



(43) International Publication Date  
05 December 2019 (05.12.2019)

(51) International Patent Classification:

G06Q 10/06 (2012.01) G06F 17/10 (2006.01)  
G06F 17/30 (2006.01)

(21) International Application Number:

PCT/SG20 18/050261

(22) International Filing Date:

28 May 2018 (28.05.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(71) Applicant: AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH [SG/SG]; 1 Fusionopolis Way, #20-10 Connexis North Tower, Singapore 138632 (SG).

(72) Inventors: SHAHAM, Eran; c/o PKM, Institute for Infocomm Research, 1 Fusionopolis Way, #20-10 Connexis North Tower, Singapore 138632 (SG). WESTERSKI, Adam Maciej; c/o PKM, Institute for Infocomm Research,

1 Fusionopolis Way, #20-10 Connexis North Tower, Singapore 138632 (SG). KANAGASABAI, Rajaraman; c/o PKM, Institute for Infocomm Research, 1 Fusionopolis Way, #20-10 Connexis North Tower, Singapore 138632 (SG).

(74) Agent: VIERING, JENTSCHURA & PARTNER LLP; P.O. Box 1088, Rochor Post Office, Rochor Road, Singapore 911833 (SG).

(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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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,

(54) Title: METHOD AND APPARATUS FOR PROCUREMENT DEMAND AGGREGATION

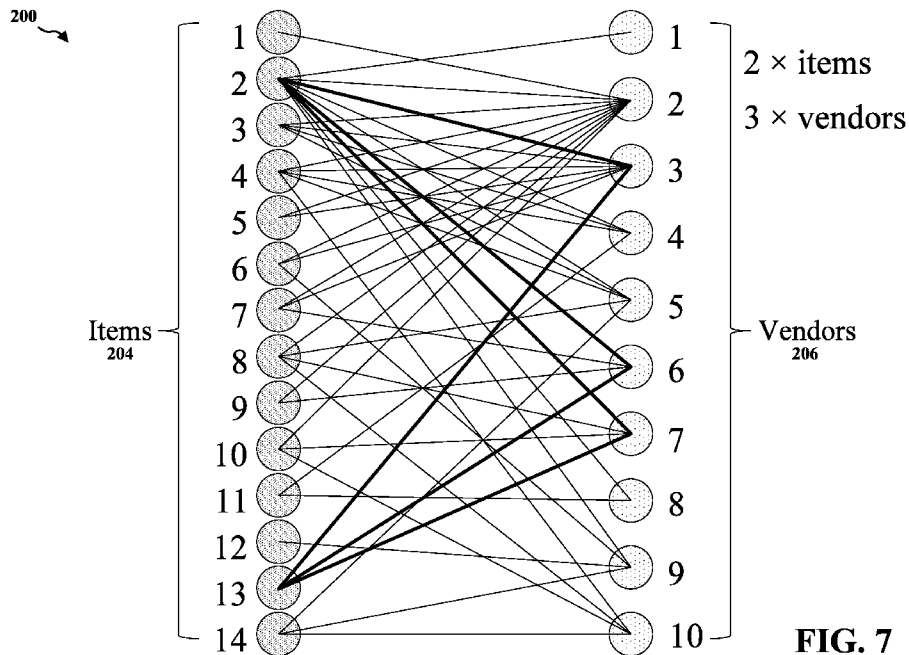


FIG. 7

(57) Abstract: A method, a computer-readable medium, and an apparatus for procurement demand aggregation are provided. The apparatus may receive a bipartite graph including a first plurality of vertices and a second plurality of vertices. Every edge of the bipartite graph may connect a vertex in the first plurality of vertices to a vertex in the second plurality of vertices. The first plurality of vertices may correspond to a plurality of items for procurement. The second plurality of vertices may correspond to a plurality of vendors. The apparatus may randomly select a set of items from the plurality of items and/or a set of vendors from the plurality of vendors. The apparatus may iteratively expand the set of items and/or the set of vendors to obtain a set of maximal patterns. The apparatus may aggregate procurement demand based on the set of maximal patterns.



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).

**Published:**

- with international search report (Art. 21(3))
- in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE

## **METHOD AND APPARATUS FOR PROCUREMENT DEMAND AGGREGATION**

### **TECHNICAL FIELD**

[0001] Various aspects of this disclosure generally relate to procurement, and more particularly, to procurement demand aggregation.

### **BACKGROUND**

[0002] Large organizations spend millions on purchases of goods and services. Procurement is an essential operation of every organization, regardless of its size, business domain, or sector (e.g., private or government). Therefore, optimizing the procurement process has a great importance. Demand Aggregation (DA) is the process of simultaneously grouping items (e.g., goods and/or services) and vendors. As a result of the demand aggregation, the grouped items may be sourced exclusively from the grouped vendors, and the grouped vendors may only provide the grouped items to the organization. Such aggregation could lead to a better value-for-money due to: (1) lower bulk prices; (2) larger vendor tendering; (3) lower shipping and handling fees; and (4) reduced legal and admin overhead. In the procurement operation of a large organization, even a small percentage of savings means substantial savings.

[0003] Several industrial attempts have been done to tackle the problem of demand aggregation for procurement. Most of them use a naive aggregation of only one dimension (e.g., item, vendor). However, such approach is far from optimal as it does not utilize the multi-item multi-vendor nature of the procurement.

### **SUMMARY**

[0004] The following presents a simplified summary in order to provide a basic understanding of various aspects of the disclosed invention. This summary is not an extensive overview of all contemplated aspects, and is intended to neither identify key or critical elements of all aspects nor delineate the scope of any or all aspects. The sole purpose is to present some concepts of one or more aspects in a simplified form as a prelude to the more detailed description that is presented later.

[0005] This disclosure focuses on the problem of finding the set of maximal edge bicliques (potentially overlapping). Each such maximal edge biclique may serve as a potential demand aggregation. An efficient Subspace Biclique Clustering for Procurement (SBCP) algorithm is provided to tackle this challenging problem.

[0006] In one aspect of the disclosure, a method, a computer-readable medium, and an apparatus for procurement demand aggregation are provided. The apparatus may receive a bipartite graph including a first plurality of vertices and a second plurality of vertices. Every edge of the bipartite graph may connect a vertex in the first plurality of vertices to a vertex in the second plurality of vertices. The first plurality of vertices may correspond to a plurality of items for procurement. Each vertex in the first plurality of vertices may represent an item for procurement. The second plurality of vertices may correspond to a plurality of vendors. Each vertex in the second plurality of vertices may represent a vendor. The apparatus may randomly select a set of items from the plurality of items and/or a set of vendors from the plurality of vendors. The apparatus may iteratively expand the set of items and/or the set of vendors to obtain a set of maximal patterns. The apparatus may aggregate procurement demand based on the set of maximal patterns.

[0007] To the accomplishment of the foregoing and related ends, the aspects disclosed include the features hereinafter fully described and particularly pointed out in the claims. The following description and the annexed drawings set forth in detail illustrate certain features of the aspects of the disclosure. These features are indicative, however, of but a few of the various ways in which the principles of various aspects may be employed, and this description is intended to include all such aspects and their equivalents.

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

[0008] FIG. 1 illustrates a simple example of a procurement bipartite graph.

[0009] FIG. 2 illustrates an example of a procurement bipartite graph.

[0010] FIG. 3 illustrates a biclique of ten items and one vendor that is identified in the procurement bipartite graph described above in FIG. 2.

[0011] FIG. 4 illustrates a biclique of six items and two vendors that is identified in the procurement bipartite graph described above in FIG. 2.

[0012] FIG. 5 illustrates a biclique of three items and four vendors that is identified in the procurement bipartite graph described above in FIG. 2.

[0013] FIG. 6 illustrates a biclique of one item and nine vendors that is identified in the procurement bipartite graph described above in FIG. 2.

[0014] FIG. 7 illustrates a biclique of two items and three vendors that is identified in the procurement bipartite graph described above in FIG. 2.

[0015] FIGS. 8A and 8B illustrates an example of a bipartite graph and its corresponding adjacency matrix.

[0016] FIG. 9 is a flowchart of a method of procurement demand aggregation.

[0017] FIG. 10 illustrates an example of an adjacency matrix of a bipartite graph for document classification based on words contained in the documents.

[0018] FIG. 11 illustrates an example of an adjacency matrix of a bipartite graph for user classification based on their movie preferences.

[0019] FIG. 12 is a conceptual data flow diagram illustrating the data flow between different means/components in an exemplary apparatus.

[0020] FIG. 13 is a diagram illustrating an example of a hardware implementation for an apparatus employing a processing system.

## **DETAILED DESCRIPTION**

[0021] The detailed description set forth below in connection with the appended drawings is intended as a description of various possible configurations and is not intended to represent the only configurations in which the concepts described herein may be practiced. The detailed description includes specific details for the purpose of providing a thorough understanding of various concepts. However, it will be apparent to those skilled in the art that these concepts may be practiced without these specific details. In some instances, well-known structures and components are shown in block diagram form in order to avoid obscuring such concepts.

[0022] Several aspects of procurement demand aggregation will now be presented with reference to various apparatus and methods. The apparatus and methods will be described in the following detailed description and illustrated in the accompanying drawings by various blocks, components, circuits, processes, algorithms, etc. (collectively referred to as

“elements”). These elements may be implemented using electronic hardware, computer software, or any combination thereof. Whether such elements are implemented as hardware or software depends upon the particular application and design constraints imposed on the overall system.

[0023] By way of example, an element, or any portion of an element, or any combination of elements may be implemented as a “processing system” that includes one or more processors, or one or more computing devices. Examples of processors include microprocessors, microcontrollers, graphics processing units (GPUs), central processing units (CPUs), application processors, digital signal processors (DSPs), reduced instruction set computing (RISC) processors, systems on a chip (SoC), baseband processors, field programmable gate arrays (FPGAs), programmable logic devices (PLDs), state machines, gated logic, discrete hardware circuits, and other suitable hardware configured to perform the various functionality described throughout this disclosure. One or more processors in the processing system may execute software. Software shall be construed broadly to mean instructions, instruction sets, code, code segments, program code, programs, subprograms, software components, applications, software applications, software packages, routines, subroutines, objects, executables, threads of execution, procedures, functions, etc., whether referred to as software, firmware, middleware, microcode, hardware description language, or otherwise.

[0024] Accordingly, in one or more example embodiments, the functions described may be implemented in hardware, software, or any combination thereof. If implemented in software, the functions may be stored on or encoded as one or more instructions or code on a computer-readable medium. Computer-readable media includes computer storage media. Storage media may be any available media that can be accessed by a computer. By way of example, and not limitation, such computer-readable media may include a random-access memory (RAM), a read-only memory (ROM), an electrically erasable programmable ROM (EEPROM), optical disk storage, magnetic disk storage, other magnetic storage devices, combinations of the aforementioned types of computer-readable media, or any other medium that can be used to store computer executable code in the form of instructions or data structures that can be accessed by a computer.

[0025] A bipartite graph is a graph whose vertices can be divided into two disjoint and independent sets  $U$  and  $V$  such that every edge connects a vertex in  $U$  to one in  $V$ .

Bipartite graphs may be useful in modelling a wide range of relationship networks. The relationship between items for procurement and vendors may be modeled by a bipartite graph.

[0026] FIG. 1 illustrates a simple example of a procurement bipartite graph 100. In the example, the procurement bipartite graph 100 includes fourteen items 102, ten vendors 104, and the relationship between item 7 and vendor 4. In some embodiments, the relationship between item 7 and vendor 4 represents that item 7 was purchased or sourced from vendor 4.

[0027] A simultaneous grouping of items and vendors within a bipartite graph may be referred to as a biclique. Given a bipartite graph and its corresponding partition into two disjoint sets of vertices, a biclique is a complete bipartite subgraph such that every vertex of the first partition is connected to every vertex of the second partition. Mathematically, the notion of biclique is defined as follows.

[0028] Definition 1. Let  $G = (U \cup V, E)$  be a bipartite graph, where  $U$  and  $V$  are two disjoint sets of vertices, and  $E$  is an edge set such that  $\forall (i, j) \in E, i \in U, j \in V$ . A biclique within  $G$  is a couple (set pair)  $(I, J)$  such that  $I \subseteq U, J \subseteq V$  and  $\forall i \in I, j \in J, (i, j) \in E$ .

[0029] FIG. 2 illustrates an example of a procurement bipartite graph 200. In the example, the procurement bipartite graph 200 includes fourteen items 204, ten vendors 206, and multiple relationships between the items 204 and the vendors 206. In some embodiments, a relationship between a particular item and a particular vendor represents that the particular item was purchased or sourced from the particular vendor.

[0030] FIG. 3 illustrates a biclique of ten items and one vendor that is identified in the procurement bipartite graph 200 described above in FIG. 2. As shown, items 1-10 of the items 204 and vendor 2 of the vendors 206 form the biclique.

[0031] FIG. 4 illustrates a biclique of six items and two vendors that is identified in the procurement bipartite graph 200 described above in FIG. 2. As shown, items 2-7 of the items 204 and vendors 2-3 of the vendors 206 form the biclique.

[0032] FIG. 5 illustrates a biclique of three items and four vendors that is identified in the procurement bipartite graph 200 described above in FIG. 2. As shown, items 2-4 of the items 204 and vendors 2-5 of the vendors 206 form the biclique.

[0033] FIG. 6 illustrates a biclique of one item and nine vendors that is identified in the procurement bipartite graph 200 described above in FIG. 2. As shown, item 2 of the items 204 and vendors 1-9 of the vendors 206 form the biclique.

[0034] FIG. 7 illustrates a biclique of two items and three vendors that is identified in the procurement bipartite graph 200 described above in FIG. 2. As shown, items 2, 13 of the items 204 and vendors 3, 6, 7 of the vendors 206 form the biclique.

[0035] The computational complexity of finding the maximum biclique depends on the exact objective function used. The maximum biclique problem has three distinct variants, with the following objective function  $\mu(I, J)$ :

[0036] (1)  $\mu(I, J) = |I| + |J|$  — known as the MAXIMUM VERTEX BICLIQUE problem. The problem may be solved in polynomial time using a minimum cut algorithm.

[0037] (2)  $\mu(I, J) = |I|$ , where  $|I| = |J|$  — known as the BALANCED COMPLETE BIPARTITE SUBGRAPH problem (also known as the balanced biclique problem). The problem was proved to be NP-complete.

[0038] (3)  $\mu(I, J) = |I| \times |J|$  — known as the MAXIMUM EDGE BICLIQUE problem. The problem was proved to be NP-complete, and challenging to approximate.

[0039] This disclosure focuses on the problem of finding the set of maximal edge bicliques (potentially overlapping). Each such maximal edge biclique may serve as a potential demand aggregation. An efficient Subspace Biclique Clustering for Procurement (SBCP) algorithm is provided to tackle this challenging problem. Extensive experimentations on artificial and real-world procurement datasets demonstrate the superiority of the SBCP algorithm over traditional techniques.

[0040] Mining maximal edge bicliques within a bipartite graph (e.g., the bipartite graph 200) may reveal potential demand aggregations. The SBCP algorithm is provided for mining such bicliques using a novel Monte Carlo subspace clustering approach.

[0041] A bipartite graph's adjacency matrix representation may be defined as follows.

[0042] Definition 2. Let  $G = (U \cup V, E)$  be a bipartite graph such that  $|U| = m$ , and  $|V| = n$ . The adjacency matrix  $X$  of graph  $G$  is a  $[m \times n]$  matrix, such that  $X_{i,j} = 1$  if  $(i, j) \in E$  and  $X_{i,j} = 0$  otherwise.



[0043] FIGS. 8A and 8B illustrates an example of a bipartite graph 800 and its corresponding adjacency matrix 850. In the example, the bipartite graph 800 includes the maximum edge biclique  $(\{i_3, i_4, i_6\}, \{j_3, j_5\})$  of six edges and five vertices.

[0044] The input of the SBCP algorithm may be an adjacency matrix  $X$  of a given bipartite graph  $G$ , consisting of only Boolean numbers, namely 0 and 1. The output of the SBCP algorithm may be a list of maximal bicliques, i.e., a list of submatrices of ones, representing maximal bicliques within  $G$ . A maximal biclique is one not contained in any larger biclique. In some embodiments, the bipartite graph may contain multiple, possibly overlapping, maximal bicliques. The SBCP algorithm may use a subspace clustering approach. This technique uses iterative random projection (i.e., a Monte Carlo strategy) to obtain the biclique's seed, which is later expanded into a maximal biclique.

[0045] SBCP algorithm for extracting a list of maximal bicliques.

Input:  $X$ , a  $[m \times n]$  matrix of Boolean numbers.

Output: List of maximal bicliques.

Initialization: Setting of  $TV$ ,  $|P|$  and  $|S|$ .

1: loop  $N$  times

2: // Seeding stage

3: choose a set of rows  $P$  uniformly at random;

4: set  $I \leftarrow P, J \leftarrow 0$ ;

5: // Interleaving row and column addition stage

6: set  $isAddRow \leftarrow \text{False}$ ;

7: set row  $i \leftarrow 1$ , column  $j \leftarrow 1$ ;

8: while row  $i \leq m$  or column  $j \leq n$  do

9:     if  $isAddRow$  then // row addition

10:         if  $X_{i,j} = 1$  then

11:             add  $i$  to  $I$ ;

12:              $i \leftarrow i + 1$ ;

13:         if  $j \leq n$  then

14:              $isAddRow \leftarrow !isAddRow$

15:         else // column addition

```

16:           if  $X_{i,j} = 1$  then
17:               add  $j$  to  $J$ ;
18:            $j \leftarrow j + 1$ ;
19:           if  $i \leq m \wedge (|J| \geq |S| \vee j > n)$  then
20:               isAddRow  $\leftarrow$  !isAddRow
21: return list of  $(I, J)$ ;

```

[0046] The structure of the SBCP algorithm may be divided into the following stages:

[0047] (i) Seeding (lines 2-4): a random selection of a set of rows to serve as a seed of the maximal biclique.

[0048] (ii) Addition of rows and columns (lines 5-20): interleaved accumulation of rows (lines 9-14) and columns (lines 15-20), which comply with the rows and columns already accumulated.

[0049] (iii) Polynomial repetition (line 1): repetition of the above two stages provides a probabilistic guarantee of acquiring a set of maximal bicliques.

[0050] The Monte Carlo nature of the SBCP algorithm is revealed in stage (i) where random seeds are generated. The subspace clustering nature of the SBCP algorithm is revealed in stage (ii), where the seed of stage (i) is expanded to form a maximal subset of rows over a maximal subset of columns, i.e., a maximal biclique.

[0051] To ease readability, the SBCP algorithm is presented as first obtaining the columns and subsequently obtaining the rows (at stage (ii)). However, the algorithm is symmetric, i.e., it may also be performed by first obtaining the rows and then obtaining the columns. In fact, running the algorithm in an interleaving fashion (i.e., alternating between first obtaining the rows, and first obtaining the columns) might help in finding imbalanced bicliques).

[0052] Due to the Monte Carlo nature of the SBCP algorithm, its iterations (line 1) are independent of each other. The algorithm may therefore be implemented in an efficient distributed computing environment to take advantage of parallel computing or special hardware in a straightforward manner.

[0053] To ease readability, lines 10 and 16 use the short notations of:  $X_{i,j} = 1$  and  $X_{i,j} = 1$ , respectively, which have the meaning of:  $\forall j \in J, \exists i \in I, X_{i,j} = 1$  and  $\forall i \in I, \exists j \in J, X_{i,j} = 1$ , respectively.

[0054] The SBCP algorithm has an inherent ability to mine multiple, possibly overlapping, bicliques by utilizing the independent random projection on each repetitive run, to reveal columns and rows relevant only to a specific biclique.

[0055] The SBCP algorithm is not designed for the enumeration of all maximal bicliques, which may be exponential in size. The algorithm has a polynomial number of iterations, and thus, the size of the return list is polynomial. However, the returned list contains, with a fixed probability, optimal bicliques.

[0056] The SBCP algorithm may be viewed as a heuristic method. For an optimal biclique,  $(I^*, J^*)$ , any randomly chosen subset  $P \subseteq I^*$  of size  $O(\log n)$  is a discriminating set for  $J^*$ , with a probability of at least 0.5. The same can be applied to using a discriminating columns set: for an optimal biclique,  $(I^*, J^*)$ , any randomly chosen subset  $S \subseteq J^*$  of size  $O(\log m)$  is a discriminating set for  $I^*$ , with a probability of at least 0.5.

[0057] The SBCP algorithm uses the above to expand the initial seed pattern (see line 4). It does so by interleaving the expansion between rows and columns addition (line 9 and 15, respectively), while using previous accumulates (columns and rows, respectively, line 10 and 16, respectively), as ever-growing discriminating sets. The later promises better discriminating results, which leads to better detection probability, which results in reduced overall run-time. The run-time of the SBCP algorithm may be polynomial:  $mn^{O(1)}$ . Thus, the SBCP algorithm is able to scale to large datasets and produce demand aggregation patterns in feasible run-time.

[0058] In some embodiments, mining algorithm may be employed for demand aggregation in procurement. In some embodiments, a novel scalable bi-clique miner may be employed. In some embodiments, constraints may be employed on mined bicliques to reduce false positive DA patterns.

[0059] The SBCP algorithm may efficiently mine patterns that are procurement relevant, namely demand aggregation patterns. The SBCP algorithm was run on a procurement database of the years 2009 - 2016. The database comprises 660,162 items, 14,834 vendors, and 1,032,275 POs (Purchasing Order). In addition, the following domain constraints have been imposed:

- Use of records of three years (i.e., years 2014 - 2016). This results in a database comprises 271,219 items, 7,319 vendors, and 391,671 POs, which constructs a bipartite graph  $G$  of  $|U(G)|=271,219$ ,  $|V(G)|=7,319$ , and  $|E(G)|=391,671$ .
- Filter out patterns with less than 20 POs a year.
- Filter out patterns with a value less than 70,000 SGD a year.
- Filter out patterns with a descending price trend (trend calculated as linear regression).
- Allow patterns to fail one of the above filters.

[0060] The SBCP miner found 643 patterns that comply with the above constraints. Evaluation of these patterns was done in the following manner:

- Flow many of the existing 33 DA patterns (i.e., patterns previously used) are identified.
- Assess new DA patterns (i.e., DAs identified by the system that differ from existing DAs) for their demand aggregation suitability.

[0061] The accuracy of the existing DAs is 71% (some of the existing DAs do not pass the above filters and thus have been excluded from the miner's list of patterns). The accuracy of the new DAs is 81%. The incurred savings is divided into two:

- Forecast saving of S\$7M due to bulk purchases.
- Expected S\$98M saving due to legal and admin cut.

[0062] This disclosure presents the SBCP algorithm, a probabilistic method for finding the maximum edge biclique using a Monte Carlo subspace clustering approach. While the SBCP algorithm can be viewed as an effective heuristic, there is a strong theoretical base for its efficacy. The SBCP algorithm promises a fixed probability of finding an optimal biclique (in the sense of balancing the number of rows and the number of columns), while giving a polynomial bound to the complexity of time and space. The SBCP algorithm achieves that with lower complexity bounds and higher detection probability. Comprehensive experimentation with real-world procurement database corroborates the usability, feasibility and scalability of the SBCP algorithm to efficiently mine relevant demand aggregation patterns.

[0063] FIG. 9 is a flowchart 900 of a method of procurement demand aggregation. In some embodiments, the method may be performed by an apparatus (e.g., the apparatus 1202/1202' shown in FIG. 12 or FIG. 13). At 902, the apparatus may receive a bipartite

graph including a first plurality of vertices and a second plurality of vertices. Every edge of the bipartite graph may connect a vertex in the first plurality of vertices to a vertex in the second plurality of vertices. The first plurality of vertices may correspond to a plurality of items for procurement. Each vertex in the first plurality of vertices may represent an item for procurement. The second plurality of vertices may correspond to a plurality of vendors. Each vertex in the second plurality of vertices may represent a vendor. In some embodiments, the bipartite graph may be represented by an adjacency matrix.

[0064] At 904, the apparatus may randomly select an initial set of items from the plurality of items and/or an initial set of vendors from the plurality of vendors. In some embodiments, the operations performed at 904 may correspond to the seeding stage of the SBCP algorithm.

[0065] At 906, the apparatus may iteratively expand the initial set of items and/or the initial set of vendors to obtain a set of maximal patterns. At each iteration, the apparatus may obtain a mined intermediate pattern that includes an expanded set of items and/or an expanded set of vendors. In some embodiments, the operations performed at 906 may correspond to the row and column addition stage of the SBCP algorithm. In some embodiments, the set of maximal patterns may contain a set of maximal bicliques. A maximal biclique is one not contained in any larger biclique. A maximal biclique is a biclique that cannot be expanded further.

[0066] Each maximal biclique may include a first set of vertices and a second set of vertices, where every vertex in the first set of vertices is connected to every vertex in the second set of vertices in the bipartite graph. In some embodiments, the first set of vertices may correspond to a resulting set of items. The resulting set of items may be expanded from the initial set of items. The second set of vertices may correspond to a resulting set of vendors. The resulting set of vendors may be expanded from the initial set of vendors.

[0067] In some embodiments, to iteratively expand the initial set of items and/or the initial set of vendors, the apparatus may interleave between adding an item into the set of items and adding a vendor into the set of vendors. In some embodiments, an item may be added into the set of items when a vertex corresponding to the item is connected to every vertex of a set of vertices corresponding to vendors that are currently within the mined intermediate pattern. In some embodiments, a vendor may be added into the set of vendors

when a vertex corresponding to the vendor is connected to every vertex of a set of vertices corresponding to items that are currently within the mined intermediate pattern.

**[0068]** At 908, the apparatus may aggregate procurement demand based on the set of maximal patterns. In some embodiments, to aggregate the procurement demand, the apparatus may procure items in the resulting set of items exclusively with the resulting set of vendors. In some embodiments, to aggregate the procurement demand, the apparatus may procure only the resulting set of items from the resulting set of vendors.

**[0069]** FIG. 10 illustrates an example of an adjacency matrix 1000 of a bipartite graph for document classification based on words contained in the documents. In the example, the bipartite graph corresponding to the adjacency matrix 1000 includes five documents A-E, six words, and elements of the adjacency matrix 1000 represent relationships between the documents and the words. If a word is contained in a document, there is a relationship between the document and the word, and the corresponding element of the adjacency matrix 1000 is set to '1'; otherwise the corresponding element is set to '0'. In some embodiments, the adjacency matrix 1000 may serve as the input to the SBCP algorithm. As a result, the SBCP algorithm may output a set of maximal bicliques. For example, documents A and B share the words 'Obama' and 'Iraq', thus forming a maximal biclique. Therefore, documents A and B probably belong to the same class of 'US politics over war in Iraq'. Documents C and D share the words 'Football' and 'FIFA', thus forming a maximal biclique. Therefore, documents C and D probably belong to the same class of 'FIFA World Cup'. In some embodiments, new documents may be classified to one or more topics based on their words similarity to other documents (e.g., being in the same maximal biclique) of the same topics.

**[0070]** FIG. 11 illustrates an example of an adjacency matrix 1100 of a bipartite graph for user classification based on their movie preferences. In the example, the bipartite graph corresponding to the adjacency matrix 1100 includes five movies, six users 1-6, and elements of the adjacency matrix 1100 represent relationships between the movies and the users. If a movie is included in the movie preferences of a user, there is a relationship between the movie and the user, and the corresponding element of the adjacency matrix 1100 is set to '1'; otherwise the corresponding element is set to '0'. In some embodiments, the adjacency matrix 1100 may serve as the input to the SBCP algorithm. As a result, the SBCP algorithm may output a set of maximal bicliques. For example, movies 'Iron Man' and 'Thor' share users 1 and 3 preferences, thus forming

a maximal biclique. Users 1 and 3 probably prefer 'Marvel' movies. Movies 'Toy Story 3' and 'Frozen' share users 2 and 6 preferences, thus forming a maximal biclique. Users 2 and 6 probably prefer 'Walt Disney' movies. In some embodiments, the maximal bicliques output by the SBCP algorithm may be used to recommend movie(s) to user(s).

[0071] In some examples, the SBCP algorithm may be used to analyze protein-protein interaction. In such examples, a first plurality of vertices in a bipartite graph may represent a first plurality of proteins, and a second plurality of vertices in the bipartite graph may represent a second plurality of proteins. If a protein in the first plurality of proteins interacts with a protein in the second plurality of proteins, there is an edge in the bipartite graph connecting corresponding vertices. The bipartite graph may serve as the input to the SBCP algorithm. As a result, the SBCP algorithm may output a set of maximal bicliques. Each maximal biclique may suggest common functionality of grouped proteins in the first plurality of proteins and/or common functionality of grouped proteins in the second plurality of proteins. Therefore, each maximal biclique may suggest common functionality of grouped proteins in the first and second plurality of proteins together (i.e., the fact that proteins of both the first plurality of proteins and the second plurality of proteins all interact).

[0072] FIG. 12 is a conceptual data flow diagram 1200 illustrating the data flow between different means/components in an exemplary apparatus 1202. The apparatus 1202 may be a computing device or a system including multiple computing devices. The apparatus 1202 may include a demand aggregation component 1210 that receive a bipartite graph.

[0073] The demand aggregation component 1210 may include a seed selection component 1204 that selects a seed set of rows and columns from an adjacency matrix. The adjacency matrix corresponds to the bipartite graph. In one embodiment, the seed selection component 1204 may perform the operations described above with reference to 904 in FIG. 9.

[0074] The demand aggregation component 1210 may include a maximal pattern identification component 1206 that expands the seed set of rows and columns received from the seed selection component 1204 to identify a set of maximal bicliques. In one embodiment, the maximal pattern determination component 1206 may perform the operations described above with reference to 906 in FIG. 9.

[0075] The apparatus 1202 may include additional components that perform each of the blocks of the algorithm in the aforementioned flowchart of FIG. 9. As such, each block in the aforementioned flowchart of FIG. 9 may be performed by a component and the apparatus may include one or more of those components. The components may be one or more hardware components specifically configured to carry out the stated processes/algorithm, implemented by a processor configured to perform the stated processes/algorithm, stored within a computer-readable medium for implementation by a processor, or some combination thereof.

[0076] FIG. 13 is a diagram 1300 illustrating an example of a hardware implementation for an apparatus 1202' employing a processing system 1314. In some embodiments, the apparatus 1202' may be the apparatus 1202 described above with reference to FIG. 12. The apparatus 1202' may include one or more computing devices. The processing system 1314 may be implemented with a bus architecture, represented generally by the bus 1324. The bus 1324 may include any number of interconnecting buses and bridges depending on the specific application of the processing system 1314 and the overall design constraints. The bus 1324 links together various circuits including one or more processors and/or hardware components, represented by the processor 1304, the component 1210 (including the components 1204, 1206), and the computer-readable medium / memory 1306. The bus 1324 may also link various other circuits such as timing sources, peripherals, voltage regulators, and power management circuits, which are well known in the art, and therefore, will not be described any further.

[0077] The processing system 1314 includes a processor 1304 coupled to a computer-readable medium / memory 1306. The processor 1304 is responsible for general processing, including the execution of software stored on the computer-readable medium / memory 1306. The software, when executed by the processor 1304, causes the processing system 1314 to perform the various functions described *supra* for any particular apparatus. The computer-readable medium / memory 1306 may also be used for storing data that is manipulated by the processor 1304 when executing software. The processing system 1314 further includes the component 1210 (including the components 1204 and 1206). The components may be software components running in the processor 1304, resident/stored in the computer readable medium / memory 1306, one or more hardware components coupled to the processor 1304, or some combination thereof.



[0078] In the following, various aspects of this disclosure will be illustrated:

[0079] Example 1 is a method or apparatus for procurement demand aggregation. The apparatus may receive a bipartite graph including a first plurality of vertices and a second plurality of vertices. Every edge of the bipartite graph may connect a vertex in the first plurality of vertices to a vertex in the second plurality of vertices. The first plurality of vertices may correspond to a plurality of items for procurement. Each vertex in the first plurality of vertices may represent an item for procurement. The second plurality of vertices may correspond to a plurality of vendors. Each vertex in the second plurality of vertices may represent a vendor. The apparatus may randomly select a set of items from the plurality of items and/or a set of vendors from the plurality of vendors. The apparatus may iteratively expand the set of items and/or the set of vendors to obtain a set of maximal patterns. The apparatus may aggregate procurement demand based on the set of maximal patterns.

[0080] In Example 2, the subject matter of Example 1 may optionally include that the set of maximal patterns may include a set of maximal bicliques.

[0081] In Example 3, the subject matter of Example 2 may optionally include that each maximal biclique may include a first set of vertices and a second set of vertices, where every vertex in the first set of vertices is connected to every vertex in the second set of vertices in the bipartite graph.

[0082] In Example 4, the subject matter of Example 3 may optionally include that the first set of vertices may correspond to a resulting set of items, the resulting set of items being expanded from the set of items, the second set of vertices corresponding to a resulting set of vendors, the resulting set of vendors being expanded from the set of vendors.

[0083] In Example 5, the subject matter of Example 4 may optionally include that, to aggregate the procurement demand, the apparatus may procure items in the resulting set of items exclusively with the resulting set of vendors.

[0084] In Example 6, the subject matter of any one of Examples 4 to 5 may optionally include that, to aggregate the procurement demand, the apparatus may procure only the resulting set of items from the resulting set of vendors.

[0085] In Example 7, the subject matter of any one of Examples 1 to 6 may optionally include that, to iteratively expand the set of items and/or the set of vendors,

the apparatus may interleave between adding an item into the set of items and adding a vendor into the set of vendors.

[0086] In Example 8, the subject matter of Example 7 may optionally include that an item may be added into the set of items when a vertex corresponding to the item is connected to every vertex of a set of vertices corresponding to the set of vendors in the bipartite graph.

[0087] In Example 9, the subject matter of any one of Examples 7 to 8 may optionally include that a vendor may be added into the set of vendors when a vertex corresponding to the vendor is connected to every vertex of a set of vertices corresponding to the set of items in the bipartite graph.

[0088] A person skilled in the art will appreciate that the terminology used herein is for the purpose of describing various embodiments only and is not intended to be limiting of the present invention. As used herein, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms "comprises" and/or "comprising," when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof.

[0089] It is understood that the specific order or hierarchy of blocks in the processes / flowcharts disclosed is an illustration of exemplary approaches. Based upon design preferences, it is understood that the specific order or hierarchy of blocks in the processes / flowcharts may be rearranged. Further, some blocks may be combined or omitted. The accompanying method claims present elements of the various blocks in a sample order, and are not meant to be limited to the specific order or hierarchy presented.

[0090] The previous description is provided to enable any person skilled in the art to practice the various aspects described herein. Various modifications to these aspects will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other aspects. Thus, the claims are not intended to be limited to the aspects shown herein, but is to be accorded the full scope consistent with the language claims, wherein reference to an element in the singular is not intended to mean "one and only one" unless specifically so stated, but rather "one or more." The word

“exemplary” is used herein to mean “serving as an example, instance, or illustration.” Any aspect described herein as “exemplary” is not necessarily to be construed as preferred or advantageous over other aspects. Unless specifically stated otherwise, the term “some” refers to one or more. Combinations such as “at least one of A, B, or C,” “one or more of A, B, or C,” “at least one of A, B, and C,” “one or more of A, B, and C,” and “A, B, C, or any combination thereof” include any combination of A, B, and/or C, and may include multiples of A, multiples of B, or multiples of C. Specifically, combinations such as “at least one of A, B, or C,” “one or more of A, B, or C,” “at least one of A, B, and C,” “one or more of A, B, and C,” and “A, B, C, or any combination thereof” may be A only, B only, C only, A and B, A and C, B and C, or A and B and C, where any such combinations may contain one or more member or members of A, B, or C. All structural and functional equivalents to the elements of the various aspects described throughout this disclosure that are known or later come to be known to those of ordinary skill in the art are expressly incorporated herein by reference and are intended to be encompassed by the claims. Moreover, nothing disclosed herein is intended to be dedicated to the public regardless of whether such disclosure is explicitly recited in the claims. The words “module,” “mechanism,” “element,” “device,” and the like may not be a substitute for the word “means.” As such, no claim element is to be construed as a means plus function unless the element is expressly recited using the phrase “means for.”

## CLAIMS

### WHAT IS CLAIMED IS:

1. A method of procurement demand aggregation, the method comprising:
  - receiving a bipartite graph comprising a first plurality of vertices and a second plurality of vertices, every edge of the bipartite graph connecting a vertex in the first plurality of vertices to a vertex in the second plurality of vertices, the first plurality of vertices corresponding to a plurality of items for procurement, each vertex in the first plurality of vertices representing an item for procurement, the second plurality of vertices corresponding to a plurality of vendors, each vertex in the second plurality of vertices representing a vendor;
  - randomly selecting at least one of a set of items from the plurality of items or a set of vendors from the plurality of vendors;
  - iteratively expanding the at least one of the set of items or the set of vendors to obtain a set of maximal patterns; and
  - aggregating procurement demand based on the set of maximal patterns.
2. The method of claim 1, wherein the set of maximal patterns comprises a set of maximal bicliques.
3. The method of claim 2, wherein each maximal biclique comprises a first set of vertices and a second set of vertices, wherein every vertex in the first set of vertices is connected to every vertex in the second set of vertices in the bipartite graph.
4. The method of claim 3, wherein the first set of vertices correspond to a resulting set of items, the resulting set of items being expanded from the set of items, wherein the second set of vertices correspond to a resulting set of vendors, the resulting set of vendors being expanded from the set of vendors.
5. The method of claim 4, wherein the aggregating of the procurement demand comprises procuring items in the resulting set of items exclusively with the resulting set of vendors.

6. The method of claim 4, wherein the aggregating of the procurement demand comprises procuring only the resulting set of items from the resulting set of vendors.
7. The method of claim 1, wherein the iteratively expanding of the at least one of the set of items or the set of vendors comprises interleaving between adding an item into the set of items and adding a vendor into the set of vendors.
8. The method of claim 7, wherein an item is added into the set of items when a vertex corresponding to the item is connected to every vertex of a set of vertices corresponding to the set of vendors in the bipartite graph.
9. The method of claim 7, wherein a vendor is added into the set of vendors when a vertex corresponding to the vendor is connected to every vertex of a set of vertices corresponding to the set of items in the bipartite graph.
10. An apparatus for procurement demand aggregation, the apparatus comprising:
  - a memory; and
  - at least one processor in at least one computing device coupled to the memory and configured to:
    - receive a bipartite graph comprising a first plurality of vertices and a second plurality of vertices, every edge of the bipartite graph connecting a vertex in the first plurality of vertices to a vertex in the second plurality of vertices, the first plurality of vertices corresponding to a plurality of items for procurement, each vertex in the first plurality of vertices representing an item for procurement, the second plurality of vertices corresponding to a plurality of vendors, each vertex in the second plurality of vertices representing a vendor;
    - randomly select at least one of a set of items from the plurality of items or a set of vendors from the plurality of vendors;
    - iteratively expand the at least one of the set of items or the set of vendors to obtain a set of maximal patterns; and
    - aggregate procurement demand based on the set of maximal patterns.

11. The apparatus of claim 10, wherein the set of maximal patterns comprises a set of maximal bicliques.

12. The apparatus of claim 11, wherein each maximal biclique comprises a first set of vertices and a second set of vertices, wherein every vertex in the first set of vertices is connected to every vertex in the second set of vertices in the bipartite graph.

13. The apparatus of claim 12, wherein the first set of vertices correspond to a resulting set of items, the resulting set of items being expanded from the set of items, wherein the second set of vertices correspond to a resulting set of vendors, the resulting set of vendors being expanded from the set of vendors.

14. The apparatus of claim 13, wherein, to aggregate the procurement demand, the at least one processor in the at least one computing device is configured to procure items in the resulting set of items exclusively with the resulting set of vendors.

15. The apparatus of claim 13, wherein, to aggregate the procurement demand, the at least one processor in the at least one computing device is configured to procure only the resulting set of items from the resulting set of vendors.

16. The apparatus of claim 10, wherein, to iteratively expand the at least one of the set of items or the set of vendors, the at least one processor is configured to interleave between adding an item into the set of items and adding a vendor into the set of vendors.

17. The apparatus of claim 16, wherein an item is added into the set of items when a vertex corresponding to the item is connected to every vertex of a set of vertices corresponding to the set of vendors in the bipartite graph.

18. The apparatus of claim 16, wherein a vendor is added into the set of vendors when a vertex corresponding to the vendor is connected to every vertex of a set of vertices corresponding to the set of items in the bipartite graph.

19. A computer-readable medium storing computer executable code, comprising instructions for:

receiving a bipartite graph comprising a first plurality of vertices and a second plurality of vertices, every edge of the bipartite graph connecting a vertex in the first plurality of vertices to a vertex in the second plurality of vertices, the first plurality of vertices corresponding to a plurality of items for procurement, each vertex in the first plurality of vertices representing an item for procurement, the second plurality of vertices corresponding to a plurality of vendors, each vertex in the second plurality of vertices representing a vendor;

randomly selecting at least one of a set of items from the plurality of items or a set of vendors from the plurality of vendors;

iteratively expanding the at least one of the set of items or the set of vendors to obtain a set of maximal patterns; and

aggregating procurement demand based on the set of maximal patterns.

20. The computer-readable medium of claim 19, wherein the set of maximal patterns contains a set of maximal bicliques.

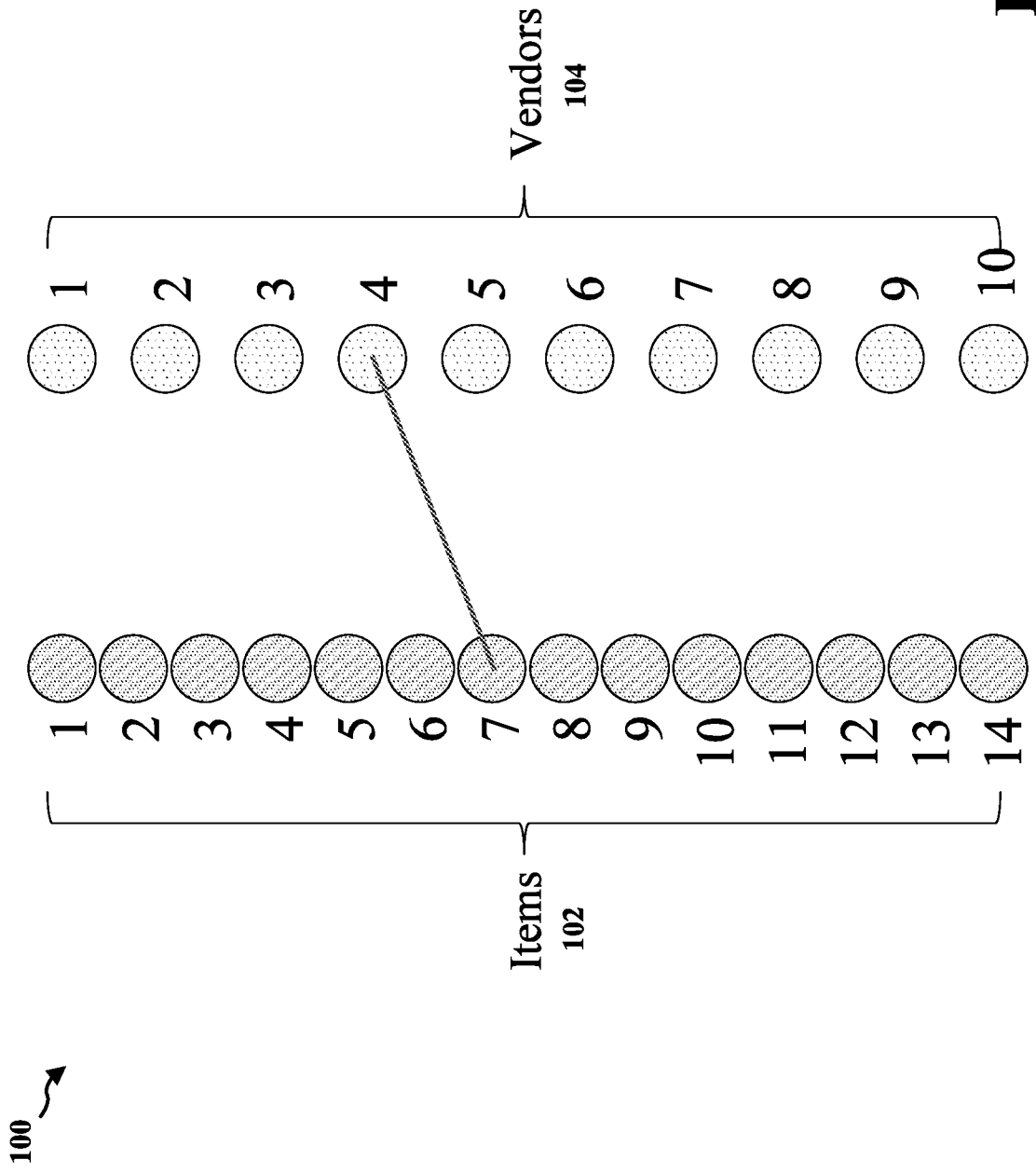


FIG. 1



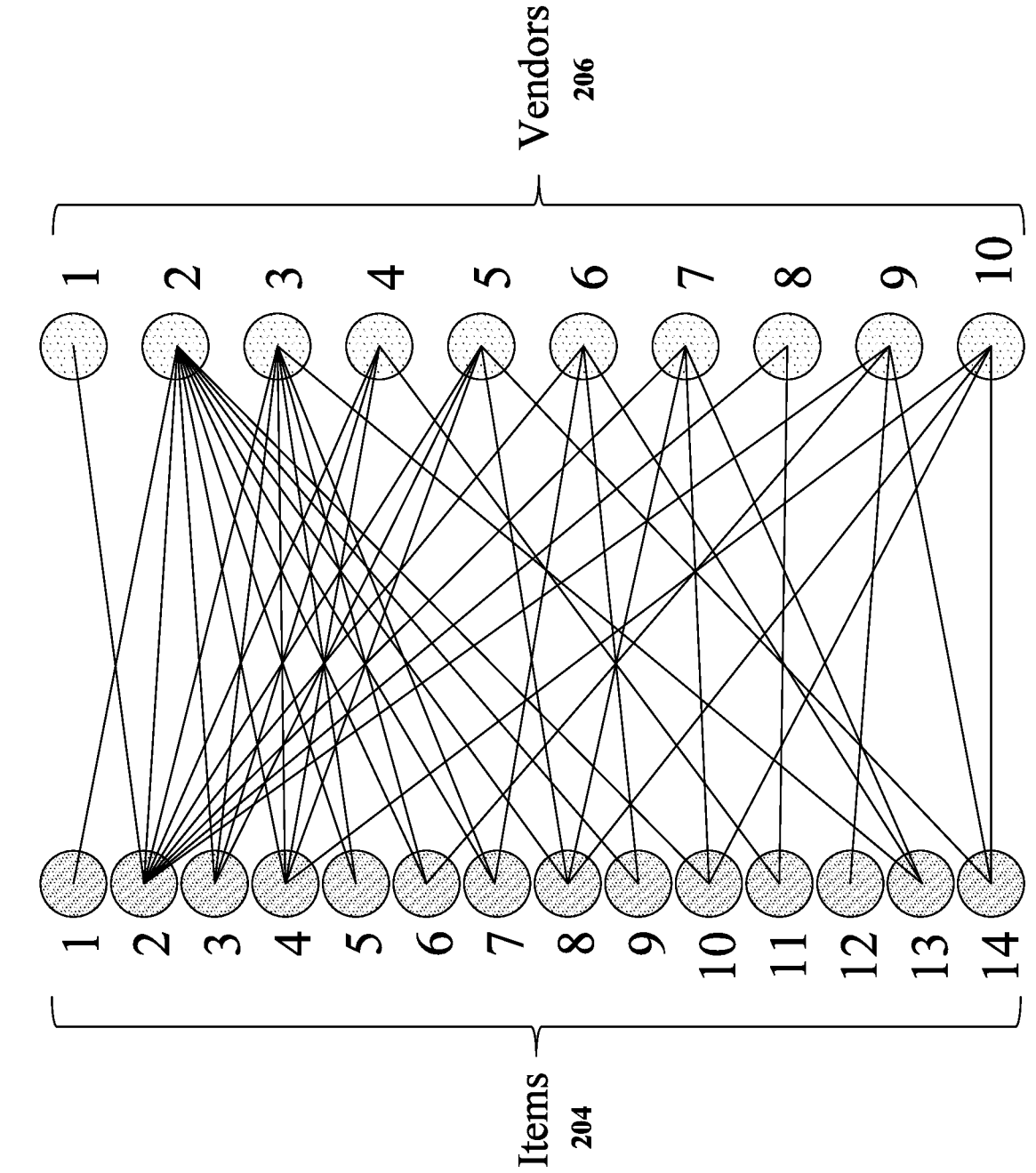


FIG. 2

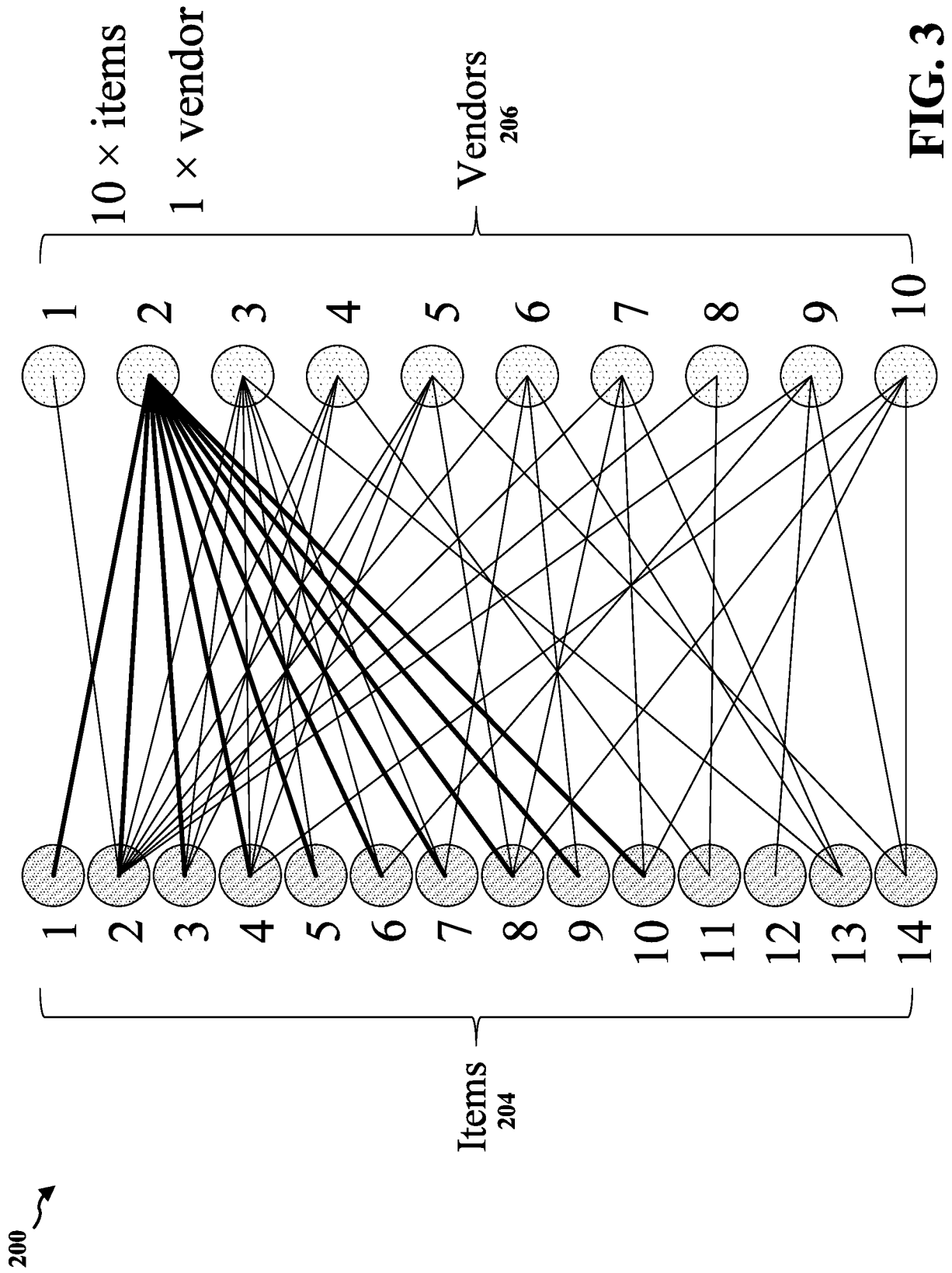
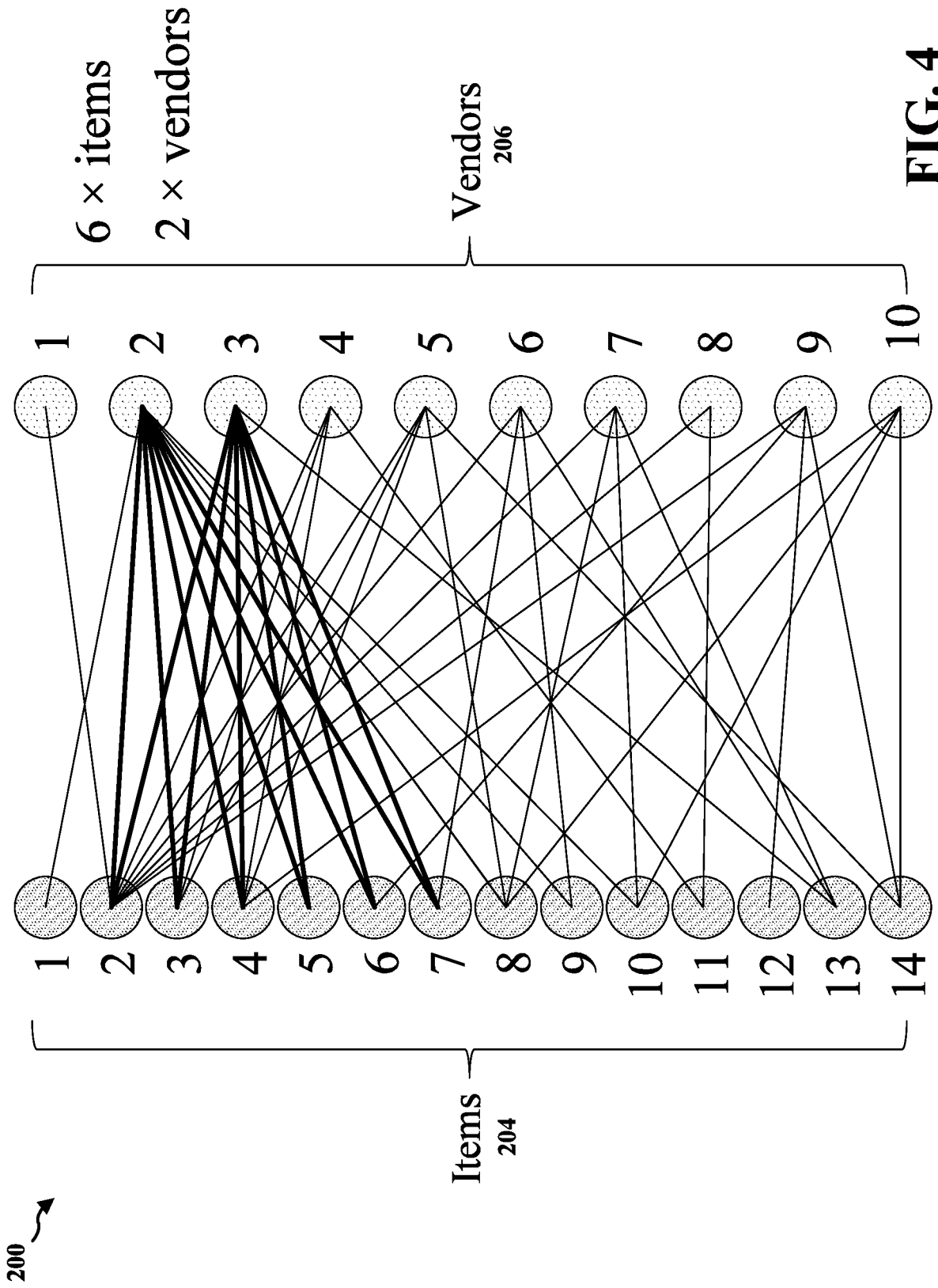


FIG. 3



**FIG. 4**

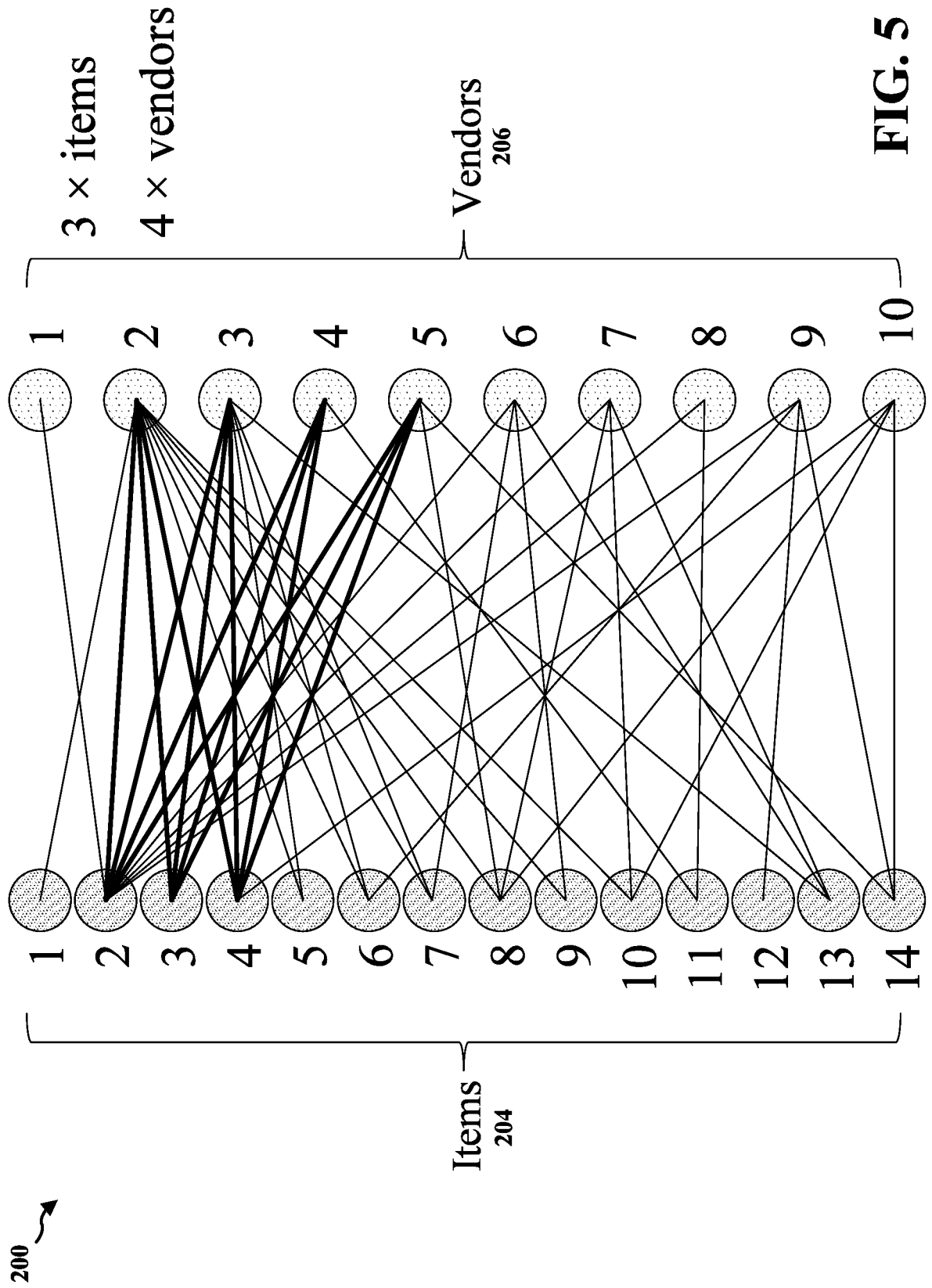


FIG. 5

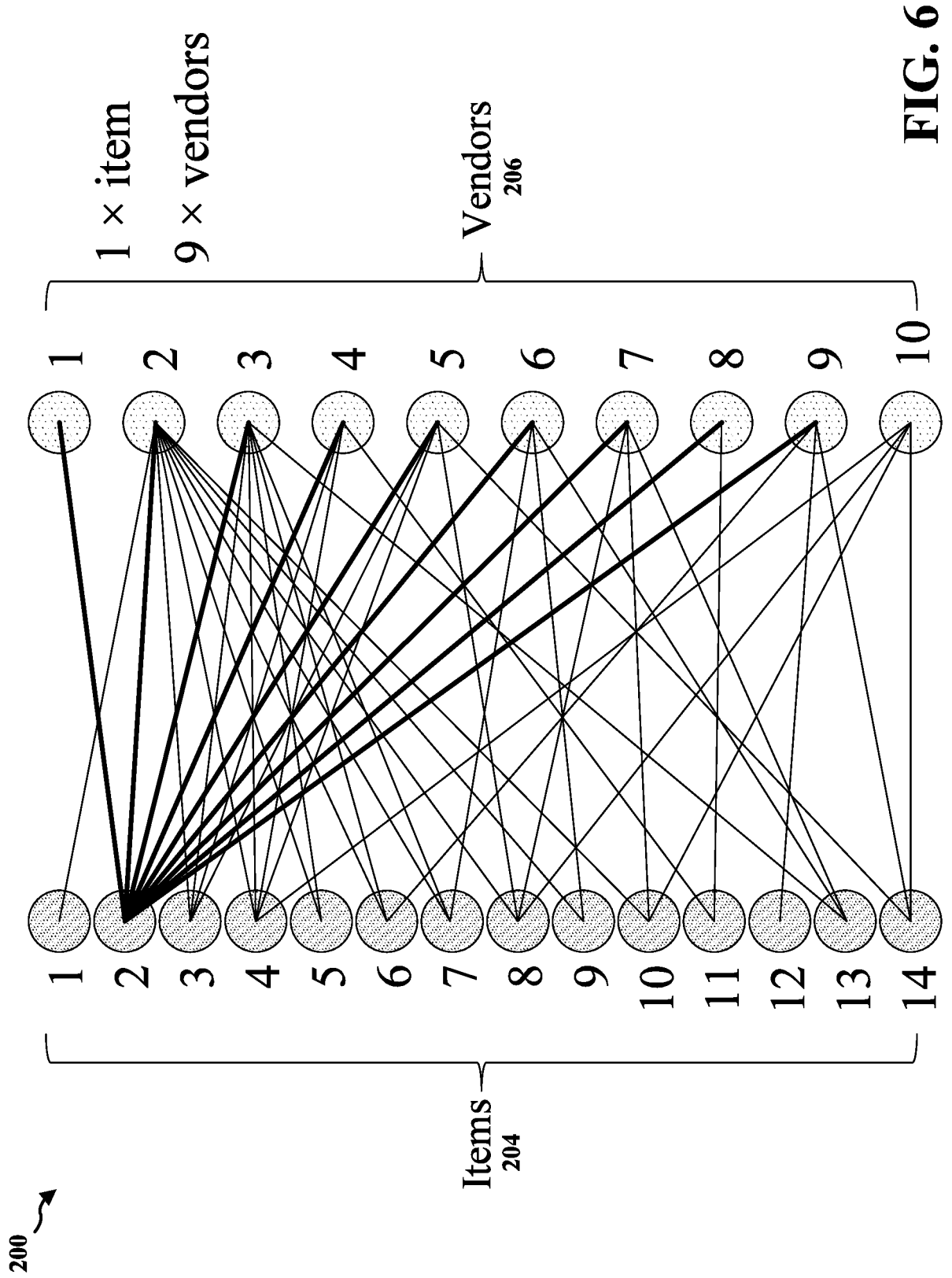


FIG. 6

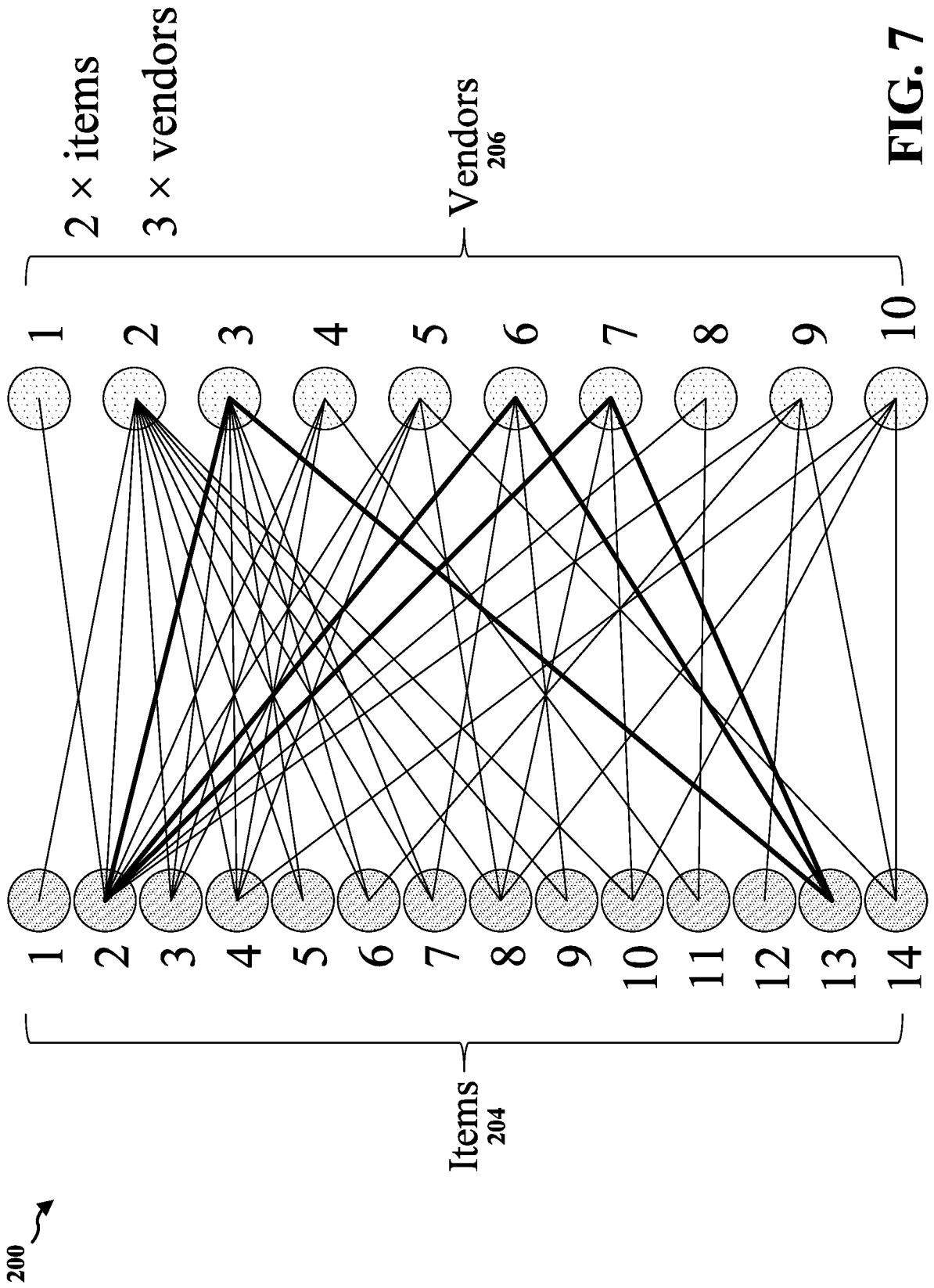


FIG. 7

200 ↗

850 ↗

	$j_1$	$j_2$	$j_3$	$j_4$	$j_5$	$j_6$	$j_7$	
$i_1$	0	1	0	0	0	0	0	0
$i_2$	0	0	0	0	0	0	0	0
$i_3$	0	1	1	0	1	0	0	0
$i_4$	0	0	1	0	1	0	1	0
$i_5$	0	0	0	0	0	0	0	0
$i_6$	0	0	1	0	1	0	0	0
$i_7$	0	0	0	0	0	0	0	0
$i_8$	0	0	0	0	1	0	1	0

800 ↗

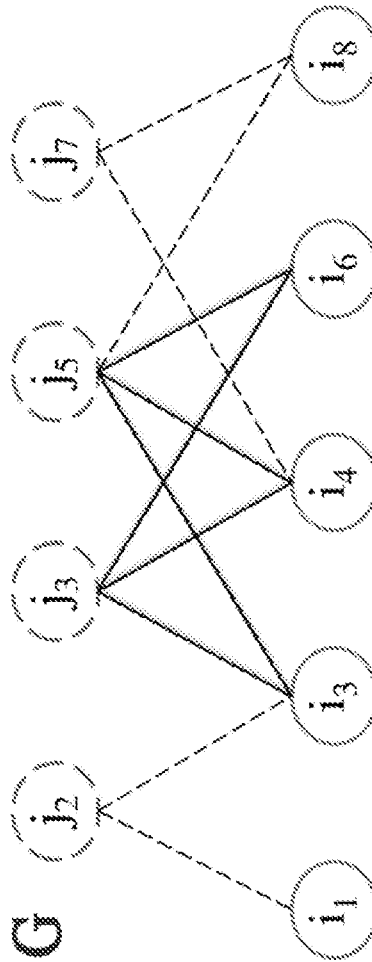
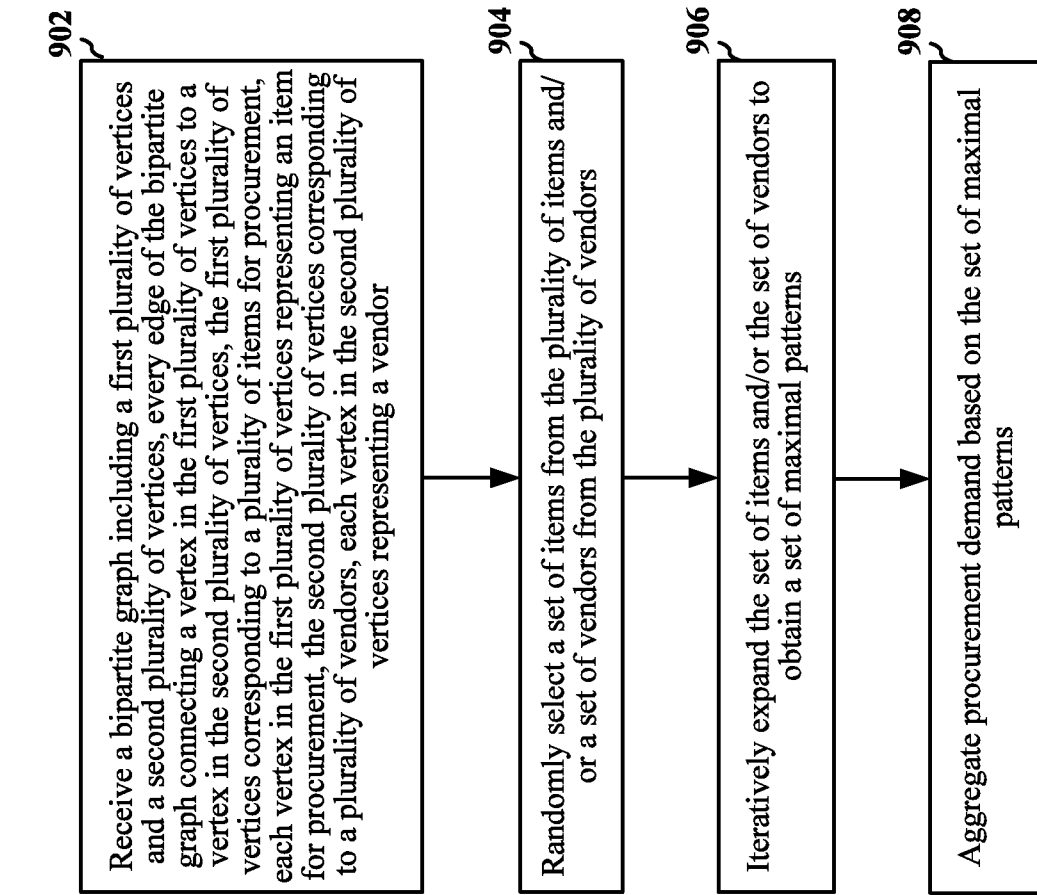


FIG. 8A

FIG. 8B




**FIG. 9**



1000 ↗

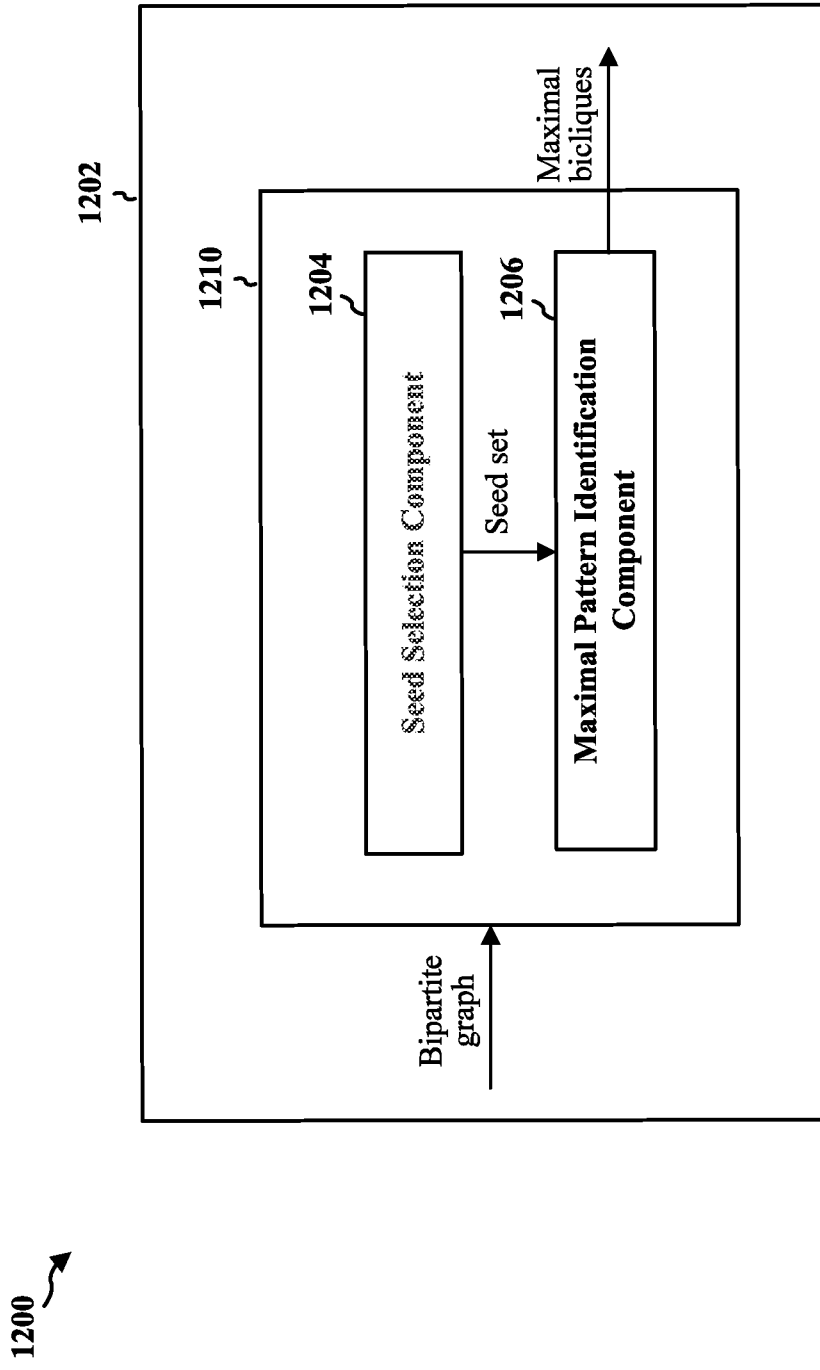
documents \ words	Obama	Football	Iraq	NASDAQ	Winter	FIFA
doc A	1	0	1	0	1	0
doc B	1	0	1	1	0	0
doc C	0	1	0	0	0	1
doc D	0	1	0	0	1	1
doc E	1	0	0	0	0	1

**FIG. 10**

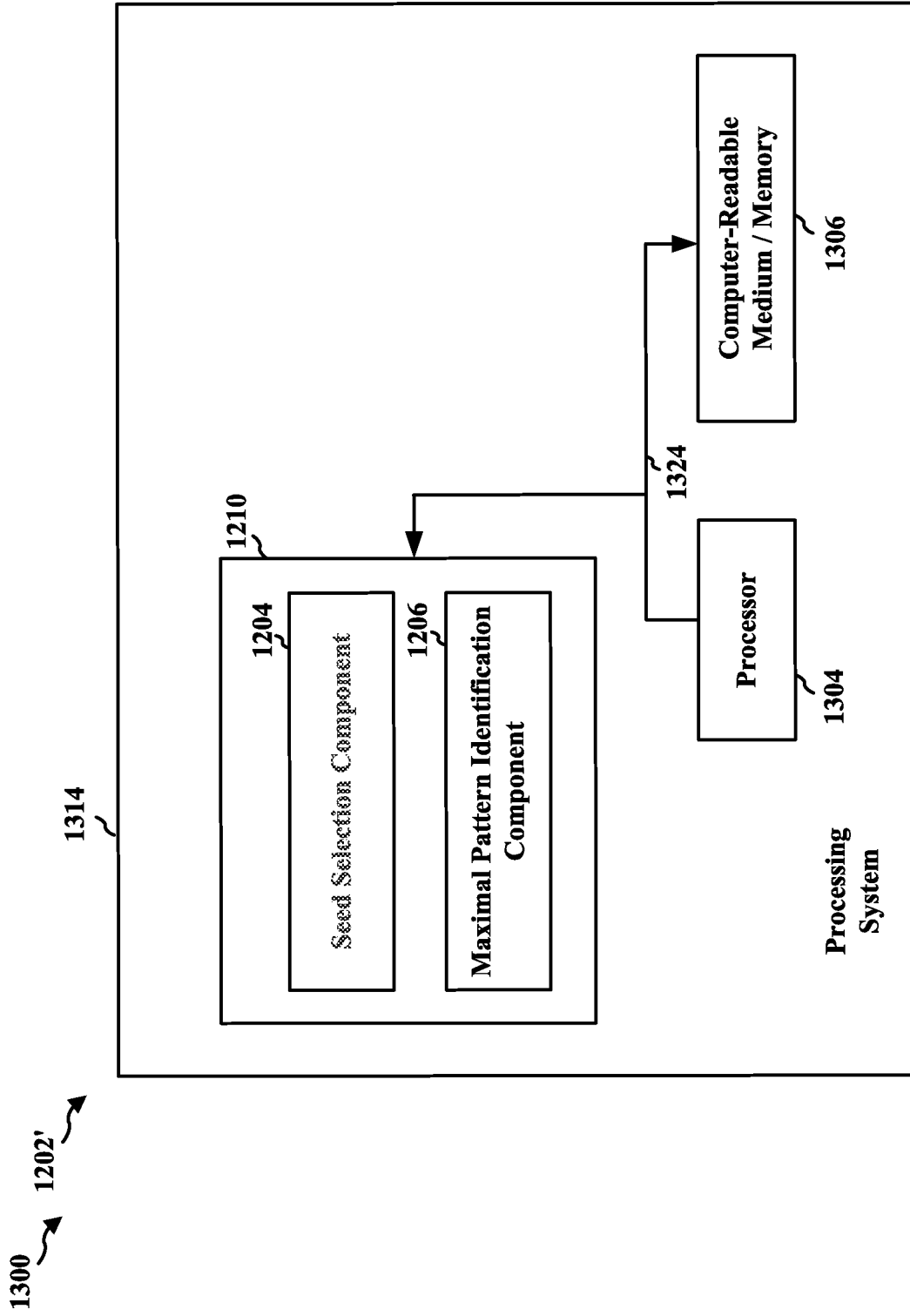
1100 

movies \ users	user_1	user_2	user_3	user_4	user_5	user_6
Iron Man	1	0	1	0	1	0
Thor	1	0	1	1	0	0
Toy Story 3	0	1	0	0	0	1
Frozen	0	1	0	0	1	1
Danger: Diabolik	1	0	0	0	0	1

**FIG. 11**



**FIG. 12**



**FIG. 13**

# INTERNATIONAL SEARCH REPORT

International application No.

**PCT/SG2018/050261**

## A. CLASSIFICATION OF SUBJECT MATTER

**G06Q 10/06 (2012.01) G06F 17/30 (2006.01) G06F 17/10 (2006.01)**

According to International Patent Classification (IPC)

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q, G06F

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)

FAMPAT, IEEE: biclique, graph, vertex, maximal, 图, 顶点, 最大 and other related terms

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MISHRA N. ET AL., On Finding Large Conjunctive Clusters. <i>Learning Theory and Kernel Machines</i> , 27 August 2003, pages 448-462 [Retrieved on 2018-07-30] <DOI:10.1007/978-3-540-45167-9_33> pg. 2 1 <sup>st</sup> three paragraphs, 4 Conjunctive Clustering Algorithm, 5 Finding Approximate $\epsilon$ -Bicliques	1-3, 7-12, 16-20
X	ACUÑA, V. ET AL., Solving the maximum edge biclique packing problem on unbalanced bipartite graphs. <i>Discrete Applied Mathematics</i> , 27 October 2011, pages 2-12 [Retrieved on 2018-07-30] <DOI:10.1016/J.DAM.2011.09.019> pg. 3 3 <sup>rd</sup> paragraph, 2. Formulation for the MEB problem, 4. Computational experiments, 4.2 Computational experiments with MEBP problem	1-20
X	US 9607104 B1 (TURNER, T. ET AL.) 28 March 2017 col. 1 line 55- col. 2 line 3, col. 9 line 39 - col. 10 line 35	1-4, 7-13, 16-20

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

30/07/2018 (day/month/year)

Date of mailing of the international search report

28/08/2018 (day/month/year)

Name and mailing address of the ISA/SG



**Intellectual Property Office of Singapore**  
51 Bras Basah Road  
#01-01 Manulife Centre  
Singapore 189554

Email: pct@ippos.gov.sg

Authorized officer

Yeo Eng Guan (Dr)

IPOS Customer Service Tel. No.: (+65) 6339 8616

**INTERNATIONAL SEARCH REPORT**

International application No.

**PCT/SG201 8/050261****C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2008/073055 A1 (AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH) 19 June 2008 para. [069], Example 2: Processing of a bipartite graph	1-4, 7-13, 16-20

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/SG201 8/050261**

*Note: This Annex lists known patent family members relating to the patent documents cited in this International Search Report. This Authority is in no way liable for these particulars which are merely given for the purpose of information.*

<b>Patent document cited in search report</b>	<b>Publication date</b>	<b>Patent family member(s)</b>	<b>Publication date</b>
US 9607104 B1	28/03/2017	US 2017/0316111 A1 WO 2017/189020 A1	02/11/2017 02/11/2017
WO 2008/073055 A1	19/06/2008	NONE	