Learning Evolving Hostilic Correlations: Quaternion-Based Temporal Knowledge Graph Embedding

Jie Chen[1,2,3], Yinlong Wang[1,2,3], Wenjing Hu[1,2,3], Shu Zhao[1,2,3,†], Sha Wei[4], Yanping Zhang[1,2,3]

[1] *School of Computer Science and Technology, Anhui University, Hefei, 230601, China*
[2] *Key Laboratory of Intelligent Computing and Signal Processing, Ministry of Education, Hefei, 230601, China*
[3] *Information Materials and Intelligent Sensing Laboratory of Anhui Province, Hefei, 230601, China*
[4] *China Academy of Information and Communications Technology (CAICT), Beijing, 100191, China*

Abstract

Temporal Knowledge Graphs (TKGs) differ from conventional static ones by incorporating dynamic facts tied to specific timestamps, thus capturing the evolution of knowledge over time. While the additional time constraints also introduce challenges for the representation and reasoning of dynamic knowledge, the field of Temporal Knowledge Graph Embedding has attracted significant attention. However, previous TKGE methods inadequately considered the correlations, which could be from both individual local quadruples (Intra-Correlations), and multiple global KG snapshots (Inter-Correlations), of which the limitation results in less expressive representation and leads to suboptimal reasoning performance. To address this, we propose TI^{2}C, an embedding-based model leveraging quaternion algebra to represent temporal factual quadruples, which adopts a holistic perspective that accounts for both Intra- and Inter- Correlations. On the one hand, we preserve timestamp information with representation of temporal entities and capture the Intra-Correlations by refined scoring function based on rotation in 3-D quaternion space. On another, we integrate a Relevant Snapshot Perceiving (ReSP) module aiming to gather potentially inter-correlated quadruples by similarity calculation, then generate the synthetic footprint embedding to trace the path along the historical timeline. Extensive experiments on ICEWS and GDELT datasets suggest that our model consistently outperforms the state-of-the-art baselines for Link Prediction task[1].

Keywords: Knowledge Graph Embedding; Temporal Knowledge Graph; Knowledge Representation and Reasoning; Quaternion; Link Prediction;

1. Introduction

Knowledge Graphs (KGs) serve as repositories for human knowledge in a structured manner. As many human facts are temporally valid or occur within specific timestamps, the static nature of KGs poses challenges, with factual triples potentially becoming outdated over time. For example, (*Cristiano Ronaldo*, *member of club*, *Real Madrid FC*) was valid only during the years 2009-2017. To address this challenge, Temporal Knowledge Graphs (TKGs) are introduced to

[†]Corresponding author: Shu Zhao (Email: zhaoshuzs2002@hotmail.com; ORCID: 0000-0001-7647-3603)
[1]Our source codes are available at `https://github.com/Revisionisim/TI2C`

represent the dynamic facts for understanding knowledge evolution, which is crucial for data intelligence. In TKGs, a fact is encoded as a quadruple (s, r, o, τ), where τ marks the specific time when the fact is true, so the former fact could be stored as: (*Cristiano Ronaldo*, *member of club*, *Real Madrid FC* , 2009-2017). This structure facilitates more accurate representation and retrieval of information, supporting applications that require tracking the temporal progression of knowledge.

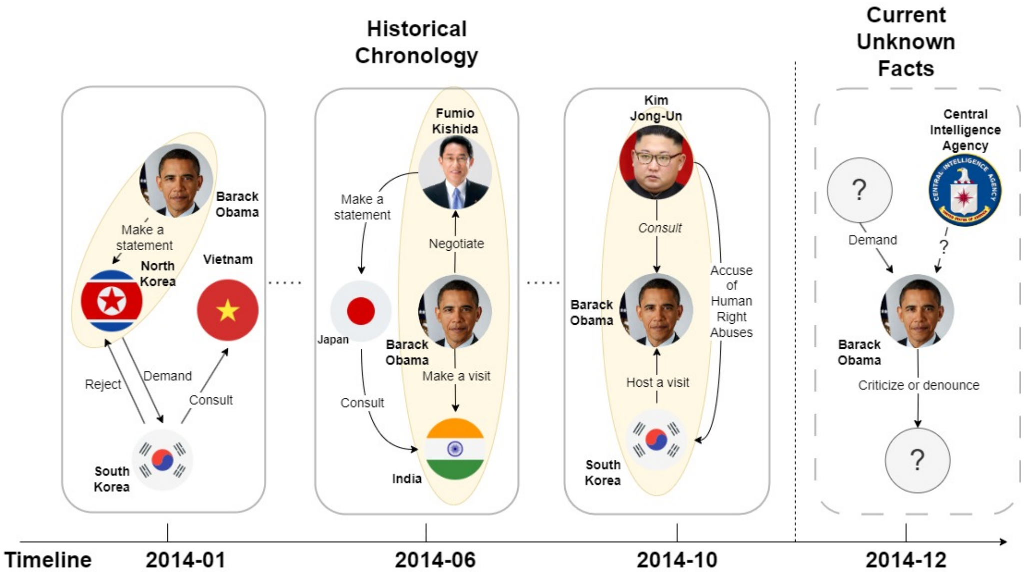


Figure 1: An example of reasoning over TKG by historical knowledge. Enitities and relations are extracted from ICEWS14 dataset. To infer the missing part at current time, our model learns the intra-correlations among entities,relaitons and timestamps within each snapshot, and capture the inter-correlated snapshots(highlighted in the figure) to trace the footprint by the proposed ReSP module. Note that we use *negotiate* to represent *engage in negotiation* for short.

Although large-scale TKGs contain billions of facts, the inherent incompleteness undermines their performance in downstream applications, including question answering, information retrieval[1]. This limitation has motivated the task of Temporal Knowledge Graph Embedding, which represents the temporal quadruples as low-dimensional vectors and infers missing links based on pre-designed score function. Many TKGE approaches [2, 3, 4] applied and extended Knowledge Graph Embedding (KGE) models [5, 6, 7], which are originally designed for static Knowledge Graphs without timestamps. Other works propose new models by introducing various methods, including Archimedean spiral [8], polynomial approximation [9], tensor decomposition [10], Heterogeneous Geometric Subspaces [11] and so on.

However, previous TKGE models struggle to capture the dual correlations in TKG adequately, as they primarily rely on simple concatenation or product operations. On the one hand, it remains challenging to capture the nuanced semantic features of individual entities, due to the rich correlations with their connecting relations. For example, consider quadruple (*Barack Obama*, *married*, *Michelle Obama*, 1992-10) and (*Barack Obama*, *host a visit*, *South Korea*, 2014-10) , entity *Barack Obama* embodies multifaceted attributes: the former denotes his family ties, while the latter emphasizes the political career. Additionally, distinguishing asynchronous occurrences of the same entities, and preserving temporal information to establish timestamp linkage poses significant challenges. Capturing such latent intra-correlations within temporal information is difficult for existing methods. On the other hand, facts in TKGs are inherently time-relevant and sequentially ordered. As shown in Figure 1, quadruples (*Kim Jung-Un*, *consult*, *Barack Obama*, 2014-10) and (*Barack Obama*, *Make a statement*, *North Korea*, 2014-01) show significant logical correlations, which are from different KG snapshots on the timeline. We believe that abundant valuable historical knowledge could be mined from the inter-correlated KG snapshots along the chronological timeline, which could provide valuable proofs for reasoning the missing facts.

To overcome above challenges, we leverage capabilities of Hamilton's quaternions and introduce an advanced **T**emporal Knowledge Graph Embedding (TKGE) model namely(**TI^{2}C**), under a thorough dual-consideration of **I**ntra- and **I**nter- **C**orrelations. TI^{2}C maps the quadruples into hyper-complex quaternion space, obtaining refined representation of entities and relations along with timestamps, and employs well-designed mathematical formulations based on quaternion rotation theory to capture the intra-correlations among them. Moreover, to capture inter-correlations within alike KG snapshots over time, we introduce an additional module called Relevant Snapshots Perceiving (ReSP), which identifies and collects likely inter-correlated KG snapshots within a manually-tunable duration of timeframe. Specifically, ReSP measures the correspondence between the current snapshot and historical ones at different timestamps by similarity function, and concatenates the entities from top-*m* inter-correlated snapshots to generate a synthetic embedding. The synthetic embedding presents valuable traces of the different chronological snapshots and provide historical knowledge to aid in reasoning tasks, thus called footprint embedding. Empirical experiments conducted on widely recognized benchmarks demonstrate that our proposed model consistently outperforms current state-of-the-art baselines for Temporal Knowledge Graph Reasoning (TKGR) tasks. Our main contributions can be summarized as follows:

1. We introduce an efficient quaternion-based Temporal Knowledge Graph Embedding (TKGE) model namely TI^2C, which efficiently leverages both intra- and inter-correlations within Temporal Knowledge Graphs (TKGs), to enhance expressiveness of knowledge representation and performance on reasoning.
2. By employing refined knowledge representation based on quaternions, TI^{2}C models intricate correlations within entities,relations and timestamps inside each quadruple, and we further develop a integrated module namely Relevant Snapshot Perceiving (ReSP) to capture potentially inter-correlated KG snapshots, under the propose of retrieving historical knowledge along the timeline.
3. Extensive experiments were conducted on three well-known TKG datasets to evaluate the effectiveness of our proposed model. The empirical results of improvements across all metrics demonstrate that our model achieves state-of-the-art performance for Link Prediction task on TKGs.

2. Related Work

2.1. TKGE Models

Knowledge Graph Embedding(KGE), which projects entities and relations into a lower-dimensional continuous space, pivots on designing expressive score functions to better assess plausibility. Based on the type of scoring function, these models could be roughly divided into

three categories: a). Translational: inspired by the translation invariance principle, TransE [12] interprets the relations as vector translation, while its simplicity motivates the following works [13, 14, 15, 6]; b). Rotational: in parallel, rotation models [16, 17, 18, 19, 5], initiated by RotatE [5], embed entities onto complex plane and treat the relations as rotation. c). Integrated: to combine translation and rotation into score functions for mutual benefit have been trending in recent years. [20, 21, 22, 23, 24, 25].
However, previous KGE models were designed for static KGs, thus struggle to deal with the temporal information in TKGs. Therefore, TKGE models extend traditional ones by incorporating additional timestamps, to represent and reason time-aware facts. Many TKGE models are built upon static KGE models [5, 26, 23, 7], including Tero [2], BoxTE [3] , RotateQVS [27], ComTR [28] and PTKE [4]. More recently, HGE [11] leverages diverse geometric subspaces to capture various temporal patterns and introduces a temporal-geometric attention mechanism. TeAST [8] utilizes an Archimedean Spiral Timeline and transforms quadruples into a third-order tensor. SToKE [29] captures structural and temporal contexts by incorporating relevant matrices and constructing tree-based evolution of events for each query. HaSa [30] addressed the issue of false negatives by the proposed contrastive learning approach and improved the performance of InfoNCE loss. IME [31] projected TKGs into multi-curvature spaces and incorporated different key properties to capture geometric structures. Our model falls under the TKGE models.

2.2. *Other Models*

Apart from embedding models, there are various methods for the Link Prediction task on TKGs. We conclude and categorize them as follows: a). Based on logical rules [32, 33, 34, 35, 36] make inferences by learning symbolic characteristics of KGs; b). [37, 38, 39] incorporate reinforcement learning to construct the reward function to perform multi-hop reasoning. c). [40, 41, 42, 43] adopt graph networks (GCN, GNN) to learn the evolutions of entities and relations at each timestamp. d). To leverage the power of pre-trained language models(PLMs), [29, 44, 45] treat TKG as textual input sequences and make predictions by the learning results of language models. More specific introduction for some models could be seen in the Baseline 5.1 section.

3. Preliminaries

3.1. *Notations*

A temporal knowledge graph $\mathcal{G}_{\mathbf{t}}$, formally defined by the entity set $\mathcal{E}$, relation set $\mathcal{R}$ and timestamp set $\mathcal{T}$, and is a structured collection of factual quadruples $\mathcal{D} = \{(s, r, o, \tau)\} \subseteq \mathcal{E} \times R \times \mathcal{E} \times \mathcal{T}$. A TKG could also be formulized as a sequence of knowledge graph snapshots ordered by timestamp, i.e., $S_{\mathcal{G}} = \{G_1, G_2, \ldots, G_t, \ldots\}$, while the snapshot at timestamp $t \in [0, \tau]$ is the set of time-relevant factual quadruples. We need to state in advance is that we will use bold items to denote vector representations in later context. More symbol definitions are listed in Table 1.

3.2. *Task Formulation*

The Link Prediction task on TKGs aims to infer the missing object entity o via answering query like (s, r, ?, τ_q) with given historical information $\left\{(s, r, o, \tau_i) : \tau_i < \tau_q\right\}$ from the earlier KG snapshot sequence: $S_{\mathcal{G}} = \left\{G_{\tau_1}, G_{\tau_2}, \ldots, G_{\tau_q}, \right\}$, with the facts in period τ_q remaining unknown. Specifically, we consider $\mathcal{E}$ as the candidates and rank them by descending order within the score function, to measure the plausibility of predicted answers for the given query. Besides, for each

Table 1: **Notations&Description in this paper**

Notation	Description
$\mathcal{G}_\mathbf{t}$	A Temporal Knowledge Graph
$\mathcal{S}_\mathcal{G}$	KG Snapshots
$s, o \in \mathcal{E}$	Infinite set of subject,object entities
$r \in \mathcal{R}$	Infinite set of relations
$\tau \in \mathcal{T}$	Infinite set of timestamps
(s, r, o, τ)	A quadruple
$\mathcal{F}$	Score function
$\mathcal{L}$	Loss function
Q	Quaternion
$W_r^{\triangleleft}$	Normalized relational rotation matrix
$\mathbb{R}^k$	k dimensional real-valued space
$\mathbb{C}^k$	k dimensional complex space
$\mathbb{H}^k$	k dimensional hypercomplex space
$\cdot$	Inner product
$\circ$	Hadmard product
$\otimes$	Hamilton product

fact in $\mathcal{G}_\mathbf{t}$, we add inverse quadruple(o, r^{-1}, s, τ_i) into $\mathcal{G}_\mathbf{t}$, correspondingly. By doing so, the reasoning task could be generalized to predicting the missing object entity, as predicting subject entity could be converted to its inverse quadruple.

3.3. Quaternion Algebra

Quaternion is an extension of complex algebra system, extending traditional complex number system to four-dimensional space. A quaternion $\mathbf{Q} \in \mathbb{H}$ is defined as $Q = a + bi + cj + dk$, where $a, b, c, d \in \mathbb{R}$ and $i, j, k \in \mathbb{C}$ are imaginary units. A quaternion could be written in the form of a scalar and a vector: $\mathbf{Q} = [a, \mathbf{v}], \mathbf{v} = (b, c, d)$. Some basic definitions and operations of the quaternion are listed as follows:

Product of *i, j, k*: The multiplication between any two quaternion units are:

$$\begin{gathered} i^2 = j^2 = k^2 = -1 \\ ij = k \quad jk = i \quad i = j \\ ji = -k \quad kj = -i \quad ik = -j \end{gathered} \tag{1}$$

where we can easily tell that the multiplication rule is non-commutative.

Conjugate: The conjugate of a quaternion Q is defined as:

$$\overline{Q} = a - bi - cj - dk \tag{2}$$

Norm: The norm of a quaternion Q is defined as:

$$|Q| = \sqrt{a^2 + b^2 + c^2 + d^2} \tag{3}$$

Inner Product: The quaternion inner product between Q_1 and Q_2 is obtained by summing up the inner products between corresponding scalar and imaginary components:

$$Q_1 \cdot Q_2 = \langle a_1, a_2 \rangle + \langle b_1, b_2 \rangle + \langle c_1, c_2 \rangle + \langle d_1, d_2 \rangle \tag{4}$$

Hamilton Product: The Hamilton product consists of multiplications of each factor in quaternions:

$$\begin{aligned} Q_1 \otimes Q_2 = &(a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) \\ &+(a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)i \\ &+(a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)j \\ &+(a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)k \end{aligned} \tag{5}$$

Quaternion Rotation: let $u^p = e^{u\frac{\theta}{2}} = \cos\frac{\theta}{2} + \sin\frac{\theta}{2}u$, where u is a unit vector. Then we can get a pure quaternion(the scalar part a = 0)rotating θ around axis u by:

$$Q' = u^p \otimes Q \otimes \overline{u}^p \tag{6}$$

4. Methodology

In this section, we introduce an innovative TKGE model called **TI**^{2}**C** (**T**ime-evolving **I**ntra-and-**I**nter-**C**orrelation), which is specifically tailored to efficiently incorporate temporal information and capture two fundamental types of correlations within TKGs, including Intra-Correlations and Inter-Correlations. This process is modeled as a quaternion rotation, as visually depicted in Figure 2.

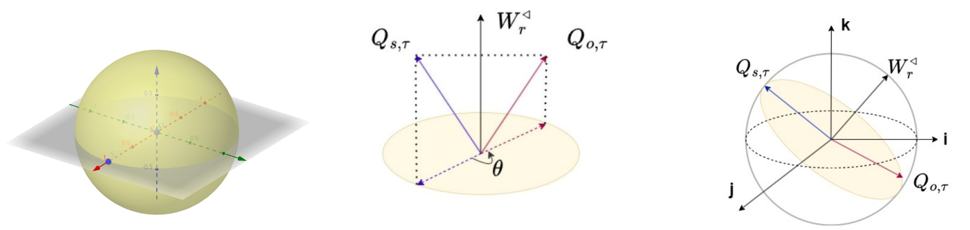


Figure 2: The overall presentation of the quaternion rotation: (a) 3D quaternion space. (b) Quaternion Rotation. (c) The illustration of TI2C model.

4.1. *Representing TKG with Quaternions*

Given a temporal knowledge graph $\mathcal{G} \subseteq \mathcal{E} \times R \times \mathcal{E} \times \mathcal{T}$, we use the quaternion matrix $\mathbf{Q}_{s,o} \in \mathbb{H}^{|\mathcal{E}| \times k}$ to represent subject and object entity embeddings, $\mathbf{W}_r \in \mathbb{H}^{|\mathcal{R}| \times k}$ to represents all relation embeddings, and $\tau \in \mathbb{H}^{|\mathcal{T}| \times k}$ to represent all timestamp embeddings, where k is the embedding dimensions. To satisfy the prerequisite of quaternion rotation, the subject and object entities are defined as pure quaternions:

$$\begin{aligned} \mathbf{Q}_s &= a_s i + b_s j + c_s k \\ \mathbf{Q}_o &= a_o i + b_o j + c_o k \end{aligned} \tag{7}$$

The primary challenge encountered by Temporal Knowledge Graph Embedding (TKGE) methods is to preserve the temporal validity of quadruples, aligning with the evolutionary nature of knowledge. Therefore, maintaining timestamp information within embeddings becomes crucial. In our model, we tackle this challenge by representing both the subject and object entities as time-evolving quaternion embeddings by following calculation:

$$\begin{aligned}\mathbf{Q}_{s,\tau} &= \tau Q_s \tau^{-1} \\ \mathbf{Q}_{o,\tau} &= \tau Q_o \tau^{-1}\end{aligned} \tag{8}$$

Where the coefficients $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}$ are rows from corresponding matrix, while $\mathbf{Q}_{s,\tau}$ and $\mathbf{Q}_{o,\tau}$ $\in \mathbb{H}$ are embeddings of subject and object entities under the current timestamp τ, respectively.

4.2. *Intra-Correlations Learning*

Definition 1. *Intra-Correlation: For a TKG, the correlations among entities,relations and timestamps within each KG snapshots is defined as Intra-Correlations*

To further encode the Intra-Correlations among entities,relations and timestamps, we employ refined representations based on quaternion rotation supported by Eq.6 :

$$\begin{aligned}\mathbf{Q}(s, r, \tau) &= W_r^{\triangleleft} \otimes \mathbf{Q}_{s,\tau} \otimes W_r^{\triangleleft^{-1}} \\ \mathbf{Q}(o, r, \tau) &= W_r^{\triangleleft} \otimes \mathbf{Q}_{o,\tau} \otimes W_r^{\triangleleft^{-1}}\end{aligned} \tag{9}$$

where $W_r^{\triangleleft}$ is the normalized relational rotation matirx, representing in Trigonometry:

$$W_r^{\triangleleft} = \cos\frac{\theta_r}{2} + u_r \sin\frac{\theta_r}{2} \tag{10}$$

Score Function: Finally, we formulated the score function of TI^{2}C as inner product of the subject and object embedding:

$$\mathcal{F}(s, r, o, \tau) = \mathbf{Q}(s, r, \tau) \cdot \mathbf{Q}(o, r, \tau) \tag{11}$$

4.3. *Inter-Correlations Learning*

Definition 2. *Inter-Correlation: For a TKG, the correlations among a sequence of KG snapshots along the timeline is defined as inter-correlation.*

Relevant Snapshot Perceiving: Existing Temporal Knowledge Graph Embedding (TKGE) models often struggle to fully exploit the correlations among Knowledge Graph (KG) snapshots over consecutive time intervals. To bridge this gap by capturing inter-correlations within sequential KG snapshots, we propose an independent module inspired by prior works [46, 41], termed Relevant Snapshot Perceiving (ReSP). ReSP aims to extract latent information from the most similar snapshots along the chronological timeline and refine the final representation. Notably, ReSP is seamlessly integrated into the score function to facilitate its implementation. The procedure of ReSP unfolds in three stages:

- *l-Temporal Horizon:* Initially, a l-Temporal Horizon is defined, within which ReSP identifies the relevant snapshots spanning a duration of length l. Subsequently, for each snapshot, ReSP ranks all candidate quadruples by computing their correspondence weights using cosine similarity:

$$\begin{aligned} W &= \{w_i | i \in [\tau - l, \tau - 1]\} \\ w_i &= cosine(\mathcal{S}_{\mathcal{G}_i}, \mathcal{S}_{\mathcal{G}_\tau}) \end{aligned} \tag{12}$$

- *Footprint Embedding Generating:* Then, we concatenate the entities embeddings retrieved from selected top-m inter-correlated snapshots, determined by the calculated correspondence weights, to obtain the footprint embedding.

$$\begin{aligned} \mathbf{Q}^f_{s,\tau} &= w_i \sum_{i=1}^{m} \mathbf{Q}_{s,\tau_i} \\ \mathbf{Q}^f_{o,\tau} &= w_i \sum_{i=1}^{m} \mathbf{Q}_{o,\tau_i} \end{aligned} \tag{13}$$

  where $w_i \in top_m(W)$ is the weight of the top-m KG snapshots, and Q_{s,τ_i}, Q_{o,τ_i} could be calculated by Eq. 8.

- *Historical Thread Weaving:* Subsequently, we employ the score function defined in Equation 11 to capture historical information at different timestamps $\{\tau_i | \in [\tau - l, \tau - 1]\}$ from the compounded footprint embeddings:

$$ReSP(s, r, o, \tau') = \mathbf{Q}^f_{s,\tau} \otimes \mathbf{Q}^f_{o,\tau} \tag{14}$$

4.4. Training Objective

We modify the margin loss function [5], which was originally used for traditional KGE models:

$$\begin{aligned} L = &- \log \sigma(\gamma - \mathcal{F}(\xi)) - \sum_{i=1}^{\eta} \log \sigma(\mathcal{F}(\xi_i') - \gamma) \\ &+ \lambda_1 ||Q||_2^2 + \lambda_2 ||W||_2^2 \end{aligned} \tag{15}$$

where η is the positive traning quadruple, γ is the number of negative samples per ξ, $\sigma(\cdot)$ is sigmoid function, and γ is a fixed margin. ξ_i' denotes the generated negative samples. Additionally, the model uses l_2 norm with ratio parameter λ_1 and λ_2 for regularization of Q and W.

5. Experiment

5.1. Experiment Setup

Datasets: We conducted experiments to perform the TKGR task on widely recognized datasets: ICEWS14 and ICEWS05-15 [47] are subsets of the Integrated Crisis Early Warning System (ICEWS) dataset, with the number indicating the occurring year of events. And GDELT is extracted from the Global Database of Events,Language and Tone, from which the events happen

Table 2: Dataset Statistics

Dataset	ICEWS14	ICEWS05-15	GDELT
$\lvert\mathcal{E}\rvert$	7128	10,488	500
$\lvert\mathcal{R}\rvert$	230	251	20
$\lvert\mathcal{T}\rvert$	365	4017	366
$\lvert\mathcal{G}_{Train}\rvert$	72,826	386,962	2,735,685
$\lvert\mathcal{G}_{Valid}\rvert$	8,963	46,092	341,961
$\lvert\mathcal{G}_{Test}\rvert$	8,941	46,275	341,961

during 2015/4/1 to 2016/3/31. Table 2 summarises the statistics of datasets, and we call tell that GDELT is much larger compared to ICEWS but has fewer entities and relations, which indicates a more data-intensive dataset.

Evaluation Protocols: Link prediction task aims to infer the missing entity for a given time-wise quer$(s, r, ?, \tau)$ or $(?, r, o, \tau)$. The model ranked all candidates by calculating the score of all possible quadruples with substitutions of the subject entity and the object entity: (s', r, o, τ) and (s, r, o', τ), where $s', o' \in \mathcal{E}$. The performance is evaluated using standard evaluation metrics: Hits@n reflects the ratio of correct predictions ranked in top n (Hits@1, Hits@3, and Hits@10), and the mean reciprocal rank (MRR) represents the average inverse ranks. Both reflect a more advanced performance with increasement on value.

Baselines: We compare our model with a wide variety of competitive baselines, including:

- **ALRE-IR [35]** combined embedding-based and logical rule-based methods for temporal knowledge graph reasoning by learning rule embeddings, enabling interpretable causal predictions and achieving superior performance over state-of-the-art models.
- **TFLEX [34]** proposed a framework of Temporal Feature-Logic Embedding to model all first-order logic operations on the entity set, and extend fuzzy logic on timestamp set to handle different temporal operators.
- **DREAM [37]** combined multi-faceted attention representation learning, which captures both semantic dependence and temporal evolution, with an adaptive reinforcement learning framework that learns reward functions for multi-hop reasoning.
- **RLAT [39]** integrated reinforcement learning with an attention mechanism, using LSTM and attention as memory components to improve reasoning path training, and employing an attention mechanism with influence factors to differentiate the importance of neighboring features.
- **IE-Evo [45]** integrated both internal structural evolution and external knowledge semantics from pre-trained language models, addressing data scarcity and capturing rich temporal dependencies.
- **L^{2}TKG [44]** used in-context learning with large language models and leveraged symbolic patterns from past facts, rather than relying on prior semantic knowledge.

- **TiRGN [40]** leverages local and global historical patterns through a recurrent graph network, capturing the sequential, repetitive, and cyclical dynamics of past events to improve the predictions of future facts.

- **RPC [41]** leveraged semantic correlations between relationships and periodic temporal patterns through novel relational and periodic correspondence units, improving the expressive ability and reasoning performance across snapshots.

- **BoxTE [3]** incorporated time stamps into box embeddings, offering full expressiveness and strong inductive capacity for predicting missing facts at different time steps.

- **PTKE [4]** embedded temporal facts into a polar coordinate system, where the modulus distinguishes different time-constrained entities, and the angular part differentiates entities with the same modulus, providing a more precise representation of temporal data within the translation-based model framework.

Implementation Details: All experiments are implemented on single NVIDIA 3090 GPU. The best models are selected by early stopping on the validation set with Hits@10. The max training epoch is 300. And we report the best hyper-parameters for each of three datasets as follows, respectively: For the dimension size of embeddings k, we set k=1000 , 1000 , 1200. The learning rate α is 0.1 for all. For the regularization rates, we set λ_1=0.002 , 0.001 and 0.0005, λ_2 = 0.01 , 0.1 and 0.01.

Table 3: Experimental results on three benchmark datasets. The best performances are in boldface, and the second best are underlined. Results of baselines are taken from original papers.

	ICEWS14				ICEWS05-15				GDELT			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
Logical-rule-based												
ALRE-IR [35]	0.540	0.428	0.612	0.718	0.602	0.490	0.677	0.775	-	-	-	-
TFLEX [34]	0.482	0.357	0.565	0.723	0.430	0.300	0.498	0.695	0.185	0.101	0.196	0.349
RL-based												
DREAM [37]	0.517	0.420	0.564	0.724	0.568	0.473	0.651	0.786	0.281	0.193	0.311	0.447
RLAT [39]	-	-	-	-	0.466	0.366	0.429	0.502	0.353	0.268	0.335	0.477
Text-based												
IE-Evo [45]	0.449	0.345	0.496	0.653	0.517	0.400	0.588	0.730	-	-	-	-
L^2TKG [48]	0.474	0.354	-	0.711	0.574	0.419	-	0.807	0.205	0.129	-	0.358
Graph-based												
TiRGN [40]	0.440	0.338	0.490	0.638	0.500	0.393	0.561	0.707	0.217	0.136	0.233	0.376
RPC [41]	0.446	0.349	0.498	0.651	0.511	0.395	0.571	0.718	0.224	0.144	0.244	0.383
Embedding-based												
Tero [2]	0.562	0.468	0.621	0.732	0.586	0.469	0.668	0.795	0.245	0.154	0.264	0.420
RotateQVS [27]	0.591	0.507	0.642	0.754	0.633	0.529	0.709	0.813	0.270	0.175	0.293	0.458
BoxTE [3]	0.613	0.528	0.664	0.763	0.667	0.582	0.719	0.820	0.352	0.269	0.377	0.511
PTKE [4]	0.554	0.444	0.624	0.760	0.572	0.454	0.649	0.789	-	-	-	-
TI^2C(Ours)	**0.628**	**0.549**	**0.675**	**0.774**	**0.678**	**0.599**	**0.726**	**0.822**	**0.367**	**0.283**	**0.391**	**0.530**

5.2. Experimental Results

Performance Benchmarking We compare our model against state-of-the-art baselines employing various approaches. The results detailed in Table 3 tell that our model consistently outperforms other baselines across all metrics. We unpack the main findings as follows.

i). Notably, our model exhibits superior performance, surpassing the best baseline by 1.6%, 2.4%, and 4.3% on Mean Reciprocal Rank (MRR) across three datasets, respectively.

ii). Particularly remarkable is our model's significant improvement on the dense GDLET dataset, attributed to its capability to extract valuable knowledge along with noisy information. Still, our model also shows great performance on the two sparser dataset ICEWS04 and ICEWS05-15, which proves its ability to deal with data sparseness.

iii). Compared to TeRo, our model demonstrates substantial enhancements over TeRo [2], which embeds the graph into a complex space and uses the Hadamard product. Specifically, we observe improvements of 6.6%, 9.2%, and 12. 2% in MRR, highlighting the superiority of quaternion embeddings and the Hamilton rotation-based scoring function.

iv). Last but not least, the notable performance of Temporal Knowledge Graph Embedding (TKGE) models, including ours, suggests the efficacy of this methodology for reasoning task.

In-depth Comparison: Compared to RotateQVS [27], our model leads by the margin of

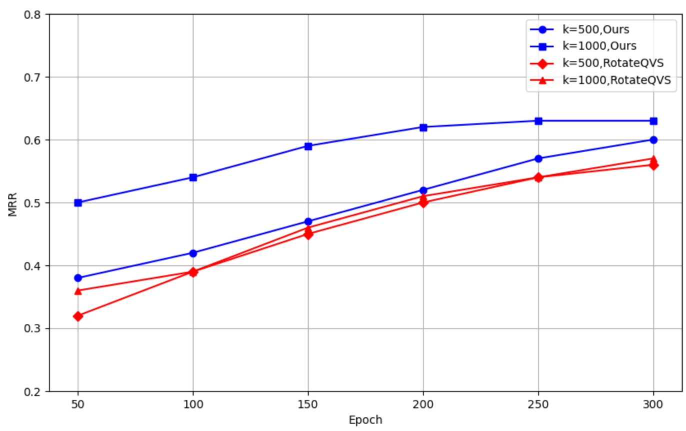


Figure 3: Comparison of RotateQVS and our model under differed embedding dimension

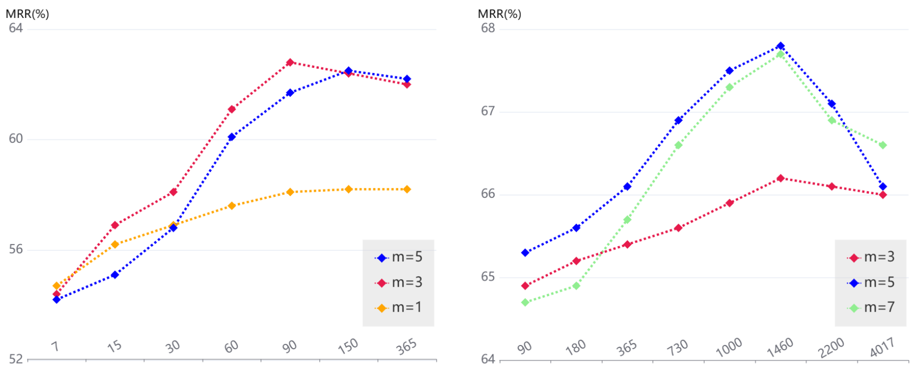
3.7%, 4.5% and 9.7%, respectively. Though our model shares similarities with it, there also exist some major differences. In this part, we explain why our model can achieve a significant performance improvement over previous RotateQVS model, and the reasons are summarized as follows:

1. **More advanced representation of temporal information**: Though our model and RotateQVS both apply quaternion representation, the scalars of imaginary parts are different. RotateQVS preserve temporal information by simply multiplication:$s_\tau = \tau s \tau^{-1}$; while our model divides embedding from timestamp matrix as scalars and models the timestamps

as different axes of quaternion space. This allows our model to show more flexibility in distinguishing the independence or strengthen the connection between them.

2. **More expressively-powerful scoring function**: Our model learns the intra-correlations by Halmiton product of quaternion representations, whose advantages have been discussed in QuatE. While RotateQVS applies translation by simply adding up the embedding of entity and relation, which lacks enough modelling ability. Besides, the score function of our proposed model significantly differs from RotateQVS, as our model calculates the similarity based on semantic matching function (dot product), while RotateQVS is based on distance translating (translation), and could be seen as the quaternion version of Tero. While for quaternion-based KGE models, semantic matching function[23, 49, 50, 51] perform generally better than distance translating one, thus are more popular.
3. **Better information retaining ability of embedding**: Another very important reason lies in the dimension of embedding, as a higher dimension can provide the embedding of the ability to retain information. The best performance of our model is achieved at 1000,1000 and 1200 for three datasets, while RotateQVS is reported to be 500 for all. To further analyze the differences, we conduct an experiment to compare RotateQVS and our model on ICEWS14 dataset with different dimension(we reproduce the results of RotateQVS by replacing the scoring function under our training framework), as shown in Figure 3. We can observe that when the dimension d=1000, the improvement of RotateQVS appears to be minimal. Therefore, the main reason for our performance lead could be allowing larger dimensions (while our model still performs better at the same degree).

5.3. Parameters Analysis



(a) ICEWS04 (b) ICEWS05-15

Figure 4: Analysis of hyperparameters m and l on ICEWS14 and ICEWS05-15 datasets

We conducted empirical analysis to investigate the impact of two hyperparameters of the ReSP module: the length of the temporal horizon l, and the number of selected inter-correlated snapshots m. Figure 4 visually illustrates the performance variance of MRR when l and m changes, from which we could observe that: For the ICEWS14 dataset, the optimal performance is achieved at l = 90 and m = 3 (see Figure 4(a)), while for ICEWS05-15, it occurs at l = 1460 and m = 5(see Figure 4(b)). In both datasets, performance initially increases and then decreases as l enlarges. This behavior may be attributed to an oversized temporal horizon: when l is very large, ReSP identifies inter-correlated snapshots from a distant past. Although the similarity score might be high, the assistance they provide could be minimal due to the remote timeframe.

5.4. Ablation Study

We conducted an ablation study to assess the individual contributions of the two correlation learning components. Table 4 presents the performance variations observed. Notably, the Intra-CL component is removed by simplifying the score function to $\mathbf{Q}s, \tau \cdot \mathbf{Q}o, \tau$. while disabling the Inter-CL component involves straightforward omission. From the findings in Table 4, it's evident that the complete TI^{2}C model with the ReSP module achieves unparalleled performance. Removal of either of the correlation learning components results in sub-optimal performance. This underscores the effectiveness of both Intra- and Inter-Correlation Learning in enhancing model performance.

Table 4: Ablation Study on the two correlations learning

	ICEWS14		ICEWS05-15	
	MRR	Hits10	MRR	Hits10
w/o Both	0.601	0.747	0.632	0.789
w/o Intra-CL	0.620	0.762	0.675	0.816
w/o Inter-CL	0.623	0.767	0.676	0.807
TI^{2}C	**0.628**	**0.774**	**0.678**	**0.822**

5.5. Case Study

To delve deeper into Inter-Correlation Learning, we conducted a case study on the ICEWS04 dataset to focus on the extracted inter-correlated quadruples. Table 5 presents the test queries and the top-m (m = 3) retrieved correlated quadruples , alongside their corresponding similarity scores, with the length of the temporal horizon l set to 150. For example, in the first query

Table 5: Case Study on ReSP capturing top-m inter-correlated quadruples from different KG snapshots along l-length timeframe. In this case, m=3 and l is set to 150 days (approximately 5 months)

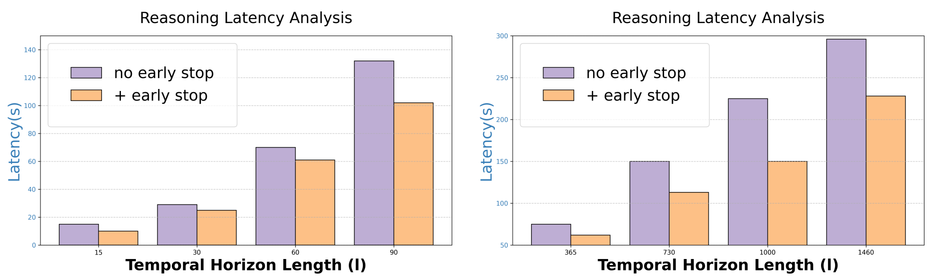
Test query	Top-3 correlated quadruples	Similarity Score
(Barack Obama, Make statement,?(Xi Jinping),2014-11-04)	(Barack Obama,Make a visit,Xi Jinping,2014-11-03)	0.87
	(Barack Obama, Meet at third location,Xi Jinping,2014-06-25)	0.80
	(Barack Obama,Make statement,Government (Syria),2014-09-12)	0.68
(?(China),Threaten, Protester(Hong Kong),2014-10-02)	(Protester(Hong Kong),Reject,China,2014-09-28)	0.79
	(China,Threaten,Vietnam,2014-05-07)	0.76
	(China,Threaten,North Korea,2014-05-20)	0.73

(Barack Obama, making a statement, ?, 2014-11-04), the task is to predict the correct object entity *Xi Jinping*. while the most correlated quadruple identified is *(Barack Obama making a visit to Xi Jinping on 2014-11-03)*, with the highest similarity score of 0.87. The finding provides significant historical information from adjacent times, illustrating the efficacy of ReSP in capturing inter-correlations between events across asynchronous KG snapshots.

5.6. *Further Analysis on ReSP*

Computational Efficiency: As ReSP requires traversing all the quadruples within a specific range of timeframe along the sequetial TKG snapshots, the potential computational cost could be a significant issue on the reasoning time. Therefore, to empirically analyze the computational efficiency, we record the reasoning delay during testing stage on ICEWS04 and ICEWS05-15 datasets. As shown in Figure 5 below, the whole reasoning delay of the best hyperparameters combination for ICEWS14(l=90,m=3) requires about 132s, while ICEWS05-15(l=1460,m=5) requires about 296s.

Additionally, to reduce the reasoning delay and achieve performance optimization, an early stop



(a) ICEWS04 (b) ICEWS05-15

Figure 5: Analysis of reasoning effiency on ICEWS14 and ICEWS05-15 datasets

method is applied to the ReSP module. To be specific, we set an m-size cache to store the top-m correlated snapshots, and record the highest score. The cache is updated when the current score is larger than the highest score, and if it has not been updated after a calculation of certain numbers of quadruples(we set the number to be 500), then the early stop is triggered. From Figure 5 below we can tell that our early stop does reduce the reasoning delay, and performs better with a larger dataset. It reduces the reasoning delay by about 3012ms and 6825ms for each dataset, respectively.

Transferability: As ReSP is an additional module, it can be applied to other TKGE model. Therefore, to demonstrate the transferability of ReSP, we replace our model with the score function of RotateQVS and reproduce the results, as shown in Table 6, which suggest the effectiveness and transferability of ReSP on other TKGE scoring functions.

Table 6: Experiments on the transferability of ReSP. * denotes our reproduced results.

	ICEWS14		ICEWS05-15	
	MRR	Hits10	MRR	Hits10
RotateQVS*	0.587	0.749	0.629	0.816
RotateQVS*+ReSP	0.601	0.767	0.640	0.837

6. Conclusion

In this paper, we introduce TI^{2}C, an efficient quaternion embedding-based model for Temporal Knowledge Graph representation and reasoning. On the one hand, we tailor the model score function by quaternion vector operations in hyper-complex space, to capture the intra-correlations among entities, relations and timestamps within KG snapshots. On the other hand, we propose a Relevant Snapshot Perceiving (ReSP) module to enhance the model's ability to retrieve relevant snapshots along the chronological timeline. By effectively capturing the Intra-and-Inter Correlations, the model refines the final representation by the structural nuances and temporal information within TKG, enhancing the model's overall performance on reasoning. Through extensive experiments on mainstream benchmark, we empirically evaluate our model, and the performance achieves state-of-the-art against recent competitive baselines.

Author Contributions

Jie Chen: Conceptualization, Writing - Review; Yinlong Wang: Methodology, Software; Wenjing Hu: Data Curation, Writing – original draft and editing; Shu Zhao: Resources, Supervision; Sha Wei: Formal analysis; Yanping Zhang: Validation.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 62376001), Anhui Province Excellent Scientific Research and Innovation Team (2024AH010004), and the China Scholarship Council, along with the sponsorship of the CCF-Zhipu AI Large Model Fund. We also appreciate the support from SMP-Zhipu.AI Large Model Cross-Disciplinary Fund (SMP-Zhipu20240210), Anhui University Smart Course Development and Pedagogical Innovation Project (2025xjjyzd003), Anhui University AI$^+$ Smart Education Curriculum Project(2024aijy044), Anhui University Teaching and Research Project (2024jyxm0081).

References

[1] C. Xu, Y.-Y. Chen, M. Nayyeri, and J. Lehmann, "Temporal knowledge graph completion using a linear temporal regularizer and multivector embeddings," in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp. 2569–2578, June 2021.

[2] C. Xu, M. Nayyeri, F. Alkhoury, H. Shariat Yazdi, and J. Lehmann, "TeRo: A time-aware knowledge graph embedding via temporal rotation," in *Proceedings of the 28th International Conference on Computational Linguistics*, pp. 1583–1593, Dec. 2020.

[3] J. Messner, R. Abboud, and I. I. Ceylan, "Temporal knowledge graph completion using box embeddings," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, pp. 7779–7787, 2022.

[4] R. Liu, G. Yin, Z. Liu, and L. Zhang, "Ptke: Translation-based temporal knowledge graph embedding in polar coordinate system," *Neurocomputing*, vol. 529, pp. 80–91, 2023.

[5] Z. Sun, Z.-H. Deng, J.-Y. Nie, and J. Tang, "Rotate: Knowledge graph embedding by relational rotation in complex space," in *International Conference on Learning Representations*, 2019.

[6] R. Abboud, I. Ceylan, T. Lukasiewicz, and T. Salvatori, "Boxe: A box embedding model for knowledge base completion," in *Advances in Neural Information Processing Systems*, vol. 33, pp. 9649–9661, 2020.

[7] Z. Zhang, J. Cai, Y. Zhang, and J. Wang, "Learning hierarchy-aware knowledge graph embeddings for link prediction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020.

[8] J. Li, X. Su, and G. Gao, "Teast: Temporal knowledge graph embedding via archimedean spiral timeline," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 15460–15474, 2023.

[9] Z. Fang, J. Qin, X. Zhu, C. Yang, and X.-C. Yin, "Arbitrary time information modeling via polynomial approximation for temporal knowledge graph embedding," in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pp. 1455–1465, 2024.

[10] F. Zhang, H. Chen, Y. Shi, J. Cheng, and J. Lin, "Joint framework for tensor decomposition-based temporal knowledge graph completion," *Inf. Sci.*, vol. 654, no. C, 2024.

[11] J. Pan, M. Nayyeri, Y. Li, and S. Staab, "Hge: Embedding temporal knowledge graphs in a product space of heterogeneous geometric subspaces," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, pp. 8913–8920, 2024.

[12] A. Bordes, N. Usunier, A. Garcia-Durán, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," in *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*, 2013.

[13] Y. Peng and J. Zhang, "Lineare: Simple but powerful knowledge graph embedding for link prediction," in *IEEE International Conference on Data Mining, ICDM* (C. Plant, H. Wang, A. Cuzzocrea, C. Zaniolo, and X. Wu, eds.), pp. 422–431, 2020.

[14] R. Li, J. Zhao, C. Li, D. He, Y. Wang, Y. Liu, H. Sun, S. Wang, W. Deng, Y. Shen, X. Xie, and Q. Zhang, "House: Knowledge graph embedding with householder parameterization," in *Proceedings of the 39th International Conference on Machine Learning*, pp. 13209–13224, 2022.

[15] Q. Zhang, R. Wang, J. Yang, and L. Xue, "Knowledge graph embedding by reflection transformation," *Knowledge-Based Systems*, vol. 238, p. 107861, 2022.

[16] C. Gao, C. Sun, L. Shan, L. Lin, and M. Wang, "Rotate3d: Representing relations as rotations in three-dimensional space for knowledge graph embedding," in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pp. 385–394, 2020.

[17] Y. Dong, L. Wang, J. Xiang, X. Guo, and Y. Xie, "Rotatect: Knowledge graph embedding by rotation and coordinate transformation in complex space," in *Proceedings of the 29th International Conference on Computational Linguistics* (N. Calzolari, C.-R. Huang, H. Kim, J. Pustejovsky, L. Wanner, K.-S. Choi, P.-M. Ryu, H.-H. Chen, L. Donatelli, H. Ji, S. Kurohashi, P. Paggio, N. Xue, S. Kim, Y. Hahm, Z. He, T. K. Lee, E. Santus, F. Bond, and S.-H. Na, eds.), pp. 4918–4932, 2022.

[18] T. Le, N. Huynh, and B. Le, "Knowledge graph embedding by projection and rotation on hyperplanes for link prediction," *Appl Intell*, vol. 53, pp. 10340–10364, 2023.

[19] T. Le, N. Le, and B. Le, "Knowledge graph embedding by relational rotation and complex convolution for link prediction," *Expert Systems with Applications*, vol. 214, p. 119122, 2023.

[20] L. Chao, J. He, T. Wang, and W. Chu, "PairRE: Knowledge Graph Embeddings via Paired Relation Vectors," in *Annual Meeting of the Association for Computational Linguistics*, pp. 4360–4369, 2021.

[21] X. Ge, Y. C. Wang, B. Wang, and C.-C. J. Kuo, "Compounding Geometric Operations for Knowledge Graph Completion," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 6947–6965, 2023.

[22] X. Zhang, Q. Yang, and D. Xu, "TranS: Transition-based knowledge graph embedding with synthetic relation representation," in *Findings of the Association for Computational Linguistics: EMNLP 2022*, pp. 1202–1208, Dec. 2022.

[23] Z. Cao, Q. Xu, Z. Yang, X. Cao, and Q. Huang, "Dual quaternion knowledge graph embeddings," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

[24] T. Le, H. Tran, and B. Le, "Knowledge graph embedding with the special orthogonal group in quaternion space for link prediction," *Knowledge-Based Systems*, vol. 266, p. 110400, 2023.

[25] J. Li and Y. Yang, "Star: Knowledge graph embedding by scaling, translation and rotation," in *Artificial Intelligence and Mobile Services – AIMS 2022: 11th International Conference, Held as Part of the Services Conference Federation, SCF 2022, Honolulu, HI, USA, December 10–14, 2022, Proceedings*, pp. 31–45, 2022.

[26] S. Zhang, Y. Tay, L. Yao, and Q. Liu, "Quaternion knowledge graph embeddings," in *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pp. 2735–2745, 2019.

[27] K. Chen, Y. Wang, Y. Li, and A. Li, "RotateQVS: Representing temporal information as rotations in quaternion vector space for temporal knowledge graph completion," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 5843–5857, May 2022.

[28] R. Yu, T. Liu, J. Yu, W. Zhang, Y. Zhao, M. Yang, M. Zhao, and J. Guo, "Combination of translation and rotation in dual quaternion space for temporal knowledge graph completion," in *2023 International Joint Conference on Neural Networks*, pp. 01–08, 2023.

[29] Y. Gao, Y. He, Z. Kan, Y. Han, L. Qiao, and D. Li, "Learning joint structural and temporal contextualized knowledge embeddings for temporal knowledge graph completion," in *Findings of the Association for Computational Linguistics: ACL 2023*, pp. 417–430, 2023.

[30] H. Zhang and J. Zhang, "Hasa: Hardness and structure-aware contrastive knowledge graph embedding," *Proceedings of the ACM on Web Conference 2024*, 2023.

[31] J. Wang, Z. Cui, B. Wang, S. Pan, J. Gao, B. Yin, and W. Gao, "Ime: Integrating multi-curvature shared and specific embedding for temporal knowledge graph completion," in *Proceedings of the ACM on Web Conference 2024*, p. 1954–1962, 2024.
[32] S. Xiong, Y. Yang, F. Fekri, and J. C. Kerce, "Tilp: Differentiable learning of temporal logical rules on knowledge graphs," in *International Conference on Learning Representations*, 2022.
[33] G. Niu and B. Li, "Logic and commonsense-guided temporal knowledge graph completion," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, pp. 4569–4577, 2023.
[34] L. Xueyuan, H. E, C. Xu, G. Zhou, H. Luo, T. Hu, F. Su, N. Li, and M. Sun, "TFLEX: Temporal feature-logic embedding framework for complex reasoning over temporal knowledge graph," in *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.
[35] X. Mei, L. Yang, X. Cai, and Z. Jiang, "An adaptive logical rule embedding model for inductive reasoning over temporal knowledge graphs," in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pp. 7304–7316, Dec. 2022.
[36] I. Singh, N. Kaur, G. Gaur, and Mausam, "NeuSTIP: A neuro-symbolic model for link and time prediction in temporal knowledge graphs," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 4497–4516, Dec. 2023.
[37] S. Zheng, H. Yin, T. Chen, Q. V. H. Nguyen, W. Chen, and L. Zhao, "Dream: Adaptive reinforcement learning based on attention mechanism for temporal knowledge graph reasoning," in *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, p. 1578–1588, 2023.
[38] R. Liu, G. Yin, Z. Liu, and Y. Tian, "Reinforcement learning with time intervals for temporal knowledge graph reasoning," *Information Systems*, vol. 120, p. 102292, 2024.
[39] L. Bai, D. Chai, and L. Zhu, "Rlat: Multi-hop temporal knowledge graph reasoning based on reinforcement learning and attention mechanism," *Knowledge-Based Systems*, vol. 269, p. 110514, 2023.
[40] Y. Li, S. Sun, and J. Zhao, "Tirgn: Time-guided recurrent graph network with local-global historical patterns for temporal knowledge graph reasoning," in *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, pp. 2152–2158, 2022.
[41] K. Liang, L. Meng, M. Liu, Y. Liu, W. Tu, S. Wang, S. Zhou, and X. Liu, "Learn from relational correlations and periodic events for temporal knowledge graph reasoning," in *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1559–1568, 2023.
[42] Y. Li, L. Chen, C. Liu, R. Zhou, and J. Li, "Efficient and effective entity alignment for evolving temporal knowledge graphs," in *2023 IEEE International Conference on Data Mining*, pp. 349–358, 2023.
[43] M. Zhang, Y. Xia, Q. Liu, S. Wu, and L. Wang, "Learning long- and short-term representations for temporal knowledge graph reasoning," in *Proceedings of the ACM Web Conference 2023*, pp. 2412–2422, 2023.
[44] D.-H. Lee, K. Ahrabian, W. Jin, F. Morstatter, and J. Pujara, "Temporal knowledge graph forecasting without knowledge using in-context learning," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pp. 544–557, 2023.
[45] K. Liu, F. Zhao, G. Xu, and S. Wu, "Ie-evo: Internal and external evolution-enhanced temporal knowledge graph forecasting," in *2023 IEEE International Conference on Data Mining*, pp. 408–417, 2023.
[46] F. Zhang, Z. Zhang, X. Ao, F. Zhuang, Y. Xu, and Q. He, "Along the time: Timeline-traced embedding for temporal knowledge graph completion," in *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pp. 2529–2538, 2022.
[47] A. García-Durán, S. Dumančić, and M. Niepert, "Learning sequence encoders for temporal knowledge graph completion," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 4816–4821, Oct.-Nov. 2018.
[48] M. Zhang, Y. Xia, Q. Liu, S. Wu, and L. Wang, "Learning latent relations for temporal knowledge graph reasoning," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 12617–12631, 2023.
[49] J. Li, X. Su, F. Zhang, and G. Gao, "Transerr: Translation-based knowledge graph embedding via efficient relation rotation," in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, and N. Xue, eds.), pp. 16727–16737, 2024.
[50] J. Yang, X. Ying, Y. Shi, X. Tong, R. Wang, T. Chen, and B. Xing, "Learning hierarchy-aware quaternion knowledge graph embeddings with representing relations as 3d rotations," in *Proceedings of the 29th International Conference on Computational Linguistics*, pp. 2011–2023, 2022.
[51] S. Zhang, X. Liang, H. Tang, X. Zheng, A. X. Zhang, and Y. Ma, "Ducape: Dual quaternion and capsule network–based temporal knowledge graph embedding," *Acm Trans. Knowl. Discovery Data*, vol. 17, pp. 104:1–104:19, 2023.