Crypto-PPML 2019 @UCSB, CA, USA

A Survey on Deep Learning Techniques for Privacy-Preserving

Kwangjo Kim, Harry Chandra Tanuwidjaja, and Rakyong Choi

Korea Advanced Institute of Science and Technology (KAIST)

August 18, 2019

KAIST

CONTENTS

KAIST

History of Deep Learning : Ideas and Milestone

Crypto-PPML 2019

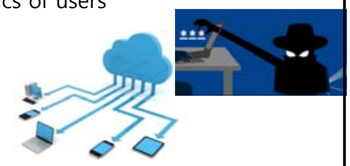
- **1943: Neural networks**
- **1957: Perceptron**
- **1974-86: Backpropagation, RBM, RNN**
- **1889-98: CNN, MNIST, Bidirectional RNN**
- **2006: Deep Learning**
- **2009: Image Net**
- **2012: AlexNet, Dropout**
- **2014: GAN (Generative Adversarial Network)**
- **2014: DeepFace**
- **2016: AlphaGo**
- **2018: AlphaZero, Capsule Networks**
- **2018 : BERT**(Bidirectional Encoder Representations from Transformers) by Google

https://deeplearning.mit.edu KAIST

Why we need Privacy-Preserving Deep Learning?

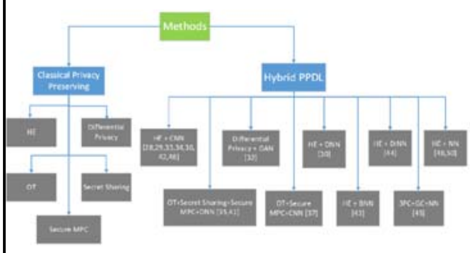
Crypto-PPML 2019

- Advances of **machine learning**
- Users (Data Owner) submit data to the trustful cloud server who want to get useful statics of users
- Data **privacy** during **training**
- Solution?
  - **Privacy Preserving Deep Learning (PPDL)**



KAIST

Our Classification



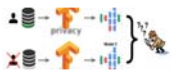
Acronym	Definition
PP	Privacy Preserving
DL	Deep Learning
HE	Homomorphic Encryption
OT	Oblivious Transfer
MPC	Multi Party Computing
CNN	Convolutional Neural Network
DNN	Deep Neural Network
BNN	Binary Neural Network

Classical Privacy-Preserving Technology

- Homomorphic Encryption
  - Support **operations** on encrypted data without private key
  - **Not directly applicable** to DL

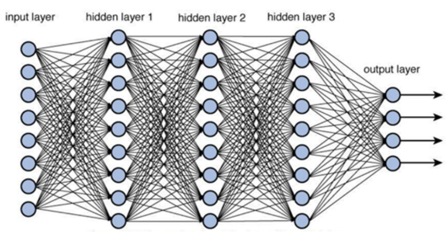
- Secure Multi-party Computation
  - Joint computation of f( ), keeping each input to be **secret**

- Differential Privacy
  - Keeping privacy before and after PP
  - Release statistics **without revealing data**



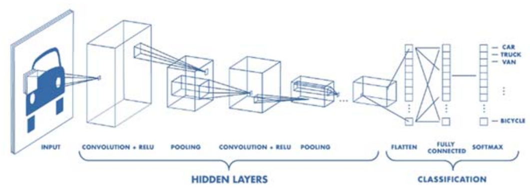
Deep Learning in Privacy-Preserving Technology(1/2)

- Deep Neural Network (DNN)



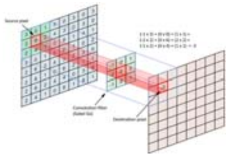
Deep Learning in Privacy-Preserving Technology(2/2)

- Convolutional Neural Network (CNN)



Deep Learning Layers(1/5)

- Convolutional Layer
  - Apply a convolution operation to the input, **passing the result** to the next layer.
  - **Dot product** operation
  - Can be used **directly** in HE

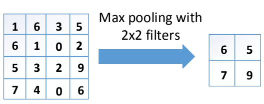


Deep Learning Layers(2/5)

- Activation Layer
  - **Non-linear** function that applies mathematical process on the output of convolutional layer.
  - Activation function: ReLU, Sigmoid, Tanh
  - Non-linear -> **high complexity**

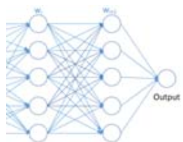
Deep Learning Layers(3/5)

- Pooling Layer
  - A sampling layer, whose purpose is to reduce the size of data
  - **Cannot** use max pooling in HE
  - Solution? **Average pooling**



Deep Learning Layers(4/5)

- Fully Connected Layer
  - **Each** neuron in this layer is **connected** to neuron in previous layer
  - The connection represents the **weight of the feature** like a compl ete graph
  - **Dot product** function
  - Can be **used directly** in HE



Deep Learning Layers(5/5)

- Dropout Layer
  - **Reduce overfitting**, act as regularizer
  - Not using all neurons
  - **Drops** some neurons **randomly**

Standard neural network Neural Network after applying dropout

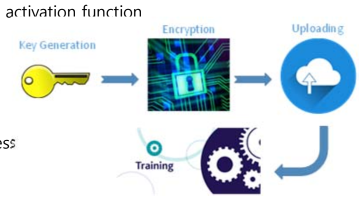
X-based Hybrid PPDL

- HE-based Hybrid PPDL
- Secure MPC-based Hybrid PPDL
- Differential Privacy-based Hybrid PPDL

HE-based Hybrid PPDL

HE-based Hybrid PPDL(1/10)

- ML Confidential: Machine Learning on Encrypted Data
  - **Polynomial approximation** as activation function
  - Cloud based scenario
  - Homomorphic encryption
  - Data is transferred to server
  - Cloud server do training process

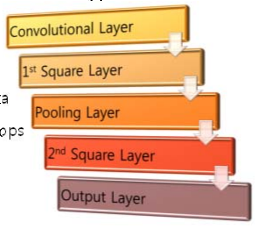


T. Graepel, K. Lauter, and M. Naehrig, "ML confidential: Machine learning on encrypted data," Internationa l Conference on Information Security and Cryptology, pp. 1-21, 2012.

HE-based Hybrid PPDL(2/10)

Crypto-PPML 2019

- Cryptonets: Applying Neural Networks to Encrypted Data with High Throughput and Accuracy
  - Protect data exchange in cloud service
  - Apply CNN to homomorphically encrypted data
  - **Weakness**: error rate increase and accuracy drops
    - When?
    - If the number of non linear layer is big

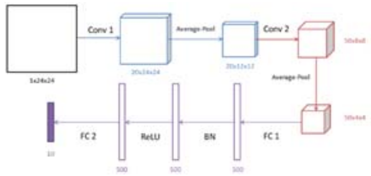


R. Gilad-Bachrach, N. Dowlin, K. Laine, K. Lauter, M. Naehrig, and J. Wernsing, "Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy," In ternational Conference on Machine Learning, pp. 201-210, 2016.

HE-based Hybrid PPDL(3/10)

Crypto-PPML 2019

- Privacy-Preserving on Deep Neural Network
  - Cloud service environment
  - Combining HE with CNN
  - **Solve Cryptonets problem**
  - Polynomial approximation
  - Batch normalization layer

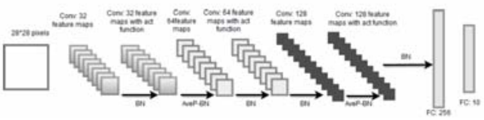


H. Chabanne, A. de Wargny, J. Milgram, C. Morel, and E. Prou, "Privacy-preserving classification on deep neural network," IACR Cryptology ePrint Archive, p. 35, 2017.

HE-based Hybrid PPDL(4/10)

Crypto-PPML 2019

- CryptoDL: Deep Neural Networks Over Encrypted Data
  - Modified CNN for encrypted data with HE
  - **Approximation technique**:
    - Taylor series (Acc 40%)
    - Chebysev polynomial (Acc 70%)
    - Derivative of activation function **(Acc 99.52%)**

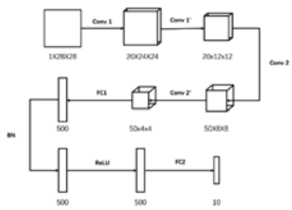


E. Hesamifard, H. Takabi, and M. Ghasemi, "Cryptodl: Deep neural networks over encrypted data," arXiv prepr int, vol. 1711.05189, 2017.

HE-based Hybrid PPDL(5/10)

Crypto-PPML 2019

- Privacy-Preserving All Convolutional Net Based on Homomorphic Encryption
  - PP technique on CNN by using HE
  - Adding **batch normalization layer**
  - **Polynomial approximation**
  - Convolution layer with **increased stride**

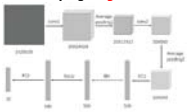


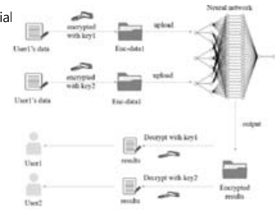
W. Liu, F. Pan, X. A.Wang, Y. Cao, and D. Tang, "Privacy-preserving all convolutional net based on homo morphic encryption," International Conference on Network-Based Information Systems, pp. 752-762, 2018
.

HE-based Hybrid PPDL(6/10)

Crypto-PPML 2019

- Distributed Privacy-Preserving Multi-Key Fully Homomorphic Encryption
  - Substituting ReLU function with low degree polynomial
  - Using batch normalization layer
  - Max pooling -> **average pooling**
  - Beneficial for classifying **large scale distributed data**



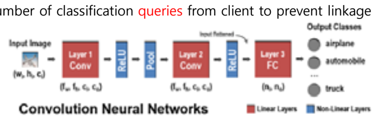


H. Xue, Z. Huang, H. Lian, W. Qiu, J. Guo, S. Wang, and Z. Gong, "Distributed large scale privacy-preserving deep mining," IEEE Third International Conference on Data Science in Cyberspace, pp. 418-422, 2018.

HE-based Hybrid PPDL(7/10)

Crypto-PPML 2019

- Gazelle: A Low Latency Framework for Secure Neural Network Inference
  - Able to switch protocol between HE and GC in PaaS scenario.
  - Structure: two convolutional layers, two ReLU layers, one pooling layer, and one fully connected layer.
  - Hide the weight, bias, and stride size in the convolutional layer.
  - Limit the number of classification queries from client to prevent linkage attack.

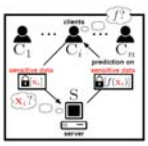


C. Juvekar, V. Vaikuntanathan, and A. Chandrakasan, "GAZELLE: A Low Latency Framework for Secure Neural Network Inference," 27th USENIX Security Symposium, pp. 1651-1669, 2018.

HE-based Hybrid PPDL(8/10)

Crypto-PPML 2019

- Tapas
  - Accelerate parallel computation using encrypted data in PaaS environment.
  - Current problem: large amount of processing time needed.
  - Main contribution:
    - New algorithm to speed up binary computation in Binary Neural Network (BNN).
  - Their technique can be parallelized by evaluating gates at the same Level for three representations at the same time -> **time improved drastically**



A. Sanyal, M.J. Kusner, A. Gascn, and V. Kanade, "TAPAS: Tricks to Accelerate (Encrypted) Prediction as a Service," arXiv preprint, arXiv:1806.03461, 2018.

HE-based Hybrid PPDL(9/10)

Crypto-PPML 2019

- FHE DiNN
  - Reduce complexity problem in HE+NN
  - Deeper network, more complexity
  - Use bootstrapping -> linear complexity of NN
  - How to do it?
    - Discretize the weight, bias value, and the domain of activation function.
    - Using sign activation function to limit the growth of signal in the range of [-1,1]

F. Bourse, M. Minelli, M. Minihold, and P. Paillier, "Fast Homomorphic Evaluation of Deep Discretized Neural Networks," Springer, Cham, 2018

HE-based Hybrid PPDL(10/10)

Crypto-PPML 2019

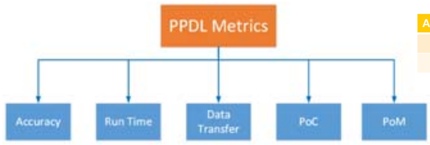
- E2DM
  - PPDL framework that performs matrix operations on HE system
  - Encrypts a matrix homomorphically, then do arithmetic operations on it.
  - Reduce complexity of matrix multiplication
    - O(d) complexity for dot product between two d x d matrices
    - instead of O(d[2]) complexity.
  - Leverage CNN with one convolutional layer, two fully connected layers, and a square activation function.

X. Jiang, M. Kim, K. Lauter, and Y. Song, "Secure Outsourced Matrix Computation and Application to Neural Networks," in Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, pp. 1209-1222, ACM, 2018.

KAIST

Metrics for Comparison

Crypto-PPML 2019



Acronym	Definition
PoC	Privacy of Client
PoM	Privacy of Model

- Accuracy: % of **correct prediction** made by used PPDL
- Run time: the total **time of encryption**, sending data from client to server, and classification process.
- Data transfer: the **amount of data** transferred from client to server.
- PoC: **neither the server or any other party knows** about **client** data.
- PoM: **neither the client or any other party knows** about the classification **model** used in server.

KAIST

Comparison of HE-based PPDL

Crypto-PPML 2019

Scenario	Proposed Schemes	DL Technique	Accuracy (%)	Run Time (s)	Data Transfer (Mbytes)	PoC	PoM
Cloud Service	ML Confidential [30]	DNN	Bad (95.00)	Bad (255.7)	-	Yes	No
	Cryptonets [33]	CNN	Good (98.95)	Bad (697)	Bad (595.5)	Yes	No
	PP on DNN [34]	CNN	Good (99.30)	-	-	Yes	No
	E2DM [46]	CNN	Good (98.10)	Good (28.59)	Good (17.48)	Yes	Yes
Image Recognition	CryptoDL [28]	CNN	Good (99.52)	Bad (320)	Bad (336.7)	Yes	No
	PP All Convolutional Net [29]	CNN	Good (98.97)	Bad (477.6)	Bad (361.6)	Yes	No
Content Sharing	Distributed PP Multi-Key FHE [38]	CNN	Good (99.73)	-	-	Yes	No
PaaS	Gazelle [42]	CNN	-	Good (0.03)	Good (0.5)	Yes	Yes
	Tapas [43]	BNN	Good (98.60)	Good (147)	-	Yes	Yes
	FHE-DNN [44]	DiNN	Bad (96.35)	Good (1.64)	-	Yes	Yes

KAIST

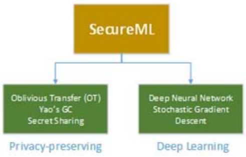
Crypto-PPML 2019

Secure MPC-based Hybrid PPDL

KAIST

MPC-based Hybrid PPDL(1/4)

- SecureML: A System for Scalable Privacy-Preserving Machine Learning
  - Based on **OT, Yao's GC, and secret sharing**
  - The sender of message remains oblivious
    - whether the receiver has got the message or not
  - **Linear regression and logistic regression**
  - Optimum value of regression?
    - Stochastic Gradient Descent (SGD)



P. Mohassel and Y. Zhang, ₩Secureml: A system for scalable privacy-preserving machine learning," pp. 19-38, 2017.

MPC-based Hybrid PPDL(2/4)

- Deepsecure: Scalable Provably-Secure Deep Learning
  - Use **OT and Yao's GC** protocol with CNN
  - Collaboration between client and server
  - **Weakness**: limited number of instance processed
  - Only able to classify one instance during each round

B. Rouhani, M. Riazi, and F. Koushanfar, "Deepsecure: Scalable provably-secure deep learning," 55th ACM /ESDA/IEEE Design Automation Conference, pp. 1-6, 2018.

MPC-based Hybrid PPDL(3/4)

- MiniONN
  - PP framework that transforms a NN into an oblivious NN.
  - Two kind of transformations:
    - piecewise linear activation function
    - oblivious transformation for smooth activation function
  - Supports all activation functions that have:
    - monotonic range
    - piecewise polynomial, or
    - can be approximated into polynomial function.

J. Liu, M. Juuti, Y. Lu, and N. Asokan, "Oblivious Neural Network Predictions via MiniONN Transformations," in Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, pp. 619-631, ACM, 2017.

MPC-based Hybrid PPDL(4/4)

- ABY3
  - PPDL framework based on three-party computation
  - Can switch between arithmetic, binary, and Yao's 3PC
  - Use binary sharing on three-party Garbled Circuit
  - Arithmetic sharing when training linear regression model
  - Outperform MiniONN by four order of magnitude faster

P. Mohassel and P. Rindal, "ABY 3: a Mixed Protocol Framework for Machine Learning," in Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, pp. 35-52. ACM, 2018.

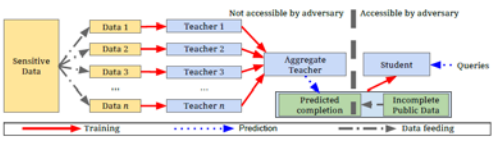
Comparison of MPC-based PPDL

Scenario	Proposed Schemes	DL Technique	Accuracy (%)	Run Time (s)	Data Transfer (Mbytes)	PoC	PoM
Cloud Service	DeepSecure [39]	CNN	Good (98.95)	Bad (10,649)	Bad (722,000)	No	Yes
Image Recognition	SecureML [35]	DNN	Bad (93.40)	-	-	No	Yes
PaaS	MiniONN [41]	NN	Good (98.95)	Good (1.04)	Good (47.60)	No	Yes
	ABY3 [45]	NN	Bad (94.00)	Good (0.01)	Good (5.20)	No	Yes

Differential Privacy-based PPDL

DP-based Hybrid PPDL

- Private Aggregation of Teacher Ensembles(PATE)
  - **Teacher phase and student phase**
  - Possible failure that reveals some part of training data



M. Abadi, U. Erlingsson, and I. Goodfellow, "On the protection of private information in machine learning systems: Two recent approaches," Computer Security Foundations Symposium, pp. 1-6, 2017.

Comparison-All

- E2DM gives the best performance:
  - High accuracy
  - Fast run time
  - Small data transfer
  - PoC
  - PoM

Proposed Schemes	DL Technique	Accuracy (%)	Run Time (s)	Data Transfer (Mbytes)	PoC	PoM
DeepSecure [39]	CNN	Good (98.95)	Bad (10,649)	Bad (722,000)	No	Yes
SecureML [35]	DNN	Bad (93.40)	-	-	No	Yes
MiniONN [41]	NN	Good (98.95)	Good (1.04)	Good (47.60)	No	Yes
ABY3 [45]	NN	Bad (94.00)	Good (0.01)	Good (5.20)	No	Yes

Proposed Schemes	DL Technique	Accuracy (%)	Run Time (s)	Data Transfer (Mbytes)	PoC	PoM
ML Confidential [30]	DNN	Bad (95.00)	Bad (255.7)	-	Yes	No
Cryptonets [33]	CNN	Good (98.95)	Bad (697)	Bad (595.5)	Yes	No
PP on DNN [34]	CNN	Good (99.30)	-	-	Yes	No
E2DM [46]	CNN	Good (98.10)	Good (28.59)	Good (17.48)	Yes	Yes
CryptoDL [28]	CNN	Good (99.52)	Bad (320)	Bad (336.7)	Yes	No
PP All Convolutional Net [29]	CNN	Good (98.97)	Bad (477.6)	Bad (361.6)	Yes	No
Distributed PP Multi-Key FHE [38]	CNN	Good (99.73)	-	-	Yes	No
Gazelle [42]	CNN	-	Good (0.03)	Good (0.5)	Yes	Yes
Tapas [43]	BNN	Good (98.60)	Good (147)	-	Yes	Yes
FHE-DNN [44]	DiNN	Bad (96.35)	Good (1.64)	-	Yes	Yes

Conclusion and Future Work

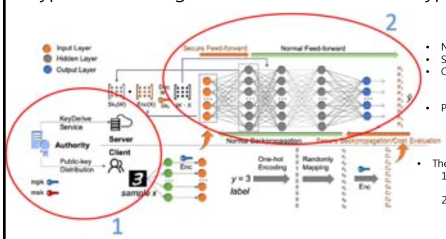
- Discussed state of the art of privacy-preserving deep learning
- Layers modified in PPDL:
  - **pooling layer, activation layer, and batch normalization layer**
- Future Work:
  - ❖Achieving more than 99% accuracy with good PoC and PoM
  - ❖Lots of Challenges still remain

Aminanto, Muhamad Erza, Rakyong Choi, Harry Chandra Tanuwidjaja, Paul D. Yoo, and Kwangjo Kim. "Deep abstraction and weighted feature selection for Wi-Fi impersonation detection." IEEE Transactions on Information Forensics and Security 13, no. 3 (2018): 621-636.

Q&A

Q&A

- CryptoNN: Training Neural Networks over Encrypted Data



- Newly published paper that improve E2DM
- Substitute HE with Functional Encryption (FE)
- Computation in:
  - HE: Encrypted form
  - CrpytoNN: Plaintext
- Performance: 95.50% accuracy

- There are 2 main parts:
  1. FE part
     - Authority manages key generation
  2. NN part
     - Feed-forward network
     - Backpropagation to minimize cost

Xu, Runhua, James BD Joshi, and Chao Li. "CryptoNN: Training Neural Networks over Encrypted Data." *arXiv preprint arXiv:1904.07303* (2019).