

Paper-2 Endling Execution Assurance of Jederated Learning on Untrusted Participonts. Keywords: Proof of work, Trusted Execution Environment (TEE). Issues: and Defenses: remove / weaken outliers from model update of pa ponts, to prevent model poisoning attacks [7] [9] [1] garente convergence of global model[8] data poisoning attacks, strattly back door altacks (2), [13], prevent lary participants TEE related issues. only executing TEE requested training rounds Defenses freshness con't be gowenteed, as F71- [9] TEE'S scheduling and I/O are controlled by host participant[20] Issues Conto old results weresponding werect proofs in the previous FL epoch to lain rewards sybil - based sheating attacks to ask for multiple rewards

sompling discions, the TEE requires to receive a commitment from the participant greshess:

"Dynomic yet deterministic" data selection luvings dynomic input data to make training results different.

Bind selecting decision with each FL's epoch's identifier

"Tompering for free identifiers to portriponts and agrainst sybil attacks.

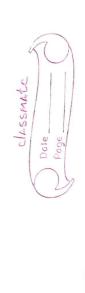
Intel's Software Gawrd Extension (SGX)

1 generates a measurement with Rash digest

2 program in inclove produces a report
including the measurement and supplementing
data (eg., output), and signs it as a quote
with a private attestation key found in the
processor.

3. can be verified by trustworthy measurement of binary and and Intel's Attestation Source (IPS)

(SKTREE » KTEE); ∑ > unforgrable signature L'key



output out is fordued by progrem of running S. vori (pKTEE, or, Gout) outputs 1 0= 5 Sign (SKTEE, 4, out) enclave generate altertation enclove ta

mit gating explicatorches by diminishing gains of adivorcing with increased computing efforts Threat Model

OVERVIEW

bosticiponts. P and SGX - enabled blatform specifice thaining selfo, newoods etc Registration Solicitution

program, lanches the on enclave with the program identifier. Then, P nevieurs and downloads the gregister with infolibe data description and

and generated on outestation report to demogratical the above there program outsto. Swelfier > E. Verill

Les

Tor unselected Swould instruct them when they S selects subset of participants [29], [33] Delection

son reconnect for participation.

Date -

2	Reporting: P gets the langus model et
	bestorms local training generales model
	wholase 1x and invokes include program
	TEE to generale brook op.
	TEE runs (SKTEE, DNN() D, Oct, Cot) - Gp
	while using randomly selected training
	Tounds.
Ŋ	Verification and Aggregation:
	S benfarms Winky (pKTHE, DININI), Oct ship Job)-140,
	S distributes secontals.
	DESIGN DETAILS
	Progo - Program outside enclose
	Pog i → inside
Ä.	Raveline Design
	Progo stores the bash digest of all model
	parameters Os at the end of each training round
	and sender commit message to TEE when all
	training rounds finish.
	Progi magnets to think one modern training
	rained, Proge retowns the input barroniters
-	the habreled of aution pana war ishondang
	training data, Then Grainfucks, Hickory
	the consistent

unsistent.

thy and

Progi seg model parameters for last round computes mudel updates chicks commitment generates a proof aslocal atteration of local execution Design of Commitment Scheme results of all training rounds and order of results arbitrary result verified according to commitment efficiently. hash from R rands 5={h,h2, hp} Straight forward: Progi randomly fetches several pairs of two successive hash values, requests raw input model parameters from Progo, and recomputes hash values to insul coincidence For constant performance: we adopt Merkle hash tree [35] based commitment MAT based: Progo exilds Merkel Tree, leaf nocles are hash values. Internal nodes are concatenation of hash values of two shildren. Progi requests of generalismos to identify the training sounds to be verified



Input parameters, hash value of output para auxilleony info (sibling nodes of nodes in the path from the requested leaf node to the tree root).

B. John guaranteling the Freshness of Results

Scheme is hard to governtee the freshness of

single training round and final results.

Y = rounds performed, R -> required no of Hounds,

L-> no, of rondomly selected rounds

detected misbehariour: 1 - (R-[Ry]+1)

1

Solution $\binom{n}{r} = \binom{n}{r}$

leverage dynamics of input training data, which make epochs differentiable.

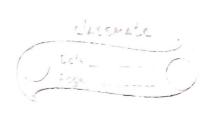
Progi megotiate with Progo about data organization in advance.

Progi generates on index and HMAC sfor each training data restored outside enclave before training.

Progi Seed Progo

Iij = PRFl seed Xitj) mod N index of jth training data at randi PRF -> psuedo rondom function

N -> total training stata.



hicks chicked integrity of outside data with HMAC, trains the DNN, verifies the output parameters.

Bind training results with used global model

seed = Progi receive light digest of warrant global

model (denoted as hot) from Prog.

Using signature and model digest, the server con ensure that whether the participant used latest model and faithfully performed local training.

sybile-based, unique identifier T, tomfur free identifier, even if on adversary registers multiple aliases, he must bounch some no of enclove inston as to generate proofs.