Random Client Selection on Contrastive Federated Learning for Tabular Data

Achmad Ginanjar[a] [*], Xue Li[a], Priyanka Singh[a], and Wen Hua[b]

[a] **School of Electrical Engineering and Computer Science**
The University of Queensland, Queensland, Australia
[b] **Department of Computing, The Hong Kong Polytechnic University Hong Kong**

Abstract

Vertical Federated Learning (VFL) has revolutionised collaborative machine learning by enabling privacy-preserving model training across multiple parties. However, it remains vulnerable to information leakage during intermediate computation sharing. While Contrastive Federated Learning (CFL) was introduced to mitigate these privacy concerns through representation learning, it still faces challenges from gradient-based attacks. This paper presents a comprehensive experimental analysis of gradient-based attacks in CFL environments and evaluates random client selection as a defensive strategy. Through extensive experimentation, we demonstrate that random client selection proves particularly effective in defending against gradient attacks in the CFL network. Our findings provide valuable insights for implementing robust security measures in contrastive federated learning systems, contributing to the development of more secure collaborative learning frameworks.

1 Introduction

Vertical Federated Learning (VFL) (Liu et al. 2024a) has emerged as a promising approach in collaborative machine learning. It allows multiple parties to jointly train models while maintaining data privacy through vertical partitioning of features (Liu et al. 2024b). This paradigm has gained significant attention in privacy-sensitive domains such as healthcare and finance, where different organisations possess distinct feature sets of the same entities (Yang et al. 2019).

Despite its potential to preserve data privacy, VFL faces inherent vulnerabilities related to information leakage during the intermediate computation sharing process. Research has shown that even partial information exchange can potentially expose sensitive data characteristics, compromising the fundamental privacy guarantees of the system (Lyu et al. 2024). These limitations have prompted researchers to seek more robust privacy-preserving solutions.

Contrastive Federated Learning (CFL) was introduced as an innovative approach to address these pri-
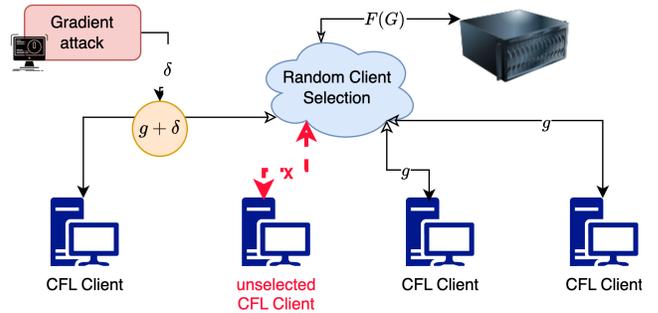

Figure 1: Random client selection within CFL network to defend poisoning gradient attack.

vacy concerns (Ginanjar et al. 2025). By incorporating contrastive learning principles, CFL reduces the need for direct feature sharing while maintaining model performance through representation learning. This method has demonstrated promising results in minimising information leakage during the training process.

However, while CFL enhances privacy preservation in feature sharing, it does not fully address the broader spectrum of security threats in federated learning, particularly internal attacks. Among these, parameter-based attacks have emerged as a significant concern, where malicious participants can exploit parameter information to reconstruct private training data or compromise model integrity (Xia et al. 2023). These attacks pose a substantial threat to the security of federated learning systems, potentially undermining their practical applications.

This paper presents three key contributions. First, it offers a systematic analysis of the impact of parameter-based attacks on contrastive federated learning (CFL) systems. Second, it demonstrates that employing a simple random client selection strategy serves as an effective defense mechanism against such attacks (Colosimo and De Rango 2024) , see Figure 1. Lastly, it quantifies the effectiveness of this defence across 10 real-world datasets, encompassing a variety of attack scenarios.

Our experimental results show that random client se-

lection can reduce attack success rates across all tested scenarios, while maintaining model performance. This finding is particularly significant for practical implementations, as it offers a computationally efficient defence mechanism (Fu et al. 2023) that can be immediately deployed in existing CFL systems.

2 Related Work

This section reviews existing research relevant to contrastive federated learning, gradient attacks in federated learning, and client selection as a defence mechanism.

2.1 Contrastive Federated Learning

Contrastive learning (Gutmann and Hyvärinen 2010; Chen et al. 2020) has been adapted to the federated setting to enhance privacy preservation. Ginanjar *et al.* (Ginanjar et al. 2025) introduce Contrastive Federated Learning (CFL) specifically designed for tabular data silos. Their work focuses on reducing the reliance on direct feature sharing by leveraging representation learning, thereby mitigating potential information leakage. This approach aims to achieve a balance between model performance and privacy protection in collaborative learning with vertically partitioned data scenarios.

2.2 Model Attacks in Federated Learning

Federated learning systems are prone to various attacks, notably those exploiting gradient and paramter information (Tolpegin et al. 2020). Xia *et al.* (Xia et al. 2023) provide a comprehensive survey of poisoning attacks in federated learning, detailing how malicious actors can manipulate parameters to compromise model integrity or infer sensitive training data. A broader perspective on privacy and robustness in federated learning is offered by Lyu *et al.* (Lyu et al. 2022). They discuss a range of attacks, including gradient-based attacks, and present various defence mechanisms to counter these threats. These studies highlight the critical need for robust security measures in federated learning.

2.3 Client Selection in Federated Learning

Lei Fu *et al.* (Fu et al. 2023) work provides a broad explanation of federated learning client selection methods. They mention that although a random sample selection does not consider heterogeneity, this method is most likely selected. This is because beside other algorithm complexity (Wu and Wang 2022; Luping, Wei, and Bo 2019; Lai et al. 2021; Zhou et al. 2022) , This is because, besides other algorithm complexity (Wu and Wang 2022; Luping, Wei, and Bo 2019; Lai et al. 2021; Zhou et al. 2022) , these approaches rely on experiments to demonstrate their effectiveness. However, there is no guarantee of their performance in the real world.

In this study, we employ random client selection.

3 Problem Formulation

3.1 Contrastive Federated Learning

Let $D = \{(x_i, y_i)\}_{i=1}^{N}$ represent the training dataset, where x_i denotes the feature vector and y_i the corresponding label. In the vertical federated learning setting, the feature space is partitioned across K parties, where each client k holds a subset of features x_i^k. The CFL framework can be formalised as:

- Client's objective is:
  $f_c(\bar{D}((\bar{E} : x; \omega^e); \omega^d)) \to x^d$

- Server's Objective is:
  $F(g) = \frac{1}{K} \sum_{k=1}^{K} (\omega^e, \omega^d) \to (\omega^{eG}, \omega^{dG})$

3.2 Model Attack

The gradient-based poisoning attack in CFL can be formulated as follows: Attack Objective:

1. A malicious party aims to reconstruct private information or compromise model performance by manipulating parameters:
   $\min_\delta L_{attack}(\omega + \delta)$ where ω represents the true parameter and δ the poisoning attack.

2. Attack constraints:
   - Parameter manipulation must remain within bounds to avoid detection: $||\delta|| \leq \epsilon$
   - The poisoned gradients should maintain statistical similarity to legitimate updates:
     $||stats(\omega + \delta) - stats(\omega)|| \leq \tau$

   Our study applies model scaling attacks. Applies $\omega + \delta = \omega \cdot \alpha$ where $\alpha \in R^+$ and α is the poison level.

3.3 Random Client Selection

The random client selection mechanism can be formalised as:

1. Selection Process
   At each training round t, a subset of clients S_t is randomly selected from the total client pool:
   $S_t \subset 1, ..., K, |S_t = m|$
   where m is the number of clients selected per round.

2. Selection probability
   Each client k has an equal probability of being selected: $P(k \epsilon S_t) = \frac{m}{K}$

3. Defence objective
   The random selection aims to minimise the attack success probability: $\min_{S_t} P(Attack_{success} || S_t$

3.4 Combined Defence Framework

The overall defence framework integrates these components:

- Select clients: $S_t \sim \text{Uniform}(K, m)$

- Update local models: $\omega_k^{t+1} = \omega_k^t \eta \bigtriangledown L_{cont}(\omega_k^t)$

- Aggregate updates: $\omega^{t+1} = \frac{1}{|S|} \sum_{k \in S_t} \omega_k^{t+1}$

Strategy	Not	Information
Client poisoning number	p_c	number of clients being poisoned.
Scaling poisoning level	p_l	$\omega \cdot \alpha$ where $\alpha = p_l$
Random client level	r_c	number of clients skipped during FL.

Table 1: Poisoning attack settings where $\{k, p_c, p_l, r_l\} \in Z$. The 'Not' is the notation.

3.5 Theoretical Analysis

Consider ω_t is the model parameter at iteration t, η learning rate, and μ strong convexity parameter. Given the poisoning probability p_c and selection ratio r_l, the attack success probability is bounded by $P(\text{attack}) \leq p_c \cdot r_l$. Under random selection ratio r_l, the expected convergence satisfies: $E[||\omega_t - \omega*||^2] \leq (1 - \eta\mu r_l)^t ||\omega_0 - \omega*||^2$.

4 Experiments

The experiments in this study use ten datasets: Adult (Becker and Kohavi 1996), Helena (Guyon et al. 2019), Jannis (Guyon et al. 2019), Higgs Small (Baldi, Sadowski, and Whiteson 2014), Aloi (Geusebroek, Burghouts, and Smeulders 2005), Epsilon (PASCAL Challenge on Large Scale Learning 2008), Cover Type (Blackard and Dean 2000), California Housing (Pace and Barry 1997), Year (Bertin-Mahieux et al. 2011), Yahoo (Chapelle and Chang 2011), and Microsoft (Qin and Liu 2013).

We perform an extensive study to challenge CFL. We use the model scaling attack proposed by Cao *et al.* (Cao et al. 2023). We intentionally do not apply any defence, such as Byzantine (So, Güler, and Avestimehr 2021) and other security enhancements (Lyu et al. 2022) to the CFL network. This was done to test the robustness of CFL. This study covers three different parameters as shown in Table 1.

The comprehensive settings for our experiments are presented in Table 2. Across all datasets, we executed a total of 19 distinct experiments.

To illustrate, consider an experiment that is assigned specific parameters: $\{n : 8, p_c : 0.2, p_l : 0.1, r_c : 0.2\}$, this indicates that the experiment is composed of 8 clients operating within a federated learning network. Among these clients, a designated subset of clients is classified as adversarial, each exhibiting a specific level of $(\omega \cdot 0.1)$ data poisoning designed to test the robustness of the learning model. The global server is responsible for aggregating the model parameters, collecting data only from a random $int(8 \cdot 0.2) = 2)$ of the clients. This methodology aims to simulate real-world scenarios where malicious clients may attempt to disrupt the learning process, allowing us to assess the robustness of CFL under various attack conditions. The details of the experiments are provided in Algorithm 4.

Table 2: Parameters used for the experiments.

Notation	values
p_c	$[0, 0.2, 0.5]$
p_l	$[0.1, 0.5, 2]$
r_c	$[1, 0.2, 0.8]$

Algorithm 1 CFL with Potential Model Poisoning

1: **Initialise:**
2:    Global model F_G
3:    Set of clients $C = \{C_1, ..., C_n\}$
4:    Poison clients $P_c \subset C$ randomly selected
5: **for** each epoch $e = 1, ..., E$ **do**
6:    **for** each batch b **do**
7:       **for** each client $c_i \in C$ **do**
8:          Train local model: $L_i = \text{Train}(c_i, \text{batch})$
9:          **if** $c_i \in P_c$ **then**
10:             Apply poisoning: $\alpha.\omega \rightarrow \omega_\alpha$
11:          **end if**
12:       **end for**
13:       $\omega_G = \text{Agg}(\{\omega_\alpha, \omega_1, ..., \omega_n\})$
14:       **for** each client $c_i \in C$ **do**
15:          Update client model: $\omega_{c_i} \leftarrow \omega_G$
16:       **end for**
17:    **end for**
18: **end for**

---

5 Results

Table 3 shows the number of failed clients to the number of poisoned clients (p_c) on different dataset., with percentages representing poisoned clients (0%, 20%, and 50%). Most dataset like "adult," "aloi," and "covtype" show no impact (0.00) when unpoisoned, but a consistent impact (15.79) when 20% or 50% clients are poisoned. The "epsilon" dataset remains unaffected (0.00) across all scenarios, while the "year" dataset shows a slightly higher impact (16.45) with 50% clients poisoned. From the Table, it is clear that CFL is able to survive with low fail rate (Fail $\leq 16\%$) in all the experiments.

Table 4 shows the number of failed clients with vary-

Table 3: Number of failed clients to the number of poisoned clients (p_c) on different datasets.

p_c	0	0.2	0.5
adult	0.00	15.79	15.79
aloi	0.00	15.79	15.79
covtype	0.00	15.79	15.79
epsilon	0.00	0.00	0.00
helena	0.00	15.79	15.79
higgs small	0.00	15.79	15.79
jannis	0.00	15.79	15.79
microsoft	0.00	15.79	15.79
yahoo	0.00	15.79	15.79
year	0.00	15.79	16.45

Table 4: Number of failed clients with different data poisoning levels (p_l).

p_l	0.1	0.5	1	2
adult	10.53	10.53	0.00	10.53
aloi	10.53	10.53	0.00	10.53
covtype	10.53	10.53	0.00	10.53
epsilon	0.00	0.00	0.00	0.00
helena	10.53	10.53	0.00	10.53
higgs small	10.53	10.53	0.00	10.53
jannis	10.53	10.53	0.00	10.53
microsoft	10.53	10.53	0.00	10.53
yahoo	10.53	10.53	0.00	10.53
year	11.18	10.53	0.00	10.53

Table 5: Number of failed clients compared to the percentage of clients being used for aggregation during Federated Learning (r_l) across different datasets.

r_l	0.2	0.8	1
adult	31.58	0.00	0.00
aloi	31.58	0.00	0.00
covtype	31.58	0.00	0.00
epsilon	0.00	0.00	0.00
helena	31.58	0.00	0.00
higgs small	31.58	0.00	0.00
jannis	31.58	0.00	0.00
microsoft	31.58	0.00	0.00
yahoo	31.58	0.00	0.00
year	31.58	0.66	0.00

ing data poisoning levels p_l (0.1, 0.5, 1, and 2) on different datasets. Most datasets, including adult, aloi, covtype, and others, maintain a value of 10.53 at lower poisoning levels (0.1 and 0.5), drop to 0.00 at level 1, and return to 10.53 at level 2. The epsilon dataset remained at 0.00 across all levels, while the year dataset varied slightly with values of 11.18 at 0.1, 10.53 at 0.5, dropping to 0.00 at level 1, and returning to 10.53 at level 2. From the Table, it is clear that CFL is able to survive with low fail rate (Fail $\leq$ 11%) in all the experiments.

Table 5 shows the Number of failed clients when a certain percentage of clients is used for aggregation during Federated Learning (r_l) across different datasets. Three r_l values were tested: 0.2, 0.8, and 1.0. Most datasets, including adult, covtype, and yahoo, utilised 31.58% of clients at r_l=0.2, with 0% at r_l=0.8 and 1.0. The epsilon dataset showed 0% client usage for all r_l values. The year dataset had 31.58% at RL=0.2, 0.66% at r_l=0.8, and 0% at r_l=1.0. Generally, lower r_l values (0.2) engage more clients than higher values (0.8 and 1.0). CFL models are not effective when only 0.2 of the total client number joins the aggregation on the server.

Figure 2 shows the mean of the standard deviation of the performance across the dataset. Based on the previous result (Table 5), this figure was calculated by removing all clients with $r_l = 0.2$. From the figure,
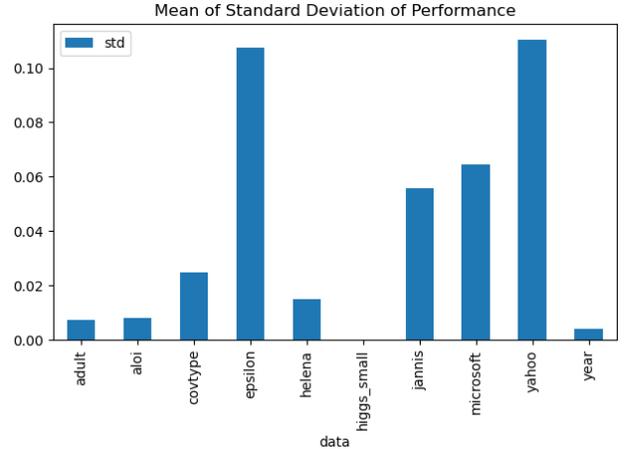

Figure 2: The mean of the standard deviation of the performance across the dataset.

we can see that the mean of the standard deviation is small $\mu_{std} \leq 0.1$. The highest mean values are in Epsilon and Yahoo, which have the largest feature sets. From this figure, it is clear that CFL is able to maintain performance across different experiment settings.

From the theoretical analysis section, our experiment findings can be explained that r_l reduce attack success probability. In addition, model convergence remains stable when $r_l \geq 0.8$.

6 Conclusion

Random client selection is an effective strategy to defend against adversarial attacks in contrastive federated learning (CFL). Our research demonstrates that CFL can effectively address model poisoning attacks through the use of random client selection in many experiments. This study focuses specifically on model scaling attacks and random selection defence, leaving other attack vectors and defence strategies for future work. Additionally, while our empirical results are promising, we provide only basic theoretical guarantees.

7 Acknowledgements

References

Baldi, P.; Sadowski, P.; and Whiteson, D. 2014. Searching for exotic particles in high-energy physics with deep learning. *Nature Communications* 5:4308.

Becker, B., and Kohavi, R. 1996. Adult. UCI Machine Learning Repository. DOI: https://doi.org/10.24432/C5XW20.

Bertin-Mahieux, T.; Ellis, D. P.; Whitman, B.; and Lamere, P. 2011. The million song dataset. In *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011)*, 591–596.

Blackard, J. A., and Dean, D. J. 2000. Comparative accuracies of artificial neural networks and discriminant analysis in predicting forest cover types from cartographic variables. *Computers and Electronics in Agriculture* 24(3):131–151.

Cao, X.; Jia, J.; Zhang, Z.; and Gong, N. Z. 2023. Fedrecover: Recovering from poisoning attacks in federated learning using historical information. In *Proceedings - IEEE Symposium on Security and Privacy*, volume 2023-May, 1366–1383. Institute of Electrical and Electronics Engineers Inc.

Chapelle, O., and Chang, Y. 2011. Yahoo! learning to rank challenge overview. In *Proceedings of the Learning to Rank Challenge*, volume 14 of *Proceedings of Machine Learning Research*, 1–24. PMLR.

Chen, T.; Kornblith, S.; Norouzi, M.; and Hinton, G. 2020. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, 1597–1607. PMLR.

Colosimo, F., and De Rango, F. 2024. Distance-statistical based byzantine-robust algorithms in federated learning. In *2024 IEEE 21st Consumer Communications & Networking Conference (CCNC)*, 1034–1035. IEEE.

Fu, L.; Zhang, H.; Gao, G.; Zhang, M.; and Liu, X. 2023. Client selection in federated learning: Principles, challenges, and opportunities. *IEEE Internet of Things Journal* 10(24):21811–21819.

Geusebroek, J.-M.; Burghouts, G. J.; and Smeulders, A. W. M. 2005. The amsterdam library of object images. *International Journal of Computer Vision* 61(1):103–112.

Ginanjar, A.; Li, X.; Hua, W.; and Pei, J. 2025. Contrastive federated learning with tabular data silos.

Gutmann, M., and Hyvärinen, A. 2010. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 297–304. JMLR Workshop and Conference Proceedings.

Guyon, I.; Sun-Hosoya, L.; Boullé, M.; Escalante, H. J.; Escalera, S.; Liu, Z.; Jajetic, D.; Ray, B.; Saeed, M.; Sebag, M.; Statnikov, A.; Tu, W.-W.; and Viegas, E. 2019. Analysis of the automl challenge series 2015-2018. In *AutoML*, Challenges in Machine Learning. Springer.

Lai, F.; Zhu, X.; Madhyastha, H. V.; and Chowdhury, M. 2021. Oort: Efficient federated learning via guided participant selection. In *15th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 21)*, 19–35.

Liu, Y.; Kang, Y.; Zou, T.; Pu, Y.; He, Y.; Ye, X.; Ouyang, Y.; Zhang, Y. Q.; and Yang, Q. 2024a. Vertical federated learning: Concepts, advances, and challenges. *IEEE Transactions on Knowledge and Data Engineering* 36:3615–3634.

Liu, Y.; Kang, Y.; Zou, T.; Pu, Y.; He, Y.; Ye, X.;

Ouyang, Y.; Zhang, Y.-Q.; and Yang, Q. 2024b. Vertical federated learning: Concepts, advances, and challenges. *IEEE Transactions on Knowledge and Data Engineering* 36(7):3615–3634.

Luping, W.; Wei, W.; and Bo, L. 2019. Cmfl: Mitigating communication overhead for federated learning. In *2019 IEEE 39th international conference on distributed computing systems (ICDCS)*, 954–964. IEEE.

Lyu, L.; Yu, H.; Ma, X.; Chen, C.; Sun, L.; Zhao, J.; Yang, Q.; and Yu, P. S. 2022. Privacy and robustness in federated learning: Attacks and defenses. *IEEE transactions on neural networks and learning systems* 35(7):8726–8746.

Lyu, L.; Yu, H.; Ma, X.; Chen, C.; Sun, L.; Zhao, J.; Yang, Q.; and Yu, P. S. 2024. Privacy and robustness in federated learning: Attacks and defenses. *IEEE Transactions on Neural Networks and Learning Systems* 35:8726–8746.

Pace, R. K., and Barry, R. 1997. Sparse spatial autoregressions. *Statistics & Probability Letters* 33(3):291–297.

PASCAL Challenge on Large Scale Learning. 2008. Epsilon Dataset: Simulated Physics Experiments. http://largescale.ml.tu-berlin.de/instructions/. Accessed: [Insert Access Date].

Qin, T., and Liu, T.-Y. 2013. Introducing LETOR 4.0 datasets. *arXiv preprint arXiv:1306.2597*.

So, J.; Güler, B.; and Avestimehr, A. S. 2021. Byzantine-resilient secure federated learning. *IEEE Journal on Selected Areas in Communications* 39(7):2168–2181.

Tolpegin, V.; Truex, S.; Gursoy, M. E.; and Liu, L. 2020. Data poisoning attacks against federated learning systems. In *Computer security–ESORICs 2020: 25th European symposium on research in computer security, ESORICs 2020, guildford, UK, September 14–18, 2020, proceedings, part i 25*, 480–501. Springer.

Wu, H., and Wang, P. 2022. Node selection toward faster convergence for federated learning on non-iid data. *IEEE Transactions on Network Science and Engineering* 9(5):3099–3111.

Xia, G.; Chen, J.; Yu, C.; and Ma, J. 2023. Poisoning attacks in federated learning: A survey. *Ieee Access* 11:10708–10722.

Yang, Q.; Liu, Y.; Chen, T.; and Tong, Y. 2019. Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)* 10(2):1–19.

Zhou, P.; Xu, H.; Lee, L. H.; Fang, P.; and Hui, P. 2022. Are you left out? an efficient and fair federated learning for personalized profiles on wearable devices of inferior networking conditions. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6(2):1–25.