Efficient Gossip Algorithms for Machine Learning

PhD Thesis

Gábor Danner
Supervisor: Márk Jelasity, PhD

Doctoral School of Computer Science
Department of Computer Algorithms and Artificial Intelligence
Faculty of Science and Informatics
University of Szeged

Szeged
2022

Introduction

In recent times, smart devices have become part of our daily lives. Their widespread presence offers numerous applications that make use of data mining. Usually, the data is collected at a central location for processing. However, this is becoming increasingly problematic due to the growing privacy concerns of the public and the policy-makers. This has motivated the need for collaborative methods that do not need the collection of sensitive data.

Google proposed federated learning to address this problem. Although this is still a centralized approach, the data remains on the device. It works in a similar way to the parameter server architecture. The server periodically sends the current model to the nodes. They perform an update step using the local data and upload it to the server for aggregation. Compression techniques can be employed to reduce communication costs.

In contrast, we proposed gossip learning, which is fully distributed. The nodes communicate directly, exchanging their models. The lack of need for a central server makes this an attractive approach for startups and communities with limited resources. It can assist the creation of non-profit intelligent smartphone services. It is suitable for other platforms as well, like smart metering and the Internet of Things.

To improve upon existing gossip learning methods, this PhD thesis presents a number of techniques that also have applications beyond this. Our research touches on many topics including secure multiparty computation, communication flow control in decentralized systems, and efficient average consensus algorithms using stateful encoder-decoder pairs.

Summary

The dissertation consists of 4 main parts, summarized below.

In Thesis 1, we focus on privacy and security issues. We propose a light-weight protocol to quickly and securely compute the sum query over a subset of participants assuming a semi-honest adversary. We apply this protocol to efficiently calculate the sum of gradients as part of a fully distributed minibatch stochastic gradient descent algorithm.

In Thesis 2, we present a systematic comparison of gossip learning and federated learning. We examine the aggregated cost for several algorithm-variants in various simulation scenarios. These experimental scenarios include different network sizes and different distributions of the training data over the devices. We also use a real smartphone trace.

In Thesis 3, we propose a family of adaptive flow control protocols that apply rate limiting inspired by the token bucket algorithm, but they also include proactive communication to prevent starvation. With the help of our traffic shaping service, some decentralized applications approach the speed of the reactive implementation, while maintaining strong guarantees regarding the total communication cost and burstiness. We perform simulation experiments in different scenarios, with a focus on machine learning applications.

In Thesis 4, we propose a communication efficient and robust algorithm for decentralized mean estimation, and adapt it to machine learning. In addition, we also rely on transfer learning for extra compression. We demonstrate these contributions through an experimental analysis.

Table 1: *Correspondence between the thesis points and the publications of the author. • and ○ indicate the core and the related publications, respectively.*

	Thesis 1	Thesis 2	Thesis 3	Thesis 4
DAIS 2015 [5]	•			
SCN 2018 [3]	•			
DAIS 2019 [10]		•		
JPDC 2021 [11]		•	•	
ECML 2019 [9]		○		
ICDCS 2018 [7]			•	
Euro-Par 2018 [6]				•
DICG 2020 [4]				•

Thesis 1: Gossip Learning with Privacy Preservation

Data mining over personal data harvested from mobile devices is a very sensitive problem due to the strong requirements of privacy preservation and security. We propose a solution that does not utilize centralized resources at all, preserving privacy by avoiding the central collection of any personal data, even in pre-processed form.

In gossip learning models perform random walks over the network and they are trained using stochastic gradient descent [12]. This involves an update step in which nodes use their local data to improve each model they receive, and then forward the updated model along the next step of the random walk. However, this method is susceptible to collusion. If the nodes before and after another node in the random walk collude they can recover private data using the two versions of the model right before and right after the local update step.

In Chapter 3 of the dissertation we address this problem, and improve gossip learning so that it can tolerate a much higher proportion of honest but curious (or semi-honest) adversaries. The key idea behind the approach is that in each step of the random walk we form a group of peers that securely compute the sum of their gradients, and the model update step is performed using this aggregated gradient. In machine learning this is called mini-batch learning, which—apart from increasing the resistance to collusion—is known to often speed up the learning algorithm as well (see, for example, [8]).

It might seem attractive to run a secure multiparty computation (MPC) algorithm within the mini-batch to compute the sum of the gradients. The goal of MPC is to compute a function of the private inputs of the parties in such a way that at the end of the computation, no party knows anything except what can be determined from the result and its own input [13]. Secure sum computation is an important application of secure MPC [2].

However, we do not only require our algorithm to be secure but also fast, light-weight, and robust, since the participating nodes may go offline at any time and they might have limited resources. One key observation is that for the mini-batch algorithm we do not need a precise sum; in fact, the sum over any group that is large enough to protect privacy will do. We propose a protocol that—using a binomial tree topology and Paillier homomorphic encryption—can produce a "quick and dirty" partial sum even in the event of failures, has adjustable capability of resisting collusion, and can be completed in logarithmic time.

Our secure sum algorithm builds upon a secret sharing scheme where a secret value is split into multiple shares such that all the shares are needed to obtain any information. The basic idea of the algorithm is to divide the local value at each node into shares, encrypt these with asymmetric additively homomorphic encryption, and send them to the root via the chain of ancestors. Although the shares travel together, they are encrypted with the public keys of different ancestors. Along the route, the arrays of shares are aggregated, and periodically re-encrypted. Finally, the root calculates the sum.

Homomorphic cryptosystems can quickly become very expensive, especially considering that in machine learning the gradients can be rather large. To achieve practical viability, we propose an extreme lossy compression, where we discretize floating point gradient values to as few as two bits. We demonstrate experimentally that this does not affect learning accuracy yet allows for an affordable cryptography cost. Our simulations are based on a real smartphone trace [1].

The contributions of the author are:

- A scalable and robust secure sum protocol that is able to securely compute a partial sum even in the event of failures and limited collusion of nodes;
- A proof about its capability of preventing the collusion attack;
- A decentralized mini-batch gradient descent method based on the building of a k-trunked binomial overlay tree and the above protocol.

The corresponding papers are:

- DAIS 2015 [5] Gábor Danner and Márk Jelasity. Fully distributed privacy preserving mini-batch gradient descent learning. In Alysson Bessani and Sara Bouchenak, editors, *Proceedings of the 15th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS 2015)*, volume 9038 of *Lecture Notes in Computer Science*, pages 30–44. Springer International Publishing, 2015.
- SCN 2018 [3] Gábor Danner, Árpád Berta, István Hegedűs, and Márk Jelasity. Robust fully distributed mini-batch gradient descent with privacy preservation. *Security and Communication Networks*, 2018:6728020, 2018.

Thesis 2: Comparison of Federated and Gossip Learning

Machine learning over distributed data stored by many clients has important applications in use cases where data privacy is a key concern or central data storage is not an option. Recently, federated learning was proposed to address this problem. In a master-worker architecture, the workers perform machine learning over their own data and the master merely aggregates the resulting models without seeing any raw data, not unlike the parameter server approach. Gossip learning is a decentralized alternative to federated learning that does not require an aggregation server or any central component. The natural hypothesis is that gossip learning is strictly less efficient than federated learning due to it relying on a more basic infrastructure: only message passing and no cloud resources.

In Chapter 4 of the dissertation, we question this hypothesis. We present a thorough comparison of the two approaches. The experimental scenarios include a real churn trace

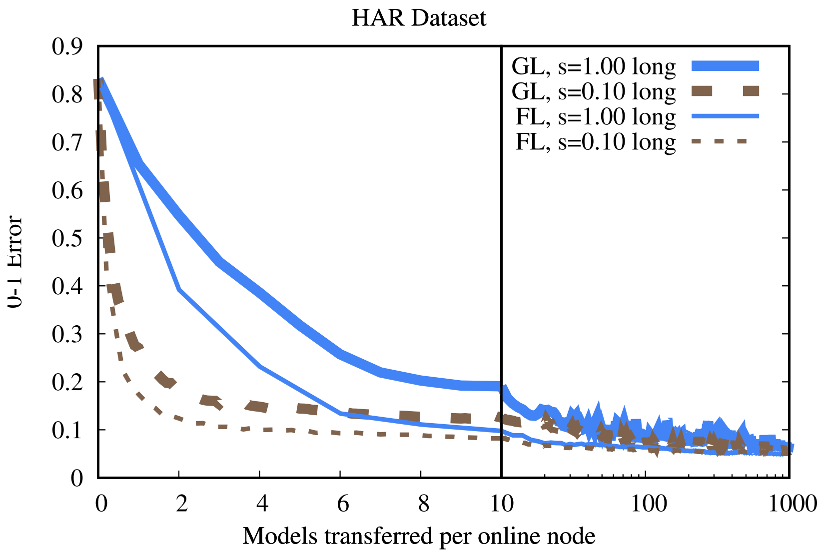

Figure 1: *Federated learning and gossip learning over the smartphone trace with long transfer time, in the 100-node scenario, with different subsampling probabilities.*

collected over mobile phones, different network sizes and different distributions of the training data over the devices. Also, we apply subsampling to reduce communication in both approaches; that is, we send only random subsets of the model parameters. Here, we introduce a new subsampling technique for gossip learning based on partitioned models where each partition has its own age parameter. Instead of sampling parameters independently, one of the partitions is chosen. This way, during model merging, the model parameters can be averaged with appropriate weights without increasing communication costs.

We compare federated and gossip learning in terms of convergence time and model quality, assuming that both approaches utilize the same amount of communication resources in the same scenarios. We also perform a systematic hyperparameter analysis. Surprisingly, the best gossip variants perform comparably with the best federated learning variants overall, thus providing a fully decentralized alternative to federated learning.

The contributions of the author are:

- The partition-based sampling technique;
- The design and development of churn-related modules of the simulator;
- Participation in the design of the improved aggregation algorithm for federated learning;
- Participation in the planning of experiments;
- The optimization of hyperparameters.

The corresponding papers are:

- DAIS 2019 [10] István Hegedűs, Gábor Danner, and Márk Jelasity. Gossip learning as a decentralized alternative to federated learning. In José Pereira and Laura Ricci, editors, *Proceedings of the 19th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS 2019)*, volume 11534 of *Lecture Notes in Computer Science*, pages 74–90. Springer International Publishing, 2019.
- JPDC 2021 [11] István Hegedűs, Gábor Danner, and Márk Jelasity. Decentralized learning works: An empirical comparison of gossip learning and federated learning. *Journal of Parallel and Distributed Computing*, 148:109–124, 2021.

A further related publication:

- ECML 2019 [9] István Hegedűs, Gábor Danner, and Márk Jelasity. Decentralized recommendation based on matrix factorization: A comparison of gossip and federated learning. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 317–332. Springer, 2019.

Thesis 3: Gossip Learning with Adaptive Flow Control

Many decentralized algorithms allow both proactive and reactive implementations. Examples include gossip protocols for broadcasting and decentralized computing, as well as chaotic matrix iteration algorithms. In proactive systems, nodes communicate at a fixed rate in regular intervals, while in reactive systems they communicate in response to certain events such as the arrival of fresh data. Although reactive algorithms tend to stabilize/converge/self-heal much faster, they have serious drawbacks: they may overload the network, and they may also cause starvation when the number of messages circulating in the system becomes too low. Proactive algorithms do not have these problems, but nodes waste a lot of time sitting on fresh information.

In Chapter 5 of the dissertation, we propose the token account framework, a novel family of adaptive protocols that apply rate limiting inspired by the token bucket algorithm to prevent uncontrolled bursts, but they also include proactive communication to prevent starvation. With the help of our traffic shaping service, some applications approach the speed of the reactive implementation, while maintaining strong guarantees regarding the total communication cost and burstiness. In a nutshell, at each node, these algorithms grant one token to the node in regular periods, and sending a message costs a token. A token can be spent immediately (proactive operation), or later, when a message is received (reactive operation). The more tokens a node has, the more eager it is to spend them, possibly sending multiple reactive messages at once. When there are too few messages circulating, the token accounts start to fill up, encouraging an increase in network activity. We perform simulation experiments in different scenarios including a real smartphone availability trace. Our results suggest up to a fourfold speedup in a broadcast application, and an order of magnitude speedup in the case of gossip learning, when compared to the purely proactive implementation.

To evaluate this token-based flow control technique with the mergeable version of gossip learning, we perform machine learning over three different datasets. We also introduce a partitioned variant of the token account algorithm, to properly make use of sampling-based compression. Here, each partition has its own token account. Our results confirm

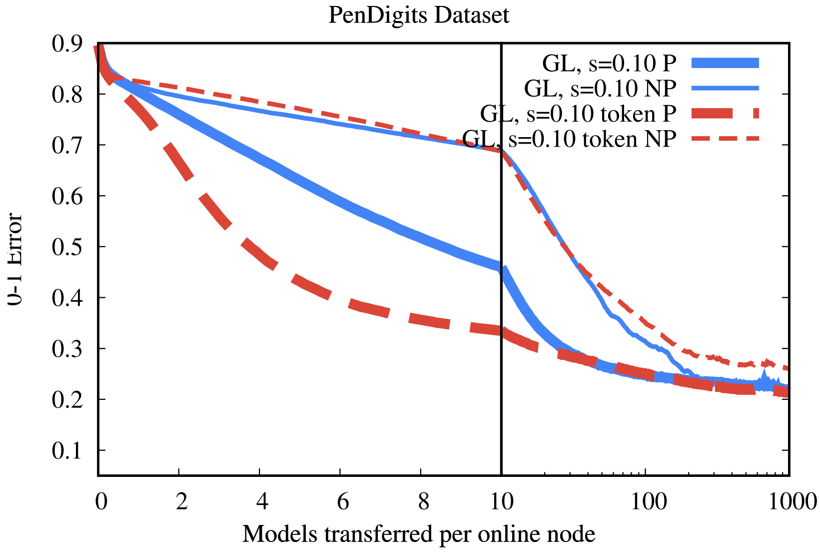

Figure 2: *Proactive and token gossip learning with one node for each example, bursty transfer, subsampling probability $s = 0.1$, smartphone trace with long transfer time. Variants with and without model partitioning are indicated by P and NP, respectively.*

that the token-based flow control approach outperforms proactive gossip learning. Furthermore, it can achieve a performance comparable to federated learning when the distribution of training examples is unbiased. However, to achieve these results, the compression mechanism must be based on partitioning, as opposed to simple subsampling. The reason is that this way, the different partitions can form "hot potato" chains separately, whereas with subsampling, these chains cannot form because sampling picks different weights for each step.

The contributions of the author are:

- The design and evaluation of the token account algorithm, which seeks to keep the number of active hot-potato message chains in dynamic equilibrium;
- An analytical derivation of the average number of tokens in the system;
- The design of the partitioned token gossip learning algorithm that combines the advantages of token-based flow control and subsampling.

The corresponding papers are:

- JPDC 2021 [11] István Hegedűs, Gábor Danner, and Márk Jelasity. Decentralized learning works: An empirical comparison of gossip learning and federated learning. *Journal of Parallel and Distributed Computing*, 148:109–124, 2021.
- ICDCS 2018 [7] Gábor Danner and Márk Jelasity. Token account algorithms: The best of the proactive and reactive worlds. In *Proceedings of The 38th International Conference on Distributed Computing Systems (ICDCS 2018)*, pages 885–895. IEEE Computer Society, 2018.

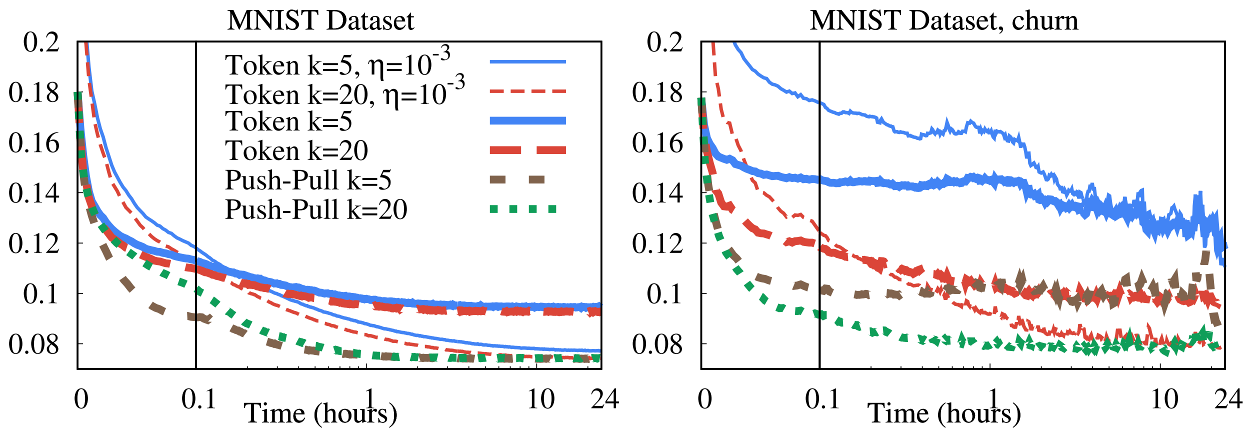

Figure 3: *Compressed push-pull learning and partitioned token gossip learning over the MNIST dataset without churn (left) and with the smartphone trace (right), with different neighborhood sizes and learning rates.*

Thesis 4: Gossip Learning Using Compressed Averaging

Mean estimation, also known as average consensus, is an important computational primitive in decentralized systems. When the average of large vectors has to be computed, as in distributed data mining applications, reducing the communication cost becomes a key design goal. One way of reducing the communication cost is to add dynamic stateful encoder-decoder pairs (codecs) to traditional mean estimation protocols. In this approach, each element of a vector message is encoded in a few bits and decoded by the recipient node. However, due to this encoding and decoding mechanism, these protocols are much more sensitive to benign failure such as message drop and message delay. Properties such as mass conservation are harder to guarantee. Hence, known approaches are formulated under strong assumptions such as reliable communication, atomic non-overlapping transactions or even full synchrony.

In Chapter 6 of the dissertation, we propose a communication efficient algorithm that supports codecs even if transactions overlap and the nodes are not synchronized. The algorithm is based on push-pull averaging, with novel features to support fault tolerance and compression. With the help of simple counters, it is able to detect whether the transferred amount (and the codec state) became inconsistent across the link due to message loss, and rolls back the state to a consistent one. As an independent contribution, we also propose a novel adaptive codec, called the pivot codec. We demonstrate experimentally that our algorithm improves the performance of existing codecs and the novel pivot codec dominates the competing codecs in the scenarios we studied.

Furthermore, we propose a novel variant of gossip learning that uses this codec-based compression to achieve a higher communication efficiency than previous methods could based on subsampling. The algorithm periodically trains the local model and performs the weighted averaging of the models in the network. Among our machine learning experiments we also include a transfer learning scenario. This means that we train a relatively small model on top of a high quality pre-trained feature set that is fixed.

The contributions of the author are:

- The design and evaluation of the compressed push-pull averaging algorithm, which is robust to message drop and overlapping transactions;
- The design and evaluation of the pivot codec, which improves the communication efficiency of the algorithm;
- The design of the compressed push-pull learning algorithm, which enables gossip learning to utilize codecs.

The corresponding papers are:

- Euro-Par 2018 [6] Gábor Danner and Márk Jelasity. Robust decentralized mean estimation with limited communication. In Marco Aldinucci, Luca Padovani, and Massimo Torquati, editors, *Euro-Par 2018*, volume 11014 of *Lecture Notes in Computer Science*, pages 447–461. Springer International Publishing, 2018.
- DICG 2020 [4] Gábor Danner, István Hegedűs, and Márk Jelasity. Decentralized machine learning using compressed push-pull averaging. In *Proceedings of the 1st International Workshop on Distributed Infrastructure for Common Good*, pages 31–36, 2020.

References

[1] Árpád Berta, Vilmos Bilicki, and Márk Jelasity. Defining and understanding smartphone churn over the internet: a measurement study. In *Proceedings of the 14th IEEE International Conference on Peer-to-Peer Computing (P2P 2014)*. IEEE, 2014.

[2] Chris Clifton, Murat Kantarcioglu, Jaideep Vaidya, Xiaodong Lin, and Michael Y. Zhu. Tools for privacy preserving distributed data mining. *SIGKDD Explor. Newsl.*, 4(2):28–34, December 2002.

[3] Gábor Danner, Árpád Berta, István Hegedűs, and Márk Jelasity. Robust fully distributed mini-batch gradient descent with privacy preservation. *Security and Communication Networks*, 2018:6728020, 2018.

[4] Gábor Danner, István Hegedűs, and Márk Jelasity. Decentralized machine learning using compressed push-pull averaging. In *Proceedings of the 1st International Workshop on Distributed Infrastructure for Common Good*, pages 31–36, 2020.

[5] Gábor Danner and Márk Jelasity. Fully distributed privacy preserving mini-batch gradient descent learning. In Alysson Bessani and Sara Bouchenak, editors, *Proceedings of the 15th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS 2015)*, volume 9038 of *Lecture Notes in Computer Science*, pages 30–44. Springer International Publishing, 2015.

[6] Gábor Danner and Márk Jelasity. Robust decentralized mean estimation with limited communication. In Marco Aldinucci, Luca Padovani, and Massimo Torquati, editors, *Euro-Par 2018*, volume 11014 of *Lecture Notes in Computer Science*, pages 447–461. Springer International Publishing, 2018.

[7] Gábor Danner and Márk Jelasity. Token account algorithms: The best of the proactive and reactive worlds. In *Proceedings of The 38th International Conference on Distributed Computing Systems (ICDCS 2018)*, pages 885–895. IEEE Computer Society, 2018.

[8] Ofer Dekel, Ran Gilad-Bachrach, Ohad Shamir, and Lin Xiao. Optimal distributed online prediction using mini-batches. *J. Mach. Learn. Res.*, 13(1):165–202, January 2012.

[9] István Hegedűs, Gábor Danner, and Márk Jelasity. Decentralized recommendation based on matrix factorization: A comparison of gossip and federated learning. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 317–332. Springer, 2019.

[10] István Hegedűs, Gábor Danner, and Márk Jelasity. Gossip learning as a decentralized alternative to federated learning. In José Pereira and Laura Ricci, editors, *Proceedings of the 19th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS 2019)*, volume 11534 of *Lecture Notes in Computer Science*, pages 74–90. Springer International Publishing, 2019.

[11] István Hegedűs, Gábor Danner, and Márk Jelasity. Decentralized learning works: An empirical comparison of gossip learning and federated learning. *Journal of Parallel and Distributed Computing*, 148:109–124, 2021.

[12] Róbert Ormándi, István Hegedűs, and Márk Jelasity. Gossip learning with linear models on fully distributed data. *Concurrency and Computation: Practice and Experience*, 25(4):556–571, 2013.

[13] Andrew C. Yao. Protocols for secure computations. In *Proceedings of the 23rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 160–164, 1982.

Összefoglalás

A pletyka alapú tanulás egy teljesen elosztott gépi tanulási keretrendszer, ahol a hálózatra kötött eszközök (csomópontok) központi szerver használata nélkül oldanak meg gépi tanulási feladatokat a különböző csomópontokon tárolt adatok összességén. A gépi tanuló modellek véletlen sétákat tesznek a hálózaton, és a helyi adatokon tanítjuk őket. Ezen disszertációban számos új módszert mutatunk be, melyek hatékonyabbá vagy biztonságosabbá teszik a pletyka tanulást.

Pletyka alapú tanulás adatvédelemmel: Egymással összejátszó csomópontok adatokat szerezhetnek meg egy másik csomópontról, ha az ott végzett tanítás előtti és utáni modellváltozat is a birtokukban van. A 3. fejezetben javaslunk egy biztonságos mini-batch gradiens módszert a felhasználók adatvédelmének elősegítése érdekében. A módszerünkben a véletlen séta minden lépésében elvégzünk egy elosztott mini-batch számítást, és az összegzett gradiensek alapján végezzük el a tanítást. Ezt az új biztonságos összegző algoritmusunkkal végezzük el. A mini-batch gradiens algoritmusnak nincs szüksége pontos összegre, és ezt kihasználva magas hibatűrést és skálázódást tudunk elérni.

A federated learning és a pletyka tanulás összehasonlítása: A federated learning központi szervert használ a gépi tanulás során a modellek kiátlagolására. A pletyka alapú tanulás egy decentralizált alternatívát kínál, mivel nem igényel szervert. Adódik a természetes feltevés, hogy a pletyka tanulás szigorúan kevésbé hatékony, mint a federated learning, mivel nem vesz igénybe felhő erőforrásokat. A 4. fejezetben megkérdőjelezzük ezt a feltevést. Összehasonlítjuk a federated és a pletyka tanulást a konvergenciaidő és a modell minősége szempontjából, feltételezve, hogy mindkét megközelítés ugyanannyi kommunikációs erőforrást használhat ugyanabban a forgatókönyvben. Meglepő módon a pletyka tanulás legjobb változatai összességében összehasonlíthatóan teljesítenek a federated learning legjobb variánsaival, ezáltal egy teljesen decentralizált alternatívát biztosítva.

Pletyka alapú tanulás adaptív áramlásvezérléssel: Sok decentralizált algoritmus esetén lehetőség van mind proaktív, mind reaktív megvalósításra. A proaktív rendszerekben a csomópontok rögzített ütemben, rendszeres időközönként kommunikálnak, míg a reaktív rendszerek bizonyos eseményekre, például friss adatok érkezésére reagálva teszik ezt. Bár a reaktív algoritmusok jellemzően sokkal gyorsabban konvergálnak, mégis komoly hátrányai vannak: egyrészt túlterhelhetik a hálózatot, másrészt a rendszerben keringő üzenetek száma túl alacsonyra is csökkenhet. A proaktív algoritmusoknak nincsenek ilyen problémái, de a csomópontok sok időt elpazarolnak friss információkon ülve. Az 5. fejezetben bemutatjuk a fedőhálózaton történő véletlen séták számát dinamikus egyensúlyban tartó algoritmusunkat, mely egyesíti a két megközelítés előnyeit. Az algoritmust alkalmazzuk gépi tanulásra is, kombinálva paraméter-mintavételezéses tömörítéssel. Az eredményeink alapján az adaptív áramlásvezérlésen alapuló megközelítés felülmúlja a proaktív pletykatanulást.

Tömörített átlagoláson alapuló pletyka tanulás: A decentralizált átlagszámítás fontos számítási primitív a decentralizált rendszerekben. Amikor a nagy vektorok átlagát kell kiszámítani, a kommunikációs költségek csökkentése kulcsfontosságú tervezési céllá válik. A 6. fejezetben javaslunk egy csökkentett kommunikációs igényű, robusztus decentralizált átlagoló algoritmust, valamint egy ezen alapuló pletyka tanuló algoritmust. Gépi tanulási kísérleteink során transzfertanulást is alkalmazunk.

Nyilatkozat

Danner Gábor "Efficient Gossip Algorithms for Machine Learning" című PhD disszertációjában a következő eredményekben Danner Gábor hozzájárulása volt a meghatározó:

- Pletyka alapú tanulás adatvédelemmel című fejezetben: egy skálázható és robusztus biztonságos összegző protokoll, amely hibák és bizonyos fokú összejátszás esetén is képes részösszeg biztonságos kiszámítására; ezen protokollról szóló bizonyítás; egy k hosszú törzzsel rendelkező binomiális fa építésén, és a fenti összegző protokollon alapuló decentralizált mini-batch módszer. [8,3]
- A federated learning és gossip learning összehasonlítása című fejezetben: új churn-kezelő peersim modulok tervezése és implementálása; meta-paraméterek optimalizálása. [5] A paraméter-mintavételezéses tömörítés egy új, partíciói-alapú változata. [6] A mátrix-faktorizáció alapú ajánló rendszerek alkalmazásában az algoritmusok finomítása a modell-inicializáció és a federated learning-beli aggregálás tekintetében. [7]
- Pletyka alapú tanulás adaptív áramlásvezérléssel című fejezetben: a fedőhálózaton történő véletlen séták számát dinamikus egyensúlyban tartó (a hálózati csomagok forgalmának limitálására használt token bucket algoritmust általánosító) algoritmus és kiértékelése; a token-szám eloszlás elméleti vizsgálata; [1] alkalmazása gépi tanulásra, kombinálása paraméter-mintavételezéses tömörítéssel. [6]
- Tömörített átlagoláson alapuló pletyka tanulás című fejezetben: egy csökkentett kommunikációs igényű, robusztus decentralizált átlagoló algoritmus; a pivot kodek; [2] az előbbieken alapuló gossip learning algoritmus. [4]

Ezek az eredmények Danner Gábor PhD disszertációján kívül más tudományos fokozat megszerzésére nem használhatók fel.

2022.01.19.

Danner Gábor

....

Danner Gábor
jelölt

....

Dr. Jelasity Márk
témavezető

Az Informatika Doktori Iskola vezetője kijelenti, hogy jelen nyilatkozatot minden társszerzőhöz eljuttatta, és azzal szemben egyetlen társszerző sem emelt kifogást.

2022.01.19.

....

Dr. Jelasity Márk
doktori iskola vezetője

[1] Gábor Danner and Márk Jelasity. Token account algorithms: The best of the proactive and reactive worlds. In Proceedings of The 38th International Conference on Distributed Computing Systems (ICDCS 2018), pages 885–895. IEEE Computer Society, 2018.
[2] Gábor Danner and Márk Jelasity. Robust decentralized mean estimation with limited communication. In Marco Aldinucci, Luca Padovani, and Massimo Torquati, editors, Euro-Par 2018, volume 11014 of Lecture Notes in Computer Science, pages 447–461. Springer International Publishing, 2018.
[3] Gábor Danner, Árpád Berta, István Hegedűs, and Márk Jelasity. Robust fully distributed mini-batch gradient descent with privacy preservation. Security and Communication Networks, 2018:6728020, 2018.
[4] Gábor Danner, István Hegedűs, and Márk Jelasity. Decentralized machine learning using compressed push-pull averaging. In Proceedings of the 1st International Workshop on Distributed Infrastructure for Common Good (DICG), pages 31–36. ACM, 2020.
[5] István Hegedűs, Gábor Danner, and Márk Jelasity. Gossip learning as a decentralized alternative to federated learning. In José Pereira and Laura Ricci, editors, Proceedings of the 19th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS 2019), volume 11534 of Lecture Notes in Computer Science, pages 74–90. Springer International Publishing, 2019.
[6] István Hegedűs, Gábor Danner, and Márk Jelasity. Decentralized learning works: An empirical comparison of gossip learning and federated learning. Journal of Parallel and Distributed Computing, 148:109–124, February 2021.
[7] István Hegedűs, Gábor Danner, and Márk Jelasity. Decentralized recommendation based on matrix factorization: A comparison of gossip and federated learning. In Peggy Cellier and Kurt Driessens, editors, ECML PKDD Workshops 2019, number 1167 in Communications in Computer and Information Science, pages 317–332. Springer Nature Switzerland AG, 2020. DMLE19.
[8] Gábor Danner and Márk Jelasity. Fully distributed privacy preserving mini-batch gradient descent learning. In Alysson Bessani and Sara Bouchenak, editors, Proceedings of the 15th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS 2015), volume 9038 of Lecture Notes in Computer Science, pages 30–44. Springer International Publishing, 2015.