「マルチメディア，分散，協調とモバイル
(DICOMO2017)シンポジウム」 平成29年6月

人工知能搭載型サイバーレンジによるシステム強靭性の検討

中山能之[1] 宮本貴義[2] 大石 恵輔[1] 岩東 佑季[1] 八槇 博史[2]

概要： 標的型攻撃を企業や組織のネットワーク環境で事前検証し，想定される被害や必要となる対策を調査することが望ましい．本研究室で開発してきた，サイバーレンジと共進化シミュレーションシステムから構成される人工知能搭載型サイバーレンジを用いてこれを実現する．標的型攻撃の対策である入り口対策では，攻撃を防げていない．こうした背景から，攻撃の早期検知が重要である．そのために，特定の端末・サーバを重点的に監視し攻撃被害を最小限にとどめることを目指す．アプローチとして，グラフ理論を用いた故障対策におけるサービス継続性の考えをサイバーセキュリティに適応させた，レジリエンス（強靭性）分析を提案する．本論文では，レジリエンス分析を行うためのシステム構成および手法を示す．

A Study on System Resilience by Cyber Range with Artificial Intelligence

Yoshiyuki Nakayama[1] Takayoshi Miyamoto[2] Keisuke Oishi[1] Yuki Iwato[1]
Hirofumi Yamaki[2]

1. はじめに

標的型攻撃による，複雑なサイバー攻撃を企業や組織のネットワーク環境で事前検証し，想定される被害や必要な対策を調査することが望ましい．

今後，発生し得るサイバー攻撃として，人工知能技術を用いて能力が向上されたマルウェアの登場が予想される．この予測のもと，人工知能技術を積極的に活用し，攻撃者に先回りする形でプロアクティブな対策を可能にする研究・開発を実現するための Super-LIFT システム[1]の開発が進行中である．本研究は，この一環として行っている研究の 1 つである．

こうした研究を行うに際して，サイバーセキュリティの研究における人工知能の利用方法を 2 つに分類した．

1 つ目は，「AI が守る」利用方法が挙げられる．近年，普及してきた人工知能を搭載したセキュリティ製品やセキュリティ管理者を支援する LIFT システム[1]がある．既存の攻撃者が対象となる．

2 つ目に，「AI から守る」利用方法が挙げられる．先ほど述べたように，人工知能を利用したマルウェアの出現が予測される．これに対抗するために，著者らが開発したサイバーレンジ（後述）と共進化シミュレーション[2]システムで先回り対策をする．共進化シミュレーションでは，攻撃を行う AI システムと防御を行う AI システムとをそれぞれ遺伝的アルゴリズムなどの進化計算を用いて自動生成し，仮想計算機と SDN とを用いて構成されたネットワークシステム上で対戦させることによって，将来的に起きうるサイバー攻撃を予測する．

次に，一般的なサイバーレンジ[3]と，本研究で扱う人工知能搭載型サイバーレンジとの違いを述べる．一般的なサイバーレンジは，サイバー攻撃を模擬して，人間の訓練のために利用することを前提にしたシステムであることが多い．一方で，人工知能搭載型サイバーレンジは，人工知能の訓練(学習)を前提としたシステムである．人工知能による学習は防御 AI と攻撃 AI の 2 つに分類される．防御 AI の場合，LIFT システムがこれに該当する．具体的には，セキュリティ管理者を支援するために必要な，攻撃事象を学習させる．攻撃 AI は先ほど述べたように，今後登場する事が予測される AI により能力が向上したマルウェアがどのような学習をしてくるか模擬する．人工知能搭載

1 東京電機大学情報環境学研究科 印西市
2 東京電機大学情報環境学部 印西市

型サイバーレンジでは，両者が学習する環境を模擬する．つまり，人工知能搭載型サイバーレンジは，防御 AI と攻撃 AI がサイバー攻撃を模擬することで，起こりうる事態の予測ができる．そこで得られるシミュレーションログから学習を行うことにより，予測されたサイバー攻撃に対抗するための防御システムの構築も可能となる．

人工知能搭載型サイバーレンジは，図 1 に示すような運用をイメージしている．

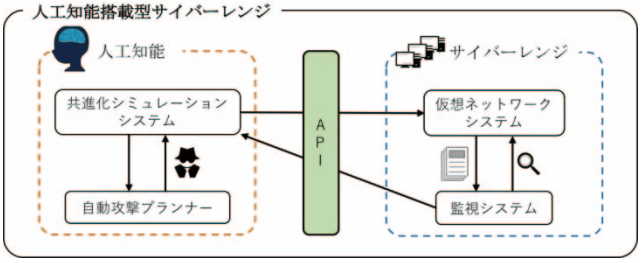

図 1 ： 人工知能搭載型サイバーレンジの運用イメージ

人工知能搭載型サイバーレンジは，大きく分けてとサイバーレンジと人工知能の 2 つに分けられる．1 つ目は，仮想ネットワークシステム[4]と監視システム[5]からなるサイバーレンジである．仮想ネットワークシステムは，NSDL(Network System Description Language)[2]に記述されたネットワーク構成に基づいてネットワークを展開する．ネットワークの展開には，Docker および OpenvSwitch を用いる．展開されたネットワーク上で攻撃実験を行う．その結果を自動集取するために，Zabbix を用いて監視システムを開発した．この 2 つで構成されるものが，図 1 のサイバーレンジにあたる．

2 つ目に人工知能として，攻撃の自動実行をサポートする自動攻撃プランナーと，シミュレーションログからの学習を実現するシステムとして，共進化シミュレーションシステムを搭載した人工知能を利用する．また，人工知能とサイバーレンジ間の制御は API で行う．人工知能搭載型サイバーレンジは，インスタンス事例を扱う．問題として，大量の事例が存在するため，インスタンス事例が発散してしまう．この問題に関しては 6 章で解決案を示す．

本研究は，標的型攻撃の対策である入り口対策で攻撃を防げていないこと．さらに，手口がより狡猾になっており，有効な対策を講じる事ができていない．こうした背景から本研究では，攻撃を水際で防ぐのではなく早期に検知することを重要視する．

早期検知のために特定の端末・サーバを重点的に監視し，攻撃被害を最小限にとどめることを目指す．また，ネットワーク上の特定の端末・サーバを重点的に守ることでセキュリティ対策のコスト削減を目指す．そのために，人工知能搭載型サイバーレンジで攻撃実験を行い実験結果の集取が必要である．集取した実験結果を利用して防御対策を施していく．

防御策として，グラフ理論を適用した故障対策におけるサービス継続性の考えを導入し，レジリエンス分析を検討した．例として，ある防御対策をネットワークで適用すると，ネットワークはより安全になる．しかし，今まで利用していたサービスが利用できなくなる可能性がある．この問題は安全性と利便性のトレードオフの関係である．攻撃事例毎に，トレードオフによる影響が少ないネットワーク構成をレジリエンス分析で見つけ出す．本論文では，レジリエンス分析のシステム構成，分析・評価手法および今後の展開を示す．

2. レジリエンス（強靭性）分析

すでに，グラフ理論を適用した故障対策の研究として，故障時でもサービスを継続可能な高信頼リンクの決定法[6]がある．こうしたグラフ理論や故障時対策の考えをサイバーセキュリティの研究に適応する．これによって，新たな防御対策としてレジリエンス(強靭性)分析を提案する．

本研究におけるレジリエンス(強靭性)とは，人工知能搭載型サイバーレンジ上に展開されたネットワークが攻撃に耐性を有し，提供されているサービスが継続可能な構成であるかを意味する．

レジリエンス分析を実施するためには，実験で集取したログデータが必要となる．このログデータを利用して，多くの攻撃を受けたサーバ・端末を見つけ出す．そして，既存のネットワーク構成から，端末・サーバの停止または，経路(回線)の削除を行い新たなネットワーク構成を生成する．こうした変更が行われた場合でも，提供中のサービスが継続可能か評価する．この一連をまとめて，レジリエンス分析とする．このレジリエンス（強靭性）分析を利用するサイバーレンジの運用手順を以下に示す．

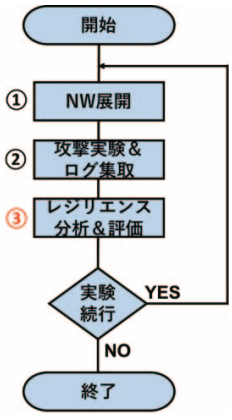

図 1：サイバーレンジ運用手順

サイバーレンジの運用手順は図 2 のとおりである．レジリエンス（強靭性）分析を行うまでに，サイバーレンジ上に実験対象のネットワークを展開し，攻撃実験を行いそのログを集取する(①, ②). この手順までに集取したログデー

タを元にレジリエンス（強靭性）分析を行う(③)．その後，分析結果を評価(③)し，攻撃と防御の共進化シミュレーション機構が実験継続の判断をする．

3. グラフ理論に基づくレジリエンスの評価

グラフ理論の点連結度および辺連結度を用いて評価する．サービス継続性に基づくとレジリエンスの評価は，端末・サーバの故障と経路(回線)故障の強靭性を評価する．だが，本研究はサイバーセキュリティの研究に故障対策におけるサービス継続性を導入したものである．これは，広義のセキュリティである．

広義のセキュリティとは，意図しない障害と攻撃に分けられる．はじめに，意図しない障害とはネットワーク上の端末・サーバの故障や経路(回線)故障を意味する．次に，攻撃とは，端末・サーバに対するクラッキング，経路(回線)への DoS 攻撃がある．これらの両方を評価対象とするのが本研究における，レジリエンス評価となる．

ネットワークをグラフと考えると，点連結度は端末・サーバの停止に相当する．また，経路の削除が辺連結度に相当する．それぞれの手法で，端末・サーバの停止，経路の削除というモデルにすることができる．端末・サーバが停止，経路の削除がされたときネットワークへの攻撃はどうなるかという問題がある．これをレジリエンスで評価する．実際にそれぞれの概念を以下に示す．

はじめに，点連結度の概念を適用するとグラフは図 3 のようになる．

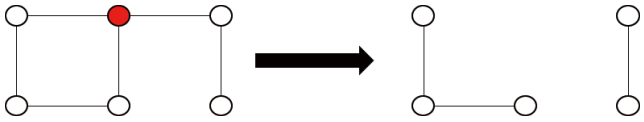

図 3：点連結度によるグラフの変化

図 3 のグラフは，ある点を削除したことで 2 つのグラフに分断された．これは，ネットワーク上のある端末・サーバの停止をした状態になる．これにより，特定の経路が利用できず，ネットワーク上のあるサービスが停止する可能性がデメリットとしてある．一方，メリットとして，攻撃の受ける経路上のサーバ・端末を停止するため，攻撃を一定の段階から深刻化することを防ぐことが期待できる．

次に，辺連結度の概念を適用するとグラフは図 4 のようになる．

図 4: 辺連結度によるグラフの変化

図 4 のグラフは，ある辺を削除したことで点連結度とは異なるグラフに分断された．ネットワークとして考えると，特定の端末・サーバ間の通信が不可能になり経路が削除された状態になる．この概念のデメリットを次に示す．特定の経路が利用できなくなる．しかし，端末やサーバを停止させないため，ネットワーク構成を変更しても攻撃を引き続き受ける可能性がある．さらに，利用できる経路が減少することで，ネットワークの利便性が損なわれる可能性がある．だが，サーバ・端末で動作しているサービスを使い続けることができる．これは，組織や企業にとって非常に重要な項目をみたすことが可能である．これまでに述べてきた，点連結度および辺連結度のデメリットとメリットを表 1 にて示す．

表 1：点連結度と辺連結度の長所と短所

手法	メリット	デメリット
点連結度	・攻撃被害の軽減，防止が期待できる	・サービスが利用できない
辺連結度	・継続してサービス利用可能 ・攻撃被害の軽減が期待できる	・利用可能な経路が限定される ・攻撃被害の防止は期待できない

これまでに示した 2 つの手法を利用して，レジリエンス評価を行う．これによって，あらゆる攻撃事象ごとに異なるネットワーク構成の変更が可能である．従って，実験によってネットワーク構成のパターンが作成される．しかし，この手法だけでは，グラフ理論に基づくモデル化では抽象的で具体例との対応付けができない．具体的な解決案は 6 章で示す．

4. システム構成

本章では，レジリエンス（強靭性）分析のシステム構成に関して検討した事項を示す．

従来までは，レジリエンス（強靭性）分析システムが存在しなかったため，共進化シミュレーション機構が実験結果を受け取っていた．今後は，レジリエンス（強靭性）分析システム同様に，実験結果を受け取り，分析を行う．分析結果を共進化シミュレーション機構へ提供することを想定している．

次にシステム構成図を図 5 として示す．この構成で，レジリエンス（強靭性）分析システムを開発する．

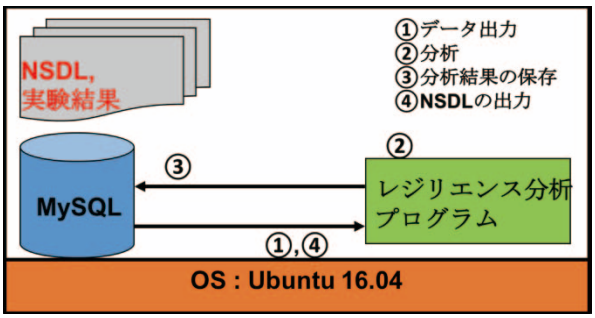

図 5：レジリエンス分析のシステム構成

MySQL と分析プログラムで構成される．MySQL は，NSDL で記述されたネットワーク構成情報と，実験結果（ログ）から抽出した攻撃被害を受けた端末・サーバ名と攻撃検知数を保持している．今回必要となるテーブルの構成例を表 2 および表 3 として示す．

表 2：攻撃検知数カウントテーブル

タグ	端末・サーバ名	攻撃検知数
vms	client1	10
vms	zabbix	0
top	top	5

表 3：ネットワーク構成情報テーブル

タグ	端末・サーバ名	表示順	IF 名
vms	client1	1	br7
vms	zabbix	1	br8
top	top	1	br1
top	top	2	br2
top	top	3	br8

ネットワーク構成情報テーブルは，タグ，端末・サーバ名，表示順の 3 つのカラムを主キーにしている．表示順以外のカラムはすべて NSDL に記述されている情報から抽出する．

次に，攻撃検知数カウントテーブルの主キーはタグ，端末・サーバ名である．攻撃検知数のカラムには，実験結果（ログ）から抽出し，集計した回数が入る．

分析プログラムは Python で実装を検討している．レジリエンス分析プログラムの動作順序は，データ取出の命令を出し，データベースからデータを取得する．取得したデータに基づき分析を行う．最後に分析結果をデータベースに保存するというプログラムである．

5. 分析・評価手順

本章では，4 章で示したシステムを用いてどのようにレジリエンス分析を行い，評価していくか示す．

はじめに，レジリエンス分析を次の手順で行うことを検討した．

Step1 実験によって得られたログデータから，頻繁に攻撃されたサーバや端末をログデータから抽出しデータベースに保存する．これを任意の回数繰り返す．繰り返し行うことで，現状のネットワーク構成における弱点を見つけ出す．

Step2 抽出結果から，グラフ理論を適用して，ネットワーク構成を変更する．具体的な手法は，①データベースから攻撃検知数カウントテーブルで集計された上位 3 つの端末・サーバ名を抽出する．②抽出したデータを利用して，ネットワーク構成情報テーブルを更新する．点連結度の場合，該当する端末・サーバをデータベースから削除する．また，辺連結度の場合は，テーブル内の IF 名のカラムに保存されている情報を削除する．削除する際に，IF 名が複数ある場合は，どれか 1 つを削除する．ただし，他の変更対象に該当する端末・サーバでも利用しているものがある場合，それを優先的に削除する．こうした条件に当てはまらない場合は，表示順の最上位を削除する．

Step3 更新された，ネットワーク構成情報テーブルから NSDL を出力する．出力された NSDL をサイバーレンジ上に展開し実験を行う．

レジリエンス分析は毎回，Step1~3 を繰り返し行う．簡易な手順を図 6 で示す．

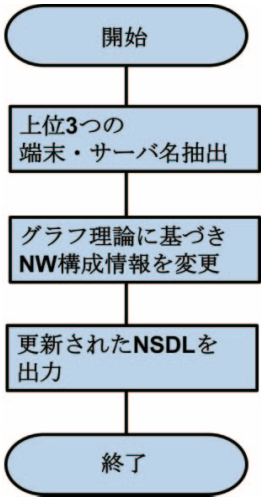

図 6：レジリエンス分析の手順

つづいて，評価方法に関して検討したことを示す．前提条件として，評価を実施するためには，1 回以上のレジリエンス分析を行っていることとする．

Step1 サービス継続性確認用サイバーレンジを起動し，Nmap を利用してサービス継続性があるか確認する．

Step2 前回(1 つ前)の分析で，攻撃検知数カウントテーブルで集計された上位 3 つの端末・サーバの攻撃検知数と最新のデータを比較する．

この評価方法は，暫定的なものであり今後，より良い評

価手法を検討・実装していく．

6. 考察

これまでに，レジリエンス分析と評価に関する具体的な手法を示してきた．

人工知能搭載型サイバーレンジは，大量の事例が存在するため，インスタンス事例が発散してしまう．また，レジリエンス分析はグラフによるモデル化では抽象的で具体例との対応付けることができない．これらの問題解決策として，両システムを統合し解決する．統合することで，人工知能搭載型サイバーレンジは，インスタンス事例をシミュレーションして，実際のシステムの状態(ログ)をレジリエンス分析システムへ渡す．レジリエンス分析システムは，受け取ったログ利用して，グラフ理論に基づいてモデル化することで，ネットワーク構成のパターンが作成される．作成されたパターンを人工知能搭載型サイバーレンジに提供する．これにより，抽象的であったパターンに具体例を対応させることが可能になり，モデル側から攻撃事象がどの程度の確立で起こるか知ることができる．また，起こり得る事象を数え上げてシミュレーションが可能になるため，発散を防ぐことができる．この効果を得るために，レジリエンス分析システムを開発する．相関図を図 7 に，図 8 に統合後の運用イメージを示す．

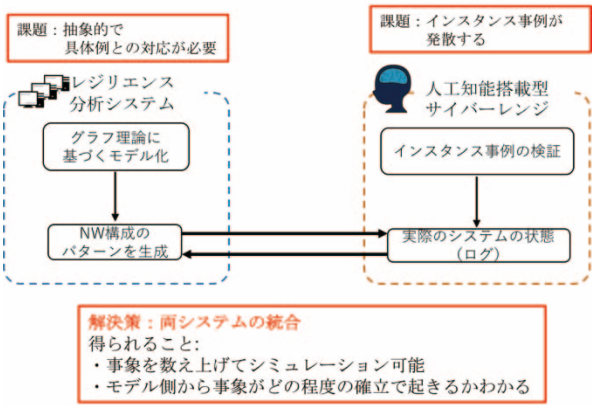

図 7：相関図

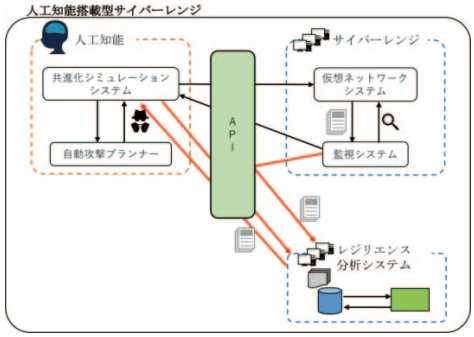

図 8：統合後の運用イメージ

統合後の人工知能搭載型サイバーレンジは，サイバーレンジでシミュレーションされたログを人工知能とレジリエンス分析システムに提供する．レジリエンス分析システムは，3，4，5 章で示した内容に従い分析・評価を行う．その結果は人工知能に提供する．

7. 今後の展開

今後は，ネットワーク構成情報テーブルから NSDL の出力を実現していく．

さらに，ネットワーク構成情報テーブルおよび，攻撃検知数カウントテーブルにあるカラム「タグ」に対して，1~5 の重みを付けることを検討している．この重みは，サービス継続性にその端末・サーバがどの程度関連があるかを意味する．すなわち，数が大きくなると，デメリットの大きい点連結度に基づく手法を利用しないようになる．サービス継続性重みテーブルを表 4 として示す．

表 4：サービス継続性重みテーブル

タグ	重み
vms	2
top	4

さらには，Web サーバをシステム構成に追加することを検討している．ブラウザ上で，ネットワーク構成の変更履歴や，攻撃検知回数の閲覧などの可視化を目指していく．

謝辞

本研究は JSPS 科研費 16K12439 の助成を受けたものです．

参考文献

[1] 佐々木良一，八槇博史“標的型攻撃に対する知的ネットワークフォレンジックシステム LIFT の開発（その 3）-今後の研究構想-”，マルチメディア，分散，協調とモバイル(DICOMO 2015)シンポジウム， 2015.

[2] 石川博也，八槇博史，“サイバー空間における攻撃と防御の共進化シミュレーション”2016 年 電子情報通信学会総合大会，DS-2-2，2016.

[3] Bernard “Chip” Ferguson，Anne Tall，“National Cyber Range Overview”，2014 IEEE Military Communications Conference，2014.

[4] 大石恵輔，八槇博史，“サイバー攻撃実験のための仮想ネットワーク自動構成方式の検討”，2016 年 電子情報通信学会総合大会，DS-2-4，2016.

[5] 中山能之，八槇博史，“仮想ネットワークシステムにおける自動攻撃と監視システムの実装”，2016 年 電子情報通信学会総合大会，DS-2-5，2016.

[6] 前田奈緒，巳波弘佳,”故障時においてもサーバへの可到達性を保証し距離増大を抑制する高信頼リンク決定法,” 2013 年 電子情報通信学会 信学技報．