(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 May 2016 (12.05.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/072839 A1

(51) International Patent Classification:
G06F 17/18 (2006.01) *G06K 9/62* (2006.01)

(21) International Application Number:
PCT/MY2015/050137

(22) International Filing Date:
30 October 2015 (30.10.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PI2014003120 6 November 2014 (06.11.2014) MY

(71) Applicant: **MIMOS BERHAD** [MY/MY]; Technology Park of Malaysia, Bukit Jalil, 57000 Kuala Lumpur (MY).

(72) Inventors: **CHEW, Wooi Keat, Martin**; c/o MIMOS Berhad, Technology Park of Malaysia, Bukit Jalil, 57000 Kuala Lumpur (MY). **FOO, Chi Hui**; c/o MIMOS Berhad, Technology Park of Malaysia, Bukit Jalil, 57000 Kuala Lumpur (MY). **KONERU, Venkaiah Chowdary**; c/o MIMOS Berhad, Technology Park of Malaysia, Bukit Jalil, 57000 Kuala Lumpur (MY). **SOMINENI, Srinu**; c/o MIMOS Berhad, Technology Park of Malaysia, Bukit Jalil, 57000 Kuala Lumpur (MY).

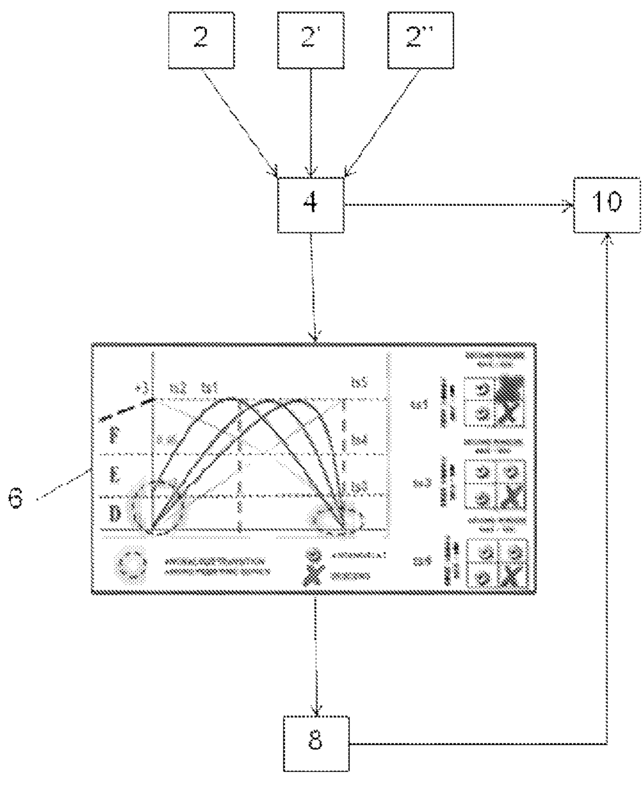
(74) Agents: **HEMINGWAY, Christopher Paul** et al.; Marks & Clerk (Malaysia) Sdn Bhd, Unit 6, Level 20, Tower A, Menara UOA Bangsar, 5 Jalan Bangsar Utama 1, Taman Bangsar, 59000 Kuala Lumpur (MY).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR IDENTIFYING POTENTIAL ANOMALIES IN TIME SERIES DATA

Figure 1



(57) Abstract: A system and method for identifying potential anomalies in time series data, the time series being normalised (100), and the unit and time scales being divided (202, 206) into sections; the time series for each time section being converted (200) into symbolic representations representing a plurality of different values within a time section; symbols representing mean and slope values for each series over time being compared such that those with smaller differences are identified (400) as peers; wherein potential anomalies are identified (500) if a transition between symbols representing maximum or minimum values for one peer is different to the other peers in the group.

WO 2016/072839 A1

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

SYSTEM AND METHOD FOR IDENTIFYING POTENTIAL ANOMALIES IN TIME SERIES DATA

Field of Invention

The invention relates to a system and method for identifying potential anomalies in time series data.

Background

In recent times data sets have become so large that it is becoming increasingly difficult to analyse them, and because of the amount of data it can be difficult to spot anomalies.

One solution to this issue is to use symbolic dynamic filtering in which time-series data are partitioned, symbol sequences are generated to generalise the same, and then compared to standard sequences to see if there are any differences. For example, flight recorder data may be analysed in this manner.

However, in very heterogeneous data, standard sequences for comparison may not be available, and therefore it is problematic to spot anomalies in such data. For example, a large amount of data is generated in tracking food from a source to a destination, so that every step can be checked if required, but if a supplier manipulates an invoice, it is hard to identify such a fraudulent entry in the mass of additional data.

Therefore an aim of the invention is to provide a method for identifying anomalies which overcomes the above problem.

Summary of Invention

In an aspect of the invention, there is provided a system for identifying potential anomalies in time series data which can be plotted as units against time, comprising receiving means for receiving time series data from one or more databases, processing means for normalising the time series and dividing unit and time scales into sections; and output means for outputting the results including the potential anomalies identified from the processing means; characterised in that the processing means converts the time series for each time section into symbolic representations representing a plurality of different values within a time section; means are provided to construct a scoring table in which two symbols representing different values for each series over time are compared such that those with smaller differences are identified as peers; and means are provided to compare two symbols representing different values within peer groups, wherein potential anomalies are identified if a transition for one peer is different to the other peers in the group.

In one embodiment the symbols used to identify peers represent mean and slope values.

In one embodiment the symbols compared within peer groups represent maximum and minimum values.

In one embodiment the unit scale is divided into sections representing substantially equal probability.

In one embodiment the symbolic representations are provided as a quadruplet codon for each series for each time section.

In one embodiment the scoring table is constructed using a Markov model.

In one embodiment a transition is the difference between a pair of symbols of a time series, each symbol being in a different time segment. Typically transitions of the same pair type are compared with peers in the same time segments.

In one embodiment the data series are normalised by calculating the mean thereof, moving the mean point to 0 of the x-axis, mapping the upper and lower boundaries of the time series to a predetermined number of standard deviations on the y-axis, and scaling each point in the time series to be between the upper and lower boundaries. Typically three standard deviations are used.

In one embodiment the anomaly detection sensitivity can be changed by adjusting the level of difference required to register an anomaly.

In one embodiment the results are stored in a database.

In one embodiment the system includes transmission means for notifying a user if a potential anomaly is detected. Typically the user is sent the results in which the potential anomaly is highlighted, for review and audit of the same.

In a further aspect of the invention, there is provided a method of identifying anomalies in time series data which can be plotted as units against time, comprising the steps of: normalising the time series; dividing unit and time scales into sections; characterised by converting the time series for each time section into symbolic representations representing a plurality of different values within a time section; constructing a scoring table in which two symbols representing different values for each series over time are compared such that those with smaller differences are identified as peers; and comparing two symbols representing different values within peer groups, potential anomalies being identified if a transition for one peer is different to the other peers in the group.

Brief Description of Drawings

It will be convenient to further describe the present invention with respect to the accompanying drawings that illustrate possible arrangements of the invention. Other arrangements of the invention are possible, and consequently the particularity of the accompanying drawings is not to be understood as superseding the generality of the preceding description of the invention.

Figure 1 is a schematic overview of an embodiment of the invention.

Figure 2 illustrates normalisation of the time series data (a) method steps; (b) dividing y-axis into equal probability sections; (b) time series after normalisation.

Figure 3 illustrates conversion of the time series data to symbolic representations (a) method steps; (b) time series prior to normalisation; (b) assigning quadruplet codons for each series for each time section.

Figure 4 illustrates the sectioning of an normal probability distribution (a) area under a normal probability distribution; (b) half the area; (c) first section; (d) second section; (e) third section; (f) final apportionment.

Figure 5 illustrates a method of generating a scoring table based on a Markov model.

Figure 6 illustrates a method of identifying a peer group of candidate time series based on Markov model scoring table.

Figure 7 illustrates a method of identifying local anomalies of candidate time series with respect to the peer group.

Figure 8 illustrates a summary of an example of the method for identifying potential anomalies: (a)

Figure 9 illustrates an example output

Detailed Description

With regard to Figure 1, there is illustrated a schematic view of the overall system.

The system comprises an electronic processing device 4, for receiving time series data from databases 2, 2', 2'', and processing the data to generate output images 6 which are stored in a database 8. If any potential anomalies are detected in the data, the relevant output images are displayed on a monitor or transmitted to a remote devicefor a user 10 to review.

With reference to Figures 2a-c, each time series are first normalised 100, by calculating 102 the mean of the time series, translating 104 the time series by moving the mean point to 0 of the x-axis, mapping 106 the upper boundary of the time series to +3 standard deviations on the y-axis, and similarly mapping 108 the lower boundary of the time series to -3 standard deviations on the y-axis., then scaling 110 each point in the time series to be between the upper and lower boundaries.

Three standard deviations are used as in a normal distribution this covers 99.73% (i.e. nearly all) of data.

With regard to Figures 3a-c, the time series are then converted 200 to symbolic representations (A-F in this example), by dividing 202 the y-axis into equal probability sections based on the normal distribution (as further described with reference to Figure 4), assigning 204 a token to each y-axis division, dividing 206 the x-axis into equal width windows representing time sections, and identifying 208 tokens corresponding to maximum, mean, minimum and slope (maximum-minimum) values of each time series within each time section, such that after all time sections have been tokenized 210, the tokens are concatenated 212 into a final symbolic representation for each time series.

As an example with reference to Figure 3c, time series ts1 in the first time section window has a minimum value (square) corresponding to unit section having symbol B, a mean value (triangle) corresponding to symbol D, a maximum value (circle) corresponding to symbol F, and a positive slope direction P (P for positive, N for negative) between start and end of the time section. Therefore the final symbolic representation representing max-mean-min-slope is FDBP. These quadruplet codons are determined for each series for each time section.

With respect to Figures 4a-f, the division of the y-axis into equal probability sections based on the normal distribution is illustrated in greater details. Figure 4a indicates the area under a normal probability distribution with three standard deviations (i.e. from -3 to +3) which sums to approximately 1.0. Figure 4b illustrates one side which sums to 0.5. Dividing one side into three equal portions makes each area equal to 0.166 (i.e. 0.5 / 3). In Figures 4c-e, first portion (i.e. symbol D) ranges from 0.00 to 0.43 on the x-axis, second portion (i.e. symbol E) ranges from 0.43 to 0.96, and third portion (i.e. symbol F) ranges from 0.97 onwards. Figure 4f illustrates the final division of the y-axis into six portions – each represented by an alphabetical token A-F.

With regard to Figure 5, a scoring table based on a Markov model is generated. A table of time series dataset symbolic representations is created 302 from which a time series is selected 304. The token probability for mean and slope columns of the selected time series is calculated 306 and summed 308. This is repeated 310 for every time series in the table until all are completed and the end scoring is determined 312.

With respect to Figure 6, a peer group of candidate times series is identified 400 based on the Markov model scoring table. A candidate time series is selected 402, and the token probability sum of the candidate time series is compared 404 with other time series. The nearest of the other time series are tagged 406 as potential peers and a different count is then performed 408 between potential peers and candidate time series for verification.

With regard to Figure 7, local anomalies of candidate time series are identified 500 in the peer group. Initially, two adjacent equal-width time section windows are selected 502 from peer time series. The ascent/descent of max/min columns are compared 504 with other time series. If an anomaly is suspected 506, the selected time section of the relevant time series is flagged 508 as being a potential anomaly. The process then repeats for other time sections. If no other time sections from the candidate time series remain 510, the number of potential local anomalies in candidate time series are counted 512.

With respect to Figures 8a-e, a summary of an example of the method for identifying potential anomalies is illustrated. In Figure 8a five time series (ts1, ts2, ts3, ts4 and ts5) are converted to symbolic representations and tabulated as shown in Figure 8b for building a simple Markov model scoring table. Given a collection of symbolic representations of time series (i.e. ts1, ts2, ts3, ts4, ts5), any given symbolic representation (e.g. ts5) in that collection are scored with respect to the whole collection by the summation of the mean and slope columnar probabilities of that given symbolic representation – mean and slope are used to represent the global similarities of a given time series. For example the columnar probability of ts5 in the slope column of the 1st window is 0.75 because ‘P’ appears three times in the combination ‘PNPP’.

As indicated in Figure 8c, the calculated scores are used to identify peers for a given time series symbolic representation (e.g. potential peers of ts1), peers being within a selected range.

As a check, a simple difference calculation is performed between the symbolic representation of the given time series and the symbolic representation of a possible peer to ascertain if the difference is within a pre-defined range, where if there is a difference, output as ‘1’, otherwise output as ‘0’, such that the final count of differences indicates the similarity (smaller numbers = greater similarity). For example:

Time Series		Symbolic Representation		
ts1	→	FFEP FEDN		
ts2	→	FFFN FEDN	(minus)	
		*0*1 *0*0	→	0 + 1 + 0 + 0 = 1
Time Series		Symbolic Representation		
ts1	→	FFEP FEDN		
ts3	→	FEDP FEDN	(minus)	
		*1*0 *0*0	→	1 + 0 + 0 + 0 = 1
Time Series		Symbolic Representation		
ts1	→	FFEP FEDN		
ts5	→	EDDP FFEP	(minus)	

*1*0 *1*1 → 1 + 0 + 1 + 1 = 3

Thus as illustrated in Figures 8d-e, local anomaly can be detected for max-max, max-min, min-min, and min-max by comparing the transitions in selected columns (i.e. ascent or descent) of the candidate time series with its peers. An anomaly is flagged when a time series makes a transition (ascend, descend, or flat) that no other of its peers make. Max and min are used because local anomaly detection focuses on the basic points and not the derived (i.e. mean and slope) points. The sensitivity of the method can be changed by adjusting the level of difference required to register an anomaly.

The method can be applied with varying user-defined numbers of y-axis tokens and x-axis equal windows to surface historical local anomalies. The minimum resolution for the x-axis is 2 data points. The more potential historical local anomalies are highlighted for a given time series, the higher the probability of data integrity issues for the candidate time series.

With regard to Figure 9, an output is produced summarising the results, in which the quadruplet codons 602 of the peer time series 606 used to identify potential local anomalies are illustrated, together with the non-peer time series 608. Any potentially anomalies 604 are highlighted for ease of understanding.

It will be appreciated by persons skilled in the art that the present invention may also include further additional modifications made to the system which does not affect the overall functioning of the system.

CLAIMS

1. A system for identifying potential anomalies in time series data which can be plotted as units against time, comprising:

receiving means for receiving time series data from one or more databases (2, 2', 2'');

processing means for normalising the time series and dividing unit and time scales into sections; and

output means for outputting the results (6) including the potential anomalies identified from the processing means;

characterised in that the processing means converts the time series for each time section into symbolic representations representing a plurality of different values within a time section;

means are provided to construct a scoring table in which two symbols representing different values for each series over time are compared such that those with smaller differences are identified as peers; and

means are provided to compare two symbols representing different values within peer groups, wherein potential anomalies are identified if a transition for one peer is different to the other peers in the group.

2. A system according to claim 1 wherein the symbols used to identify peers represent mean and slope values.

3. A system according to claim 1 wherein the symbols compared within peer groups represent maximum and minimum values.

4. A system according to claim 1 wherein the unit scale is divided into sections representing substantially equal probability.

5. A system according to claim 1 wherein the scoring table is constructed using a Markov model.

6. A system according to claim 1 wherein a transition is the difference between a pair of symbols of a time series, each symbol being in a different time segment.

7. A system according to claim 6 wherein transitions of the same pair type are compared with peers in the same time segments.

8. A system according to claim 1 wherein the data series are normalised by calculating (102) the mean thereof, moving (104) the mean point to 0 of the x-axis, mapping (106, 108) the upper and lower boundaries of the time series to three standard deviations on the y-axis, and scaling (110) each point in the time series to be between the upper and lower boundaries.

9. A system according to claim 1 wherein the anomaly detection sensitivity can be changed by adjusting the level of difference required to register an anomaly.

10. A method of identifying anomalies in time series data which can be plotted as units against time, comprising the steps of:

normalising the time series;

dividing (202, 206) unit and time scales into sections;

characterised by converting (200) the time series for each time section into symbolic representations representing a plurality of different values within a time section;

constructing a scoring table in which two symbols representing different values for each series over time are compared such that those with smaller differences are identified as peers; and

comparing two symbols representing different values within peer groups, potential anomalies being identified if a transition for one peer is different to the other peers in the group.

Figure 1

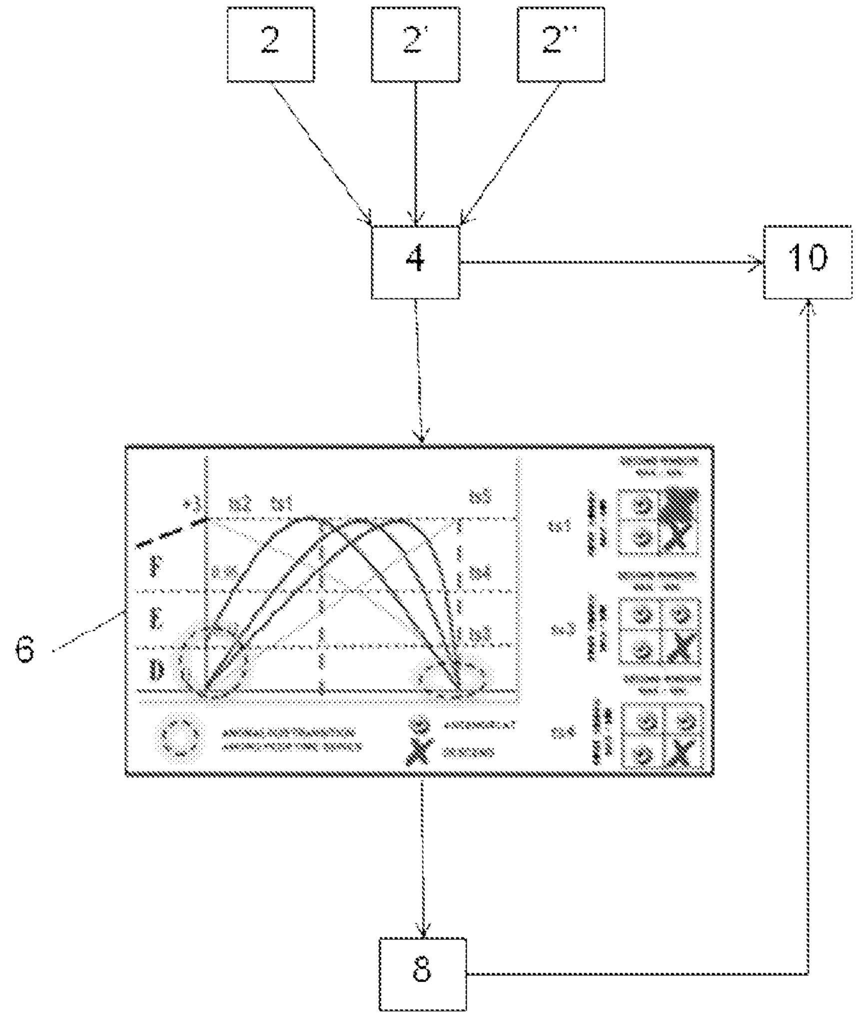


Figure 2a

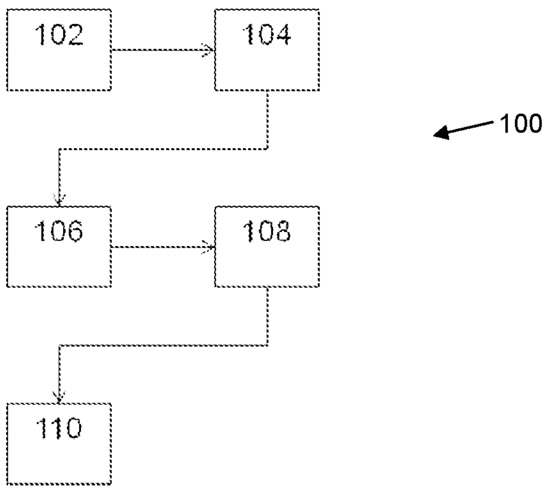


Figure 2b

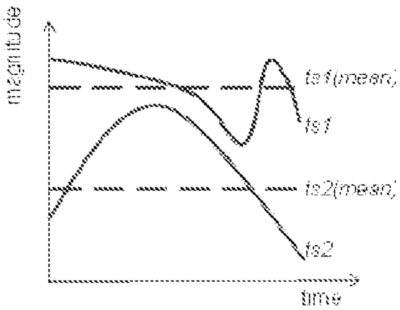


Figure 2c

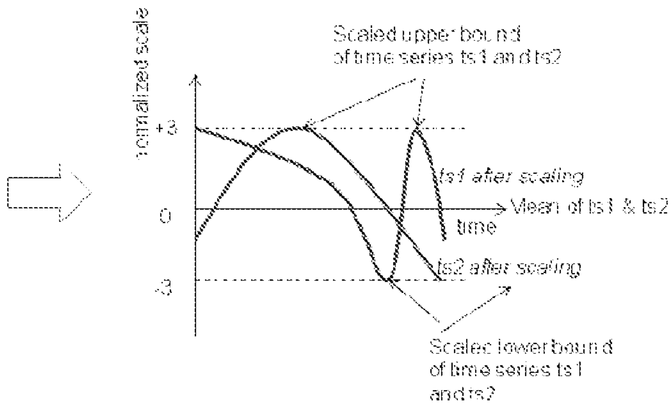


Figure 3a

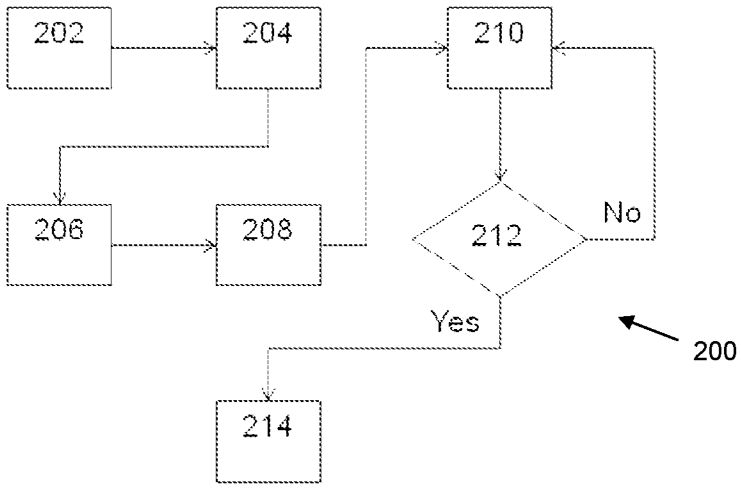


Figure 3b

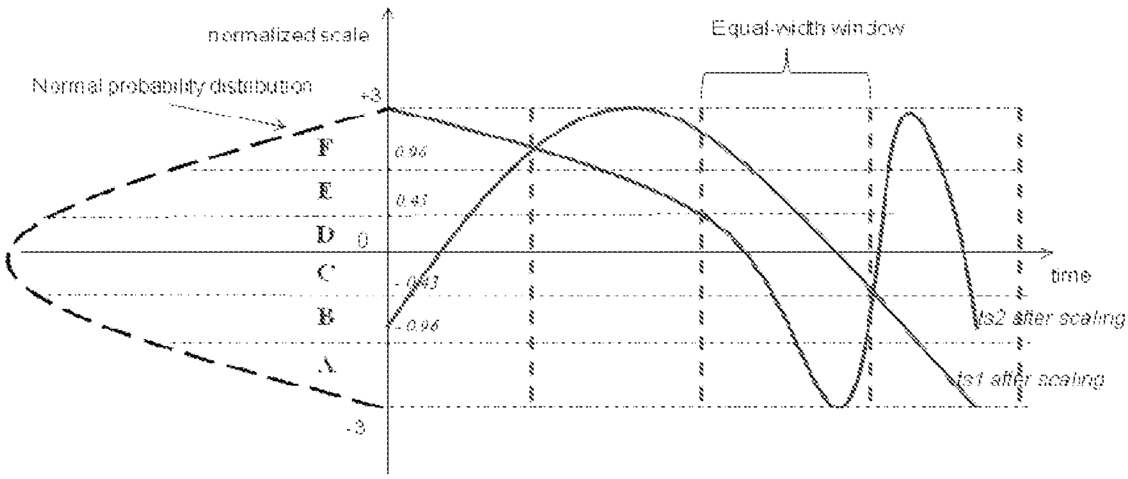
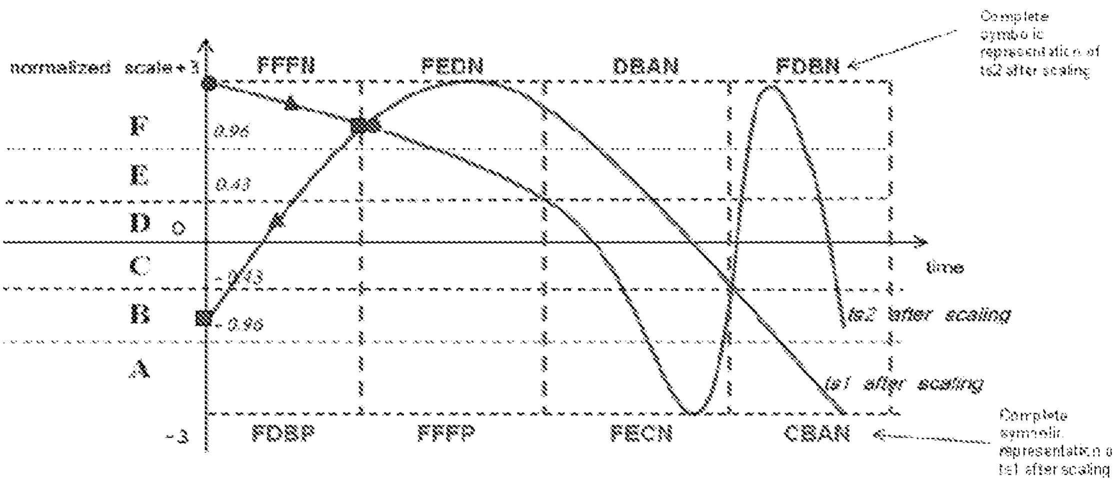


Figure 3c



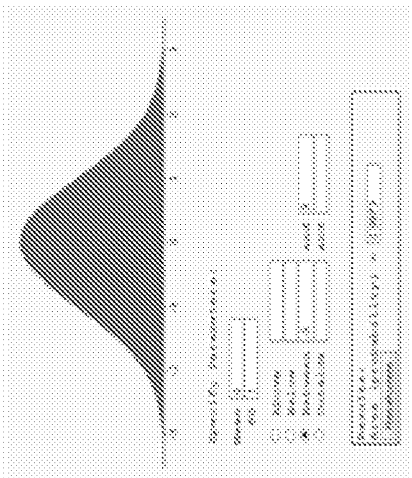


Figure 4a

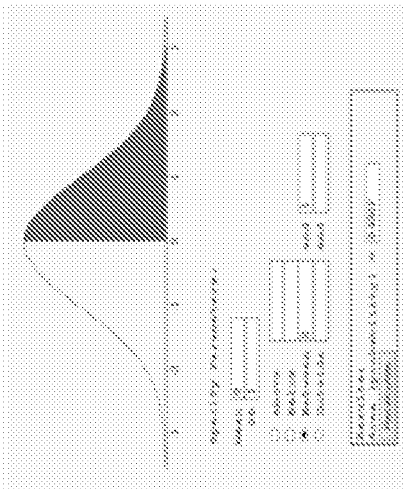


Figure 4b

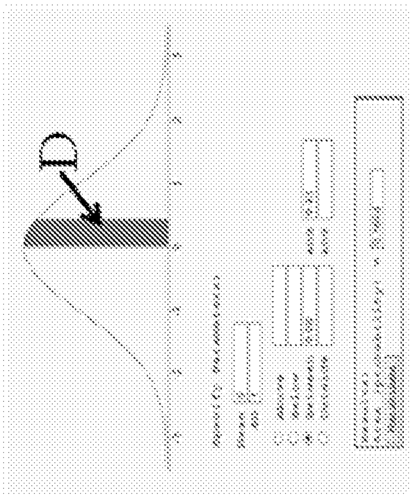


Figure 4c

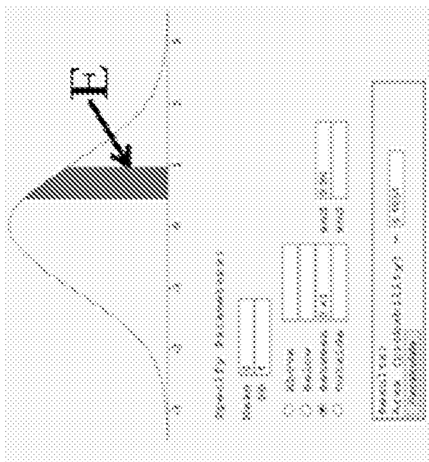


Figure 4d

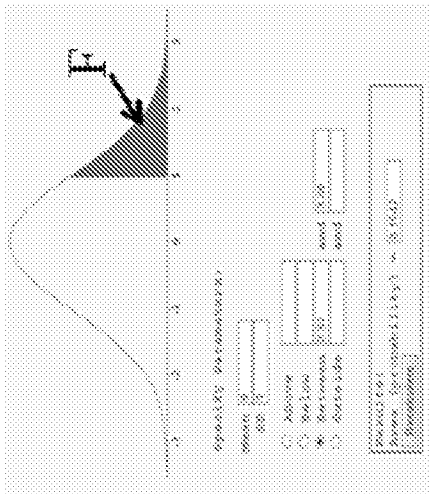


Figure 4e

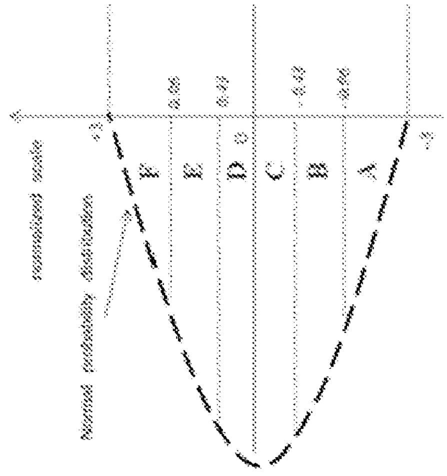


Figure 4f

Figure 5

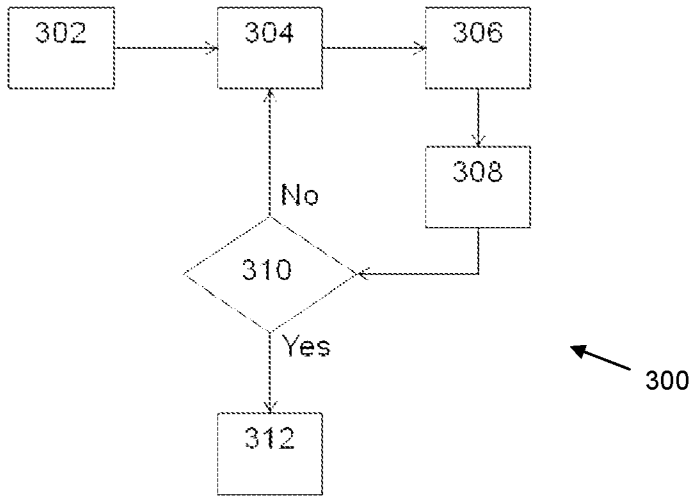


Figure 6

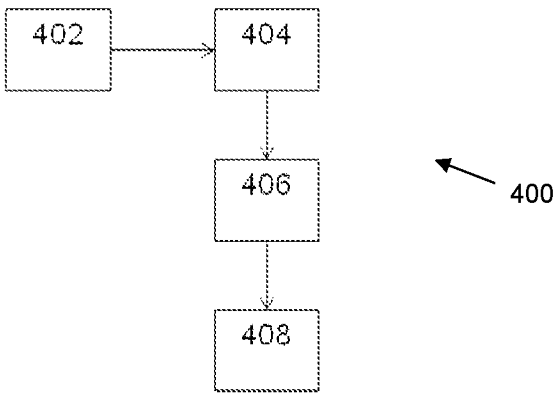


Figure 7

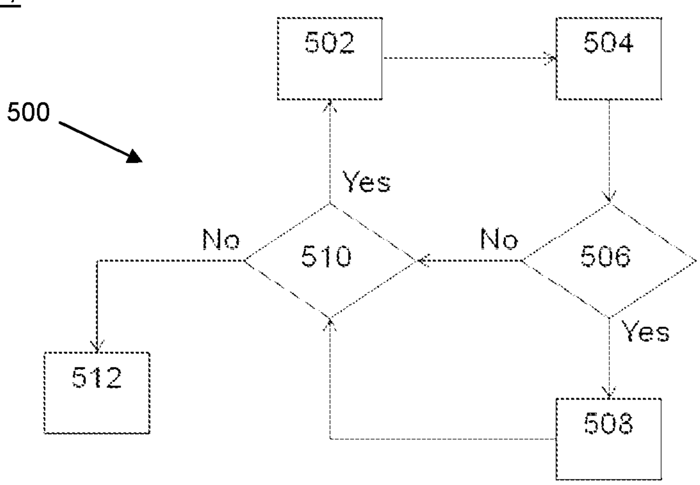


Figure 8a

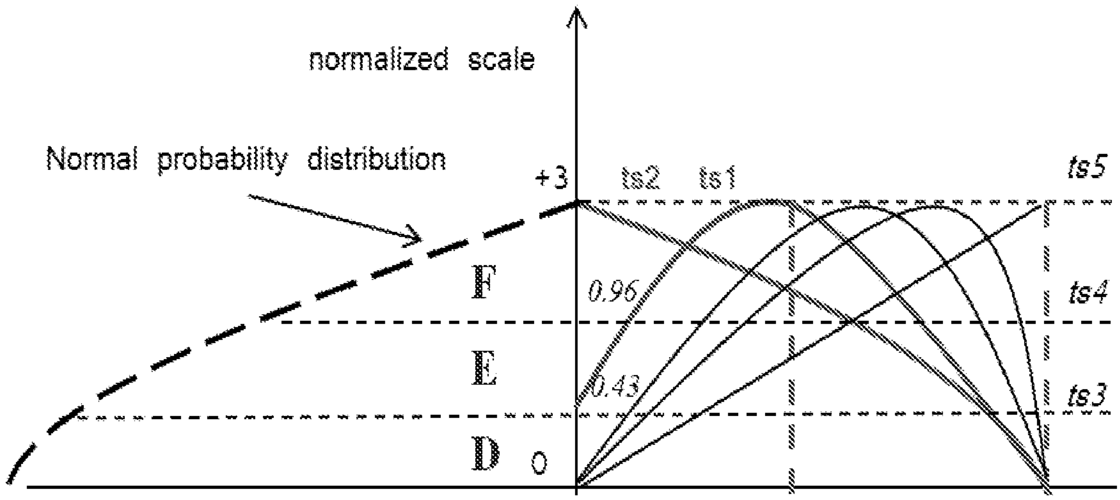


Figure 8b

SERIES	1st WINDOW				2nd WINDOW			
	MAX	MEAN	MIN	SLOPE	MAX	MEAN	MIN	SLOPE
ts1	F	F	E	P	F	E	D	N
ts2	F	F	F	N	F	E	D	N
ts3	F	E	D	P	F	E	D	N
ts4	F	E	D	P	F	E	D	N
ts5	E	D	D	P	F	F	E	P
Columnar probability of ts5	-	0.00	-	0.75	-	0.00	-	0.00

Figure 8c

SERIES	1st WINDOW				2nd WINDOW				SCORE
	MAX	MEAN	MIN	SLOPE	MAX	MEAN	MIN	SLOPE	
ts1	F	F	E	P	F	E	D	N	2.50
ts2	F	F	F	N	F	E	D	N	1.75
ts3	F	E	D	P	F	E	D	N	2.50
ts4	F	E	D	P	F	E	D	N	2.50
ts5	E	D	D	P	F	F	E	P	0.75

Potential peer time series for ts1

Figure 8d

SERIES	1ST WINDOW				2ND WINDOW				SCORE	PEERS OF ts1?
	MAX	MEAN	MIN	SLOPE	MAX	MEAN	MIN	SLOPE		
ts1	F	F	E	P	F	E	D	N	2.50	Candidate
ts2	F	F	F	N	F	E	D	N	1.75	No
ts3	F	E	D	P	F	E	D	N	2.50	Yes
ts4	F	E	D	P	F	E	D	N	2.50	Yes
ts5	E	D	D	P	F	F	E	P	0.75	No

Possible anomaly as the previous min-to-min of ts1 "descends" with no descent seen in the peers

Figure 8e

SERIES	1ST WINDOW				2ND WINDOW			
	MAX	MEAN	MIN	SLOPE	MAX	MEAN	MIN	SLOPE
ts1	F	F	E	P	F	E	D	N
ts2	F	F	F	N	F	E	D	N
ts3	F	E	D	P	F	E	D	N
ts4	F	E	D	P	F	E	D	N
ts5	E	D	D	P	F	F	E	P

SERIES	1ST WINDOW				2ND WINDOW			
	MAX	MEAN	MIN	SLOPE	MAX	MEAN	MIN	SLOPE
ts1	F	F	E	P	F	E	D	N
ts2	F	F	F	N	F	E	D	N
ts3	F	E	D	P	F	E	D	N
ts4	F	E	D	P	F	E	D	N
ts5	E	D	D	P	F	F	E	P

SERIES	1ST WINDOW				2ND WINDOW			
	MAX	MEAN	MIN	SLOPE	MAX	MEAN	MIN	SLOPE
ts1	F	F	E	P	F	E	D	N
ts2	F	F	F	N	F	E	D	N
ts3	F	E	D	P	F	E	D	N
ts4	F	E	D	P	F	E	D	N
ts5	E	D	D	P	F	F	E	P

SERIES	1ST WINDOW				2ND WINDOW			
	MAX	MEAN	MIN	SLOPE	MAX	MEAN	MIN	SLOPE
ts1	F	F	E	P	F	E	D	N
ts2	F	F	F	N	F	E	D	N
ts3	F	E	D	P	F	E	D	N
ts4	F	E	D	P	F	E	D	N
ts5	E	D	D	P	F	F	E	P

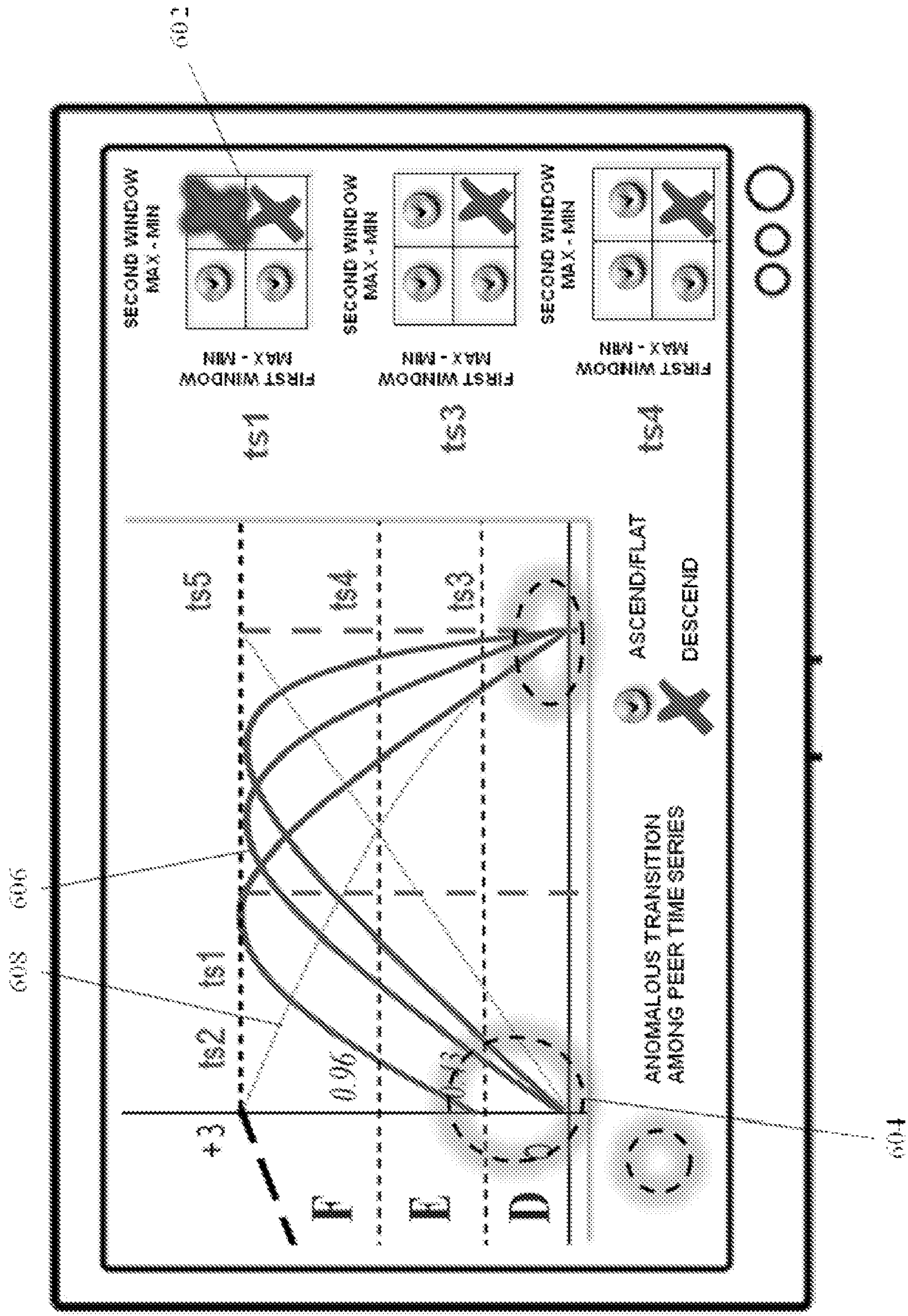

602
SECOND WINDOW
MAX - MIN
FIRST WINDOW
MAX - MIN
ts1
ts3
ts4
ts5
ts2
0.96
+3
F
E
D
ASCEND/FLAT
DESCEND
ANOMALOUS TRANSITION
AMONG PEER TIME SERIES
606
608
604

Figure 9

INTERNATIONAL SEARCH REPORT

International application No
PCT/MY2015/050137

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F17/18 G06K9/62
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F G06K G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LIN J ET AL: "Experiencing SAX: a novel symbolic representation of time series", DATA MINING AND KNOWLEDGE DISCOVERY, vol. 15, no. 2, 3 April 2007 (2007-04-03), pages 107-144, XP019525926, ISSN: 1573-756X, DOI: 10.1007/S10618-007-0064-Z section 3: introduction, 3.1-3.3 sections 4.4.1, 4.5.2 -----	1-10
X	LI G ET AL: "TSX: A novel symbolic representation for financial time series", PRICAI 2012; LECTURE NOTES IN ARTIFICIAL INTELLIGENCE, vol. 7458, 3 September 2012 (2012-09-03), pages 262-273, XP047015307, ISBN: 978-3-642-32694-3 sections 1, 3, 4 -----	1-10
	-/--	

[X] Further documents are listed in the continuation of Box C.

[ ] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 February 2016	29/02/2016

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Domingo Vecchioni, M

1

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MALINOWSKI S ET AL: "1d-SAX: A novel symbolic representation for time series", CORRECT SYSTEM DESIGN; LECTURE NOTES IN COMPUTER SCIENCE, vol. 8207, no. 558, 17 October 2013 (2013-10-17), pages 273-284, XP047043383, ISSN: 0302-9743 ISBN: 978-3-642-22877-3 sections 2-4 -----	1-10
X	ESMAEL B ET AL: "Multivariate time series classification by combining trend-based and value-based approximations", COMPUTATIONAL SCIENCE AND ITS APPLICATIONS ICCSA 2012, LECTURE NOTES IN COMPUTER SCIENCE, vol. 7336, 18 June 2012 (2012-06-18), pages 392-403, XP047007443, ISBN: 978-3-642-31127-7 sections 1, 2-4 -----	1-10
X	LKHAGVA B ET AL: "New time series data representation ESAX for financial applications", PROCEEDINGS OF THE 22ND INTERNATIONAL CONFERENCE ON DATA ENGINEERING WORKSHOPS (ICDEW'06), 3-7 APRIL 2006, ATLANTA, GA, USA, 3 April 2006 (2006-04-03), pages 17-22, XP010912010, DOI: 10.1109/ICDEW.2006.99 ISBN: 978-0-7695-2571-6 sections 1-3, 5 -----	1-10

1