# 先端科学技術研究科　修士論文要旨

| | | | |
|---|---|---|---|
| 所属研究室<br>（主指導教員） | サイバーレジリエンス構成学<br>（門林　雄基　（教授）） | | |
| 学籍番号 | 2411284 | 提出日 | 令和 8年 1月 18日 |
| 学生氏名 | 村田　百合菜 | | |
| 論文題目 | エッジにおけるドリフトとデータ汚染攻撃の判別によるセキュアなモデル学習の提案 | | |

要旨

エッジコンピューティング環境では，環境変化に起因するデータドリフトに加え，悪意のあるデータ汚染攻撃が発生し得る．そのような状況下で従来のドリフト検知手法を適用すると，自然なドリフトと汚染攻撃に起因するドリフトを区別できず，誤って汚染データを用いてモデルを再学習してしまう危険性がある．本研究では，この問題を解決するため，モデル更新前に汚染攻撃によるドリフトを識別し，エッジ環境における安全なモデル再学習を可能にする手法を提案する．提案手法は，汚染攻撃によりIoTデバイス間およびデバイス内部のセンサー間の相関構造，ならびに時間的な依存関係において特有の変化が生じるという仮説に基づく．まず，デバイス間の相関と相互相関に基づくラグ（相互相関関数が最大値をとる時間遅れ）の変化量を算出し，閾値判定により汚染攻撃によるドリフトの発生を検知する．次に，欠損率を考慮した上で，Ward法による階層的クラスタリングとデバイス内部センサー間相関の変動量を統合し，攻撃を受けたデバイス群を自動的に特定する．実験の結果，過半数のデバイスが広域に汚染される状況においても，汚染源となるデバイス群を再学習から除外し，モデル性能の劣化を抑制できることを確認した．これにより，エッジコンピューティング環境におけるセキュアなモデル再学習の実現可能性を示した．