生物進化モデルによるサイバー脅威の発見
系統樹を用いた攻撃進化の可視化

SecHack365 思索駆動コース　幸松豊拓

Introduction

背景

サイバー攻撃は高度化・多様化しており、
従来の「検知→対応」の
事後対応セキュリティでは限界がある。
特に0-day攻撃に対しては、既存の攻撃発生から検知して対策をすることは無力である。

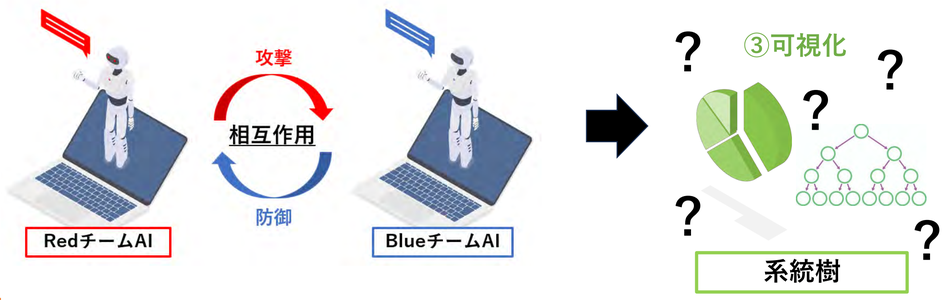
課題

- 攻撃者より「後」に対策する構造
- 既知パターンにしか対応できない
- 新しい攻撃手法の発見が人間に依存

生物学の知識を応用して、サイバーセキュリティの課題を解決したい

Purpose

目的　生物進化の系統樹モデルは、サイバー攻撃の進化を可視化・予測できるか？

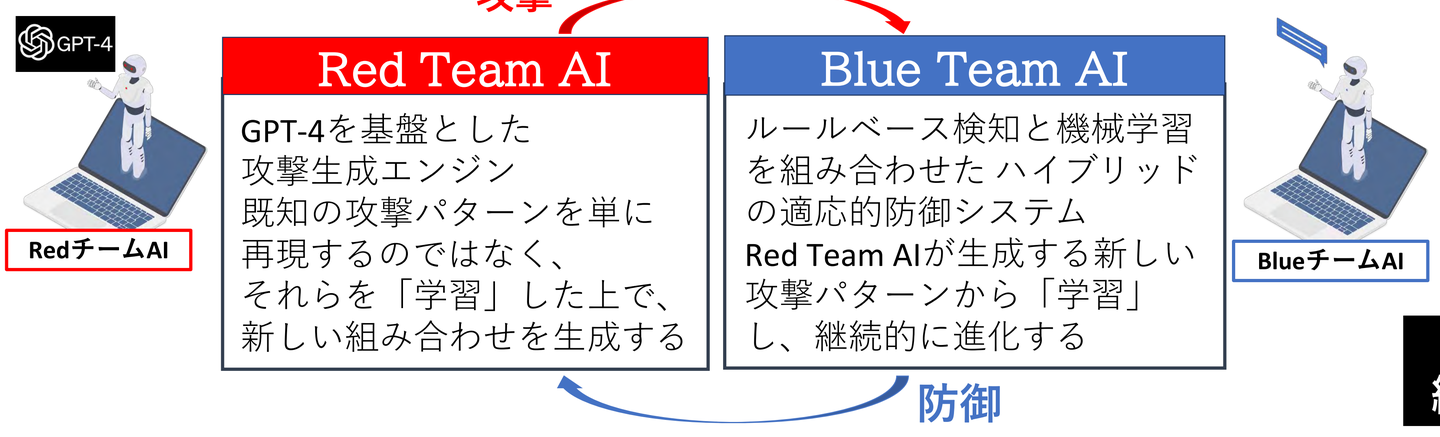


3つの研究課題

Q1: AI同士の共進化で未知攻撃を生成できるか？
Q2: 生成された攻撃は本当に「新しい」のか？
Q3: 攻撃ドメインを変えるとパターンも適応的に変化するか？

Methodology

システム全体アーキテクチャ

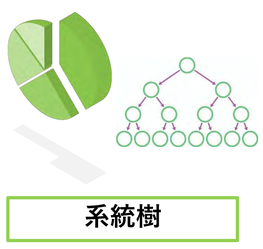


①と②が攻撃防御を繰り替えすことで共進化

Visualization

生物学における系統樹を
サイバー攻撃に応用した可視化
既知の攻撃パターンから、
どのように新しい攻撃が
「進化」したかを、
視覚的かつ定量的に追跡

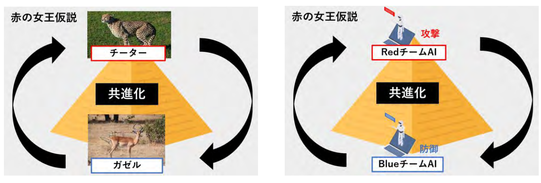
③可視化



研究課題を検証する3つの方法論

Q1. 未知攻撃の生成

生物学における捕食者-被食者の共進化モデルを応用し、
Red Team AI（GPT-4、攻撃生成役）と
Blue Team AI（機械学習、検知役）を対峙させた。
・Red Teamは既知攻撃150パターンを学習後、
新しい組み合わせを20秒間隔で生成。
・Blue Teamは検知失敗した攻撃を訓練データに追加。
時系列での検知率変化、生成された攻撃の世代数、
進化タイプ（組み合わせ・派生・変異）を測定することで、
共進化メカニズムを検証した。



Q2. 新規性の検証

Step1:既知攻撃DBとの 編集距離を計算し、
類似度85%未満を 「別種候補」
Step2:コマンドの構文的特徴12次元を抽出し、
特徴空間上での距離を測定。
Step3: 主要のセキュリティツールで検知テストを実施。
これら3指標を統合した 新規性スコア（0-1）を算出し、
0.7以上を新種（0-day級）と 分類

Q3. 汎用性の実証

ドメイン切り替え→各ドメインでの実験実施
この実験のドメイン間比較分析を行うことで、
進化パターンの可視化比較定量化

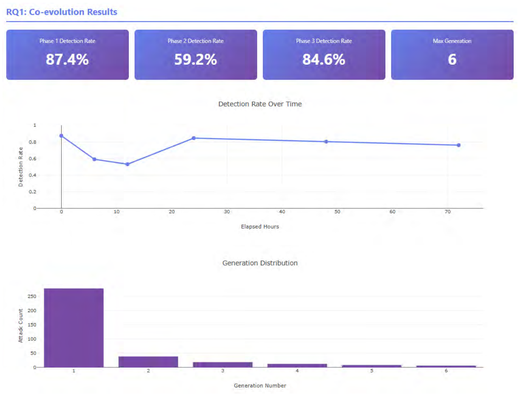
各ドメイン：
[PowerShell]
[Web App]
[IoT/OT]
[Cloud]

- Branch factor (分岐係数)
- Tree depth (樹の深さ)
- Leaf count (葉ノード数)

Googleの画像生成AIを使用して作成

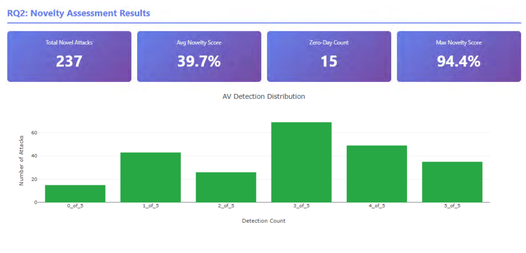
Results

3つの課題の結果



共進化を確認
検知率の一時的低下 → 攻撃の進化
その後の回復 → 防御側の適応
6世代にわたる継続的な進化

進化的アプローチは有効
攻撃・防御の共進化を定量的に実証
系統学的指標で新規性を評価
セキュリティ研究への生物学的視点の応用可能性を示唆



系統的多様性を確認
Zero-Day: 4.1% → 検知隔離に成功
検知分散 → 多様な進化経路
高新規性スコア → 既知パターンからの乖離

適応放散は限定的
ドメイン間で一様な進化パターン
統計的有意差なし (p>0.05)
進化メカニズムはドメイン非依存

Outlook

まとめ

・生物学 × サイバーセキュリティ
→ 従来の分野境界を越えた融合研究ができた
本研究は、サイバーセキュリティ分野に生物学的進化の視点を導入し、攻撃と防御の共進化を定量的に実証した。進化系統樹と系統学的指標により、攻撃の新規性と進化動態を可視化・評価する枠組みを確立した。

今後の発展の方向性

長期実験: 数週間で深い進化観測
実環境適用: テストデータ → 実攻撃データ
自動防御: 検知パターンの自動更新

NICT National Institute of Information and Communications Technology

https://isome-lab.com/