(19) CONFÉDÉRATION SUISSE
INSTITUT FÉDÉRAL DE LA PROPRIÉTÉ INTELLECTUELLE

(11) CH 718 977 A2

(51) Int. Cl.: G06F 21/56 (2013.01)

Demande de brevet pour la Suisse et le Liechtenstein
Traité sur les brevets, du 22 décembre 1978, entre la Suisse et le Liechtenstein

(12) DEMANDE DE BREVET

(21) Numéro de la demande: 00794/22

(22) Date de dépôt: 01.07.2022

(43) Demande publiée: 31.03.2023

(30) Priorité: 30.09.2021 US 17/449,616

(71) Requérant:
Acronis International GmbH, Rheinweg 9
8200 Schaffhausen (CH)

(72) Inventeur(s):
Stanislav Protasov, 098185 Singapore (SG)
Alexey Malanov, 399947 Singapore (SG)
Serg Bell, 469986 Singapore (SG)

(74) Mandataire:
IP Partners SA, A-One Business Center La Pièce 1 - A5
1180 Rolle (CH)

(54) Analyse du comportement basée sur une machine à états finis pour la détection de logiciels malveillants.

(57) Un système et un procédé sont divulgués pour identifier une activité malveillante sur un dispositif cible en fonction de l'analyse du comportement du dispositif cible. Le système comprend un analyseur comportemental exécuté sur une machine virtuelle connectée au dispositif cible. La machine virtuelle collecte les événements et les paramètres du système à partir du dispositif cible et exécute un script, indépendant du dispositif cible, pour détecter une menace. Le script est un ensemble d'instructions exécutées pour analyser le comportement d'un objet en traitant et en corrélant les événements. Le script comprend une structure de règles qui stocke les signatures et les expressions des malwares connus. En corrélant les paramètres des événements sélectionnés avec les paramètres des logiciels malveillants connus, il est déterminé si l'événement impose une menace ou non. Une machine à états finis est utilisée pour la table de transition des états.

Description

DOMAINE DE L'INVENTION

[0001] La présente divulgation concerne généralement la gestion de la sécurité des données, et en particulier, mais sans s'y limiter, un système et un procédé permettant d'identifier une activité malveillante sur un dispositif cible par le biais de l'analyse de son comportement.

HISTORIQUE

[0002] Avec les progrès stupéfiants de la technologie numérique, la sécurité des données est devenue une préoccupation majeure. La mise à jour constante des formes de logiciels malveillants est un problème omniprésent qui entraîne des attaques réussies de logiciels malveillants, la modification des définitions des menaces, la compromission de données critiques, etc. La sécurité des données est normalement défendue à l'aide de diverses unités de détection des logiciels malveillants, généralement classées en composants programmables de détection des logiciels malveillants et en composants de sécurité. L'analyse des données peut être effectuée par le composant programmable de détection des logiciels malveillants à différents niveaux du système informatique, tels que l'espace utilisateur, l'espace noyau et l'émulateur. L'analyse peut être effectuée sur des processus individuels, des applications spécifiques, des fonctions système et l'ensemble du système d'exploitation. Les fichiers, les threads et les opérations de registre sont pris en charge comme types d'analyse. Si l'attaque du logiciel malveillant est détectée, les composants de sécurité peuvent effectuer des actions correctives pour bloquer toute communication avec le logiciel malveillant.

[0003] Traditionnellement, l'approche générale de l'analyse comportementale comprend les étapes de crochetage et de collecte des événements système, après quoi les paramètres des événements système sont analysés pour détecter tout modèle caractéristique d'une activité malveillante. Un aspect désavantageux des analyseurs conventionnels est la consommation élevée de ressources informatiques ainsi que la faible vitesse d'exécution de l'analyse. L'une des raisons de ces inconvénients est que pour effectuer l'analyse, l'analyseur conventionnel doit effectuer une pile d'opérations par utilisateur, par processus, par application, puis les faire correspondre à une signature de comportement connue. Un autre aspect désavantageux des analyseurs conventionnels est qu'au cours de la surveillance du système, l'analyseur de comportement conventionnel lui-même n'a pas de fonctionnalités pour prédire une menace ou suivre un risque de menace. Un autre aspect désavantageux des analyseurs conventionnels est le manque de souplesse des analyseurs de comportement pour mettre à jour les signatures, les règles de détection et les règles de traitement des événements, dans lesquels la réception des mises à jour se fait sous forme de bibliothèques ou de code compilé et présente des limitations en termes d'applicabilité, de traitement des erreurs, de mises à jour différentielles, etc.

[0004] Par conséquent, il existe un besoin pour un système et une méthode d'identification d'une activité malveillante sur un dispositif cible qui surmonte les aspects désavantageux susmentionnés de l'art antérieur.

DESCRIPTION SOMMAIRE

[0005] Des modes de réalisation et des aspects techniques divulgués ici concernent un système et un procédé d'identification d'une activité malveillante sur un dispositif cible via une analyse du comportement.

[0006] La présente divulgation envisage un procédé pour identifier une activité malveillante sur un dispositif cible en se basant sur l'analyse du comportement du dispositif cible. Le dispositif cible est connecté à une machine virtuelle sur laquelle un analyseur comportemental est implémenté. L'analyseur comportemental est un ensemble d'instructions de script exécutées pour analyser le comportement d'un objet ou des événements du dispositif cible. La machine virtuelle interprète les instructions de script de l'analyseur de comportement et fonctionne avec les objets de l'analyseur de comportement. Le procédé comprend la récupération d'événements système du dispositif cible au niveau de la machine virtuelle ; le filtrage des événements système en appliquant des filtres aux paramètres d'événements associés aux événements du dispositif cible pour identifier au moins un événement intéressant ; le changement d'un état d'une machine à états finis en un état attendu conformément à la table de transition à états finis ; et l'exécution d'instructions définies pour l'état attendu au niveau d'un gestionnaire de menaces déterminant une instance d'activité malveillante sur le dispositif cible.

[0007] Dans un autre mode de réalisation non limitatif, un référentiel d'analyseur de comportement est connecté à l'analyseur de comportement qui comprend des informations associées à des états finis attendus correspondant à une pluralité d'événements.

[0008] Dans un autre mode de réalisation non limitatif, l'au moins un événement d'intérêt est un événement qui expose le dispositif cible à une chance d'attaque par un logiciel malveillant ou à une activité malveillante.

[0009] Dans un autre mode de réalisation non limitatif, le ou les paramètres d'événement comprennent des paramètres de clé de registre, de permission d'accès et de session réseau.

[0010] La présente divulgation envisage également un système permettant d'identifier une activité malveillante sur un dispositif cible en fonction de l'analyse du comportement du dispositif cible. Le système comprend une machine virtuelle connectée au dispositif cible pour récupérer les événements système du dispositif cible, la machine virtuelle exécute un module d'analyse du comportement. Le module d'analyse du comportement comprend un module de traitement d'événe-

ments configuré pour filtrer les paramètres d'événements associés aux événements du dispositif cible pour identifier au moins un événement intéressant ; un module de corrélation d'événements configuré pour changer un état de la machine à états finis en un état attendu conformément à la machine à états finis ; et un module de traitement des menaces configuré pour identifier une instance d'activité malveillante sur le dispositif cible.

[0011] Dans un autre mode de réalisation non limitatif, le système comprend en outre un référentiel d'analyseur de comportement connecté à l'analyseur de comportement, dans lequel le référentiel d'analyseur de comportement comprend des informations associées à des états finis attendus correspondant à une pluralité d'événements.

[0012] Dans un autre mode de réalisation non limitatif, le au moins un événement d'intérêt est un événement qui expose le dispositif cible à une chance d'attaque par un logiciel malveillant ou à une activité malveillante.

[0013] Dans un autre mode de réalisation non limitatif, le ou les paramètres d'événement comprennent des paramètres de clé de registre, de permission d'accès et de session réseau.

BRÈVE DESCRIPTION DES DESSINS

[0014] Une meilleure compréhension des modes de réalisation de la présente divulgation, y compris les alternatives ou les variations de celles-ci, peut être obtenue en se référant à la description détaillée des modes de réalisation avec les dessins suivants, dans lesquels :

La figure 1 illustre un schéma fonctionnel représentant un système permettant d'identifier une activité malveillante sur un dispositif cible en fonction de l'analyse du comportement du dispositif cible, conformément à un mode de réalisation de la présente divulgation.

La figure 2 illustre un schéma fonctionnel représentant un procédé pour un système d'identification d'une activité malveillante sur un dispositif cible basé sur l'analyse du comportement du dispositif cible, conformément à un mode de réalisation de la présente divulgation.

La figure 3 illustre un procédé de préparation des règles d'un analyseur de comportement, conformément à un mode de réalisation de la présente divulgation.

La figure 4 illustre un schéma fonctionnel représentant un procédé pour un système permettant d'identifier une activité malveillante sur un dispositif cible en fonction de l'analyse du comportement du dispositif cible, conformément à un mode de réalisation de la présente divulgation.

DESCRIPTION DÉTAILLÉE

[0015] Les modes de réalisation seront maintenant décrits plus en détail ci-après en référence aux figures d'accompagnement, dans lesquelles des modes de réalisation préférés sont représentés. Toutefois, ce qui précède peut également être mis en oeuvre dans des configurations alternatives qui sont conformes à la présente divulgation.

[0016] La figure 1 illustre un schéma fonctionnel représentant un système 100, ci-après dénommé système 100, pour identifier une activité malveillante sur un dispositif cible en fonction de l'analyse du comportement du dispositif cible 102, conformément à un mode de réalisation de la présente divulgation. Le système 100 est implémenté sur le dispositif cible 102 pour identifier une activité malveillante dans le système d'exploitation et les applications 103 fonctionnant sur le dispositif cible 102. Le système comprend principalement une machine virtuelle 104 avec un analyseur comportemental mis en oeuvre sur le dispositif cible 102. Les exemples du dispositif cible 102 peuvent inclure au moins un ou plusieurs ordinateurs portables, ordinateurs de bureau, serveurs, passerelles, ports de communication, dispositifs portables ou mobiles, pare-feu, et autres.

[0017] Dans une mise en oeuvre, la machine virtuelle 104 peut être configurée sur le dispositif cible 102. La machine virtuelle est une ressource informatique qui utilise son propre logiciel à la place d'une machine informatique physique pour exécuter des programmes et des fonctions. Par exemple, la machine virtuelle exécute son propre système d'exploitation. Plusieurs machines virtuelles peuvent être exécutées sur un système informatique, où chaque machine virtuelle utilise son propre système d'exploitation et fonctionne séparément des autres machines virtuelles fonctionnant sur la même machine informatique. Les machines virtuelles peuvent être utilisées pour la virtualisation des serveurs, ce qui permet à l'utilisateur de consolider ses ressources informatiques et d'améliorer son efficacité. En outre, les machines virtuelles peuvent effectuer des tâches spécifiques qui sont considérées comme une menace pour l'exécution dans un dispositif cible hôte. Des exemples de telles tâches peuvent inclure l'accès à des données infectées par des virus, le test de systèmes d'exploitation, l'exécution de fonctions inconnues et non identifiées, etc. Comme la machine virtuelle est séparée du reste du dispositif cible hôte, elle peut exécuter les fonctions sans imposer une quelconque menace au système cible hôte. Tout dysfonctionnement survenant dans la machine virtuelle ne peut pas affecter les performances du dispositif cible hôte. Ces aspects techniques avantageux des machines virtuelles rendent le présent sujet plus efficace, plus robuste et plus sûr par rapport aux arts connus actuels dans le domaine. Un autre avantage de l'utilisation de machines virtuelles pour le traitement de l'analyse du comportement est la compatibilité des opérations de traitement des événements, des opérations

de traitement des menaces et des autres opérations de l'analyseur de comportement avec les applications, les services système et le matériel du dispositif cible.

[0018] La machine virtuelle dont il est question à la figure 1 fonctionne sur le dispositif cible 102. La machine virtuelle peut accéder au système d'exploitation et aux composants du système du dispositif cible. La machine virtuelle peut être connectée au dispositif cible via des méthodes de transport pour le transport et la communication des données. Parmi les méthodes de transport, on peut citer notamment le socket réseau, la communication interprocessus (IPC) et la mémoire partagée. Une prise réseau est une structure à l'intérieur d'un noeud de réseau d'un réseau informatique qui sert de point d'extrémité pour envoyer et recevoir des données sur le réseau. Les sockets ne sont créés que pendant la durée de vie d'un processus d'une application exécutée dans le noeud. L'IPC implique la communication d'un processus avec un autre processus. L'IPC peut servir entre des processus apparentés initiés à partir d'un seul processus, tels que des processus parents et enfants, et entre des processus non apparentés, ou deux ou plusieurs processus différents. D'autres exemples de méthodes de transport peuvent inclure le réseau de stockage (SAN) pris en charge pour le stockage directement connecté à l'aide des protocoles Fibre Channel (FC) ou Internet SCSI (iSCSI), HotAdd qui est un agent de serveur virtuel installé sur une machine virtuelle résidant sur un serveur ESX, Réseau local (NBD et NBDSSL) qui transmet les données sur la connexion TCP/IP entre le serveur ESX et le dispositif cible/ordinateur proxy, Network Attached Storage (NAS) qui permet à l'ordinateur proxy de l'agent de serveur virtuel (VSA) de lire les données directement à partir du serveur de fichiers du réseau (NFS), sans passer par un hôte ESX ou transférer les données sur le LAN, et ainsi de suite. La machine virtuelle peut ainsi accéder à des processus liés ou non liés en cours de traitement sur le système d'exploitation du dispositif cible, y compris les journaux et les crochets du système. Dans une autre mise en oeuvre, la machine virtuelle peut être connectée au dispositif cible via un serveur distant.

[0019] Conformément au mode de réalisation du présent objet, le dispositif cible 102 comprend un système d'exploitation et d'autres applications. Le dispositif cible 102 peut être tout dispositif intelligent ou un système informatique qui comprend un ou plusieurs processeurs et une ou plusieurs mémoires. Le système d'exploitation comporte un ou plusieurs composants de système d'exploitation. L'accès aux composants du système est partagé avec la machine virtuelle pour traiter les paramètres d'événements de tous les événements se produisant sur le dispositif cible 102. Conformément à un mode de réalisation exemplaire possible, un accès direct à distance à la mémoire du dispositif cible 102 peut être fourni à la machine virtuelle 104 pour garantir que la machine virtuelle 104 a accès à la mémoire vive (RAM) du dispositif cible 102 pour récupérer, de manière non invasive, les paramètres d'événement du système à partir de celle-ci.

[0020] Les dispositifs cibles 102 sont des dispositifs informatiques dotés d'un processeur, tels que des ordinateurs portables, des ordinateurs de bureau, des serveurs, des passerelles, des ports de communication ou des dispositifs mobiles. Le réseau peut être réparti entre une pluralité de dispositifs informatiques et une pluralité d'emplacements géographiques. La menace de logiciel malveillant peut provenir de menaces de réseau, de menaces de proximité physique, de menaces de localisation secondaire, etc. Le dispositif cible 102 capte la menace lorsqu'il se connecte à un serveur non protégé via Internet.

[0021] La machine virtuelle 104, conformément à un mode de réalisation de la présente divulgation, est configurée pour exécuter un analyseur de comportement 106. L'analyseur de comportement, dans une mise en oeuvre sans limitation, est un outil construit à l'aide d'instructions de programmation de haut niveau pour détecter et arrêter les processus présentant un comportement malveillant et suspect concernant les événements de niveau noyau ainsi que les événements provenant du système d'exploitation. L'analyseur de comportement est configuré pour analyser le comportement d'un objet, et dans certains cas, son comportement prévu ou potentiel pour détecter des activités suspectes. La détection de logiciels malveillants basée sur l'analyse du comportement évalue un objet de dispositif cible en fonction de ses actions prévues avant qu'il ne puisse réellement exécuter ce comportement. Par exemple, les tentatives d'exécution d'actions qui ne sont pas autorisées pour un utilisateur particulier, les actions qu'il est anormal d'exécuter sur ce dispositif cible particulier, et autres, indiquent que le comportement de l'objet ou l'activité est potentiellement malveillant. D'autres exemples de comportements intentionnels qui peuvent être considérés comme malveillants ou suspects peuvent inclure des tentatives de découverte d'un environnement sandbox, la désactivation des contrôles de sécurité, l'installation de rootkits, l'enregistrement pour un démarrage automatique et autres. L'évaluation d'un tel comportement malveillant est considérée comme une analyse dynamique.

[0022] Comme mentionné précédemment, l'analyseur de comportement est un outil construit à l'aide de langages de programmation de haut niveau. Les instructions exécutables codées dans les langages de programmation de haut niveau sont appelées scripts ou bytecodes. Pour exécuter le script, la machine virtuelle sert d'interprète. Les langages de haut niveau utilisent des interprètes au lieu de compilateurs. L'objectif ici est de créer un code intermédiaire qui peut comprendre le langage et l'exécuter sans avoir à traduire le programme en code machine. Par conséquent, la machine virtuelle agit comme un code intermédiaire qui peut comprendre et exécuter le jeu d'instructions pour l'analyse du comportement codé dans un langage de haut niveau et exécuter les instructions sans que le script soit compilé par le processeur du dispositif cible. Le jeu d'instructions est le script qui est désigné comme l'analyseur de comportement dans la présente description.

[0023] L'analyseur de comportement est construit en utilisant un langage de programmation de haut niveau. Par exemple, dans une mise en oeuvre, le script Lua est utilisé pour construire l'ensemble d'instructions. Lua est un langage de programmation léger, de haut niveau et multi-paradigme. Lua est un langage multiplateforme, car l'interpréteur si le bytecode compilé est écrit en ANSI C. L'analyseur de comportement comprend un module de détection de comportement 106A qui

initialise la détection de comportement. La détection du comportement est basée sur une structure de règles. La structure de règles est exécutée par la machine virtuelle. La structure de règles est essentiellement un mécanisme de structuration des données qui représente différents types de tableaux et de dictionnaires. Chaque règle est une table avec des tableaux qui peuvent être indexés non seulement avec des nombres mais aussi avec des chaînes de caractères. Comme une règle globale standard, les règles à l'intérieur des machines virtuelles contiennent des tableaux avec des structures de données associées à différents événements à détecter. Dans une mise en oeuvre du mode de réalisation, les règles contiennent un ensemble de tableaux avec des structures de données associées à des événements malveillants à détecter. Par exemple, RuleEmotet. L'Emotet est un cheval de Troie bancaire qui tente de se faufiler sur votre ordinateur et de voler des informations sensibles et privées. Comme le malware Emotet est connu, les signatures et les instructions associées à celui-ci sont stockées dans la table RuleEmotet. De même, un ensemble de tables est créé en Lua pour identifier de tels événements malveillants.

[0024] La règle est divisée de manière proéminente en deux étapes, d'abord une signature-automate et ensuite, un bloc logique. L'analyseur de comportement extrait les signatures associées aux événements. La signature est une empreinte ou un motif typique associé à une attaque malveillante sur un réseau ou un système informatique. Ce motif peut être une série d'octets dans un fichier (séquence d'octets) dans le trafic réseau. Il peut également prendre la forme d'une exécution logicielle non autorisée, d'un accès non autorisé au réseau, d'un accès non autorisé à un répertoire ou d'anomalies dans l'utilisation des privilèges du réseau.

[0025] L'analyseur de comportement extrait ensuite des expressions, qui sont des combinaisons logiques de signatures, et combine les expressions en un bloc logique. Les expressions logiques peuvent être un ensemble d'opérateurs booléens identifiant l'événement de signature. L'événement de signature peut être un type de paquet, une séquence de types de paquets, ou l'un quelconque d'un certain nombre d'événements liés à la signature, tels qu'un compte ou une période de temps. Des opérateurs logiques sont utilisés pour décrire les relations entre les événements de signature, par exemple si un compte dépasse une certaine valeur. Pour chaque signature, un ou plusieurs de ces identificateurs et opérateurs sont combinés pour fournir une expression régulière décrivant cette signature.

[0026] L'analyseur de comportement extrait ensuite les gestionnaires qui sont des fonctions liées à la signature et stocke les gestionnaires dans le bloc logique. Plus précisément, il s'agit d'une fonction avec deux arguments, context et event et les deux arguments sont des tableaux. Le contexte est initialement vide, et certains éléments sont couplés à partir de l'événement CtxCreate. Event contient les paires clé-valeur des événements actuellement traités. Le gestionnaire est appelé lorsque l'expression correspondante devient vraie. Des exemples de gestionnaire incluent Print(a, b, c...). La fonction print est appelée pour imprimer quelque chose sur la console, particulièrement utile pendant le débogage des règles. Quelques autres exemples incluent trace(level, message), get_context(puid), on_event(context) et autres.

[0027] Le module de détection de comportement 106A initialise ainsi la structure de règles pour traiter les signatures, les expressions logiques et les handlers liés à l'événement. L'événement est traité par un module de traitement d'événements 106D configuré pour filtrer les paramètres d'événements associés aux événements du dispositif cible 102 pour identifier au moins un événement d'intérêt. Selon un mode de réalisation, l'au moins un événement d'intérêt est un événement qui expose le dispositif cible à une chance d'attaque par un logiciel malveillant et/ou à une activité malveillante et suspecte. Parmi les exemples d'événements intéressants, citons, sans s'y limiter, l'interface avec un serveur non protégé via Internet ou la connexion d'une clé USB au dispositif cible, etc. Dans un mode de réalisation alternatif non limitatif, le ou les paramètres d'événement comprennent des paramètres de clé de registre, de permission d'accès et de session réseau, etc.

[0028] L'analyseur comportemental génère un contexte pour le nouveau processus, conformément à la mise en oeuvre du présent mode de réalisation. L'événement spécial CtxCreate est envoyé à ce contexte avec une ou plusieurs propriétés enroulées. Les exemples de propriétés comprennent cmdline, image_filename, pid, parent_pid, sid. Les nouveaux événements sont regroupés en une seule chaîne. Dans un exemple, la chaîne est gen-line. Ce ne sont pas toutes les propriétés, mais seulement certaines propriétés spécifiques avec un ordre fixe qui sont regroupées dans la chaîne. Un ou plusieurs filtres peuvent être appliqués pour sélectionner les propriétés spécifiques de l'événement. La chaîne est ensuite vérifiée par rapport à toutes les signatures des règles. Seules les propriétés sélectionnées sont combinées dans la chaîne, et elles sont séparées par des octets avec les valeurs x01, x02 pour plus de clarté.

[0029] Le module d'analyse de comportement 106 comprend en outre un module de corrélation d'événements 106D, selon la mise en oeuvre du mode de réalisation. Les expressions pertinentes sont évaluées pour corréler l'événement. Le module de corrélation d'événements 106D utilise une machine à états finis pour la transition d'un état à un autre sur la base de la table de transition d'état. L'automate fini est un module de calcul qui peut être utilisé pour simuler une logique séquentielle. Fondamentalement, il représente et contrôle le flux d'exécution. Sur la base de la séquence d'événements, des paramètres d'événements et des caractéristiques d'une activité suspecte ou malveillante se rapportant à l'événement en cours de traitement, l'état de l'automate fini est modifié par le module de corrélation d'événements. Le module de corrélation d'événements est configuré pour changer l'état de la machine à états finis à un état attendu correspondant aux événements d'intérêt conformément à la table de transition à états finis. Le tableau A est une démonstration exemplaire des séquences d'événements et de l'évaluation des expressions. Le tableau A est statique et global. Comme indiqué dans le tableau, plusieurs événements sont pris en compte, les événements a, b et c. Les expressions correspondantes sont indiquées dans la ligne 4-15 qui comprend „et, ou, alors, sauf“. Chaque événement a plusieurs déclencheurs qui sont

exécutés lorsque cet événement se produit. Par exemple, lors de l'occurrence de l'événement A, les actions suivantes sont exécutées :

1. Le gestionnaire 1 est exécuté
2. si l'événement 2(b) est déjà activé, l'événement 4 (a et B) se produit
3. l'événement 6 (a et b) se produit automatiquement, puis l'événement 12 (a, b ou c) se produit. Le handler7 est exécuté en conséquence
4. Les conditions 'ou' sont évaluées de manière triviale
5. Les conditions „et“ et „alors“ nécessitent des déclencheurs avec deux opérandes. Les conditions complexes sont divisées en conditions triviales. Un vecteur d'événement contextuel est créé pour chaque processus surveillé, classé avec des zéros.

Even-tId	Meaning				Triggers			
	a	EXECUTE Handler1	IF 2 then SET 4	SET 6		IF 2 then SET 9		EXCEPT 2 SET 13
2	b		IF 1 then SET 4	SET 6		IF 1 then SET 8		
3	c		IF 4 then SET 5	SET 12	if 6 then SET 7	IF 9 then SET 10	IF 6 then SET 11	
4	a and b		IF 3 then SET 5					
5	4 and c	EXECUTE Handler2						
6	a or b		IF 3 then SET 7	SET 12				
7	6 and c	EXECUTE Handler3 once						
8	a then b	EXECUTE Handler4						
9	b then a							
10	9 then c	EXECUTE Handler5						
11	6 then c	EXECUTE Handler 6						
12	6 or c	EXECUTE Handler7 once						
13	a except b							EXCEPT 3 SET 14
14	13 except c	EXECUTE Handler8						

Table A

[0030] Le module d'analyse de comportement 106 comprend en outre un module de traitement de menace 106D configuré pour identifier une menace comportementale. Sur la base d'une corrélation d'événements déterminée en fonction de l'état généré par une machine à états finis. Par exemple, pendant la corrélation d'événements, si l'événement correspond à

une tentative d'accès à un fichier de données non autorisé, le module de traitement de la menace peut déterminer une instance d'activité malveillante sur le dispositif cible. Si le module de traitement des menaces 106D détermine qu'un cas d'activité malveillante s'est produit sur le dispositif cible 102, le module de traitement des menaces 106D peut en outre être configuré pour informer un micrologiciel de solutions de sécurité anti-malware sur le dispositif cible 102 du cas d'activité malveillante pour une action supplémentaire.

[0031] Le système 100 comprend en outre un référentiel d'analyseur de comportement 108 connecté à l'analyseur de comportement via un interpréteur de machine virtuelle. Le référentiel de l'analyseur de comportement 108 comprend des informations associées à des états finis attendus correspondant à une pluralité d'événements. Dans un mode de réalisation, le référentiel de l'analyseur de comportement 108 comprend une ou plusieurs règles. Ces règles sont décrites plus en détail en relation avec la figure 3 ci-dessous.

[0032] La figure 2 illustre un schéma fonctionnel représentant un processus 200 pour un système d'identification d'une activité malveillante sur un dispositif cible en fonction de l'analyse du comportement du dispositif cible, conformément à un mode de réalisation de la présente divulgation. Il est à noter que le processus 200, conformément à un mode de réalisation, peut être exécuté par le système 100. Au bloc 202, le processus 200 comprend le chargement de mises à jour comprenant une liste d'événements, des attributs d'événements pour le filtrage d'événements, une table de transition d'état et des instructions sous la forme de scripts vers le module d'analyse de comportement 106 de la machine virtuelle 104. Conformément à un mode de réalisation, les mises à jour comprenant la liste d'événements, les attributs d'événements pour le filtrage d'événements, la table de transition d'état et les instructions sous la forme de scripts sont fournies par le référentiel d'analyseur de comportement 108. Dans un mode de réalisation, ces informations peuvent être fournies au module d'analyse de comportement 106 en temps réel, c'est-à-dire que les informations fournies par le référentiel d'analyse de comportement 108 peuvent correspondre à l'événement d'intérêt en cours de traitement par le module d'analyse de comportement 106.

[0033] Au bloc 204, le processus 200 comprend l'initialisation du module de traitement d'événements 106D avec la liste d'événements et les attributs d'événements spécifiés pour enregistrer les événements d'intérêt. Plus précisément, le module de traitement d'événements 106D reçoit les paramètres d'événements de tous les événements qui fonctionnent sur le dispositif cible 102.

[0034] Au bloc 206, le processus 200 comprend l'enregistrement et le filtrage d'un ou plusieurs événements spécifiés. Les événements spécifiés signifient ici l'événement d'intérêt. Les termes événement(s) spécifié(s) et événement d'intérêt sont utilisables de manière interchangeable dans la présente divulgation. Dans un mode de réalisation, l'étape est exécutée par le module de traitement d'événements 106D du module d'analyse de comportement 106.

[0035] Au bloc 208, le processus 200 comprend le changement de l'état de la machine à états finis (ou de la machine virtuelle 104) conformément à une table de transition d'état, dans laquelle chaque état est lié à la séquence d'événements, aux attributs d'événements et aux conditions d'état spécifiques. Dans une mise en oeuvre, le module de corrélation d'événements 106D exécute cette étape. Plus précisément, le référentiel d'analyse de comportement 108 fournit les informations associées à l'état lié à la séquence d'événements, aux attributs d'événements et aux conditions d'état spécifiques au module de corrélation d'événements 106D pour faciliter le changement d'état de la machine virtuelle 104 vers celui fourni par le référentiel d'analyse de comportement 108. Comme mentionné précédemment, cet état est l'état attendu du dispositif cible 102 qui est simulé sur la machine virtuelle 104. Cet état attendu du dispositif cible 102 simulé sur la machine virtuelle 104 est analysé et comparé à un état en temps réel du dispositif cible 102.

[0036] Au bloc 210, le processus 200 comprend la demande de savoir si l'état identifie une menace. Plus précisément, le module de traitement des menaces 106D est configuré pour identifier un comportement malveillant ou suspect sur le dispositif cible. Si le module de gestion des menaces 106D détermine qu'un cas d'activité malveillante s'est produit sur le dispositif cible 102, le module de gestion des menaces 106D peut en outre être configuré pour informer un micrologiciel de solutions de sécurité anti-malware sur le dispositif cible 102 du cas d'activité malveillante pour une action supplémentaire, comme illustré au bloc 212 du processus 200.

[0037] Si aucune menace n'est détectée, le processus 200 passe au bloc 214, où le processus demande si les instructions de script sont définies pour l'état actuel du dispositif cible 102. Si les instructions de script sont définies, alors le processus passe au bloc 216 où les instructions de script sont exécutées. Si les instructions de script ne sont pas définies, le processus passe au bloc 218, où le prochain événement intéressant ou le prochain événement spécifié est pris en charge pour être traité par le module de corrélation d'événements 106D.

[0038] Une fois que les instructions de script sont exécutées au bloc 216, le processus passe au bloc 220 où le processus demande si une vérification de condition d'état spécifique est définie pour l'état actuel après l'exécution des instructions de script au bloc 216. Si une telle vérification pour cet état actuel est disponible dans le référentiel de l'analyseur de comportement 108, alors le processus passe au bloc 222 où la vérification est effectuée par le module de corrélation d'événements 106D, après quoi le processus est redirigé vers le bloc 208 et les étapes susmentionnées sont répétées. Si une telle vérification n'est pas disponible, le processus passe au bloc 218, où le prochain événement intéressant ou le prochain événement spécifié est pris en charge pour être traité par le module de corrélation d'événements 106D.

[0039] La figure 3 illustre un procédé de préparation des règles d'un analyseur de comportement. La méthode est exécutée sur un processeur. Le procédé comprend principalement l'exécution d'une classe d'échantillons de logiciels malveillants connus au bloc 302. Au bloc 304, le procédé comprend la collecte de tous les événements liés à l'exécution des échantillons. En outre, au bloc 306, le procédé comprend l'exécution d'une classe de logiciels de confiance. Au bloc 308, le procédé comprend la collecte de tous les événements liés au logiciel de confiance. Au bloc 310, le procédé comprend l'identification d'un ensemble d'événements et de conditions caractérisant la classe d'échantillons de logiciels malveillants connus. Au bloc 312, le procédé comprend la génération d'une table de transition d'états d'événements et de conditions pour l'analyse de logiciels malveillants comportementaux avec un taux attendu de faux positifs. Au bloc 314, le procédé comprend l'ajout d'instructions aux états du graphique pour répondre à une activité suspecte et malveillante. En outre, au bloc 316, le procédé comprend la mise à jour des règles de l'analyseur de comportement sur un système informatique cible comprenant une liste d'événements, des attributs d'événements pour le filtrage d'événements, une table de transition d'états et des instructions sous forme de scripts. Au bloc 318, le procédé comprend l'exécution des scripts sur l'interpréteur de script d'analyse de comportement sur une machine virtuelle dédiée pour détecter les logiciels malveillants sur un dispositif cible.

[0040] La figure 4 illustre un schéma fonctionnel représentant un procédé 400 d'identification d'une activité malveillante sur un dispositif cible en fonction de l'analyse du comportement du dispositif cible, conformément à un mode de réalisation de la présente divulgation. Il est à noter que l'ordre dans lequel le procédé 400 est décrit n'est pas destiné à être interprété comme une limitation, et n'importe quel nombre de blocs du procédé décrit peut être combiné dans n'importe quel ordre pour mettre en oeuvre le procédé ou le système.

[0041] Le procédé 400, conformément à un mode de réalisation du présent objet, peut être mis en oeuvre par le système 100. Au bloc 402, le procédé 400 comprend la récupération d'événements du dispositif cible 102 au niveau d'une machine virtuelle 104. La machine virtuelle 104 peut être couplée de manière communicative au dispositif cible 102 via Internet. En outre, dans une mise en oeuvre, la machine virtuelle 104 peut être dotée d'un accès direct à distance à la mémoire du dispositif cible 102 pour en extraire des paramètres d'événements système sans avoir d'impact sur la vitesse de fonctionnement et les ressources informatiques du dispositif cible 102.

[0042] Au bloc 404, le procédé 400 comprend le filtrage des événements par des paramètres associés aux événements du dispositif cible 102 pour identifier au moins un événement d'intérêt. Dans un mode de réalisation, l'événement d'intérêt est un événement qui expose le dispositif cible à une chance d'attaque par un logiciel malveillant ou à une activité malveillante. Le terme événement d'intérêt est également désigné de manière interchangeable comme un événement spécifié dans la présente divulgation. Dans un mode de réalisation, cette étape est effectuée au niveau du module de traitement d'événements 106D.

[0043] Au bloc 406, le procédé 400 comprend le changement d'un état de la machine virtuelle à un état attendu correspondant à l'au moins un événement d'intérêt pour une comparaison avec un état en temps réel du dispositif cible correspondant à l'au moins un événement d'intérêt. Dans un mode de réalisation, cette étape est exécutée par le module de corrélation d'événements 106.

[0044] Au bloc 408, le procédé 400 comprend l'analyse de l'état attendu et de l'état en temps réel. Dans un mode de réalisation, cette étape est exécutée par le module de corrélation d'événements 106D.

[0045] Au bloc 410, le procédé 400 comprend l'identification des divergences entre l'état attendu et l'état en temps réel pour déterminer une instance d'activité malveillante sur le dispositif cible. Dans un mode de réalisation, le module de traitement des menaces 106D est configuré pour identifier les écarts entre l'état attendu et l'état en temps réel pour déterminer un cas d'activité malveillante sur le dispositif cible. Si le module de gestion des menaces 106D détermine qu'un cas d'activité malveillante s'est produit sur le dispositif cible 102, le module de gestion des menaces 106D peut en outre être configuré pour informer un micrologiciel de solutions de sécurité anti-malware sur le dispositif cible 102 du cas d'activité malveillante pour une action supplémentaire, comme illustré au bloc 212 du procédé 200.

[0046] Ainsi, des mises en oeuvre particulières de l'objet ont été décrites. D'autres mises en oeuvre entrent dans le cadre des revendications suivantes. Dans certains cas, les actions récitées dans les revendications peuvent être exécutées dans un ordre différent tout en obtenant des résultats souhaitables. Dans certaines mises en oeuvre, le multitâche et le traitement parallèle peuvent être avantageux.

Revendications

1. Procédé pour identifier une activité malveillante sur un dispositif cible sur la base d'une analyse du comportement du dispositif cible, le procédé étant exécuté sur un processeur d'un dispositif cible en utilisant une machine virtuelle fonctionnant sur un dispositif cible pour exécuter un ensemble d'instructions de script, le procédé comprenant :
   a) la récupération d'un ou plusieurs événements système du dispositif cible au niveau d'une unité de traitement d'événements, dans laquelle les événements système caractérisent le comportement du système ;
   b) filtrer les événements système par des paramètres d'événements système au niveau d'une unité de traitement d'événements pour identifier au moins un événement intéressant, dans lequel l'événement intéressant est associé à la table de transition d'état ;

c) changer un état de la machine à états finis au niveau d'un corrélateur d'événements à un état attendu correspondant à l'au moins un événement d'intérêt conformément à la table de transition d'état ;
d) exécuter des instructions définies pour l'état attendu au niveau d'un gestionnaire de menaces ; et
e) répéter les étapes a-d, jusqu'à ce que l'état attendu caractérise une activité malveillante sur un dispositif cible.

2. Procédé selon la revendication 1, dans lequel l'unité de traitement d'événements récupère des événements système à partir d'au moins une source parmi le journal des événements système, le journal des activités de l'utilisateur, les crochets système, l'intercepteur de pilotes, les journaux externes.

3. Procédé de la revendication 1, comprenant l'étape initiale consistant à initialiser l'unité de traitement d'événements, le corrélateur d'événements et le gestionnaire de menaces avec des règles d'analyseur de comportement provenant d'un référentiel de règles d'analyseur de comportement.

4. Le procédé de la revendication 1, comprenant en outre la mise à jour d'un référentiel de règles d'analyseur de comportement.

5. Le procédé de la revendication 4, dans lequel l'analyseur de comportement continue à analyser le comportement du système à partir de l'état antérieur à la mise à jour.

6. La méthode selon la revendication 1, dans laquelle le au moins un événement d'intérêt est un événement qui expose le dispositif cible à une chance d'une attaque de logiciel malveillant ou d'une activité malveillante.

7. Système pour identifier une activité malveillante sur un dispositif cible en fonction de l'analyse du comportement du dispositif cible, dans lequel au moins une machine virtuelle fonctionne sur le dispositif cible et exécute un ensemble d'instructions de script, le système comprenant:
un référentiel de mise à jour d'analyse comportementale; et
un analyseur de comportement comprenant:
a) un module de traitement d'événements pour récupérer des événements système du dispositif cible et pour filtrer l'événement système en appliquant un ensemble de filtres aux paramètres système pour identifier au moins un événement intéressant, dans lequel l'événement intéressant est associé à la table de transition d'état;
c) un module de corrélation d'événements configuré pour faire transiter un état de la machine à états finis vers un état attendu correspondant à l'au moins un événement d'intérêt conformément à la table de transition d'état, dans lequel le corrélateur d'événements traite une séquence d'événements d'intérêt et détecte l'activité malveillante lors de la transition vers un état prédéfini comme une menace ;
d) un gestionnaire de menace configuré pour exécuter des instructions définies pour l'état attendu.

8. Le système de la revendication 7, comprenant en outre un référentiel d'analyseur de comportement couplé de manière communicative à la machine virtuelle, dans lequel le référentiel d'analyseur de comportement comprend des informations associées à des états finis attendus correspondant à une pluralité d'événements.

9. Le système selon la revendication 7, dans lequel l'au moins un événement d'intérêt est un événement qui expose le dispositif cible à une chance d'une attaque de logiciel malveillant ou d'une activité malveillante.

10. Le procédé, exécuté sur un processeur, pour préparer un ensemble de règles d'analyseur de comportement comprenant :
a. l'exécution d'un ou plusieurs échantillons d'une classe connue de logiciels malveillants ;
b. récupérer des événements système associés à l'exécution d'un ou plusieurs échantillons ;
c. l'exécution d'un ou plusieurs échantillons d'au moins une application sur liste blanche ;
d. récupération d'un ou plusieurs événements système associés à l'exécution de l'au moins une application de la liste blanche ;
e. identifier un ensemble d'événements et de conditions caractérisant la classe d'échantillons de logiciels malveillants connus;
f. générer une table de transition d'états d'événements et de conditions pour l'analyse de logiciels malveillants comportementaux avec un taux attendu de faux positifs;
g. définir des instructions aux états de la table de transition pour répondre à une activité suspecte et malveillante;
h. la mise à jour des règles de l'analyseur de comportement sur un système informatique cible, y compris la liste des événements, les attributs des événements pour le filtrage des événements, la table de transition d'état et les instructions sous forme de scripts.

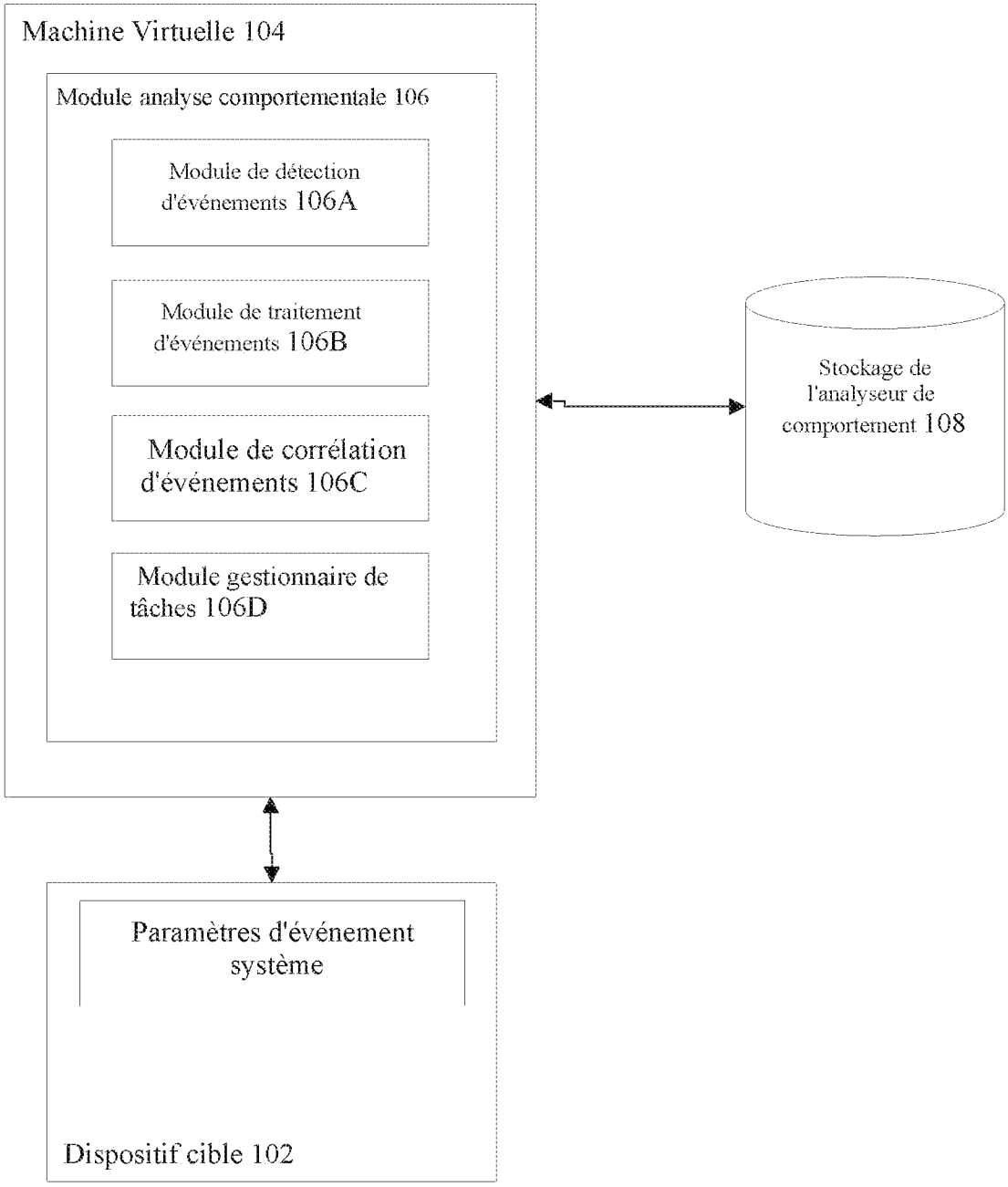
Machine Virtuelle 104
Module analyse comportementale 106
Module de détection d'événements 106A
Module de traitement d'événements 106B
Module de corrélation d'événements 106C
Module gestionnaire de tâches 106D
Stockage de l'analyseur de comportement 108
Paramètres d'événement système
Dispositif cible 102

FIG. 1

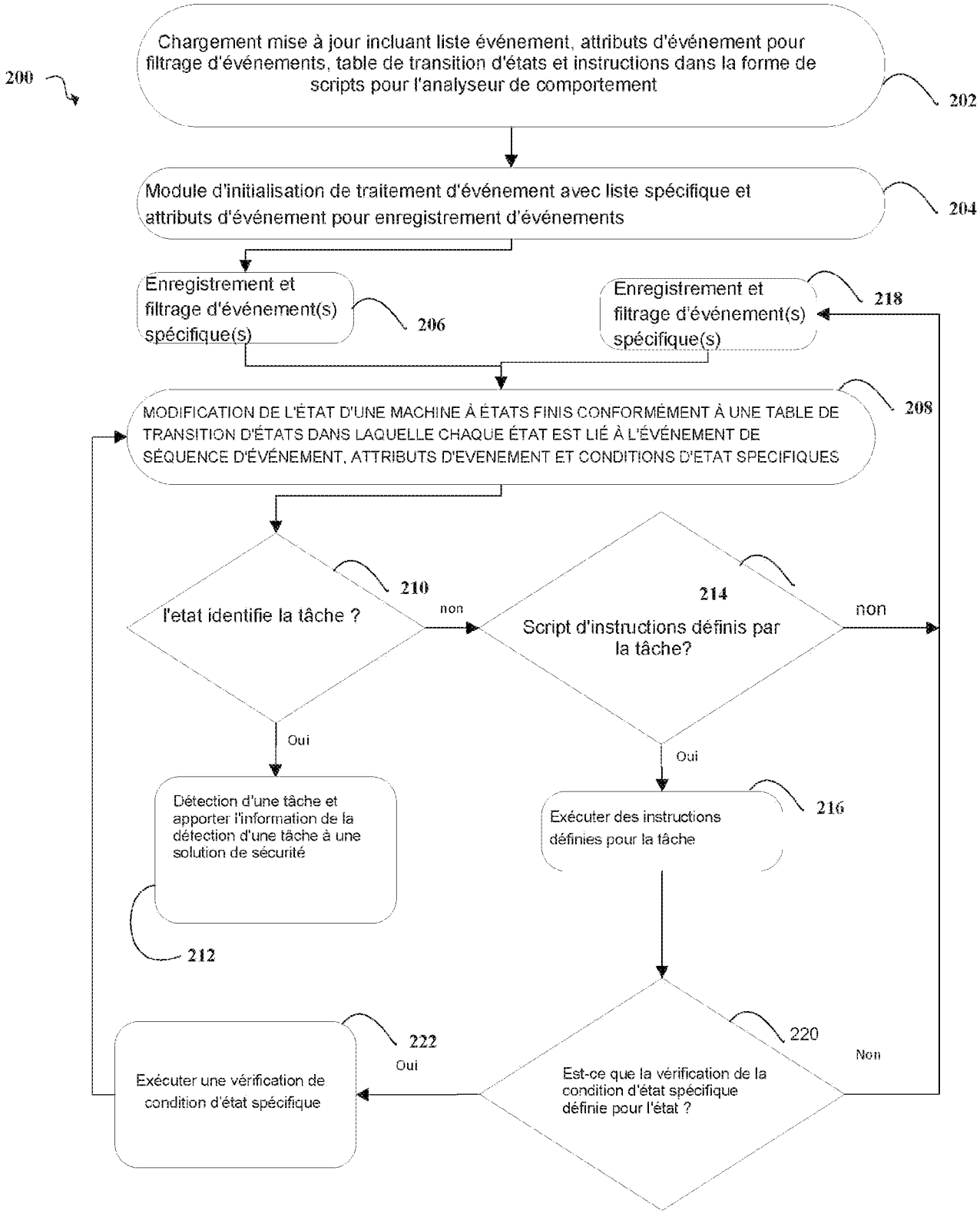
200
Chargement mise à jour incluant liste événement, attributs d'événement pour filtrage d'événements, table de transition d'états et instructions dans la forme de scripts pour l'analyseur de comportement
202
Module d'initialisation de traitement d'événement avec liste spécifique et attributs d'événement pour enregistrement d'événements
204
Enregistrement et filtrage d'événement(s) spécifique(s)
206
Enregistrement et filtrage d'événement(s) spécifique(s)
218
MODIFICATION DE L'ÉTAT D'UNE MACHINE À ÉTATS FINIS CONFORMÉMENT À UNE TABLE DE TRANSITION D'ÉTATS DANS LAQUELLE CHAQUE ÉTAT EST LIÉ À L'ÉVÉNEMENT DE SÉQUENCE D'ÉVÉNEMENT, ATTRIBUTS D'EVENEMENT ET CONDITIONS D'ETAT SPECIFIQUES
208
l'etat identifie la tâche ?
210
non
Script d'instructions définis par la tâche?
214
non
Oui
Oui
Détection d'une tâche et apporter l'information de la détection d'une tâche à une solution de sécurité
212
Exécuter des instructions définies pour la tâche
216
Est-ce que la vérification de la condition d'état spécifique définie pour l'état ?
220
Non
Oui
Exécuter une vérification de condition d'état spécifique
222

FIG. 2

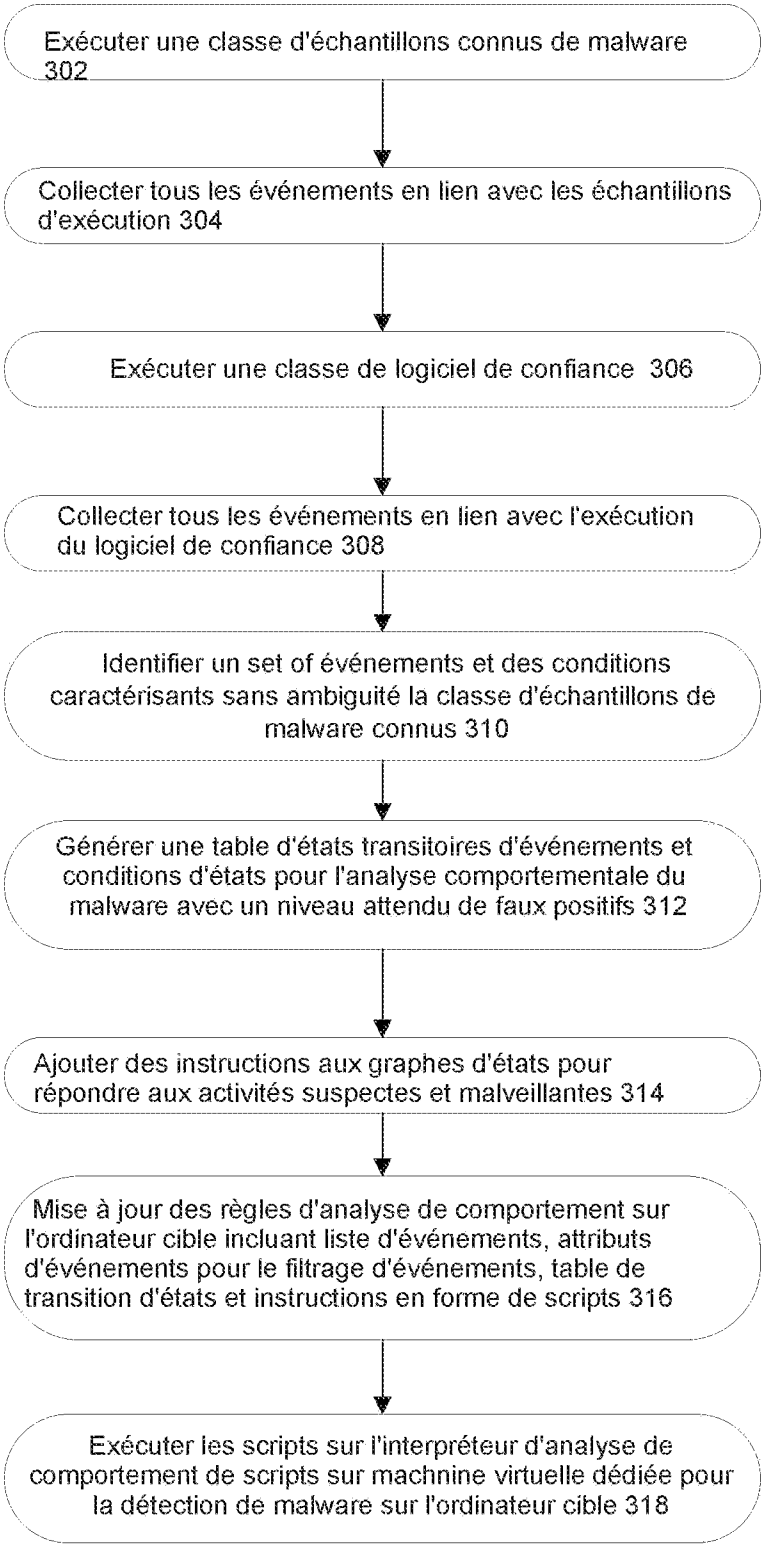
Exécuter une classe d'échantillons connus de malware 302
Collecter tous les événements en lien avec les échantillons d'exécution 304
Exécuter une classe de logiciel de confiance 306
Collecter tous les événements en lien avec l'exécution du logiciel de confiance 308
Identifier un set of événements et des conditions caractérisants sans ambiguité la classe d'échantillons de malware connus 310
Générer une table d'états transitoires d'événements et conditions d'états pour l'analyse comportementale du malware avec un niveau attendu de faux positifs 312
Ajouter des instructions aux graphes d'états pour répondre aux activités suspectes et malveillantes 314
Mise à jour des règles d'analyse de comportement sur l'ordinateur cible incluant liste d'événements, attributs d'événements pour le filtrage d'événements, table de transition d'états et instructions en forme de scripts 316
Exécuter les scripts sur l'interpréteur d'analyse de comportement de scripts sur machnine virtuelle dédiée pour la détection de malware sur l'ordinateur cible 318

FIG. 3

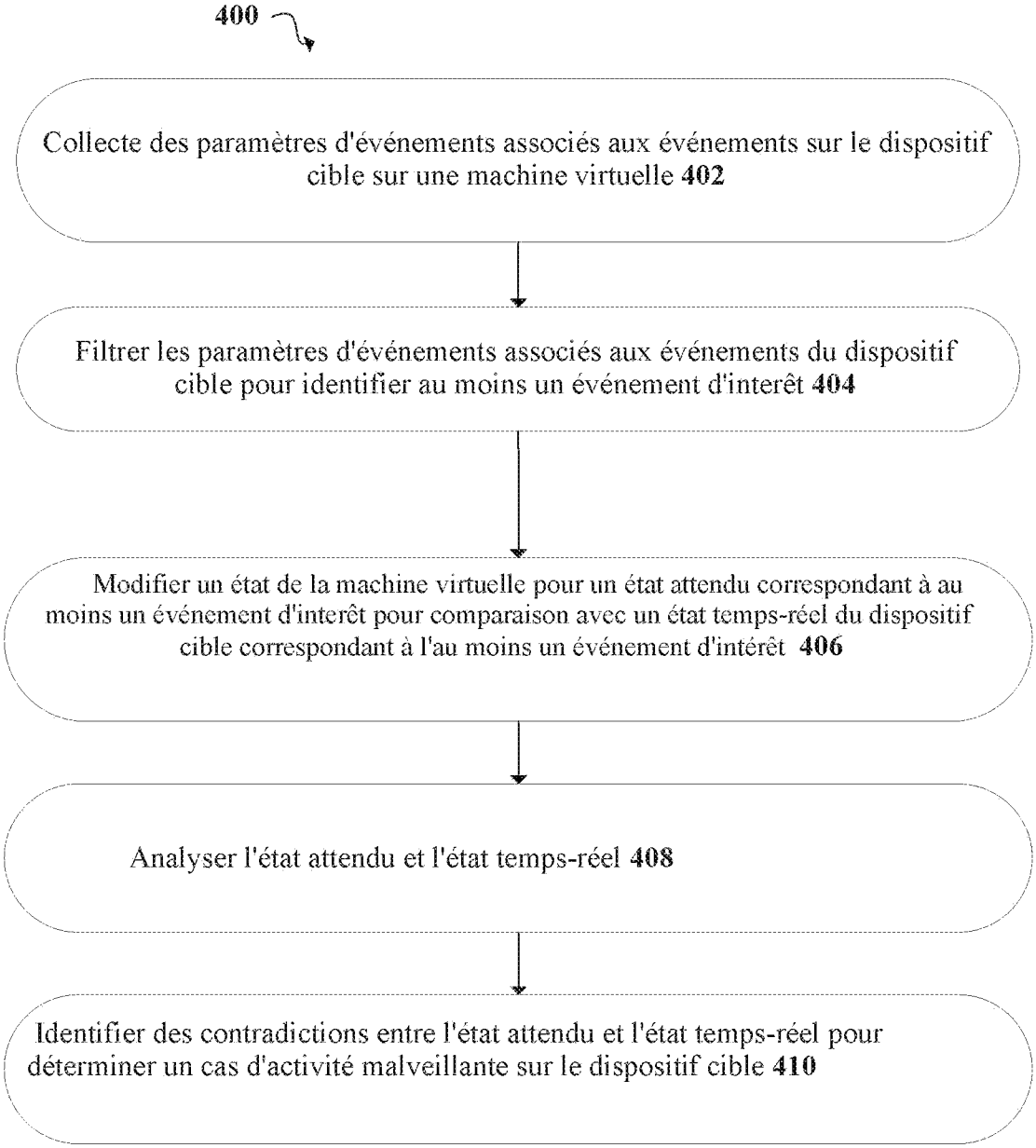
400
Collecte des paramètres d'événements associés aux événements sur le dispositif cible sur une machine virtuelle 402
Filtrer les paramètres d'événements associés aux événements du dispositif cible pour identifier au moins un événement d'interêt 404
Modifier un état de la machine virtuelle pour un état attendu correspondant à au moins un événement d'interêt pour comparaison avec un état temps-réel du dispositif cible correspondant à l'au moins un événement d'intérêt 406
Analyser l'état attendu et l'état temps-réel 408
Identifier des contradictions entre l'état attendu et l'état temps-réel pour déterminer un cas d'activité malveillante sur le dispositif cible 410

FIG. 4