



(51) International Patent Classification:  
*H04L 12/26* (2006.01)

(21) International Application Number:  
PCT/US2015/043020

(22) International Filing Date:  
30 July 2015 (30.07.2015)

(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: **SHARMA, Puneet**; 1501 Page Mill Rd., Palo Alto, California 94304-1100 (US). **MALBOUBI, Mehdi**; 1501 Page Mill Rd., Palo Alto, California 94304-1100 (US).

(74) Agents: **FERGUSON, Chris** 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).

**Declarations under Rule 4.17:**

- *as to the identity of the inventor (Rule 4.17(i))*
- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

**Published:**

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

(54) Title: TRAFFIC ON DEFINED NETWORK HAVING AGGREGATIONS OF NETWORK DEVICES

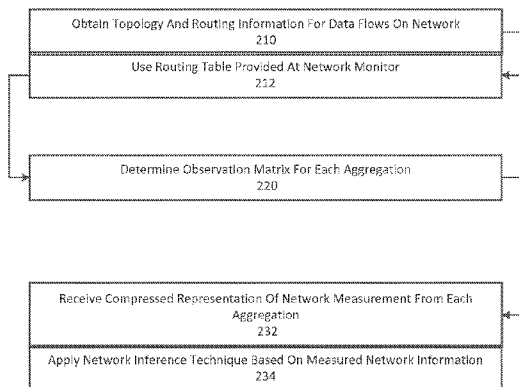


**FIG. 2**

(57) Abstract: An observation matrix is determined for each aggregation of network devices (e.g., rack with servers) using topology and routing information. The corresponding observation matrix can be communicated to each of multiple aggregations of network devices in order to obtain, from each aggregation, a compressed representation of the network traffic through individual network devices of that aggregation.



## **TRAFFIC ON DEFINED NETWORK HAVING AGGREGATIONS OF NETWORK DEVICES**

### **BACKGROUND**

**[0001]** For many types of networks, network flow measurements are used for a variety of applications, including managing utilization of the network and optimizing traffic throughput. In many applications, network measurements are primarily obtained from Simple Network Management Protocol ("SNMP") resources such as link counters. Many data centers utilize dedicated network monitoring links between servers and or racks of servers and network monitor system in order to communicate SNMP link load data, representing direct network measurements at an endpoint of the link (e.g., server, rack). However, the volume of monitoring traffic resulting from the number of data flows, as well as end-to-end network paths such flows utilize, typically far exceeds the bandwidth allocated for monitoring in the network.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

**[0002]** FIG. 1 illustrates an example system for estimating traffic on a defined network in which multiple aggregations of network devices are provided.

**[0003]** FIG. 2 determines an example method for determining optimized compressed network measurements from each aggregation of a defined network.

**[0004]** FIG. 3 illustrates an example of a data center for implementing one or more examples.

### **DETAILED DESCRIPTION**

**[0005]** Examples described herein estimate traffic in a defined network using compressed network measurements which are reflective of data flows through a given aggregation of network devices. According to examples described, an observation matrix is determined for each aggregation of network devices (e.g., servers or rack with servers) using topology and routing information. The corresponding observation matrix can be communicated to each of multiple aggregations of network devices in order to obtain, from each aggregation, a compressed representation of the

network traffic through individual networked devices such as servers of that aggregation.

**[0006]** As used herein, the term "optimal" or variations thereof (e.g., "optimized") means an outcome configured to induce or augment an objective or characteristic at the expense of another characteristic or objective.

**[0007]** Additionally, the term "aggregation" (and variants thereof) in context of network devices is intended to including a physical and/or logical aggregation (e.g., defined group) of network devices. Examples of aggregations include physically clustered network devices (e.g., see racks as described with FIG. 1 and 3), as well as logical groupings of physically clustered devices (e.g., two or more racks) and/or multiple devices which are logically defined and/or physically interconnected into individual aggregations.

**[0008]** In estimation theory in general,  $Y=AX$  represents a fundamental relationship in which  $Y$  is an actual measurement,  $X$  is that is to be estimated, and  $A$  is an observation matrix which maps the actual and estimated measurements. With regard to network traffic monitoring, examples recognize that the actual measurements are made by servers or groups of servers, but such measurements typically cannot be shared (e.g., with a network controller) because of bandwidth constraints. Thus, network elements often contain local resources for making actual measurements, but the local measurements cannot be communicated directly to a central location where all flows within the network are known, as such communication would require far too much bandwidth than allocated for monitoring purpose. Rather, examples such as described provide for the communication of information that represents a compressed and optimal profile of the traffic being exchanged through individual devices of the node.

**[0009]** According to examples, a network is identified in terms of aggregations of network devices (e.g., edge devices), from which data flows of the network originate. Each aggregation of edge devices is provided an observation matrix which (i) accounts for routing and traffic flow across the entire network, (ii) is specific to the particular aggregation of devices, and (iii) when applied to actual measurements of each edge device of the

aggregation, generate an aggregate set of measurements which are representative of the traffic exchanged through the aggregation.

**[0010]** Examples described herein provide the methods, techniques, and actions performed by a computing device that are performed programmatically, or as a computer-implemented method. Examples may be implemented as hardware, or a combination of hardware (e.g., a processor(s)) and executable instructions (e.g., stored on a machine-readable storage medium). These instructions can be stored in one or more memory resources of the computing device. A programmatically performed step may or may not be automatic.

**[0011]** Examples described herein can be implemented using programmatic modules or components. The programmatic modules or components may be any combination of hardware (e.g., processor(s)) and programming to implement the functionalities of the modules or components described herein. In examples described herein, such combinations of hardware and programming may be implemented in a number of different ways. For example, the programming for the components may be processor executable instructions stored on at least one non-transitory machine-readable storage medium and the hardware for the components may include at least one processing resource to execute those instructions. In such examples, the at least one machine-readable storage medium may store instructions that, when executed by the at least one processing resource, implement the components.

**[0012]** Some examples described herein can generally involve the use of computing devices, including processing and memory resources. For example, examples described herein may be implemented, in whole or in part, on computing devices such as desktop computers, cellular or smart phones, personal digital assistants (PDAs), laptop computers, printers, digital picture frames, and tablet devices. Memory, processing, and network resources may all be used in connection with the establishment, use, or performance of any example described herein (including with the performance of any method or with the implementation of any system).

**[0013]** Furthermore, examples described herein may be implemented through the use of instructions that are executable by one or more processors. These instructions may be carried on a computer-readable

medium. Machines shown or described with figures below provide examples of processing resources and computer-readable mediums on which instructions for implementing examples described herein can be carried and/or executed. In particular, the numerous machines shown with examples include processor(s) and various forms of memory for holding data and instructions. Examples of computer-readable mediums include permanent memory storage devices, such as hard drives on personal computers or servers. Other examples of computer storage mediums include portable storage units, such as CD or DVD units, flash memory (such as carried on smart phones, multifunctional devices or tablets), and magnetic memory. Computers, terminals, network enabled devices (e.g., mobile devices, such as cell phones) are all examples of machines and devices that utilize processors, memory, and instructions stored on computer-readable mediums. Additionally, examples may be implemented in the form of computer-programs, or a computer usable carrier medium capable of carrying such a program.

**[0014]** SYSTEM DESCRIPTION

**[0015]** FIG. 1 illustrates an example system for estimating traffic on a defined network in which multiple aggregations of network devices are provided. As shown, a defined network 10 includes a network monitoring system 110 which communicates with multiple aggregations 120 of network devices 122 using network channels of the defined network 10. Among other functions, the network monitoring system 110 can estimate traffic flow throughout the defined network 10, using actual measurements made by sensors on edge devices and other traffic measuring resources that are distributed on the network 10. As described in greater detail, the network monitoring system 110 can utilize routing information and network topology to determine parameters for enabling network traffic measurements at each aggregation to be compressed in a manner that is meaningful and representative of the network traffic incident on or passing through all of the network devices 122 of the aggregation 120, but the compressed monitoring data has significantly less bandwidth requirements than direct network measurement data which can be obtained over the same set of data channels.

**[0016]** In an example of FIG. 1, defined network 10 can represent, for example, a data center network in which the aggregations of network devices include racks of servers. While some examples are described in the context of data center networks, examples as described can be applicable other kinds of networks in which groups or clusters of network devices are utilized, including other kinds of software defined networks and Ethernet type networks.

**[0017]** Depending on the type of network, each aggregation 120 includes a cluster, group or stack of networked or edge devices 122 and a set of aggregation components 124. The network devices 122 can be physically and/or logically aggregated to share a set of physical and/or network resources for operating on the network 10. By way of example, each aggregation 120 can correspond to a group of servers which share cabling or physical channels for communicating on the network 10, as well as data switches for enabling individual servers of the aggregation to access the network channels of the defined network 10 to communicate with other servers on other aggregations. The aggregations 120 can each also include one or more aggregation resources 124, which can include switches which are shared amongst the network devices 122 to provide access to data channels of the network 10. The aggregation resources 124 can also include local resources to make network measurements at each aggregation 120. Numerous other kinds of resources can also be shared at each aggregation 120, including ports, physical housing structures, and logical resources. Each aggregation 120 can include a set of links 121 for communicating network measurements to the network monitoring system 110.

**[0018]** The network monitoring system 110 includes functionality for predicting and/or analyzing various aspects of the defined network 10. In an example of FIG. 1, network monitoring system 110 includes a network inference engine 112, a statistical component for performing path analysis 114, and a compression parameter determination component 116. The network inference engine 112 can develop, for example, a network inference model or data structure to analyze or predict traffic patterns and characterizations of data flows, as well as other aspects of the defined network 10 (e.g., network traffic on data channels, network ingress/egress of aggregations 120, and/or network devices 122). For example, the network

inference engine 112 can be used to develop network traffic matrices or models which can in turn, be used to identify information which affects the health, performance and/or efficiency of the defined network 10 as a whole. For example, an output of the network inference engine 112 can be used to determine (i) a cause of a network anomaly, (ii) predict future network anomalies and traffic patterns, (iii) identify "heavy hitters" which utilize a disproportionate amount of bandwidth on the defined network 10, and/or (iv) determine traffic volume and throughput for various data channels.

**[0019]** The relevance or accuracy of the network inference engine 112 can be based on the quantity, depth and accuracy of network measurements which the network inference engine 112 uses as input. In many applications, network measurements which are made at each aggregation 120 and communicated to the network monitoring system 110 via links 121. However, in many kinds of defined networks 10, there is insufficient bandwidth for communicating full sets of network measurements (or other network measurement sources) which would reflect traffic or state of individual data paths within the defined network 10. The defined network 10 may typically use the links 121, which are dedicated for network measurement communications, but the links 121 are far less in quantity and bandwidth than data paths of the defined network 10. For example, in data centers, network traffic measurement data can be obtained through standard Simple Network Management Protocol (SNMP) which is resident on individual racks of servers. While such network traffic measurement data may be readily available locally at the rack, the number of data flows which have an end point at the rack are exponentially greater than the number of servers present within each rack. More generally, given a first aggregation 120 with  $n$  network devices, and a second aggregation 120 with  $m$  network devices, the total number of data flows just between the first and second aggregations is  $n \times m$ . Thus, as shown by an example of FIG. 1, when links 121 between network measurement sources and the network monitor system 110 are provided, the bandwidth allocated for monitoring on these links is limited and generally is not enough to carry monitoring data for the number of data flows, or paths for such data flows. As a result, locally obtained network measurement data, such as show by directly measured data ("DMD") 123 (such as SNMP link data) from a particular aggregation 120 at a moment has

a limited range of use, as the link bandwidth constrains the amount of direct measurement (e.g., SNMP link load data) which the network monitoring system 110 can ultimately receive, as compared to the data flows and end-to-end network paths.

**[0020]** Accordingly, the network monitoring system 110 can operate to obtain a compressed or reduced, but representative, form of network measurements from the individual aggregations 120, for use with the network inference engine 112. Numerous techniques exist for obtaining compressed network measurements as a mechanism to augment or supplement directly obtained network measurement data 123. By way of example, the directly obtained network measurement data 123 can correspond to SNMP link load data, as communicated by network measuring resources (e.g. aggregate resources 124) of each aggregation 120. In order to obtain supplementary network measurement data, examples determine and utilize a set (or matrix) of compression parameters ("compression parameter data set 115") to optimize a compression of network measurements ("compressed network measurements 125" or "CNM 125") made at each aggregation 120 to accurately reflect the network measurements which affect all of the data flows of the respective aggregation 120. Each set of compression parameter data set 115 can also be determined to be specific to each aggregation 120, so that the resulting compression parameter data set 115 reflects the measured network data at that aggregation 120. The optimization provided through the use of compression parameter data set 115 can be provided at expense of bandwidth. In some implementations, the compressed network measurements 125 can supplement SNMP link loads, provided over the links between aggregations and the network monitor system.

**[0021]** The optimization provided through the use of compression parameter data set 115 can be provided at expense of bandwidth. In this respect, examples such as provided with FIG. 1 are distinct from other approaches, some of which, for example, seek to obtain supplemental network measurement through sampling in order to determine more significant data flows. In contrast to such approaches, an example of FIG 1 generates the compressed network measurement 125 to be reflective of all of the data flows of the aggregation 120.

**[0022]** In an example of FIG. 1, the network monitoring system 110 implements operations to optimize the relevance and accuracy of the compression operator implemented on the network measurements which are made at a given location (e.g., aggregation 120), with available bandwidth serving as the cost for optimizing the network measurements. Moreover, in some implementations, the compressed network measurements are representative of all the data flows of the corresponding aggregation 120. Still further, the compressed network measurements 125 can also take into account multipath routing of data flows.

**[0023]** In an example of FIG. 1, the network monitoring system 110 implements processes for determining the optimal observation matrix (i.e. compression operator or compression parameter data set) 115 based in part on topology information 117 and the routing information (as determined from the routing matrix 119). Multiple compression parameter data sets 115 can be determined, with each compression parameter data set 115 being determined for optimization of network measurements made at a corresponding aggregation 120. As described with other examples, the compression parameter data set 115 can be determined as an observation matrix which provides coefficients to each aggregation 120. The aggregation resource 124 can include network measurement functionality to receive and apply the compression parameter data set 115 for that aggregation 120. The values of the compression parameter data set 115 enable each aggregation 122 to return a set of network measurements 125 which are compressed by coefficients of the compression parameter data set 115, but the compressed set of network measurements 125 are also highly representative of the traffic profile or state of the corresponding aggregation 120. As further described, examples provide that the calculations used to obtain the compression parameter data sets 115 and corresponding sets of compressed network measurements 125 are relatively light computationally, while the representation of the individual aggregations 120 provided by the corresponding compressed sets of network measurements 125 have characteristics of being both relatively fine in granularity and representative of network traffic characteristics of all of the data flows that are formed through that aggregation 120.

**[0024]** In some implementations, network monitoring system 110 generates compression parameter data set 115 which account for multipath routing of individual data flows. In order to count for multipath routing, network monitoring system 110 can include components and functionality such as shown with statistical path analysis logic 114. The statistical path analysis logic 114 can determine statistical expectation of data traffic from individual flows on separate paths of the defined network 10. The compression parameter determination logic 116 can use the statistical expectations in determining the compression parameter data sets 115 of each aggregation 120. In determining the statistical expectations the statistical path analysis logic 114 can utilize the probability distribution function of the routing matrix or different realization of the routing matrix can be used.

**[0025]** Among other benefits, examples such as described provide a set of linear combination of per-flow measurements which are supplementary to directly measured network information (such as provided by SNMP link loads). Accordingly, an example of FIG. 1 can utilize compressed network measurements 125 to increase the estimation accuracy of network data flows, which are recognized as being highly variable over time/space, using compressed sensing techniques. As described with some other examples, the compression parameter data sets 115 for each aggregation 120 can be designed as observation matrices, which can be distributed to the aggregations 120 in an offline manner. Examples as described can improve the estimation accuracy resulting from the compressed network measurements 125, while also achieving high Compression Ratio (CR) under limited link bandwidths.

**[0026]** METHODOLOGY

**[0027]** FIG. 2 determines an example method for determining optimized compressed network measurements from each aggregation of a defined network. In describing an example of FIG. 2, reference may be made to elements of FIG. 1 for purpose of illustrating suitable components for performing a step or sub-step being described.

**[0028]** With reference to an example of FIG. 2, network topology and routing information can be obtained for the defined network 10 (210). In

some examples, an example method of FIG. 2 can be performed at network monitor system 110 or similar network element, where a routing table for the network is located (212).

**[0029]** With the routing and topology information, an observation matrix is determined for each aggregation 120 (220). Each observation matrix may include or correspond to a compression parameter data set 115 for a particular aggregation 120. In many networks, including data centers, an assumption can be made that the network exhibits sparse behavior, with large fluctuations of data flows present. Given sparseness and fluctuations, a traffic matrix optimization problem can be posed as:

$$\hat{X} = \min_X \|Y^O - A^O X\|_2 + \lambda \|X\|_1 \text{ where}$$

$$Y = HX$$

$$Y^O = \begin{bmatrix} Y \\ Y_1^O \\ Y_2^O \\ \vdots \\ Y_L^O \end{bmatrix}, A^O = \begin{bmatrix} H \\ A_1^O \\ A_2^O \\ \vdots \\ A_L^O \end{bmatrix}, \begin{bmatrix} Y_1^O \\ Y_2^O \\ \vdots \\ Y_L^O \end{bmatrix} = \begin{bmatrix} A_1^O X \\ A_2^O X \\ \vdots \\ A_L^O X \end{bmatrix}$$

**[0030]** In the equations provided,  $X$  is a vector presentation of the traffic matrix which is unknown and must be estimated,  $A^O$  is compression operator, also called the observation matrix,  $Y^O$  is the compressed network measurements (or linear combinations of unknowns), and  $H$  is the routing matrix. For any given aggregation 120, the set of compression operator 115 can be termed as an Optimal Local Observation Matrices (OLOM), which can be determined and applied independently of the OLOM of other aggregations.

**[0031]** A smaller traffic matrix estimation problem can be formulated by considering local observation matrix as:

$$\begin{bmatrix} H \\ A_i^O \end{bmatrix} \text{ for } i = 1, \dots, L$$

based on the assumption that local observation matrices  $A_i^O$  can be designed, independently. In designing each OLOM for each grouping 120,  $L$  represents the number of racks (or aggregations),  $r$  a sub-set of link-loads

(or rows of  $H$ ), can be assumed to have contributions from flows observed by  $A^o_i$ . In some examples, the local observation matrices can be pre-distributed among racks/servers in an offline manner. The optimal observation matrix  $A^o_i$  is thus designed for each aggregation 120, and aggregation-specific observation matrices are distributed among servers in each aggregation 120, where measurement compression/aggregation modules are available and can be utilized to generate a desirable set of linear combinations of  $X$  (shown with  $Y^o$ ). These aggregated measurements are then communicated to the network monitoring system where actual traffic matrix estimation is performed using a network inference technique of choice.

**[0032]** Assuming the compressed sensing network inference, a good local observation matrix can be designed by minimizing the sum of all off-diagonal elements of the corresponding Gram matrix ( $G$ ) defined as below where  $T$  denotes the Transpose operation.

$$G = A^{oT} A^o \text{ where } A^o := \begin{bmatrix} H \\ A^o_i \end{bmatrix} \text{ for } i = 1, \dots, L$$

**[0033]** In some examples, the routing matrix  $H$  includes coefficients or parametric values to reflect multi-path routing of individual data flows, and the coefficients can be used with the observation matrix. Specifically, to consider the randomness of some entries of  $H$  (due to multi-path routing in many data center network), a statistical expectation over  $H$  can be considered in the objective function. Note, for notation simplicity,  $A^o_i$  is considered as  $A$ :

$$\hat{A} = \underset{A}{\operatorname{argmin}} F = \underset{A}{\operatorname{argmin}} \|A^{oT} A^o - I\|_F^2 = \underset{A}{\operatorname{argmin}} E_H \left\{ \left\| \begin{bmatrix} H \\ A \end{bmatrix}^T \begin{bmatrix} H \\ A \end{bmatrix} - I \right\|_F^2 \right\} = \underset{A}{\operatorname{argmin}} E_H \left\{ \operatorname{Trace} \left( \left( \begin{bmatrix} H \\ A \end{bmatrix}^T \begin{bmatrix} H \\ A \end{bmatrix} - I \right) \left( \begin{bmatrix} H \\ A \end{bmatrix}^T \begin{bmatrix} H \\ A \end{bmatrix} - I \right)^T \right) \right\}$$

**[0034]** A computational process, such as provided by the Newton method, can be used to solve the resulting optimization problem where the gradient of objective function  $F$  is:

$$\nabla F = 4A(E_H[H^T H] + A^T A - I)$$

E denotes the statistical expectation operator, and the iterations of Newton

algorithm(for  $i_{th}$  column of  $\begin{bmatrix} H \\ A \end{bmatrix}$ ) is continued until convergence as:

$$A(:, i)_{k+1} = A(:, i)_k - \eta \nabla F \quad \text{for } i = 1, \dots, N$$

**[0035]** The network monitoring system 110 may receive a compressed representation of network measurements made by each aggregation 120 (232). The compressed information can serve as supplemental information to directly measured network information, which for example can be communicated over dedicated data links.

**[0036]** The network monitoring system 110 can apply a selected network inference technique based on the received measured network information, including the compressed representation of network measurements (234).

**[0037]** DATA CENTER EXAMPLE

**[0038]** FIG. 3 illustrates an example of a data center for implementing one or more examples. In an example of FIG. 3, a data center 300 includes a plurality of racks, represented by racks 320, 330 and a network management system 310. Each rack 320, 330 can include an aggregation of servers 322, 332, a set of top of rack switches 324, 334, and one or more aggregation switches 326, 336 and core switches 328, 338. In an example shown, the aggregation switches 326, 336 can obtain and communicate network measurements to the network monitoring system 310 through a set of links 311. While FIG. 3 illustrates an example in which an aggregation is provided by a rack, in other examples, an aggregation can be defined to extent to multiple physically connected and/or logically defined racks.

**[0039]** The network monitoring system 310 can include an optimization component 312 for determining a rack specific observation matrix 315, and a network inference engine 314. As described with other examples, the network monitoring system 310 can generate the observation matrix 315 for the particular rack 320. In some variations, the observation matrix 315 can be communicated offline to the respective rack 320. The aggregation switch 326 of the receiving rack can implement the observation matrix 315 to return

a compressed set of network measurements 317 on the link 311. The compressed set of measurements 317 can supplement directly measured values 313 which can also be communicated using the links 311. The network measuring system 310 can use the directly measured values 313 and the compressed set of measurements 317 as input for the inference engine 314.

**[0040]** Although illustrative embodiments have been described in detail herein with reference to the accompanying drawings, variations to specific embodiments and details are encompassed by this disclosure. It is intended that the scope of embodiments described herein be defined by claims and their equivalents. Furthermore, it is contemplated that a particular feature described, either individually or as part of an embodiment, can be combined with other individually described features, or parts of other embodiments. Thus, absence of describing combinations should not preclude the inventor(s) from claiming rights to such combinations.

**What is claimed is:**

1. A method for estimating traffic on a defined network, the method being implemented by one or more processors and comprising:
  - obtaining information which is indicative of (i) a topology of the defined network, the topology including multiple aggregations of network devices, and (ii) routing information that identifies a plurality of data flows amongst network devices of the defined network;
  - determining, using at least the topology and the routing information, a corresponding observation matrix for each of the multiple aggregations of network devices; and
  - communicating the corresponding observation matrix to each of the multiple aggregations of network devices in order to obtain, from each aggregation, a compressed representation of the network traffic through individual network devices of that aggregation.
2. The method of claim 1, further comprising generating a representation of the network traffic through the defined network using the compressed representation of each of the multiple aggregations of network devices.
3. The method of claim 1, wherein communicating the corresponding observation matrix to each of the multiple aggregations includes distributing the corresponding observation matrix to each of the multiple aggregations of network devices.
4. The method of claim 1, wherein determining the corresponding observation matrix for the multiple aggregations of network devices includes determining a statistical expectation, using at least the routing information, in order to account for multi-path routing of data provided in part through the network devices of each aggregation.
5. The method of claim 4, wherein determining the corresponding observation matrix for the multiple aggregations of network devices

includes determining, from the statistical expectation, a set of non-binary coefficients for the corresponding observation matrix of each aggregation.

6. The method of claim 4, wherein determining the statistical expectation includes using at least one of a probability distribution function of the routing matrix or a different set of realizations of the routing matrix.

7. The method of claim 1, wherein the corresponding observation matrix of each aggregation is optimized to improve estimation accuracy.

8. The method of claim 1, wherein the defined network corresponds to a data center, each aggregation of network devices corresponds to a rack of servers, and the method is performed at a network monitoring center.

9. A computer system comprising:  
a set of memory resources to store a set of instructions to monitor a network of the computer system;  
one or more processors to execute the set of instructions to:  
obtain information which is indicative of (i) a topology of the defined network, the topology including multiple aggregations of network devices, and (ii) routing information that identifies a plurality of data flows amongst network devices of the defined network;  
determine, using at least the topology and the routing information, a corresponding observation matrix for each of the multiple aggregations of network devices; and  
communicate the corresponding observation matrix to each of the multiple aggregations of network devices in order to obtain, from each aggregation, a compressed representation of the network traffic through individual network devices of that aggregation.

10. The system of claim 9, wherein the one or more processors execute the set of instructions to generate a representation of the network traffic through the defined network using the compressed representation of each of the multiple aggregations of network devices.

11. The system of claim 9, wherein the one or more processors execute the set of instructions to communicate the corresponding observation matrix to each of the multiple aggregations by distributing the corresponding observation matrix to each of the multiple aggregations of network devices.

12. The system of claim 9, wherein the one or more processors execute the set of instructions to (i) determine a statistical expectation, using at least the routing information, in order to account for multi-path routing of data provided in part through the network devices of each aggregation, and (ii) from the statistical expectation, determine a set of non-binary coefficients for the corresponding observation matrix of each aggregation, the set of non-binary coefficients representing multi-path routing of data for individual flows.

13. The system of claim 9, wherein the one or more processors optimize the observation matrix of each aggregation for reduction of sparseness.

14. The system of claim 9, wherein the one or more processors are implemented on a network monitor computer, and wherein each of the multiple aggregations corresponds to a rack of servers in a data center.

15. A data center system comprising:  
a network monitor computer;

a plurality of racks, each rack including an aggregation of multiple servers;

wherein the network monitor computer uses a network connection to each rack to:

determine, using at least a topology and routing information, a corresponding observation matrix for each of the multiple aggregations of network devices; and

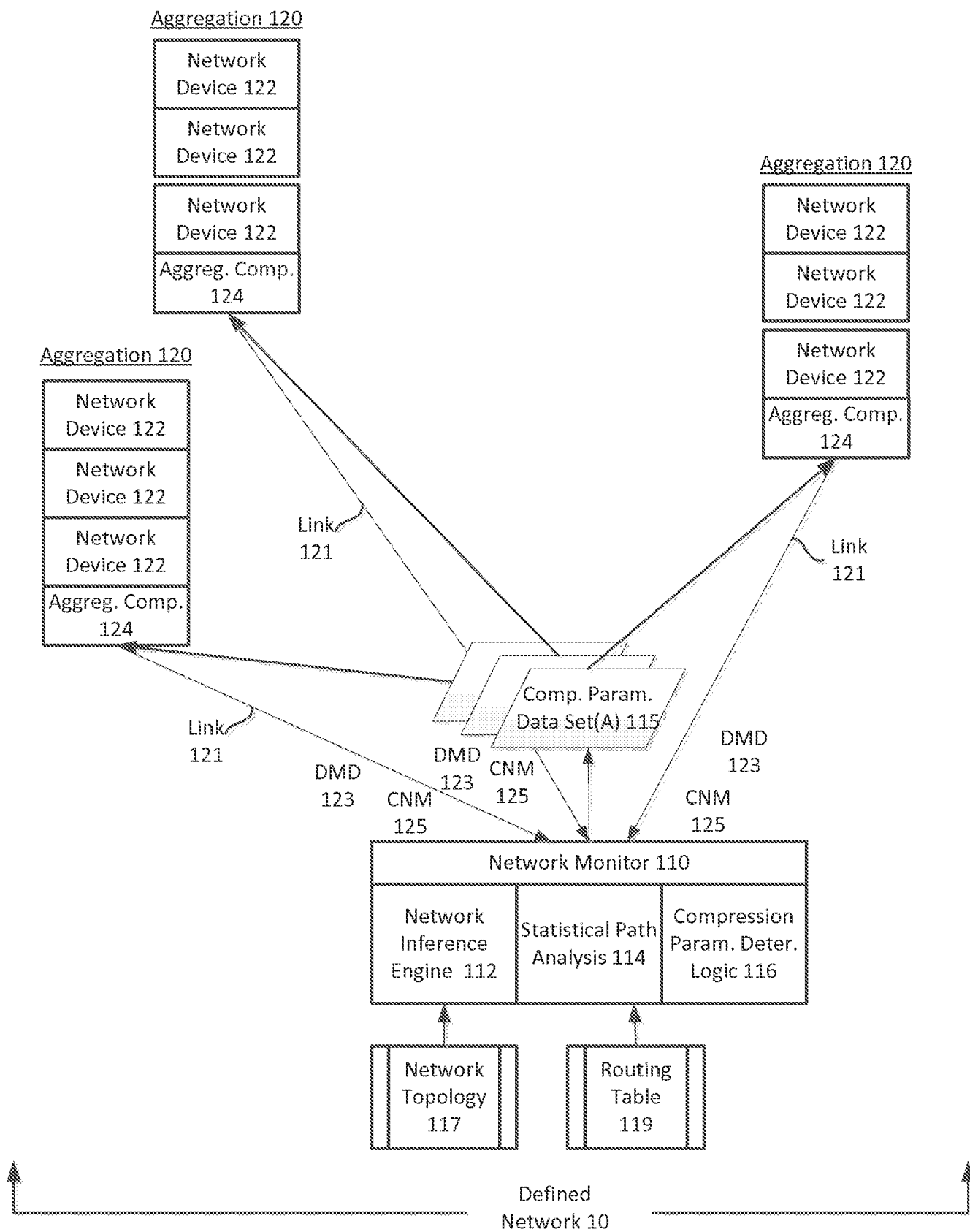
communicate the corresponding observation matrix to each of the multiple aggregations of network devices in order to obtain, from each aggregation, a compressed representation of the network traffic through individual network devices of that aggregation; and

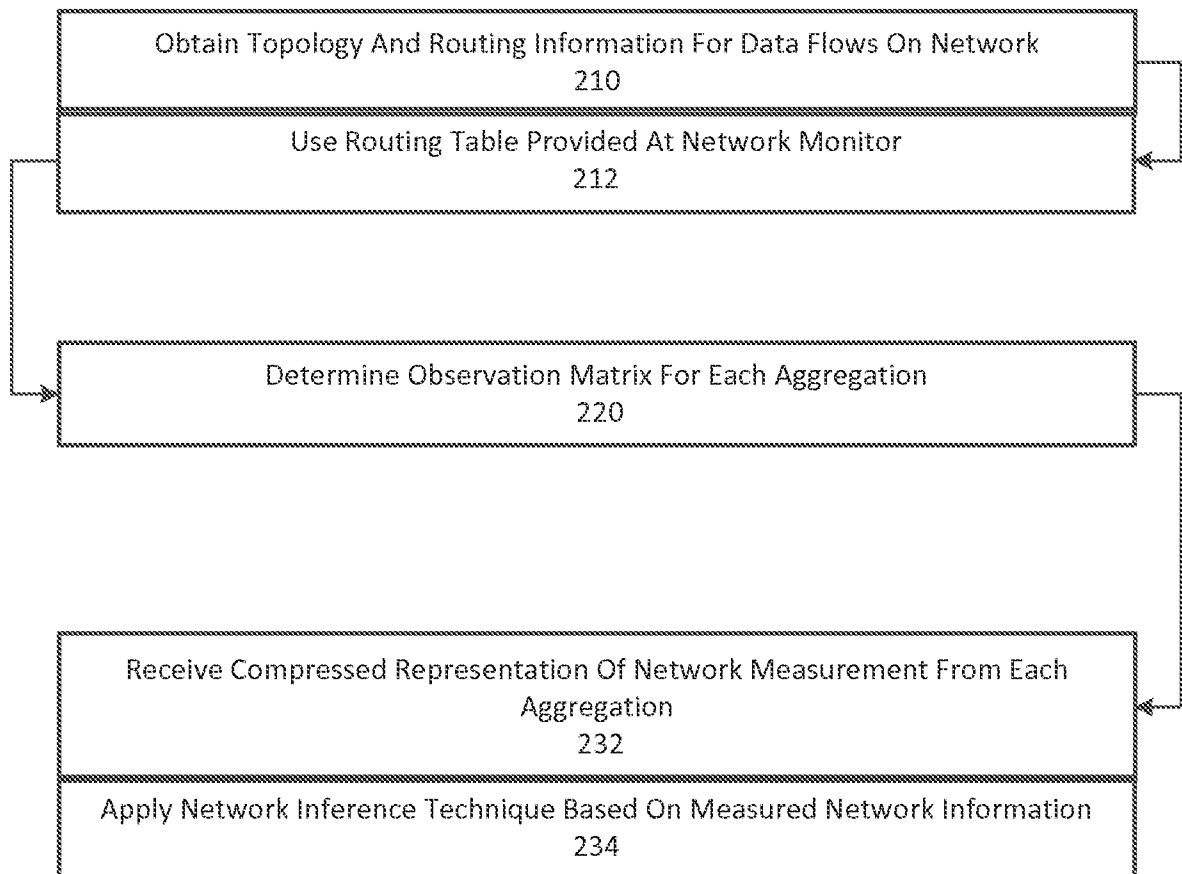
wherein each rack of the plurality of racks operates to:

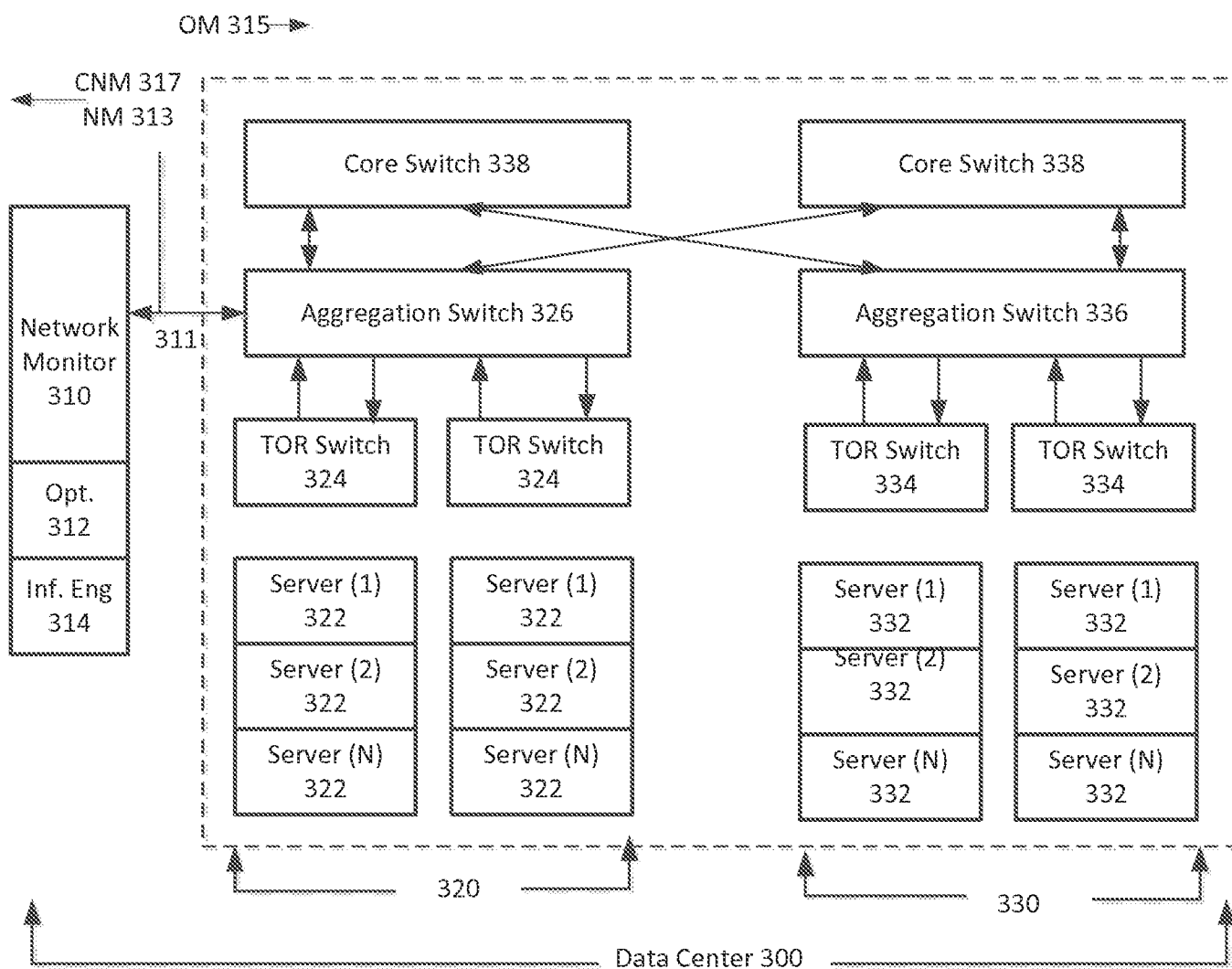
receive the corresponding observation matrix from the network monitor computer;

apply the observation matrix to network traffic measurements made on the rack for the aggregation of servers, and

send a determination from applying the observation matrix to the traffic measurements to the network monitor computer.

**FIG. 1**

**FIG. 2**

**FIG. 3**

## INTERNATIONAL SEARCH REPORT

International application No.  
**PCT/US2015/043020****A. CLASSIFICATION OF SUBJECT MATTER****H04L 12/26(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; G06F 15/16; G06F 15/173; H04L 12/24; H04L 12/28; H04L 12/56

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: network traffic, topology, routing information, aggregation, data flow, observation matrix****C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                            | Relevant to claim No. |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| A         | JP 2012-244302 A (NIPPON TELEGR & TELEPH CORP. <NTT>) 10 December 2012<br>See paragraphs [0021]–[0063] and figures 1–5.                                       | 1–15                  |
| A         | US 2005-0190695 A1 (PETER PHAAL) 01 September 2005<br>See paragraphs [0028]–[0034], [0072]–[0074] and figures 1–4.                                            | 1–15                  |
| A         | JP 2008-085812 A (OKI ELECTRIC IND. CO., LTD. et al.) 10 April 2008<br>See paragraphs [0014]–[0089] and figures 1–7.                                          | 1–15                  |
| A         | US 2010-0281388 A1 (JOHN KANE et al.) 04 November 2010<br>See paragraphs [0084]–[0136] and figures 8–14.                                                      | 1–15                  |
| A         | KR 10-2007-0120737 A (UNIVERSITY-INDUSTRY COOPERATION GROUP OF KYUNG HEE UNIVERSITY et al.) 26 December 2007<br>See paragraphs [0042]–[0046] and figures 6–7. | 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

"&amp;" document member of the same patent family

Date of the actual completion of the international search

18 April 2016 (18.04.2016)

Date of mailing of the international search report

**30 May 2016 (30.05.2016)**

Name and mailing address of the ISA/KR

International Application Division

Korean Intellectual Property Office

189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea

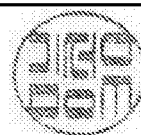


Facsimile No. +82-42-481-8578

Authorized officer

KIM, Seong Woo

Telephone No. +82-42-481-3348



**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/US2015/043020**

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s)                                                                                                                                                                                                                                                                                                                                                                                                                | Publication<br>date                                                                                                                                                                                                                                                                                                            |
|-------------------------------------------|---------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| JP 2012-244302 A                          | 10/12/2012          | JP 5362769 B2                                                                                                                                                                                                                                                                                                                                                                                                                             | 11/12/2013                                                                                                                                                                                                                                                                                                                     |
| US 2005-0190695 A1                        | 01/09/2005          | US 6894972 B1<br>US 7164657 B2                                                                                                                                                                                                                                                                                                                                                                                                            | 17/05/2005<br>16/01/2007                                                                                                                                                                                                                                                                                                       |
| JP 2008-085812 A                          | 10/04/2008          | JP 4425255 B2                                                                                                                                                                                                                                                                                                                                                                                                                             | 03/03/2010                                                                                                                                                                                                                                                                                                                     |
| US 2010-0281388 A1                        | 04/11/2010          | CA 2751211 A1<br>CA 2751211 C<br>CA 2751214 A1<br>CA 2751214 C<br>CN 102369690 A<br>CN 102379102 A<br>CN 102379102 B<br>EP 2392101 A1<br>EP 2392103 A1<br>EP 2392103 A4<br>JP 2012-517148 A<br>JP 2012-517149 A<br>JP 5327915 B2<br>JP 5403440 B2<br>KR 10-1223059 B1<br>KR 10-1353369 B1<br>KR 10-2011-0112460 A<br>US 2010-0195516 A1<br>US 2015-0039755 A1<br>US 8531978 B2<br>US 8838780 B2<br>WO 2010-088590 A1<br>WO 2010-088622 A1 | 05/08/2010<br>21/04/2015<br>05/08/2010<br>17/03/2015<br>07/03/2012<br>14/03/2012<br>10/09/2014<br>07/12/2011<br>07/12/2011<br>15/04/2015<br>26/07/2012<br>26/07/2012<br>30/10/2013<br>29/01/2014<br>18/01/2013<br>20/01/2014<br>12/10/2011<br>05/08/2010<br>05/02/2015<br>10/09/2013<br>16/09/2014<br>05/08/2010<br>05/08/2010 |
| KR 10-2007-0120737 A                      | 26/12/2007          | KR 10-0819543 B1                                                                                                                                                                                                                                                                                                                                                                                                                          | 07/04/2008                                                                                                                                                                                                                                                                                                                     |