CJ, Fransoo JC, Tan T (eds) Sustainable supply chains: a research-based textbook on operations and strategy. Springer International Publishing, New York. doi:10.1007/978-3-319-29791-0_6

Baum C, Damgård I, Orlandi C (2014) Publicly auditable secure multiparty computation. In: Proceedings of the 9th conference on security and cryptography for networks (SCN 2014). https://eprint.iacr.org/2014/075

Curran MA (1996) Environmental life-cycle assessment. McGraw-Hill Professional Publishing, New York

Dwork C (2006) Differential privacy. In: Bugliesi M, Preneel B, Sassone V, Wegener I (eds) Automata, languages and programming: 33rd international colloquium, ICALP 2006, Venice, Italy, July 10–14, 2006, Proceedings, Part II, pp 1–12. Springer Berlin Heidelberg, Berlin, Heidelberg. doi:10.1007/11787006_1

Finnveden G, Hauschild MZ, Ekvall T, Guinée J, Heijungs R, Hellweg S, Koehler A, Pennington D, Suh S (2009) Recent developments in life cycle assessment. J Environ Manage 91(1):1–21. doi:10.1016/j.jenvman.2009.06.018

Franklin Associates (2007) Cradle-to-gate life cycle inventory of nine plastic resins and two polyurethane precursors. Appendix F. Tech. rep., American Chemistry Council

Frischknecht R (2004) Transparency in LCA-a heretical request? Int J Life Cycle Assess 9(4):211–213. doi:10.1007/BF02978595

Fung BCM, Wang K, Chen R, Yu PS (2010) Privacy-preserving data publishing. CSUR 42(4):1–53. doi:10.1145/1749603.1749605

Goryczka S, Xiong L, Sunderam V (2013) Secure multiparty aggregation with differential privacy. In: Proceedings of the joint EDBT/ICDT 2013 workshops on—EDBT 13. Association for Computing Machinery (ACM). doi:10.1145/2457317.2457343

Heijungs R, Suh S (2002) The computational structure of life cycle assessment, vol 11. Springer, Berlin

Hunsager EA, Bach M, Breuer L (2014) An institutional analysis of EPD programs and a global PCR registry. Int J Life Cycle Assess 19(4):786–795. doi:10.1007/s11367-014-0711-8

Hunt RG, Franklin WE (1996) LCA—How it came about. Int J Life Cycle Assess 1(1):4–7

ISO (2006) ISO 14044. Environmental management—Life cycle assessment—Requirements and guidelines. ISO, Geneva, Switzerland

Kaenzig J, Friot D, Saadé M, Margni M, Jolliet O (2010) Using life cycle approaches to enhance the value of corporate environmental disclosures. Bus Strategy Environ 20(1):38–54. doi:10.1002/bse.667

Kantarcioglu M (2008) A survey of privacy-preserving methods across horizontally partitioned data. In: Privacy-preserving data mining, pp 313–335. Springer Science and Business Media. doi:10.1007/978-0-387-70992-5_13

Katz J (2007) Universally composable multi-party computation using tamper-proof hardware. In: Advances in cryptology-EUROCRYPT 2007, pp 115–128. Springer

Kerschbaum F, Strüker J, Koslowski T (2011) Confidential information-sharing for automated sustainability benchmarks. In: Proceedings of the 32nd international conference on information systems ICIS 2011

Koffler C (2016) Transparency at any cost? LinkedIn Pulse. https://www.linkedin.com/pulse/transparency-any-cost-christoph-koffler. Accessed 14 Nov 2016

Kuczenski B (2015) Partial ordering of life cycle inventory databases. Int J Life Cycle Assess 20(12):1673–1683. doi:10.1007/s11367-015-0972-x

Lindell Y, Pinkas B (2009) Secure multiparty computation for privacy-preserving data mining. J Priv Confid 1(1):5. https://eprint.iacr.org/2008/197

Menezes AJ, Vanstone SA, Oorschot PCV (1996) Handbook of applied cryptography, 1st edn. CRC Press Inc., Boca Raton

Nakano K, Hirao M (2011) Collaborative activity with business partners for improvement of product environmental performance using LCA. J Clean Prod 19(11):1189–1197. doi:10.1016/j.jclepro.2011.03.007

Paillier P (1999) Public-key cryptosystems based on composite degree residuosity classes. In: Stern J (ed) Advances in cryptology—EUROCRYPT'99: international conference on the theory and application of cryptographic techniques Prague, Czech Republic, May 2–6, 1999 Proceedings, pp 223–238. Springer Berlin Heidelberg, Berlin, Heidelberg. doi:10.1007/3-540-48910-X_16

Pinkas B, Schneider T, Smart NP, Williams SC (2009) Secure two-party computation is practical. In: Advances in cryptology—ASIACRYPT 2009, pp 250–267. Springer Science and Business Media. doi:10.1007/978-3-642-10366-7_15

Solér C, Bergström K, Shanahan H (2010) Green supply chains and the missing link between environmental information and practice. Bus Strategy Environ 19(14–15):14–25. doi:10.1002/bse.655

UNEP (2016) Global LCA data access network. http://www.scpclearinghouse.org/working-group/54-global-lca-data-access-network.html. Accessed 17 Oct 2016

UNEP/SETAC (2011) Global guidance principles for life cycle assessment databases. Tech. rep., United Nations Environment Programme

Weidema BP, Bauer C, Hischier R, Mutel C, Nemecek T, Reinhard J, Vadenbo CO, Wernet G (2013) Overview and methodology. Data quality guideline for the ecoinvent database version 3. Tech. rep., The ecoinvent Centre, St. Gallen

World Steel Association (2011) Life cycle assessment methodology report. World Steel Association, Brussels, Belgium

Yao AC (1982) Protocols for secure computations. In: 23rd annual symposium on foundations of computer science (SFCS 1982). Institute of Electrical and Electronics Engineers (IEEE). doi:10.1109/SFCS.1982.38