【ネットワーク LSI システム（藤野・吉田）研究室】
～ハードウェア＆AI セキュリティ～

RITSUMEIKAN NTSOC LAB.

研究室の基礎データ

所在地：ローム記念館3F 南側（RO301　藤野個研室，並びでエレベータに向かって第１～第４研究室）
Web ページ：http://www.ritsumei.ac.jp/se/re/fujinolab/
連絡先：【教授　藤野　毅】fujino@se.ritsumei.ac.jp 【助教　吉田　康太】y0sh1d4@fc.ritsumei.ac.jp
2025 年度予定学生数：修士 2 回生：9 名　修士 1 回生：4 名　学部生：本年度配属予定者

デジタル AI 社会と研究テーマ

2000 年以降，インターネットとスマホの普及で社会のデジタル化は急速に進んできました．今後はこれに AI 技術が加わって，生活や仕事も大きく変化していくと予想されます．図1に情報データの流れの観点から見たデジタル AI 社会の姿を示しますが，サイバー空間（インターネットおよびその上にあるクラウド計算資源）と物理空間（皆さんが生活する場所にあるスマホ，自動運転車など）のコンピュータが相互にデータ通信を行い AI を使ってデータの解析を行います．解析した結果を物理空間にフィードバックして役立てることで，豊かな未来社会を実現しようというのが，政府が提唱した「Society5.0」という考え方です．最近ではこの考え方を拡張し，「デジタル田園国家都市構想」という名称で，デジタル AI 技術を活用して地方を含めた日本全体を活性化することを目指しています．しかしながらこのような社会を実現するためには，悪意あるサイバー攻撃や障害に伴う脅威からこれらの技術を守り，安全に運用する技術の開発とそれに向けた人材の育成が必須です．

当研究室では，電子機器間のデータの暗号通信および AI によるデータの解析時のセキュリティに焦点を当てた研究を行っています．具体的には暗号通信における暗号鍵の生成や安全な管理（ハードウェアセキュリティ）や AI 処理時のセキュリティやプライバシーに焦点を当てて以下のような研究を行っています．

(1) ハードウェアセキュリティ
　例） PUF 技術．サイドチャネル攻撃技術
(2) AI セキュリティとプライバシー
　例） 敵対的サンプル，連合学習

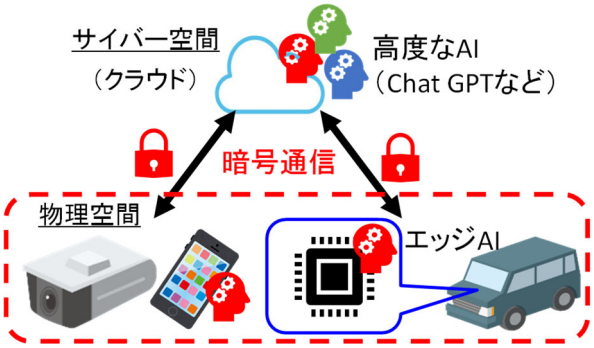

図１．デジタル AI 社会の姿

研究テーマ概要の解説

(1) PUF（Physical Unclonable Functions）技術：物理空間に置かれた多数の電子機器にユニークな（他の機器とは異なる）暗号鍵を安全に格納するために有効な技術です．電子機器に使用されている LSI（集積回路）は与えられたプログラムやデータが同一であればどのチップでも同じ動作をするように作られています．一方 PUF は図2に示したように LSI のチップごとの製造ばらつき（トランジスタの性能や配線の寄生容量など）を利用して，同一の入力に対してチップごとに異なる出力（レスポンス）を得ることができます．このレスポンスは人間の指紋のように複製できない固有の ID となっていることから，暗号鍵の生成に利用できます．従来使われていた，外部で暗号鍵を生成して不揮発性メモリに保管する手法と比較して，PUF 技術では低コストで複製が困難な鍵生成を行える点で優れています．一方で，生成したレスポンスを安定化する技術や，過去のレスポンスから未知のレスポンスを予測する攻撃に対抗する技術が必要で，誤り訂正や深層学習を用いた攻撃の研究を進めています．

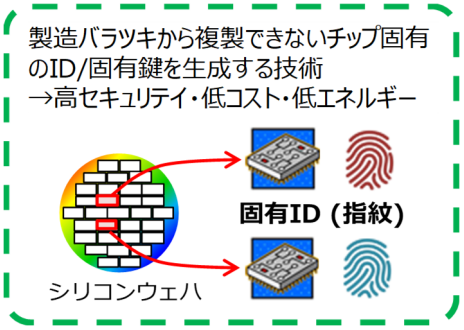

図２．PUF 技術

(2) サイドチャネル攻撃技術:暗号通信において秘匿が必要なデータ(平文)は暗号鍵を用いて暗号データに変換されますが，攻撃者が暗号鍵を知っていれば，容易に暗号データを解読することが可能です．図3に示したように，現在広く使われている暗号アルゴリズム(AES 暗号など)では，攻撃者が平文と暗号文のペアを集めても暗号鍵を全数探索しない限り暗号鍵推定できないように作られています．しかしながら，暗号処理を行っている際の処理時間・消費電力，漏洩電磁波(これらをサイドチャネル情報といいます)を使って暗号鍵を推定する方法がサイドチャネル攻撃です．物理空間に配置される機器では攻撃者がサイドチャネル情報を入手することが容易であるため，対策が必要です．対策として，乱数を使って暗号鍵とサイドチャネル情報の相関をなくす「マスキング」などが有効とされていましたが，深層学習を使って，マスキング対策を無効化できるため，これらの対策手法の研究を行っています．

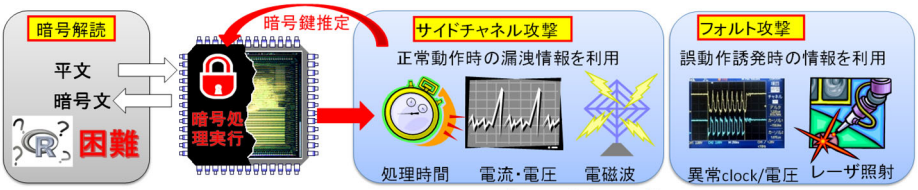

図3．サイドチャネル攻撃技術

(3) AI のセキュリティとプライバシー保護技術:図4に示したように，AI では大量のデータから推論を行うための学習モデルを作成し，そのモデルを用いて推論を行います．この学習および推論のプロセスの中で，(a)～(d)に示すような様々なセキュリティ脅威が存在し，それらの脅威の評価や対策が必要です．例えば，図中(b)の脅威事例として敵対的サンプル(Adversarial Examples)と呼ばれる画像改ざん手法があり，人間にはわからないほどの巧妙なノイズを画像中に印加することで，AI の推論を誤らせることが可能であり，画像を解析して運転支援を行うシステムにとって大きな脅威となります．また，AI の応用分野が広がり，医療情報などを扱う場合には，学習データ中に個人を特定する情報が含まれないようにする必要があります．このための技術として図5に示すようなサーバーにデータを集約せずにモデルを学習させる連合学習や匿名化などが使われるようになりました．このような AI のセキュリティやプライバシー保護に関する研究を行ってきましたが，今後生成 AI を用いた応用技術や著作権保護などについても検討をすすめていきたいと考えています．

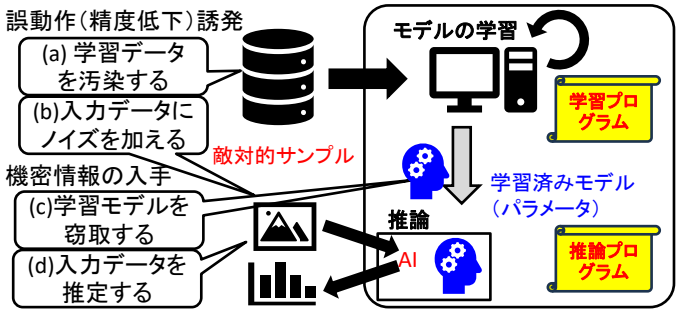

図4．AI のセキュリティ脅威

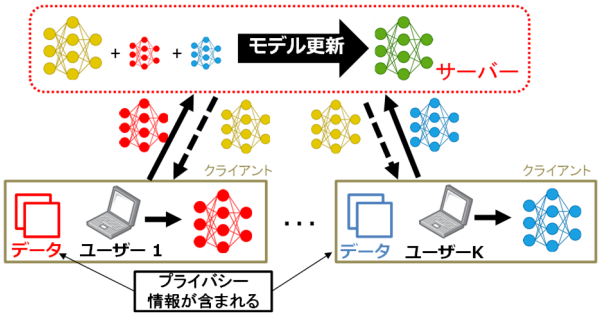

図5．連合学習技術

応用演習・卒業研究の進め方

応用演習では，機械学習の基礎や深層学習の入門知識を，Python を用いてプログラミング演習しながら学ぶとともに，就職活動に向けた研究説明等のプレゼン演習を行う予定です．4 月頃から，暗号技術に関する本の輪講を週 1 回行いつつ，大学院生と一緒に，PUF シミュレーション，サイドチャネル攻撃，深層学習プログラミングなどの実践的な実習を行います．就職活動が収束した学生は卒業研究テーマに本格的に着手し，7月の上旬に発表会を行います．夏休み以降は自主的に研究を進め，12 月の発表会を経て卒業論文を作成します．

2025 年度 4 月に当研究室に本配属された学部生が大学院に進学したい場合は，大学院から他の研究室に移籍することが必要となります．このため，原則として学部で卒業する予定の学生が当研究室を志望してください．学部生の間の 1 年間とはなりますが，(1)深層学習技術，(2)暗号などのセキュリティ技術，(3)FPGA ボードや回路などの LSI 設計技術などを学びたい，(4)生成 AI などの新しい技術に触れてみたい，というような意欲があり，デジタル AI 社会で活躍できる人材を目指して，卒業まで真面目に努力をする学生を歓迎します．