



- (51) **International Patent Classification:**
H04L 12/26 (2006.01) *H04L 12/24* (2006.01)
- (21) **International Application Number:**
PCT/US2015/023491
- (22) **International Filing Date:**
31 March 2015 (31.03.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (71) **Applicant:** HEWLETT PACKARD ENTERPRISE DEVELOPMENT LP [US/US]; 11445 Compaq Center Drive West, Houston, TX 77070 (US).
- (72) **Inventors:** FRAENKEL, Noam; Shabazi 26, 56100 Yehud (IL). SHAKED, Doron; Kiryat Technion Technion, 32000 Haifa (IL). COHEN HAREL, Maya; Einstein 10, Kiryat Weizmann, Ness Ziona, 76101 Nes Ziona (IL).
- (74) **Agents:** KHAN, Saif M et al.; Hewlett Packard Enterprise, 3404 E. Harmony Road, Mail Stop 79, Fort Collins, CO 80528 (US).

- (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.
- (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).

[Continued on next page]

(54) **Title:** INFERRING A NETWORK TOPOLOGY

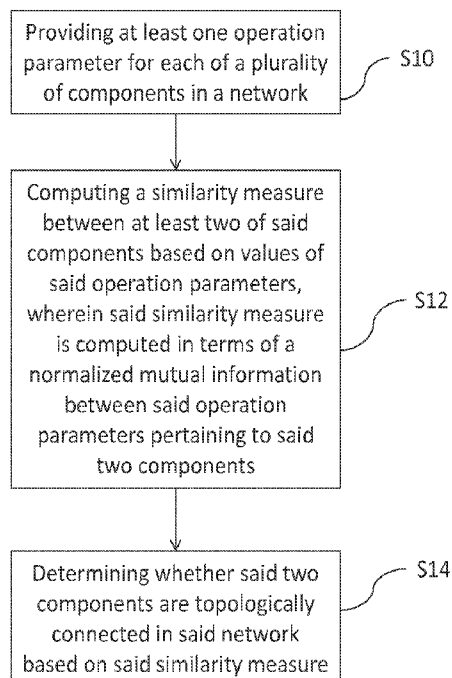


Fig. 3

(57) **Abstract:** In a method for inferring a topology of components in a network, at least one operation parameter is provided for each of a plurality of components in a network, and a similarity measure is computed between at least two of said components based on values of said operation parameters. Based on said similarity measure, it is determined whether said two components are topologically connected, wherein said similarity measure is computed in terms of a normalized mutual information between said operation parameters pertaining to said two components



Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

Published:

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

5

Inferring a Network Topology

Background

10 Topology, in information technology (IT) management terms, characterizes the way computers, network devices, software components, computer systems, etc. are connected in an IT network or environment. Topological information is paramount in providing sophisticated information technology management capabilities and analysis, such as change management, event management, problem resolution, and operation analysis.

15

Brief Description of the Drawings

Fig. 1 illustrates a system according to an example;

20 Fig. 2 is a schematic diagram of a network environment in which the method and system according to an example of the present disclosure may be employed;

Fig. 3 is a flow diagram for illustrating an example of a method for inferring a
25 topology of components in a network;

Fig. 4 is a flow diagram for illustrating another example of a method for inferring a topology of components in a network; and

30 Fig. 5 is a schematic diagram of the network environment in Fig. 2, but in addition shows the discovered network connections according to an example.

Detailed Description

Establishing and maintaining the topological model of the network environment may be intricate and difficult, in particular for large networks. Some common techniques include IP scans, port scans, sniffing network communication, probing live systems for known ports, or scanning file systems and configuration files. A common trait of these methods is that they may be active processes that are initiated solely for the purpose of discovering the environment, and these processes may run periodically, or even constantly, if real-time updates are desired, and may require considerable configuration, maintenance and management resources.

Accordingly, the present disclosure provides examples of methods and systems which may, in some examples, allow inferring a topology of components in a network with less effort and resources.

In an example, a method for inferring a topology of components in a network comprises providing at least one operation parameter for each of a plurality of components in a network, computing a similarity measure between at least two of said components based on values of said operation parameters, and determining whether said two components are topologically connected in said network based on said similarity measure. Said similarity measure is computed in terms of a normalized mutual information between said operation parameters pertaining to said two components.

Normalized mutual information may, in some examples, yield reliable traces of the topology of the components in a network based on statistical data of operation parameters collected from the plurality of components in the network. In particular, mutual information may allow to leverage data already collected and maintained in many IT management applications to infer and maintain the topology model of the network with minimal additional resources and investment.

Mutual information is generally known in information theory as a measure to quantify the

resemblance of probability distributions, and corresponds to the relative entropy between the joint distribution and the product distribution. This measure may reliably be used in the context of network analysis with a large variety of different input data, and hence is versatile and may provide a particularly reliable and sound assessment of the network topology based on the leverage data that is available for the components in a given network.

In the sense of the present disclosure, components may be understood to refer to any units or entities that provide operation parameters that can be monitored. In an IT environment, these may typically be network devices, computer systems, or software components.

A network, in the sense of the present disclosure, denotes connections or links between the components. The links may correspond to a wired or wireless connection between the components. However, the links may also be virtual links in the sense that they merely indicate some correlation or synchronization of components in the network, or synchronization in the way these components work.

A network may generally be represented in terms of a graph, with the nodes of the graph corresponding to the components and the edges corresponding to the connections or links. A graph represents a certain topology, and the present disclosure is concerned with inferring the topology, i.e. the connections, based on an analysis of the operation parameters that are available from the components.

In some examples, the normalization of the mutual information may enhance the suitability of the mutual information as similarity measure.

In particular, in some examples, said mutual information may be normalized with a joint entropy of said operation parameters pertaining to said two components, in particular by dividing said mutual information by said joint entropy.

In some examples, normalization with the joint entropy allows to provide a bounded similarity measure.

The normalization may also comprise a normalization factor over a sum of said mutual information, said sum over said components.

5 Normalization by dividing the mutual information by a sum of mutual information, over at least part of said components, allows to reduce the adverse effect of random processes or, conversely, monotone signals on the similarity measure.

In an example, a plurality of operation parameters may be provided.

10 As assessment of the similarity based on a plurality of operation parameters may enhances the accuracy.

In general, said similarity measure, in particular said mutual information, may involve a comparison of different operation parameters relating to different devices. For instance, a
15 similarity may be determined between a processor reading of the first device, and error statistics of a second device.

However, in some examples, said mutual information is given in terms of corresponding or the same operation parameters pertaining to said two components.

20 In an example, said similarity measure is computed in terms of a weighted normalized mutual information.

Weights may be assigned to said operation parameters.

25 With a suitable weighting, accidental synchronized behavior that might otherwise distort the results can be efficiently suppressed.

Said weights of said weighted mutual information may be given in terms of a probability
30 distribution.

In some examples, said weights of said weighted mutual information are given in terms of a standard deviation of said mutual information, and/or a median of said mutual information.

5 In particular, contributions in the similarity measure for the mutual information of operation parameters that deviate significantly from the standard deviation or median may be suppressed with small weights. This may allow to reduce the effect of accidental synchronized behavior, which might otherwise be overemphasized by the normalization.

10 The operation parameters may correspond to any kind of values or readings that are available for the components, such as performance metrics, logs, or events. In some examples, a large variety of operation parameters can be employed to infer the network topology, regardless of the nature of these parameters.

15 For instance, in case the components in the network are computing devices, the operation parameters may correspond to readings of processors, to job statistics, to parameters quantifying a memory allocation, or to error statistics.

20 Providing said operation parameters may involve counting events or measurement data collected from said components in time slots.

Providing said plurality of operation parameters may also comprise counting relative frequencies, and/or assigning a probability distribution to said operation parameters.

25 In this sense, an operation parameter may be understood to comprise a collection of values, and an associated probability distribution.

In some examples, said normalized mutual information pertaining to a pair of components (i,j) and operation parameters (k,l) is given as

30

$$tr(M_i^k, M_j^l) := \frac{NMI(M_i^k, M_j^l)}{\sum_i NMI(M_i^k, M_j^l) \times \sum_j NMI(M_i^k, M_j^l)}, \quad (1)$$

wherein

$$NMI(M_i^k, M_j^l) := \frac{I(M_i^k, M_j^l)}{H(M_i^k, M_j^l)}, \quad (2)$$

5

wherein M_i^k and M_j^l denote said operation parameters and associated probability distributions for said components (i, j), $I(X, Y)$ denotes a mutual information of variables X and Y, and $H(X, Y)$ denotes a joint entropy of said variables X and Y.

10

In Equation (2), NMI denotes a mutual information I normalized by the joint entropy H. As can be taken from Equation (1), a normalized mutual information is then again normalized with two normalization factors, when each of the normalization factors is a sum of a normalized mutual information over components i and j, respectively.

15

In some examples, said similarity measure between a pair of components (i, j) is given as

$$S_{i,j} := \sum_k \sum_l w_{k,l} \operatorname{tr}(M_i^k, M_j^l), \quad (3)$$

wherein $w_{k,l}$ denote said weights.

20

The weights $w_{k,l}$ correspond to a probability distribution over operation parameters (k, l).

In particular, one may choose a diagonal distribution, i.e.,

25

$$w_{k,l} := w_k \delta_{k,l} \quad (4)$$

wherein w_k denotes a set of weights, and $\delta_{k,l} = 1$ if $k = l$, and $\delta_{k,l} = 0$ otherwise.

The inventors found that merely taking into account corresponding operation parameters

of different components (rather than cross terms) allows for a particularly efficient computation of the similarity measure, and still provides a reliable assessment of the network topology.

5 In some examples, determining whether said two components are topologically connected in said network comprises comparing said similarity measure against a predetermined threshold value.

10 Said threshold value may be a threshold value that depends on the pair of components (i, j) under consideration, or may alternatively be a universal threshold value.

15 The disclosure also relates to a system for inferring a topology of components in a network, said system comprising a receiving unit to receive at least one operation parameter for each of a plurality of components in a network, and a computing unit to compute a similarity measure between at least two of said components based on values of said operation parameters. Said computing unit determines whether said two components are topologically connected in said network based on a similarity measure, and said computing unit computes said similarity measure in terms of a normalized mutual information between said operation parameters pertaining to said two components.

20 Said system, in particular said computing unit, may be adapted to implement a method with some or all of the features described above.

25 The present disclosure further relates to a computer program or to a computer program product comprising computer-readable instructions, such that, when said instructions are run on a computing device connected to a system with some or all of the features described above, implement on said system a method with some or all of the features described above.

30 Implementing examples will now be described in greater detail with reference to the accompanying drawings.

Fig. 1 illustrates a system 10 for inferring a topology of a network according to an example of the disclosure. The system 10 comprises a receiving unit 18 to receive at least one operation parameter for each of the plurality of the components in a network, and a computing unit 20 to compute a similarity measure between at least two of said components based on values of said operation parameters. Said values of said operation parameters may be provided to the computing unit 20 by the receiving unit 18, as indicated by the arrow in Fig. 1.

Fig. 2 illustrates the system 10 according to an example in conjunction with a network 12 and a network monitoring unit 16.

The network 12 comprises a plurality of network components 14₁ to 14₆. For instance, the network 12 may be a computer network in an office environment, and the network components 14₁ to 14₆ may comprise desktop computers that are distributed in the computing network 12. However, the present disclosure is not so limited and can generally be applied to any kind of network and network components.

Fig. 2 shows a small network 12 with six network components 14₁ to 14₆. However, it will be understood that this is a mere example chosen for the ease of presentation, and in general the present disclosure may be applied to large networks with possibly hundreds of network components, or any number of network components. In fact, in some examples the method according to the present disclosure provides an efficient characterization of the network topology in large networks with many network components.

The components 14₁ to 14₆ may be interconnected in the network 12 in various ways, such as by wire over individual direct connections or via a communication bus, or wirelessly. The connections may indicate a flow of communication between the network components 14₁ to 14₆. For the case of presentation, the network links/connections are not shown in Fig. 2. In some applications, these connections may be unknown, and in one example it

may be the objective of the present disclosure to derive or determine the connections by monitoring the network components 14₁ to 14₆.

5 In another example, the physical network connections between the network components 14₁ to 14₆ may be known, but may be insufficient or inappropriate to characterize the correlations in the operations of the network components 14₁ to 14₆, or may not properly represent these correlations.

10 As can be taken from Fig. 2, the network components 14₁ to 14₆ provide surveillance data or monitoring data (indicated by dashed arrows) to a network monitoring unit 16. The surveillance data may comprise data that is routinely collected and maintained in many conventional IT management applications. In case the network components 14₁ to 14₆ are computing units, the surveillance data may relate to logs, events, or performance metrics, which may take numerical, categorical, or Boolean values. For instance, surveillance data
15 may relate to a processor load of a central processing unit (CPU) of the computing units 14₁ to 14₆, to memory allocation, job statistics, or error statistics. In some examples the present disclosure allows to process a large variety of surveillance data and employ them for determining a similarity measure, irrespective of the nature and meaning of this data. Each network component 14₁ to 14₆ may provide a single operation parameter, or a plu-
20 rality of operation parameters.

As can be taken from Fig. 2, the receiving unit 18 is adapted to receive a pre-determined selection of surveillance parameters from the network monitoring unit 16, and forwards
25 this data to the computing unit 20, which in turn computes a similarity measure between at least a pair of network components 14₁ to 14₆ based on the surveillance parameters.

Fig. 1 and 2 show the receiving unit 18 and the computing unit 20 of the system 10 as two separate units. However, in other examples these units could also be combined into a single unit, or could be integrated into a conventional network monitoring unit 16.

Methods of computing a similarity measure between at least two network components 14₁ to 14₆ based on the surveillance data according to an example of the present disclosure will now be described with reference to the flow diagrams of Fig. 3 and 4.

5 In S10, at least one operation parameter for each of a plurality of components in a network may be provided.

In S12, a similarity measure between at least two of said components based on values of said operation parameters may be computed. Said similarity measure may be computed in
10 terms of a normalized mutual information between said operation parameters pertaining to said two components.

In S14, it may be determined whether said two components are topologically connected in said network based on said similarity measure.

15 An example of a method of computing the similarity measure based on surveillance data received from the network components 14₁ to 14₆ of the network 12 will now be described in greater detail with reference to the flow diagram of Fig. 4.

20 In S20, the computing unit 20 may collect suitable surveillance parameters provided from the network components 14₁ to 14₆ via the network monitoring unit 16. As described above, the surveillance data may comprise a large variety of values and data formats, such as numeric, categorical or Boolean values.

25 In S22, in order to reduce this inherent complexity, the surveillance data may be converted to monitoring data, such as by counting events in given pre-determined time slots. This pre-processing results in a plurality of operation parameters M_i^k , wherein $i = 1, \dots, N$ denotes the network components (in this case $N = 6$), and $k = 1, \dots, P$ denotes the various parameters that are monitored.

30

The pre-processing S22 may also involve assigning probability distributions to each of the operation parameters M_i^k , such as based on a priori information or based on relative frequencies obtained from counting said events in said time slots.

- 5 Based on the monitoring data in S24, the computing unit 20 computes the similarity measure S_{ij} between pairs of components $i, j = 1, \dots, N$ based on a weighted normalized mutual information, as will be described in further detail below.

10 An underlying idea of the method is that when systems and components are connected with one another and interacting with one another in a network 12, these interactions will leave identifiable traces in the collected surveillance data. Similarly to a disturbance analysis, if network components 14₁ to 14₆ are connected in a topology, then a disturbance occurring in one of them will often cause a related disturbance in a connected component. In some examples, a weighted normalized mutual information is suitable to reveal these
15 kind of connections and similarities. On the one hand, mutual information is agnostic to the type of monitoring data, and on the other hand is superior to correlation techniques that may assume an underlying linear monitor dependence.

20 Assuming two random variables X and Y with a joint probability distribution $p(x,y)$ and marginal probability distributions $p(x)$ and $p(y)$, the mutual information $I(x,y)$ may be defined as

$$I(X, Y) = \sum_x \sum_y p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (5)$$

25 Alternatively, the mutual information may be understood to denote a relative entropy between the joint distribution $p(x,y)$ and the product distribution $p(x) \times p(y)$.

The mutual information may be expressed in terms of the entropy

$$H(X) = - \sum_x p(x) \log p(x) \quad (6)$$

as follows:

$$I(X, Y) = H(X) + H(Y) - H(X, Y) \quad (7)$$

5 The mutual information $I(X, Y)$ is non-negative, but is not a priori bounded. Good results can be obtained by normalizing mutual information to the unit interval by dividing by the joint entropy,

$$NMI(X, Y) := \frac{I(X, Y)}{H(X, Y)} \quad (8)$$

10 Signals that may have many different values, such as random noise contributions or conversely monotone signals, typically have a high entropy, and their normalized mutual information NMI with any other signal is approximately the entropy of that other signal. Hence, when a given monitor has both varied and non-varied signals, the varied normalized mutual information dominates. These issues can be overcome, and good results can
15 be obtained by further normalizing NMI by dividing by the sum of normalized mutual information over the respective components i and j , respectively.

Hence, for operation parameters M_i^k , where k denotes the monitors and i denotes the components, we define

$$20 \quad tr(M_i^k, M_j^l) := \frac{NMI(M_i^k, M_j^l)}{\sum_i NMI(M_i^k, M_j^l) \times \sum_j NMI(M_i^k, M_j^l)} \quad (9)$$

Based on the normalized mutual information given in Equation (9), a similarity measure that quantifies the similarity or connection or resemblance between the pair of components
25 i, j may be given as

$$S_{i,j} := \sum_k \sum_l w_{k,l} tr(M_i^k, M_j^l) \quad (10)$$

wherein $w_{k,l}$ denote the probability distribution of weights.

The weights may, for example, be chosen to suppress pairs of operation parameters with accidental synchronized behavior.

5

Equation (10) provides a similarity measure based on all combinations of pairs of operation parameters M_i^k, M_j^l , $k, l = 1, \dots, P$. However, a reliable assessment of the network topology can be achieved by comparing corresponding or identical operation parameters for different network components, thereby achieving a significant reduction of the complexity.

10

Hence, for example,

$$S_{i,j} = \sum_k w_k \operatorname{tr}(M_i^k, M_j^k) \quad (11)$$

15

wherein

$$\operatorname{tr}(M_i^k, M_j^k) = \frac{NMI(M_i^k, M_j^k)}{\sum_i NMI(M_i^k, M_j^k) \times \sum_j NMI(M_i^k, M_j^k)} \quad (12)$$

20

As described above, the weights w_k can be chosen with a view to suppress accidental synchronized behavior that may be observed in some operation parameters, for instance by taking into account the standard deviation S_k of all terms $\operatorname{tr}(M_i^k, M_j^k)$, computing the median μ of all standard deviations S_k , and defining a parabola through $(0,0)$ and through $(\mu,1)$ as a raw weight factor:

25

$$v_k := \max \{0, S_k \cdot (2\mu - S_k)\} \quad (13)$$

The weight factors w_k may then be obtained from normalization,

14

$$w_k := \frac{v_k}{\sum_k v_k} \quad (14)$$

In case a single operation parameter k is provided per component, i.e., $P = 1$, there is a single weight $w_k = 1$.

5

Once the similarity measure S_{ij} has been computed, the computing unit 20 may determine whether two given components i, j are topologically connected in S16 based on a comparison with a pre-determined threshold.

10

For instance, the components i, j may be determined to be connected if and only if $S_{ij} \geq \theta$, for some given threshold θ .

The results may be output in S28, such as in the form of a list of connections or a graph of connections.

15

Resulting topological connections that may be the outcome of the method according to the present disclosure are illustrated in Fig. 5.

20

Fig. 5 largely corresponds to Fig. 2, but in addition shows respective connections 22₁ to 22₅ (represented by solid lines) between the network components 14₁ to 14₆ whenever the computing unit 20 determines that the corresponding similarity measure between these components is no smaller than the pre-determined threshold θ .

25

In the example of Fig. 5, one can identify two sub-groups of components, a first sub-group comprising the network components 14₁, 14₂, 14₃ and 14₆, and a second sub-group comprising the network components 14₄ and 14₅. None of the network components in the first sub-group is connected to any of the network components in the second sub-group, and vice-versa.

Among the network components in the first sub-group, the network components 14₁, 14₂, and 14₆ are all interconnected, whereas the remaining network component 14₃ is connected (via a connection 22₄) to the network component 14₂ but not to other network components.

5

The method according to the disclosure hence may provide helpful and valuable insight on the topological connections of the given computing system based on monitoring data.

10

The computing unit 20 may be implemented on a desktop or laptop computer, but may also be integrated into a conventional network monitoring unit 16.

15

The techniques described above were applied on data from 64 different performance metrics over a test period of eight servers and five weeks. The eight servers worked in two clusters of four, wherein each cluster was split between two different data centers. In the test network, different stages of a job were run on different servers, and all stages used a meta data service. To discover the topology, the data was split into five week-long sets, and it was found that all the clusters and the two sub-clusters could be identified correctly. The number of inter-cluster links that were missed, i.e. not identified, as well as the number of erroneous links, was negligible.

20

The description of the examples and the Figures are merely illustrative, and should not be understood to imply any limitation. The scope of the present disclosure is to be determined solely from the appended claims.

5

Reference Signs

	10	system for inferring a network topology
	12	network
10	14 ₁ – 14 ₆	network components of network 12
	16	network monitoring unit
	18	receiving unit
	20	computing unit
	22 ₁ – 22 ₅	topological connections between network components 14 ₁ – 14 ₆

15

5

Claims

1. A method for inferring a topology of components in a network, comprising:
providing at least one operation parameter for each of a plurality of components
10 in a network;
computing a similarity measure between at least two of said components based
on values of said operation parameters; and
determining whether said two components are topologically connected in said
network based on said similarity measure;
15 wherein said similarity measure is computed in terms of a normalized mutual in-
formation between said operation parameters pertaining to said two components.
2. The method according to claim 1, wherein said normalized mutual information is ob-
tained by normalizing with a joint entropy of said operation parameters pertaining to
20 said two components.
3. The method according to claim 1, wherein said mutual information is obtained by
normalizing over a sum of said mutual information, said sum over said components.
- 25 4. The method according to claim 1, wherein said normalized mutual information is a
weighted normalized mutual information.
5. The method according to claim 4, wherein weights of said weighted normalized mu-
tual information are given in terms of a standard deviation or a median of said mu-
30 tual information.

6. The method according to claim 1, wherein providing said at least one operation parameter comprises assigning a probability distribution to said operation parameter.
7. The method according to claim 1, wherein said mutual information pertaining to a pair of components (i,j) and operation parameters (k,l) is given as

$$tr(M_i^k, M_j^l) := \frac{NMI(M_i^k, M_j^l)}{\sum_l NMI(M_i^k, M_j^l) \times \sum_j NMI(M_i^k, M_j^l)},$$

wherein

$$NMI(M_i^k, M_j^l) := \frac{I(M_i^k, M_j^l)}{H(M_i^k, M_j^l)},$$

wherein M_i^k, M_j^l denote said operation parameters and associated probability distributions for said components (i,j), $I(X,Y)$ denotes a mutual information of variables X and Y, and $H(X,Y)$ denotes a joint entropy of said variables X and Y.

8. The method according to claim 7, wherein said similarity measure between a pair of components (i,j) is given as

$$S_{i,j} := \sum_k \sum_l w_{k,l} tr(M_i^k, M_j^l),$$

wherein $w_{k,l}$ denote said weights.

9. The method according to claim 8, wherein

$$w_{k,l} := w_k \delta_{k,l},$$

wherein w_k denotes a set of weights, and $\delta_{k,l} = 1$ if $k = l$, and $\delta_{k,l} = 0$ otherwise.

10. The method according to claim 1, wherein determining whether said two components are topologically connected in said network comprises comparing said similarity measure against a predetermined threshold value.
- 5 11. The method according to claim 1, wherein said operation parameters are parameters quantifying a status or operation of said components in said network.
12. The method according to claim 1, wherein said components are computing devices, and said network is a communication network linking said computing devices.
- 10 13. A system for inferring a topology of components in a network, comprising:
a receiving unit to receive at least one operation parameter for each of a plurality of components in a network; and
a computing unit to compute a similarity measure between at least two of said
15 components based on values of said operation parameters;
wherein said computing unit determines whether said two components are topologically connected in said network by comparing said similarity measure with a pre-determined threshold value; and
wherein said computing unit computes said similarity measure in terms of a
20 weighted and normalized mutual information between said operation parameters pertaining to said two components.
14. The system according to claim 13, wherein said computing unit obtains said normalized mutual information by normalizing with a joint entropy of said operation parameters
25 pertaining to said two components.
15. A non-transitory computer-readable storage medium, comprising computer-readable instructions to implement a method comprising:
receiving at least one operation parameter for each of a plurality of components
30 in a network;

computing a similarity measure between at least two of said components based on values of said operation parameters; and
determining whether said two components are connected in said network by comparing said similarity measure with a threshold value;
5 wherein said similarity measure is computed in terms of a normalized mutual information between said operation parameters pertaining to said two components.

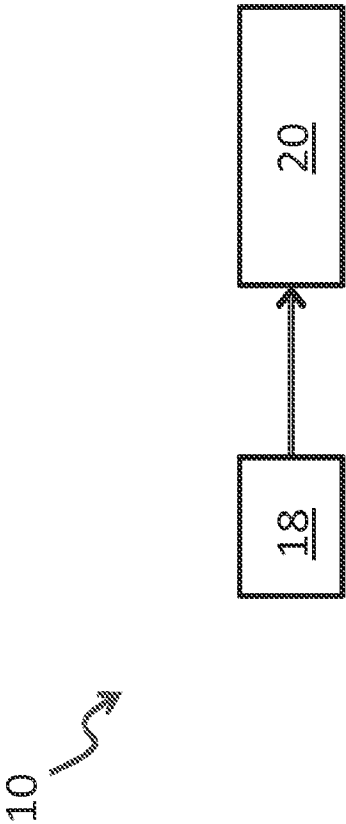


Fig. 1

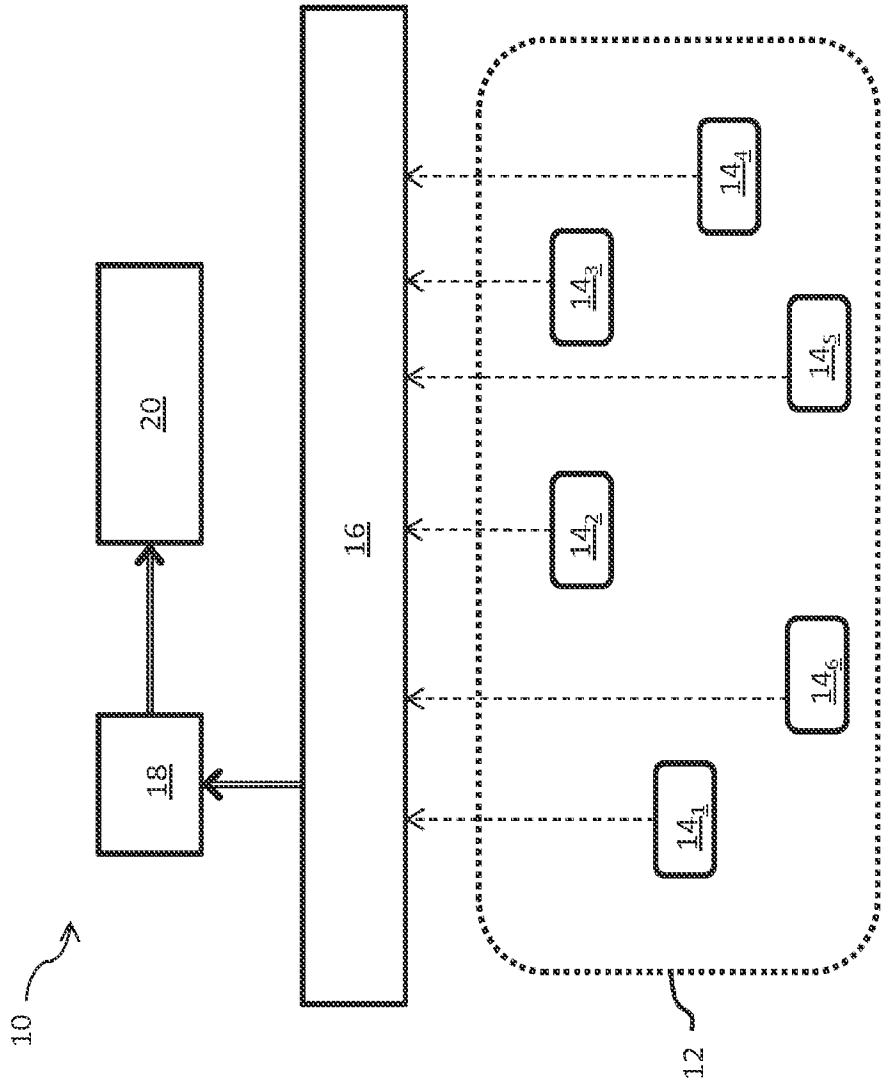


Fig. 2

3/5

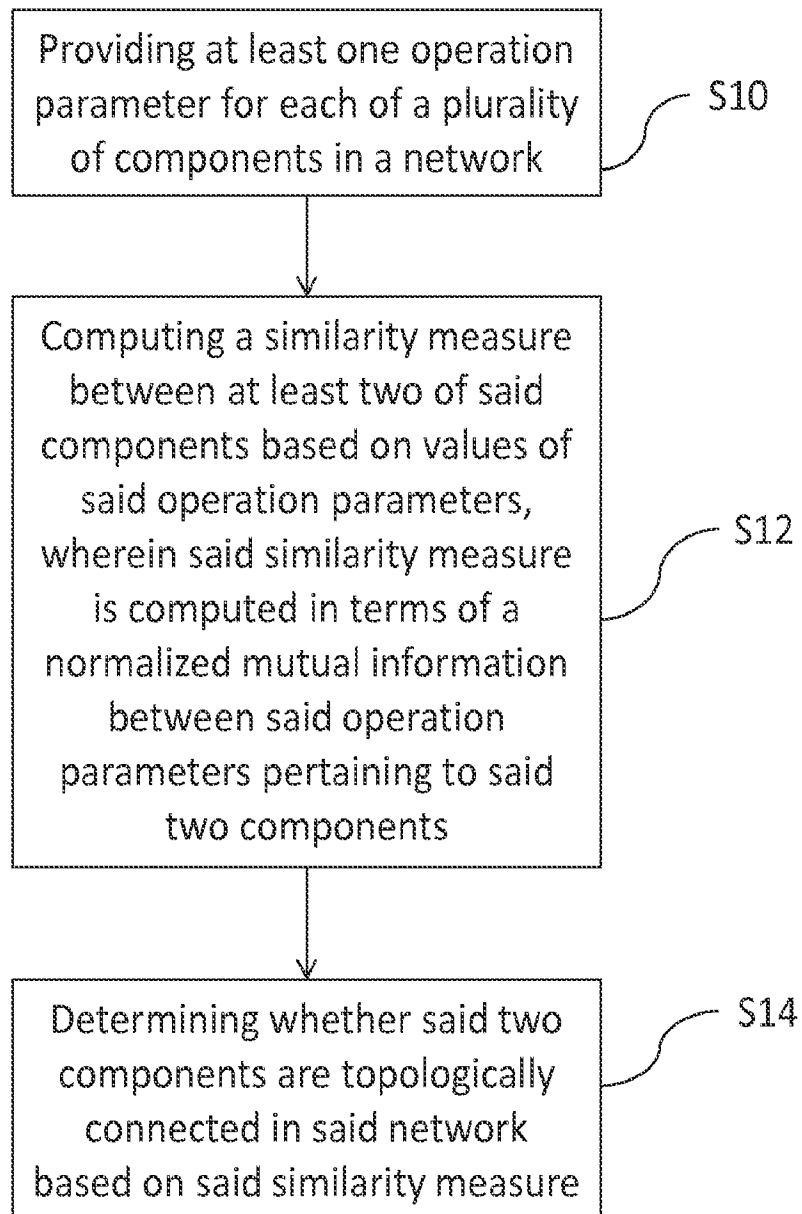


Fig. 3

4/5

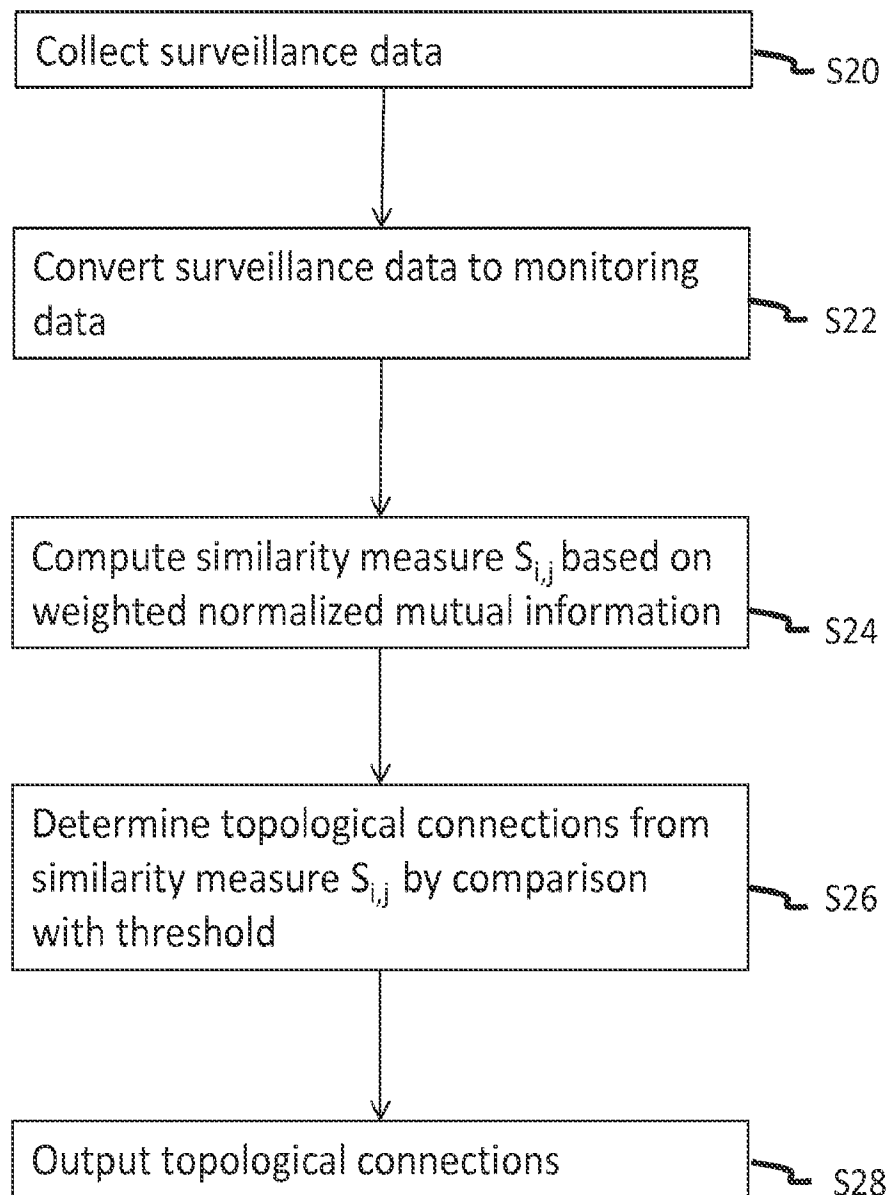


Fig. 4

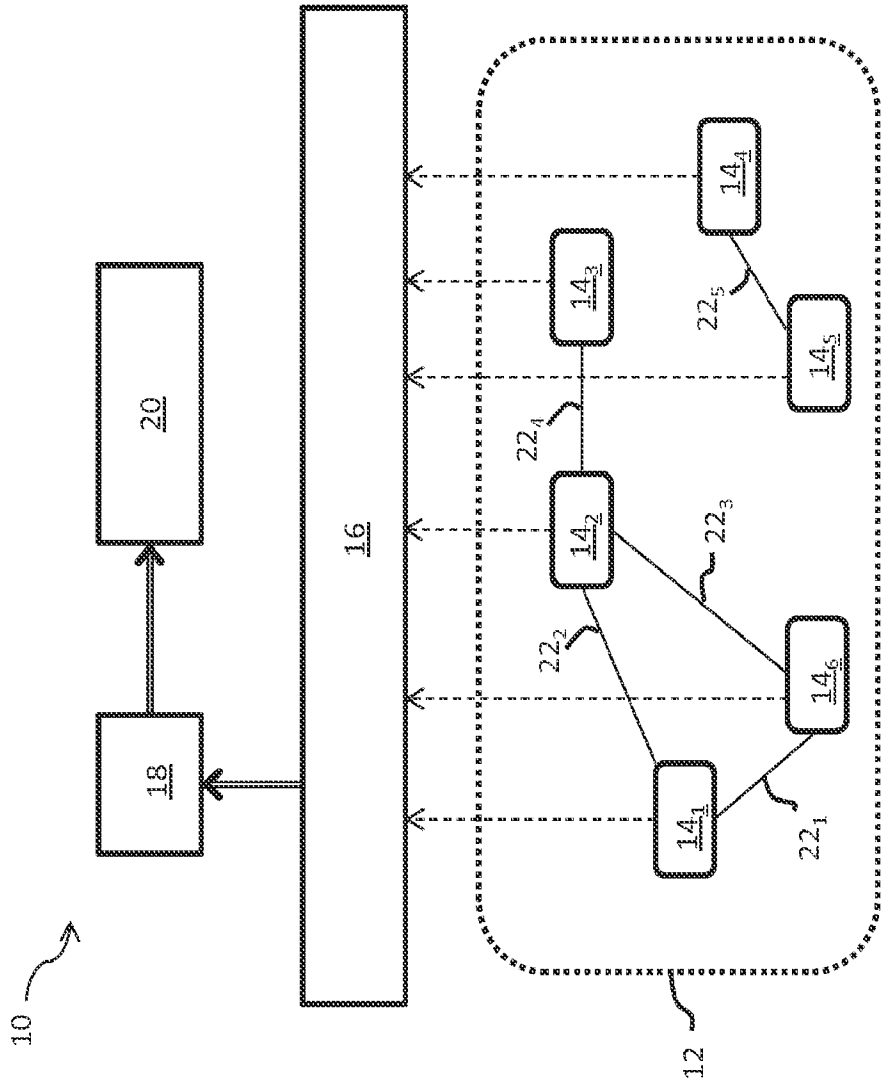


Fig. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2015/023491**A. CLASSIFICATION OF SUBJECT MATTER****H04L 12/26(2006.01)i, H04L 12/24(2006.01)i**

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)

H04L 12/26; H04J 1/16; G06F 15/173; G06F 15/00; G06K 9/32; G06T 15/50; H04L 12/24

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

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) & Keywords: inferring, topology, network, component, operation, parameter, similarity, measure, normalized, mutual, information

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2013-0103822 A1 (LAWRENCE WOLCOTT et al.) 25 April 2013 See paragraphs [0003], [0027], [0037]; and figure 1.	1-15
A	US 2003-0097693 A1 (MINEO MOROHASHI et al.) 22 May 2003 See paragraphs [0086]-[0118]; and figures 4-7.	1-15
A	US 7317693 B1 (MARTIN ROESCH et al.) 08 January 2008 See column 17, line 4 - column 20, line 12; and figures 11-14.	1-15
A	US 2010-0027911 A1 (ALBAN LEFEBVRE et al.) 04 February 2010 See paragraphs [0031]-[0042]; and figures 1-3.	1-15
A	US 2013-0077891 A1 (SOMKID NIMNUAL et al.) 28 March 2013 See paragraphs [0015]-[0103]; and figures 1-2.	1-15



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

23 December 2015 (23.12.2015)

Date of mailing of the international search report

02 March 2016 (02.03.2016)

Name and mailing address of the ISA/KR

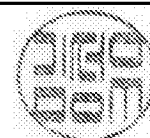
International Application Division
Korean Intellectual Property Office
189 Cheongsu-ro, Seo-gu, Daejeon Metropolitan City, 35208,
Republic of Korea

Facsimile No. +82-42-472-7140

Authorized officer

KIM, Seong Woo

Telephone No. +82-42-481-3348



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2015/023491

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2013-0103822 A1	25/04/2013	CA 2793261 A1 EP 2584737 A2 US 09021086 B2	21/04/2013 24/04/2013 28/04/2015
US 2003-0097693 A1	22/05/2003	EP 1248231 A1 EP 1248231 A4 JP 2001-188768 A WO 2001-048690 A1	09/10/2002 30/08/2006 10/07/2001 05/07/2001
US 7317693 B1	08/01/2008	US 07467205 B1	16/12/2008
US 2010-0027911 A1	04/02/2010	US 08731334 B2	20/05/2014
US 2013-0077891 A1	28/03/2013	EP 2761594 A1 US 08611692 B2 WO 2013-048846 A1	06/08/2014 17/12/2013 04/04/2013