(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

(43) Date de la publication internationale
18 janvier 2007 (18.01.2007)

PCT

(10) Numéro de publication internationale
WO 2007/006994 A2

(51) Classification internationale des brevets : Non classée

(21) Numéro de la demande internationale :
PCT/FR2006/050669

(22) Date de dépôt international : 4 juillet 2006 (04.07.2006)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0552098 7 juillet 2005 (07.07.2005) FR

(71) Déposant *(pour tous les États désignés sauf US)* : **FRANCE TELECOM** [FR/FR]; 6 Place d'Alleray, F-75015 Paris (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants *(pour US seulement)* : **SIBERT, Hervé** [FR/FR]; 17 rue Robert le Magnifique, F-14000 Caen (FR). **BESSON, Emmanuel** [FR/FR]; 6 allée de la Fosse au Loup, F-14200 Herouville Saint-Clair (FR). **GOUGET, Aline** [FR/FR]; 22 rue Le Brun, F-75013 Paris (FR).

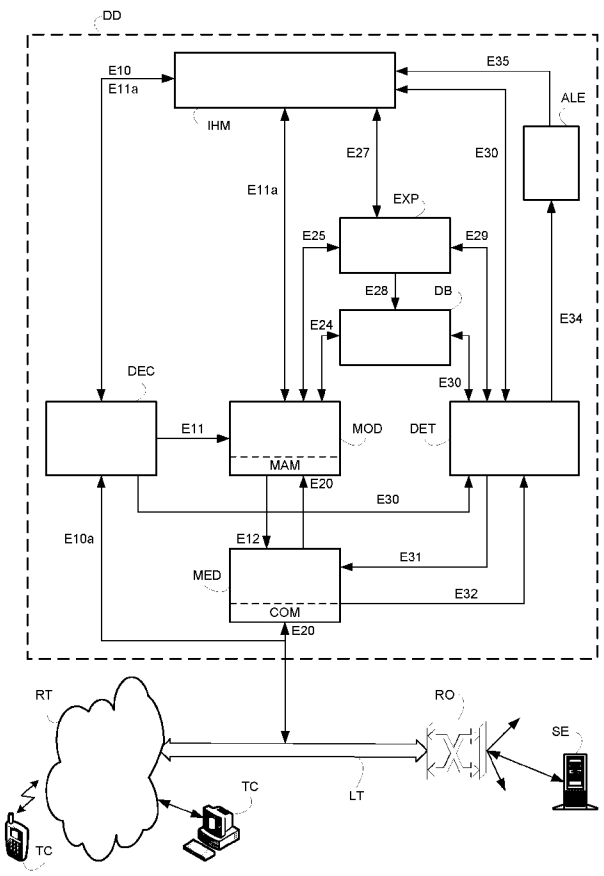
(74) Mandataire : **LAPOUX, Roland**; CABINET MARTINET & LAPOUX, 43 boulevard Vauban - BP 405 Guyancourt, F-78055 Saint Quentin Yvelines Cedex (FR).

(81) États désignés *(sauf indication contraire, pour tout titre de protection nationale disponible)* : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Suite sur la page suivante]

(54) **Title:** STATIC DETECTION OF ANOMALIES IN TRAFFIC CONCERNING A SERVICE ENTITY

(54) **Titre :** DETECTION STATIQUE D'ANOMALIES DANS LE TRAFIC RELATIF A UNE ENTITE DE SERVICE



(57) **Abstract:** The invention concerns a device for fast detection of anomalies in the traffic (LT) concerning at least one service entity (SE) following an attack of denial of service by flooding, wherein a module (MOD) provides a model of the normal activity of the entity through models for volume components of the traffic. Each model comprises a period of validity, statistical values and a conformity threshold dependent on the statistical values. A module (DET) determines for at least one evaluation of volume component at a later date a deviation of the volumic component relative to the model of the volume component and having a period of validity including the evaluation date. A module (ALE) determines a global alarm based on the deviation of the volume components for evaluation and signals an abnormal activity if the global alarm value exceeds a predetermined alarm value.

(57) **Abrégé :** Dans un dispositif détectant rapidement des anomalies dans le trafic (LT) relatif à au moins une entité de service (SE) suite à des attaques de type déni de service par inondation, un module (MOD) modélise l'activité normale de l'entité par des modèles pour des composantes volumiques du trafic. Chaque modèle comprend une période de validité, des valeurs statistiques et un seuil de conformité dépendant des valeurs statistiques. Un module (DET) détermine pour au moins une évaluation de composante volumique à une date ultérieure une déviation de la composante volumique par rapport au modèle de la composante volumique et ayant une période de validité incluant la date d'évaluation. Un module (ALE) détermine une alerte globale selon des déviations des composantes volumiques pour l'évaluation et signale une activité anormale si la valeur

[Suite sur la page suivante]

WO 2007/006994 A2

(84) États désignés *(sauf indication contraire, pour tout titre de protection régionale disponible)* : ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publiée :

— *sans rapport de recherche internationale, sera republiée dès réception de ce rapport*

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

Détection statique d'anomalies dans le trafic relatif à une entité de service

La présente invention se rapporte au domaine de la sécurité des réseaux et aux attaques de type déni de service par inondation. Plus particulièrement elle a trait à une détection d'anomalies dans le trafic supporté par une liaison de transmission et relatif à au moins une entité de service.

Les réseaux de télécommunications, comme l'internet, transmettent des données entre différentes entités de service, via une infrastructure commune. Une entité de service connectée à un tel réseau répond à des requêtes de terminaux de clients en leur fournissant un service, c'est-à-dire en effectuant des actions bien définies et demandées par les clients. Des entités de service sont par exemple un serveur web, un serveur de contenus ou de streaming proposant un téléchargement de fichiers multimédias ou diffusant des fichiers multimédias, un serveur de messagerie électronique qui relaie des messages, ou un serveur de noms de domaine DNS qui fournit des adresses IP correspondant à des noms de domaine. Ces entités de service sont, dans certains cas, des extrémités du réseau et sont localisées chez des clients d'un opérateur, tandis que d'autres entités de service, tels les serveurs DNS, sont gérées par l'opérateur du réseau lui-même.

Une attaque par déni de service est une attaque qui vise à rendre indisponible une entité de service. Il existe plusieurs types d'attaques par déni de service, par exemple des requêtes spécifiques qui attaquent directement le fonctionnement de l'entité de service en lui demandant d'effectuer une action

"non conforme". Parmi les attaques par déni de service, les attaques de type déni de service par inondation consistent à dépasser, et donc à "inonder" la capacité réseau de l'entité de service ou de la liaison de transmission par laquelle l'entité de service est reliée au réseau. Dans les deux cas, des caractéristiques volumiques du trafic réseau à destination de l'entité de service augmentent soudainement.

Afin de détecter les attaques par déni de service, il existe deux grandes familles de procédés de détection.

La première famille est relative aux détections par signature. Elle consiste à observer de manière continue le trafic à proximité d'une entité de service potentielle, et à comparer les observations avec des motifs de trafic conservés en mémoire et qui caractérisent des attaques connues. Les procédés de détection de la première famille sont particulièrement adaptés à la détection d'intrusion et la détection de dénis de service non basés sur l'inondation des entités de service.

L'invention concerne davantage la deuxième famille qui est relative aux détections d'anomalies. Une anomalie est une évaluation de trafic non conforme à un ensemble d'évaluations de trafic normal admissibles. L'ensemble des évaluations admissibles est déterminé a priori à l'aide de règles et de connaissances expertes. Il existe plusieurs types de règles telles qu'une liste de contrôle d'accès et des politiques de sécurité, fixées par l'opérateur ou le client qui possède une entité de service. Ces règles définissent des caractéristiques que le trafic normal doit satisfaire ou au contraire ne pas satisfaire. Cependant, ces règles sont insuffisantes : les

entités de service les plus sensibles et les plus visées par la cyber-criminalité sont aussi les plus difficiles à protéger a priori ; en particulier, on traite le trafic d'entités de service dont les clients ne sont pas connus à l'avance, et les règles que l'on peut fixer a priori sont limitées. Une attaque peut en effet se conformer à la plupart des règles fixées a priori dans la forme du trafic que l'attaque envoie pour inonder l'entité de service. Pour éviter cela, il est nécessaire de fixer en plus des règles de type "seuil" sur des composantes volumiques du trafic. Pour une bonne qualité de détection, ces seuils doivent prendre en compte le trafic de l'entité de service en temps normal, ce qui implique une phase d'apprentissage du trafic normal par le procédé de détection, le succès de l'attaque étant fondé sur la quantité de trafic d'attaque effectivement parvenu à l'entité de service pour traitement. Si les règles sont des seuils de trafic, leur efficacité et leur pertinence dépend de la différence entre ces seuils et l'activité du service au moment de l'attaque. Il est donc important que de telles règles prennent en compte l'activité du service, et des règles définies a priori sont insuffisantes.

La détection d'anomalies comportementales consiste à modéliser l'activité de l'entité de service à protéger en l'observant et en modélisant son comportement. Peu de procédés de détection d'anomalies connus sont orientés spécifiquement vers les attaques par inondation. Etant donné des évaluations et un modèle de comportement de l'entité de service, ces procédés font appel à des variables statistiques pour prédire la ou les prochaines évaluations. Si les prochaines évaluations effectives

divergent significativement des évaluations prédites, alors une alerte est signalée. Cependant ces procédés de détection d'anomalies détectent également des attaques autres que les attaques par inondation, induisant une perte de temps de traitement par l'opérateur de sécurité.

Par ailleurs sont connus des procédés de détection d'anomalies dédiés à la détection d'attaques par inondation. Mais leur coût est élevé et leurs algorithmes sont maintenus confidentiels. Certains de ces procédés de lutte contre les attaques par inondation sont conçus pour être installés au cœur de réseau et/ou se basent sur des informations basées sur un mécanisme de comptage installé dans des routeurs, lesquelles informations ne caractérisent que très grossièrement les attaques.

L'invention a pour **objectif** de détecter statiquement près d'une entité de service à protéger des anomalies comportementales aussi bien dans le trafic dirigé vers cette entité de service que dans le trafic qui en provient, sans recourir nécessairement à des règles prédéterminées pour trafic normal, ni à une prédiction d'évaluations de variables statistiques, de manière à signaler très rapidement et précisément une attaque de type déni de service par inondation.

Pour atteindre cet objectif, un procédé pour détecter des anomalies dans le trafic supporté par une liaison de transmission et relatif à une entité de service, incluant préalablement une modélisation de l'activité normale de l'entité de service, est caractérisé en ce que

la modélisation fournit au moins un modèle pour une composante volumique du trafic, chaque modèle comprenant une période de validité, des valeurs statistiques relatives à ladite composante volumique pendant ladite période de validité et un seuil de conformité dépendant des valeurs statistiques,

ledit procédé comprenant pour au moins une évaluation de composante volumique à une date ultérieure, une détermination d'une déviation de la composante volumique évaluée par rapport au modèle relatif à la composante volumique et ayant une période de validité incluant sensiblement la date de l'évaluation, et une détermination d'une valeur d'alerte en fonction de la déviation d'au moins une composante volumique.

L'invention détecte et caractérise une attaque de type déni de service par inondation afin de protéger les entités de service d'un opérateur ou de ses clients par exemple. Lorsqu'une attaque de type déni de service par inondation est détectée, l'invention la caractérise par des déviations des composantes volumiques évaluées qui représentent une dispersion anormale des composantes volumiques évaluées comparativement à des hétérogénéités des composantes volumiques dans les modèles d'activité normale et par suite aux moyennes des composantes volumiques dans ces modèles. L'invention quantifie ensuite l'anomalie à détecter par la valeur d'alerte dépendant des déviations des composantes volumiques pour l'évaluation et des évaluations précédentes à celle-ci.

Ainsi, l'invention permet de détecter une activité anormale dans le trafic de l'entité de service lorsque la valeur d'alerte excède un niveau

d'alerte prédéterminé. La conformité d'une évaluation de composante volumique est appréciée, contrairement à la technique antérieure, en fonction de la date de cette évaluation ; on peut ainsi avantageusement tenir compte de la variation normale du trafic sur une certaine période, par exemple au cours d'une journée.

On notera que, dans le cadre de l'invention, les évaluations de composantes volumiques peuvent porter sur le trafic à destination de l'entité de service à protéger, ou sur le trafic provenant de cette entité, ou sur ces deux trafics à la fois.

La détection d'anomalies comportementales selon l'invention est "statique" en ce sens qu'elle repose sur des périodes d'activité stables pour l'entité de service surveillée, et qu'elle détermine des modèles pertinents pour la détection comprenant des valeurs statistiques relatives aux composantes volumiques constituant des variables statiques et donc stationnaires dans le temps.

Selon une réalisation préférée, la modélisation comprend

des évaluations périodiques de composantes volumiques relatives au trafic de l'entité de service pendant plusieurs durées prédéterminées,

une détermination de périodes d'activité stable de l'entité de service de manière récursive en déterminant des valeurs statistiques de chaque composante volumique et en partitionnant les valeurs des composantes volumiques pendant chaque durée prédéterminée en fonction des valeurs statistiques en des clusters dont le plus hétérogène est sélectionné et partitionné en d'autres clusters et ainsi de suite jusqu'à ce que le dernier cluster sélectionné ait une

hétérogénéité inférieure à un seuil d'hétérogénéité prédéterminé, et

une détermination de plusieurs valeurs statistiques et un seuil de conformité pour chacune des composantes volumiques et pour chacun des clusters,

chaque modèle rassemblant ainsi un cluster d'évaluations d'une composante volumique respective pendant une période de validité respective.

Selon des dispositions particulières, afin d'assurer une détection efficace des anomalies lorsque le trafic normal relatif à l'entité de service évolue, l'invention prévoit une actualisation de modèles pour remplacer périodiquement des modèles anciens courants relatifs à l'entité de service par des modèles actualisés en fonction de valeurs récentes des composantes volumiques associées à l'entité de service. Avantageusement, les valeurs récentes des composantes volumiques évaluées sont épurées de toute valeur ayant conduit au signalement d'une activité anormale.

L'invention concerne également un dispositif pour détecter des anomalies dans le trafic supporté par une liaison de transmission et relatif à une entité de service, l'activité normale de l'entité de service étant préalablement modélisée, et une sonde de trafic incluant le dispositif pour détecter des anomalies. Le dispositif est caractérisé en ce qu'il comprend

un moyen pour modéliser l'activité normale de l'entité de service par au moins un modèle pour une composante volumique du trafic, chaque modèle comprenant une période de validité, des valeurs statistiques relatives à ladite composante volumique

pendant ladite période de validité et un seuil de conformité dépendant des valeurs statistiques,

un moyen pour déterminer pour au moins une évaluation de composante volumique à une date ultérieure, une déviation de la composante volumique évaluée par rapport au modèle relatif à la composante volumique et ayant une période de validité incluant sensiblement la date de l'évaluation, et

un moyen pour déterminer une valeur d'alerte en fonction de la déviation d'au moins une composante volumique.

Enfin, l'invention se rapporte à un programme d'ordinateur apte à être mis en œuvre dans un dispositif de détection d'anomalies selon l'invention. Le programme comprend des instructions qui, lorsque le programme est chargé et exécuté dans ledit dispositif, réalisent les étapes du procédé de détection d'anomalies selon l'invention.

D'autres caractéristiques et avantages de la présente invention apparaîtront plus clairement à la lecture de la description suivante de plusieurs réalisations préférées de l'invention, données à titre d'exemples non limitatifs, en référence aux dessins annexés correspondants dans lesquels :

- la figure 1 est un bloc-diagramme schématique d'un dispositif de détection d'anomalies statique dans un réseau de type internet, selon l'invention; et
- la figure 2 est un algorithme du procédé de détection d'anomalies statique selon l'invention.

En référence à la figure 1, un dispositif de détection d'anomalies statique DD selon l'invention est inclus dans une sonde de trafic et fait office

d'agent essentiellement logiciel pour inhiber des attaques contre des entités de service SE dans un réseau de télécommunications par paquets à débit élevé RT géré par un opérateur selon le protocole IP ("Internet Protocol"). Des données sont transmises dans une liaison de transmission LT du réseau RT constituant une partie de l'internet. Les données sont par exemple contenues dans des paquets transmis par des terminaux de client TC et destinées aux entités de service SE comprenant des serveurs web. La liaison de transmission LT est située à proximité des entités de service SE dont le trafic est à surveiller par la sonde afin que celle-ci détecte des anomalies dans le trafic qui caractérisent des comportements anormaux des clients des entités de service, particulièrement des attaques par inondation des entités de service.

Par exemple, la liaison de transmission LT à laquelle la sonde est connectée en écoute selon la figure 1, est située dans le réseau RT au point d'entrée d'une ou plusieurs entités de service sur le réseau, entre le réseau RT et le dernier routeur RO de l'opérateur du réseau RT précédant un routeur relié à une entité de service SE. La sonde n'est pas connectée à une liaison du client entre le routeur RO et l'entité de service SE dont le trafic pourrait être déjà congestionné à cause de la capacité limitée de la liaison du client.

En variante, la sonde est connectée en coupure à la liaison de transmission LT de manière également à retirer des paquets anormaux destinés à une ou plusieurs entités de service surveillées.

Le dispositif de détection DD signale une alerte à partir de l'observation du trafic relatif à l'entité ou les entités de service dans la liaison LT

afin que l'opérateur prenne des mesures adéquates pour contrer une attaque dirigée vers l'entité ou les entités de service.

Quelle que soit l'implémentation de la sonde, limitée ou non au dispositif de détection statique de l'invention, la sonde émet des alertes à partir de l'écoute du trafic en un point du réseau RT.

Comme montré à la figure 1, le **dispositif de détection statique DD** peut être sous la forme d'un ordinateur ou d'une station de travail, ou d'un système informatique local ou réparti. En relation avec l'invention, le dispositif DD comprend les modules suivants :

une interface homme - machine de management IHM, locale ou éloignée, incluant notamment un clavier et un écran, pour activer le dispositif automatiquement ou manuellement par un opérateur et saisir notamment des identificateurs et caractéristiques d'entité de service, sélectionner des données pour la détection et lire des alertes;

un module de déclaration d'entité de service DEC pour sélectionner une partie du trafic à surveiller dans la liaison LT qui est destinée à des entités de service SE à protéger;

un module de médiation MED connecté en écoute selon la figure 1, ou en coupure en variante, à la liaison de transmission LT afin d'évaluer des composantes volumiques de trafic;

un module de modélisation MOD qui construit des modèles d'activité normale pour les entités de service à protéger en fonction de valeurs de composantes volumiques de trafic évaluées par le module de médiation MED afin de produire des modèles

de comportement statique pour chaque entité de service protégée et pour chaque composante volumique;

une base de données DB pour enregistrer des modèles relatifs aux composantes volumiques des entités de service;

un module expert EXP pour introduire des connaissances expertes sur les modèles relatifs à chaque entité de service;

un module de détection d'anomalies DET détectant des anomalies dans le trafic observé destiné aux entités de service à protéger, anomalies qui sont décelées sous forme de déviations anormales de composantes volumiques de trafic évaluées par le module de médiation MED par rapport à des modèles d'activité normale; et

un module d'alerte ALE délivrant des alertes suite à des déviations anormales de composantes volumiques.

Le **procédé de détection statique d'anomalies de trafic** selon l'invention est mis en œuvre dans le dispositif de détection DD et comprend trois étapes principales E1, E2 et E3, comme montré à la figure 2. Les étapes E1, E2 et E3 comprennent respectivement des sous-étapes E10 à E12, E20 à E29 et E30 à E35. Des sous-étapes sont également indiquées dans la figure 1 au niveau de liens entre des modules du dispositif de détection DD intervenant pour l'exécution de ces sous-étapes.

La première étape E1 comporte une déclaration des entités de service SE à protéger explicitement par l'opérateur, ce qui définit un ensemble de services à protéger dispensés par ces entités.

La deuxième étape E2 modélise l'activité statique des entités de service SE à protéger. Elle

construit des modèles de comportement qui reflète une activité normale des entités de service suite à une utilisation normale de celles-ci par leurs clients. L'étape E2 considère donc des spécificités de chacune des entités de service à protéger. Des composantes volumiques prédéterminées pour la détection d'anomalies sont prédéterminées pour chaque entité de service à protéger. Puis, l'activité normale de chaque entité de service est modélisée selon les composantes volumiques prédéterminées qui sont évaluées en dépendance du trafic réel et conservées dans la base DB. Cette première modélisation considère chaque "évaluation" ou "agrégation d'évaluations" comme un point dans un espace ayant pour dimension le nombre de composantes volumiques considérées pour une entité de service donnée. Puis, l'activité de chaque entité de service est "découpée" en plusieurs modèles temporels pendant lesquels l'activité de l'entité de service est considérée comme stable, et qui sont associés à des clusters (nuages) séparés de points présents dans l'espace de dimension P.

La troisième étape E3 détecte une activité anormale dans le trafic surveillé dans la liaison LT relatif aux entités de service. L'étape E3 teste si les évaluations des composantes volumiques de trafic correspondent bien à des modèles. L'étape E3 peut aussi tester si l'apparition des évaluations des composantes volumiques notamment par rapport à un ordre des évaluations dans les modèles et des repères temporels, par exemple exprimées en heure, jour de la semaine et mois, est admissible ou non à l'aide de graphes définissant des transitions admissibles ou non entre deux modèles et de données contenues dans les modèles produits à l'étape de modélisation E2.

L'étape de détection E3 fournit un vecteur de déviation par rapport aux valeurs modélisées.

A la première sous-étape E10 de l'étape de **déclaration d'entité de service** E1, l'opérateur définit des entités de service SE qu'il souhaite protéger dans le module de déclaration DEC via l'interface homme - machine IHM. Le module DEC sélectionne ainsi une partie du trafic dans la liaison LT qui est destinée à chaque entité de service déclarée.

L'opérateur définit explicitement des entités de service en saisissant des identificateurs et des caractéristiques des entités de service à protéger transmis au module DEC, ou bien en saisissant des identificateurs des entités de service transmis au module DEC qui lit des caractéristiques des entités de service sélectionnées en correspondance aux identificateurs des entités dans une liste préenregistrée dans une base en relation avec la liaison LT. Les caractéristiques identifiant une entité de service sont par exemple des triplets. Typiquement le triplet de caractéristiques d'une entité de service inclut une adresse de destination ou classe d'adresse de destination IP, un protocole de transport et un port. Les identificateurs d'entité de service peuvent aussi être spécialisés, par exemple être les triplets de caractéristiques eux-mêmes.

Par exemple, une entité de service déclarée a pour triplet de caractéristiques (adresse de destination IP, protocole(s) de transport T, port(s) P) :

une adresse de destination IP qui est au format 'w.x.y.z/m', avec w, x, y et z compris entre 0 et

255, et m compris entre 0 et 32, et éventuellement affectée d'un masque selon la notation CIDR (Classless Internet Domain Routing); par exemple: 159.151.254.0/25 signifie de l'adresse 159.151.254.0 à l'adresse 159.151.254.127 puisque $2^{32} - 2^{25} - 1 = 127$;

une liste de protocoles de transport qui est au format 'p1;p2' avec p1, p2, … des valeurs en nombre fini prises dans des mots-clés tels que 'tcp', 'udp', 'ip', 'icmp', la valeur 'ip' recouvrant à la fois 'tcp' et 'udp'; et

une liste de ports qui est au format 'p1;p2;p3-p4' avec p1, p2, p3, p4, … des valeurs entières comprises entre 0 et 65535 et pouvant être incluses dans une plage de valeurs entières, par exemple '3200-3299', ou dans une liste de valeurs entières, par exemple '3200; 3202; 3208'.

Des entités de service "complémentaires" peuvent être implicitement déclarées à la suite de la déclaration explicite des entités de service à protéger. Une entité de service complémentaire implicite peut être une entité de service déclarée par défaut afin de recouvrir des attaques sur des protocoles spécifiques, comme les trafics indésirables selon le protocole de paquet de commande 'icmp' (Internet Control Message Protocol). Une entité de service complémentaire implicite peut être construite comme un complément à au moins une entité de service déclarée ayant une adresse IP, et relative à des trafics autres que ceux qui concernent des services hébergés par l'entité de service déclarée mais ayant ladite adresse IP et par exemple des ports différents.

Les entités de service "complémentaires" sont inférées automatiquement des caractéristiques des

entités de service déclarées pour synthétiser l'activité de la liaison LT du réseau qui n'est pas explicitement déclarée par l'opérateur. Les entités de service complémentaires à protéger ont des triplets de caractéristiques (adresse de destination IP, protocole(s) de transport T, tous ports sauf P) et (adresse de destination IP, protocole(s) de transport sauf T, port(s) P).

En variante, à une étape E10a pouvant être conjointe à l'étape E10 ou la remplacer pour certaines entités de service, la déclaration d'entités de service est automatique grâce à une écoute du trafic dans la liaison LT par le module de déclaration DEC. Dans cette variante, le module DEC est connecté en écoute selon la figure 1. Le module DEC établit une liste des entités de service demandées par les paquets qui passent dans la liaison LT et classe ces entités de service par fréquence d'apparition des paquets destinés à ces entités, afin d'obtenir une pré-liste des entités de service. Ensuite, cette liste est réduite automatiquement afin par exemple de ne conserver en mémoire du module DEC que les entités de service demandées par plus de 1% des paquets, ou bien est modifiée par l'opérateur. La liste ainsi établie est la liste des entités de service à protéger.

Après l'étape E10 ou E10a, les entités de service à protéger par le dispositif DD sont bien identifiées.

Le module de modélisation MOD est initialisé en recevant la liste des identificateurs des entités de service à protéger fournie par le module de déclaration DEC. Le module MOD reçoit la liste automatiquement à la sous-étape E11, ou en variante

après validation de l'opérateur via l'interface IHM à la sous-étape E11a.

A chaque entité de service à protéger, le module de modélisation MOD associe des **paramètres d'évaluation de trafic** par défaut qui sont une granularité et une liste par défaut de composantes volumiques de trafic à destination, ou éventuellement en provenance, de l'entité de service SE.

La granularité définit une période d'évaluation du trafic de l'entité de service et peut être une valeur par défaut dépendant d'une caractéristique de l'entité de service. Par exemple, la granularité est 10 secondes si le port P de l'entité de service est 80 (HTTP), ou 1 minute si le port est 23 (Telnet).

Les **composantes volumiques de trafic** dans la liste par défaut correspondent à des caractéristiques pouvant être relevées à partir d'informations des couches de réseau et de transport et agrégées sous forme de comptes de compteurs remis à zéro à chaque période d'évaluation de trafic définie par la granularité. Les composantes volumiques sont par exemple :

au niveau de la couche de réseau IP :

la volumétrie exprimée par un débit de trafic en octets par seconde,

la volubilité exprimée par un débit de trafic en paquets par seconde,

la connexité exprimée par un nombre d'adresses de source différentes dans des paquets destinés à l'entité de service, et

la fragmentation exprimée par un taux de paquets fragmentés en fonction des bits DF et MF dans le champ drapeau de l'entête d'un paquet IP; et

au niveau de la couche de transport TCP ("Transmission Control Protocol", mots anglais signifiant "Protocole de Commande de Transmission") :

l'ouverture de connexion exprimée par un taux de paquets incluant un bit de contrôle SYN à "1",

le "stress" exprimé par un taux de paquets incluant un bit de contrôle PUSH à "1" indiquant que les données sont à transférer à la couche supérieure,

l'urgence exprimée par un taux de paquets incluant un bit de contrôle de message urgent URG à "1", et

la volatilité exprimée par un nombre de ports différents sollicités.

A la liste par défaut de composantes volumiques de trafic peut être ajoutée une liste d'autres composantes volumiques spécifiques qui dépendent de l'entité de service, comme des grandeurs propres à un protocole applicatif utilisé. Par exemple une composante volumique à évaluer sur un service HTTP (HyperText Transfer Protocol) est le nombre de requêtes sur une méthode spécifique, comme la méthode 'GET', le nombre de requêtes vers des pages CGI (Common Gateway Interface), PHP (Personal Home Page), JSP (Java Server Page), ou un code d'erreur de type 2xx, 3xx, 4xx ou 5xx renvoyé par le serveur impliquant une mesure du trafic retour, ou la version du protocole utilisée (0.9, 1.0, 1.1).

Cependant à la sous-étape E11a, via l'interface IHM, l'opérateur peut modifier manuellement la liste de composantes volumiques ainsi que tout autre paramètre d'évaluation qui a pris une valeur par défaut, comme la granularité.

A la sous-étape E12, le module MOD transmet une liste des entités de service, éventuellement après

validation de l'opérateur, au module de médiation MED du dispositif DD. Pour l'activité normale de chaque entité de service à protéger que le module MOD doit modéliser, la liste d'entités de service comprend l'identificateur et les caractéristiques de l'entité de service et les composantes volumiques de trafic nécessaires à la modélisation afin qu'à partir de ces composantes volumiques, le module MOD produise un modèle de comportement statique de l'entité de service.

Au début E20 de l'étape de **modélisation d'activité d'entité de service** E2, le module de médiation MED extrait de chaque paquet pertinent capturé dans la liaison LT et destiné à une entité de service identifiée à protéger, notamment l'adresse de source IP, l'adresse de destination IP, la valeur du champ de protocole de transport, le port source, le port destination, la longueur totale du paquet en octets, le champ drapeau avec les bits de fragmentation, et la liste des drapeaux TCP, afin d'évaluer les composantes volumiques.

Pour chaque entité de service, le module MED comprend des compteurs COM qui respectivement comptent par exemple des octets, des paquets, des adresses de source prédéterminées et des bits de contrôle prédéterminés pour exprimer les composantes volumiques, et qui sont ainsi assignés à l'évaluation des composantes volumiques de l'entité de service. Les compteurs sont réinitialisés régulièrement à la période d'évaluation selon la granularité, par exemple de l'ordre de quelques secondes, typiquement 10 secondes. Chaque évaluation des composantes volumiques est horodatée avec la date de début de la période. Au cours de la sous-étape E20 et à

l'expiration de chaque période d'évaluation, les comptes des compteurs constituent des valeurs des composantes volumiques de trafic demandées qui sont agrégées, formatées et fournies par le module MED au module MOD.

Les composantes volumiques fournies par le module MED sont conservées temporairement dans une base d'évaluations incluse dans le module MOD. Pour une entité de service donnée, le module MOD traite une suite d'ensembles de valeurs de composantes volumiques qui ont été évaluées pendant une durée prédéterminée qui est très supérieure à la période d'évaluation de granularité retenue pour l'agrégation des composantes volumiques, ou bien définie par l'opérateur. En particulier, les modèles étant établis avec une période de validité de l'ordre de l'heure dans la journée, voire de la journée dans le mois, la durée prédéterminée pour relever les composantes volumiques est comprise entre une semaine et au moins un mois. L'ensemble des durées prédéterminées pendant lesquelles les comptes des compteurs COM sont lus et fournis au module MOD pour effectuer la première modélisation lors de la mise en marche du dispositif de détection DD s'appelle phase d'**apprentissage**. A partir de ces composantes volumiques évaluées pendant la phase d'apprentissage pour les entités de service à protéger, le module MOD produit des modèles de comportement statique de chaque entité de service protégée, comme expliqué ci-après.

Pour **modéliser** le comportement du trafic à destination de l'entité de service donnée SE et ainsi l'activité normale de celle-ci, le module MOD détermine récursivement des points de rupture dans la

suite d'ensembles de composantes volumiques pendant la phase d'apprentissage, afin de déterminer des périodes d'activité stable appelées clusters, à la sous-étape E21. Pour chaque entité de service, le module de modélisation MOD définit des périodes d'activité stable sur lesquelles les clusters s'étendent.

Par exemple, la modélisation est fondée sur une approche hybride reposant à la fois sur une définition a priori et une définition a posteriori de la période de comportement des clusters. Les composantes volumiques évaluées sont d'abord regroupées a priori en clusters distincts en fonction de leur jour de semaine. Puis pour chaque jour de la semaine, une analyse par partitionnement découvre les périodes d'activité distinctes, avec une limite élevée sur le nombre de clusters ainsi découverts qui sont suffisamment distincts les uns des autres. Cette limite découle de la finesse de modélisation souhaitée. Ainsi, le nombre de clusters possibles dans un jour peut être limité à un nombre minimum, et/ou la durée de chaque cluster au niveau temporel peut être limitée à une durée minimum.

La récursivité de la modélisation comprend par exemple un centrage - réduction de chacune x des composantes volumiques évaluées dans le cluster initial d'une journée en déterminant des valeurs statistiques de la composante volumique, comme la moyenne des valeurs évaluées de la composante volumique et l'écart-type de la composante volumique, et en remplaçant chaque valeur de la composante volumique par un rapport d'évaluation de la différence entre ladite valeur et la moyenne sur l'écart-type afin que la moyenne des rapports relatifs à la composante volumique soit nulle et leur

écart-type soit égal à 1. Des distances euclidiennes entre des rapports de deux évaluations pour toutes les composantes volumiques sont déterminées.

Puis de manière récursive, le cluster initial est partitionné en des clusters dont le plus hétérogène est partitionné en d'autres clusters, et ainsi de suite. Pour chaque cluster à partitionner, des fonctions signées sont évaluées chacune égale à une somme des rapports pour toutes les composantes volumiques relatives aux évaluations incluses dans le cluster. Chaque changement de signe dans la série chronologique des fonctions signées relatives aux évaluations incluses dans le cluster est caractérisé par un produit des fonctions signées relatives à deux évaluations successives qui est négatif. Le changement de signe indique une séparation rompant la série entre la fin d'un nouveau cluster et le début d'un nouveau cluster suivant dans le cluster à partitionner. Pour chaque nouveau cluster, une hétérogénéité est estimée égale à la somme des carrés des distances euclidiennes entre le centre de ce nouveau cluster et les rapports relatifs à toutes les composantes volumiques et aux évaluations dans ce nouveau cluster. Le cluster ayant l'hétérogénéité la plus élevée est alors sélectionné pour être partitionné en des clusters afin d'y sélectionner comme précédemment le cluster le plus hétérogène, jusqu'à ce que l'hétérogénéité d'un cluster sélectionné soit inférieure à un seuil d'hétérogénéité prédéterminé.

Le module MOD produit un ensemble de clusters qui comprennent des ensembles de composantes volumiques évaluées consécutivement ayant une hétérogénéité admissible. La reconnaissance préliminaire des comportements stables sépare ainsi

le cluster initial des valeurs des composantes volumiques évaluées en phase d'apprentissage en des clusters les plus homogènes possibles. Chaque cluster correspond à une période de validité incluse dans une période prédéterminée telle qu'une journée. En conséquence, pour une phase apprentissage sur plusieurs semaines, chaque journée dispose d'un ensemble de clusters.

A la sous-étape E22, pour chacun des clusters ainsi déterminés et pour chacune x des composantes volumiques, le module MOD calcule des valeurs statistiques, par exemple la moyenne μ_x, l'écart-type σ_x, la population égale au nombre d'évaluations dans le cluster et un seuil de conformité S_x dépendant des valeurs statistiques précédentes calculées. Le seuil de conformité S_x pour une composante volumique x est par exemple égal à la somme de la moyenne μ_x de la composante volumique et du produit de l'écart-type σ_x de la composante volumique par la racine carrée du rapport de l'hétérogénéité du cluster à la fin du partitionnement sur le nombre d'évaluations dans le cluster.

Finalement à la sous-étape E23, le module MOD rassemble chaque cluster et chaque composante volumique x pour produire un **modèle** qui est une liste comprenant l'identificateur de l'entité de service, une période de validité calculée à partir des dates de début et de fin des évaluations qui se trouvent dans le cluster, par exemple le lundi entre 9 h et 10 h 30, la granularité dans le cluster, la date de création du modèle, la désignation de la composante volumique x et les valeurs statistiques, comme la moyenne μ_x et l'écart-type σ_x, relatives à la composante volumique x dans le cluster, et le seuil de conformité S_x dépendant des valeurs statistiques

précédentes et relatif à la composante volumique. Le cluster est ainsi défini par les données des modèles des différentes composantes volumiques, et les évaluations qui ont présidé à sa création peuvent être effacées dans le module MOD. Les modèles de tous les clusters sont transférés à la base de données DB qui les enregistre à la sous-étape E24 et/ou au module expert EXP décrit plus loin à la sous-étape E25.

Le module de modélisation MOD comprend un sous-module **d'actualisation de modèles** MAM qui périodiquement lit les modèles anciens courants relatifs à une entité de service donnée dans la base DB pour fournir des modèles actualisés remplaçant les modèles anciens courants. Ce remplacement de modèles vise à refléter l'évolution potentielle des usages des clients en matière de communication avec les entités de service protégées SE et donc l'évolution du trafic réel dans la liaison LT entre les terminaux TC et les entités de service protégées. La période d'actualisation peut être une durée sensiblement égale à celle de la phase d'apprentissage.

Pour une **actualisation de modèles** E26 répétant au moins les sous-étape E20 à E25, le module MOD produit des modèles actualisés en fonction de valeurs récentes des composantes volumiques associées à l'entité de service, évaluées avec la granularité définie dans lesdits modèles courants et fournies par le module MED depuis la dernière actualisation, comme à l'étape E20. Les valeurs récentes des composantes volumiques évaluées sont d'abord conservées temporairement dans la base d'évaluations, puis épurées de toute valeur ayant conduit à la génération d'une alerte et donc au signalement d'une activité

anormale de l'entité de service, en vérifiant que ces valeurs récentes sont conformes aux modèles anciens existants, c'est-à-dire qu'elles ne déclenchent pas d'alerte en leur appliquant une fonction de détection détaillée plus loin.

L'épuration peut cependant être sautée par décision de l'opérateur ou conformément à la définition de différents modes d'actualisation : par exemple un mode "obsolète" pour lequel l'épuration est sautée, et un mode "normal" pour lequel l'épuration est exécutée.

Ensuite, le sous-module MAM introduit la valeur récemment évaluée de la composante volumique dans l'ensemble des valeurs de la composante volumique de tout modèle associé dont la période de validité inclut sensiblement la date d'évaluation de la valeur récente de la composante volumique. Les valeurs statistiques de la composante volumique, comme la moyenne, la dispersion représentée par l'écart-type et la population, et le seuil de conformité relatifs à la composante volumique associée au modèle sont actualisés par le sous-module MAM en fonction de la valeur récente de la composante volumique.

De préférence le sous-module d'actualisation de modèles MAM actualise les valeurs statistiques y compris le seuil de conformité en appliquant une pondération aux valeurs de la composante volumique du modèle ancien courant. La pondération dépend de la population des valeurs de la composante volumique du modèle ancien qui est a priori différente de la population des valeurs de composante volumique récemment évaluées, à cause de l'épuration. Avantageusement, la pondération dans le sous-module MAM dépend de la date de création du modèle ancien courant pour conférer moins de poids au modèle ancien

courant; par exemple, le paramètre "population" est divisé par un coefficient qui croît avec un âge du modèle ancien courant.

En conséquence le module MOD établit un âge de chaque modèle enregistré exprimé en nombre de jours, en association à l'identificateur de l'entité de service, à la période de validité (heures de début et fin d'un jour de la semaine ou du mois) et à un code entier de phase indiquant si le modèle est en cours de construction pendant la phase apprentissage, en cours d'utilisation pour la détection, ou invalide temporairement, ou encore a généré une alerte récente.

Après l'actualisation, les modèles actualisés sont transférés à la base DB qui les enregistre comme à la sous-étape E24 et/ou au module expert EXP comme à la sous-étape E25, pour remplacer et supprimer les modèles anciens courants.

A la fin de chaque modélisation E23 sont éventuellement introduites des **connaissances expertes** par le module expert EXP. A une sous-étape E27 succédant à la sous-étape E25, le module expert EXP traite les modèles transmis par module MOD après la modélisation et concernant une entité de service ou plusieurs entités de service ayant des caractéristiques communes comme par exemple une adresse IP. Le module EXP crée, à partir des modèles courants, des schémas de conformité avancés qui sont utilisés conjointement aux modèles dans l'étape de détection E3.

Par exemple, pour chaque entité de service et chaque granularité, le module EXP crée un automate dont les sommets du graphe d'admissibilité sont les modèles relatifs à l'entité de service et la

granularité et dont les arêtes orientées du graphe d'admissibilité constituent une base de règles. Par exemple une arête entre des états de modèle ETa et ETb est associée à une règle du type "il y a une transition de l'état ETa vers l'état ETb si la période de validité du modèle associé à l'état ETa précède immédiatement la période de validité associée à l'état ETb". A travers l'interface homme - machine IHM, l'opérateur peut modifier l'automate en ajoutant d'autres arêtes. L'automate ainsi obtenu est associé à chaque modèle de l'entité de service pour la granularité et est enregistré dans la base de données DB à une sous-étape E28.

Alternativement à la sous-étape E27, le module EXP n'est activé que lorsque l'opérateur active le module de détection DET. Le module EXP lit alors dans la base DB les modèles associés aux paramètres d'évaluation de la détection demandée, et requiert les modifications et la validation du graphe d'admissibilité par l'opérateur, avant d'envoyer les automates associés au module DET à la sous-étape E29.

L'étape de **détection d'activité anormale** E3 comprend des sous-étapes d'initialisation E30 à E32 et des sous-étapes de comparaison aux modèles E33 à E35.

Le module de détection d'anomalies DET est activé soit automatiquement après les sous-étapes E24 et E25 de la phase d'apprentissage, c'est-à-dire après la première modélisation, soit par l'opérateur via l'interface homme - machine IHM. A la sous-étape E30, le module DET reçoit du module de déclaration DEC et/ou de l'opérateur via l'interface IHM la liste des identificateurs et caractéristiques des entités de service à protéger, et éventuellement la liste des

paramètres d'évaluation, notamment la granularité des évaluations, pour chaque entité de service. Le module DET lit dans la base de données DB tous les modèles courants correspondant aux paramètres, et si pour une entité de service aucune granularité n'est mentionnée, le module DET appelle dans la base DB les modèles courants et donc les plus récents pour cette entité de service et lit leurs granularités. Ainsi, le module DET dispose de la liste des entités de service à protéger et des modèles courants pertinents pour la détection. Le module DET charge et conserve en mémoire vive tous ces modèles courants.

Pour chaque entité de service SE que le dispositif DD doit protéger, l'identificateur et les caractéristiques de l'entité de service et les paramètres d'évaluation de trafic de cette entité nécessaires à la détection et contenus dans les modèles sont, après modification et validation éventuelles par l'opérateur, appliqués ensuite par le module DET au module de médiation MED, à la sous-étape E31.

En réponse aux caractéristiques et paramètres de l'entité de service, le module de médiation MED délivre périodiquement au module de détection DET une évaluation du trafic destiné à l'entité de service, à la sous-étape E32. Cette évaluation comprend les comptes des compteurs COM qui sont assignés à l'évaluation des composantes volumiques de l'entité de service et qui sont réinitialisés régulièrement à la période d'évaluation selon la granularité demandée pour l'entité de service.

Les composantes volumiques évaluées de l'entité de service ont des valeurs "instantanées" exprimées par les comptes des compteurs respectifs COM qui sont délivrés par le module MED pour être traités par le

module de détection DET. Suite à l'évaluation, le module DET appelle les modèles courants relatifs aux composantes volumiques. Chaque composante volumique peut être associée à au moins un modèle courant d'activité normale transféré de la base BD dans la mémoire vive. Le modèle courant d'activité normale est considéré comme pertinent s'il a une période de validité incluant la date de l'évaluation avec plus ou moins une fenêtre temporelle et donc pendant laquelle l'évaluation a été réalisée à la fenêtre temporelle près. La fenêtre temporelle permet de pallier une activité de trafic a priori normale qui intervient de manière sensiblement décalée dans le temps, et ainsi évite des fausses alertes.

A la sous-étape E33, pour chaque composante volumique x, le module DET détermine une déviation D_x de l'évaluation par rapport à chaque modèle d'activité courant pertinent. La déviation D_x est par exemple le rapport entre la distance (valeur absolue de la différence) entre la valeur instantanée x de la composante volumique évaluée et sa moyenne μ_x dans le modèle, et la distance (valeur absolue de la différence) entre le seuil de conformité S_x dans le modèle et la moyenne μ_x dans le modèle. Le module DET détermine aussi une déviation globale DG en fonction des déviations pour les composantes volumiques évaluées, par exemple égale à la moyenne quadratique des valeurs des déviations.

Si des connaissances expertes ont été introduites selon la sous-étape E29, le module DET bénéficie des connaissances expertes pour réduire le nombre de faux positifs qui devraient conduire à des alertes qui n'en sont pas. Par exemple, lorsqu'un graphe d'admissibilité a été créé par le module expert EXP comme décrit à la sous-étape E27, les

connaissances expertes dépendent des valeurs statistiques de modèles proches, c'est-à-dire dont la distance dans le graphe audit modèle courant est faible, comme par exemple les modèles séparés du modèle courant par une arête, et dont les périodes de validité sont peu éloignées de la date d'évaluation de la valeur de déviation de la composante volumique associée au modèle courant, afin que les connaissances expertes soient utilisées pour affiner, par exemple diminuer, la valeur de déviation. Pour le calcul du rapport entre les distances pour la valeur de déviation D_X, la valeur du seuil de conformité S_C du modèle peut être par exemple remplacée par le plus grand des seuils de conformité dans les modèles proches dudit modèle courant.

Puis à la sous-étape E34, pour chaque évaluation, le module DET regroupe les valeurs des déviations D_X des composantes volumiques et de la déviation globale DG dans un vecteur d'alerte qui représente ainsi la similarité plus ou moins prononcée de l'évaluation aux modèles courants pertinents pour cette évaluation. Le vecteur d'alerte est délivré au module d'alerte ALE chargé de la sortie des alertes.

Finalement à la sous-étape E35, le module d'alerte ALE détermine une valeur d'alerte globale en examinant ledit vecteur d'alerte. Si la valeur d'alerte globale excède un niveau d'alerte prédéterminé, le module ALE signale une activité anormale dans le trafic de l'entité de service SE en transmettant une alerte pour l'opérateur via l'interface IHM, et/ou vers un dispositif externe. L'alerte transmise est accompagnée des valeurs des composantes volumiques de l'évaluation qui a déclenché l'alerte, des seuils de conformité

pertinents au moment de l'évaluation, et d'un type d'alerte dépendant des valeurs des déviations des composantes volumiques.

En pratique, le procédé de détection d'anomalies selon l'invention est prévu pour pouvoir détecter des anomalies relatives au trafic de plusieurs entités de service supporté par la liaison de transmission. Préalablement, chaque entité de service est déclarée par une adresse de destination, au moins un protocole de transport et au moins un port et une liste de composantes volumiques à évaluer selon une période d'évaluations prédéterminée.

L'invention décrite ici concerne un procédé et un dispositif informatique DD pour détecter des anomalies dans le trafic supporté par la liaison de transmission et relatif à une ou plusieurs entités de service SE. Selon une implémentation préférée, les étapes du procédé de l'invention sont déterminées par les instructions d'un programme d'ordinateur incorporé dans le dispositif informatique. Le programme comporte des instructions de programme qui, lorsque ledit programme est chargé et exécuté dans le dispositif, dont le fonctionnement est alors commandé par l'exécution du programme, réalisent les étapes du procédé selon l'invention.

En conséquence, l'invention s'applique également à un programme d'ordinateur, notamment un programme d'ordinateur sur ou dans un support d'informations, adapté à mettre en œuvre l'invention. Ce programme peut utiliser n'importe quel langage de programmation, et être sous la forme de code source, code objet, ou de code intermédiaire entre code source et code objet tel que dans une forme

partiellement compilée, ou dans n'importe quelle autre forme souhaitable pour implémenter le procédé selon l'invention.

Le support d'informations peut être n'importe quelle entité ou dispositif capable de stocker le programme. Par exemple, le support peut comporter un moyen de stockage ou support d'enregistrement, tel qu'une ROM, par exemple un CD ROM ou une ROM de circuit microélectronique, ou encore une clé USB, ou encore un moyen d'enregistrement magnétique, par exemple une disquette (floppy disc) ou un disque dur.

D'autre part, le support d'informations peut être un support transmissible tel qu'un signal électrique ou optique, qui peut être acheminé via un câble électrique ou optique, par radio ou par d'autres moyens. Le programme selon l'invention peut être en particulier téléchargé sur un réseau de type internet.

Alternativement, le support d'informations peut être un circuit intégré dans lequel le programme est incorporé, le circuit étant adapté pour exécuter ou pour être utilisé dans l'exécution du procédé selon l'invention.

REVENDICATIONS

1 - Procédé pour détecter des anomalies dans le trafic supporté par une liaison de transmission (LT) et relatif à une entité de service (SE), incluant préalablement une modélisation de l'activité normale de l'entité de service, caractérisé en ce que

la modélisation (E2) fournit au moins un modèle pour une composante volumique du trafic, chaque modèle comprenant une période de validité, des valeurs statistiques relatives à ladite composante volumique pendant ladite période de validité et un seuil de conformité dépendant des valeurs statistiques,

ledit procédé comprenant pour au moins une évaluation de composante volumique à une date ultérieure, une détermination (E31 - E33) d'une déviation de la composante volumique évaluée par rapport au modèle relatif à la composante volumique et ayant une période de validité incluant sensiblement la date de l'évaluation, et une détermination (E34, E35) d'une valeur d'alerte en fonction de la déviation d'au moins une composante volumique.

2 - Procédé conforme à la revendication 1, selon lequel la déviation de la composante volumique dans un modèle est le rapport entre la distance entre la composante volumique évaluée et la moyenne de la composante volumique dans le modèle, et la distance entre le seuil de conformité dans le modèle et la moyenne.

3 - Procédé conforme à la revendication 1 ou 2, selon lequel la modélisation comprend

des évaluations périodiques (E20) de composantes volumiques relatives au trafic de l'entité de service (SE) pendant plusieurs durées prédéterminées,

une détermination de périodes d'activité stable de l'entité de service de manière récursive en déterminant (E21) des valeurs statistiques de chaque composante volumique et en partitionnant les valeurs des composantes volumiques pendant chaque durée prédéterminée en fonction des valeurs statistiques en des clusters dont le plus hétérogène est sélectionné et partitionné en d'autres clusters et ainsi de suite jusqu'à ce que le dernier cluster sélectionné ait une hétérogénéité inférieure à un seuil d'hétérogénéité prédéterminé, et

une détermination (E22) de plusieurs valeurs statistiques et un seuil de conformité pour chacune des composantes volumiques et pour chacun des clusters,

chaque modèle rassemblant ainsi un cluster d'évaluations d'une composante volumique respective pendant une période de validité respective.

4 - Procédé conforme à la revendication 3, selon lequel le seuil de conformité pour une composante volumique et un cluster est égal à la somme de la moyenne de la composante volumique et du produit de l'écart-type de la composante volumique par la racine carrée du rapport de l'hétérogénéité du cluster sur le nombre d'évaluations dans le cluster.

5 - Procédé conforme à l'une quelconque des revendications 1 à 4, selon lequel la déviation de chaque composante volumique est affinée par des connaissances expertes (E27) comprenant au moins une valeur parmi lesdites valeurs statistiques et lesdits

seuils de conformité de modèles dont les périodes de validité sont peu éloignées de la date d'évaluation de la valeur de déviation.

6 - Procédé conforme à l'une quelconque des revendications 1 à 5, comprenant une actualisation de modèles (E26) pour remplacer périodiquement des modèles anciens courants relatifs à l'entité de service par des modèles actualisés en fonction de valeurs récentes des composantes volumiques associées à l'entité de service.

7 - Procédé conforme à la revendication 6, comprenant une épuration des valeurs récentes des composantes volumiques évaluées de toute valeur ayant conduit au signalement d'une activité anormale.

8 - Procédé conforme à la revendication 6 ou 7, selon lequel les valeurs statistiques et le seuil de conformité dans un modèle sont actualisés en appliquant une pondération dépendant d'une date de création du modèle aux valeurs de la composante volumique du modèle.

9 - Procédé conforme à l'une quelconque des revendications 1 à 8, détectant des anomalies relatives au trafic de plusieurs entités de service (SE) supporté par la liaison de transmission (LT), et comprenant préalablement une déclaration de chaque entité de service par une adresse de destination, au moins un protocole de transport et au moins un port et une liste de composantes volumiques à évaluer selon une période d'évaluations prédéterminée.

10 - Dispositif (DD) pour détecter des anomalies dans le trafic supporté par une liaison de transmission (LT) et relatif à une entité de service (SE), l'activité normale de l'entité de service étant préalablement modélisée, caractérisé en ce qu'il comprend

un moyen (MOD) pour modéliser l'activité normale de l'entité de service par au moins un modèle pour une composante volumique du trafic, chaque modèle comprenant une période de validité, des valeurs statistiques relatives à ladite composante volumique pendant ladite période de validité et un seuil de conformité dépendant des valeurs statistiques,

un moyen (DET) pour déterminer pour au moins une évaluation de composante volumique à une date ultérieure, une déviation de la composante volumique évaluée par rapport au modèle relatif à la composante volumique et ayant une période de validité incluant sensiblement la date de l'évaluation, et

un moyen (ALE) pour déterminer une valeur d'alerte en fonction de la déviation d'au moins une composante volumique.

11 - Sonde de trafic incluant un dispositif (DD) conforme à la revendication 10 pour détecter des anomalies.

12 - Programme d'ordinateur apte à être mis en œuvre dans un dispositif (DD) pour détecter des anomalies dans le trafic supporté par une liaison de transmission (LT) et relatif à une entité de service (SE), l'activité normale de l'entité de service étant préalablement modélisée, ledit programme comprenant des instructions qui, lorsque le programme est chargé

et exécuté dans ledit dispositif, réalisent les étapes consistant à:

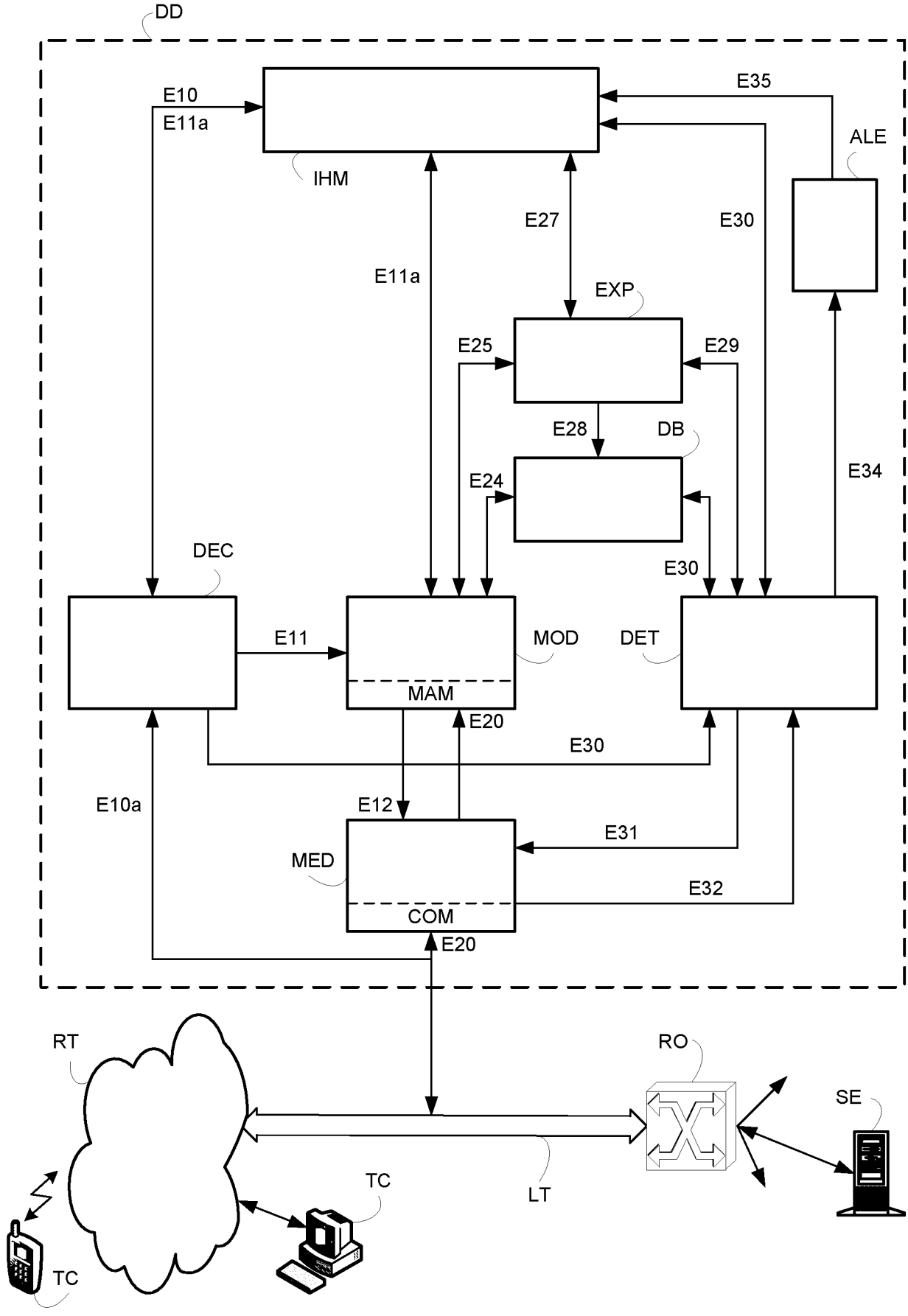
modéliser (E2) l'activité normale de l'entité de service par au moins un modèle pour une composante volumique du trafic, chaque modèle comprenant une période de validité, des valeurs statistiques relatives à ladite composante volumique pendant ladite période de validité et un seuil de conformité dépendant des valeurs statistiques,

déterminer (E31 - E33) pour au moins une évaluation de composante volumique à une date ultérieure, une déviation de la composante volumique évaluée par rapport au modèle relatif à la composante volumique et ayant une période de validité incluant sensiblement la date de l'évaluation, et

déterminer (E34, E35) une valeur d'alerte en fonction de la déviation d'au moins une composante volumique.

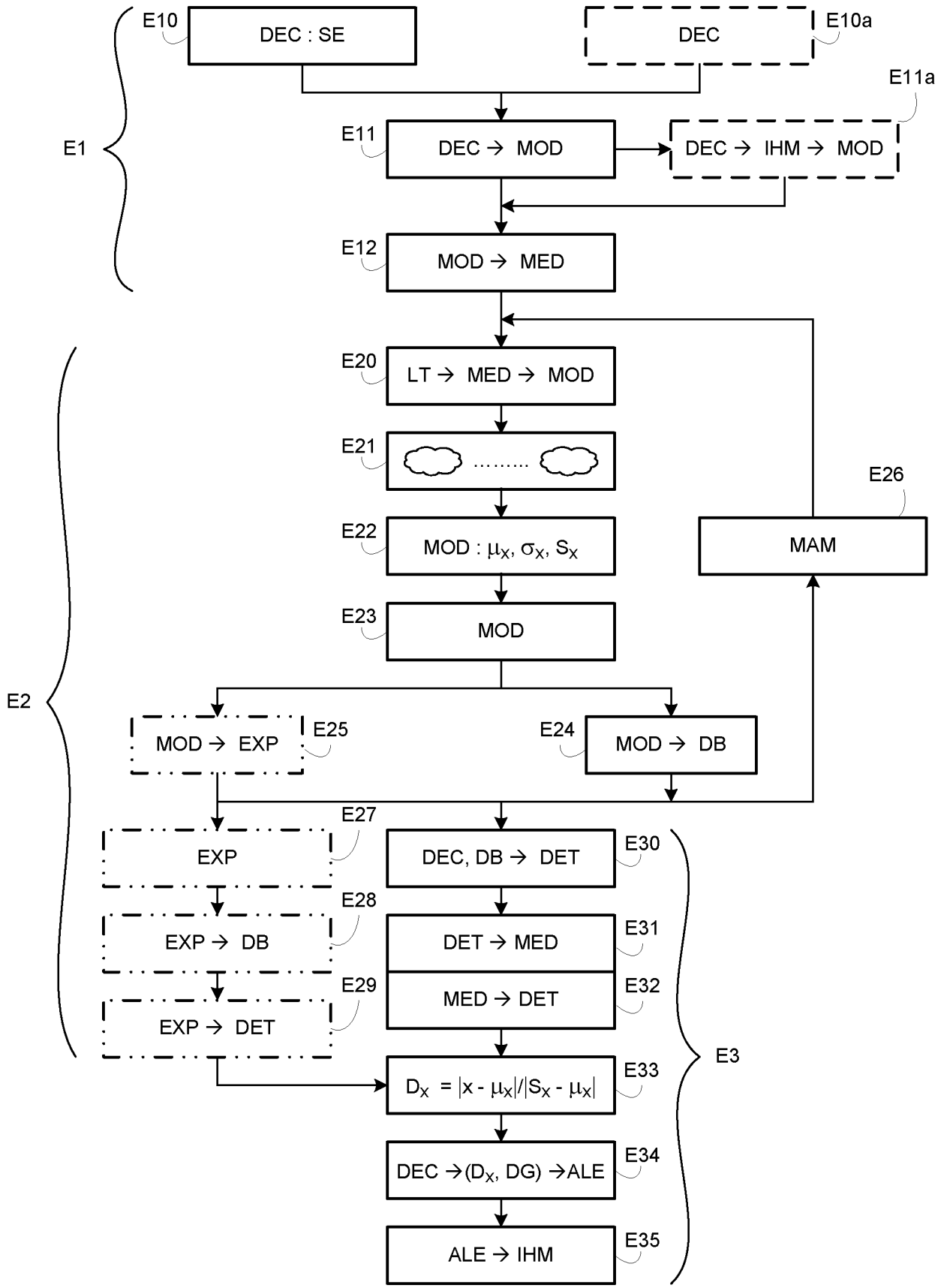
1/2

FIG. 1


DD
E10
E11a
IHM
E35
ALE
E27
E30
E11a
EXP
E25
E29
E28
DB
E24
E34
DEC
E30
E11
MOD
DET
MAM
E20
E30
E10a
E12
E31
MED
E32
COM
E20
RT
RO
SE
TC
LT
TC

2/2

FIG. 2


E10
DEC : SE
E10a
DEC
E11a
E1
E11
DEC → MOD
DEC → IHM → MOD
E12
MOD → MED
E20
LT → MED → MOD
E21
……..
E26
E22
MOD : μX, σX, SX
MAM
E23
MOD
E2
MOD → EXP
E25
E24
MOD → DB
E27
EXP
E30
DEC, DB → DET
E28
EXP → DB
E31
DET → MED
E29
EXP → DET
E32
MED → DET
E3
E33
DX = |x - μX|/|SX - μX|
E34
DEC →(DX, DG) →ALE
E35
ALE → IHM