様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　　元 年　　６ 月 ２７ 日現在

機関番号：３２６５７
研究種目：挑戦的萌芽研究
研究期間：2016～2018
課題番号：１６Ｋ１２４３９

研究課題名（和文）人工知能技術を用いたサイバー攻撃と対策の自動共進化による「先回り」の実現

研究課題名（英文）Study on Automatic Coevolution between Cyber Attacks and Countermeasures Based on AI Technologies

研究代表者
八槇　博史（YAMAKI, Hirofumi）

東京電機大学・システムデザイン工学部・教授

研究者番号：１０３２２１６６

交付決定額（研究期間全体）：（直接経費）　　2,600,000円

研究成果の概要（和文）：標的型攻撃が，人工知能技術を搭載したマルウェアを用いて高度化されるという将来像を前提として，攻撃者に先回りする形で攻撃内容を自動合成してシミュレーションを行い，防御方法までを自動的に生成するための技術について研究を行った．
具体的には，マルウェアの攻撃とそれに対する防御を模擬するためのネットワークシミュレータをクラウドコンピュータ上に構築し，攻撃内容や防御内容を共進化計算により変化させて様々な局面を再現するシステムを構成した．また，人工知能を搭載したマルウェアのモデルとして，プランニング計算に基づいて攻撃計画を行うものや，人間との対話を通じて詐欺をはたらくものについて検討した．

研究成果の学術的意義や社会的意義
本研究は，(1)マルウェアや攻撃対策システムを自律的なソフトウェアであるエージェントとしてモデル化し，(2)共進化計算を通じてマルウェアと攻撃対策システムとのあいだの「イタチごっこ」をシミュレーションすることにより(3)将来高度化すると予想されるマルウェアの攻撃に先んじて対策を構築するというものである．
　この技術の開発により，新しい攻撃の出現を対策が後追いするというセキュリティ対策技術の宿命から脱し，攻撃者の先回りをして防御をするという，真の意味でプロアクティブなセキュリティ対策が実現する．

研究成果の概要（英文）：Based on the future image that targeted attacks are advanced using malware equipped with artificial intelligence technology, the attack contents are automatically synthesized and simulated in a form ahead of the attackers, and defense methods are automatically generated. We researched the technology to generate in.
Specifically, we constructed a network simulator to simulate malware attack and defense against it on a cloud computer, and constructed a system to reproduce various aspects by changing attack content and defense content by co-evolutionary calculation . In addition, we examined the attack model based on the planning calculation and the one that works the fraud through the dialogue with human being as a model of the artificial intelligence loaded malware.

研究分野：情報セキュリティ

キーワード：人工知能　共進化学習　標的型攻撃　マルウェア

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９、ＣＫ－１９（共通）

１．研究開始当初の背景

特定の企業や組織を攻撃対象とする標的型攻撃が問題となっている.現在の標的型攻撃では,攻撃者によるマルウェアの作成や，侵入後の攻撃内容の決定は，多くのケースで人間である攻撃者が，様々な脆弱性情報をもと構成し実施している．この意味で，マルウェアは攻撃者が遠隔操作するマニピュレータにすぎず，それ自体が高度な処理を行うわけではない.

他方で，近年人工知能技術が急速に発達しており，物理世界においてロボットなどが自律的に行動できるように，サイバー世界においてもソフトウェア，中でもマルウェアが人工知能を搭載し，人間である攻撃者の遠隔操作を必要とせず，自律的にサイバー攻撃を実施する時代が近く到来するであろう．そのようなマルウェアは，人間が遠隔操作する場合よりもはるかに持続的・迅速・複雑な攻撃が可能である.

機械学習に代表される人工知能技術を用いて攻撃対策を行うアプローチはこれまでも多数提案されており，それを謳うセキュリティ製品も多く販売されている．その一方で，攻撃そのものが人工知能により強化される事態への対処は世界的にも緒についたところである．本研究で注目するのは人工知能技術を用いた攻撃手法生成に関する先行研究であり，たとえば以下のものでは進化アルゴリズムを用いて既存マルウェアの亜種の自動生成に成功している.

Noreen, S., et al. (2009, July). Evolvable malware. In Proceedings of the 11th Annual conference on Genetic and evolutionary computation (pp. 1569-1576). ACM.

２．研究の目的

本研究の目的は，上記のビジョンのもと，人工知能を搭載したソフトウェアとしてマルウェアをモデル化し，それが実施しうる攻撃をシミュレーションし，さらに様々な攻撃内容を進化アルゴリズムにより自動生成することで，起こりうるサイバー攻撃を予見するための技術を開発することである.

研究目的を達成するため，以下の各技術を開発した.

(1)　マルウェアの動作を自動合成し，クラウド上に構成する仮想ネットワークシステムの中で評価するシミュレーションを実施する.

(2)　このシミュレーションの結果をもとに，遺伝的アルゴリズムなどの進化計算をクラウド上で適用することにより，危険性の高いマルウェアの動作内容を予測する.

(3)　同様の手法により防御システムの側も進化させる.

(4)　攻撃側と防御側が互いに適応する共進化を発生させることにより，個々のネットワークシステムについて有効かつ現実的な攻撃と防御の組合せを発見する手法について研究する.

３．研究の方法

本研究で採った手法は,(1)疑似サイバー攻撃の自動生成技術,(2)ネットワークシステムシミュレータ,(3)共進化学習機構といった各サブシステムの実現を通じて,(4)サイバー攻撃・防御の評価技術の確立をめざす，というものであった.(1)についてはサイバー攻撃のテストツールとサイバー攻撃データベースをもとに様々な攻撃の組合せを生成する.(2)については仮想計算機と仮想ネットワークの構築技術を組合わせることで実現する.(3)については提案者がこれまで開発してきた進化計算機構を発展させることで実現する.これらを組合わせることによって,それぞれのネットワークシステムにおいてどのようなサイバー攻撃が有効で，それに対する最適な防御を自動的に発見することができるようになる.

これらのうち，(1) については，検討対象となるネットワークシステムをすべてソフトウェアとして構築することによって，多様なシミュレーションを同時並列的に行えるようになる.そこで，近年急速に技術の進んだ仮想化技術を用いた．具体的には，サーバや端末類は kvm と Docker により仮想計算機として構築し，それらをつなぐネットワークに関しては OpenFlow に基づいて定義した．シミュレーションを実施するためにはこれらを統合した記述を行う必要があるが,それを行うための規格が十分に成熟していなかったため,独自に Network Description Language (NDL) とよぶ記述系を開発し，それを用いてネットワークシステムを定義できるようにした.

(2) に関して，以上でのべた記述系に基づいた，与えられた NDL をもとに仮想ネットワークを商用クラウドである Amazon Web Services 上に自動構築するシステムを開発した．本研究で想定するシミュレーションのためには，仮想ネットワークシステムを多数そして何度も構築する必要があるため，自動構築に要する時間を短縮する必要がある．このため，共通部分を仮想ネットワークシステム間で共有することによって，自動構築に関わるオーバーヘッドを全体として短縮する手法を採った.

(3) に関しては，以前より Amazon Web Services 上のシミュレーション制御システムである GPGCloud を，本研究で利用可能なように改造し，上記の仮想ネットワークシステム上での実験を制御できるようにした．さらに，本研究の主眼である共進化シミュレーションを実現するた

め，攻撃ベクトルや防御ベクトルの生成機構を新規に開発し，さらに共進化シミュレーションが可能となるよう，進化対象とする遺伝子ベクトルを複数もつことのできる遺伝的アルゴリズム制御を追加した．

４．研究成果

本研究を実施するためのネットワークシミュレータと，共進化シミュレーションをクラウド上で行うための機構とを開発し，これらを結合したシステムを開発した．また，人工知能を搭載したマルウェアの挙動を検証するため，各種の攻撃手法について検討を行った．

ネットワークシミュレータは，研究期間の前半においては Docker を用い，Amazon Web Services の上に仮想ネットワークシステムを展開するシステムを構築したが，その段階ではネットワークを記述するための規格が十分整っておらず，独自の記述言語 NSDL を定義していた．その後，同様の機能が Kubernetes を用いて実現可能になってきたため，現在は Kubernetes 上への移植を行っている．

また，共進化シミュレーションのための機構としては，研究代表者が以前から開発していた GPGCloud をもとに，ネットワーク攻撃シミュレーションに必要となる機能およびインタフェースを追加することで目的のシステムを構築した．

人工知能を搭載したマルウェアの挙動解析としては，ヒューリスティック探索に基づいて侵入活動を継続的かつ自律的に行う手法の開発と評価を行った．また，研究開始当初は想定していなかったが，人間である利用者に働きかけるタイプのマルウェアについても検討するため，詐欺を自動的に行うマルウェアとその対策について研究を行っている．

これらの成果は，以下で示す学会発表を通じて公表している．

５．主な発表論文等

〔雑誌論文〕（計　0 件）

〔学会発表〕（計　14 件）

1. 大崎康太，八槇博史，“生体認証における識別器の検証機構，” 第８１回情報処理学会全国大会，7ZA-07，福岡市，2019 年 3 月．
2. 平野　誠，八槇博史，“機械学習を用いた攻撃検知に関する学習手法の精度評価，” 第８１回情報処理学会全国大会，5ZA-05，福岡市，2019 年 3 月．
3. Saurav Brahma，齊藤悠希，八槇博史，“動的プランニングを用いたサイバー攻撃経路自動生成機構の Kubernetes 上での実装，” 第８１回情報処理学会全国大会，2ZA-04，福岡市，2019 年 3 月．
4. 山越祐希，八槇博史，“詐欺プログラム対策のための詐欺プロセスモデルの検討，” 第８１回情報処理学会全国大会，1ZA-06，福岡市，2019 年 3 月．
5. Saurav Brahma，八槇博史，“動的プランニングを用いたサイバー攻撃経路の自動生成，” 第 17 回情報科学技術フォーラム（FIT2018），L-020，福岡市，2018 年 9 月．
6. 渋谷健太，久山真宏，松本隆，八槇博史，佐々木良一，“標的型に対する知的ネットワークフォレンジックシステム LIFT の開発と機能拡張 （その 4） －将来起こりうる攻撃方法の推定－，” コンピュータセキュリティシンポジウム 2017(CSS2017)，1D3-5，山形市，2017 年 10 月．
7. ファン　タァン　クァン，八槇博史，“感情解析に基づく誘導型サイバー攻撃検知の検討，” コンピュータセキュリティシンポジウム 2017（CSS2017），2C1-3，山形市，2017 年 10 月．
8. 大石 恵輔，中山 能之，岩東 佑季，石川 博也，宮本 貴義，八槇 博史，“サイバー攻撃対策のための人工知能搭載型サイバーレンジの検討，” マルチメディア，分散，協調とモバイル(DICOMO2017)シンポジウム，pp．1635-1639，札幌市，2017 年 6 月．
9. 山越祐希，八槇博史，“対話エージェントを通じた誘導型サイバー攻撃の検討，” 電子情報通信学会 2017 年総合大会，D-19-4，名古屋市，2017 年 3 月．
10. 中山 能之，宮本 貴義，大石 恵輔，岩東 佑季，八槇 博史，“人工知能搭載型サイバーレンジによるシステム強靱性の検討，” マルチメディア，分散，協調とモバイル(DICOMO2017)シンポジウム，pp．1635-1639，札幌市，2017 年 6 月．
11. 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，” コンピュータセキュリティシンポジウム(CSS) 2016，3F4-4，秋田市，2016 年 10 月．
12. 齊藤悠希，八槇博史，“自動プランニングを用いたサイバー攻撃手順の生成，” コンピュータセキュリティシンポジウム(CSS) 2016，3F4-1，秋田市，2016 年 10 月．
13. 鈴木 文仁，佳山 こうせつ，八槇 博史，佐々木 良一，“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－ベイジアンネットワークの適用－，” マルチメディア、分散、協調とモバイル DICOMO2016 シンポジウム，pp．1075-1080，三重県鳥羽市，2016 年 7 月．
14. 岸有哉，江口健，石川毅，鈴木竜生，宮口侑己，大澤一生，與五澤守，佐野香，八槇博史，

上野洋一郎，佐々木良一，小林　浩，“自律分散型インターネットセキュリティ基盤を模擬したテストベッドでの DDoS 攻撃の遮断実験，” 信学技報，vol．116，no．79，IA2016-9，pp．45-50，2016 年 6 月．

〔図書〕（計　0 件）

〔産業財産権〕
○出願状況（計　0 件）

名称：
発明者：
権利者：
種類：
番号：
出願年：
国内外の別：

○取得状況（計　0 件）

名称：
発明者：
権利者：
種類：
番号：
取得年：
国内外の別：

〔その他〕
ホームページ等

６．研究組織

(1)研究分担者
研究分担者氏名：
ローマ字氏名：
所属研究機関名：
部局名：
職名：
研究者番号（8 桁）：

(2)研究協力者
研究協力者氏名：
ローマ字氏名：