(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 1 820 170 B1**

(12) **FASCICULE DE BREVET EUROPEEN**

(45) Date de publication et mention de la délivrance du brevet:
16.04.2008 Bulletin 2008/16

(21) Numéro de dépôt: **05819246.9**

(22) Date de dépôt: **24.11.2005**

(51) Int Cl.:
G08B 15/00 *(2006.01)*

(86) Numéro de dépôt international:
PCT/FR2005/050983

(87) Numéro de publication internationale:
WO 2006/056721 (01.06.2006 Gazette 2006/22)

(54) **SUPPRESSION DE FAUSSES ALERTES PARMI LES ALERTES PRODUITES DANS UN SYSTEME D'INFORMATIONS SURVEILLE**

UNTERDRÜCKUNG VON FALSCHALARMEN UNTER IN EINEM ÜBERWACHTEN INFORMATIONSSYSTEM PRODUZIERTEN ALARMEN

SUPPRESSION OF FALSE ALARMS AMONG ALARMS PRODUCED IN A MONITORED INFORMATION SYSTEM

(84) Etats contractants désignés:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI SK TR

(30) Priorité: **26.11.2004 FR 0412559**

(43) Date de publication de la demande:
22.08.2007 Bulletin 2007/34

(73) Titulaire: **France Telecom S.A.**
75015 Paris (FR)

(72) Inventeurs:
- **MORIN, Benjamin**
  F-35000 RENNES (FR)
- **VIINIKKA, Jouni**
  F-14000 CAEN (FR)

(74) Mandataire: **Joly, Jean-Jacques et al**
Cabinet Beau de Loménie
158, rue de l'Université
75340 Paris Cédex 07 (FR)

(56) Documents cités:
EP-A- 0 856 826
GB-A- 2 258 311
US-A1- 2002 161 763

EP 1 820 170 B1

Description

Arrière-plan de l'invention

[0001] L'invention concerne un système et un procédé de suppression de fausses alertes parmi les alertes produites dans un système d'informations surveillé.

[0002] La sécurité des systèmes d'informations passe par le déploiement de systèmes de détection d'intrusions. Ces systèmes de détection d'intrusions se situent en amont des systèmes de prévention d'intrusions. Ils permettent de détecter des activités allant à l'encontre de la politique de sécurité d'un système d'informations.

[0003] Les systèmes de détection d'intrusions sont entre autres constitués de sondes de détection d'intrusions « SDI » qui émettent des alertes vers des systèmes de gestion d'alertes « SGA ».

[0004] En effet, les sondes de détection d'intrusions sont des composants actifs du système de détection d'intrusions qui analysent une ou plusieurs sources de données à la recherche d'événements caractéristiques d'une activité intrusive et émettent des alertes vers les systèmes de gestion d'alertes. Un système de gestion d'alertes SGA centralise les alertes provenant des sondes et effectue éventuellement une analyse de l'ensemble de ces alertes.

[0005] Les SGA sont constitués de plusieurs modules de traitement d'alertes « MTA », chargés de traiter les alertes en aval de leur production par les SDI. Les MTA produisent eux-mêmes des alertes de plus haut niveau traduisant leur traitement sur les alertes.

[0006] Une fois traitées par les modules du SGA, les alertes sont présentées à un opérateur de sécurité du système d'information dans une console de présentation des alertes « CPA ».

[0007] Parmi les MTA, on peut distinguer les modules de suppression de fausse alertes « MSFA », qui sont chargés d'identifier les alertes qui sont de fausses alertes « faux positifs », c'est à dire les alertes qui sont produites par les SDI alors qu'aucune activité intrusive n'a eu lieu. A l'inverse, les alertes produites suite à une activité intrusive qui a effectivement eu lieu sont des vraies alertes « vrais positifs ».

[0008] Les sondes de détection d'intrusions génèrent un très grand nombre d'alertes qui peut comprendre plusieurs milliers d'alertes par jour en fonction des configurations et de l'environnement.

[0009] Cet excès d'alertes est majoritairement lié aux fausses alertes. En général, parmi les milliers d'alertes générées quotidiennement dans un système d'informations, 90 à 99% des alertes sont des fausses alertes.

[0010] L'analyse de la cause de ces fausses alertes montre qu'il s'agit très souvent de comportements erratiques d'entités (par exemple des serveurs) du réseau surveillé, mais qui ne sont pas pertinents du point de vue de la sécurité du système d'informations. Il peut aussi s'agir de comportements normaux d'entités, dont l'activité ressemble à une activité intrusive, si bien que les sondes de détection d'intrusions (SDI) émettent des alertes par erreur.

[0011] Comme les comportements normaux constituent la majorité de l'activité d'une entité, les fausses alertes engendrées sont récurrentes et participent pour une grande part à l'excès global d'alertes.

[0012] Une mauvaise prise en compte de la politique de sécurité du système d'informations dans la configuration des SDI peut aussi engendrer de fausses alertes. Dans tous les cas, la nature (vrai ou faux positif) d'une alerte dépend en grande partie des propriétés intrinsèques du système d'informations surveillé.

[0013] Dans les SGA actuels, la qualification d'une alerte de vrai ou faux positif est laissée à l'appréciation de l'opérateur de sécurité en charge de l'analyse des alertes car c'est lui qui connaît les propriétés de son système d'informations. Comme le nombre d'alertes générées est grand, le temps consacré par l'opérateur à chaque alerte est réduit.

[0014] Il existe des techniques probabilistes de traitement des alertes issues de sondes de détection d'intrusions pour détecter de fausses alertes. Ces techniques reposent sur des connaissances préalables des propriétés des attaques qui sont référencées dans les alertes, ainsi que sur les propriétés du système d'informations surveillé.

Objet et résumé de l'invention

[0015] L'invention a pour but de remédier à ces inconvénients, et de fournir un procédé simple de suppression de fausses alertes qui ne nécessite pas de connaissances préalables et qui permette un diagnostique réel, aisé et rapide de ces alertes.

[0016] Ces buts sont atteints grâce à un procédé de suppression de fausses alertes parmi les alertes produites dans un système d'informations surveillé, dans lequel les alertes sont classées automatiquement au moyen d'un module de suppression de fausses alertes suivant deux catégories constituées de fausses et vraies alertes selon des critères déterminés basés sur un apprentissage progressif dudit module à partir de l'expertise d'un opérateur humain en charge d'un classement initial manuel des alertes, ledit apprentissage progressif comportant les phases suivantes :

- une phase d'apprentissage initial dans laquelle ledit module de suppression de fausses alertes procède à un enregistrement de diagnostics de l'opérateur humain concernant un nombre déterminé d'alertes initiales et comprenant pour une alerte initiale donnée, une extraction de l'ensemble de mots composant ladite alerte initiale donnée, et

une association à chaque mot dudit ensemble de mots, d'un compteur désignant le nombre cumulé d'occurrences dudit mot dans l'une des deux catégories, et

- une phase de validation dans laquelle ledit module de suppression de fausses alertes procède à la classification de nouvelles alertes en fonction dudit enregistrement de diagnostics et d'une supervision de l'opérateur humain qui confirme ou corrige les classifications de nouvelles alertes.

[0017] Ainsi, le procédé selon l'invention facilite le travail de l'opérateur de sécurité en permettant au MSFA d'apprendre progressivement le travail de ce dernier pour lui proposer au bout de cet apprentissage, des diagnostics automatiques sur la nature des alertes sans aucune connaissance préalable. L'apprentissage progressif et supervisé du MSFA permet une prise en compte optimale des modifications pouvant être apportées par l'opérateur de sécurité tout en permettant de mesurer de manière simple la fréquence d'apparition des mots dans les catégories de fausses et vraies alertes.

[0018] Ces critères déterminés comportent une comparaison des probabilités d'appartenance des alertes à l'une et à l'autre des deux catégories.

[0019] Ainsi, une technique probabiliste de comparaison garantie un apprentissage efficace et mesurable.

[0020] Avantageusement, dans la phase de validation, les confirmations ou corrections des classifications de nouvelles alertes apportées par l'opérateur humain sont utilisées par le module de suppression de fausses alertes pour minimiser un taux de correction lui permettant d'augmenter la fiabilité de toute classification ultérieure de nouvelles alertes.

[0021] Ainsi, le taux de correction permet de quantifier la fiabilité du classement des alertes et par conséquent d'améliorer toute classification ultérieure de nouvelles alertes.

[0022] Selon une autre particularité de l'invention, le procédé comporte une phase opérationnelle dans laquelle la classification des nouvelles alertes est effectuée de manière autonome si le taux de correction de la classification des nouvelles alertes de la phase de validation devient inférieur à un nombre seuil déterminé.

[0023] Ainsi, le taux de correction fournit un filtrage fiable pour le passage vers une phase de classification autonome des nouvelles alertes.

[0024] Selon encore une autre particularité de l'invention, dans la phase opérationnelle, les fausses alertes peuvent être supprimées ou stockées dans un moyen de stockage et seules les vraies alertes sont envoyées à une console de présentation des alertes (CPA).

[0025] Ainsi, le volume d'alertes devant être présenté à l'opérateur humain en charge de la sécurité est considérablement réduit.

[0026] La classification des alertes durant la phase de validation et la phase opérationnelle comporte, pour une nouvelle alerte donnée, les étapes suivantes :

- extraction de l'ensemble de mots composant ladite nouvelle alerte donnée,
- comparaison des probabilités d'appartenance de la nouvelle alerte donnée à l'une et à l'autre desdites catégories,
- classement de la nouvelle alerte donnée dans l'une des deux catégories selon le résultat de la comparaison de l'étape précédente,
- incrémentation des compteurs selon la catégorie de la nouvelle alerte donnée, et
- transmission de la nouvelle alerte donnée ainsi classée à la console de présentation des alertes.

[0027] Ainsi, en utilisant l'ensemble de mots constituants les alertes et leurs compteurs associés selon les étapes ci-dessus, les alertes peuvent être classées avec une fiabilité continuellement croissante.

[0028] La comparaison des probabilités d'appartenance de la nouvelle alerte donnée à l'une et à l'autre desdites catégories comporte les étapes suivantes :

- calculer pour chaque mot de l'ensemble de mots de ladite nouvelle alerte, la probabilité que chaque mot soit présent dans des alertes appartenant à l'une ou à l'autre des catégories en déterminant le rapport entre le compteur désignant le nombre cumulé d'occurrences de chaque mot dans des alertes de l'une ou l'autre des catégories et le nombre total d'occurrences de mots dans l'une ou l'autre des catégories respectivement,
- calculer la probabilité de chaque catégorie en déterminant le rapport entre le nombre total d'occurrences de mots dans des alertes de chaque catégorie et le nombre total de mots,
- calculer le produit, sur l'ensemble des mots composant l'alerte, des probabilités que chaque mot respectif de l'alerte soit présent dans des alertes appartenant à chaque catégorie multiplié par la probabilité de chaque catégorie, et
- comparer le résultat de l'étape précédente selon les deux catégories.

[0029] Ainsi, les étapes ci-dessus tirent profit des compteurs pour comparer les probabilités d'appartenance d'une alerte donnée à l'une et à l'autre des catégories avec un nombre optimal d'étapes de calcul, minimisant par conséquent le temps de calcul.

[0030] Avantageusement, la correction par le module de suppression de fausses alertes, de la classification des

nouvelles alertes durant la phase de validation comporte les étapes suivantes :

- correction de la catégorie d'une nouvelle alerte précédemment classée par ledit module, s'il reçoit une notification de l'opérateur humain indiquant que ledit précédent classement de ladite nouvelle alerte est faux,
- décrémentation des compteurs désignant les nombres cumulés d'occurrences de mots dans la catégorie faussement classée,
- incrémentation des compteurs désignant les nombres cumulés d'occurrences de mots dans la catégorie corrigée.

[0031] Ainsi, le réglage des compteurs est un moyen efficace et rapide pour améliorer l'apprentissage du MSFA.

[0032] L'invention vise aussi un module de suppression de fausses alertes comportant :

- des moyens de traitement de données permettant de classer automatiquement les alertes suivant deux catégories constituées de fausses et vraies alertes selon des critères déterminés basés sur un apprentissage progressif à partir de l'expertise d'un opérateur humain en charge d'un classement initial manuel des alertes, et
- des moyens mémoires permettant d'enregistrer, lors d'une phase d'apprentissage initial de l'apprentissage progressif, des diagnostics de l'opérateur humain concernant un nombre déterminé d'alertes initiales en permettant pour une alerte initiale donnée d'extraire l'ensemble de mots composant ladite alerte initiale donnée et en associant à chaque mot dudit ensemble de mots, un compteur désignant le nombre cumulé d'occurrences dudit mot dans l'une des deux catégories,

les moyens de traitement de données permettant en outre de procéder à la classification de nouvelles alertes en fonction dudit enregistrement de diagnostics et d'une supervision de l'opérateur humain qui confirme ou corrige les classifications de nouvelles alertes.

[0033] Avantageusement, lors d'une phase opérationnelle, les moyens de traitement de données sont destinés en outre à classifier de manière autonome de nouvelles alertes si un taux de correction de la classification des nouvelles alertes de la phase de validation devient inférieur à un nombre seuil déterminé.

[0034] De préférence, le module comporte en outre un moyen de stockage permettant de stocker, lors de la phase opérationnelle, les fausses alertes de sorte que seules les vraies alertes sont envoyées à une console de présentation des alertes.

[0035] L'invention vise aussi un système d'informations surveillé comportant un réseau interne à surveiller, des sondes de détection d'intrusions, un système de gestion d'alertes, une console de présentation d'alertes, et un module de suppression de fausses alertes selon les caractéristiques ci-dessus.

Brève description des dessins

[0036] D'autres particularités et avantages de l'invention ressortiront à la lecture de la description faite, ci-après, à titre indicatif mais non limitatif, en référence aux dessins annexés, sur lesquels :

- la figure 1 est une vue très schématique d'un système d'informations surveillé comportant un module de suppression de fausses alertes selon l'invention ; et
- la figure 2 est un organigramme très schématique illustrant les étapes d'un procédé de suppression de fausses alertes parmi les alertes produites dans un système de sécurité d'informations, selon l'invention.

Description détaillée de modes de réalisation

[0037] La figure 1 illustre un exemple très schématique d'un réseau ou d'un système d'informations surveillé 1 comportant un système de sécurité d'informations 3, une console de présentation d'alertes « CPA » 5 et un réseau interne 7 à surveiller comprenant un ensemble d'entités, par exemple des stations de travail 7a, 7b, 7c, des serveurs 7d, des proxy web 7e etc.

[0038] Le système de sécurité d'informations 3 comporte un ensemble 11 de sondes de détection d'intrusions « SDI » 11a, 11b et 11c destinées à émettre des alertes lorsque des attaques sont détectées, et un système de gestion d'alertes « SGA » 15 destiné à analyser les alertes émises par les sondes 11a, 11b et 11c et comprenant des modules de traitement d'alertes « MTA » 15a, 15b.

[0039] En outre, conformément à l'invention, le système de sécurité d'informations 3 comporte un module de suppression de fausses alertes « MSFA » 17 connecté aux sondes de détection d'intrusions 11a, 11b et 11c, au système de gestion d'alertes 15, et à la console de présentation d'alertes 5, par l'intermédiaire d'un routeur 19 d'aiguillage.

[0040] En effet, le routeur 19 est connecté au MSFA 17 via des liaisons 18a et 18b, aux sondes de détection d'intrusions 11a, 11b et 11c via des liaisons 13a, 13b et 13c, au SGA 15 via des liaisons 16a et 16b, et à la CPA 5 via une liaison 6.

[0041] Le module de suppression de fausses alertes 17 comporte des moyens de traitement 21 de données permettant de classer (c'est-à-dire marquer) automatiquement les alertes suivant deux catégories constituées de fausses et vraies alertes selon des critères déterminés basés sur un apprentissage progressif du MSFA 17 à partir de l'expertise d'un opérateur humain 23 en charge d'un classement initial manuel des alertes. Ces critères déterminés comportent une comparaison des probabilités d'appartenance des alertes à l'une et à l'autre des deux catégories. Ainsi, un programme informatique conçu pour mettre en oeuvre un procédé de suppression de fausses alertes selon la présente invention peut être exécuté par les moyens de traitement 21 du MSFA 17.

[0042] Le MSFA 17 selon l'invention est adaptatif en ce sens qu'il intègre progressivement l'expertise de l'opérateur humain 23 en charge de la qualification initiale manuelle des fausses alertes et présente trois phases successives de fonctionnement (voir aussi figure 2).

[0043] La première phase P1 est une phase d'apprentissage initial dans lequel le MSFA 17 ne marque pas les alertes, il se contente d'enregistrer les diagnostics de l'opérateur humain 23. En effet, le MSFA 17 comporte des moyens mémoires 25 permettant aux moyens de traitement 21 d'enregistrer des diagnostics de l'opérateur humain 23 concernant un nombre déterminé d'alertes initiales.

[0044] La deuxième phase P2 est une phase de validation dans laquelle les moyens de traitement 21 de données du MSFA 17 procèdent à la classification de nouvelles alertes en fonction de l'enregistrement de diagnostics et d'une supervision de l'opérateur humain 23 qui confirme ou corrige les classifications de nouvelles alertes. En effet, lorsque le nombre d'alertes ayant transité par le MSFA 17 a atteint un nombre suffisant, par exemple supérieur à un seuil fixé par l'opérateur humain 23, le MSFA 17 commence à marquer les alertes, qui lui sont présentées à travers la liaison 18a.

[0045] Avantageusement, dans la phase de validation, les confirmations ou corrections des classifications de nouvelles alertes apportées par l'opérateur humain 23 sont utilisées par le module de suppression de fausses alertes 17 pour minimiser un taux de correction lui permettant d'augmenter la fiabilité de toute classification ultérieure de nouvelles alertes.

[0046] Ainsi, les première et deuxième phases d'apprentissage initial et de validation forment un apprentissage progressif du MSFA 17.

[0047] Par ailleurs, la troisième phase P3 est une phase opérationnelle dans laquelle la classification des nouvelles alertes est effectuée de manière autonome par les moyens de traitement 21 du MSFA 17 si le taux de correction de la classification des nouvelles alertes de la phase de validation devient inférieur à un nombre seuil déterminé. Ainsi, le MSFA 17 marque les alertes et envoie seulement les vraies alertes à la console de présentation des alertes CPA 5. Les fausses alertes sont soit directement supprimées, soit stockées dans les moyens mémoires 25 ou de préférence dans un moyen de stockage 27 annexe via une liaison 26. Le choix entre la suppression ou le stockage des fausses alertes peut être déterminé par l'opérateur humain 23.

[0048] Ainsi, le MSFA 17 est destiné à traiter les alertes provenant directement des SDI 11a, 11b, 11c via les liaisons 13a, 13b, 13c et 18a ou éventuellement des autres MTA 15a, 15b via les liaisons 16b et 18a. Chaque alerte générée par un SDI 11a, 11b, 11c ou un MTA 15a, 15b est soumise au MSFA 17 pour analyse. Le MSFA 17 marque les alertes qu'il juge être de fausses alertes et les soumet (liaisons 18b, 16a) au SGA 15. L'alerte munie de son marquage est ensuite envoyée (liaisons 16b, 6) à la CPA 5 pour y être consultée par l'opérateur humain 23 de sécurité.

[0049] On notera que, au cours des deuxième et troisième phases, l'opérateur humain 23 peut toujours intervenir, par exemple via une liaison directe 8 avec le MSFA 17 pour réviser les diagnostics de ce dernier. En effet, en cas d'erreur de qualification d'une alerte par le MSFA 17, l'opérateur humain 23 a la possibilité de corriger le diagnostic du MSFA 17 a posteriori via la CPA 5. Cette correction est transmise (liaison 8) au MSFA 17, qui révise ainsi ses diagnostics ultérieurs en prenant en compte la correction apportée par l'opérateur humain 23.

[0050] Ainsi, l'apprentissage du MSFA 17 est supervisé par l'opérateur humain 23 de sécurité qui apprend à ce dernier à classifier les alertes. De plus, cet apprentissage est progressif car au début, le MSFA 17 commet des erreurs de marquage et au fur et à mesure que l'opérateur humain 23 de sécurité confirme ou infirme les marquages, le diagnostic du MSFA 17 devient plus fiable. Finalement, lorsque le filtrage du MSFA 17 est suffisamment fiable, c'est à dire que son taux d'erreur de classification est tolérable, les alertes identifiées comme étant de fausses alertes (faux positifs) peuvent être soit directement supprimées, soit stockées dans le moyen de stockage 27 annexe de manière à ce que seules les vraies alertes (vrais positifs) soient présentées à l'opérateur humain 23. Le travail de l'opérateur humain 23 se trouve donc facilité car le volume d'alertes qui lui est présenté est très réduit.

[0051] Ainsi, le critère permettant de décider si une alerte est un vrai ou un faux positif est basé sur une comparaison des probabilités d'appartenance des alertes à l'une et à l'autre des deux catégories.

[0052] D'une manière générale, un « message d'alerte » a (ou plus simplement une alerte a) peut être défini comme un ensemble de n mots $m_{i\in\{1,\cdots.n\}}$, n étant un nombre entier qui peut varier d'une alerte à une autre : $a=(m_1,...,m_n)$.

[0053] Les mots d'une alerte désignent par exemple la nature d'une attaque contre le système d'informations 1, l'identité des victimes, l'identité présumée des attaquants, le type de faille exploitée, et la date.

[0054] Conformément à l'invention, étant donné une alerte a, le problème à résoudre Q consiste à déterminer si la probabilité que l'alerte a soit un faux positif est supérieure à la probabilité que l'alerte a soit un vrai positif. Si c'est le cas, alors l'alerte a est marquée comme « faux positif » sinon l'alerte a est inchangée (c'est-à-dire considérée comme

« vrai positif »).

[0055] On désigne par $P(vp|m_1,...,m_n)$ la probabilité qu'une alerte *a* contenant les mots $m_1,...,m_n$ soit un vrai positif *vp* et $P(fp|m_1,...,m_n)$ la probabilité qu'une alerte *a* contenant les mots $m_1,...,m_n$ soit un faux positif *fp*.

[0056] Le problème *Q* consiste donc à déterminer si

$$P(fp \mid m_1,...,m_n) \leq P(vp \mid m_1,...,m_n)$$

Cependant, d'après la définition des probabilités conditionnelles :

$$P(fp \mid m_1,...,m_n) = \frac{P(fp, m_1,...,m_n)}{P(m_1,...,m_n)} \text{ et } P(vp \mid m_1,...,m_n) = \frac{P(vp, m_1,...,m_n)}{P(m_1,...,m_n)}.$$

[0057] Par conséquent, le problème *Q* se réduit à déterminer si

$$P(fp, m_1,...,m_n) \leq P(vp, m_1,...,m_n) .$$

En outre, étant donné que $P(fp|m_1,...,m_n)=P(m_1,...,m_n|fp).P(fp)$ et $P(vp|m_1,...,m_n)=P(m_1,...,m_n|vp).P(vp)$ et en faisant l'hypothèse que les variables m_i sont conditionnellement indépendantes entre elles, il s'ensuit :

$$P(fp, m_1,...,m_n) = P(m_1 \mid fp).....P(m_n \mid fp).P(fp)$$

et

$$P(vp, m_1,...,m_n) = P(m_1 \mid vp).....P(m_n \mid vp).P(vp) .$$

[0058] Les valeurs $P(m_i|C)$ représentent la probabilité qu'un mot m_i soit présent dans une alerte qui appartient à la classe ou catégorie $C \in \{vp,fp\}$.

[0059] Au cours de l'apprentissage du MSFA 17, les moyens de traitement 21 construisent des compteurs H_C qui indiquent la fréquence des différents mots dans les deux catégories $C \in \{vp,fp\}$. En effet, le MSFA 17 construit une première table de hashage H_{fp} qui associe à chaque mot m_i la valeur $H_{fp}(m_i)$ qui désigne le nombre cumulé d'occurrences du mot m_i dans des alertes qui sont des faux positifs, ainsi qu'une seconde table de hashage H_{vp} qui associe à chaque mot m_i la valeur $H_{vp}(m_i)$ qui désigne le nombre cumulé d'occurrences du mot m_i dans des alertes qui sont des vrais positifs. Dans la suite, la notation $mots(H_C)$ désigne le domaine de définition de la table de hachage H_C, c'est-à-dire l'ensemble des mots correspondant à la catégorie $C \in \{vp,fp\}$.

[0060] Par conséquent, le nombre total d'occurrences de mots dans des vrais positifs est donné par la formule suivante :

$$N_{vp} = \sum_{m_i \in mots(H_{vp})} H_{vp}(m_i).$$

[0061] De même, le nombre total d'occurrences de mots dans des faux positifs est donné par la formule suivante :

$$N_{fp} = \sum_{m_i \in mots(H_{fp})} H_{fp}(m_i)$$

[0062] Par ailleurs, la probabilité qu'un mot m_i soit présent dans une alerte qui appartient à la classe $C \in \{vp,fp\}$ est donnée par la formule suivante :

$$P(m_i \mid C) = \frac{H_C(m_i)}{N_C}$$

En outre, la probabilité d'une classe C est donnée par la formule suivante :

$$P(C) = \frac{N_C}{N_{vp} + N_{fp}}$$

Par conséquent, les deux dernières formules permettent de calculer les probabilités $P(fp,m_1,...,m_n)$ et $P(vp,m_1,...,m_n)$ et ainsi de résoudre le problème Q ci-dessus.

[0063] La figure 2 est un organigramme très schématique illustrant les étapes du procédé de suppression de fausses alertes parmi les alertes produites dans un système de sécurité d'informations 3.

[0064] Les étapes E1 à E3 décrivent l'enregistrement dans les moyens mémoires 25 du MSFA 17 des diagnostics de l'opérateur humain 23 durant la phase d'apprentissage initial P1.

[0065] A l'étape E1, le MSFA 17 reçoit une alerte initiale donnée a'.

[0066] A l'étape E2, le MSFA 17 procède à l'extraction de l'ensemble de mots m'_i composant cette alerte initiale donnée : $a'=(m'_1,...,m'_n)$.

[0067] A l'étape E3, le MSFA 17 associe à chaque mot m'_i de l'ensemble de mots $\{m'_1,...,m'_n\}$, un compteur $H_C(m'_i)$ désignant le nombre cumulé d'occurrences du mot m'_i dans l'une des deux catégories $C \in \{vp,fp\}$.

[0068] L'étape E4 est un test destiné à vérifier si le nombre d'alertes ayant transité par le MSFA 17 a atteint un nombre suffisant. Ainsi, lorsque le nombre d'alertes est inférieur à un nombre seuil fixé par exemple par l'opérateur humain 23, on reboucle à l'étape E1.

[0069] En revanche, si le nombre d'alertes n'est pas inférieur au nombre seuil alors on passe aux étapes E5 à E14 décrivant la classification des alertes, par le MSFA 17 durant la phase de validation P2.

[0070] On notera que dans la phase d'apprentissage initiale P1, aucun marquage n'est effectué par le MSFA 17.

[0071] L'étape E5 indique la réception par le MSFA 17 d'une nouvelle alerte donnée notée a.

[0072] A l'étape E6, le MSFA 17 procède à l'extraction de l'ensemble de mots m_i composant cette nouvelle alerte donnée $a = (m_1,...,m_n)$.

[0073] A l'étape E7, les probabilités d'appartenance de la nouvelle alerte donnée a à l'une et à l'autre des catégories sont comparées :

$$P(fp \mid m_1,...,m_n) \leq P(vp \mid m_1,...,m_n) \ .$$

Cette comparaison peut comporter les sous étapes E71 à E74 suivantes :

[0074] A l'étape E71, le MSFA 17 calcule pour chaque mot m_i de l'ensemble de mots $\{m_1,...,m_n\}$ de la nouvelle alerte a, la probabilité que chaque mot m_i soit présent dans des alertes appartenant à l'une ou à l'autre des catégories C ($C \in \{vp,fp\}$) en déterminant le rapport entre le compteur $H_C(m_i)$ désignant le nombre cumulé d'occurrences de chaque mot m_i dans des alertes de l'une ou l'autre des catégories C et le nombre total N_C d'occurrences de mots dans l'une ou l'autre des catégories respectivement, c'est-à-dire $P(m_i \mid C) = \frac{H_C(m_i)}{N_C}$.

[0075] A l'étape E72, le MSFA 17 calcule la probabilité de chaque catégorie en déterminant le rapport entre le nombre total N_C d'occurrences de mots dans des alertes de chaque catégorie C et le nombre total de mots $N_{vp} + N_{fp}$, c'est-à-dire $P(C) = \frac{N_C}{N_{vp} + N_{fp}}$.

[0076] A l'étape E73, le MSFA 17 calcule le produit, sur l'ensemble des mots $\{m_1,...,m_n\}$ composant la nouvelle alerte donnée *a*, des probabilités $P(m_i|C)$ que chaque mot respectif de cette alerte soit présent dans des alertes appartenant à chaque catégorie multiplié par la probabilité de chaque catégorie $P(C)$, c'est-à-dire $\left(\prod_i P(m_i \mid C)\right).P(C)$.

[0077] A l'étape E74, le MSFA 17 compare le résultat de l'étape précédente selon les deux catégories, c'est-à-dire :

$$\left(\prod_i P(m_i \mid fp)\right).P(fp) \leq \left(\prod_i P(m_i \mid \mathrm{vp})\right).\mathrm{P}(\mathrm{vp}) \ .$$

[0078] A l'étape E8, la nouvelle alerte donnée *a* est classée par le MSFA 17 dans l'une des deux catégories selon le résultat de la comparaison de l'étape précédente E7.

[0079] A l'étape E9, le MSFA 17 incrémente les compteurs $H_C(m_i)$ selon la catégorie $C \in \{vp,fp\}$ de la nouvelle alerte donnée.

[0080] Ensuite, à l'étape E10, le MSFA 17 transmet la nouvelle alerte donnée *a* ainsi classée (marquée) à la console de présentation des alertes CPA 5.

[0081] Eventuellement, l'opérateur humain 23 interagit avec le MSFA 17 via la CPA 5 pour corriger un diagnostic erroné effectué par le MSFA 17.

[0082] En effet, à l'étape E11, si le MSFA 17 reçoit une notification de l'opérateur humain 23 indiquant que le précédent classement *C* de la nouvelle alerte *a* est faux, alors le MSFA 17 procède à la correction du diagnostic selon les étapes E12 à E14, sinon on passe directement à l'étape E15.

[0083] A l'étape E12, le MSFA 17 corrige la catégorie de la nouvelle alerte *a* selon la notification de l'opérateur humain 23. Autrement dit, le MSFA 17 marque la nouvelle alerte *a* par un classement $\overline{C}$ contraire au classement antérieur *C*.

[0084] A l'étape E13, le MSFA 17 décrémente les compteurs $H_C(m_i)$ désignant les nombres cumulés d'occurrences de mots dans la catégorie *C* faussement classée.

[0085] A l'étape E14, le MSFA 17 incrémente les compteurs $H_{\overline{C}}(m_i)$ désignant les nombres cumulés d'occurrences de mots dans la catégorie corrigée $\overline{C}$.

[0086] L'étape E15 est un test destiné à vérifier si le taux d'erreur de classification est tolérable.

[0087] En effet, tant que le taux de correction de la classification des nouvelles alertes de la phase de validation P2 n'est pas inférieur à un nombre seuil déterminé, on reboucle à l'étape E5.

[0088] Sinon, on passe aux étapes E16 à E22 décrivant la classification des alertes, par le MSFA 17 durant la phase opérationnelle P3. Les étapes E16 à E21 sont similaires aux étapes E5 à E10 de la phase de validation P2.

[0089] En effet, à la réception d'une autre nouvelle alerte *a* à l'étape E16, le MSFA 17 procède à extraire à l'étape E17, l'ensemble de mots m_i composant cette nouvelle alerte $a=(m_1,...,m_n)$. L'étape E18, est une comparaison des probabilités d'appartenance de cette nouvelle alerte à l'une et à l'autre des catégories. L'étape E19 est le classement de la nouvelle alerte. L'étape E20, consiste à incrémenter les compteurs selon la catégorie de classement de la nouvelle alerte. A l'étape E21, le MSFA 17 transmet la nouvelle alerte ainsi classée à la CPA 5.

[0090] Finalement, à l'étape E22 les fausses alertes sont stockées dans le moyen de stockage 27. En variante à l'étape E22 les fausses alertes peuvent être supprimées.

[0091] Ainsi, le MSFA 17 selon l'invention, évalue la probabilité qu'une alerte soit un faux positif en fonction des mots qui la composent. Le MSFA 17 marque les alertes qu'il juge être des faux positifs et transmet l'alerte munie de son marquage à l'opérateur humain 23 de sécurité. Ce dernier a la possibilité de modifier le diagnostic effectué par le MSFA 17 si celui-ci est erroné via la console de présentation des alertes CPA 5. Dans ce dernier cas, la modification est prise en compte par le MSFA 17 pour réviser ses diagnostics ultérieurs.

[0092] De cette manière, la fiabilité du MSFA 17 dans le traitement des alertes va croissant à mesure que l'opérateur humain 23 corrige ses diagnostics.

Revendications

1. Procédé de suppression de fausses alertes parmi les alertes produites dans un système d'informations surveillé (1), **caractérisé en ce que** les alertes sont classées automatiquement au moyen d'un module de suppression de fausses alertes (17) suivant deux catégories constituées de fausses et vraies alertes selon des critères déterminés basés sur un apprentissage progressif dudit module (17) à partir de l'expertise d'un opérateur humain (23) en charge d'un classement initial manuel des alertes, ledit apprentissage progressif comportant les phases suivantes :

- une phase d'apprentissage initial (P1) dans laquelle ledit module de suppression de fausses alertes (17) procède à un enregistrement de diagnostics de l'opérateur humain (23) concernant un nombre déterminé d'alertes initiales et comprenant pour une alerte initiale donnée, une extraction de l'ensemble de mots composant ladite alerte initiale donnée, et une association à chaque mot dudit ensemble de mots, d'un compteur désignant le nombre cumulé d'occurrences dudit mot dans l'une des deux catégories, et
- une phase de validation (P2) dans laquelle ledit module de suppression de fausses alertes (17) procède à la classification de nouvelles alertes en fonction dudit enregistrement de diagnostics et d'une supervision de l'opérateur humain (23) qui confirme ou corrige les classifications de nouvelles alertes.

2. Procédé selon la revendication 1, **caractérisé en ce que** lesdits critères déterminés comportent une comparaison des probabilités d'appartenance des alertes à l'une et à l'autre desdites deux catégories.

3. Procédé selon la revendication 1, **caractérisé en ce que** dans la phase de validation (P2), lesdites confirmations ou corrections des classifications de nouvelles alertes apportées par l'opérateur humain (23) sont utilisées par le module de suppression de fausses alertes (17) pour minimiser un taux de correction lui permettant d'augmenter la fiabilité de toute classification ultérieure de nouvelles alertes.

4. Procédé selon la revendication 3, **caractérisé en ce qu'**il comporte une phase opérationnelle (P3) dans laquelle la classification des nouvelles alertes est effectuée de manière autonome, si le taux de correction de la classification des nouvelles alertes de la phase de validation (P2) devient inférieur à un nombre seuil déterminé.

5. Procédé selon la revendication 4, **caractérisé en ce que** dans la phase opérationnelle (P3), les fausses alertes sont supprimées ou stockées dans un moyen de stockage (27) et seules les vraies alertes sont envoyées à une console de présentation des alertes (5).

6. Procédé selon l'une quelconque des revendications 1 à 5, **caractérisé en ce que** la classification des alertes durant la phase de validation (P2) et la phase opérationnelle (P3) comporte, pour une nouvelle alerte donnée, les étapes suivantes :

- extraction de l'ensemble de mots composant ladite nouvelle alerte donnée,
- comparaison des probabilités d'appartenance de la nouvelle alerte donnée à l'une et à l'autre desdites catégories,
- classement de la nouvelle alerte donnée dans l'une des deux catégories selon le résultat de la comparaison de l'étape précédente,
- incrémentation des compteurs selon la catégorie de la nouvelle alerte donnée, et
- transmission de la nouvelle alerte donnée ainsi classée à la console de présentation des alertes (5).

7. Procédé selon la revendication 6, **caractérisé en ce que** la comparaison des probabilités d'appartenance de la nouvelle alerte donnée à l'une et à l'autre desdites catégories comporte les étapes suivantes :

- calculer pour chaque mot de l'ensemble de mots de ladite nouvelle alerte, la probabilité que chaque mot soit présent dans des alertes appartenant à l'une ou à l'autre des catégories en déterminant le rapport entre le compteur désignant le nombre cumulé d'occurrences de chaque mot dans des alertes de l'une ou l'autre des catégories et le nombre total d'occurrences de mots dans l'une ou l'autre des catégories respectivement,
- calculer la probabilité de chaque catégorie en déterminant le rapport entre le nombre total d'occurrences de mots dans des alertes de chaque catégorie et le nombre total de mots,
- calculer le produit, sur l'ensemble des mots composant l'alerte, des probabilités que chaque mot respectif de l'alerte soit présent dans des alertes appartenant à chaque catégorie multiplié par la probabilité de chaque catégorie, et
- comparer le résultat de l'étape précédente selon les deux catégories.

8. Procédé selon l'une quelconque des revendications 1 à 7, **caractérisé en ce que** la correction par ledit module de suppression de fausses alertes (17) de la classification des nouvelles alertes durant la phase de validation (P2) comporte les étapes suivantes :

- correction de la catégorie d'une nouvelle alerte précédemment classée par ledit module (17) s'il reçoit une notification de l'opérateur humain (23) indiquant que ledit précédent classement de ladite nouvelle alerte est faux,
- décrémentation des compteurs désignant les nombres cumulés d'occurrences de mots dans la catégorie

faussement classée,
- incrémentation des compteurs désignant les nombres cumulés d'occurrences de mots dans la catégorie corrigée.

9. Module de suppression de fausses alertes, **caractérisé en ce qu'**il comporte :

- des moyens de traitement (21) de données permettant de classer automatiquement les alertes suivant deux catégories constituées de fausses et vraies alertes selon des critères déterminés basés sur un apprentissage progressif à partir de l'expertise d'un opérateur humain (23) en charge d'un classement initial manuel des alertes, et
- des moyens mémoires (25) permettant d'enregistrer, lors d'une phase d'apprentissage initial de l'apprentissage progressif, des diagnostics de l'opérateur humain (23) concernant un nombre déterminé d'alertes initiales en permettant pour une alerte initiale donnée d'extraire l'ensemble de mots composant ladite alerte initiale donnée et en associant à chaque mot dudit ensemble de mots, un compteur désignant le nombre cumulé d'occurrences dudit mot dans l'une des deux catégories,

les moyens de traitement (21) de données permettant en outre de procéder à la classification de nouvelles alertes en fonction dudit enregistrement de diagnostics et d'une supervision de l'opérateur humain (23) qui confirme ou corrige les classifications de nouvelles alertes.

10. Module selon la revendication 9, **caractérisé en ce que** lors d'une phase opérationnelle (P3), les moyens de traitement (21) de données sont destinés en outre à classifier de manière autonome de nouvelles alertes si un taux de correction de la classification des nouvelles alertes de la phase de validation (P2) devient inférieur à un nombre seuil déterminé.

11. Module selon la revendication 10, **caractérisé en ce qu'**il comporte en outre un moyen de stockage (27) permettant de stocker, lors de la phase opérationnelle, les fausses alertes de sorte que seules les vraies alertes sont envoyées à une console de présentation des alertes (5).

12. Système d'informations surveillé comportant un réseau interne à surveiller (7), des sondes de détection d'intrusions (11a, 11b, 11c), un système de gestion d'alertes (15), et une console de présentation d'alertes (5) **caractérisé en ce qu'**il comporte en outre un module de suppression de fausses alertes (17) selon l'une quelconque des revendications 9 à 11.

Claims

1. Method for suppressing false alerts among the alerts produced in a monitored information system (1), **Characterized in that** the alerts are categorized automatically, by means of a module (17) for suppressing false alerts, into two categories formed of false and true alerts according to determined criteria based on progressive training of the said module (17) through the expertise of a human operator (23) responsible for an initial manual categorization of alerts, the said progressive training including the following phases:

- an initial training phase (P1) in which the said module (17) for suppressing false alerts records diagnostics of the human operator (23) relating to a determined number of initial alerts and comprising, for a given initial alert, an extraction of the set of words forming the said given initial alert, and an association with each word of the said set of words, of a counter indicating the cumulative number of occurrences of the said word in one of the two categories, and
- a validation phase (P2) in which the said module (17) for suppressing false alerts categorizes new alerts according to the said diagnostics recording and according to a supervision action by the human operator (23) who confirms or corrects the categorizations of new alerts.

2. Method according to Claim 1, **characterized in that** the said determined criteria include a comparison of probabilities that the alerts belong to one and to the other of the said two categories.

3. Method according to Claim 1, **characterized in that**, in the validation phase (P2), the said confirmations or corrections of categorizations of new alerts introduced by the human operator (23) are used by the module (17) for suppressing false alerts to minimize a correction rate, enabling it to improve the reliability of any further categorization of new alerts.

4. Method according to Claim 3, **characterized in that** it includes an operational phase (P3) in which new alerts are categorized autonomously if the correction rate of the categorization of new alerts in the validation phase (P2) becomes less than a determined threshold number.

5. Method according to Claim 4, **characterized in that**, in the operational phase (P3), the false alerts are suppressed or stored in a storage means (27) and only the true alerts are sent to an alert presentation console (5).

6. Method according to any one of Claims 1 to 5, **characterized in that** the categorization of alerts during the validation phase (P2) and the operational phase (P3) includes, for a given new alert, the following steps:

- extracting the set of words forming the said given new alert,
- comparing probabilities that the given new alert belongs to one and to the other of the said categories,
- categorizing the given new alert in one of the two categories according to the result of the comparison in the previous step,
- incrementing counters according to the category of the given new alert, and
- transmitting the given new alert thus categorized to the alert presentation console (5).

7. Method according to Claim 6, **characterized in that** the comparison of probabilities that the given new alert belongs to one and to the other of the said categories includes the following steps:

- calculating for each word of the set of words of the said new alert the probability that each word is present in alerts belonging to one or to the other of the categories by determining the ratio between the counter indicating the cumulative number of occurrences of each word in alerts of one or the other of the categories and the total number of occurrences of words in one or the other of the categories respectively,
- calculating the probability of each category by determining the ratio between the total number of occurrences of words in alerts of each category and the total number of words,
- calculating the product, on the set of words forming the alert, of the probabilities that each respective word of the alert is present in alerts belonging to each category multiplied by the probability of each category, and
- comparing the result of the previous step according to the two categories.

8. Method according to any one of Claims 1 to 7, **characterized in that** the correction, by the said module (17) for suppressing false alerts, of the categorization of new alerts during the validation phase (P2) includes the following steps:

- correcting the category of a new alert previously categorized by the said module (17) if it receives a notification from the human operator (23) indicating that the said previous categorization of the said new alert is false,
- decrementing counters indicating the cumulative numbers of occurrences of words in the falsely categorized category,
- incrementing counters indicating the cumulative numbers of occurrences of words in the corrected category.

9. Module for suppressing false alerts, **characterized in that** it includes:

- data processing means (21) for automatically categorizing the alerts into two categories formed of false and true alerts according to determined criteria based on progressive training through the expertise of a human operator (23) responsible for an initial manual categorization of alerts, and
- memory means (25) for recording, during an initial training phase of the progressive training, diagnostics of the human operator (23) relating to a determined number of initial alerts, enabling for a given initial alert the extraction of the set of words forming the said given initial alert and associating with each word of the said set of words a counter indicating the cumulative number of occurrences of the said word in one of the two categories,

the data processing means (21) additionally providing for categorizing new alerts according to the said diagnostics recording and according to a supervision action by the human operator (23) who confirms or corrects the categorizations of new alerts.

10. Module according to Claim 9, **characterized in that**, during an operational phase (P3), the data processing means (21) are additionally intended to autonomously categorize new alerts if a correction rate of the categorization of new alerts in the validation phase (P2) becomes less than a determined threshold number.

11. Module according to Claim 10, **characterized in that** it additionally includes a storage means (27) for storing, during the operational phase, the false alerts such that only the true alerts are sent to an alert presentation console (5).

12. Monitored information system including an internal network to be monitored (7), intrusion detection probes (11a, 11b, 11c), an alert management system (15) and an alert presentation console (5), **characterized in that** it additionally includes a module (17) for suppressing false alerts according to any one of Claims 9 to 11.

Patentansprüche

1. Verfahren zum weglassen falscher Warnsignale bei den Warnsignalen, die in einem überwachten Informationssystem (1) erzeugt werden, **dadurch gekennzeichnet, dass** die Warnsignale durch ein Modul zum weglassen falscher Warnsignale (17) automatisch in zwei Kategorien eingeteilt werden, die gemäß von Kriterien, die aufgrund von schrittweisem Lernen des Moduls (17) anhand der Erfahrung eines menschlichen Bedieners (23), der mit einer anfänglichen manuellen Einteilung der Warnsignale betraut ist, festgelegt sind, aus falschen und richtigen Warnsignalen gebildet sind, wobei das schrittweise Lernen folgende Phasen aufweist:

- eine Phase des anfänglichen Lernens (P1), in der das Modul zum weglassen falscher Warnsignale (17) eine Eintragung von Diagnosen des menschlichen Bedieners (23) betreffend eine bestimmte Anzahl von anfänglichen Warnsignalen vornimmt, und die bei einem gegebenen anfänglichen Warnsignal eine Entfernung sämtlicher Worte aufweist, die das gegebene anfängliche Warnsignal bilden, sowie eine Verknüpfung eines Zählers, der die kumulierte Anzahl des Vorkommens des Wortes in einer der beiden Kategorien angibt, mit jedem Wort der Gesamtheit der Wörter, und
- eine Phase der Validierung (P2), in der das Modul zum weglassen falscher Warnsignale (17) die Einteilung neuer Warnsignale in Abhängigkeit von der Eintragung von Diagnosen und von einer Überwachung des menschlichen Bedieners (23), der die Einteilung neuer Warnsignale bestätigt oder korrigiert, vornimmt.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** die festgelegten Kriterien einen Vergleich der Wahrscheinlichkeit der Zugehörigkeit der Warnsignale zu einer und zur anderen der beiden Kategorien aufweisen.

3. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** die Bestätigungen oder Korrekturen der Einteilungen neuer Warnsignale, die durch den menschlichen Bediener (23) erfolgen, in der Phase der Validierung (P2) von dem Modul zum weglassen falscher Warnsignal (17) verwendet werden, um eine Korrekturrate auf ein Mindestmaß zu senken, was ihm ermöglicht, die Zuverlässigkeit jeglicher späteren Einteilung neuer Warnsignale zu steigern.

4. Verfahren nach Anspruch 3, **dadurch gekennzeichnet, dass** es eine operationelle Phase (P3) aufweist, in der die Einteilung der neuen Warnsignale eigenständig erfolgt, falls die Korrekturrate der Einteilung der neuen Warnsignale der Validierungsphase (P2) unter einen bestimmten Schwellenwert sinkt.

5. Verfahren nach Anspruch 4, **dadurch gekennzeichnet, dass** in der operationellen Phase (P3) die falschen Warnsignale weggelassen oder in einem Speichermittel (27) gespeichert werden und nur die echten Warnsignale zu einem Bedienungsplatz zur Darstellung der Warnsignale (5) gesendet werden.

6. Verfahren nach einem der Ansprüche 1 bis 5, **dadurch gekennzeichnet, dass** die Einteilung der Warnsignale während der Validierungsphase (P2) und der operationellen Phase (P3) bei einem neuen gegebenen Warnsignal folgende Schritte aufweist:

- Entfernung sämtlicher Wörter, die das neue gegebene Warnsignal bilden,
- Vergleich der Wahrscheinlichkeit der Zugehörigkeit des neuen gegebenen Warnsignals zu der einen und der anderen der Kategorien,
- Einteilung des neuen gegebenen Warnsignals in eine der beiden Kategorien gemäß dem Ergebnis des Vergleichs aus dem vorhergehenden Schritt,
- Inkrementierung der Zähler gemäß der Kategorie des neuen gegebenen Warnsignals, und
- Übertragung des auf diese Weise eingeteilten neuen gegebenen Warnsignals zu dem Bedienplatz zur Darstellung der Warnsignale (5).

7. Verfahren nach Anspruch 6, **dadurch gekennzeichnet, dass** der Vergleich der Wahrscheinlichkeit der Zugehörigkeit des neuen gegebenen Warnsignals zu der einen und der anderen der Kategorien folgende Schritte aufweist:

- Berechnen, für jedes Wort der Gesamtheit der Wörter des neuen Warnsignals, der Wahrscheinlichkeit, dass jedes Wort in Warnsignalen vorhanden ist, die zu der einen oder der anderen der Kategorien gehören, durch das Bestimmen des Verhältnisses zwischen dem Zähler, der die kumulierte Anzahl des Vorkommens jedes Worts in Warnsignalen der einen oder der anderen der Kategorien angibt, und der Gesamtzahl des Vorkommen von Wörtern in der einen beziehungsweise der anderen der Kategorien.
- Berechnen der Wahrscheinlichkeit jeder Kategorie durch Bestimmen des Verhältnisses zwischen der Gesamtzahl des Auftretens von Wörtern in Warnsignalen jeder Kategorie und der Gesamtzahl der Wörter,
- Berechnen des Produkts, aus der Gesamtheit der Wörter, die das Warnsignal bilden, der Wahrscheinlichkeit, dass jedes Wort des Warnsignals in Warnsignalen vorhanden ist, die zu jeder Kategorie gehören, multipliziert mit der Wahrscheinlichkeit jeder Kategorie, und
- Vergleichen des Ergebnisses des vorhergehenden Schritts gemäß den beiden Kategorien.

8. Verfahren nach einem der Ansprüche 1 bis 7, **dadurch gekennzeichnet, dass** die Korrektur der Einteilung der neuen Warnsignale durch das Modul zum weglassen falscher Warnsignale (17) während der Validierungsphase (P2) folgende Schritte aufweist:

- Korrektur der Kategorie eines neuen, vorher eingeteilten Warnsignals durch das Modul (17), falls es eine Mitteilung vom menschlichen Bediener (23) erhält mit der Angabe, dass die vorherige Einteilung des neuen Warnsignals falsch ist,
- Dekrementierung der Zähler, die die kumulierte Anzahl des Auftretens von Wörtern in der falsch zugeordneten Kategorie bestimmen,
- Inkrementierung der Zähler, die die kumulierte Anzahl des Auftretens von Wörtern in der korrigierten Kategorie bestimmen,

9. Modul zum weglassen falscher Warnsignale, **dadurch gekennzeichnet, dass** es Folgendes aufweist:

- Mittel zur Verarbeitung (21) von Daten, die es ermöglichen, die Warnsignale automatisch in zwei Kategorien einzuteilen, die gemäß Kriterien, die bestimmt werden aufgrund eines schrittweisen Lernens anhand der Erfahrung eines menschlichen Bedieners (23), der mit einer anfänglichen manuellen Einteilung der Warnsignale betraut ist, aus falschen und richtigen Warnsignalen bestehen, und
- Speichermittel (25), die es ermöglichen, in einer Phase des anfänglichen Lernens des schrittweisen Lernens Diagnosen des menschlichen Bedieners (23) betreffend eine bestimmte Anzahl von anfänglichen Warnsignalen aufzuzeichnen, indem es bei einem gegebenen anfänglichen Warnsignal ermöglicht wird, sämtliche Wörter, die das gegebene anfängliche Warnsignal bilden, zu entnehmen und indem mit jedem Wort der Gesamtheit der Wörter ein Zähler verknüpft wird, der die kumulierte Anzahl des Vorkommens des Wortes in einer der beiden Kategorien bestimmt,
- Mittel zur Verarbeitung (21) von Daten, mit denen es ferner möglich ist, die Einteilung neuer Warnsignale in Abhängigkeit von der Eintragung von Diagnosen und von einer Überwachung des menschlichen Bedieners (23), der die Einteilungen neuer Warnsignale bestätigt oder korrigiert, durchzuführen.

10. Modul nach Anspruch 9, **dadurch gekennzeichnet, dass** die Mittel zur Verarbeitung (21) von Daten in einer operationellen Phase (P3) ferner dazu bestimmt sind, eigenständig neue Warnsignale einzuteilen, falls eine Korrekturrate der Einteilung der neuen Warnsignale der Validierungsphase (P2) unter einen bestimmten Schwellenwert sinkt.

11. Modul nach Anspruch 10, **dadurch gekennzeichnet, dass** es ferner ein Speichermittel (27) aufweist, das es ermöglicht, während der operationellen Phase die falschen Warnsignale zu speichern, so dass nur die richtigen Warnsignale zu einem Bedienplatz zur Darstellung der Warnsignale (5) gesendet werden.

12. Überwachtes Informationssystem, aufweisend ein internes Netz zum Überwachen (7), Sonden zum Erkennen von Eindringungen (11a, 11b, 11c), ein System zur Verwaltung von Warnsignalen (15) und ein Bedienplatz zur Darstellung von Warnsignalen (5), **dadurch gekennzeichnet, dass** es ferner ein Modul zum weglassen falscher Warnsignale (17) gemäß einem der Ansprüche 9 bis 11 aufweist.

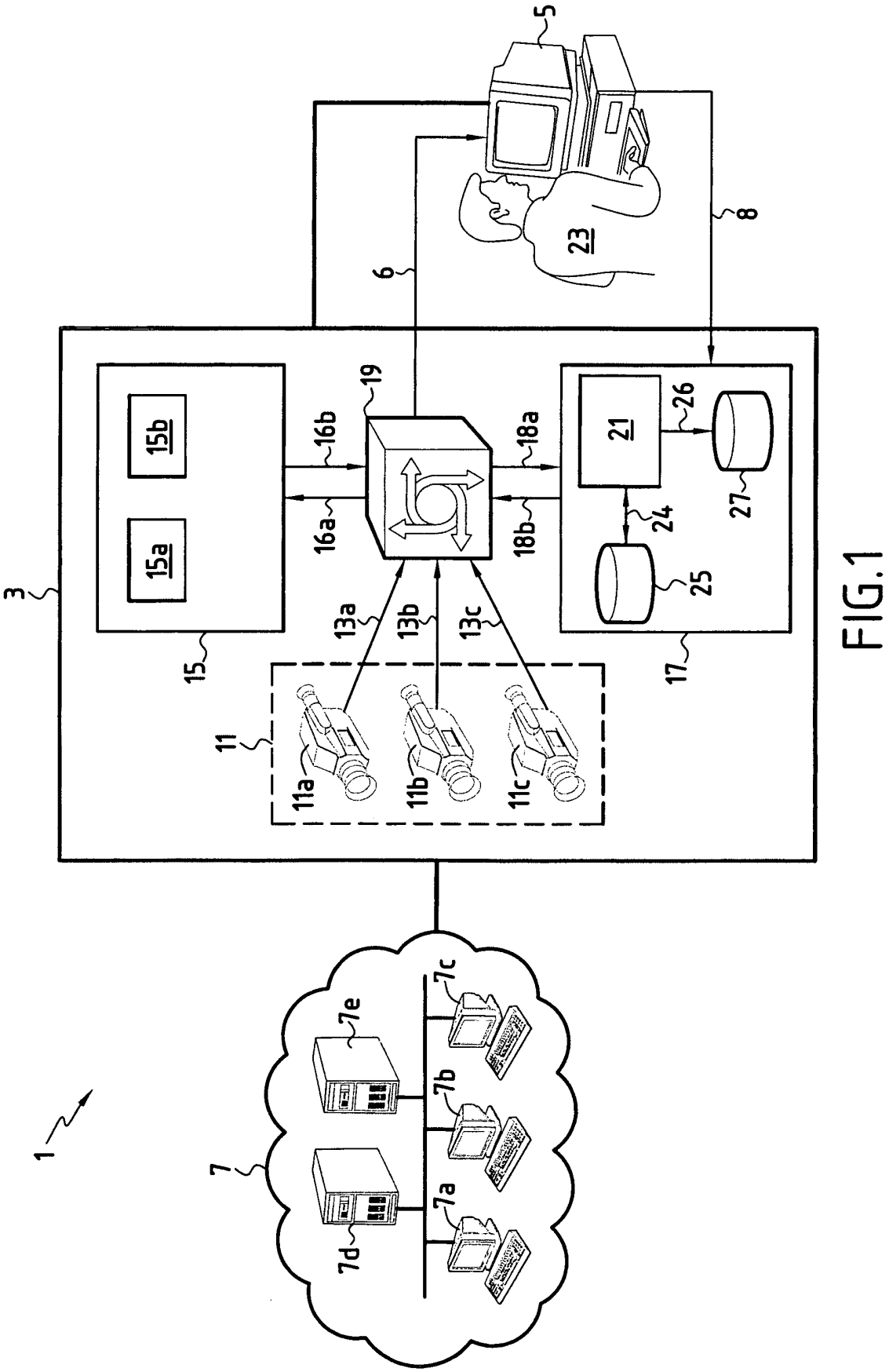
1
3
5
6
7
8
11
11a
11b
11c
13a
13b
13c
15
15a
15b
16a
16b
17
18a
18b
19
21
23
24
25
26
27
7a
7b
7c
7d
7e

FIG.1

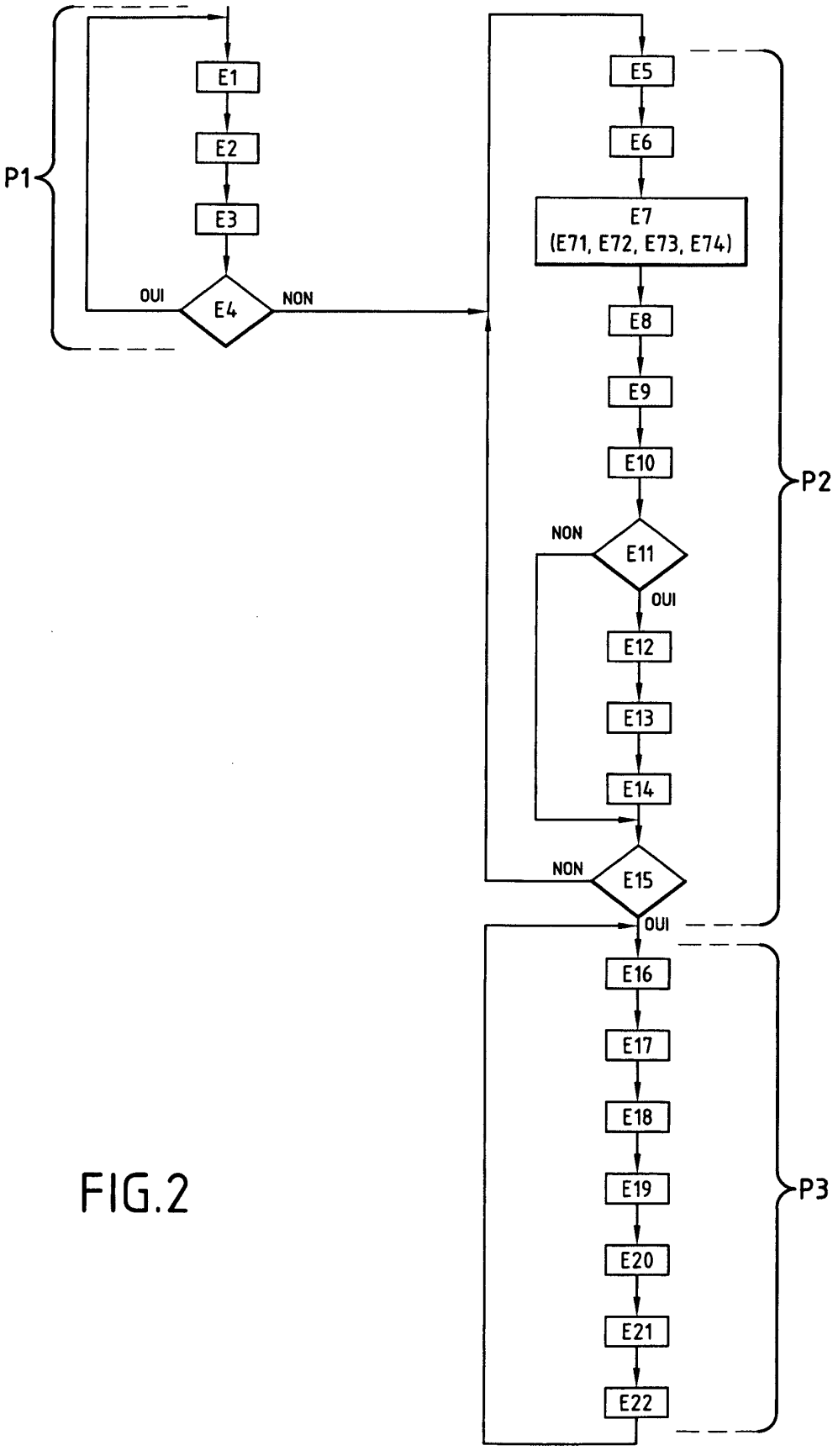
P1
E1
E2
E3
E4
OUI
NON
E5
E6
E7
(E71, E72, E73, E74)
E8
E9
E10
E11
NON
OUI
E12
E13
E14
E15
NON
OUI
P2
E16
E17
E18
E19
E20
E21
E22
P3

FIG.2