Social-Aware Federated Learning: Challenges and Opportunities in Collaborative Data Training

Abdul-Rasheed Ottun, *University of Tartu, 51009, Tartu, Estonia*
Pramod C. Mane, *Indian Institute of Management Rohtak, Rohtak, 124010, India*
Zhigang Yin, Souvik Paul, and Mohan Liyanage, *University of Tartu, 51009, Tartu, Estonia*
Jason Pridmore, *Erasmus University Rotterdam, 3062, Rotterdam, The Netherlands*
Aaron Yi Ding, *Delft University of Technology, 2628, CD Delft, The Netherlands*
Rajesh Sharma, *University of Tartu, 51009, Tartu, Estonia*
Petteri Nurmi, *University of Helsinki, 00100, Helsinki, Finland*
Huber Flores, *University of Tartu, 51009, Tartu, Estonia*

Federated learning (FL) is a promising privacy-preserving solution to build powerful AI models. In many FL scenarios, such as healthcare or smart city monitoring, the user's devices may lack the required capabilities to collect suitable data, which limits their contributions to the global model. We contribute social-aware federated learning as a solution to boost the contributions of individuals by allowing outsourcing tasks to social connections. We identify key challenges and opportunities, and establish a research roadmap for the path forward. Through a user study with N = 30 participants, we study collaborative incentives for FL showing that social-aware collaborations can significantly boost the number of contributions to a global model provided that the right incentive structures are in place.

Federated learning (FL) has emerged as a potent mechanism for training powerful AI models in a decentralized and privacy-preserving way.[1] Federated learning is particularly powerful for emerging smartphone and smart device scenarios, such as healthcare or smart cities,[2] as it enables individuals to take advantage of their devices to collect data and to train the model without needing to release data from the devices. Another benefit of federated learning is that it can take the advantage of parallelization and decentralization to minimize the resource demands of individual devices.

Digital Object Identifier 10.1109/MIC.2022.3219263
Date of publication 3 November 2022; date of current version 14 March 2023.

The performance of FL models is intrinsically linked with the availability of training contributions from individuals, which in turn depends on the data they can collect. Indeed, if the data used for training are limited, the final model may suffer from poor generality as it fails to capture the true distribution of the data.[3] In the worst case, the model may even fail to converge if the data are too heterogeneous.[4] Given the heterogeneity of smart device ecosystems, the risk of failing to access sufficient amounts of the right data is significant as the devices may lack the right capabilities or may produce substandard contributions due to device limitations. In addition, contributors may attempt to act as free riders by refusing to spend resources on training the model. Instead, they try to benefit solely from the contributions of other users.[5] While many federated learning models demand a sufficient level of participation for training, even this approach is not sufficient as the free riders may simply send random parameter updates, which can actually harm the overall model.[6] Ensuring sufficient quality for the federated learning model and overcoming these limitations requires new ways to boost the contributions of individuals, while preserving the quality of the data. At the same time, the anonymous nature of the contributions can give malicious entities an opportunity to hamper the global AI model. Thus, an additional layer of social trust can improve the resilience of the system and help to overcome misuse.

The present article contributes social-aware federated learning, the use of social connections to boost the training contributions in FL. These connections can either be known people, which implies a trusted relationship with the contributing person, or opportunistic contacts that are within the range of device-to-device connections.[7] Social-aware FL can simultaneously prevent situations where a client seeks to obtain benefits without contributing (i.e., free-rider problem), introduce trust in training process through social connections,[8] and mitigate the risk of malicious contributions. Our work harnesses social connections for fostering participation in the learning process, whereas previous works have been limited to using social connections to improve security and to avoid malicious users by determining with whom to share model updates or who to use as the aggregator.[9] In terms of improving the rate of FL contributions, the main alternatives for social-aware FL are to crowdsource the contributions and to rely on a centralized authority to coordinate model updates or to offer incentives to motivate individuals to contribute.[10,11] We reflect on the state-of-the-art to identify key challenges and open issues, provide ways to overcome these challenges, and establish a research roadmap with the aim of acting as a catalyst for further research.

To understand the potential and the limitations of social-aware FL, we conduct an experiment with $N = 30$ participants to investigate the willingness of users to outsource tasks to other users. The results of our study show that individuals are interested in delegating tasks to others, and that the users are willing to execute the tasks for other users, provided that suitable incentive mechanisms are in place. Our work paves the way toward innovative social mechanisms for boosting contributions to training FL models and enables users to benefit from the FL model even when they lack the necessary capability to contribute to the model training themselves.

CHALLENGES AND OPPORTUNITIES

Collaborative compensation: Social links alone are not sufficient for supporting FL as people are prone to churn, i.e., their willingness to contribute wanes over time. Incentives are a potential way to overcome—or at least mitigate this issue.[10,11] Incentives for FL need to account for the complexity of the contributions as they affect the overall ecosystem in improving the global model instead of benefiting the initiator directly. The incentives should take these roles into consideration and potentially compensate both the person executing the task and the person serving as intermediary. At the same time, the compensation may be drawn from other users of the FL system as they all potentially benefit from the contribution to the model.

Data poisoning: Robust training of FL models requires multiple individuals to contribute aggregated data. This, however can be exploited by malicious actors that exploit the system or compromise it through other forms of misuse. For example, so-called data poisoning can be used to hamper the AI inference process. Despite the several methods for detecting data poisoning[12] or other attacks, enforcing them with each model update is difficult. Incorporating an additional layer of trust based on social connections can reduce the possibility of aggregating poisoned updates to the global model. While social links are expected to increase the level of trust in the data providers, social links can also become a source of attacks when digital identities are stolen. For example, smartphones can have exploits (e.g., malware) without the users noticing them. These vulnerabilities can be used to poison the data that contributes to the global model or even the model itself. Overcoming this issue requires solutions that analyze the influence of individual contributions to the global model. For malicious actors, reputation mechanisms can offer a way to disqualify users that poison the data, e.g., by offering a way to rank the users based on the quality of their contributions.

Training moments: Ensuring high accuracy for an FL model requires multiple training rounds—at least until the model starts to converge. The processing time that is needed for these rounds can be significant and hamper the normal functionality of the device. As the key benefit of FL is avoiding data disclosure, this process cannot even be offloaded. Ensuring the training does not hamper the user's everyday activities requires a mechanism that allows suspending and later resuming the training on the individual devices. Alternatively, methods that quantify the duration of time in which users can dedicate time and resources for a training task can be adopted, e.g., it is possible to quantify and predict the stability of a user's stay at a given location.[7] Probing times can also be considered when outsourcing

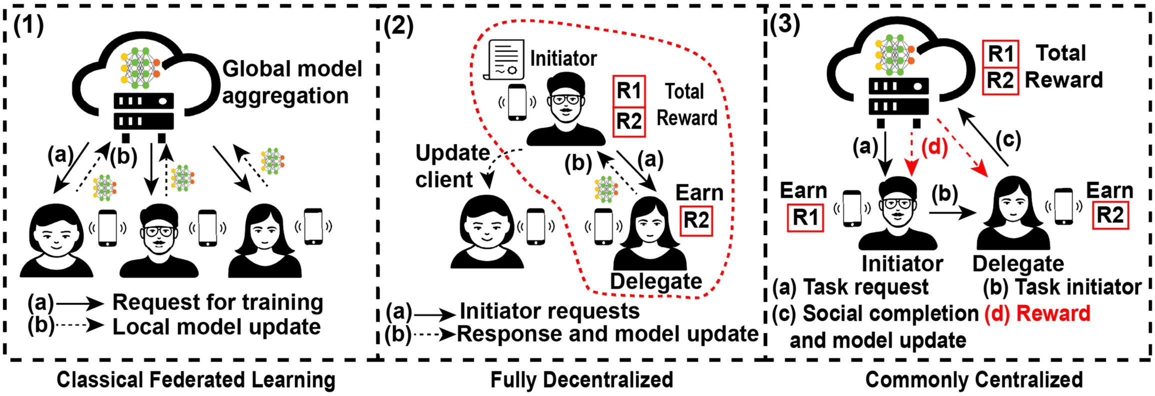

FIGURE 1. Design alternatives to extend federated learning with social-aware capabilities. (1) Classical federated learning. (2) Social-aware over a decentralized FL architecture. (3) Social-aware over a centralized FL architecture.

an FL task to social connections to guarantee that tasks are not rejected by the delegate.[13]

Recurring issues: There are also challenges that recur in all kinds of social-aware systems.[4] For example, besides participation and engagement, the distribution opportunities are governed by social contacts that are limited. This may require solutions where further distribution is allowed, i.e., the social contact further propagates the request to one of his/her own contacts until the task can be executed. There are also recurring issues that affect the implementation of the FL model itself. For example, data can be highly heterogeneous, relying on sensors that only some devices have, or on multimodal data that requires contributions from different sources. This requires the underlying FL model to be generic so that it can aggregate the different types of data. The data may also contain dependencies and come from different distributions (i.e., the data are non-I.I.D.), which requires separate mechanisms, such as using data augmentation or local tuning, to ensure convergence.[14] Another recurring issue is privacy. Even if FL itself has been designed to be privacy preserving, there are attacks that can violate the user's privacy (e.g., data or model inversion). At the same time, user's may not be fully aware of the privacy protection they are offered—especially as they may not be aware of what happens to the data once it has been collected. Improving the user's level of trust requires methods that foster transparency, e.g., by offering explanations of the inner functions of the FL pipeline and provide insights into the processing that is happening on the background.[15] Trustworthiness of FL models is difficult to achieve as a successful attack can be easily propagated to all the involved devices.[12] This then requires instrumenting each client and having them troubleshoot the model. Thus, further mechanisms and architectures that prevent attacks are also required.

SOCIAL-AWARE FEDERATED LEARNING

Model and Assumptions: We consider a federated learning scenario where people with smartphones or other IoT devices collect data and use that to train a local model, which is then shared with an aggregator in exchange for a global model that can be used to improve services on the local device. We assume that the FL application can access the social contacts of the participant's friends or other social contacts that have the same application installed. Each participant is expected to contribute a certain number of updates to the model and a separate coordinator is responsible for requesting these updates. The coordinator is typically provided by a centralized authority but it is also possible to choose the coordinator in a decentralized way using social voting. Each time the device sends valid updates, it receives a compensation from an incentive mechanism used by the FL algorithm. Social-aware FL extends the basic FL by allowing the device to delegate the request to one of their social contacts. In this case, the device serves as initiator and the social contact as a delegate. When tasks are delegated, the initiator is assumed to share all or part of the compensation they receive from the incentive mechanism with the delegate. Whenever a new user downloads the application, they are shared the current model parameters to ensure their training contributions are most useful for the current model.

Implementations: As shown in Figure 1, there are different ways to implement social-aware federated learning in practice and the specifics depend on the overall implementation of the federated learning system. This also determines which information needs to be exchanged between devices. In a fully distributed case, the initiator needs to share their local model with the delegate who then needs to load the model into memory and update it before sending the model parameters back to the initiator, which then sends them to the aggregator. Depending on the nature of the contracts, this type of system may require a separate reputation system to ensure the delegate gets compensated by the initiator. Alternatively, the initiator and delegate can establish a contract that is verified by the device(s) being responsible for aggregation, and the compensation can be then handled in line with this contract. In case a centralized aggregator exists, as is common in federated learning scenarios, the compensation scheme would be coordinated through the coordinator. The initiator can then either inform the delegate of the task descriptor and its own id or send the full model parameters to the delegate as in the distributed case. The delegate can then send the updated parameters directly to the aggregator and, if needed, share with the initiator.

Communications: The communication between the devices depends on the characteristics of the federated learning task. Most FL tasks correspond to horizontal FL where each device shares the same feature space but has access to different samples. In this case the model is trivial to share as all devices have exactly the same structure. In many smart device scenarios, including our experiments, devices have access to different sensors that need to be integrated to learn a common model. Thus, the feature space of the devices is different. This is known as vertical FL, which typically requires different architecture.[2] One possibility is to rely on a hierarchical model where each sensor (type) has separate convolutional structure, and a secondary convolutional structure maps the contributions of individual sensors into a unified format.[16] Finally, regardless of the implementation, the communications between the initiator and delegate should naturally be secured. This can be accomplished using a secure association mechanism to establish the communication channel and to encrypt the communications that take place. Social-aware FL assumes prior trust relationship between the delegate and initiator, but the security of the mechanism can be further improved by integrating a mechanism on the initiator to detect possible malicious updates, e.g., by examining prediction performance before and after the update.[17]

EXPERIMENTAL SETUP

We study the potential of social collaboration in federated learning by conducting an experiment that evaluates the user's perceptions of collaboration, their willingness to hand out or to execute tasks, and the valuations that the users place on different tasks. We focus on tasks that involve training on diverse sensor measurements as sensor data are the important source of data for emerging AI models and as the data they provide are highly heterogeneous.

Experimental Design and Methodology: Our study consists of two parts. The first part uses a Vickrey auction[18] and the second part is designed following a between-subjects methodology with two conditions: detached and attached. In the detached condition, the initiator hands out tasks to another user who is then given a fixed compensation (1€). The attached condition is otherwise the same but the initiator can freely choose to keep a fraction of the compensation, and the rest is distributed to the person carrying out the task. In both conditions, the initiator is always the same and the only difference is how the compensation from the incentive mechanism is shared with the social contact acting as a delegate. The monetary compensation is given once the execution of the task is completed. Tasks that are handed out to others can be accepted or declined by the receiving party. If the task is rejected, it goes back to the initiator so that it can be handed out to others. Tasks that are accepted cannot be handed out again to avoid recurring tasks.

Application prototype: We implemented a mobile application for the study that allows the user to perform tasks or to distribute them to other users. The app is designed as a web application, implemented using the Adalo platform, and can be executed on any smartphone. The app uses notifications and alarms to make the participants aware about tasks received from other users. The app also integrates a database that captures the interactions of users with the tasks, e.g., rejecting a task, accepting a task, and task execution time.

Apparatus and task: We used three different smartphone models. Each smartphone is assigned to a specific task depending on the sensors it has: iPhone (HDR camera), Caterpillar CAT S61 (thermal camera), and Redmi Note 8 (air quality). We consider both generic tasks that can be performed on any device and specific tasks that can only be performed on devices having the appropriate sensor or other instrumentation. As generic task we consider GPS location, and as specific tasks we consider high-resolution imaging, thermal imaging, and air quality monitoring, in line with the devices considered in the experiment. All tasks are listed in each smartphone, and participants use the application to complete them individually or by handing out the task to another participant. The

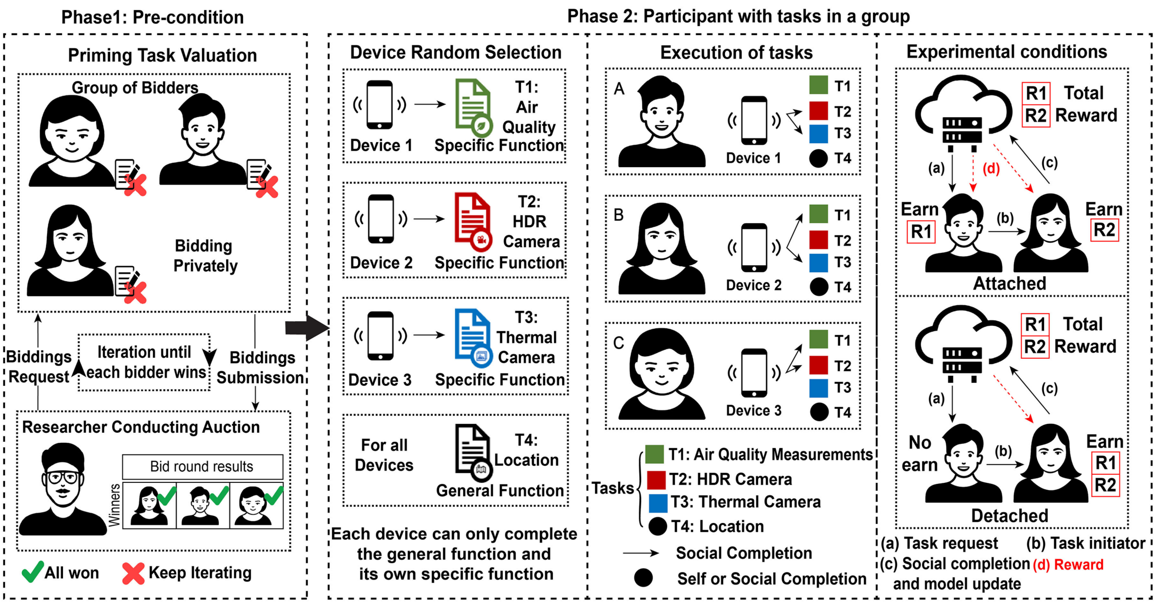

FIGURE 2. Overview of the experimental procedure (Phases 1 and 2).

nature of the tasks effectively corresponds to a vertical federated learning scenario[2] as the devices do not necessarily share the same feature space. We chose this type of scenario as it is representative of FL scenarios for smart devices and as it is a scenario that benefits from collaboration.

Participants: We recruited 30 participants for the experiment. The participants were divided into groups of three to test social-aware collaboration. The participants are of different nationalities and were recruited through mailing lists and social media posts. As the study was designed as a between-subjects study, this means that 15 users (five groups of three) were allocated to both experimental conditions. To ensure the experiment involved social connections, we mentioned that the experiment was a group experiment that required three participants who share relationships to jointly test a social application and the participants were encouraged to bring along another person they know who they will pair with for the testing. Most participants came with friends, acquaintances, colleagues, or flatmates. When participants were unknown to each other, trust was ensured first by not using the personal device of the individual but rather one provided by the researcher, and second, by stating early during the experiment the reward obtained by participating, which has been used in other studies to encourage engagement in collaborative activities.[19]

Procedure: The study procedure is summarized in Figure 2 and is split into two parts. Prior to start the study, the recruited participants are assigned into a group. Each group is allocated to either one of the two experimental conditions (detached or attached) following a counterbalanced design. In the first part of the study (Phase 1), participants are preconditioned (or primed) to calibrate their valuations of different sensor data. This is done to ensure that people have reasonable understanding of the costs and valuations associated with different data, which forms their basis for deciding how to split the payments in the detached condition. The priming was achieved using a second-price Vickrey auction. We relied on this type of auction as it captures the most realistic perceptions of valuations from users over time.[18] After participants understood the auction type and signed a consent form, they were presented with a list of 20 distinct sensing tasks. Each sensing task focused on a different sensor to allow people to understand the potential task and to establish a valuation for tasks with diverse requirements. The sensors that were considered in the task were: camera, Bluetooth, microphone, GPS, humidity sensor, thermal camera, temperature sensor, touch sensor, and WiFi. Bids were elicited using questions that linked the sensor with a specific application. As an example, in one task the participants were asked *"how much would you take to perform a task requiring the use of your phone's microphone to record a five-second sound clip to measure noise level in your current room."* Participants then wrote their bids privately on a piece of paper. A researcher collected the bids and announced the winner. The first phase of the experiment concluded once every participant in the group had won the auction at least once.

The second part (Phase 2) introduces the participants to the mobile application. The researcher responsible for conducting the study explained the mobile application's functionality to the participants, who were given time to familiarize themselves with the application. Next, the three smartphones were distributed among the people in a group. Participants were then presented with tasks they needed to perform, and they were given 5 minutes to perform the them. We presented four tasks, one for each of the four sensors (HDR camera, thermal camera, air quality sensor, and GPS), to the participants and ask them to complete the task or to distribute it to another user. Given the differences in functionality and the design of the experiment, participants were able to carry out two tasks by themselves (GPS and the one task for which they had the right sensor on their device) and the two other tasks always required delegating the task to others. Once the experiment is finished, the smartphone was collected back and participants were compensated with a monetary reward for performing the tasks. To get the compensation, we stated at the beginning of the experiment that at least three tasks should be completed. If the participants did not finish the tasks, no compensation was given. The overall experiment lasted around 30 minutes. In addition, a tea/coffee mug was given to each participant at the end of the experiment.

RESULTS

Results of Priming Experiment

In the priming phase, in total 100 bidding rounds were executed and 297 bids were received. Figure 3(a) shows the distribution of the bid values. We applied a multivariate outlier detection (based on Mahalanobis distance) to remove bid values at both extremes (i.e., among the smallest and largest). Figure 3(a)-1 shows the overall distribution of the bids and 3(a)-2 shows the resulting distribution after the outliers are removed. The results show that initially many users place higher bids, but rapidly recalibrated and started to accept lower valued bids as they were exposed to other user's bids.

We also separately assessed how privacy considerations factor into the user's valuations. Previous studies have shown that the privacy implications of sensors affect user's perceptions[20] and thus we would expect to see these also reflected in the bids. However, also other factors have been shown to affect users, e.g., resource consumption is an important determinant. To isolate the effects of privacy, we chose three sensors that have similarly high resource consumption but different privacy implications: GPS (Personal), Camera (Public), and Microphone (Social). Figure 3(b) shows the overall distribution of the task valuations for these sensors. The valuations reflect differences in privacy implications, which can be observed both from the mean valuations and the variance of the bids. The average valuations are 10.00€ for GPS, 5.00€ for camera, and 5.00€ for microphone. As the relative ordering reflects the differences in privacy, the results of the bids are in line with previous findings, suggesting that the valuations resulting from the priming experiment are realistic.

Results From Application Use

Earnings in the two conditions: We first assess the earnings of the participants across the two conditions: attached and detached. As the attached condition resulted in the payment being split between the initiator and the task executor, we would expect the detached condition to result in higher overall payment. The results also confirm this but only show a marginal difference. Specifically, Figure 3(c) shows that the difference is merely 0.60€ between the two conditions (detached: mean $=$ 3.00€, SD $=$ 1.18, attached: mean $=$ 2.40€, SD $=$ 1.21). A Mann–Whitney U-test confirmed that no significance was found between the conditions $(U = 114, p > 0.05)$. This result thus shows that the experimental design worked as intended and that the splitting the payments resulted in a marginal loss of compensation. The payment differences were dependent on the role that the person was acting as. We also compared the similarity of the payment distributions across the two conditions and roles (the values correspond to the test statistic of the Anderson–Darling test, which measures similarity of distributions). When the user acts as initiator, the payments are higher for self completion, suggesting that users are willing to carry out the tasks themselves—at least in exchange of compensation. No differences are found across the two conditions, suggesting that whether people can keep parts of the other user's compensation or otherwise does not affect the user's willingness to carry out the task themselves. In contrast, when the user acts as task executor, the distributions of payments depend heavily on the condition with the user's more likely to take the advantage of social collaboration whenever they can keep some of the payment (detached: 13.10, attached: 4.52). Overall, the results thus show that the compensation mechanism has desired impact on the payment structure, also that payment structure affects how willing people are to take the advantage of social collaboration with the best results obtained

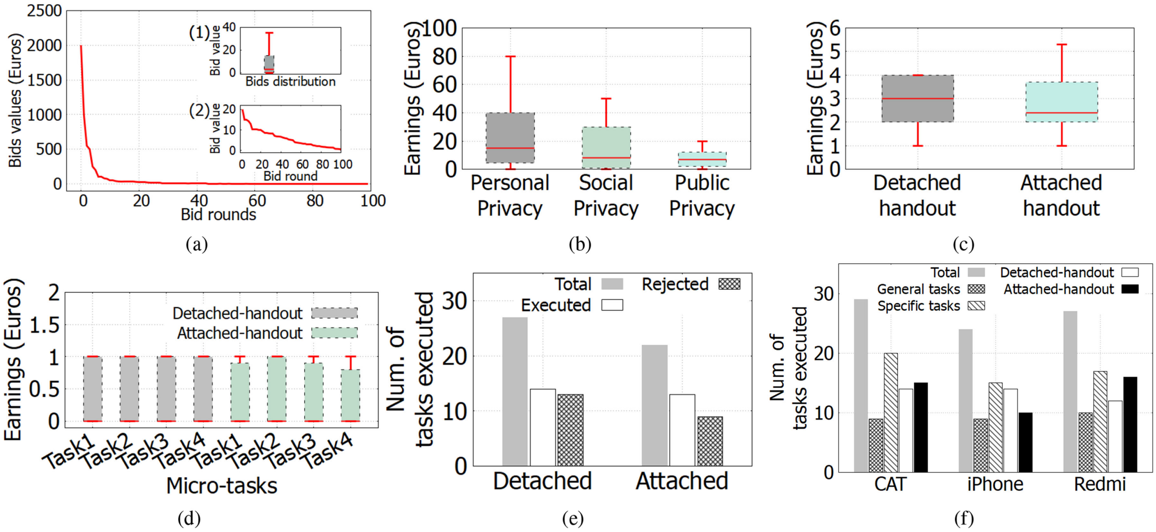

FIGURE 3. [a–b] Priming results of Phase 1. (a) Auction priming and bidding performed by participants. (b) Quantifiable valuate of tasks based on sensor type and privacy data considerations. (c–f) Results of hand out tasks using our prototype application in Phase 2. (c) Distribution of earnings in both conditions. (d) Earnings obtained per each task in the experiment, both conditions. (e) Dissected actions of outsourced tasks. (f) Influence of device usage when performing tasks.

when both the task initiator and task executor can be compensated for their effort.

Task completion: Figure 3(d) shows the earnings per task for the two conditions. The total earnings in the attached condition are smaller for most tasks, also for generic tasks that could be executed on any device, suggesting that the compensation structure also had some influence on user's willingness to execute tasks. Figure 3(e), in turn, compares the total number of tasks that were outsourced and the number of tasks that were accepted for execution or rejected by the person receiving them. The detached condition resulted in higher number of tasks being outsourced. Of these tasks, roughly the same number were accepted as in the attached condition, indicating that task executors were more likely to reject the task when they only receive partial payment. The results thus show, on one hand, that giving a compensation to the initiator is necessary to ensure as many tasks as possible are outsourced. On the other hand, the results show that the payment structure has to be carefully designed to motivate those receiving the tasks to accept and execute the tasks also.

Device usage: Next, we analyze whether the device type influences the outsourcing of tasks in the two conditions. Figure 3(f) shows that for the generic tasks not much differences can be observed. In contrast, for the specific tasks a clear difference can be observed depending on the task, device, and condition. On the CAT S61 and the Redmi smartphones, users were willing to execute more tasks than on the iPhone. This can potentially be explained by privacy considerations (see the following) and the differences in tasks completions also support this view. Specifically, in line with the valuations in the priming experiment, tasks with higher impact on personal privacy were less likely to be executed (i.e., the generic GPS task) than tasks involving social or public privacy sphere. The differences may also result from perceptions of the devices and the data. For example, thermal images often appear less privacy intrusive than HDR images. As for the conditions, for the thermal imaging task (i.e., CAT S61) no differences could be observed. For the air quality task (i.e., Redmi), a higher number of tasks were executed in the attached than in the detached condition, whereas for the HDR imaging the result was reversed. The differences in the HDR imaging task mirror the differences in the number of tasks that were outsourced in the two conditions and thus the difference is likely simply a result of higher number of tasks being possible to execute. In contrast, the result for the air quality task suggests that people were likely more prone to rejecting air quality monitoring tasks that were outsourced with partial compensation.

DISCUSSION

Stakeholders and adoption: Our experiments demonstrated that social mechanism can improve acceptance rate from the user's perceptive. Our method can be generalized not just to FL contexts, e.g., it can also be used to improve data collection in crowdsensing and crowdsourcing platforms. Our social mechanism to divide the compensation to execute a task (from finding the best suitor to task execution) can supplement existing platforms and provide new opportunities to augment the scope of data collection.

Room for improvement: We demonstrated how social connections can be harnessed to increase the rate of contributions in federated learning. This is particularly useful in scenarios where users lack a specific sensor or where they temporarily run low of resources. Our core focus was on exploring user perceptions, further work is needed to quantify the effects on overall performance of the FL models. This would require running the experiment for several rounds as the performance of federated learning depends on the number of samples that are provided and their capability to represent the overall distribution of data. When multiple clients provide data, a smaller amount of rounds is usually sufficient as the inclusion of data from different clients improves coverage of the overall data distribution. We are also interested in performing performance analysis focused on extracting system-oriented metrics, such as the amount of data transfer, training payload, training time per device, and the number of rounds for model convergence.

Autonomous social agents: Social relations can be exploited by agents to automatize the outsourcing of a task to social connections. Digital agents assigning tasks on behalf of a user can be useful to perform optimal decisions for assigning tasks to social connections. It is also possible that the execution of a task can be performed solely by agents interacting through social connections. For instance, one agent can request another agent to measure the air quality of a room while the users are unaware of this interaction. Naturally, this can also open back doors to possible cyberattacks if the digital agent of a user is compromised. To mitigate this problem, blockchain and smart contract solutions can be adopted.

Potential applications: Besides the large number of applications that can rely on FL support,[2] it is also possible to envision new use cases and applications that require the social intervention of users to improve models. For instance, to speed up the convergence of models to accurately detect new illnesses, e.g., COVID-19, digital contact tracing applications can benefit from obtaining representative samples from infected individuals through social connections.

SUMMARY AND CONCLUSIONS

We proposed extending federated learning with social collaborative features to boost training contributions, particularly in situations where contributors lack the means to contribute themselves (e.g., lack of required sensor). We studied the potential and the limitations of social-aware collaboration through a user study with $N = 30$ participants and two conditions (attached and detached) and showed that individuals are indeed willing to contribute, but that the degree of contributions depends on the way the incentives to contribute are structured. Giving the people outsourcing the task a compensation results in the highest number of task requests, but can decrease the number of tasks that are accepted by those receiving the tasks—unless the incentives account for this. Our work paves the way toward integrating social collaboration into federated learning and offers insights into the design of effective compensation mechanisms for boosting the contributions to federated learning applications.

ACKNOWLEDGMENTS

This research was part of the SPATIAL project that has received funding from the European Union's Horizon 2020 research and innovation programme under Grant 101021808. European Social Fund has also financed this research via "ICT programme" measure, CHIST-ERA Grant CHIST-ERA-19-XAI-010, and the Academy of Finland under Projects 317875 and 339614.

REFERENCES

1. Q. Li et al., "A survey on federated learning systems: Vision, hype and reality for data privacy and protection," *IEEE Trans. Knowl. Data Eng.*, to be published, doi: 10.1109/TKDE.2021.3124599.
2. Q. Yang et al., "Federated machine learning: Concept and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, 2019.
3. T. Pulkkinen, J. K. Nurminen, and P. Nurmi, "Understanding WiFi cross-technology interference detection in the real world," in *Proc. IEEE 40th Int. Conf. Distrib. Comput. Syst.*, 2020, pp. 954–964.
4. Y. Ruan et al., "Towards flexible device participation in federated learning," in *Proc. Int. Conf. Artif. Intell. Statist.*, 2021, pp. 3403–3411.
5. Y. Fraboni et al., "Free-rider attacks on model aggregation in federated learning," in *Proc. Int. Conf. Artif. Intell. Statist.*, 2021, pp. 1846–1854.
6. L. Lyu, X. Xu, Q. Wang, and H. Yu, "Collaborative fairness in federated learning," in *Federated Learning*. Berlin, Germany: Springer, 2020, pp. 189–204.
7. H. Flores et al., "Collaboration stability: Quantifying the success and failure of opportunistic collaboration," *IEEE Comput. Mag.*, vol. 55, no. 8, pp. 70–81, Aug. 2022.

8. X. Gong et al., "Incentivizing truthful data quality for quality-aware mobile data crowdsourcing," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2018, pp. 161–170.
9. L. U. Khan, Z. Han, D. Niyato, and C. S. Hong, "Socially-aware-clustering-enabled federated learning for edge networks," *IEEE Trans. Netw. Serv. Manage.*, vol. 18, no. 3, pp. 2641–2658, Sep. 2021.
10. H. Yu et al., "A fairness-aware incentive scheme for federated learning," in *Proc. AAAI/ACM Conf. AI, Ethics, Soc.*, 2020, pp. 393–399.
11. L. U. Khan et al., "Federated learning for edge networks: Resource optimization and incentive mechanism," *IEEE Commun. Mag.*, vol. 58, no. 10, pp. 88–93, Oct. 2020.
12. V. Tolpegin et al., "Data poisoning attacks against federated learning systems," in *Proc. Eur. Symp. Res. Comput. Secur.*, 2020, pp. 480–501.
13. P. Kairouz et al., "Advances and open problems in federated learning," *Found. Trends Mach. Learn.*, vol. 14, no. 1/2, pp. 1–210, 2021.
14. H. Zhu et al., "Federated learning on non-iid data: A survey," *Neurocomputing*, vol. 465, pp. 371–390, 2021.
15. P. P. Angelov et al., "Explainable artificial intelligence: An analytical review," *Data Mining Knowl. Discov.*, vol. 11, no. 5, 2021, Art. no. e1424.
16. S. Yao et al., "DeepSense: A unified deep learning framework for timeseries mobile sensing data processing," in *Proc. 26th Int. Conf. World Wide Web*, 2017, pp. 351–360.
17. J. Chen, X. Zhang, R. Zhang, C. Wang, and L. Liu, "De-Pois: An attack-agnostic defense against data poisoning attacks," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 3412–3425, 2021.
18. S. Parsons et al., "Auctions and bidding: A guide for computer scientists," *ACM Comput. Surv.*, vol. 43, no. 2, pp. 1–59, 2011.
19. I. Bohnet, "Trust in experiments," in *Behavioural and Experimental Economics*. Berlin, Germany: Springer, 2010, pp. 253–257.
20. S. Hosio et al., "Monetary assessment of battery life on smartphones," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2016, pp. 1869–1880.

ABDUL-RASHEED OTTUN is a doctoral candidate at the University of Tartu, 51009, Tartu, Estonia. His research interests include autonomous vehicles, federated learning, and explainable AI. Contact him at rasheed.ottun@ut.ee.

PRAMOD C. MANE is an assistant professor at the Indian Institute of Management Rohtak, Rohtak, 124010, India. His research interests include socially aware computing, game theory, and social networks. Contact him at pramod.mane@iimrohtak.ac.in.

ZHIGANG YIN is a doctoral candidate at the University of Tartu, 51009, Tartu, Estonia. His research interests include light sensing, autonomous vehicles, and pervasive systems. Contact him at zhigang.yin@ut.ee.

SOUVIK PAUL is a doctoral candidate at the University of Tartu, 51009, Tartu, Estonia. His research interests include federated learning, data poisoning, and device-to-device communications. Contact him at souvik.paul@ut.ee.

MOHAN LIYANAGE is a lecturer at the University of Tartu, 51009, Tartu, Estonia. His research interests include Internet of Things, edge computing, and networking. Contact him at mohan.liyanage@ut.ee.

JASON PRIDMORE is an associate professor at Erasmus University Rotterdam, 3062, Rotterdam, The Netherlands. His research interests include communications, AI, and digital services. Contact him at pridmore@eshcc.eur.nl.

AARON YI DING is an associate professor at the Delft University of Technology, 2628, CD Delft, The Netherlands. His research interests include edge AI, edge systems, and IoT networking services. He is a Member of the IEEE. Contact him at aaron.ding@tudelft.nl.

RAJESH SHARMA is an associate professor at the University of Tartu, 51009, Tartu, Estonia. His research interests include social networks, device-to-device, and AI. Contact him at rajesh.sharma@ut.ee.

PETTERI NURMI is an associate professor at the University of Helsinki, 00100, Helsinki, Finland. His research interests include distributed systems, pervasive data science, and sensing systems. Contact him at petteri.nurmi@helsinki.fi.

HUBER FLORES is an associate professor at the University of Tartu, 51009, Tartu, Estonia. His research interests include distributed, mobile, and pervasive computing systems. Contact him at huber.flores@ut.ee.