MPC於雲端AI的資安應用

組員:

林靖軒408410002
潘彥蓉408410045
吳承濬408410059
鄭宇軒408410062

指導教授:陳鵬升

目錄

一、專題摘要

研究動機

近年來隨著AI與雲端運算的發展，越來越多雲端服務會要求使用者將包含個人隱私的資料上傳至適當的伺服器，方能享受它們的服務，然而一旦將資料上傳至雲端就很難再保證該資料的安全，無法避免資料可能被雲端紀錄或洩漏。

為了在過程中保證資料的安全，我們選用多方安全計算**(MPC**, **secure multiparty computation)**來解決此問題，透過多方安全計算任一方無法取得其他參與者資料的前提來保證資料安全。

我們利用可進行多方安全計算的ABY framework [參考文獻6] 來建立 Fast Super-Resolution CNN (FSRCNN) model [參考文獻2]，來驗證我們的想法。

多方安全計算(MPC, secure multiparty computation)可以定義為在一個分布式網絡且不存在可信第三方的情況下，多個參與實體各自持有秘密輸入，並希望共同完成對某函數的計算並得到結果，前提是要求每個參與實體均不能得知除自身外其他參與實體任何輸入信息。

ABY是一個使用混合協議的多方安全運算框架，用來在不洩漏參與方隱私輸入的前提下合作計算出任意函數或成果，由協議表示為算術或 boolean 電路，並且可以使用 arithmetic sharing、Boolean sharing、Yao's garbled circuits。這些協議也可以自由組合，而ABY則使他們彼此間能有效轉換，如圖1所示。

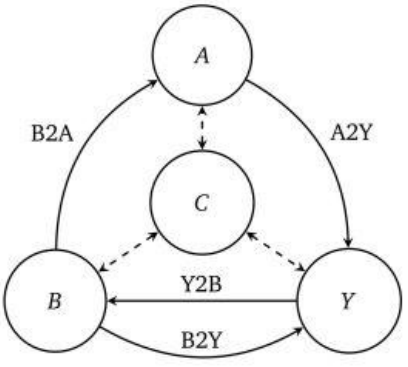


圖1

二、研究技術與方法

我們利用實驗室吳家瑋學長提供的 SHR wrapper[參考文獻1]去建構pre-trained FSRCNN model [參考文獻2]，目標是將**YUV420**標準的**qcif(176x144)**檔轉成**cif(352x288)**檔。

Fast Super-Resolution CNN，是SRCNN的改良版，用來將低解析率(LR)的圖像重建成高解析率圖像(HR)的一種技術如圖2所示，整個model可分為以下5個步驟:

1. 特徵提取 Feature Extraction:
   透過5x5的卷積核提取低解析率圖像的特徵
2. 收縮Shrinking:
   用1x1的卷積核進行降維，降低計算複雜度
3. 學習映射關係Non-linear Mapping:
   多個3X3卷積實現非線性映射
4. 擴展Expanding:
   因shrinking時降低LR的維度，故加入此層將其擴展到原本的維度使結果變好
5. 反卷積Deconvolution:
   將之前的特徵合併，使用upsamling將圖像變為HR

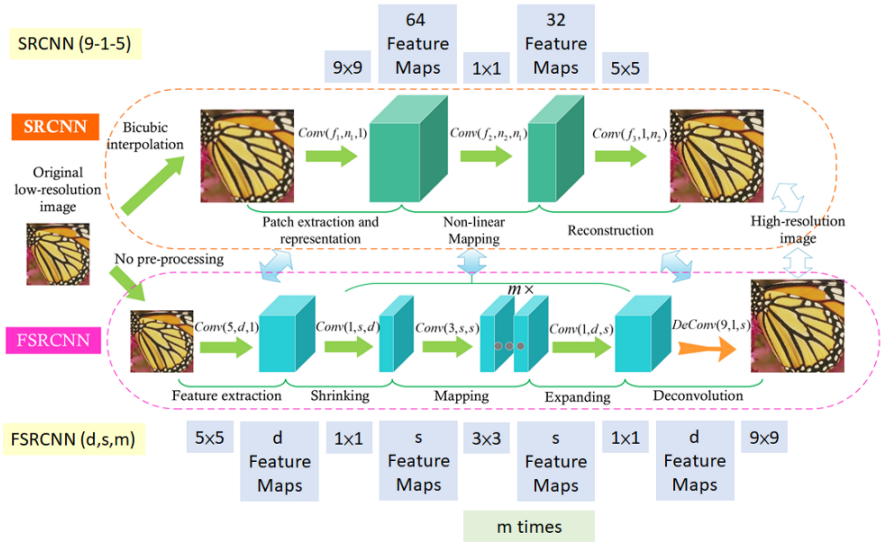


圖2

為了驗證多方安全計算在經過大量運算後，資料仍能保證其準確性，同時CLIENT端(YUV檔案)與SERVER端(AI模型的權重參數)均無法存取對方的資料。考量到時間成本與硬體限制，分為以下三個版本討論:

- 加速版(如圖3所示):只透過ABY進行資料交換，速度為三者最快，但資料安全性為三者最差，用來測試其極限速度。

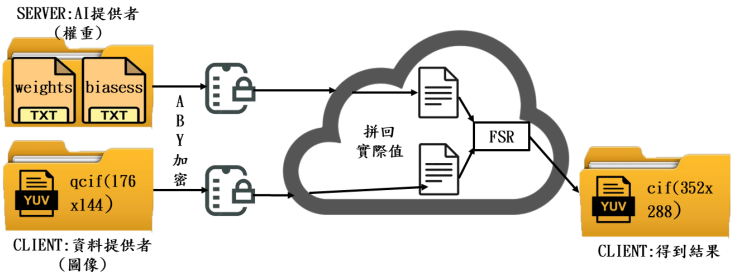


圖3

- 正式版(如圖4所示):使其中一個layer進行ABY計算，模擬大量運算後的結果，具有一定的安全性，驗證進行大量的多方安全計算後，不影響AI的結果。

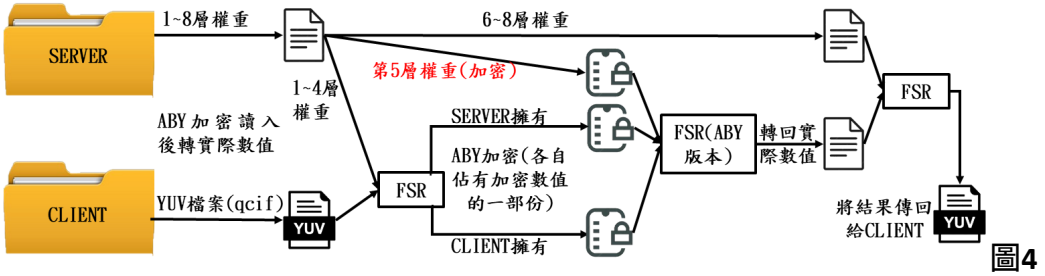


圖4

- 概念版(如圖5所示):所有的layer都進行ABY計算，可保證資料的安全，但由於預估耗時過久且記憶體不足等硬體限制，我們並未將此版本跑出。

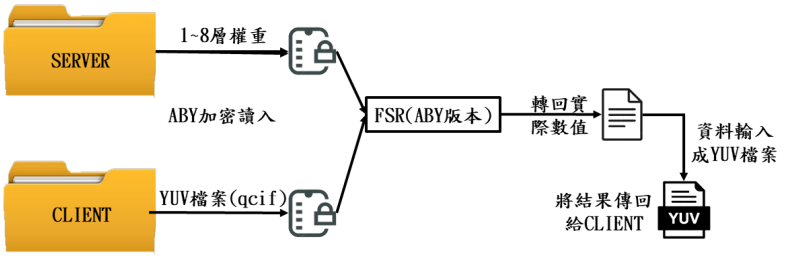


圖5

三、成果與討論

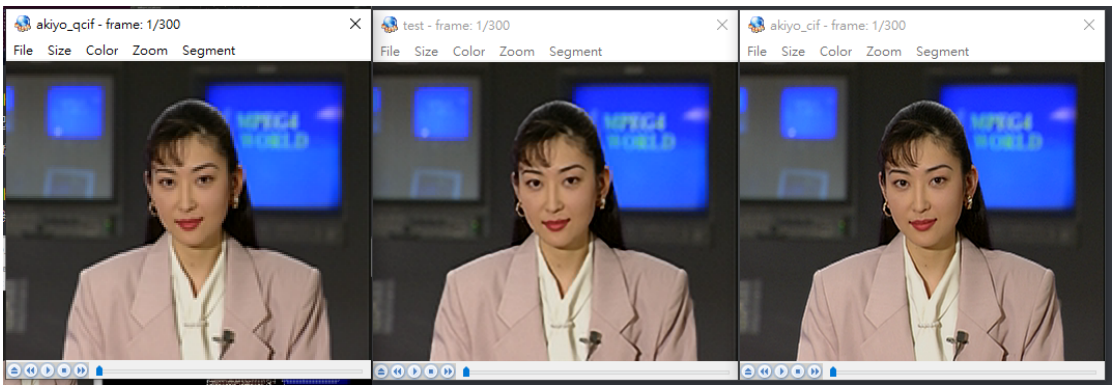
在實驗部分，我們使用一台電腦在不同的資料夾底下模擬CLIENT與SERVER兩端進行資料的ABY安全運算，並檢查是否能在保障安全性的情況下從中取得正確結果。

- 執行環境:
  - Ubuntu 20.04
  - 16GB RAM 10GB swap memory
  - CPU:**Intel(R) Core(TM) i7-9700 CPU @ 3.00GHz 8C8T**

- 需安裝的軟體:
  - make
  - cmake
  - git
  - **Install C compilers**
  - g++ 9.0

  Install libraries:
  - libssl-dev
  - libgmp-dev
  - libboost-all-dev
  - **Clone ABY** https://github.com/encryptogroup/ABY.git --recursive

- **原始程式(無ABY加密版):**
  執行時間 : **2sec / 1frame** 直接將資料交出，沒有安全性



原圖ZOOM in 2倍　　轉換結果　　原圖(高清版)

- **ABY加速版:**
  執行時間: **1min / 1frame**　與原版相比多了些雜訊


轉換結果　　原轉換結果

- **ABY正式版:**
  執行時間: **16days / 1frame** 與加速版相比再多了些雜訊


轉換結果　　原轉換結果

- **ABY概念版:**
  執行時間: **1year/ 1frame**　由於耗時過久，並未將結果跑出。

結論

經比較，雖然我們所設計的3個版本花費的時間較原版本多了數倍，然而其呈現結果與原圖比較僅有些許差異，且此差異可視為誤差。透過部分程式流程進行安全運算，依然能使另一端難以透過逆向工程的方式去取得原始資料，故其資料安全性仍可以保證，證明MPC具有應用在雲端AI的潛力。

四、達成之效益及未來方向

- 達成效益:
  期望透過此次範例拋磚引玉，使多方安全運算能用於加密雲端計算領域。

- 實用性、未來方向:
  在未來的實用性上，例如各種雲端計算、車用電子系統、隱私推薦系統...等領域都可以應用。故我們期望，更多的人參與到多方安全運算的應用中，鼓勵社會保護數據隱密但流通，且使社會總體利益最大化。

  目前受限於多方安全運算之程式本身需要的資源較多，且沒有受到硬體加速的支援，導致執行時間過慢，無法於實際的場域執行。期望在速度增加後，能夠更廣泛的被大家所應用，並作為未來雲端運算的標準加密資訊交換方法之一。

五、團隊分工方式

在製作過程中我們遇到許多難題，謝謝許多人的幫助讓我們能完成，在此感謝指導教授 陳鵬升教授以及吳家瑋學長和所有幫助過我們的人。

Work	鄭宇軒	林靖軒	潘彥蓉	吳承濬
程式流程設計	V			
副函式	V	V	V	
正式版	V	V		
加速版	V			
概念版	V	V	V	V
程式檢查	V	V	V	V
簡報製作	V	V	V	V
貢獻百分比	35%	30%	25%	10%

六、參考文獻與網站

[1]吳家瑋 **,"A General-Purpose C Compiler for Secure Two-Party Computation ," 2022.**

[2]Milad Abdollahzadeh
Implementing Deep Convolutional Neural Networks in C without External Libraries | by Milad Abdollahzadeh | Towards Data Science , 2021.

[3] B. Pinkas, T. Schneider, N. P. Smart, and S. C. Williams, "Secure two-party computation is practical," 2009.

[4] M. Franz, B. Deiseroth, K. Hamacher, S. Jha, S. Katzenbeisser, and H. Schröder, "Secure computations on non-integer values," 2010 IEEE International Workshop on Information Forensics and Security, 2010.

[5] A. C.-C. Yao,"Protocols for secure computations,"23rd Annual Symposium on Foundations of Computer Science (sfcs 1982), 1982.

[6] D. Demmler, T. Schneider, and M. Zohner, "Aby - a framework for efficient mixed-protocol secure two-party computation," 2015.

[7] Chao Dong, Chen Change Loy, and Xiaoou Tang,"Accelerating the Super-Resolution Convolutional Neural Network,"2016