

# 第六回留学報告書

Funai Overseas Scholarship 2020 年度奨学生  
古賀樹

2022 年 12 月

2020 年度から University of California San Diego の Computer Science 専攻 Ph.D. 課程に在籍している古賀樹と申します。夏のインターンが終了し Ph.D. 生活の 3 年目が始まりました。この報告書では夏のインターンと 3 年目秋学期の研究や生活についてのご報告をさせていただきます。

## 1 研究

### 1.1 インターン

夏学期は Apple でインターンをし、3 ヶ月間プライバシー保護下での Federated Learning (連合学習) に取り組みました。Apple では iPhone などのデバイスに貯蔵されているデータを用いて機械学習を行うアプリケーションが多々存在します。しかし、データが Apple のサーバーに保存されることはプライバシーの観点で望ましくありません。そこで、データを各デバイスから出さずに学習を行うことのできる Federated Learning が用いられます。(更なるプライバシー保護のためにはさらに中央サーバーでノイズを加える必要があります。)

インターンの期間中は、学習対象のデータを保持しているデバイスが少ない際に学習が極端に遅く、また結果が悪化する問題を解決する手法に関する研究を行いました。Federated Learning の設定の中で対象に似ていて保持デバイスが多いデータをうまく活用することで、学習速度の大幅な向上と精度の向上を達成しました。論文のドラフトを秋学期に書き終えたので、年明け締切の会議に提出予定です。

インターンでは論文として成果を出すことができただけでなく、大企業ならではの問題設定、課題解決の道筋を体感できたことが個人的な収穫でした。今

回の研究を大学で行おうとしたら数十倍のリソース(時間、計算資源、人)が必要になっていたでしょう。また組織がどのように運営されているか、特に PhD をとった人材がその経験、知識をプロダクトに還元する仕組みを間近で見ることができたことも収穫の一つでした。来年の夏は Microsoft でインターンをするようになりました。企業間でどのように文化が違うのかを体感するのが楽しみです。

### 1.2 大学

秋学期は走っている二つのプロジェクトのうち、複数機関のデータを用いる設定での因果推論におけるプライバシーの研究に注力しました。夏前には、複数機関からのデータや推定値をプライバシー保護下で集計する方法に焦点を当てていました。その際、各機関での推定は比較的シンプルな設定のもと行われることを仮定していました。

一方秋学期は、提案手法の汎用性を示すため、因果推論でよく用いられる Matching を用いた推定方法でも同様な集計を行えることを示しました。因果推論において推定したい対象として、Average Treatment Effect (ATE) があります。例えばある薬の効果調べる際に、薬を投与したときと投与しなかったときに結果が平均的にどの程度異なるかを推定するときに ATE を用います。ここで問題になるのは、あ

る個人については薬を投与したときと投与しなかったときのどちらかの結果しか観測できない、という点です。Matching では薬の投与の有無が異なる似た個人（例：年齢や性別が同じ）の結果を観測できなかった側の結果として扱うことでこの問題を解決します。

今回はさらにプライバシー保護下での Matching を考えます。一般にプライバシー保護を行うためには、一人のデータ有無を打ち消す程度のノイズを最終的な推定に加える必要がありますが、Matching は用いるデータによっては一人のデータが最終的な推定に大きく影響することがあります。通常加えられるノイズは最悪ケースのデータを元に計算しますが、その場合 Matching ではノイズが支配的になってしまいます。そのため、今回は推定に用いられるデータ依存のノイズを加える手法を採用し、プライバシー保護下での Matching の有用性を大幅に改善しました。現在執筆中の論文は年明け締め切りの国際会議に提出予定です。

## 2 授業

秋学期は CSE207B: Applied Cryptography という講義を取りました。これで卒業要件を満たすまで授業は残すところ一つとなりました。

講義の中では Security の定義と暗号技術がそれぞれどのように満たし、実装がどのように行われているかを学びました。ユニークだったのは課題の構成です。毎回、まずは講義で扱った暗号の脆弱性を突いて暗号の復号をコーディングで行います。すると残りの課題（主に証明）が記載された PDF が手に入ります。実際に手を動かすことで暗号の脆弱性について理解を深めることができる一方、その作業が長引くと残りの宿題の内容すらわからず工数の見積りもできないため、非常にストレスフルでした。内容としては満足だったので、将来 UCSD の CSE に来る人はぜひ比較的余裕のある学期に受講することをお勧めします。

## 3 生活

インターン期間中は奨学生の方々が同じベイエリアに集まっていたこともあり、食事や旅行を楽しみました。秋学期は週末を使ってインターンの成果をまとめていたので、これまでほどアクティビティには時間を費やしませんでした。夏と秋学期の間にはロードトリップで Grand Canyon や Sedona を訪れました。自然の満喫はリフレッシュ方法として非常に気に入っているので、在学中に西側の国立公園はできる限り回りたいところです。



図 1: Grand Canyon



図 2: Sedona

また秋学期の途中からは University of California 全体の大学院生・ポスドクによるストライキが行われました。目的は全体的な待遇の改善で、その中には物価上昇に伴う賃金の上昇も含まれています。実際、物価・家賃は上がるばかりでこのままだと PhD 学生としても金銭的に UC に進学しにくくなる未来も想像ができます。CS の学生は夏のインターンな

どの抜け道がありますが、そうでない場合は QoL を維持するのが難しくなっていくでしょう。ストライキに参加する学生の数を見て、自らの権利を声を大にして主張するという点で、アメリカに来てから最も日本との文化の違いを感じた期間でした。

## 4 最後に

PhD 生活 3 年目になり、徐々にスキルセットの成熟と研究の加速を感じながら充実した生活を送っています。ストライキが起こるような状況の中でも財団の支援のお陰で研究に集中することができています。改めて様々な面での支援をしてくださっている財団の皆様に心よりお礼を申し上げます。