【公報種別】特許法第１７条の２の規定による補正の掲載
【部門区分】第６部門第３区分
【発行日】令和3年2月25日(2021.2.25)

【公表番号】特表2020-505707(P2020-505707A)
【公表日】令和2年2月20日(2020.2.20)
【年通号数】公開・登録公報2020-007
【出願番号】特願2019-541304(P2019-541304)
【国際特許分類】
Ｇ０６Ｆ　21/55　　(2013.01)
Ｇ０６Ｎ　20/00　　(2019.01)
【ＦＩ】
Ｇ０６Ｆ　21/55　　３２０
Ｇ０６Ｎ　20/00　　１３０

【手続補正書】
【提出日】令和3年1月14日(2021.1.14)
【手続補正１】
【補正対象書類名】特許請求の範囲
【補正対象項目名】全文
【補正方法】変更
【補正の内容】
【特許請求の範囲】
【請求項１】
継続的に学習するモデルを介してネットワーク上で提供されるオンラインサービスを保護するための方法であって、
前記オンラインサービスから１つまたは複数のネットワークセッションに関連するセキュリティ信号のセットを収集するステップであって、前記セキュリティ信号のセットは、時間のローリングウィンドウで収集される、ステップと、
前記セキュリティ信号のセットのそれぞれのセキュリティ信号が悪意のあるものか無害であるかを識別するステップと、
前記時間のローリングウィンドウの均衡させた訓練データセットを、
悪意のある信号それぞれについて識別された攻撃タイプに基づいて、前記セキュリティ信号のセットのうち悪意のある信号を均衡させるステップと、
無害の信号がそれぞれ受信されるデバイスのタイプに基づいて、前記セキュリティ信号のセットのうち無害の信号を均衡させるステップと、
前記悪意のある信号を前記無害の信号と交差結合することにより、前記悪意のある信号を前記無害の信号と均衡させるステップと
により、生成するステップと、
前記均衡させた訓練データセットに基づいて予測モデルを生成するステップと、
を含み、前記オンラインサービスから新しいネットワークセッションに関連する追加のセキュリティ信号を受信することに応答して、前記追加のセキュリティ信号が悪意のあるものか無害のものかが判定するために前記予測モデルが適用される、
方法。
【請求項２】
前記セキュリティ信号のセットのそれぞれのセキュリティ信号が悪意のあるものか無害であるかを識別するステップが、
それぞれのセキュリティ信号を生成モデルで検査するステップであって、前記生成モデルは、前記均衡させた訓練データセットに従ってモデルトレーナーによって生成され、所与のセキュリティ信号が悪意のあるものか無害であるかの検出結果を生成するよう構成さ

れる、ステップと、
前記検出結果を分析者ユーザに送信するステップと、
前記検出結果に関して前記分析者ユーザからアクションを受け取るのに応答して、前記所与のセキュリティ信号が悪意のあるものか無害であるかを示すように前記検出結果を更新するステップと、
をさらに含む、請求項１に記載の方法。
【請求項３】
自動化アタッカーが、前記オンラインサービスに対する攻撃を模擬し、前記セキュリティ信号のセットのそれぞれのセキュリティ信号が悪意のあるものか無害であるかを識別するステップが、
前記攻撃に応答して生成された前記セキュリティ信号を識別する通知を、前記自動化アタッカーから受け取るステップと、
前記攻撃に応答して生成された前記セキュリティ信号を、前記検出結果に関わらず悪意のあるものとして扱うステップと、
をさらに含む、請求項２に記載の方法。
【請求項４】
前記セキュリティ信号のセットのそれぞれのセキュリティ信号が悪意のあるものか無害であるかを識別するステップが、
前記所与のセキュリティ信号から特徴を抽出するステップと、
前記所与のセキュリティ信号から抽出された前記特徴が、ある攻撃タイプを定義するものとして管理ユーザによって指定された特徴セットを満たすかどうかを判定するステップと、
抽出された前記特徴が前記特徴セットを満たすと判定するのに応答して、前記所与のセキュリティ信号を悪意のあるものと指定するステップと、
抽出された前記特徴が前記特徴セットを満たさないと判定するのに応答して、前記所与のセキュリティ信号を無害と指定するステップと、
をさらに含む、請求項２に記載の方法。
【請求項５】
悪意のある信号それぞれについて識別された攻撃タイプに基づいて、前記セキュリティ信号のセットのうち前記悪意のある信号を均衡させるステップは、
前記悪意のある信号に対して観察された攻撃タイプのセットについて、攻撃タイプの総対数を、
前記攻撃タイプのセット中で実際よりも少なく示された攻撃タイプの相対量を増加させるステップと、および
前記攻撃タイプのセット中で実際よりも多く示された攻撃タイプの相対量を減少させるステップと、
のうち少なくとも一方によって均衡させるステップと、
をさらに含む、請求項４に記載の方法。
【請求項６】
前記特徴セットは、前記管理ユーザから与えられる構造化文書中で識別され、前記構造化文書は、前記セキュリティ信号のセット中で観察すべき特徴タイプおよびデータフィールドを特定し、前記特徴セットの特徴は、コードを修正する必要なしに、前記構造化文書に基づいて前記セキュリティ信号のセットから動的に抽出される、請求項４に記載の方法。
【請求項７】
過去のデータ信号が前記セキュリティ信号のセットに含められる、請求項１に記載の方法。
【請求項８】
前記無害の信号がそれぞれ受信されるデバイスのタイプに基づいて、前記セキュリティ信号のセットのうち無害の信号を均衡させるステップは、

前記ローリングウィンドウ内で少なくとも１つの悪意のある信号が収集された前記オンラインサービス内のデバイスを識別するステップと、
前記デバイスに関連する無害の信号を、前記セキュリティ信号のセットから除去するステップと、
をさらに含む、請求項１に記載の方法。
【請求項９】
前記悪意のある信号を無害の信号と交差結合することにより、前記悪意のある信号を前記無害の信号と均衡させるステップは、
前記均衡させた訓練データセットについて複数の攻撃シナリオを生成するために前記悪意のある信号を前記無害の信号と交差結合するステップをさらに含み、前記複数の攻撃シナリオは、前記悪意のある信号と前記無害の信号とを含む、
請求項８に記載の方法。
【請求項１０】
前記ローリングウィンドウ内の前記セキュリティ信号のセットを収集するステップは、
現在の時刻からの前記ローリングウィンドウを定義する時間フレームを識別するステップと、
前記現在の時刻で前記オンラインサービスから新しいセキュリティ信号を受信するステップと、
前記現在の時刻で受信した前記新しいセキュリティ信号を前記セキュリティ信号のセットに追加するステップと、
前記現在の時刻からの前記時間フレームの外側で前記セキュリティ信号のセットに追加された既存のセキュリティ信号を識別するステップと、
前記セキュリティ信号のセットから、前記現在の時刻からの前記時間フレームの外側で前記セキュリティ信号のセットに追加されたと識別された前記既存のセキュリティ信号を除去するステップと
をさらに含む、請求項１に記載の方法。
【請求項１１】
継続的に学習するモデルを介してネットワーク上で提供されるオンラインサービスをセキュリティ保護するためのシステムであって、
プロセッサと、
命令を含むメモリ記憶装置と、を備え、前記命令は、前記プロセッサによって実行された時に、
前記オンラインサービス内のデバイスからセキュリティ信号を受信することであって、前記セキュリティ信号は１つまたは複数のネットワークセッションと関連付けられる、受信することと、
前記セキュリティ信号のそれぞれから特徴ベクトルを抽出することであって、所与の特徴ベクトルは、所与のセキュリティ信号が受信される所与のデバイスの状態を表す数値を提供する、抽出することと、
関連付けられた予測モデルを介して前記特徴ベクトルそれぞれの検出結果を生成することであって、所与の検出結果は、所与のネットワークセッションと関連付けられた前記所与のセキュリティ信号が前記所与のデバイス上で悪意のある活動を示唆するか無害の活動を示唆するかを識別する、生成することと、
ローリングウィンドウを定義することであって、前記ローリングウィンドウは、現在の時刻からのある時間フレーム内に受信された複数のセキュリティ信号および関連付けられた検出結果を含む、定義することと、
前記ローリングウィンドウについて均衡させた訓練データセットを生成することと、
を行うように動作可能であり、
前記均衡させた訓練データセットを生成するために、前記システムがさらに、
悪意のある活動を示唆するものと識別された前記ローリングウィンドウ内の前記セキュリティ信号それぞれの攻撃タイプを識別し、

前記ローリングウィンドウ内で実際よりも少なく示された攻撃タイプで識別されたセキュリティ信号の量を、実際よりも多く示された攻撃タイプで識別されたセキュリティ信号の量に対して増加させ、

悪意のある活動を示唆するものと識別された前記セキュリティ信号を、無害の活動を示唆するものと識別されたセキュリティ信号と交差結合して、前記ローリングウィンドウについての攻撃シナリオを生成し、

機械学習アルゴリズムに従って、前記均衡させた訓練データセットに基づいて前記関連付けられた予測モデルを更新する、

ように構成される、システム。

【請求項１２】

前記関連付けられた予測モデルを更新することは、

前記均衡させた訓練データセットに従って、前記セキュリティ信号が前記デバイス上で悪意のある活動を示唆するか無害の活動を示唆するかを、開発モデルの方が正確に識別することを、前記機械学習アルゴリズムが示すのに応答して、前記検出結果を生成するために使用される生成モデルを、前記機械学習アルゴリズムに従って前記均衡させた訓練データセットから開発された前記開発モデルに置き換えることを含む、請求項１１に記載のシステム。

【請求項１３】

過去の信号が前記ローリングウィンドウに含められ、前記過去の信号は、前記時間フレームの外側で収集されたセキュリティ信号を含む、請求項１１に記載のシステム。

【請求項１４】

前記オンラインサービス内の前記デバイスから受信された前記セキュリティ信号は、自動化アタッカーが前記オンラインサービスに対して既知の悪意のある活動を行うのに応答して発生したセキュリティ信号を含み、

前記自動化アタッカーが前記既知の悪意のある活動を行うのに応答して発生した前記セキュリティ信号に対して生成された前記検出結果は、前記自動化アタッカーからの通知に基づいて、前記所与のセキュリティ信号が悪意のある活動を示唆することを示すように設定される、請求項１１に記載のシステム。

【請求項１５】

前記均衡させた訓練データセットを生成するために、前記システムがさらに、

無害の活動を示唆するものと識別されたセキュリティ信号が受信された特定のデバイスが、前記ローリングウィンドウ内で悪意のある活動を示唆するものと識別された１つまたは複数のセキュリティ信号と関連付けられていることを識別するのに応答して、無害の活動を示唆するものと識別された前記セキュリティ信号を前記ローリングウィンドウから除去するように構成される、請求項１１に記載のシステム。

【請求項１６】

前記均衡させた訓練データセットを生成するために、前記システムがさらに、

無害の活動を示唆するものと識別された前記ローリングウィンドウ内の前記セキュリティ信号それぞれが受信されたデバイスのタイプを識別し、

前記ローリングウィンドウ内で実際よりも少なく示されたデバイスタイプで識別されたセキュリティ信号の量を、実際よりも多く示されたデバイスタイプで識別されたセキュリティ信号に対して増加させる、

ように構成される、請求項１１に記載のシステム。

【請求項１７】

継続的に学習するモデルを介してネットワーク上で提供されるオンラインサービスをセキュリティ保護するためのプロセッサ実行可能命令を含むコンピュータ可読記憶装置であって、前記プロセッサ実行可能命令は、

前記オンラインサービスから１つまたは複数のネットワークセッションに関連するセキュリティ信号のセットを収集することであって、前記セキュリティ信号のセットは時間のローリングウィンドウ内で収集される、収集することと、

所与のネットワークセッションと関連するそれぞれのセキュリティ信号が悪意のあるものか無害であるかを識別するために、予測モデルを介して前記セキュリティ信号のセットのそれぞれのセキュリティ信号を検査することであって、前記予測モデルは、管理ユーザによって定義された特徴ベクトルに基づいて、所与のセキュリティ信号が悪意のあるものか無害であるかの検出結果を生成するように構成される、検査することと、
前記セキュリティ信号のセットを悪意のある信号または無害の信号として識別するために、前記セキュリティ信号のセットを前記検出結果に関連付けることと、
均衡させた訓練データセットを生成するために、前記悪意のある信号を前記無害の信号と均衡させることと、
前記悪意のある信号それぞれの攻撃タイプを識別することと、
前記無害の信号のそれぞれが収集されたデバイスのタイプを識別することと、
攻撃事例のセットを生成するために、前記ローリングウィンドウ内の悪意のある信号の各相対数を、識別された攻撃タイプに基づいて等しくすることと、
識別された無害事例のセットを生成するために、前記ローリングウィンドウ内の無害の信号の各相対数を、識別されたデバイスタイプに基づいて等しくすることと、および
前記攻撃事例のセット中の攻撃事例の数を、前記無害事例のセット中の無害事例の数に対して均衡させるために、前記攻撃事例のセットを前記無害事例のセットの少なくとも一部分と交差結合すること、
を含むことと、
前記均衡させた訓練データセットおよび機械学習アルゴリズムに基づいて前記予測モデルを精緻化することと、
を含む、コンピュータ可読記憶装置。
【請求項１８】
前記悪意のある信号を前記無害の信号と均衡させることは
前記ローリングウィンドウ内で少なくとも１つの悪意のある信号が収集された前記オンラインサービス内でデバイスを識別することと、
前記デバイスと関連する無害の信号を、前記セキュリティ信号のセットから除去することと、
をさらに含む、請求項１７に記載のコンピュータ可読記憶装置。
【請求項１９】
前記予測モデルを介してそれぞれのセキュリティ信号を検査することは、
前記検出結果を分析者ユーザに送信することと、
前記検出結果に関連する分析者ユーザからのアクションを受け取るのに応答して、前記分析者ユーザからの前記アクションに照らして、前記所与のセキュリティ信号が悪意のあるものか無害であるかを示すように前記検出結果を更新することと、
をさらに含む、請求項１７に記載のコンピュータ可読記憶装置。
【請求項２０】
前記特徴ベクトルは、1つまたは複数の攻撃タイプを識別するよう構成された予測モデルを定義するために前記管理ユーザによって与えられる構造化文書中で識別されるデバイスのステータスの特徴セットを含み、前記構造化文書は、前記セキュリティ信号のセット中で観察すべき特徴タイプおよびデータフィールドを特定し、前記特徴セットの特徴は、コードを修正する必要なしに、前記構造化文書に基づいて前記セキュリティ信号のセットから動的に抽出される、請求項１７に記載のコンピュータ可読記憶装置。