BESCHWERDEKAMMERN
DES EUROPÄISCHEN
PATENTAMTS

BOARDS OF APPEAL OF
THE EUROPEAN PATENT
OFFICE

CHAMBRES DE RECOURS
DE L'OFFICE EUROPÉEN
DES BREVETS

Internal distribution code:

(A) [-] Publication in OJ
(B) [-] To Chairmen and Members
(C) [-] To Chairmen
(D) [X] No distribution

Datasheet for the decision of 5 November 2019

Case Number:	T 0340/14 - 3.5.06
Application Number:	10176516.2
Publication Number:	2306357
IPC:	G06F21/00
Language of the proceedings:	EN

Title of invention:
Method and system for detection of previously unknown malware

Applicant:
Kaspersky Lab, ZAO

Headword:
Detection of previously unknown malware/KASPERSKY

Relevant legal provisions:
EPC Art. 56, 123(2)

Keyword:
Inventive step - (no)
Amendments - added subject-matter (yes)

Decisions cited:

This datasheet is not part of the Decision.
It can be changed at any time and without notice.

Catchword:

Beschwerdekammern

Boards of Appeal

Chambres de recours

Boards of Appeal of the
European Patent Office
Richard-Reitzner-Allee 8
85540 Haar
GERMANY
Tel. +49 (0)89 2399-0
Fax +49 (0)89 2399-4465

Case Number: T 0340/14 - 3.5.06

DECISION
# of Technical Board of Appeal 3.5.06
of 5 November 2019

Appellant: (Applicant)	Kaspersky Lab, ZAO 39A/3 Leningradskoe Shosse Moscow 125212 (RU)
Representative:	Sloboshanin, Sergej V. Füner, Ebbinghaus, Finck, Hano Mariahilfplatz 3 81541 München (DE)
Decision under appeal:	**Decision of the Examining Division of the European Patent Office posted on 21 October 2013 refusing European patent application No. 10176516.2 pursuant to Article 97(2) EPC.**

Composition of the Board:

Chairman	M. Müller
Members:	A. Teale
	B. Müller

Summary of Facts and Submissions

I. The appeal is against the decision, dispatched with reasons on 21 October 2013, to refuse European patent application No. 10 176 516.2 due to lack of inventive step, Article 56 EPC, in view of the following documents:

D2: US 2008/0027891 A1 and
D4: US 7 392 544 B1.

II. A notice of appeal and the appeal fee were received on 18 December 2013. The appellant requested that the decision be set aside and made an auxiliary request for oral proceedings.

III. With a statement of grounds of appeal, received on 23 January 2014, the appellant submitted claims according to a main and an auxiliary request.

IV. In an annex to a summons to oral proceedings the board set out its preliminary opinion that *inter alia* claim 1 of the main request seemed to lack inventive step, Article 56 EPC, in view of D2 and D4. Claim 4 of the main request and the feature added to *inter alia* claim 1 of the auxiliary request with respect to that of the main request appeared to contain added subject-matter, Article 123(2) EPC. The added feature also made these claims unclear, Article 84 EPC, and appeared unable to lend inventive step, Article 56 EPC, to the claims.

V. With a response received on 7 October 2019 the appellant submitted amended pages of the description and sets of claims according to a main and first and second auxiliary requests. The appellant requested that the decision be set aside and that a patent be granted on the basis of said new requests.

VI. At the oral proceedings, held on 5 November 2019, the appellant requested that the decision under appeal be set aside and that a patent be granted in the following version:

Claims:
- No. 1 to 9 according to the main request, or
- No. 1 to 8 according to the first auxiliary request or
- No. 1 to 7 according to the second auxiliary request;

all claim sets filed with the letter of 7 October 2019.

Description (for all requests):
- Pages 1 to 4 filed with the letter of 7 October 2019,
- page 18 filed with a letter of 24 May 2011 and
- pages 5 to 17 as originally filed.

Drawings (for all requests):
Sheets 1/7 to 7/7 with figures 1 to 7 as originally filed.

VII. At the end of the oral proceedings the board announced its decision.

VIII. Claim 1 according to the main request reads as follows:

"A computer-implemented method for detection of unknown malware, the method comprising:
(a) receiving metadata information about files executed on a remote computer and information about events of execution of the files, including information about actions performed on the files;

(b) filtering out known metadata information and known information about events on [the] basis of information stored in WhiteList and BlackList knowledge databases; characterized by
(c) determining an invocation sequence of the files based on the information about events of file execution, including at least events of file download, dropping, linking and invocation;
(d) constructing a parent-child hierarchy based on the invocation sequence of the files, wherein the parent-child hierarchy describes parent-child relationships between the files;
(e) performing a risk analysis of the files based on the received, not filtered information and based at least in part on one of the following criteria: statistics of the events of execution of each file, name stability of the file source, address stability of the file source, activity of the file and a behavioral pattern of the file;
(f) performing a risk assessment of the files based on the parent-child hierarchy to determine a level of danger of the related files, wherein a risk assessed to a parent file is based at least in part on the risk associated with one or more children files of the parent file; and
(g) determining whether the files are malicious based on the risk analysis and risk assessment, wherein the risk analysis of a file is based on calculating a level of activity of the file (["]Activity" parameter), a degree of danger of the file ("Danger" parameter), and a significance for the file ("Significance" parameter), wherein the Significance parameter is computed as Significance = (Activity * Danger), and wherein the Danger parameter is calculated based on a decision tree of weight coefficients built up for the criteria used in the risk analysis of the file."

IX. Claim 1 of the first auxiliary request differs from that of the main request in the addition of the following expression at the end:

", wherein a majority of the weight coefficients is dynamically changing depending on the accumulated information in the knowledge databases".

X. Claim 1 of the second auxiliary request differs from that of the main request in the addition of the following expression at the end:

", wherein the decision tree of weight is expandable by adding new criteria, wherein a maximum weight of a criterion associated with a higher level tree node is not changed by weights of added criteria at lower tree node levels".

Reasons for the Decision

1. The admissibility of the appeal

The appeal fulfills the admissibility requirements under the EPC.

2. A summary of the invention

2.1 The invention addresses the problem of detecting previously unknown computer threats, such as malicious programs, before they enter a computer system; see page 2, lines 11 to 13, and page 4, lines 16 to 26. To do this, an antivirus client program running on the user's computer sends event information and metadata relating to a file to the "Kaspersky Security Network" (KSN) server system. The event information can relate to actions such as file downloading or file dropping and event statistics; see page 3, lines 3 to 6, and page 5, lines 11 to 15. The file metadata can be a file name or the URL from which the file was downloaded; see paragraph bridging pages 2 and 3 and page 5, lines 15 to 20.

2.2 The server system filters (figure 1; step 101) the information from the user's computer to identify known malware using a "Blacklist" (BL) (101). Similarly software known not to be malware is identified using a "Whitelist" (WL) (101); see page 3, lines 10 to 11. What remains is treated as potential unknown malware, and the system subjects it to a risk analysis and risk assessment (102) to calculate a "danger factor", the danger factor of a particular object being calculated as the aggregate value for that object and related objects in a "parent-child" hierarchical "Downward-Starter" "DS-chart" (also referred to as a "DS-graph"; see page 6, line 15, and figure 1) based on the invocation of files; see figure 2. The KSN updates both the blacklist and the whitelist based on its findings (105A, 105B); see page 5, lines 25 to 30.

2.3 The description gives four examples of calculating the "danger" parameter using a "decision tree" and aggregating the results; see page 8, six lines from the bottom, to page 14, line 11, and figures 3 to 5.

2.4 The "danger" parameter of a particular object is calculated (see figure 1; step 102) by building a decision tree depending on the circumstances (see figures 3 to 5), the decision tree criteria having weight coefficients, the majority of which dynamically change depending on the accumulated information in the knowledge base; see page 8, five lines beneath table 1. For instance, "Example 1" (see page 8) gives a list of criteria, illustrated by the tree of weight coefficients in figure 3, that the system checks, for instance the host name, when an executable file is downloaded from the Internet. If, for example, the host name (**soho.com.server911.ch) includes the name (soho.com) of a host on the whitelist then the weight coefficient for the "masking" criterion is set to 100; see paragraph bridging pages 8 and 9. Figure 5 illustrates new criteria being added to an existing decision tree, the new criteria relating to "Driver installation" and "Direct write to disk"; see page 12, "Example 4".

2.5 Using the DS-chart, the system calculates a danger factor for each parent node in the DS-chart; see figure 1; step 103. The danger posed by a parent is assessed based on the danger posed by its children; see page 7, lines 14 to 18.

2.6 Using the danger factor, the system then calculates the "significance" parameter according to the formula "significance = danger x activity", the activity parameter being based on the number of downloads of an object or the number of times a given object is invoked over a certain period of time; see page 7, line 24, to page 8, line 5.

3. Clarity, Article 84 EPC

3.1 In a section of the decision entitled "Obiter dictum" the examining division stated that claims 4 to 8 were unclear, Article 84 EPC, since the meanings of the parameters "Activity", "Danger" and "Significance" were vague. In particular, the definition of "Danger" in claims 4 and 8 was unclear and did not allow a skilled person to identify which technical features were necessary for calculating the degree of danger of a file. It was also not clear how a skilled person should use and/or combine said parameters when performing the risk analysis of a file.

3.2 The appellant has argued that the term "activity" is clear in view of the passages in the description (last paragraph on page 7 and first on page 8), explaining that activity is determined by a number of downloads or the number of times a given file is invoked over a period of time. The term "Danger" was clear in view of figure 4 and the examples described on pages 8 to 14 and referred to a calculation based on "a decision tree of weight coefficients built up for the criteria used in the risk analysis of the file", the criteria being stated in claim 1. The term "significance" was explained on page 15, fourth paragraph.

3.3 In the light of the cited explanations in the description and drawings, the board finds that the terms objected to by the examining division are sufficiently clear for the assessment of inventive step. Whether all requirements of Article 84 EPC are fulfilled need not be decided.

4. The prior art on file

4.1 Document D2

4.1.1 D2 relates to detecting previously unknown malicious entities (see [6, 8]), such as files, in a computer system. Figure 7 shows a client-server structure in which client processing systems (710) identify a starting entity or group of related entities and send data on these entities to a central server processing system (720) for analysis and a decision on whether the entities are malicious; see [202].

4.1.2 Entities to be assessed by the system, termed "starting entities", are identified according to specific actions defined by a set of rules; see [86-122]. For each starting entity, characteristic threat values (CTVs) are calculated for recorded entity behaviour (243) regarding one or more characteristics (420; see [132-3]) of the entity (see [73-80]); see figure 4; step 430, and [137]. An entity threat value (ETV) is then calculated as the sum of the CTVs (step 440; see [149], in particular the equation at the end) and compared with an entity threat threshold (ETT); see step 450 and [152]. If the ETV is greater than or equal to the threshold then the starting entity is identified as being malicious.

4.1.3 By analysing the relationships between entities (see figures 5A and 5C; 230), the system also derives a group threat value (GTV)(210) for a group of related entities, which is then compared (see [82] and [169]) with a group threat threshold (GTT). If the threshold is exceeded then the entity/entity group is designated malicious.

4.1.4 Related entries are represented as nodes in a tree structure, shown in figure 5A; see [157-8]. The threat value of a parent node (500) is calculated by summing the ETVs of its child nodes (510-540); see [160-163]. Figures 6A and 6B illustrate a method for identifying related entities according to a set of rules; see [175-185 and 192] and page 9, table 2.

4.1.5 The appellant has argued that D2 does not disclose the calculation of a danger parameter based on a decision tree of weight coefficients, since the summation in paragraph [149] was flat, lacking hierarchy. The board disagrees and finds that the summation at the end of paragraph [149] in D2 can be seen as a "decision tree" within the meaning of that term in the claims. In the board's view, a decision tree as claimed cannot be distinguished from the tree representation of an algebraic term, such as (b+c) * d, the root of which has two child nodes, namely d and the sum (b+c), the latter in turn having two child nodes, namely b and c, and the entire tree comprising three "weights": b, c and d. The board sees the summation at the end of paragraph [149] in D2 yielding the entity threat value (ETV) as representing a calculation of a danger parameter based on a "decision tree" (the summation term) of weight coefficients (the CTVs).

4.1.6 Hence, in the terms of claim 1 of the main request, D2 discloses a computer-implemented method for detection of unknown malware (see [8]), the method comprising:

a. receiving (see figure 7; server processing system 720) metadata information about files executed on a remote computer and information about events of execution of the files, including information about actions performed on the files (see [77, 132 and 202]);

c. determining an invocation sequence of the files based on the information about events of file execution, including at least events of file download, dropping, linking and invocation (see "related entity" rule "vii" in [181]);

d. constructing a parent-child hierarchy (see figure 5A) based on the invocation sequence of the files, wherein the parent-child hierarchy describes parent-child relationships between the files (see [158]);

e. performing a risk analysis of the files based on the received, not filtered information and based at least in part on one of the following criteria: statistics of the events of execution of each file, name stability of the file source, address stability of the file source, activity of the file and a behavioral pattern of the file (see [82] regarding "behaviour 253 of the entity"]);

f. performing a risk assessment of the files based on the parent-child hierarchy to determine a level of danger of the related files, wherein a risk assessed to a parent file is based at least in part on the risk associated with one or more children files of the parent file (see figure 5A and [158-162]); and

g. determining whether the files are malicious based on the risk analysis and risk assessment (see [13, 82 and 168-9]),

h. wherein the risk analysis of a file is based on calculating a degree of danger of the file ("Danger" parameter) calculated based on a decision tree of weight coefficients built up for the criteria used in the risk analysis of the file; see figure 5A and [149, 158].

4.2 Document D4

4.2.1 D4 was cited in the decision as evidence that it was known in the prior art to improve the performance of malware detection systems to recognise and quickly deal with known objects, whilst more exhaustive checking could be reserved for unknown objects. The board agrees with this assessment.

4.2.2 According to its abstract and column 1, lines 35 to 47, D4 discloses carrying out a signature check to identify "known" files, be they known to be malicious (on the blacklist) or known not to be malicious (on the whitelist) (see column 2, lines 51 to 67), whereas a risk analysis and a risk assessment are carried out for unknown files, involving deciding which malware detection algorithms, in addition to signature detection, need to be used for the file in order to decide whether it is malware. According to column 2, lines 47 to 51, when a new piece of software appears on the Internet "it takes anywhere from 15 minutes to 2 hours to update the databases of the anti-virus software vendors".

5. Inventive step, Article 56 EPC

5.1 The main request

5.1.1 Claim 1 has been restricted with respect to that in the decision by adding the features (based on page 7, line 24, to page 8, line 23) of several dependent claims in the decision, namely claims 4 (activity, danger and significance parameters), 5 (significance = activity * danger) and 8 (calculation of danger parameter based on a decision tree of weight coefficients for different types of risk-analysis criteria).

5.1.2 The appealed decision found that the subject-matter of claim 1 differed from the disclosure of D2 in the step of

b. filtering out known metadata information and known information about events on [the] basis of information stored in WhiteList and BlackList knowledge databases.

According to the decision, this feature described a known technique for improving performance in malware detection systems by differentiating between known and unknown objects. For the known objects the antivirus check could be relatively short, while for unknown objects the antivirus check was more exhaustive. This technique was disclosed in D4. The person skilled in the art would normally combine two documents in the same technical field in order to achieve an advantage, indicated in D4, column 3, lines 22 to 25. Claim 1 thus lacked inventive step.

5.1.3 In the board's view, the subject-matter of present claim 1 differs from the disclosure of D2 due to the features added (see point 5.1.1 above) not only in feature "b", set out above, but also in the following step:

i. wherein the risk analysis of a file is also based on calculating a level of activity of the file ("Activity" parameter) and a significance for the file ("Significance" parameter), wherein the Significance parameter is computed as Significance = (Activity * Danger).

The appellant has argued that neither D2 nor D4 discloses the calculation of a "danger" level using a decision tree of weights built up for the criteria used in the analysis or the claimed "significance" parameter, taking into account both the level of activity and the degree of danger of file events. D2 merely compared a sum of threat values with a predetermined threshold value to establish whether an entity was malware. The decision tree of the invention had the advantage over the summation of D2 that criteria could be added to, or deleted from, the tree at any time without changing the overall tree structure, and the addition of new criteria at a certain tree level did not influence the weights of the higher tree levels. For instance, the addition of the criteria "Driver installation" and "Direct write to disk" to the tree did not change the weight of their parent node "File system"; see page 12, lines 15 to 21, and figure 5. D4 did not disclose the weighting of the entity threat value (ETV) with an activity parameter, nor did it explain how the risk analysis and risk assessment were carried out. The decision tree used by the invention was more flexible and detailed and thus provided a more accurate approach to determining the "danger" posed by a file.

5.1.4 At the oral proceedings the appellant argued that the expression in claim 1 of each request "detection of unknown malware" meant detection going beyond signature detection in using events and metadata, such as file names and URLs. The appellant also argued that the summation in paragraph [149] of D2 could not be considered a decision tree. The invention did not involve a threshold, unlike D2; see Event Threat Threshold ETT [152].

5.1.5 As stated above, the board regards the summation in paragraph [149] of D2 as a term falling within the expression "decision tree" in claim 1.

5.1.6 Difference features "b" and "i" concern independent stages of the malware detection method, and either one could be implemented without the other, there being no synergistic effect. Hence their contributions to inventive step must be considered separately. The board agrees with the reasoning in the decision (point 4.4) regarding the obviousness of difference "b"; see above. Moreover D4 is evidence that it was known at the priority date to use a "whitelist" and a "blacklist", set out in feature "b", to decide whether a known file is malware not; see column 2, lines 51 to 67.

5.1.7 Turning to feature "i", this feature involves using the "activity" characteristic of an entity as a weighting factor for the "danger" parameter (ETV) already derived in D2. D2 discloses weighting factors; see, for example, the frequency of entity behaviour being used as a multiplier, i.e. a weighting factor, albeit of a constant (0.01), in calculating a characteristic threat value (CTV); see [140]. Moreover figure 2B shows a "Behaviour Recordal Module" (243) and a "Characteristic Analysis Module" (240) for monitoring the behaviour of an entity; see [77, 134]. Under these circumstances the skilled person starting from D2 would have realised these modules to collect "activity" statistics, for instance the number of downloads of an object (see the references to remote network connections in [133, 140], and used the activity factor as a weighting factor for the danger factor as a usual matter of design, thus adding feature "i". Moreover, any subsequent normalization needed to bring the ETV into the required range between 0 and 1 would, in the board's view, fall within the mathematical competence of the person skilled in the art.

5.1.8 Consequently the subject-matter of claim 1 does not involve an inventive step, Article 56 EPC, in view of D2 and D4.

5.2 The first auxiliary request

5.2.1 Claim 1 of the first auxiliary request differs from the independent claims of the main request in the feature (based on page 8, lines 19 to 21) that:

j. a majority of the weight coefficients is dynamically changing depending on the accumulated information in the knowledge databases.

5.2.2 The appellant has argued that the dynamic adjustment provided a better real-time detection of previously unknown malicious entities and questioned whether it would have been obvious for the skilled person to add this feature. In contrast, D2 (see [138, 140-144]) used constant values or constant values weighted by a frequency factor, rather than values depending on database values. Such dynamic adjustment provided more accurate risk analysis.

5.2.3 The board understands feature "j" to mean that a majority of the weight coefficients change if the pertinent content of the knowledge databases changes.

5.2.4 In the board's view, updating databases which are simultaneously responding to queries belongs to the normal operation of a database, which would have been within the competence of the skilled person. Such database updates might cause dynamic changes in the weight coefficients, and sometimes in a majority of the weight coefficients. D4 refers (column 2, lines 47 to 51) to the databases of the anti-virus software vendors being updated in response to the appearance of a new piece of software on the Internet, the update occurring within 15 minutes to 2 hours. Whenever updates to the databases in D2, the whitelist and blacklist, had an impact on the weight coefficients, it would have been obvious for the skilled person to consider changing the weights accordingly, and doing it, if desired, would have been straightforward (feature "j").

5.2.5 This feature therefore cannot lend inventive step to the subject-matter of claim 1 of the first auxiliary request.

6. The second auxiliary request

6.1 Compared to claim 1 of the main request, claim of this request sets out the following expression at the end:

", wherein the decision tree of weight is expandable by adding new criteria, wherein a maximum weight of a criterion associated with a higher level tree node is not changed by weights of added criteria at lower tree node levels".

6.2 In the oral proceedings the board raised an objection under Article 123(2) EPC that the maximum value of a higher level tree node not being changed by weights of added criteria at lower tree levels was not directly and unambiguously derivable from the original application, since it was broader than the specific disclosure in figure 5 of the addition of two new criteria 5.3.7 (Driver Installation) and 5.3.8 (Direct write to disk) at the lowest level in the tree not causing the value of the tree node directly above them - "5.3 File system" - to change.

6.3 The appellant argued that the basis for this amendment was provided by figure 5 and "Example 4", in particular page 12, lines 17 to 21, which states that:

"An addition of two new criteria, 'Driver Installation' and 'Direct write to Disk', will lead to the modification of the decision tree highlighted in **FIG. 5**. After this modification of the decision tree, the maximum weight of the criterion 5.3 will not change. However, this criterion will become more informative and precise, which in turn will positively affect the quality of the decisions."

The skilled person would have understood this passage in context to disclose the feature added to claim 1.

6.4 The board finds that the cited passage only relates to the case of additional criteria being added at the lowest tree level without changing the maximum tree node (5.3) directly above them. Contrary to Article 123(2) EPC, there is no basis in the original application, in particular page 12, lines 19 to 21, for generalising this feature to the maximum value of a higher level tree node not being changed by weights of added criteria at lower tree levels. In particular, the description fails to state explicitly that the disclosed embodiment is just one example of a more general concept, and it does not disclose the method by which the invention achieves the claimed effect (that the weights at higher levels do not change in response to lower level changes), from which the skilled person might be able to derive a more general principle.

6.5 As a consequence, claim 1 does not comply with Article 123(2) EPC.

7. Conclusion

As none of the requests complied with the EPC, the appeal could not be allowed.

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar: The Chairman:

Beschwerdekammern des Europäischen Patentamts
Boards of Appeal of the European Patent Office
Chambres de recours de l'Office européen des brevets

I. Aperribay M. Müller

Decision electronically authenticated