



- (51) International Patent Classification:
G06Q 30/02 (2012.01) G06Q 50/00 (2006.01)
- (21) International Application Number:
PCT/US2016/046082
- (22) International Filing Date:
8 August 2016 (08.08.2016)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
62/201,738 6 August 2015 (06.08.2015) US
- (71) Applicant: HRL LABORATORIES, LLC [US/US];
3011 Malibu Canyon Road, Malibu, CA 90265 (US).
- (72) Inventors: XU, Jiejun; 13058 5th Street, Chino, CA 91710 (US). LU, Tsai-Ching; 107 Mcafee Court, Thousand Oaks, CA 91360 (US).
- (74) Agent: TOPE-MCKAY, Cary, R.; Tope-McKay & Associates, 30745 Pacific Coast Highway #420, Malibu, CA 90265 (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).

Published:

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

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR IDENTIFYING USER INTERESTS THROUGH SOCIAL MEDIA

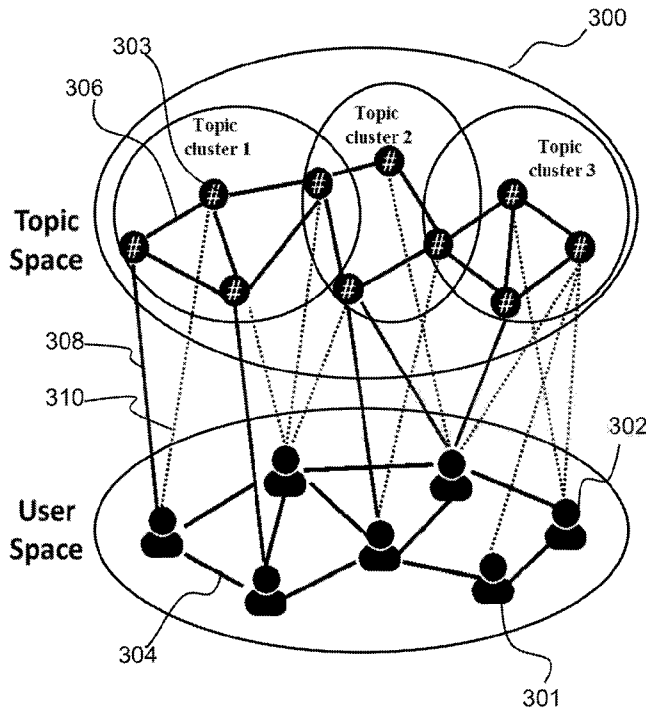


FIG. 3

(57) Abstract: Described is a system for discovering user interests through online social media, and more specifically, to a way of doing so by means of a bi-directional graph model. During operation, the system generates a confidence matrix F based on user interactions and co-occurring tags on a social media platform. The confidence matrix F indicates a likelihood of the users in the social media platform as being interested in a particular topic. Based on such likelihoods, an action can be initiated regarding a particular topic for those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold. For example, the system generates and presents an online advertisement to users regarding a particular topic to those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold.

WO 2017/024316 A1

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

[0001] SYSTEM AND METHOD FOR IDENTIFYING USER INTERESTS
THROUGH SOCIAL MEDIA

[0002] GOVERNMENT RIGHTS

5 [0003] This invention was made with government support under U.S. Government Contract Number D12PC00285 issued by IARPA. The government has certain rights in the invention.

[0004] CROSS-REFERENCE TO RELATED APPLICATIONS

10 [0005] This is a non-provisional patent application of U.S. Provisional Application No. 62/201,738, filed on August 06, 2015, the entirety of which is hereby incorporated by reference.

[0006] BACKGROUND OF INVENTION

15 [0007] (1) Field of Invention

[0008] The present invention relates to a system for discovering user interests and, more specifically, to a system for discovering user interests through online social media using a bi-directional graph model.

20 [0009] (2) Description of Related Art

[00010] There has been a growing interest on discovering user interests and topics from online social media (See the list of Incorporated Literature References, References Nos. 3 and 4). A common approach is to use a vector representation generated from the text of all the posts by a user to represent his interest. Then
25 the similarity between two users can be measured by the similarity scores of feature vectors of the two users. This is also known as the bag-of-words approach. However, this type of approach is quite susceptible to noisy text. This is more severe in the social media context as users are free to publish any posts about their lives, which may not reflect their true topics of interests.
30 Another well-studied method for hidden user topics discovery is the LDA-based

method (Latent Dirichlet Allocation). Some of the studies that have used the LDA-based method can be seen in Literature Reference Nos. 1, 4, and 8. Since LDA relies on the bag-of-words assumption, it suffers similar shortcomings. In addition, the computational requirement for LDA is usually high, and it puts
5 significant bottleneck on the scalability of the approach.

[00011] Another approach to identifying interests is to analyze network topologies as constructed in both social and topic space. In Literature Reference No. 2, the authors looked into communities of users in the reciprocal Twitter follower
10 network and summarized user interests into several categories. In Literature Reference No. 5, the authors proposed a graph-based framework to link entities mentions in tweets posted by a user via modeling the users' topics of interest. One of the commonality of the aforementioned approaches is that both methods focused on only one type of network topology (e.g., either user-centric or topic-
15 centric network) in their analysis, which does not allow for reviewing bi-relational aspects in multiple networks in a unified manner.

[00012] Thus, a continuing need exists for a system that can be used to efficiently
20 and effectively discover user interests through online social media by leveraging topologies of both (user and topic) networks in a unified manner for user interest modeling.

[00013] SUMMARY OF INVENTION

25 [00014] This disclosure provides a system for discovering user interests through online social media. The system includes one or more processors and associated memory (e.g., hard drive, etc.) with instructions encoded thereon. Upon execution of the instructions, the one or more processors perform several operations. For example, during operation, the system generates a confidence
30 matrix **F** based on user interactions and co-occurring tags on a social media

platform (e.g., Twitter, Tumblr, or any other social media platform). The confidence matrix \mathbf{F} indicates a likelihood of the users in the social media platform as being interested in a particular topic. Based on such likelihoods, an action can be initiated regarding a particular topic for those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold. For example, the system can generate and present an online advertisement to users regarding a particular topic to those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold (e.g., greater than 50% or any other predetermined threshold as deemed appropriate by an operator).

[00015] In another aspect, the system performs operations of constructing a user interaction network \mathbf{W} based on a collection of user interactions on a social media platform; constructing a tag co-occurrence network \mathbf{R}_h based on a collection of co-occurring tags on the social media platform; constructing a topic correlation network \mathbf{R} based on the tag co-occurrence network \mathbf{R}_h ; generating a user graph Laplacian \mathbf{L}_g from the user interaction network \mathbf{W} ; generating a topic graph Laplacian \mathbf{L}_c from the topic correlation network \mathbf{R} ; and generating an initial label assignment matrix \mathbf{Y} based on initial known user-topic associations.

[00016] Further, in generating a topic correlation network \mathbf{R} , the topic correlation network is generated by applying Louvain community detection on \mathbf{R}_h .

[00017] In yet another aspect, the rows of confidence matrix \mathbf{F} represent users, and the columns represent topics, such that each entry of the confidence matrix \mathbf{F} indicates the likelihood of a user as being interested in a particular topic.

[00018] Finally, the present invention also includes a computer program product and a computer implemented method. The computer program product includes computer-readable instructions stored on a non-transitory computer-readable medium that are executable by a computer having one or more processors, such that upon execution of the instructions, the one or more processors perform the operations listed herein. Alternatively, the computer implemented method includes an act of causing a computer to execute such instructions and perform the resulting operations.

10 [00019] BRIEF DESCRIPTION OF THE DRAWINGS

[00020] The objects, features and advantages of the present invention will be apparent from the following detailed descriptions of the various aspects of the invention in conjunction with reference to the following drawings, where:

15 [00021] FIG. 1 is a block diagram depicting the components of a system according to various embodiments of the present invention;

[00022] FIG. 2 is an illustration of a computer program product embodying an aspect of the present invention;

20 [00023] FIG. 3 is an illustration of a bi-relational graph for user-interest modeling according to various embodiments of the present invention;

[00024] FIG. 4A is an illustration of an example tag network;

25 [00025] FIG. 4B is an illustration of an example topic network as associated with the tag network depicted in FIG. 4A;

[00026] FIG. 4C is an illustration of an example tag network;

30

[00027] FIG. 4D is an illustration of an example topic network as associated with the tag network depicted in FIG. 4C; and

[00028] FIG. 5 is a flowchart illustrating a process for identifying user interests according to various embodiments of the present invention.

[00029] DETAILED DESCRIPTION

[00030] The present invention relates to a system for discovering user interests and, more specifically, to a system for discovering user interests through online social media using a bi-directional graph model. The following description is presented to enable one of ordinary skill in the art to make and use the invention and to incorporate it in the context of particular applications. Various modifications, as well as a variety of uses in different applications will be readily apparent to those skilled in the art, and the general principles defined herein may be applied to a wide range of aspects. Thus, the present invention is not intended to be limited to the aspects presented, but is to be accorded the widest scope consistent with the principles and novel features disclosed herein.

[00031] In the following detailed description, numerous specific details are set forth in order to provide a more thorough understanding of the present invention. However, it will be apparent to one skilled in the art that the present invention may be practiced without necessarily being limited to these specific details. In other instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

25

[00032] The reader's attention is directed to all papers and documents which are filed concurrently with this specification and which are open to public inspection with this specification, and the contents of all such papers and documents are incorporated herein by reference. All the features disclosed in this specification,

(including any accompanying claims, abstract, and drawings) may be replaced by alternative features serving the same, equivalent or similar purpose, unless expressly stated otherwise. Thus, unless expressly stated otherwise, each feature disclosed is one example only of a generic series of equivalent or similar features.

5

[00033] Furthermore, any element in a claim that does not explicitly state “means for” performing a specified function, or “step for” performing a specific function, is not to be interpreted as a “means” or “step” clause as specified in 35 U.S.C. Section 112, Paragraph 6. In particular, the use of “step of” or “act of” in the claims herein is not intended to invoke the provisions of 35 U.S.C. 112, Paragraph 6.

10

[00034] Before describing the invention in detail, first a list of cited references is provided. Next, a description of the various principal aspects of the present invention is provided. Subsequently, an introduction provides the reader with a general understanding of the present invention. Finally, specific details of various embodiment of the present invention are provided to give an understanding of the specific aspects.

15

20

[00035] (1) List of Incorporated Literature References

[00036] The following references are cited throughout this application. For clarity and convenience, the references are listed herein as a central resource for the reader. The following references are hereby incorporated by reference as though fully set forth herein. The references are cited in the application by referring to the corresponding literature reference number, as follows:

25

1. Harvey, M., Crestani, F., & Carman, M. J. (2013). Building User Profiles from Topic Models for Personalised. *Conference on Information and Knowledge Management (CIKM)*. San Francisco.

2. Java, A., Song, X., Finin, T., & Tseng, B. (2007). Why we twitter: Understanding microblogging usage and communities. *In Proc. 9th WebKDD and 1st SNA-KDD Workshop on Web Mining and Social Network Analysis.*
- 5 3. Michelson, M., & Macskassy, S. A. (2010). Discovering Users' Topics of Interest on Twitter: A First Look. *Proceedings of the fourth workshop on Analytics for noisy unstructured text data (AND)*. Toronto.
4. Ovsjanikov, M., & Chen, Y. (2010). Topic modeling for personalized recommendation of volatile items. *European conference on Machine learning and knowledge discovery in databases: Part II.*
- 10 5. Shen, W., Wang, J., Luo, P., & Wang, M. (2013). Linking Named Entities in Tweets with Knowledge Base via User Interest Modeling. *ACM SIGKDD international conference on Knowledge discovery and data mining*. Chicago.
- 15 6. Wang, H., Huang, H., & Ding, C. (2009). Image annotation using multi-label correlated Green's function. *IEEE 12th International Conference on Computer Vision*. Kyoto.
7. Weng, L., & Menczer, F. (2014). Topicality and Social Impact: Diverse Messages but Focused Messengers. *CoRR abs/1402.5443*.
- 20 8. Xu, J., Compton, R., Lu, T.-C., & Allen, D. (2014). Rolling through Tumblr: Characterizing Behavioral Patterns of the Microblogging Platform. *ACM Web Science*. Bloomington.
9. Xu, J., Jagadeesh, V., & Manjunath, B. (2014). Multi-label Learning with Fused Multimodal Bi-relational Graph. *IEEE Transaction on Multimedia*.
- 25 10. Xu, Z., Lu, R., Xiang, L., & Yang, Q. (2011). Discovering User Interest on Twitter with a Modified Author-Topic Model . *IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*.

11. Jiejun Xu, Tsai-Ching Lu. Toward Precise User-Topic Alignment in Online Social Media. In IEEE International Conference on Big Data (IEEE BigData), Santa Clara, California, 2015.
12. D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Scholkopf. Learning with local and global consistency. In *NIPS*. MIT Press, 2004.
13. X. Zhu. Semi-supervised learning literature survey. In *University of Wisconsin Madison, Computer Sciences TR- 1530*, 2008.
14. R. Compton, D. Jurgens, and D. Allen. Geotagging one hundred million twitter accounts with total variation mini- mization. In *IEEE International Conference on Big Data*, volume abs/1404.7152, 2014.
15. R. Ottoni, D. B. L. Casas, J. P. Pesce, W. M. Jr., C. Wilson, A. Mislove, and V. Almeida. Of pins and tweets: Investigating how users behave across image- and text-based social net- works. In *Proceedings of the Eighth International Conference on Weblogs and Social Media (ICWSM)*, 2014.
16. L. Weng and F. Menczer. Topicality and impact in social media: Diverse messages, focused messengers. *PLoS ONE*, 10(2):e0118410, 02 2015.
17. Y. Yamaguchi, T. Amagasa, and H. Kitagawa. Tag-based user topic discovery using twitter lists. In *International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Kaohsiung, Taiwan, 25-27 July 2011.
18. V. Blondel, J. Guillaume, R. Lambiotte, and E. Mech. Fast unfolding of communities in large networks. *J. Stat. Mech*, page P10008, 2008.

[00037] (2) Principal Aspects

[00038] Various embodiments of the invention include three “principal” aspects.

The first is a system for discovering user interests through online social media, and more specifically, to a way of doing so by means of a bi-directional graph model. The system is typically in the form of a computer system
5 operating software or in the form of a “hard-coded” instruction set. This system may be incorporated into a wide variety of devices that provide different functionalities. The second principal aspect is a method, typically in the form of software, operated using a data processing system (computer). The third principal aspect is a computer program product. The computer program product
10 generally represents computer-readable instructions stored on a non-transitory computer-readable medium such as an optical storage device, e.g., a compact disc (CD) or digital versatile disc (DVD), or a magnetic storage device such as a floppy disk or magnetic tape. Other, non-limiting examples of computer-readable media include hard disks, read-only memory (ROM), and flash-type
15 memories. These aspects will be described in more detail below.

[00039] A block diagram depicting an example of a system (i.e., computer system 100) of the present invention is provided in FIG. 1. The computer system 100 is configured to perform calculations, processes, operations, and/or functions
20 associated with a program or algorithm. In one aspect, certain processes and steps discussed herein are realized as a series of instructions (e.g., software program) that reside within computer readable memory units and are executed by one or more processors of the computer system 100. When executed, the instructions cause the computer system 100 to perform specific actions and
25 exhibit specific behavior, such as described herein.

[00040] The computer system 100 may include an address/data bus 102 that is configured to communicate information. Additionally, one or more data processing units, such as a processor 104 (or processors), are coupled with the

address/data bus 102. The processor 104 is configured to process information and instructions. In an aspect, the processor 104 is a microprocessor.

Alternatively, the processor 104 may be a different type of processor such as a parallel processor, application-specific integrated circuit (ASIC), programmable logic array (PLA), complex programmable logic device (CPLD), or a field programmable gate array (FPGA).

[00041] The computer system 100 is configured to utilize one or more data storage units. The computer system 100 may include a volatile memory unit 106 (e.g., random access memory ("RAM"), static RAM, dynamic RAM, etc.) coupled with the address/data bus 102, wherein a volatile memory unit 106 is configured to store information and instructions for the processor 104. The computer system 100 further may include a non-volatile memory unit 108 (e.g., read-only memory ("ROM"), programmable ROM ("PROM"), erasable programmable ROM ("EPROM"), electrically erasable programmable ROM "EEPROM"), flash memory, etc.) coupled with the address/data bus 102, wherein the non-volatile memory unit 108 is configured to store static information and instructions for the processor 104. Alternatively, the computer system 100 may execute instructions retrieved from an online data storage unit such as in "Cloud" computing. In an aspect, the computer system 100 also may include one or more interfaces, such as an interface 110, coupled with the address/data bus 102. The one or more interfaces are configured to enable the computer system 100 to interface with other electronic devices and computer systems. The communication interfaces implemented by the one or more interfaces may include wireline (e.g., serial cables, modems, network adaptors, etc.) and/or wireless (e.g., wireless modems, wireless network adaptors, etc.) communication technology.

[00042] In one aspect, the computer system 100 may include an input device 112 coupled with the address/data bus 102, wherein the input device 112 is configured to communicate information and command selections to the processor 100. In accordance with one aspect, the input device 112 is an alphanumeric input device, such as a keyboard, that may include alphanumeric and/or function keys. Alternatively, the input device 112 may be an input device other than an alphanumeric input device. In an aspect, the computer system 100 may include a cursor control device 114 coupled with the address/data bus 102, wherein the cursor control device 114 is configured to communicate user input information and/or command selections to the processor 100. In an aspect, the cursor control device 114 is implemented using a device such as a mouse, a track-ball, a track-pad, an optical tracking device, or a touch screen. The foregoing notwithstanding, in an aspect, the cursor control device 114 is directed and/or activated via input from the input device 112, such as in response to the use of special keys and key sequence commands associated with the input device 112. In an alternative aspect, the cursor control device 114 is configured to be directed or guided by voice commands.

[00043] In an aspect, the computer system 100 further may include one or more optional computer usable data storage devices, such as a storage device 116, coupled with the address/data bus 102. The storage device 116 is configured to store information and/or computer executable instructions. In one aspect, the storage device 116 is a storage device such as a magnetic or optical disk drive (e.g., hard disk drive ("HDD"), floppy diskette, compact disk read only memory ("CD-ROM"), digital versatile disk ("DVD")). Pursuant to one aspect, a display device 118 is coupled with the address/data bus 102, wherein the display device 118 is configured to display video and/or graphics. In an aspect, the display device 118 may include a cathode ray tube ("CRT"), liquid crystal display ("LCD"), field emission display ("FED"), plasma display, or any other display

device suitable for displaying video and/or graphic images and alphanumeric characters recognizable to a user.

[00044] The computer system 100 presented herein is an example computing
5 environment in accordance with an aspect. However, the non-limiting example of the computer system 100 is not strictly limited to being a computer system. For example, an aspect provides that the computer system 100 represents a type of data processing analysis that may be used in accordance with various aspects described herein. Moreover, other computing systems may also be
10 implemented. Indeed, the spirit and scope of the present technology is not limited to any single data processing environment. Thus, in an aspect, one or more operations of various aspects of the present technology are controlled or implemented using computer-executable instructions, such as program modules, being executed by a computer. In one implementation, such program modules
15 include routines, programs, objects, components and/or data structures that are configured to perform particular tasks or implement particular abstract data types. In addition, an aspect provides that one or more aspects of the present technology are implemented by utilizing one or more distributed computing environments, such as where tasks are performed by remote processing devices
20 that are linked through a communications network, or such as where various program modules are located in both local and remote computer-storage media including memory-storage devices.

[00045] An illustrative diagram of a computer program product (i.e., storage device)
25 embodying the present invention is depicted in FIG. 2. The computer program product is depicted as floppy disk 200 or an optical disk 202 such as a CD or DVD. However, as mentioned previously, the computer program product generally represents computer-readable instructions stored on any compatible non-transitory computer-readable medium. The term "instructions" as used with

respect to this invention generally indicates a set of operations to be performed on a computer, and may represent pieces of a whole program or individual, separable, software modules. Non-limiting examples of “instruction” include computer program code (source or object code) and “hard-coded” electronics (i.e. computer operations coded into a computer chip). The “instruction” is stored on any non-transitory computer-readable medium, such as in the memory of a computer or on a floppy disk, a CD-ROM, and a flash drive. In either event, the instructions are encoded on a non-transitory computer-readable medium.

10

[00046] (3) Introduction

15

[00047] This disclosure describes a technique to discover user interests from online social media (e.g., Tumblr, etc.) based on a bi-relational graph. Specifically, the graph model contains two sub-structures: a network of users and a network of topics (represented by tags). The former is used to capture user interaction (e.g., reblog, etc.) in the social space, and the latter is used to capture tag co-occurrence in the topic space. Subsequently, the user interest discovery problem is formulated as a multi-label learning problem on the proposed bi-relational graph. Given some initial associations of users and tags, the system can estimate the associations for the rest of the user nodes and tag nodes across the two sub-networks.

20

25

[00048] In some embodiments, a purpose of the system and method is to discover the topics of interest for a particular social media user. This allows for better clustering and search of users based upon their interests. As an example, focus was put on the Tumblr platform with an aim to generate a set of “topic tags” for each user based on what the user posts or reblogs about, and how the user interacts with others. The bi-relational graph representation allows for effective

exploitation of user similarity and topic correlation simultaneously. This contrasts with previous work where the two factors are considered in isolation.

5 [00049] As can be appreciated by those skilled in the art, the system and method can be used, for example, for scientific technology analysis (e.g., to predict future collaboration among users based on their interests), for building user profiles from interest models for personalized or marketing services, and other data collection uses.

10 [00050] (4) Specific Details of Various Embodiments

[00051] As noted above, this disclosure provides a unique bi-relational graph-based model for user interest discovery. This has a broad range of applications, including accurate user profiling and personalized recommendation. Topics or interests are treated as 'labels' in this context, and the problem of user interests discovery is formulated as the multi-label classification problem on graphs. The general process of multi-label classification has been studied extensively in the image annotation domain (see Literature Reference Nos. 6 and 9). The graph-based multi-label classification technique according to various embodiments of the present invention represents a transductive semi-supervised learning process that diffuses the label information (i.e., interests, topics) from a small subset of users to the rest in the graph. Through careful construction of a bi-relational graph, user similarity and label correlation are exploited jointly in the diffusion process. Currently analysis is being conducted on the Tumblr data. The choice of platform is inspired by Literature Reference No. 8, as it shows Tumblr is heavily driven by user interests.

15
20
25

[00052] (4.1) Formulation

[00053] An example construction of the bi-relational graph is shown in FIG. 3. As shown, there are at least two networks, a topic space 300 and a social or user

space 302. User space solid lines 304 indicate affinity relationship among user nodes 301 (i.e., user similarity), and topic space solid lines 306 indicate affinity relationship among topic nodes 303 (i.e., topic correlation). The cross network solid lines 308 across the two networks denote the initial label (i.e.,
5 topic/interest) association, and the cross network dotted lines 310 denote the label assignment to be estimated. Thus, solid lines 304 and 306 within each of the two sub-graphs indicate social homophily relation and topic correlation, while the solid black lines 308 across two sub-graphs denote the initial known user-topic assignments.

10

[00054] In terms of classification, most existing graph-based semi-supervised learning frameworks attempt to minimize a cost function which takes into account two properties: smoothness of the data (i.e., user) graph and the deviation of initial assignments. Here a third property is introduced into the regularization framework, smoothness in the label (topic) graph. The process for
15 constructing the graph is provided in further detail below.

[00055] Suppose that there is a collection of N users $U = \{u_1, u_2, \dots, u_N\}$ and K topics of interest $= \{t_1, t_2, \dots, t_K\}$. Assuming that some of the users in U are
20 (partially) labeled for their topics of interest, a goal is to predict the topics of interest for the remaining unlabeled users u_i in the collection with the label subset $l_i \subseteq T$.

[00056] The graph-based multi-label learning technique according to various
25 embodiments of the present invention represents a transductive semi-supervised learning process that diffuses the label information from a small subset of nodes to the rest based on the intrinsic graph structure. Note that the terms “topics of interest” and “labels” may be used interchangeably. The basic step in conventional graph-based learning is to construct a graph where vertices
30 represent data instances and edge weights represent affinity between them. The

key to graph-based multi-label learning is the prior assumption of consistency: nearby data instances or data instances that lie on the same structure are likely to share the same label. Generally it is formulated in a regularization framework as follows:

5

$$\mathbf{F}^* = \operatorname{argmin}_{\mathbf{F}} \{ \Omega_{\text{smooth}}(\mathbf{F}) + \Omega_{\text{prior}}(\mathbf{F}) \},$$
 where \mathbf{F} is the to-be-learned matrix containing the label assignments of the graph nodes.

[00057] The first term corresponds to a loss function which reflects the consistency assumption by imposing the smoothness constraint on the neighboring labels. The second term is a regularizer for the fitting constraint, which means that initial assigned labels should be changed as little as possible (see Literature Reference Nos. 12 and 13).

15 [00058] In the context of the present system, data instances correspond to users, and their affinity can be characterized by the social interactions or computed based on any other similarity measures such as user demographics and geolocations. Note that the first term of the above regularization framework is in accordance to the social homophily assumption. In addition to the user graph, the
20 conventional graph-based learning framework is augmented by introducing a new graph to emphasize the correlation among topics. In conjunction, the two graphs make up the bi-relational graph model as illustrated in FIG. 3.

[00059] Given label association for a small set of data (i.e. initial assignments
25 between user nodes and topic nodes), a goal is to estimate the hidden links between the two types of node in the remaining part. Such a model allows for effective exploitation of the smoothness constraints on both sub-graphs as well as the interplay between them.

30 [00060] (4.2) Graph Construction

[00061] The construction of the user graph in this work is based on the primary interactions in social media platforms. For instance, one can focus on the *@mention* action in Twitter. Twitter users often “@mention” each other by prepending an “@” to the mentioned users name. Although there are other types of interaction such as *like* and *retweet*, *@mention* has been shown to indicate social ties (see Literature Reference No. 14). Similarly, the system focuses on the *reblog* action on Tumblr (which is the official mechanism to republish the content of another user’s posts in Tumblr), as it has been shown to indicate common hobbies and interests among users (see Literature Reference No. 8). In order to obtain strong social ties, in various embodiments, the systems focus on *@mention* and *reblog* that are reciprocated (note although the *@mention* and *reblog* are used, they are provided as non-limiting examples and the system is not limited to such cues). In other words, a bidirectional edge is only introduced between user i and j if u_i *@mentions* (*reblogs*) u_j and u_j *@mentions* (*reblogs*) u_i at some point in time. The weight of an edge is determined based on the minimum number of reciprocated frequency (i.e., *@mentions* (*reblogs*)) between the two users.

[00062] The construction of the topic graph is based on the co-occurrence among topics. However, topics are usually not explicitly defined in microblogging platforms, with a few rare exceptions as in Literature Reference No. 15. Alternatively, the system can be devised to consider user defined tags as channels to study topics in social media. This strategy has been studied in existing literature (see Literature Reference Nos. 16 and 17).

[00063] For illustrative purposes, FIGs. 4A and 4C show snapshots of tag co-occurrence networks constructed with Twitter and Tumblr data, while FIGs. 4B and 4D depict corresponding topic networks, respectively. As an example, the

size and/or color of a node is proportional to its degree; the width of an edge is proportion to the co-occurrence frequency. The “degree” of a node in the network is the number of connections it has to other nodes. For example, the nodes can be illustrated such that their color changes gradually from, for
5 example, green to purple to white. In this non-limiting example, the greener a node is, the higher degree (i.e., connected to many others, or center node) it is; on the other hand, white/purple colors indicate the corresponding nodes are less connected to others (i.e., peripheral nodes). As another example, the larger the
10 node is, the higher degree it is, while smaller nodes indicate that they are less connected.

[00064] As can be seen, the tags in each of the networks are related to a single coherent topic. For instance, the tags in the Twitter network are related to
15 “Marvel”, which is a popular comic publisher. The nodes in the graph include comic titles (and their name variations), comic characters, and cast members of comic book adaptation movies. The same observation can be seen from the sample tag network derived from the Tumblr platform, where nodes related to “football” often co-occur together.

20 [00065] Since tags on social media sites are invented autonomously by millions of content generators, there is no predefined consensus on how to group them into topics. Multiple duplicate tags may be developed to represent the same event, theme, or object. For instance, #loki, #thor, #odin, #asgard are all related to the
25 fictional characters in a Marvel movie; #worldcup2014, #brazilwc2014, #wc2014, #fifawc14 are all about the major soccer event that occurred in June 2014. In order to reduce duplication and noise, raw tags can be aggregated and abstracted to a more general level clusters of semantically related tags, referred to as topics. These clusters are detected by finding communities in the tag-based
30 co-occurrence network. For example, the Louvain community detection method

(see Literature Reference No. 18) can be used to identify the topic clusters because of its computational efficiency. The basic idea of the Louvain method is to repeatedly find small communities by optimizing modularity locally on all nodes, then group each of these small communities into a single node. FIGs. 4B and 4D show examples of the resulting topic graphs. Strong topic locality can be observed.

[00066] (4.3) Multi-label Learning on the Bi-relational Graph

[00067] As mentioned above, conventional graph-based learning framework minimizes a cost function with two terms. Introducing a new topic graph to the framework leads to the updated regularization framework regarding \mathbf{F} as follows:

[00068] Let \mathbf{W} be a $N \times N$ affinity matrix denoting the data graph constructed with N data points (users), and \mathbf{R} be a $K \times K$ affinity matrix denoting the label graph constructed for K topics. The frequency-based weights in \mathbf{W} and \mathbf{R} are normalized to the same dynamic range. Let $\mathbf{F} = (F_1, \dots, F_N)^T = (C_1, \dots, C_K)$ be a $N \times K$ matrix denoting the final association between every user topic pairs. (C_1, \dots, C_K) are the columns of \mathbf{F} , corresponding to the \mathbf{K} labels. Similarly let $\mathbf{Y} = (Y_1, \dots, Y_N)^T$ be an $N \times K$ matrix denoting the initial label assignments. Each Y_{ij} has 1 or 0 as the possible values: 1 if user i is labeled with topic j , 0 if it is unlabeled. The overall cost function is expressed as:

$$\Omega(\mathbf{F}) = \underbrace{\frac{1}{2} \eta \sum_{i,j} W_{ij} \left\| \frac{F_i}{\sqrt{D_{ii}}} - \frac{F_j}{\sqrt{D_{jj}}} \right\|^2}_{\text{Smoothness on user graph}} + \underbrace{\frac{1}{2} \mu \sum_{i,j} W_{ij} \left\| \frac{C_i}{\sqrt{D_{ii}}} - \frac{C_j}{\sqrt{D_{jj}}} \right\|^2}_{\text{Smoothness on topic graph}} + \underbrace{\sum_i \|F_i - Y_i\|^2}_{\text{prior constraint}}, \tag{1}$$

Smoothness on user graph Smoothness on topic graph prior constraint

where D and D' are both diagonal matrix whose (i, i) entries equal to the sum of

the i -th row of W and R , i.e., $D_{ii} = \sum_{j=1}^N W_{ij}$ and $D'_{ii} = \sum_{j=1}^K R_{ij}$. The solution of F can be found by minimizing the above cost function.

5

[00069] The first term of the above equation (1) is the smoothness constraint on the user graph. Minimizing it means neighboring vertices should share similar labels. For instance, if two users are close to each other based on their frequent re-blog activities (e.g., *@mention*, *reblog*), they will probably have common interests (thus with similar labels). The second term is the smoothness constraint on the topic or label graph. Minimizing it means neighboring vertices should include similar users. For instance, if two topics are highly correlated with each other, then they are likely to be of interest to the same set of users. The third term indicates that the initially known user topic pairs should be changed as little a possible.

10

[00070] η and μ are two constants controlling the trade-off of the regularization terms. If μ is set to zero, it means to ignore the correlation among topics, and the formulation is reduced to traditional multi-label learning on a single (social) graph.

15

[00071] The first term of the above cost function can be rewritten as:

$$\begin{aligned} & \frac{1}{2} \eta \sum_{i,j} W_{ij} \left\| \frac{F_i}{\sqrt{D_{ii}}} - \frac{F_j}{\sqrt{D_{jj}}} \right\|^2 \\ &= \frac{1}{2} \eta \left(\sum_{i=1}^N F_i^T F_i + \sum_{j=1}^N F_j^T F_j - 2 \sum_{i,j=1}^N \frac{W_{ij} F_i^T F_j}{\sqrt{D_i D_j}} \right) \\ &= \eta \left(\sum_{i=1}^N F_i^T F_i - \sum_{i,j=1}^N \frac{W_{ij} F_i^T F_j}{\sqrt{D_i D_j}} \right) \\ &= \eta \text{tr}(F^T (I - D^{-1/2} W D^{-1/2}) F) \end{aligned} \quad 25$$

(2)

5

[00072] Similarly, the second and third terms of the cost function can be rewritten in a matrix form with several algebraic steps. Thus, the original cost function above can be written in a more concise form as:

10

$$\Omega(\mathbf{F}) = \eta \text{tr}(\mathbf{F}^T \mathbf{L}_g \mathbf{F}) + \mu \text{tr}(\mathbf{F} \mathbf{L}_c \mathbf{F}^T) + \text{tr}((\mathbf{F} - \mathbf{Y})^T (\mathbf{F} - \mathbf{Y})), \quad (3)$$

where $\mathbf{L}_g = \mathbf{I} - \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}$ and $\mathbf{L}_c = \mathbf{I} - \mathbf{D}'^{-1/2} \mathbf{R} \mathbf{D}'^{-1/2}$. They are the Normalized Laplacian of user graph and topic graph respectively.

15

[00073] By applying the following matrix properties:

$$\frac{\partial \text{tr}(\mathbf{X}^T \mathbf{A} \mathbf{X})}{\partial \mathbf{X}} = (\mathbf{A} + \mathbf{A}^T) \mathbf{X}, \quad \frac{\partial \text{tr}(\mathbf{X} \mathbf{A} \mathbf{X}^T)}{\partial \mathbf{X}} = \mathbf{X} (\mathbf{A} + \mathbf{A}^T), \quad (4)$$

20

the equation can be differentiated with respect to \mathbf{F} as follows:

$$\frac{\partial \Omega(\mathbf{F})}{\partial \mathbf{F}} = \eta \mathbf{L}_g \mathbf{F} + \mu \mathbf{F} \mathbf{L}_c + (\mathbf{F} - \mathbf{Y}). \quad (5)$$

[00074] This is because both \mathbf{L}_g and \mathbf{L}_c are symmetric matrices. The solution for \mathbf{F}

can be obtained by requiring $\frac{\partial \Omega(\mathbf{F})}{\partial \mathbf{F}}$ to zero. With some simple algebraic steps, it becomes apparent that

$$(\eta \mathbf{L}_g + \mathbf{I})\mathbf{F} + \mu \mathbf{F} \mathbf{L}_c = \mathbf{Y},$$

5 which is essentially a matrix equation with the form of $\mathbf{AX} + \mathbf{XB} = \mathbf{C}$. Solution to the equation can be easily obtained from existing numerical libraries, such as Linear Algebra PACKage (LAPACK) and Matlab. LAPACK is a software package provided by Univ. of Tennessee; Univ. of California, Berkeley; Univ. of Colorado Denver; and NAG Ltd.. Note that F_{ij} is essentially a confidence

10 value of user u_i being interested in topic t_j .

[00075] Once \mathbf{F} or F_{ij} is found, labels can be assigned (i.e., topics of interest) to users using simple thresholds. Basically a user with a higher value can be assigned to the corresponding topic with higher confidence. The overall process for

15 inferring user's topics of interest is summarized in the Algorithm below.

[00076] Input: Set $\mathbf{E} = \{(e_i^a, e_i^b, w_i) | i = 1, 2, \dots, N_{|\mathbf{E}|}\}$ containing the collection of user interactions, e.g., e_i^a reblogs e_i^b for w_i times. Set $\mathbf{H} = \{(h_j^1, h_j^2, \dots, h_j^{m_j}) | j = 1, 2, \dots, N_{|\mathbf{H}|}\}$ containing the collection of co-occurring tags, e.g., $h_j^1, \dots, h_j^{m_j}$ are

20 associated to the j^{th} social media post. Output: Confidence matrix \mathbf{F} , where F_{ij} is indicates the probability of user u_i interested in topic t_j . As shown in FIG. 5, the algorithm proceeds according to the following steps:

1. Construct (or generate) a user interaction network \mathbf{W} 500 from \mathbf{E} .
- 25 2. Construct tag co-occurrence network \mathbf{R}_h 502 from \mathbf{H} .

3. Construct topic correlation network \mathbf{R} 504 by applying Louvain community detection on \mathbf{R}_h .
4. Compute user graph Laplacian \mathbf{L}_g 506 from \mathbf{W} .
5. Compute topic graph Laplacian \mathbf{L}_c 508 from \mathbf{R} .
- 5 6. Compute \mathbf{Y} 510 based on the initial known user-topic associations.
7. Compute \mathbf{F} 512 by minimizing the cost function in Eq. (3), i.e., solve the following matrix equation:

$$\eta \mathbf{L}_g \mathbf{F} + \mu \mathbf{F} \mathbf{L}_c + (\mathbf{F} - \mathbf{Y}) = 0.$$
- 10 8. Return the most confidence user-topic pairs by sorting and ranking entries in \mathbf{F} .

[00077] The system can then be used to characterize social media users' topics of interest by estimating the \mathbf{F} matrix using information derived from online social network as described in the above algorithm. The rows of the \mathbf{F} matrix represent users, and the columns represent topics. Each entry of the matrix indicates the likelihood of a user interested in a particular topic.

[00078] This invention is important because the research outcome allows for better clustering and search of online users, and it has direct impacts on personalization, recommendation, and many other aspects of online experience enhancement. The system has been applied on characterizing online users' topics of interested on two social media platforms -- Twitter and Tumblr. In both cases, substantial improvements were obtained compared to existing methods. For example, the process as described herein is supported by the experimental studies as described in Literature Reference No. 11.

[00079] As noted above, there are several applications in which the system can be implemented by automatically initiating an action regarding a particular topic

for those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold (e.g., greater than 50% likelihood). For example, based on the user-topic pairs and ranked entries in **F**, the system can then be used to market services or products to particular individuals based on their interests, such as by automatically generating and presenting an online advertisement 514 to users regarding a particular topic to those users whose likelihood of being interested in the particular topic exceeds the predetermined threshold. As a non-limiting example, if a particular user has a high likelihood of interest (e.g., greater than 50%) in topics associated with Marvel characters, then banner ads for upcoming movies associated with cartoon characters can be presented through the internet to the user. As another non-limiting example, if a particular user has a high likelihood of interest in topics associated with football games, such as the World Cup, then banner ads for travel packages to various football games can be presented to the user (e.g., a banner ad for flights and hotel accommodations to the host city of an international football event). As yet another non-limiting example, if a user has a high likelihood of interest in topics associated with automobile performance, then mailings or banner ads can be presented to the user regarding new vehicles.

20 [00080] Finally, while this invention has been described in terms of several embodiments, one of ordinary skill in the art will readily recognize that the invention may have other applications in other environments. It should be noted that many embodiments and implementations are possible. Further, the following claims are in no way intended to limit the scope of the present invention to the specific embodiments described above. In addition, any recitation of “means for” is intended to evoke a means-plus-function reading of an element and a claim, whereas, any elements that do not specifically use the recitation “means for”, are not intended to be read as means-plus-function elements, even if the claim otherwise includes the word “means”. Further, while

particular method steps have been recited in a particular order, the method steps may occur in any desired order and fall within the scope of the present invention.

CLAIMS

What is claimed is:

1. A system for identifying user interests through social media, the system
5 comprising:
 - one or more processors and a memory, the memory being a non-transitory computer-readable medium having executable instructions encoded thereon, such that upon execution of the instructions, the one or more processors perform operations of:
 - 10 generating a confidence matrix \mathbf{F} based on user interactions and co-occurring tags on a social media platform, the confidence matrix \mathbf{F} indicating a likelihood of the users in the social media platform as being interested in a particular topic; and
 - initiating an action regarding a particular topic for those users
15 whose likelihood of being interested in the particular topic exceeds a predetermined threshold.
2. The system as set forth in Claim 1, further comprising operations of:
 - 20 constructing a user interaction network \mathbf{W} based on a collection of user interactions on a social media platform;
 - constructing a tag co-occurrence network \mathbf{R}_h based on a collection of co-occurring tags on the social media platform;
 - constructing a topic correlation network \mathbf{R} based on the tag co-occurrence network \mathbf{R}_h ;
 - 25 generating a user graph Laplacian \mathbf{L}_g from the user interaction network \mathbf{W} ;
 - generating a topic graph Laplacian \mathbf{L}_c from the topic correlation network \mathbf{R} ; and

generating an initial label assignment matrix \mathbf{Y} based on initial known user-topic associations.

3. The system as set forth in Claim 2, wherein in generating a topic correlation network \mathbf{R} , the topic correlation network is generated by applying Louvain community detection on \mathbf{R}_h .
5
4. The system as set forth in Claim 3, wherein the rows of confidence matrix \mathbf{F} represent users, and the columns represent topics, such that each entry of the confidence matrix \mathbf{F} indicates the likelihood of a user as being interested in a particular topic.
10
5. The system as set forth in Claim 4, wherein initiating an action further comprises operations of generating and presenting an online advertisement to users regarding a particular topic to those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold.
15
6. The system as set forth in Claim 1, wherein the rows of confidence matrix \mathbf{F} represent users, and the columns represent topics, such that each entry of the confidence matrix \mathbf{F} indicates the likelihood of a user as being interested in a particular topic.
20
7. The system as set forth in Claim 1, wherein initiating an action further comprises operations of generating and presenting an online advertisement to users regarding a particular topic to those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold.
25
8. A method for identifying user interests through social media, the method comprising acts of:

- generating, with one or more processors, a confidence matrix \mathbf{F} based on user interactions and co-occurring tags on a social media platform, the confidence matrix \mathbf{F} indicating a likelihood of the users in the social media platform as being interested in a particular topic; and
- 5 initiating, with the one or more processors, an action regarding a particular topic for those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold.
9. The method as set forth in Claim 8, further comprising operations of:
- 10 constructing a user interaction network \mathbf{W} based on a collection of user interactions on a social media platform;
- constructing a tag co-occurrence network \mathbf{R}_h based on a collection of co-occurring tags on the social media platform;
- constructing a topic correlation network \mathbf{R} based on the tag co-
- 15 occurrence network \mathbf{R}_h ;
- generating a user graph Laplacian \mathbf{L}_g from the user interaction network \mathbf{W} ;
- generating a topic graph Laplacian \mathbf{L}_c from the topic correlation network \mathbf{R} ; and
- 20 generating an initial label assignment matrix \mathbf{Y} based on initial known user-topic associations.
10. The method as set forth in Claim 9, wherein in generating a topic correlation network \mathbf{R} , the topic correlation network is generated by applying Louvain
- 25 community detection on \mathbf{R}_h .
11. The method as set forth in Claim 10, wherein the rows of confidence matrix \mathbf{F} represent users, and the columns represent topics, such that each entry of the

confidence matrix **F** indicates the likelihood of a user as being interested in a particular topic.

- 5 12. The method as set forth in Claim 11, wherein initiating an action further comprises acts of generating and presenting an online advertisement to users regarding a particular topic to those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold.
- 10 13. The method as set forth in Claim 8, wherein the rows of confidence matrix **F** represent users, and the columns represent topics, such that each entry of the confidence matrix **F** indicates the likelihood of a user as being interested in a particular topic.
- 15 14. The method as set forth in Claim 8, wherein initiating an action further comprises acts of generating and presenting an online advertisement to users regarding a particular topic to those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold.
- 20 15. A computer program product for identifying user interests through social media, the computer program product comprising:
a non-transitory computer-readable medium having executable instructions encoded thereon, such that upon execution of the instructions by one or more processors, the one or more processors perform operations of:
25 generating a confidence matrix **F** based on user interactions and co-occurring tags on a social media platform, the confidence matrix **F** indicating a likelihood of the users in the social media platform as being interested in a particular topic; and

initiating an action regarding a particular topic for those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold.

5 16. The computer program product as set forth in Claim 15, further comprising operations of:

constructing a user interaction network \mathbf{W} based on a collection of user interactions on a social media platform;

10 constructing a tag co-occurrence network \mathbf{R}_h based on a collection of co-occurring tags on the social media platform;

constructing a topic correlation network \mathbf{R} based on the tag co-occurrence network \mathbf{R}_h ;

generating a user graph Laplacian \mathbf{L}_g from the user interaction network \mathbf{W} ;

15 generating a topic graph Laplacian \mathbf{L}_c from the topic correlation network \mathbf{R} ;

generating an initial label assignment matrix \mathbf{Y} based on initial known user-topic associations.

20 17. The computer program product as set forth in Claim 16, wherein in generating a topic correlation network \mathbf{R} , the topic correlation network is generated by applying Louvain community detection on \mathbf{R}_h .

25 18. The computer program product as set forth in Claim 17, wherein the rows of confidence matrix \mathbf{F} represent users, and the columns represent topics, such that each entry of the confidence matrix \mathbf{F} indicates the likelihood of a user as being interested in a particular topic.

- 5 19. The computer program product as set forth in Claim 18, wherein initiating an action further comprises operations of generating and presenting an online advertisement to users regarding a particular topic to those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold.
- 10 20. The computer program product as set forth in Claim 15, wherein the rows of confidence matrix F represent users, and the columns represent topics, such that each entry of the confidence matrix F indicates the likelihood of a user as being interested in a particular topic.
- 15 21. The computer program product as set forth in Claim 15, wherein initiating an action further comprises operations of generating and presenting an online advertisement to users regarding a particular topic to those users whose likelihood of being interested in the particular topic exceeds a predetermined threshold.

1/8

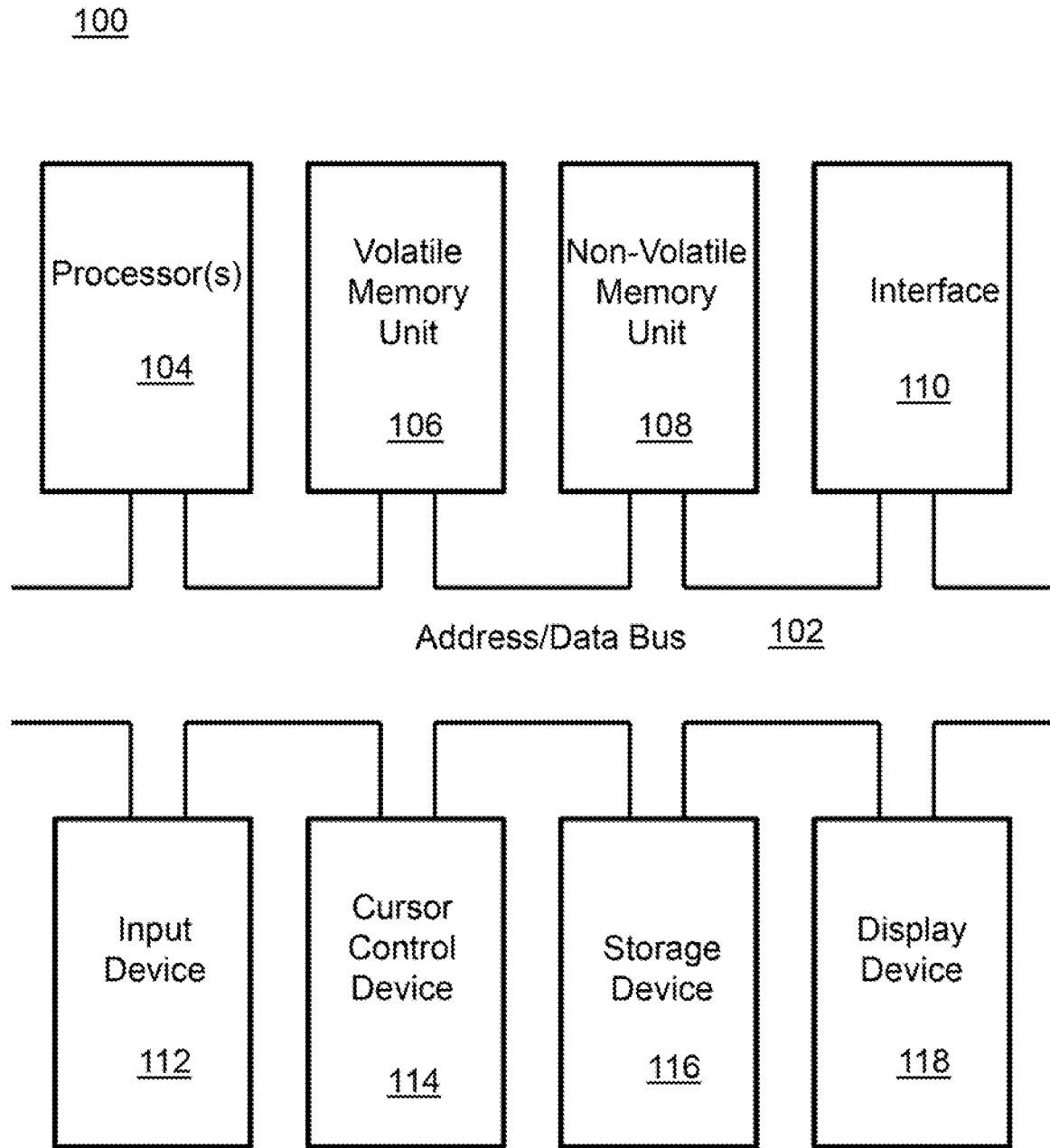


FIG. 1

2/8

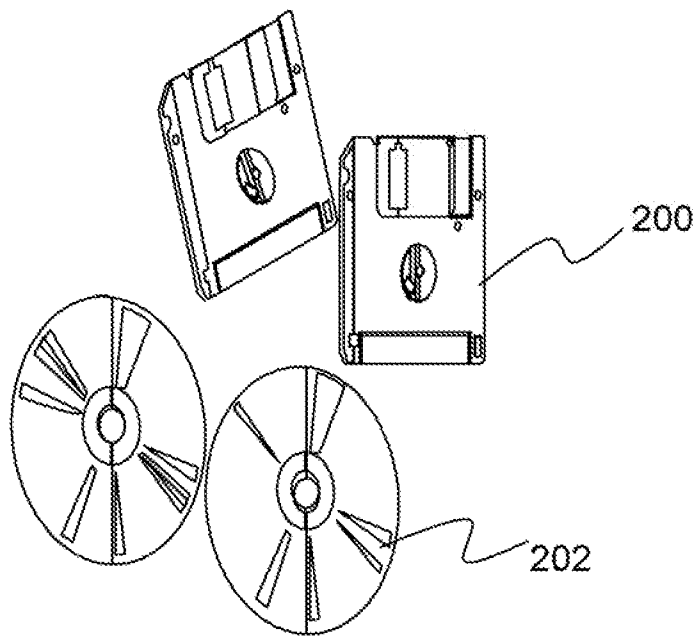


FIG. 2

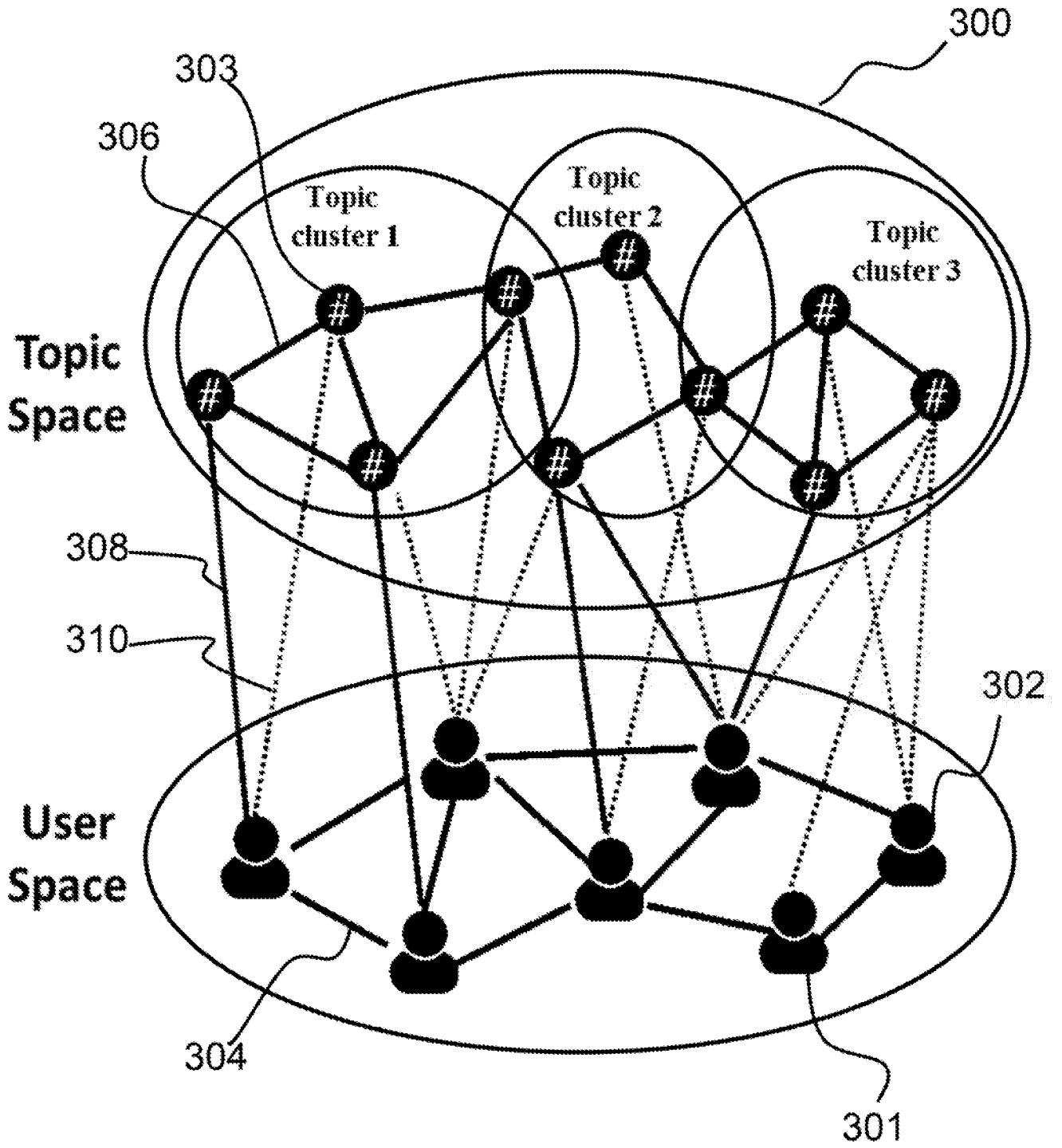


FIG. 3

TWITTER

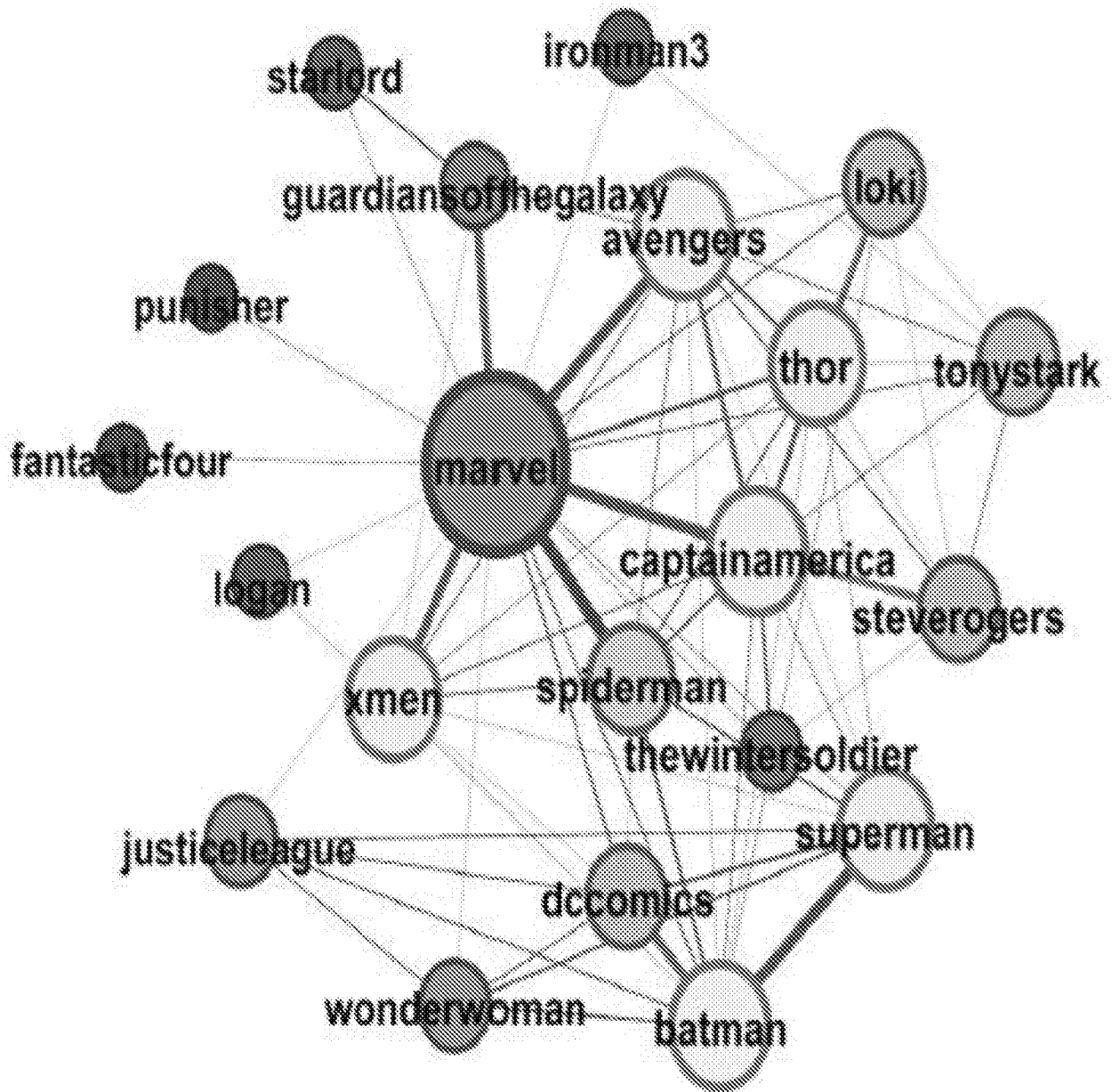


FIG. 4A

Tag Network

TWITTER

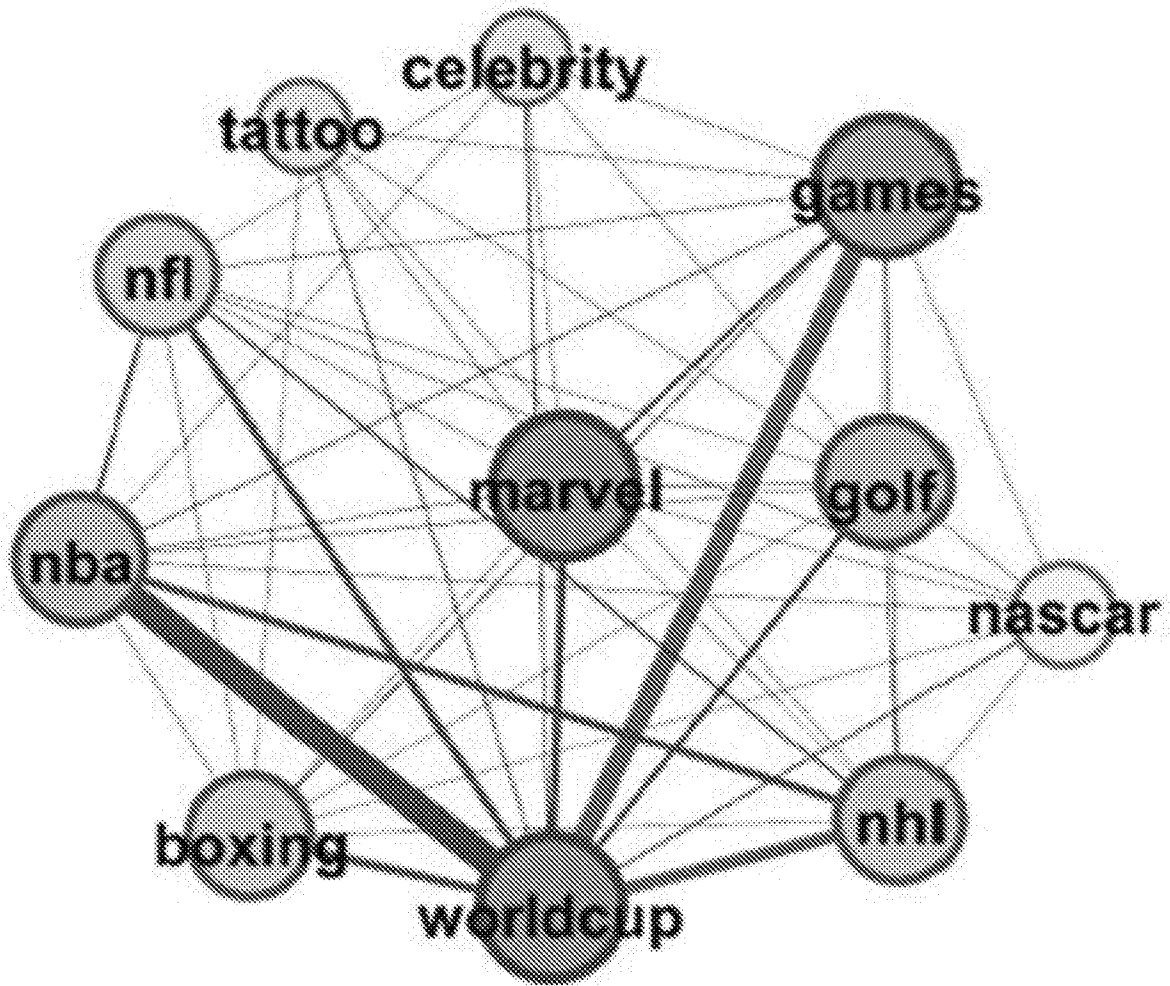


FIG. 4b

Topic Network

Tumblr

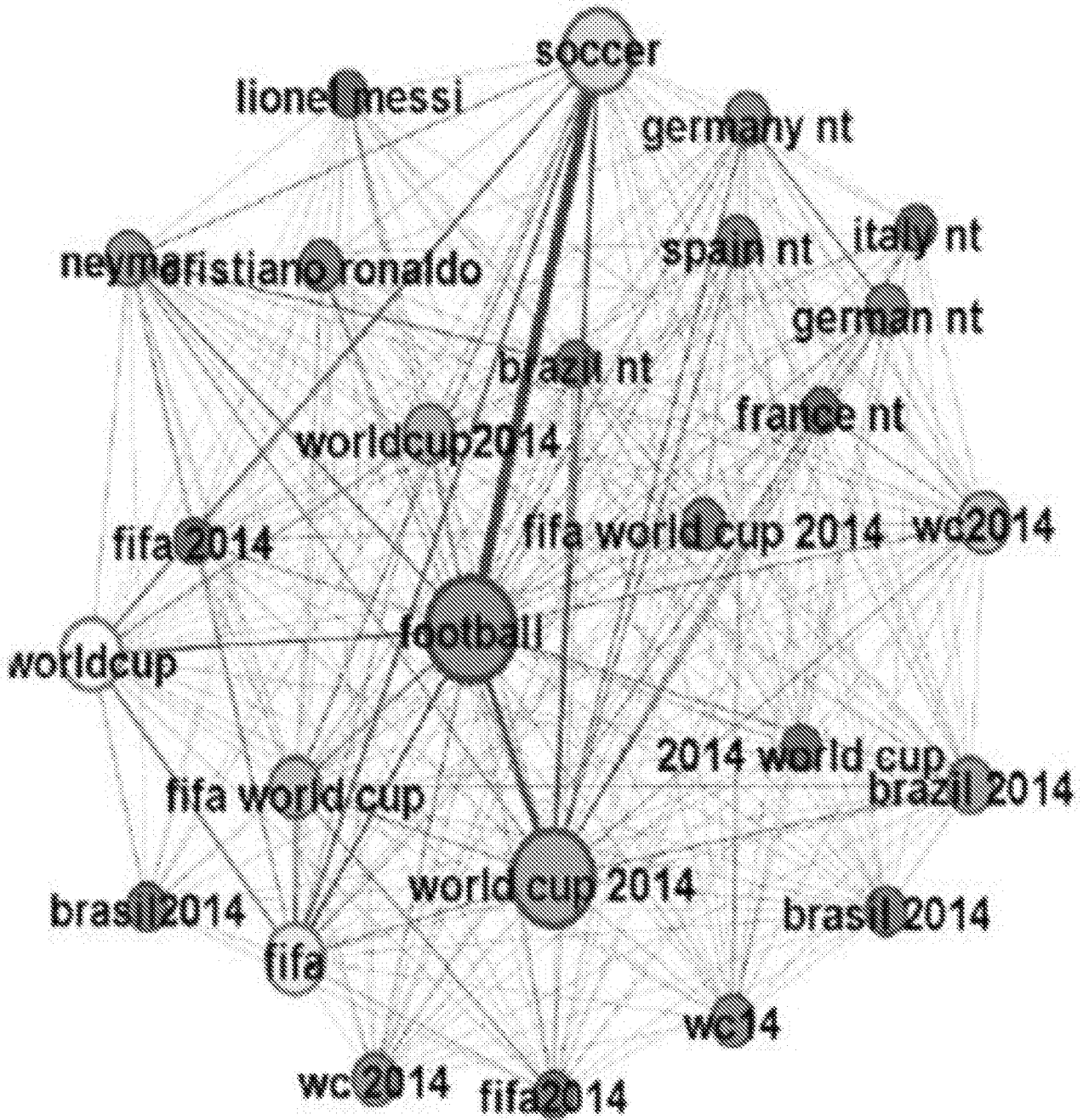


FIG. 4C

Tag Network

7/8

Tumblr

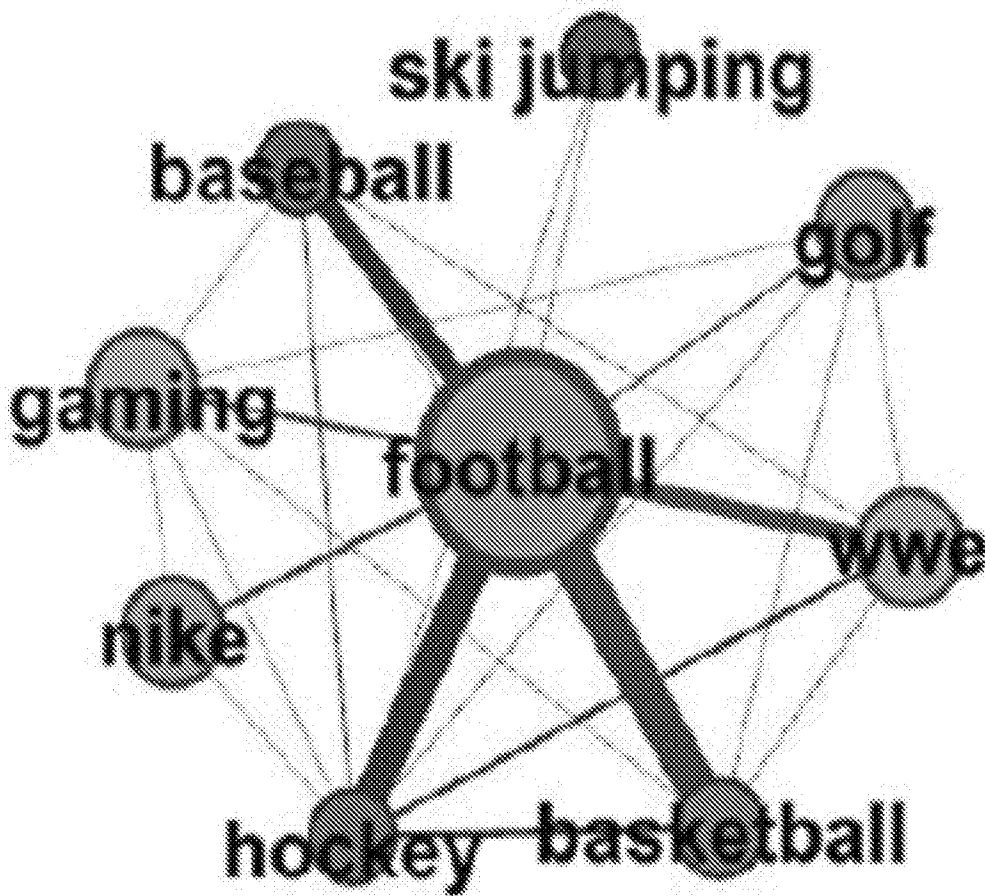


FIG. 4D

Topic Network

8/8

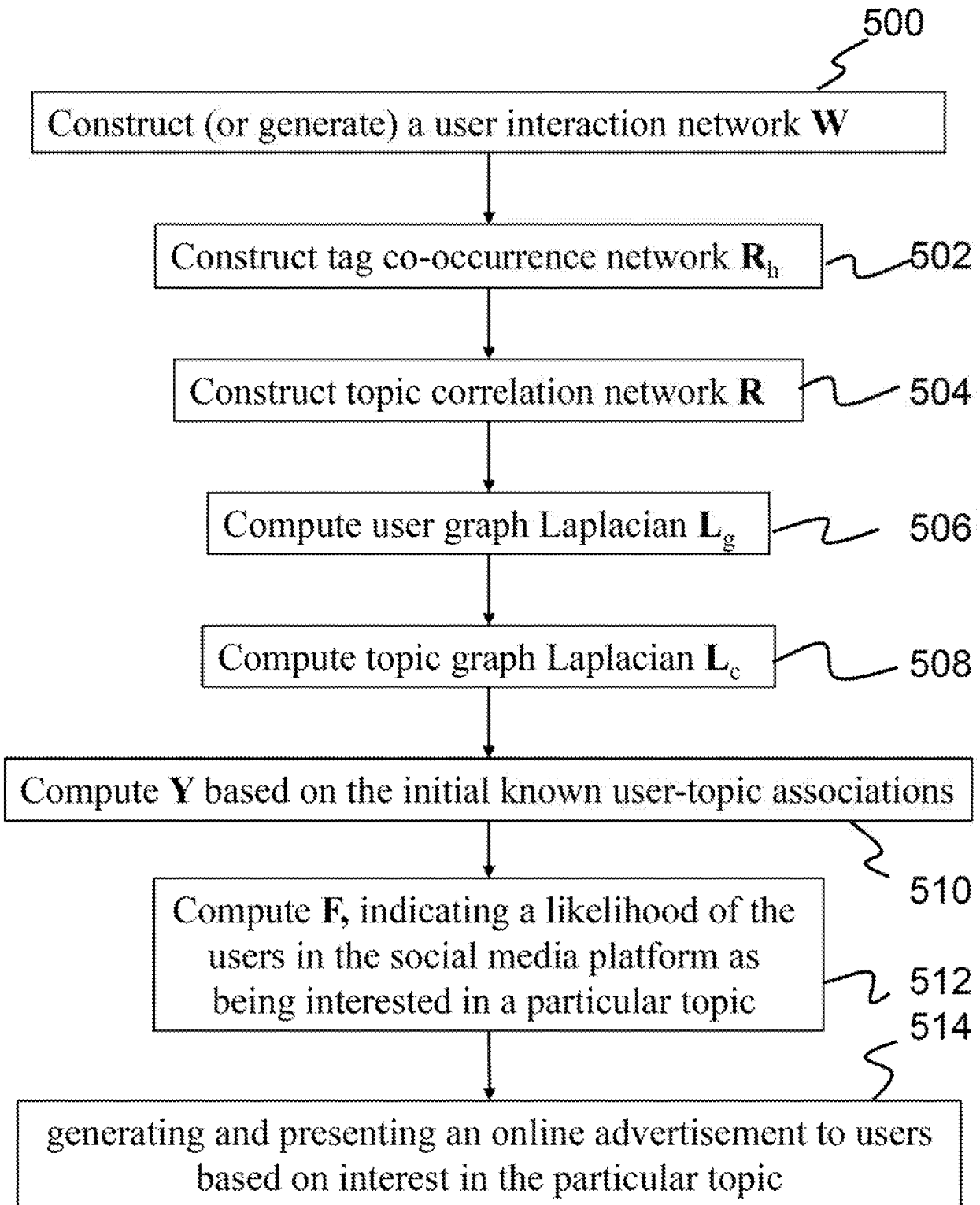


FIG. 5

A. CLASSIFICATION OF SUBJECT MATTER**G06Q 30/02(2012.01)i, G06Q 50/00(2006.01)i**

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

B. FIELDS SEARCHEDMinimum documentation searched (classification system followed by classification symbols)
G06Q 30/02; G06F 15/16; G06Q 99/00; H04L 29/08; G06F 17/30; G06Q 50/00Documentation 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 modelsElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)
eKOMPASS(KIPO internal) & Keywords: social network, interest, interaction, tag, likelihood, advertisement, topic**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2013-0110641 A1 (JUSTIN ORMONT et al.) 02 May 2013 See abstract, paragraphs [0027], [0057], [0091]-[0096], claims 1, 3, 5, 12, 16 and figure 3.	1,6-8,13-15,20-21
Y		2-5,9-12,16-19
Y	US 2014-0040370 A1 (TAGGED, INC.) 06 February 2014 See abstract, paragraphs [0016], [0083] and figure 4.	2-5,9-12,16-19
A	US 2012-0066073 A1 (VENKATACHARI DILIP et al.) 15 March 2012 See abstract, claims 1-2, 4, 8 and figures 3, 5-6.	1-21
A	US 2012-0054205 A1 (SHUMEET BALUJA et al.) 01 March 2012 See abstract, claims 1-3, 6, 15 and figures 2-3, 5.	1-21
A	US 2009-0228296 A1 (EMIL ISMALON) 10 September 2009 See abstract, claims 1-5, 9 and figures 2-4.	1-21

 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

25 November 2016 (25.11.2016)

Date of mailing of the international search report

25 November 2016 (25.11.2016)

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

KANG, Min Jeong

Telephone No. +82-42-481-8131



INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2016/046082

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2013-0110641 A1	02/05/2013	US 2013-0110583 A1 WO 2013-063537 A1	02/05/2013 02/05/2013
US 2014-0040370 A1	06/02/2014	EP 2693669 A2	05/02/2014
US 2012-0066073 A1	15/03/2012	WO 2012-031239 A2 WO 2012-031239 A3	08/03/2012 07/06/2012
US 2012-0054205 A1	01/03/2012	BR PI0811488 A2 KR 10-1426933 B1 KR 10-2010-0051593 A US 2008-0275861 A1 US 8055664 B2 US 8473500 B2 WO 2008-134772 A1	04/11/2014 08/08/2014 17/05/2010 06/11/2008 08/11/2011 25/06/2013 06/11/2008
US 2009-0228296 A1	10/09/2009	None	