



(51) International Patent Classification:

G06F 16/9535 (2019.01) G06N 3/048 (2023.01)
G06N 3/045 (2023.01) G06N 3/0895 (2023.01)

(21) International Application Number:

PCT/US2023/012368

(22) International Filing Date:

06 February 2023 (06.02.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

202210397811.4 15 April 2022 (15.04.2022) CN

(71) Applicant: **MICROSOFT TECHNOLOGY LICENSING, LLC** [US/US]; One Microsoft Way, Redmond, Washington 98052-6399 (US).

(72) Inventors: **GONG, Ming**; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US). **ZHANG, Xingyao**; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US). **YANG, Ze**; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US). **SHOU, Linjun**; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US). **JIANG, Daxin**; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US).

(74) Agent: **CHATTERJEE, Aaron C.** et al.; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM,

(54) Title: TEMPORAL CO-CONTRASTIVE LEARNING-BASED NODE REPRESENTATION GENERATION

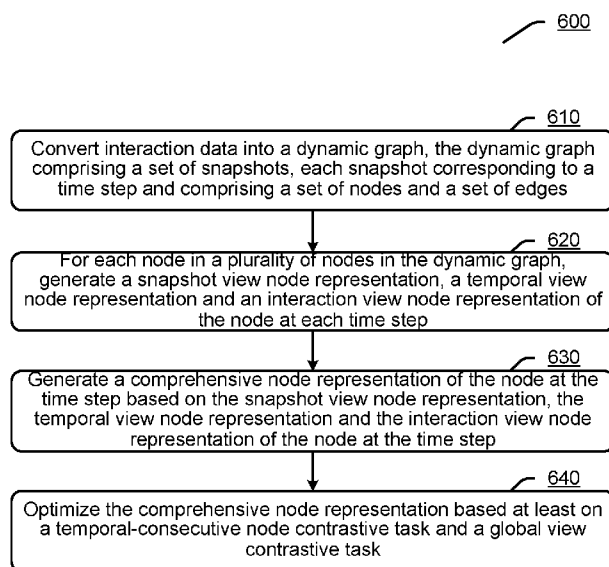


FIG 6

(57) Abstract: This disclosure proposes methods, apparatuses, computer program products and computer-readable mediums for temporal co-contrastive learning-based node representation generation. Interaction data may be converted into a dynamic graph which comprises a set of snapshots, each snapshot corresponding to a time step and comprising a set of nodes and a set of edges. For each node in a plurality of nodes in the dynamic graph, a snapshot view node representation, a temporal view node representation and an interaction view node representation of the node at each time step may be generated. A comprehensive node representation of the node at the time step may be generated based on the snapshot view node representation, the temporal view node representation and the interaction view node representation of the node at the time step. The comprehensive node representation may be optimized based at least on a temporal-consecutive node contrastive task and a global view contrastive task.



AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

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

TEMPORAL CO-CONTRASTIVE LEARNING-BASED NODE REPRESENTATION GENERATION

BACKGROUND

5 A dynamic graph may be usually represented as a sequence of graph snapshots, and these graph snapshots correspond to different time steps. Compared with a static graph, a dynamic graph is more suitable for recommendation system modeling. This is because, in a recommendation system, nodes may have dynamic semantics that change over time, and a dynamic graph may characterize such temporal evolution characteristic more effectively.

10 SUMMARY

This Summary is provided to introduce a selection of concepts that are further described below in the Detailed Description. It is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used to limit the scope of the claimed subject matter. Embodiments of the present disclosure propose methods, apparatuses, computer program products
15 and computer-readable mediums for temporal co-contrastive learning-based node representation generation. Interaction data may be converted into a dynamic graph, the dynamic graph comprising a set of snapshots, each snapshot corresponding to a time step and comprising a set of nodes and a set of edges. For each node in a plurality of nodes in the dynamic graph, a snapshot view node representation, a temporal view node representation and an interaction view node
20 representation of the node at each time step may be generated. A comprehensive node representation of the node at the time step may be generated based on the snapshot view node representation, the temporal view node representation and the interaction view node representation of the node at the time step. The comprehensive node representation may be optimized based at least on a temporal-consecutive node contrastive task and a global view contrastive task.

25 It should be noted that the above one or more aspects comprise the features hereinafter fully described and particularly pointed out in the claims. The following description and the drawings set forth in detail certain illustrative features of the one or more aspects. These features are only indicative of the various ways in which the principles of various aspects may be employed, and this disclosure is intended to include all such aspects and their equivalents.

30 BRIEF DESCRIPTION OF THE DRAWINGS

The disclosed aspects will hereinafter be described in conjunction with the appended drawings that are provided to illustrate and not to limit the disclosed aspects.

FIG.1 illustrates an exemplary process of temporal co-contrastive learning-based node representation generation according to an embodiment.

35 FIG.2 illustrates an example of converting interaction data into a dynamic graph according to an

embodiment.

FIG.3 illustrates an example of generating a snapshot view node representation according to an embodiment.

FIG.4 illustrates an example of generating a temporal view node representation according to an
5 embodiment.

FIG.5 illustrates an example of generating an interaction view node representation according to an embodiment.

FIG.6 illustrates a flowchart of an exemplary method for temporal co-contrastive learning-based node representation generation according to an embodiment.

10 FIG.7 illustrates an exemplary apparatus for temporal co-contrastive learning-based node representation generation according to an embodiment.

FIG.8 illustrates an exemplary apparatus for temporal co-contrastive learning-based node representation generation according to an embodiment.

DETAILED DESCRIPTION

15 The present disclosure will now be discussed with reference to several example implementations. It should be understood that these implementations are discussed only for enabling those skilled in the art to better understand and thus implement the embodiments of the present disclosure, rather than suggesting any limitations on the scope of the present disclosure.

Dynamic graph modeling may broadly refer to the learning of a graph representation of a dynamic
20 graph or node representations of nodes in the dynamic graph, which may also be referred to as, e.g., dynamic graph embedding learning, etc. Existing dynamic graph modeling takes network dynamics and snapshot structures into account, which typically first models node representations within each snapshot separately and then aggregates dynamics across snapshots. The existing dynamic graph modeling does not consider temporal message passing among snapshots. Since
25 each snapshot only includes interactions among nodes occurred within a corresponding time window, separate modeling of each snapshot will also cause interactions among nodes even sparser. Moreover, separate modeling of each snapshot will also face the problem of lacking supervision signals, which reduces accuracy of node representations. A dynamic graph-based recommendation system may perform prediction for links among nodes, and the links reflect
30 predicted interaction behaviors among the nodes. Due to the above limitations of the existing dynamic graph modeling, the current modeling of a dynamic graph-based recommendation system does not treat short-term interaction behaviors and long-term interaction behaviors differently, but treats short-term interaction behaviors and long-term interaction behaviors in the same approach. The embodiments of the present disclosure propose an architecture or mechanism of adopting
35 temporal co-contrastive learning-based node representation generation in dynamic graph

modeling, so as to generate more accurate node representations for a dynamic graph. Accordingly, when the node representation generation architecture according to the embodiments of the present disclosure is applied to a dynamic graph-based recommendation system, the recommendation system may perform temporal-sensitive recommendation tasks with higher effectiveness and robustness, wherein the recommendation tasks may include, e.g., prediction for links among nodes, etc.

The embodiments of the present disclosure apply contrastive learning to dynamic graph modeling, so as to provide supervision signals from data of a dynamic graph itself, and thereby effectively alleviate the data dispersion problem of dynamic graph. Contrastive learning can also preserve dynamic graph semantics in terms of short-term user interests and long-term user interests.

The embodiments of the present disclosure propose enhanced contrastive views in different granularities for node-level representation learning, e.g., snapshot view, temporal view, interaction view, etc. These views are task-specific, e.g., specific to recommendation tasks performed by a recommendation system. Benefiting from node representations generated under these views respectively, the embodiments of the present disclosure may effectively capture both local structural semantics and global structural semantics.

The embodiments of the present disclosure fully consider characteristics of dynamic graphs, e.g., rich temporal information and sequence relationship among snapshots, in the design of contrastive tasks for contrastive learning. In an aspect, the embodiments of the present disclosure propose a temporal-consecutive node contrastive task for explicitly learning short-term user interests from short-term interaction behaviors. For example, the temporal-consecutive node contrastive task may be established for snapshot view, which may provide additional supervision signals, enhance message passing among consecutive snapshots, learn short-term node representations, etc. for snapshot-level node representation learning. In another aspect, the embodiments of the present disclosure propose a global view contrastive task for explicitly learning long-term user interests from long-term interaction behaviors. For example, the global view contrastive task may be established for temporal view and interaction view, which enables the temporal view and the interaction view to cooperatively supervise each other and learn long-term node representations by performing contrast between the temporal view and the interaction view. Herein, a node representation may also be referred to as, e.g., node embedding, etc. Moreover, herein, user interest may refer to a user's interest in an item, which may cause the user to perform an interaction behavior to the item. The temporal-consecutive node contrastive task and the global view contrastive task may be used together for implementing temporal-based co-contrastive learning.

Through adopting at least the snapshot view, the temporal view and the interaction view as described above and adopting the temporal-consecutive node contrastive task and the global view

contrastive task as described above, the embodiments of the present disclosure can capture both short-term user interests and long-term user interests, thereby achieving more consistent performance over different time spans.

The embodiments of the present disclosure may implement high-quality dynamic graph modeling through the temporal co-contrastive learning-based node representation generation architecture, and therefore may be applied to a wider range of scenarios, e.g., recommendation systems, etc. In a dynamic graph-based recommendation system, nodes in a dynamic graph may be abstracted into user nodes corresponding to users and item nodes corresponding to items. Items may widely refer to various objects with which users interact. Items may have different meanings in different specific application scenarios, e.g., items in a news recommendation scenario may refer to news recommended to users, items in a movie recommendation scenario may refer to movies recommended to users, items in a product recommendation scenario may refer to products recommended to users, etc. The recommendation system may perform prediction for links among nodes, e.g., prediction for links or edges between user nodes and item nodes. A predicted link or edge may reflect an interaction behavior between two nodes, e.g., reflecting an interaction behavior between a user node and an item node. In different specific application scenarios, interaction behaviors may have different meanings, e.g., users' behaviors of clicking, browsing, purchasing, commenting, adding to favorites, etc., to items. Benefiting from the more accurate node representations provided by the temporal co-contrastive learning-based node representation generation architecture according to the embodiments of the present disclosure, a dynamic graph-based recommendation system adopting the embodiments of the present disclosure can perform more accurate prediction and recommendation with higher effectiveness and robustness.

FIG.1 illustrates an exemplary process 100 of temporal co-contrastive learning-based node representation generation according to an embodiment. The process 100 may be performed for implementing parameter learning of temporal co-contrastive learning-based node representation generation architecture, thereby effectively accomplishing dynamic graph modeling.

Interaction data 102 may be existing data about interactions between a plurality of users and a plurality of items. For example, the interaction data 102 may refer to temporal-sensitive information about user-item interactions. The interaction data 102 may include an interaction record at each time point in a sequence of time points, and the interaction record may indicate, e.g., there exist interactions between which users and which items at that time point, there exists no interaction between which users and which items at that time point, etc. The interaction data 102 may adopt various formats, e.g., table, chart, text, etc.

At 110, the interaction data 102 may be converted into a corresponding dynamic graph. The dynamic graph may include a set of snapshots. Each snapshot may correspond to a time step,

wherein a time step may also be referred to as a time point. For example, a timeline may be divided into a plurality of time steps through time windows of predetermined length, and a snapshot at each time step may be used for characterizing interactions occurred within a time window associated with this snapshot, wherein the time window associated with the snapshot may refer to a time window immediately before this time step. Each snapshot may be a static graph including a set of nodes and a set of edges. The set of nodes may include nodes involved in interactions occurred within a time window associated with the snapshot, e.g., user nodes and item nodes. Each edge in the set of edges may indicate that there exists interaction between a user node and an item node linked by the edge within a time window associated with the snapshot. Therefore, the set of snapshots in the dynamic graph may characterize node interaction information observed at different time steps, respectively.

At 120, node representations of different views may be generated for the dynamic graph obtained at 110. Herein, "view" may broadly refer to a processing perspective or processing approach for information or data in a dynamic graph.

In an implementation, the embodiments of the present disclosure propose a snapshot view. A snapshot view may refer to performing semantic extraction to each snapshot individually so as to capture local semantics. Exemplarily, a snapshot view may use a graph attention (GAT) network for capturing local structure of nodes within a snapshot and generate local semantic representations of the nodes. Accordingly, a node representation generated for a node under a snapshot view may be referred to as a snapshot view node representation. The operations at 120 may include generating a snapshot view node representation 122 of each node in the dynamic graph at each snapshot or time step. For example, for a target node, at a target time step, a snapshot view node representation of the target node at the target time step may be generated through applying a graph attention network to a snapshot corresponding to the target time step.

In an implementation, the embodiments of the present disclosure propose a temporal view which is a global view. A temporal view may refer to capturing time-aware global semantics over a temporal sequence of a plurality of snapshots. Exemplarily, a temporal view may generate a global semantic representation of a node through aggregating a plurality of snapshot view node representations of the same one node at a plurality of snapshots or time steps. Accordingly, a node representation generated for a node under a temporal view may be referred to as a temporal view node representation. The operations at 120 may include generating a temporal view node representation 124 of each node in the dynamic graph at each time step. For example, for a target node, at a target time step, a temporal view node representation of the target node at the target time step may be generated through aggregating at least one historical snapshot view node representation of the target node. The at least one historical snapshot view node representation

may include at least one snapshot view node representation of the target node at at least one time step before the target time step and optionally a snapshot view node representation of the target node at the target time step.

In an implementation, the embodiments of the present disclosure propose an interaction view which is a global view. An interaction view may refer to capturing global semantics from global node interaction information in a task-specific approach. Exemplarily, an interaction view may, under guidance by global user-item interaction information, generate a global semantic representation of a node from snapshot view node representations of interacted nodes having historical interactions with the node. Accordingly, a node representation generated for a node under an interaction view may be referred to as an interaction view node representation. The operation at 120 may include generating an interaction view node representation 126 of each node in the dynamic graph at each time step. For example, for a target node, at a target time step, at least one interacted node having historical interaction with the target node may be firstly identified. The interacted node may refer to a node that has interaction with the target node at at least one time step before the target time step and optionally at the target time step, e.g., the target node and the interactive node have a link or an edge in at least a certain snapshot. Then, an interaction view node representation of the target node at the target time step may be generated through aggregating the at least one snapshot view node representation of the at least one interacted node at the target time step.

At 130, a comprehensive node representation may be generated based on node representations of different views. A comprehensive node representation is a node representation which is obtained through combining a snapshot view, a temporal view and an interaction view and merges local semantics and global semantics. For example, a comprehensive node representation may be generated based on a snapshot view node representation, a temporal view node representation and an interaction view node representation. In an implementation, e.g., a multi-layer perceptron (MLP) may be adopted for generating a comprehensive node representation based on a snapshot view node representation, a temporal view node representation and an interaction view node representation. The operation at 130 may include generating a comprehensive node representation of each node in the dynamic graph at each time step. For example, for a target node, at a target time step, a comprehensive node representation of the target node at the target time step may be generated through applying a multi-layer perceptron to a snapshot view node representation, a temporal view node representation and an interaction view node representation of the target node at the target time step.

According to the process 100, contrastive learning tasks, such as a temporal-consecutive node contrastive task 140 and a global view contrastive task 150, may be established for providing

additional supervision signals from the perspective of local structures and long-term interests respectively. The temporal-consecutive node contrastive task 140 and the global view contrastive task 150 may be used together for implementing temporal-based co-contrastive learning.

The temporal-consecutive node contrastive task 140 may be established for a snapshot view. The same node should have similar semantic meanings in two consecutive snapshots, which forms the basis of short-term user interest. Two snapshot view node representations of the same node at two consecutive snapshots may form a positive sample pair for contrastive learning. The difference between the two snapshot view node representations of the node at the two consecutive snapshots is a good perturbation result for performing contrastive learning to acquire temporal changes among different snapshots. The temporal-consecutive node contrastive task 140 may encourage message passing among consecutive snapshots and facilitate to learn short-term node representations. For example, for a target node, at a target time step, a snapshot view node representation of the target node at the target time step and a previous snapshot view node representation of the target node at the previous time step may be taken as a positive sample pair for contrastive learning.

The global view contrastive task 150 may be established for a temporal view and an interaction view, which is used for learning a long-term node representation within the temporal view and the interaction view, and capturing long-term user interests based on these two different global views such as the temporal view and the interaction view. The temporal view and the interaction view may not only enhance information diversity, but also cooperatively supervise each other to form a positive sample pair for contrastive learning. For example, for a target node, at a target time step, a temporal view node representation and an interaction view node representation of the target node at the target time step may be taken as a positive sample pair for contrastive learning.

At 170, the comprehensive node representation may be optimized based at least on the temporal-consecutive node contrastive task 140 and the global view contrastive task 150. Herein, optimization for comprehensive node representation may widely refer to optimization for various parameters in the node representation generation architecture as reflected by, e.g., the process 100, so as to improve the quality or accuracy of the comprehensive node representation. In an implementation, a loss function may be established based at least on the temporal-consecutive node contrastive task 140 and the global view contrastive task 150, and the comprehensive node representation may be optimized based on the loss function. For example, the comprehensive node representation may be optimized through minimizing a loss calculated via the loss function.

Optionally, the process 100 may also include establishing a graph context prediction task 160. The graph context prediction task 160 may be used for preserve temporal structure around a node from consecutive snapshots across a plurality of time steps. The graph context prediction task 160 may

also be used for optimizing the comprehensive node representation. For example, the above-mentioned loss function established based on the temporal-consecutive node contrastive task 140 and a global view contrastive task 150 may be further based on the graph context prediction task 160.

- 5 Through the process 100, parameter learning of the temporal co-contrastive learning-based node representation generation architecture may be implemented, thereby accomplishing dynamic graph modeling. Further, the node representation generation architecture may be applied for recommendation tasks in a dynamic graph-based recommendation system. In an implementation, for two nodes in a dynamic graph, it may be predicted, based on comprehensive node representations of the two nodes at the current time step, whether there exists interaction between the two nodes at the next time step, e.g., whether there exists a link or an edge between the two nodes. The two nodes may be a user node and an item node respectively. As an example, an external logistic regression classifier may be trained for predicting, based on a comprehensive node representation of a first node at the current time step and a comprehensive node representation of a second node at the current time step, whether there exists interaction between the first node and the second at the next time step.

More exemplary details of the temporal co-contrastive learning-based node representation generation according to the embodiments of the present disclosure will be described below.

- A dynamic graph may be represented as a set of snapshots, e.g., a sequence of static graph snapshots. Each snapshot corresponds to a time step. As an example, a dynamic graph G may be represented as $G = \{g^1, g^2, \dots, g^T\}$, wherein T is the number of time steps, $g^t \mid_{t \in [1, T]}$ represents a snapshot at the t -th time step. g^t includes information about interactions among nodes that occurred in a time window $[t - 1, t]$ associated with the time step t . g^t may be represented as $g^t = \{V_t, E_t\}$, wherein V_t represents a set of nodes included in g^t , and E_t represents a set of edges or links included in g^t .

- In a recommendation scenario such as a recommendation system, interaction data may include information about interactions between users and items, e.g., a user clicks on an item, a user purchases an item, etc. Accordingly, g^t may be further represented as $g^t = \{(u, y_{ui}, i) \mid u, i \in V_t\}$, wherein u represents a user node corresponding to a user, i represents an item node corresponding to an item, $y_{ui} \in [0, 1]$ represents whether there exists an edge between u and i , $y_{ui} = 0$ represents whether there exists no edge or link and accordingly there exists no interaction between u and i , and $y_{ui} = 1$ represents there exists an edge or link and accordingly there exists interaction between u and i . In the recommendation scenario, given u , it is required to predict whether there will exist interaction between u and i or to generate a probability that interaction will occur between u and i .

FIG.2 illustrates an example 200 of converting interaction data into a dynamic graph according to an embodiment.

It is assumed that interaction data 210 includes information about interactions between a set of users $\{u_1, u_2, \dots, u_{V1}\}$ and a set of items $\{i_1, i_2, \dots, i_{V2}\}$, wherein $V1$ is the number of users and $V2$ is the number of items. For example, the interaction data 210 may include records about there exist interactions between which users and which items at each time point. It should be understood that although the interaction data 210 is shown as a chart in FIG.1, the interaction data 210 may adopt any other formats in practice.

The interaction data 210 may be converted into a corresponding dynamic graph 220. The dynamic graph 220 includes a set of snapshots $\{g^1, g^2, \dots, g^T\}$. Each snapshot corresponds to a time step. For example, the snapshot g^1 corresponds to a time step $t_1 = 1$, the snapshot g^2 corresponds to a time step $t_2 = 2$, and so on. Moreover, each snapshot may have an associated time window. For example, a time window associated with the snapshot g^1 is a time window $[t_0, t_1]$ immediately before the time step t_1 . Each snapshot includes information about interactions among nodes that occurred within an associated time window. For example, the snapshot g^1 includes information about interactions among nodes that occurred within the time window $[t_0, t_1]$. Interaction information is presented visually in a snapshot in the form of nodes and edges. As an example, the snapshot g^1 includes user nodes u_1, u_2 and item nodes i_1, i_2 . The snapshot g^1 also includes edges indicating there exists interaction between two linked nodes, e.g., an edge between u_1 and i_1 , an edge between u_1 and i_2 , an edge between u_2 and i_1 , etc. It should be understood that, preferably, a time window associated with each snapshot may have the same predetermined length.

According to the embodiments of the present disclosure, a snapshot view node representation of each node at each time step may be generated.

Taking a snapshot g^t as an example, initial node representations of all V_t nodes in a D-dimensional continuous space at a time step t are represented by $H^0 = \{h_1^0, h_2^0, \dots, h_{|V_t|}^0\}$, $h_n^0 \in \mathbb{R}^D$, wherein a node n may denote a user node and an item node. The initial node representations may be generated through any approaches, e.g., randomly generated, previously designated, etc. The initial node representation H^0 may be converted into a higher-level snapshot view node representation $H^t = \{h_1^t, h_2^t, \dots, h_{|V_t|}^t\}$, $h_n^t \in \mathbb{R}^D$ through, e.g., a graph attention network. The graph attention network is a neural network architecture operating on graph-structured data, which leverages self-attention, multi-head attention, etc. A snapshot view node representation h_n^t of a node n at a time step t may be calculated as:

$$h_n^t = \sigma \left(\frac{1}{K} \sum_{k=1}^K \sum_{v \in N_n} \alpha_{nv}^k W^k h_v^t \right) \quad \text{Equation (1)}$$

wherein N_n is a set of node n 's one-hop neighbor nodes, σ is an activation function used in the graph attention network (e.g., LeakyReLU), W^k is a learnable parameter matrix for the k -th multi-head attention, K is the number of multi-head attentions, α_{nv}^k is a correlation coefficient between node n and node v and delegates the influence weight of neighbor node v on node n . Through Equation (1), a snapshot view node representation of each node at each time step may be obtained. In the recommendation scenario, a snapshot view node representation may reflect static user interests. It should be understood that the generation of a snapshot view node representation is performed for each snapshot individually. For example, a snapshot view node representation of the node n at the time step t is generated within the snapshot g^t corresponding to the time step t . FIG.3 illustrates an example 300 of generating a snapshot view node representation according to an embodiment. Assuming that a graph attention network 310 is applied to a set of snapshots $\{g^1, g^2, \dots, g^T\}$ respectively, a set of snapshot view node representations may be generated for each snapshot respectively according to Equation (1). Taking a snapshot g^2 as an example, through the graph attention network 310, snapshot view node representations $h_{u_1}^2, h_{u_2}^2, h_{i_1}^2$ and $h_{i_2}^2$ may be generated for nodes u_1, u_2, i_1 and i_2 included in the snapshot g^2 , respectively.

According to the embodiments of the present disclosure, a temporal view node representation of each node at each time step may be generated.

The purpose of generating a temporal view node representation is to capture time-aware changes in a graph structure over a plurality of snapshots. Taking a node n as an example, for a time step t , a sequence $\{h_n^1, h_n^2, \dots, h_n^t\}$ represents a set of historical snapshot view node representations of the node n , which includes t snapshot view node representations of the node n from a time step 1 to a time step t , wherein the node n may denote a user node and an item node. A temporal view node representation $\overline{h_n^t}$ of the node n at the time step t may be calculated as:

$$e_n^{tj} |_{tj \in [1, t]} = \text{MLP} \left(h_n^{tj} \right) \quad \text{Equation (2)}$$

$$\alpha_n^{ttj} = \frac{\exp(e_n^t \cdot e_n^{tj})}{\sum_{tj \in [1, t]} \exp(e_n^t \cdot e_n^{tj})} \quad \text{Equation (3)}$$

$$\overline{h_n^t} = \sum_{tj \in [1, t]} \alpha_n^{ttj} h_n^{tj} \quad \text{Equation (4)}$$

wherein MLP denotes a linear layer or a multi-layer perceptron, α_n^{ttj} is a correlation coefficient and denotes the influence weight of time step tj on time step t . The process of Equation (2) to Equation (4) may implement the aggregation of historical snapshot view node representations of

the node n , so as to generate a temporal view node representation of the node n at the time step t . Through Equation (2) to Equation (4), a temporal view node representation of each node at each time step may be obtained.

FIG.4 illustrates an example 400 of generating a temporal view node representation according to an embodiment. Taking a node u_1 as an example, it is assumed that it is desired to generate a temporal view node representation of the node u_1 at a time step t . First, historical snapshot view node representations of the node u_1 may be extracted, e.g., snapshot view node representations $\{h_{u_1}^1, h_{u_1}^2, \dots, h_{u_1}^t\}$ of u_1 at time steps $\{1, 2, \dots, t\}$. Then, according to Equation (2) to Equation (4), a temporal view node representation $\overline{h_{u_1}^t}$ of the node u_1 at the time step t may be generated through aggregating $\{h_{u_1}^1, h_{u_1}^2, \dots, h_{u_1}^t\}$.

According to the embodiments of the present disclosure, an interaction view node representation of each node at each time step may be generated.

In an interaction view, an interaction view node representation of a target node may be generated with snapshot view node representations of interacted nodes having historical interactions with the target node. For example, an interaction view node representation of a user node may be generated through aggregating snapshot view node representations of item nodes having historical interactions with the user node, and an interaction view node representation of an item node may be generated through aggregating snapshot view node representations of user nodes having historical interactions with the item node. Taking a user node u as an example, for a time step t , a set of item nodes having historical interactions with the user node u may be represented as $S_t = \{i_1, i_2, \dots, i_N\}$ which includes N item nodes having historical interactions with the user node u until the time step t . An interaction view node representation $\overline{h_u^{st}}$ of the user node u at the time step t may be calculated as:

$$e_{i_k}^t = MLP(h_{i_k}^t) \quad \text{Equation (5)}$$

$$\alpha_{i_k}^t = \frac{\exp(e_{i_k}^t \cdot W^e)}{\sum_{i_k \in S_t} \exp(e_{i_k}^t \cdot W^e)} \quad \text{Equation (6)}$$

$$\overline{h_u^{st}} = \sum_{i_k \in S_t} \alpha_{i_k}^t h_{i_k}^t \quad \text{Equation (7)}$$

wherein MLP denotes a linear layer or a multi-layer perceptron, W^e is a learnable parameter, $\alpha_{i_k}^t$ is a correlation coefficient and denotes the influence weight of item node i_k on user node u . The process of Equation (5) to Equation (7) may implement the aggregation of snapshot view node representations of item nodes having historical interactions with the user node u at the time step t , so as to generate an interaction view node representation of the user node u at the time step t . Through Equation (5) to Equation (7), an interaction view node representation of each node at

each time step may be obtained.

Moreover, through an approach similar to that for calculating $\overline{h_u^{st}}$ as described above, an interaction view node representation $\overline{h_i^{st}}$ of an item node i at a time step t may be calculated. For example, an interaction view node representation of the item node i at the time step t may be generated through aggregating snapshot view node representations of user nodes having historical interactions with the item node i at the time step t . Optionally, $\overline{h_u^{st}}$ and $\overline{h_i^{st}}$ may be uniformly represented as $\overline{h_n^{st}}$, wherein the node n may denote a user node and an item node.

FIG.5 illustrates an example 500 of generating an interaction view node representation according to an embodiment. Taking a node u_1 as an example, it is assumed that it is desired to generate an interaction view node representation of a node u_1 at a time step t . At 510, interacted nodes having historical interactions with the node u_1 may be identified, e.g., nodes i_1 , i_2 and i_3 . Interacted nodes may be identified from snapshots $\{g^1, g^2, \dots, g^t\}$ corresponding to time steps $\{1, 2, \dots, t\}$, e.g., taking those nodes having edges or links with the node u_1 as interacted nodes. At 520, according to Equation (5) to Equation (7), an interaction view node representation $\overline{h_{u_1}^{st}}$ of the node u_1 at the time step t may be generated through aggregating snapshot view node representations $\{h_{i_1}^t, h_{i_2}^t, h_{i_3}^t\}$ of the interacted nodes i_1 , i_2 , and i_3 at the time step t .

According to the embodiments of the present disclosure, a temporal-consecutive node contrastive task may be established.

A dynamic graph involves rich information in different snapshots, and thus it is important to learn temporal-sensitive semantics in a self-supervised approach. The same node should have similar semantic meanings in two consecutive snapshots. As an example, for two consecutive snapshots g^t and g^{t+1} , the same node n should have similar semantic meanings in the two consecutive snapshots. Thus, snapshot view node representations h_n^t and h_n^{t+1} of the node n in the two snapshots may form a positive sample pair for contrastive learning. The temporal-consecutive node contrastive task encourages message passing between consecutive snapshots, for the generation of short-term node-level representations. Temporal-consecutive node contrastive task-based learning may include adopting the following loss function:

$$\mathcal{L}_{TCL} = \sum_{n \in V} -\log \left(\frac{\exp(h_n^t \cdot h_n^{t+1})}{\sum_{v \in P^{t+1}(n) + \{n\}} \exp(h_n^t \cdot h_v^{t+1})} \right) \quad \text{Equation (8)}$$

wherein $\mathcal{L}_{TCL}$ is a loss calculated for the temporal-consecutive node contrastive task, V represents a set of all nodes, and $P^{t+1}(n)$ denotes a set of negative samples of the node n at a time step $t + 1$. Exemplarily, $P^{t+1}(n)$ may include nodes randomly sampled, as negative samples, from nodes different from the node n , which may be used for establishing negative sample pairs. The temporal-consecutive node contrastive task may provide additional supervision signals with the

established positive sample pairs and negative sample pairs. Moreover, the above loss function may at least learn short-term interest bias through temporal-consecutive edge changes.

According to the embodiments of the present disclosure, a global view contrastive task may be established.

- 5 As described above, in order to preserve global semantic context on the dynamic graph, for a given node n , at a time step t , a temporal view node representation $\overline{h}_n^t$ and an interaction view node representation $\overline{h}_n^{st}$ may be generated from a temporal view and an interaction view. These two node representations of the node n at the time step t may form a positive sample pair for contrastive learning. Moreover, nodes, which are taken as negative samples, may be randomly
10 sampled from nodes that do not have direct links with the node n , for establishing negative sample pairs.

In order to convert a temporal view node representation from a temporal view and an interaction view node representation from an interaction view into the same semantic space, the following mapping may be performed:

$$\begin{aligned} \overline{h}_n^{t'} &= W^{g_2} \sigma(W^{g_1} \overline{h}_n^t + b^{g_1}) + b^{g_2} \\ \overline{h}_n^{s^{t'}} &= W^{g_2} \sigma(W^{g_1} \overline{h}_n^{st} + b^{g_1}) + b^{g_2} \end{aligned} \quad \text{Equation (9)}$$

wherein W^{g_1} , W^{g_2} , b^{g_1} and b^{g_2} are learnable parameters.

Global view contrastive task-based learning may include adopting the following loss function:

$$\mathcal{L}_{GCL} = \sum_{n \in V} -\log \left(\frac{\exp(\overline{h}_n^{t'} \cdot \overline{h}_n^{s^{t'}})}{\sum_{v \in P^{t+1}(n) + \{n\}} \exp(\overline{h}_n^{t'} \cdot \overline{h}_v^{s^{t'}})} \right) \quad \text{Equation (10)}$$

- 20 wherein $\mathcal{L}_{GCL}$ is a loss calculated for the global view contrastive task, V denotes a set of all nodes, and $P^{t+1}(n)$ denotes a set of negative samples of the node n . The global view contrastive task may provide additional supervision signals with the established positive sample pairs and negative sample pairs. Moreover, the above loss function may capture global semantic information and long-term interests through two types of global view representations.

- 25 According to the embodiments of the present disclosure, a comprehensive node representation may be generated based on node representations of different views. Assuming that a snapshot view node representation h_n^t , a temporal view node representation $\overline{h}_n^t$, and an interaction view node representation $\overline{h}_n^{st}$ of a node n at a time step t have been calculated, a comprehensive node representation $\widetilde{h}_n^t$ of the node n at the time step t may be calculated as:

$$\widetilde{h}_n^t = MLP(h_n^t; \overline{h}_n^t; \overline{h}_n^{st}) \quad \text{Equation (11)}$$

wherein MLP denotes a multi-layer perceptron.

According to the embodiments of the present disclosure, optionally, a graph context prediction task may be established.

The graph context prediction task may motivate each node to perceive neighbor nodes in adjacent time steps, so as to preserve temporal structure around the node from consecutive snapshots across a plurality of time steps. The learning of the graph context prediction task may include adopting the following loss function:

$$\mathcal{L}_{GCP} = \sum_{n \in V} \left(\sum_{v \in N_{walk}^{t+1}(n)} -\log \left(\sigma(\widetilde{h}_n^t \cdot \widetilde{h}_v^t) \right) - \omega \cdot \sum_{v' \in P^{t+1}(n)} \log \left(1 - \sigma(\widetilde{h}_n^t \cdot \widetilde{h}_{v'}^t) \right) \right)$$

Equation (12)

wherein $\mathcal{L}_{GCP}$ is a loss calculated for the graph context prediction task, V denotes a set of all nodes, σ is a sigmoid function, $N_{walk}^{t+1}(n)$ denotes a set of nodes that co-occur with the node n on fixed-length random walks, ω is a negative sampling ratio and is a tunable hyper-parameter for balancing positive samples and negative samples, and $P^{t+1}(n)$ denotes a set of negative samples of the node n . In the graph context prediction task, nodes, which are taken as positive samples, may be those nodes that have interaction relationship with the node n determined based on, e.g., random walks, and nodes, which are taken as negative samples, may be nodes randomly sampled from nodes that do not have interaction relationship with the node n .

According to the embodiments of the present disclosure, a comprehensive node representation may be optimized at least with a loss function.

In order to capture short-term and long-term temporal-sensitive semantic information, the following loss function may be defined for a dynamic graph:

$$\mathcal{L} = \mathcal{L}_{GCP} + \lambda(\mathcal{L}_{TCL} + \mathcal{L}_{GCL})$$

Equation (13)

wherein $\mathcal{L}$ is a calculated comprehensive loss, and λ is a balancing coefficient. Since $\mathcal{L}_{TCL}$ and $\mathcal{L}_{GCL}$ are co-contrastive learning tasks and have the same scale in calculation, $\mathcal{L}_{TCL}$ and $\mathcal{L}_{GCL}$ may be combined together in Equation (13). Moreover, it should be understood that $\mathcal{L}_{GCP}$ in Equation (13) is optional, and thus the above loss function may also be established only based on $\mathcal{L}_{TCL}$ and $\mathcal{L}_{GCL}$.

The loss function in Equation (13) may be used for optimizing a comprehensive node representation. For example, parameters in the node representation generation architecture according to the embodiments of the present disclosure may be learned through minimizing the loss calculated via Equation (13). Exemplarily, the above optimizing process may adopt, e.g., Adam optimizer, etc. Moreover, for each iteration, parameter updating may be performed according to Equation (13).

FIG.6 illustrates a flowchart of an exemplary method 600 for temporal co-contrastive learning-

based node representation generation according to an embodiment.

At 610, interaction data may be converted into a dynamic graph, the dynamic graph comprising a set of snapshots, each snapshot corresponding to a time step and comprising a set of nodes and a set of edges.

5 At 620, for each node in a plurality of nodes in the dynamic graph, a snapshot view node representation, a temporal view node representation and an interaction view node representation of the node at each time step may be generated.

At 630, a comprehensive node representation of the node at the time step may be generated based on the snapshot view node representation, the temporal view node representation and the
10 interaction view node representation of the node at the time step.

At 640, the comprehensive node representation may be optimized based at least on a temporal-consecutive node contrastive task and a global view contrastive task.

In an implementation, the interaction data may be data about interaction between a plurality of users and a plurality of items. The set of nodes may comprise user nodes and item nodes. Each
15 edge in the set of edges may indicate that there exists interaction between a user node and an item node linked by the edge within a time window associated with the snapshot.

In an implementation, the generating a snapshot view node representation may comprise: generating the snapshot view node representation of the node at the time step through applying a graph attention network to a snapshot corresponding to the time step.

20 In an implementation, the generating a temporal view node representation may comprise: generating the temporal view node representation of the node at the time step through aggregating at least one historical snapshot view node representation of the node.

In an implementation, the generating an interaction view node representation may comprise: identifying at least one interacted node having historical interaction with the node; and generating
25 the interaction view node representation of the node at the time step through aggregating at least one snapshot view node representation of the at least one interacted node at the time step.

In an implementation, the generating a comprehensive node representation may comprise: generating the comprehensive node representation of the node at the time step through applying a multi-layer perceptron to the snapshot view node representation, the temporal view node
30 representation and the interaction view node representation.

In an implementation, the temporal-consecutive node contrastive task may at least comprise: taking the snapshot view node representation of the node at the time step and a previous snapshot view node representation of the node at the previous time step as a positive sample pair for contrastive learning.

35 In an implementation, the global view contrastive task may at least comprise: taking the temporal

view node representation and the interaction view node representation of the node at the time step as a positive sample pair for contrastive learning.

In an implementation, the optimizing the comprehensive node representation may comprise: establishing a loss function based at least on the temporal-consecutive node contrastive task and the global view contrastive task; and optimizing the comprehensive node representation based on the loss function.

In an implementation, the comprehensive node representation may be optimized further based on a graph context prediction task.

In an implementation, the method 600 may further comprise, for two nodes in the plurality of nodes: predicting, based on comprehensive node representations of the two nodes at the current time step, whether there exists interaction between the two nodes at the next time step. The two nodes may be a user node and an item node respectively.

It should be understood that the method 600 may further comprise any steps/processes for temporal co-contrastive learning-based node representation generation according to the embodiments of the present disclosure described above.

FIG.7 illustrates an exemplary apparatus 700 for temporal co-contrastive learning-based node representation generation according to an embodiment.

The apparatus 700 may comprise: a converting module 710, for converting interaction data into a dynamic graph, the dynamic graph comprising a set of snapshots, each snapshot corresponding to a time step and comprising a set of nodes and a set of edges; a view node representation generating module 720 for, for each node in a plurality of nodes in the dynamic graph, generating a snapshot view node representation, a temporal view node representation and an interaction view node representation of the node at each time step; a comprehensive node representation generating module 730, for generating a comprehensive node representation of the node at the time step based on the snapshot view node representation, the temporal view node representation and the interaction view node representation of the node at the time step; and an optimizing module 740, for optimizing the comprehensive node representation based at least on a temporal-consecutive node contrastive task and a global view contrastive task. Moreover, the apparatus 700 may further comprise any other modules that are configured for performing any steps/processes of the methods for temporal co-contrastive learning-based node representation generation according to the embodiments of the present disclosure described above.

FIG.8 illustrates an exemplary apparatus 800 for temporal co-contrastive learning-based node representation generation according to an embodiment.

The apparatus 800 may comprise at least one processor 810. The apparatus 800 may further comprise a memory 820 connected with the at least one processor 810. The memory 820 may

store computer-executable instructions that, when executed, cause the at least one processor 810 to: convert interaction data into a dynamic graph, the dynamic graph comprising a set of snapshots, each snapshot corresponding to a time step and comprising a set of nodes and a set of edges; for each node in a plurality of nodes in the dynamic graph, generate a snapshot view node representation, a temporal view node representation and an interaction view node representation of the node at each time step; generate a comprehensive node representation of the node at the time step based on the snapshot view node representation, the temporal view node representation and the interaction view node representation of the node at the time step; and optimize the comprehensive node representation based at least on a temporal-consecutive node contrastive task and a global view contrastive task.

In an implementation, the generating a snapshot view node representation may comprise: generating the snapshot view node representation of the node at the time step through applying a graph attention network to a snapshot corresponding to the time step.

In an implementation, the generating a temporal view node representation may comprise: generating the temporal view node representation of the node at the time step through aggregating at least one historical snapshot view node representation of the node.

In an implementation, the generating an interaction view node representation may comprise: identifying at least one interacted node having historical interaction with the node; and generating the interaction view node representation of the node at the time step through aggregating at least one snapshot view node representation of the at least one interacted node at the time step.

In an implementation, the temporal-consecutive node contrastive task may at least comprise: taking the snapshot view node representation of the node at the time step and a previous snapshot view node representation of the node at the previous time step as a positive sample pair for contrastive learning.

In an implementation, the global view contrastive task may at least comprise: taking the temporal view node representation and the interaction view node representation of the node at the time step as a positive sample pair for contrastive learning.

In an implementation, the optimizing the comprehensive node representation may comprise: establishing a loss function based at least on the temporal-consecutive node contrastive task and the global view contrastive task; and optimizing the comprehensive node representation based on the loss function.

Moreover, the at least one processor 810 may be further configured for performing any other steps/processes of the methods for temporal co-contrastive learning-based node representation generation according to the embodiments of the present disclosure described above.

The embodiments of the present disclosure propose a computer program product for temporal co-

contrastive learning-based node representation generation. The computer program product may comprise a computer program that is executed by at least one processor for: converting interaction data into a dynamic graph, the dynamic graph comprising a set of snapshots, each snapshot corresponding to a time step and comprising a set of nodes and a set of edges; for each node in a plurality of nodes in the dynamic graph, generating a snapshot view node representation, a temporal view node representation and an interaction view node representation of the node at each time step; generating a comprehensive node representation of the node at the time step based on the snapshot view node representation, the temporal view node representation and the interaction view node representation of the node at the time step; and optimizing the comprehensive node representation based at least on a temporal-consecutive node contrastive task and a global view contrastive task. Moreover, the computer program may be further executed by the at least one processor for performing any other steps/processes of the methods for temporal co-contrastive learning-based node representation generation according to the embodiments of the present disclosure described above.

The embodiments of the present disclosure may be embodied in a non-transitory computer-readable medium. The non-transitory computer-readable medium may comprise instructions that, when executed, cause one or more processors to perform any steps/processes of the methods for temporal co-contrastive learning-based node representation generation according to the embodiments of the present disclosure described above.

It should be understood that all the operations in the methods described above are merely exemplary, and the present disclosure is not limited to any operations in the methods or sequence orders of these operations, and should cover all other equivalents under the same or similar concepts.

Additionally, the articles "a" and "an" as used in this description and appended claims, unless otherwise specified or clear from the context that they are for the singular form, should generally be interpreted as meaning "one" or "one or more."

It should also be understood that all the modules in the apparatuses described above may be implemented in various approaches. These modules may be implemented as hardware, software, or a combination thereof. Moreover, any of these modules may be further functionally divided into sub-modules or combined together.

Processors have been described in connection with various apparatuses and methods. These processors may be implemented using electronic hardware, computer software, or any combination thereof. Whether such processors are implemented as hardware or software will depend upon the particular application and overall design constraints imposed on the system. By way of example, a processor, any portion of a processor, or any combination of processors

presented in the present disclosure may be implemented with a micro-processor, micro-controller, digital signal processor (DSP), a field-programmable gate array (FPGA), a programmable logic device (PLD), a state machine, gated logic, discrete hardware circuits, and other suitable processing components configured to perform the various functions described throughout the present disclosure. The functionality of a processor, any portion of a processor, or any combination of processors presented in the present disclosure may be implemented with software being executed by a microprocessor, micro-controller, DSP, or other suitable platform.

Software shall be construed broadly to mean instructions, instruction sets, code, code segments, program code, programs, subprograms, software modules, applications, software applications, software packages, routines, subroutines, objects, threads of execution, procedures, functions, etc.

The software may reside on a computer-readable medium. A computer-readable medium may include, by way of example, memory such as a magnetic storage device (e.g., hard disk, floppy disk, magnetic strip), an optical disk, a smart card, a flash memory device, random access memory (RAM), read only memory (ROM), programmable ROM (PROM), erasable PROM (EPROM), electrically erasable PROM (EEPROM), a register, or a removable disk. Although a memory is shown as being separate from the processor in various aspects presented in this disclosure, a memory may also be internal to the processor (e.g., a cache or a register).

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. All structural and functional equivalents to the elements of the various aspects described throughout the present disclosure that are known or later come to be known to those of ordinary skilled in the art are intended to be encompassed by the claims.

CLAIMS

1. A method for temporal co-contrastive learning-based node representation generation, comprising:

converting interaction data into a dynamic graph, the dynamic graph comprising a set of snapshots, each snapshot corresponding to a time step and comprising a set of nodes and a set of edges;

for each node in a plurality of nodes in the dynamic graph, generating a snapshot view node representation, a temporal view node representation and an interaction view node representation of the node at each time step;

generating a comprehensive node representation of the node at the time step based on the snapshot view node representation, the temporal view node representation and the interaction view node representation of the node at the time step; and

optimizing the comprehensive node representation based at least on a temporal-consecutive node contrastive task and a global view contrastive task.

2. The method of claim 1, wherein

the interaction data is data about interaction between a plurality of users and a plurality of items,

the set of nodes comprises user nodes and item nodes, and

each edge in the set of edges indicates that there exists interaction between a user node and an item node linked by the edge within a time window associated with the snapshot.

3. The method of claim 1, wherein the generating a snapshot view node representation comprises:

generating the snapshot view node representation of the node at the time step through applying a graph attention network to a snapshot corresponding to the time step.

4. The method of claim 1, wherein the generating a temporal view node representation comprises:

generating the temporal view node representation of the node at the time step through aggregating at least one historical snapshot view node representation of the node.

5. The method of claim 1, wherein the generating an interaction view node representation comprises:

identifying at least one interacted node having historical interaction with the node; and

generating the interaction view node representation of the node at the time step through aggregating at least one snapshot view node representation of the at least one interacted node at the time step.

6. The method of claim 1, wherein the generating a comprehensive node representation

comprises:

generating the comprehensive node representation of the node at the time step through applying a multi-layer perceptron to the snapshot view node representation, the temporal view node representation and the interaction view node representation.

7. The method of claim 1, wherein the temporal-consecutive node contrastive task at least comprises:

taking the snapshot view node representation of the node at the time step and a previous snapshot view node representation of the node at the previous time step as a positive sample pair for contrastive learning.

8. The method of claim 1, wherein the global view contrastive task at least comprises:

taking the temporal view node representation and the interaction view node representation of the node at the time step as a positive sample pair for contrastive learning.

9. The method of claim 1, wherein the optimizing the comprehensive node representation comprises:

establishing a loss function based at least on the temporal-consecutive node contrastive task and the global view contrastive task; and

optimizing the comprehensive node representation based on the loss function.

10. The method of claim 1, wherein

the comprehensive node representation is optimized further based on a graph context prediction task.

11. The method of claim 1, further comprising, for two nodes in the plurality of nodes:

predicting, based on comprehensive node representations of the two nodes at the current time step, whether there exists interaction between the two nodes at the next time step.

12. The method of claim 11, wherein

the two nodes are a user node and an item node respectively.

13. An apparatus for temporal co-contrastive learning-based node representation generation, comprising:

at least one processor; and

a memory storing computer-executable instructions that, when executed, cause the at least one processor to:

convert interaction data into a dynamic graph, the dynamic graph comprising a set of snapshots, each snapshot corresponding to a time step and comprising a set of nodes and a set of edges,

for each node in a plurality of nodes in the dynamic graph, generate a snapshot view node representation, a temporal view node representation and an interaction view node

representation of the node at each time step,

generate a comprehensive node representation of the node at the time step based on the snapshot view node representation, the temporal view node representation and the interaction view node representation of the node at the time step, and

optimize the comprehensive node representation based at least on a temporal-consecutive node contrastive task and a global view contrastive task.

14. The apparatus of claim 13, wherein the optimizing the comprehensive node representation comprises:

establishing a loss function based at least on the temporal-consecutive node contrastive task and the global view contrastive task; and

optimizing the comprehensive node representation based on the loss function.

15. A computer program product for temporal co-contrastive learning-based node representation generation, comprising a computer program that is executed by at least one processor for:

converting interaction data into a dynamic graph, the dynamic graph comprising a set of snapshots, each snapshot corresponding to a time step and comprising a set of nodes and a set of edges;

for each node in a plurality of nodes in the dynamic graph, generating a snapshot view node representation, a temporal view node representation and an interaction view node representation of the node at each time step;

generating a comprehensive node representation of the node at the time step based on the snapshot view node representation, the temporal view node representation and the interaction view node representation of the node at the time step; and

optimizing the comprehensive node representation based at least on a temporal-consecutive node contrastive task and a global view contrastive task.

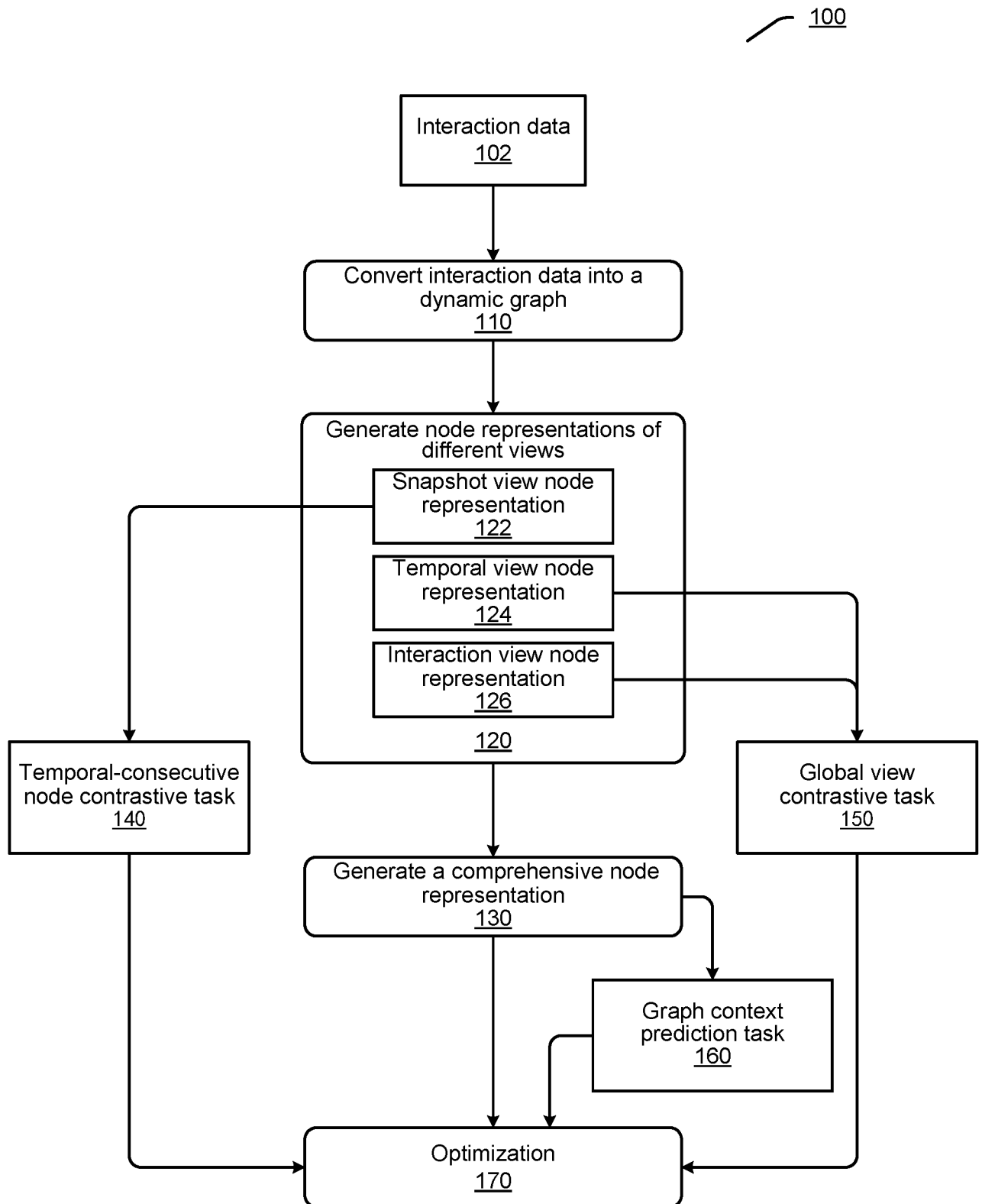


FIG 1

200

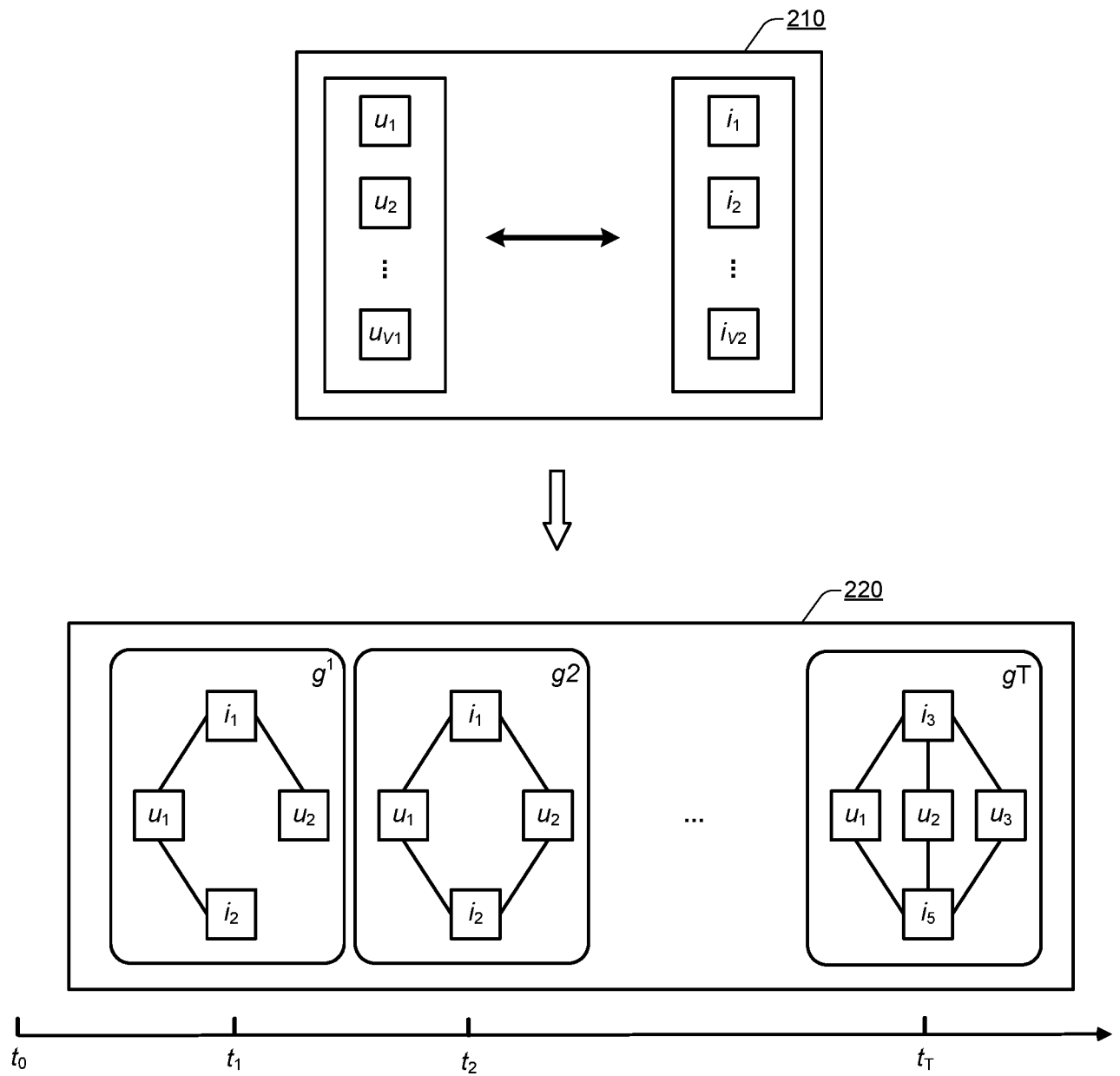


FIG 2

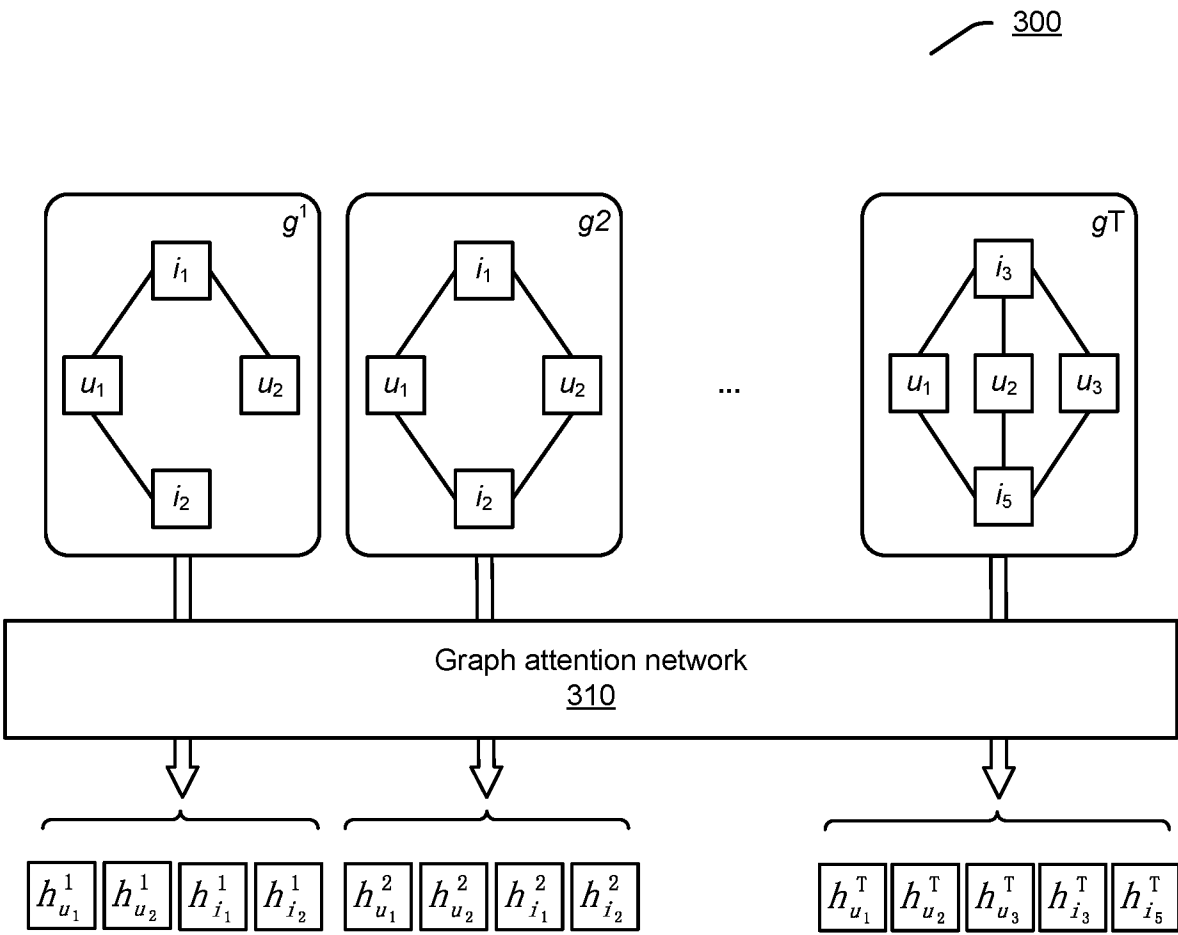


FIG 3

400

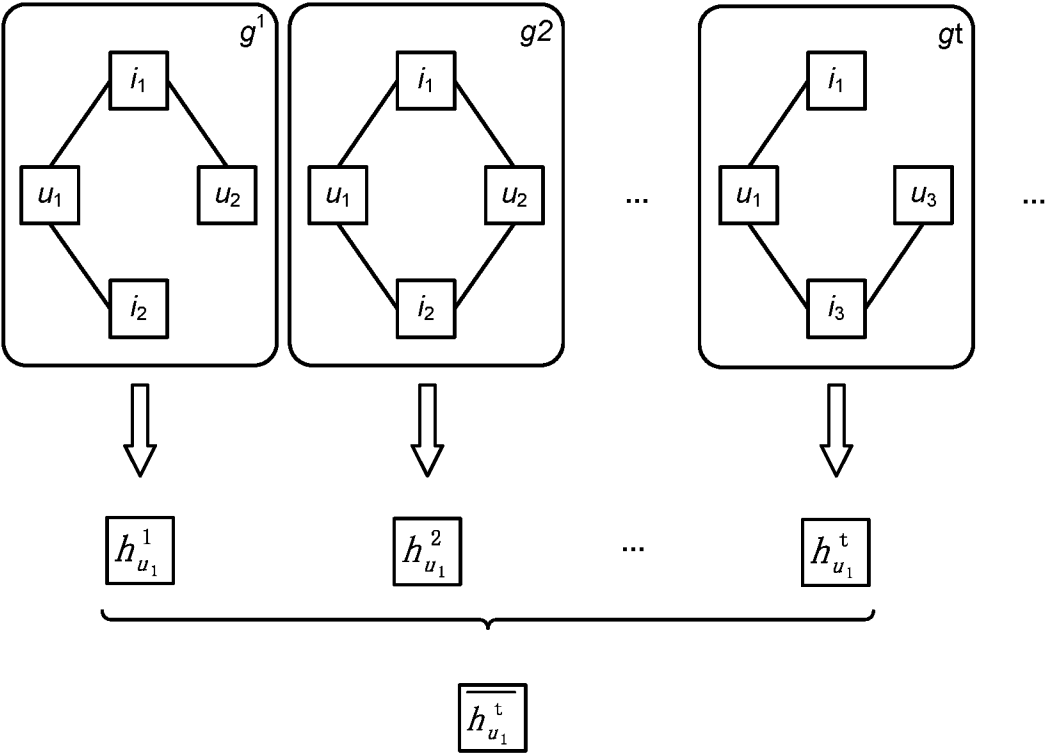


FIG 4

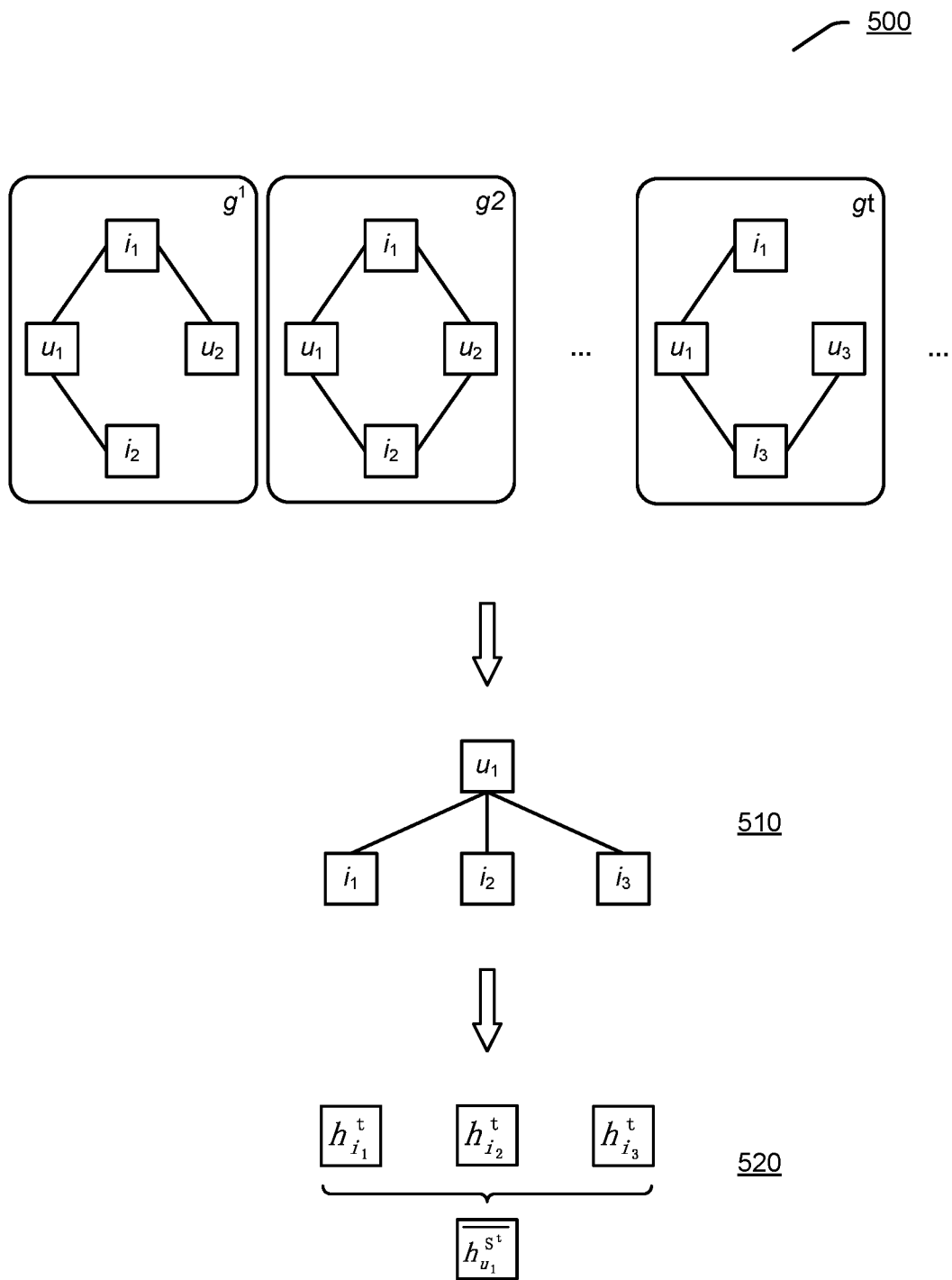
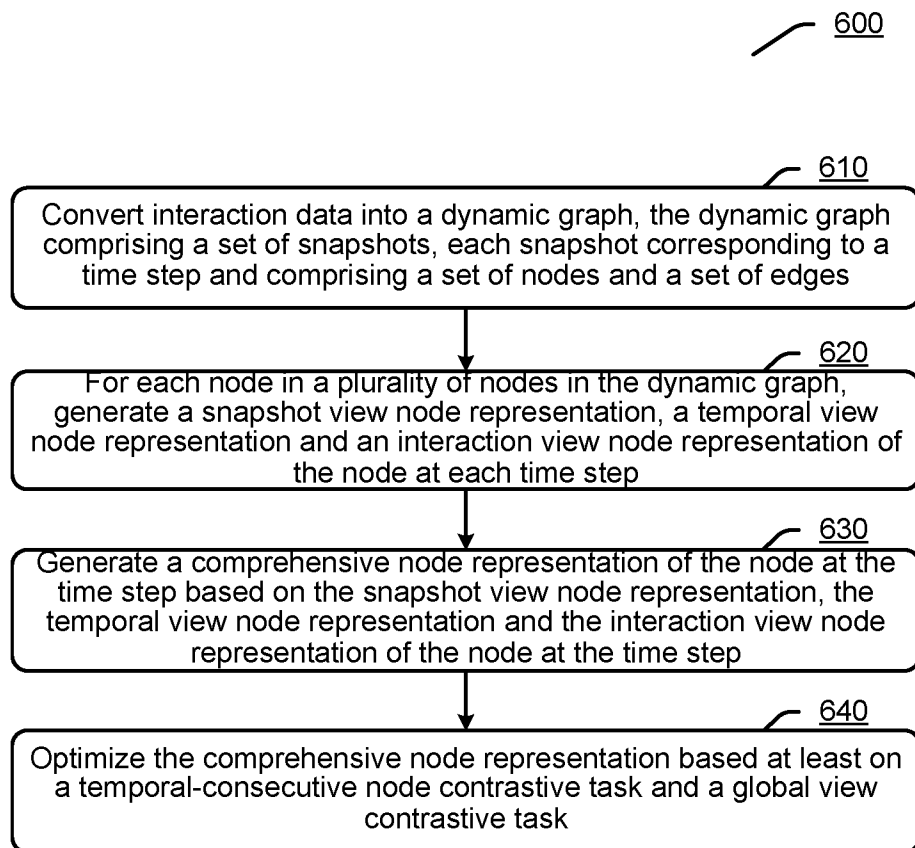
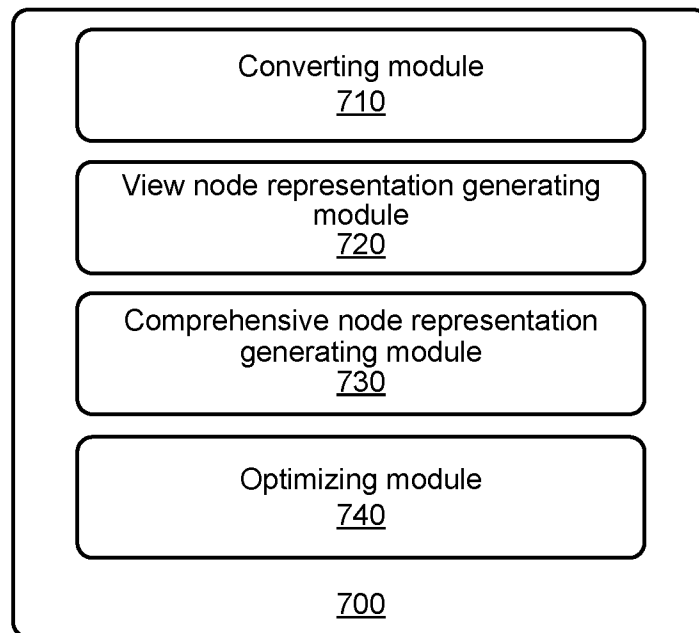
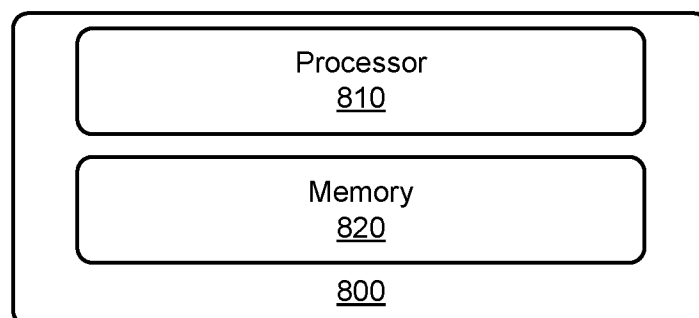


FIG 5

**FIG 6**

**FIG 7****FIG 8**

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2023/012368

A. CLASSIFICATION OF SUBJECT MATTER INV. G06F16/9535 G06N3/045 G06N3/048 G06N3/0895 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06N 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) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LU WANG ET AL: "TCL: Transformer-based Dynamic Graph Modelling via Contrastive Learning", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 17 May 2021 (2021-05-17), XP081964257, the whole document -----	1-15
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 4 May 2023		Date of mailing of the international search report 12/05/2023
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Ntarlagiannis, V