Dual Quaternion Decoupling and Enhanced Temporal Embeddings for Temporal Knowledge Graph Completion

Yue Zhao[1,3], Wenbin Zhang[3], Jiazheng Guo[2],Tianyi Xu[2], Mei Yu[2], Mankun Zhao[2,*]
[1] School of New Media and Communication, Tianjin University, Tianjin, 300350, China.
[2] College of Intelligence and Computing, Tianjin University, Tianjin, 300350, China.
[3] Information and Network Center, Tianjin University, Tianjin, 300350, China.
{zhao_yue, zhangwenbin, gjz426, tianyi.xu, zmk, yumei}@tju.edu.cn

***Abstract*—Temporal Knowledge Graph Completion models are proposed to predict facts at specific times. However, existing models treat temporal information as complementary to the semantic information of entity (relation) and ignore independence between them, which leads to the loss of details and accuracy of temporal information. Furthermore, temporal semantic information is intrinsically linked within a time range but differs markedly between different time ranges. To address the above problems, we propose a Dual Quaternion Decoupling and Enhanced Temporal Embeddings for Temporal Knowledge Graph Completion model (DQD-ETE). To decouple semantic and temporal information, we model the semantic information as the real part and the temporal information as the dual part of the dual quaternion. In addition, we use the combination of prototype representations and contrastive learning to get more discriminative temporal semantic information embeddings. Extensive experiment results show that DQD-ETE produces state-of-the-art performances on well-known benchmark datasets for Temporal Knowledge Graph Completion.**

***Index Terms*—Temporal Knowledge Graph, Knowledge Graph Completion, Dual Quaternion, Knowledge Representation Learning.**

I. Introduction

Static Knowledge Graphs (SKGs) store facts about the real world in a constructed directed graph, typically represented in the form of triples(s, r, o), where s is subject, r is relation, and o is object. However, some facts only exist at specific times. For example, the fact (Joe Biden, Discuss by telephone, France) is valid on April 20, 2023. Specifically, Temporal Knowledge Graphs (TKGs) introduce temporal information into the triples, transforming the facts into quadruples (s, r, o, τ), where τ (timestamp) indicates the specific time the fact occurred. Therefore, this fact is represented as (Joe Biden, Discuss by telephone, France, 2023-4-20). Large-scale KGs (e.g., DBPedia [1], Freebase [2]) are widely used in recommendation [3], information retrieval [4], and other tasks for their structured knowledge. Yet they face completeness and complexity challenges with growing data volume. Consequently, Knowledge Graph Completion (KGC), which infers missing facts from existing knowledge, garnered significant attention. However, Static KGC models (e.g., RotatE [5] and QuatE [6]) that ignore temporal information perform well on SKGs but poorly on TKGs. This is mainly because the Temporal KGC task needs to consider the dynamic properties and interrelationships of entity and relation over time and the intrinsic connection between temporal information. Recent Temporal KGC models that focus on capturing the temporal information and dynamic nature of real-world facts (e.g., TBDRI [7], HERCULES [8] and SPA [9]) demonstrated very impressive completion performance on TKGs.

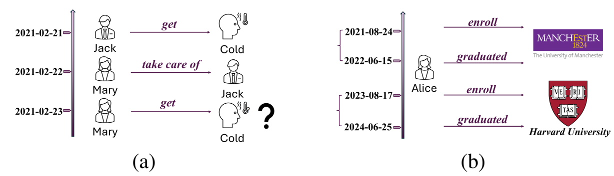


Fig. 1. Examples of the above problem of Temporal Knowledge Graph.

However, there are two problems with these existing TKGC models. (1) Currently existing models (e.g., RotateQVS [10], ComTR [11]) treat temporal information as complementary to the semantic information of entity (relation) during the construction process of temporal entity (relation), lacking an independent approach to model temporal semantic information. This makes it difficult for these models to interpret the independent semantic information and temporal information of temporal entity (relation), reducing the interpretability of these models. It also leads to the loss of details and accuracy of temporal information. Therefore, modelling semantic information and temporal information independently is necessary. (2) Although existing TKGC models focus on the complexity and variability of temporal information, they neglect the intrinsic connections of temporal information within a time range. This temporal correlation results in interconnected facts within a specific time range. As shown in Fig. 1(a), existing models struggle to accurately predict the fact (Mary, get, Cold, 2021-2-23) when they neglect the intrinsic connections. Moreover, the temporal semantic information associated with the same entity (relation) across different time ranges can vary significantly.

* Corresponding author.

As shown in Fig. 1(b), Alice's temporal semantics differ in these two time ranges: she is a student at different schools. Therefore, if we do not distinguish different time ranges, we cannot determine which school Alice attends, and cannot accurately predict her other behaviors. Thus, we need to pay attention to the correlation of temporal information within time ranges and the differentiation of temporal information between different time ranges. To our knowledge, very few of the existing TKGC models can capture these aspects.

To solve the above problems, we propose the model DQD-ETE. In this framework, to realise the decoupling of semantic information and temporal information, we embed temporal entity (relation) in the form of dual quaternions. It models the semantic information related to entity (relation) as the real part and the temporal information associated to temporal information as the dual part, respectively. Specifically, the temporal entity describes its initial position and pose in space, including the entity's translation and rotation information, while temporal relations represent transformations that alter entities' spatial states. Thus, we can enhance the model's interpretability. Next, to learn the connections among facts within time ranges, we generate prototype representations for the common feature of each time range and then update the embeddings of the temporal information. To enhance the distinctiveness between different time ranges, we introduce contrastive learning to improve the separation between the features of different time ranges.

The main contributions of this paper are summarized as follows:

- We propose a TKGC model DQD-ETE based on Dual Quaternion space, which considers the independence between semantic information and temporal information.
- We update the temporal information embeddings using prototype representations and contrastive learning, resulting in more distinctive timestamp embeddings.
- We evaluate DQD-ETE on existing TKGC datasets and experimental results indicate that DQD-ETE outperforms other models.

We will make our code available upon acceptance.

II. Related Work

Roughly speaking, we can divide into two categories: SKGC models and TKGC models.

A. SKGC models

Existing SKGC models typically embed entities and relations into low-dimensional vectors and construct corresponding scoring functions. Due to the construction of scoring functions, we can roughly divide SKGC models into translational distance models and semantic matching models.

Translational distance models models model relations as geometric transformations from subject to object. TransE [12] the first and most widely used, represents relations as translations with $s+r=o$ for triples. TransH [13]projects entities onto relation-specific hyperplanes. TransR [14]represents entities and relations into separate spaces and projects entities from entity space to corresponding relation space. RotatE [5] defines each relation as a rotation operation in complex space from the subject to the object, which can address various relation patterns. HAKE [15], which is an extension of RotatE, considers the hierarchical attributes of relations and unifies modeling complex relation patterns and hierarchical semantics.

Semantic matching models measure the authenticity by calculating the similarity of triples. RESCAL [16] represents each relation as a full-rank matrix for scoring via matrix multiplication. DistMult [17] simplifies it with diagonal relation matrices, while ComplEx [18] extends to complex space for asymmetric relations. QuatE [6] advances to quaternion space, treating relations as rotations. DualE [19] further introduces dual quaternion space to unify rotations and translations.

B. TKGC models

TKGC models incorporate temporal information into entities or relations to learn corresponding time-specific representations. Similar to SKGC models, we can roughly divide TKGC models into temporal translational distance models and temporal semantic matching models.

For **temporal translational distance models**, TTransE [20] extends TransE to TKGs, representing the timestamp τ as a translation operation similar to relations. HYTE [21] applies the TransH to TKGs, associating temporal information with entities and relations by generating specific-time hyperplanes for timestamps. Tero [22] which combines TransE and RotatE, defines the temporal information as rotations in complex space from the initial time to the current time for entities and relation as translation. RotateQVS [10] extends the complex embedding space to quaternion space and defines the temporal changes as rotation operations on entities.

For **temporal semantic matching models**, TA-DistMult [17] extends DisMult with recurrent neural networks to incorporate temporal information. TComplEx [23]and TNTComplEx model the knowledge graph via fourth-order tensor decomposition with temporal supplements. Chronor [24]learns relation- and timestamp-parameterized rotations based on RotatE. TeLM [25] adopts linear time regularizers and multivector embeddings for tensor decomposition. TLT-KGE [26] uses timeline-tracking to distinguish cross-timestamp entity/relation representations, and TBDRI [7] proposes relational interaction-based block decomposition with core tensors binding timestamps.

III. PRELIMINARIES

A. Quaternions

A quaternion q is defined as $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, where a, b, c, d are real numbers [27]. $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are square roots of -1, satisfying the Hamilton's rules: $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk}$ = -1. Correspondingly, a n-dimensional quaternion vector $q \in \mathbb{H}^n$ is defined as: $q = q_r + q_i\mathbf{i} + q_j\mathbf{j} + q_k\mathbf{k}$, where $q_r, q_i, q_j, q_k \in \mathbb{R}^n$. Some basic definitions of quaternions are defined as follows: **Quaternion-Inner product:** Given two quaternions $q_1 = a_1 + b_1\mathbf{i} + c_1\mathbf{j} + d_1\mathbf{k}$ and $q_2 = a_2 + b_2\mathbf{i} + c_2\mathbf{j} + d_2\mathbf{k}$, the inner product between q_1 and q_2 is the sum of the product of the corresponding part:

$$q_1 \cdot q_2 = \langle a_1, a_2 \rangle + \langle b_1, b_2 \rangle + \langle c_1, c_2 \rangle + \langle d_1, d_2 \rangle \quad (1)$$

where $\langle \cdot, \cdot \rangle$ represents vector inner product.
Quaternion-Hamilton product: The multiplication product of two quaternions $q_1$and q_2 follows Hamilton's rules, denoted as $q_1 \otimes q_2$.

B. Dual Quaternions

Dual quaternion algebra extends dual-number theory with Hamilton's quaternion algebra [28]. A dual quaternion $Q = a + b\epsilon$, (where a,b are quaternions and $\epsilon^2 = 0$) unifies rotations and translations. Its multiplication is defined as $Q_1 \underline{\otimes} Q_2 = a_1 \otimes a_2 + (a_1 \otimes b_2 + a_2 \otimes b_1)\,\epsilon$.

IV. DQD-ETE

A. Dual Embedding

Suppose that a temporal knowledge graph $\mathcal{G}$ consists of N entities, M relations, and T timestamps. In DQD-ETE, we learn quaternion embedding vectors for entities and relations to represent their semantic information and use the quaternion matrix $Q \in \mathbb{H}^{N \times k}$ to denote all entity embeddings, quaternion matrix $R \in \mathbb{H}^{M \times k}$ to denote all relation embeddings, where k is the dimension of embeddings. To have a more comprehensive and targeted representation of entity and relation at the same timestamp, temporal embeddings are designed for entity and relation, respectively. We set two quaternion matrices $W_T^e \in \mathbb{H}^{T \times k}$ and $W_T^r \in \mathbb{H}^{T \times k}$ to denote temporal information embeddings of all timestamps for entity and relation.

Given a quadruple (s, r, o, τ), the semantic information embedding of subject s, relation r and object o is defined as q_s, q_r, q_o, where $q_s, q_o \in Q, q_r \in R$. While the temporal information embedding for entity and relation respectively under timestamp τ is defined as t_τ^e, t_τ^r, where $t_\tau^e \in W_T^e, t_\tau^r \in W_T^r$. Specifically, they are embedded in the following form:

$$\begin{cases} q_s = \{\ a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} : a_0, a_1, a_2, a_3 \in \mathbb{R}^k\ \} \\ q_o = \{\ b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k} : b_0, b_1, b_2, b_3 \in \mathbb{R}^k\ \} \\ q_r = \{\ c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k} : c_0, c_1, c_2, c_3 \in \mathbb{R}^k\ \} \\ t_\tau^e = \{\ d_0 + d_1\mathbf{i} + d_2\mathbf{j} + d_3\mathbf{k} : d_0, d_1, d_2, d_3 \in \mathbb{R}^k\ \} \\ t_\tau^r = \{\ e_0 + e_1\mathbf{i} + e_2\mathbf{j} + e_3\mathbf{k} : e_0, e_1, e_2, e_3 \in \mathbb{R}^k\ \} \end{cases} \quad (2)$$

In DQD-ETE, our goal is to construct embeddings with temporal semantic representations of temporal entity and temporal relation, while considering the independence of semantic information and temporal information. We find that the dual quaternion unify dual numbers and quaternions, enabling joint representation of rotations and translations for complex spatial transformations. Therefore, we consider the embeddings of temporal entity and temporal relation as dual quaternion vectors, where the semantic embeddings of the entity and the relation serve as the real part of the dual quaternion, and the corresponding temporal information embeddings serve as the dual part of the dual quaternion. Thus, the embeddings of entity s and o and the relation r under the timestamp τ can be represented as follows:

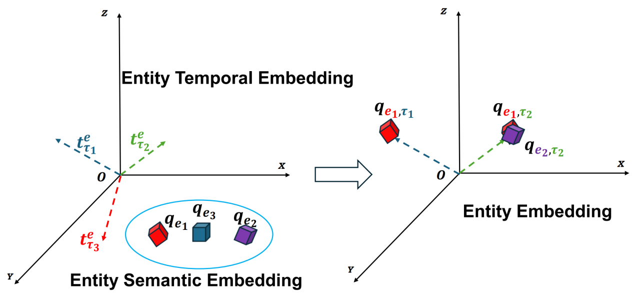


Fig. 2. Illustration of the full representation combining semantic information and temporal information in DQD-ETE .

$$\begin{aligned} q_{s,\tau} &= q_s + t_\tau^e \epsilon \\ q_{o,\tau} &= q_o + t_\tau^e \epsilon \\ q_{r,\tau} &= q_r + t_\tau^r \epsilon \end{aligned} \quad (3)$$

where, $q_{s,\tau}$, $q_{o,\tau}$ and $q_{r,\tau}$ denote the embeddings of subject, object and relation at current timestamp τ.

Fig. 2 illustrates this method, which combines the semantic information part and temporal information part. DQD-ETE models the temporal information as the position of the element in space and the semantic information as the pose embodied in space (rotation angle and orientation in space). This method not only distinguishes the independence of semantic information and temporal information and clarifies the position and pose of temporal entity in space, but also allows them to be embedded in a unified framework.

B. Timestamp-Prototype

In this section, we use the prototype network computation approach to learn common features for each time range, thus increasing the similarity of temporal information embeddings in the same time range. Meanwhile, we use contrastive learning to differentiate the temporal information embeddings of different time ranges.

First, we use mean-based prototypes to portray the distribution of temporal information in the time range, so that these prototypes can approximate the common features of temporal information in that time range. Specifically, We define the time range that the timestamp τ belongs to as I. The time range corresponding to each timestamp is shown below:

$$I_\tau = \left\lfloor \frac{iR}{|T|} \right\rfloor \quad (4)$$

where i represents the index of timestamp τ, R represents the number of time ranges, $|T|$ represents the number of timestamps, and I_τ represents the time range corresponding to timestamp τ.

In each time range, we calculate the time range prototype representation by averaging the embedding of internal temporal information as follows:

$$p_I^e = \frac{1}{K}\sum_{t_\tau^e \in I} t_\tau^e \qquad p_I^r = \frac{1}{K}\sum_{t_\tau^r \in I} t_\tau^r \tag{5}$$

where p_I^e and p_I^r denote prototype representations of temporal information embeddings for entity and relation in the time range, and K is the number of timestamps in the time range.

To capture the common features of temporal information in the time range, we fuse the obtained prototype representation of the time range with the original temporal information embedding to obtain the updated temporal information embedding, computed as follows:

$$t_\tau^{e\,'} = t_\tau^e + p_{I_\tau}^e \qquad t_\tau^{r\,'} = t_\tau^r + p_{I_\tau}^r \tag{6}$$

where $p_{I_\tau}^e$ and $p_{I_\tau}^r$ are prototype representations of temporal information embeddings in entity and relation under the time range corresponding to the timestamp τ, respectively. $t_\tau^{e\,'}$ and $t_\tau^{r\,'}$ are the updated temporal information embeddings.

For timestamps in the same time range, we expect their temporal information embedding to be closer together. And we expect the embedding to be more decentralized for timestamps between different time ranges. We use the contrastive learning framework [29], and the contrastive loss is represented as follows:

$$L = -\frac{1}{B}\sum_{i=1}^{B} \log \frac{\sum_{z_i^+ \in P(i)} e^{\mathrm{sim}\left(\mathbf{z}_i, \mathbf{z}_i^+\right)/\tau}}{\sum_{z_j \in N(i)} e^{\mathrm{sim}\left(\mathbf{z}_i, \mathbf{z}_j\right)/\tau}} \tag{7}$$

where z_i is an instance, B is the number of samples in the batch, z_i^+ represents its positive samples, $P(i)$ is the set of all positive samples in the minibatch, z_j represents its positive samples, and $N(i)$ is the set of all negative instances in the batch. $\mathrm{sim}(\cdot,\cdot)$ represents the similarity function and τ is temperature parameter.

For entity temporal information embedding and relation temporal information embedding, we use the (7) to calculate their contrastive loss respectively, and the total contrastive loss obtained is:

$$L_{con} = \lambda_e L_{ent} + \lambda_r L_{rel} \tag{8}$$

where L_{ent} and L_{rel} represent the contrastive loss of entity temporal information and relation temporal information, respectively, and λ_e and λ_r denote the weights of the two contrastive loss.

C. Training

After updating the embedding vectors in the above two modules, the complete embedding of s, o and r under the timestamp τ that we end up with can be represented as follows:

$$q_{s,\tau}^{'} = q_s + t_\tau^{e\,'}\epsilon \quad q_{o,\tau}^{'} = q_o + t_\tau^{e\,'}\epsilon \quad q_{r,\tau}^{'} = q_r + t_\tau^{r\,'}\epsilon \tag{9}$$

Then, we unitize the temporal relation embedding to obtain the unitized temporal relation embedding $q_{r,\tau}^{'\,\triangleleft}$; Then, we translate and rotate the subject by performing dual quaternion multiplication between the subject $q_{s,\tau}^{'}$ and the temporal relation $q_{r,\tau}^{'\,\triangleleft}$. The calculation process is as follows:

$$q_{s,\tau}^{\diamondsuit} = q_{s,\tau}^{'} \underline{\otimes}\, q_{r,\tau}^{'} \tag{10}$$

Scoring Function: We apply the Dual Quaternion inner product between $q_{s,\tau}^{\diamondsuit}$ and $q_{o,\tau}^{'}$ as the scoring function:

$$d(s,r,o,\tau) = q_{s,\tau}^{\diamondsuit} \cdot q_{o,\tau}^{'} \tag{11}$$

Loss Function: TKGC can be regarded as a multi-classification task; we use full multiclass log-softmax loss function as a quadruple loss. The final training objective is to minimize the following:

$$\begin{aligned}\mathcal{L} = \sum_{(s,r,o,\tau)\in\mathcal{G}} \Bigg(&-log\left(\frac{exp\left(d(s,r,o,\tau)\right)}{\sum_{s'\in\mathcal{E}} exp\left(d(s',r,o,\tau)\right)}\right) \\ &-log\left(\frac{exp\left(d(s,r,o,\tau)\right)}{\sum_{o'\in\mathcal{E}} exp\left(d(s,r,o',\tau)\right)}\right)\Bigg) + L_{con}\end{aligned} \tag{12}$$

where, s' and o' is the object of the negative quadruple (s',r,o,τ) and the subject of the negative quadruple (s,r,o',τ).

V. Experiments and Results

A. Datasets and Setting

To evaluate our proposed model, we perform link prediction task on three commonly used temporal knowledge graph benchmark datasets, namely ICEWS14, ICEWS05-15 [30] and GDELT [31]. We train with Adagrad (lr=0.1), batch size=20000, and max epochs=200. Hyperparameters are optimized via grid search on the validation set.

B. Evaluation Protocol and Baseline

To evaluate our model, we follow the popular evaluation indicators of previous TKGC models (MRR, Hits@1, Hits@3, and Hits@10). MRR is the mean inverse rank of the correct quadruple. Hit@n measures the proportion of correct quadruples among the top n quadruples, where n=1, 3, 10. All the higher metrics (Hits@1, Hits@3, Hits@10 and MRR) indicate better performance.

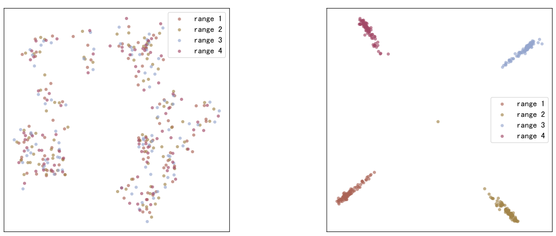
We compare DQD-ETE with the state-of-the-art SKGC and TKGC models. For the SKGC models, we compare our model with TransE [12], DistMult [32], RotatE [5] and QuatE [6]. For the TKGC models, we compare our model with TeRo [22], TComplEx [23], TeLM [25], RotateQVS [10], TLT-KGE [26], SPA [9], HERCULES [8], TBDRI [7], ComTR [11] and SANe [33].

C. Main Results

The experimental results on ICEWS14, ICEWS05-15 and GDELT datasets are reported in Table I. Overall, the proposed DQD-ETE outperforms all existing models, which significantly demonstrates the effectiveness of our model. Since DQD-ETE is based on Dual Quaternion space, this considers the independence of semantic information and temporal

TABLE I
EXPERIMENTAL RESULTS ON THREE EXPERIMENTED DATASETS. THE BEST SCORE IS IN BOLD AND THE SECOND BEST SCORE IS UNDERLINED.

	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
TransE	28.0	9.4	-	63.7	29.4	9.0	-	66.3	11.3	0.0	15.8	31.2
DistMult	43.9	32.3	-	67.2	45.6	33.7	-	69.1	19.6	11.7	20.8	34.8
RotatE	41.8	29.1	47.8	69.0	30.4	16.4	35.5	59.5	-	-	-	-
QuatE	47.1	35.3	53.0	71.2	48.2	37.0	52.9	72.7	-	-	-	-
TeRo	56.2	46.8	62.1	73.2	58.6	46.9	66.8	79.5	24.5	15.4	26.4	42.0
TComplEx	61.0	53.0	66.0	77.0	66.0	59.0	71.0	80.0	29.8	21.3	32.3	46.4
TeLM	62.5	54.5	67.3	77.4	67.8	59.9	72.8	82.3	-	-	-	-
RotateQVS	59.1	50.7	64.2	75.4	63.3	52.9	70.9	81.3	27.0	17.5	29.3	45.8
TLT-KGE	63.4	55.1	68.4	78.6	69.0	60.9	74.1	83.5	35.8	26.5	38.8	54.3
SPA	65.8	54.4	73.7	**85.7**	71.3	58.0	82.0	**93.3**	36.0	28.2	38.4	51.0
HERCULES	69.4	65.0	71.4	77.9	73.5	68.6	76.1	82.9	-	-	-	-
TBDRI	65.2	55.2	69.7	78.5	70.9	64.6	75.7	82.1	26.9	16.4	28.0	44.1
ComTR	71.5	66.4	73.8	80.2	75.2	70.4	78.0	84.0	-	-	-	-
SANe	63.8	55.8	68.8	78.2	68.3	60.5	73.4	82.3	30.1	21.2	32.6	47.6
DQD-ETE	**74.3**	**71.5**	**77.5**	81.9	**81.7**	**78.7**	**83.4**	87.3	**40.3**	**30.7**	**43.6**	**59.2**



(a) Temporal Embedding in TeLM (b) Temporal Embedding in DQD-ETE

Fig. 3. Visualisations of the learned Temporal embeddings on ICEWS14. Different colours indicate different time ranges.

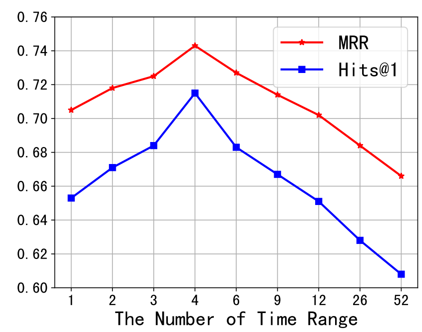


Fig. 4. The change of Hits@1 and MRR with the number of time range R on ICEWS14.

information. The results demonstrate that modelling with consideration for independence effectively improves the learning of embeddings. In addition, because GDELT has a large degree of temporal variability, some facts last for multiple consecutive timestamps. It is even more important to consider the intrinsic connections of temporal information within a time range to be able to make a prediction. The results of DQD-ETE on GDELT proved to be a better solution to this problem. Moreover, the method achieves better results compared to baselines on all three of these datasets with varying levels of sparsity, indicating that the proposed model can overcome the problem of data sparsity to a certain extent.

D. Analysis

Capacity of Distinguishing temporal information embedding: In this section, we discuss the ability of DQD-ETE to distinguish temporal information embeddings. we utilize PCA to visualize the trained temporal information embeddings of TeLM and DQD-ETE. The visualization results are shown in Fig. 3. We observe that the distribution of temporal embeddings in the time range in TeLM is scattered. It cannot learn the intrinsic connection of temporal information within a time range. On the contrary, DQD-ETE can learn temporal information embedding effectively, resulting in time range clusters. This demonstrates the effectiveness of DQD-ETE in learning the intrinsic connection of temporal information.

Parameter Study: We test the the impact of the number of time ranges R on ICEWS14 (Fig. 4). Performance peaks at $R = 4$ and declines afterward, because too few ranges introduce low-correlation noise, while an excessive number of ranges leads to sparse temporal information within each range, limiting relevant feature extraction.

VI. CONCLUSION

In this paper, we propose a novel TKGC model DQD-ETE, which embeds temporal entity (relation) in dual quaternion, where the real part is the semantic information and the dual part is the temporal information. Furthermore, to learn the connections among facts within time ranges, we generate prototype representations for the common feature of each time range and then update the feature embeddings of the temporal information. Then, we evaluate DQD-ETE and find that DQD-ETE significantly improves on the three benchmark datasets compared to the previous model and better captures the correlation of temporal information within the same time range.

References

[1] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, "Dbpedia: A nucleus for a web of open data," in *international semantic web conference*. Springer, 2007, pp. 722–735.

[2] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor, "Freebase: a collaboratively created graph database for structuring human knowledge," in *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, 2008, pp. 1247–1250.

[3] K. Tu, P. Cui, D. Wang, Z. Zhang, J. Zhou, Y. Qi, and W. Zhu, "Conditional graph attention networks for distilling and refining knowledge graphs in recommendation," in *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 2021, pp. 1834–1843.

[4] Z. Li, H. Liu, Z. Zhang, T. Liu, and N. N. Xiong, "Learning knowledge graph embedding with heterogeneous relation attention networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, pp. 3961–3973, 2021.

[5] Z. Sun, Z. Deng, J. Nie, and J. Tang, "Rotate: Knowledge graph embedding by relational rotation in complex space," in *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. [Online]. Available: https://openreview.net/forum?id=HkgEQnRqYQ

[6] S. Zhang, Y. Tay, L. Yao, and Q. Liu, "Quaternion knowledge graph embeddings," *Advances in neural information processing systems*, vol. 32, 2019.

[7] M. Yu, J. Guo, J. Yu, T. Xu, M. Zhao, H. Liu, X. Li, and R. Yu, "Tbdri: block decomposition based on relational interaction for temporal knowledge graph completion," *Applied Intelligence*, vol. 53, no. 5, pp. 5072–5084, 2023.

[8] S. Montella, L. M. Rojas-Barahona, and J. Heinecke, "Hyperbolic temporal knowledge graph embeddings with relational and time curvatures," in *Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021, Online Event, August 1-6, 2021*, ser. Findings of ACL, C. Zong, F. Xia, W. Li, and R. N. and, Eds., vol. ACL/IJCNLP 2021. Association for Computational Linguistics, 2021, pp. 3296–3308. [Online]. Available: https://doi.org/10.18653/v1/2021.findings-acl.292

[9] Z. Wang, H. Du, Q. Yao, and X. Li, "Search to pass messages for temporal knowledge graph completion," in *Findings of the Association for Computational Linguistics: EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, Y. Goldberg, Z. Kozareva, and Y. Zhang, Eds. Association for Computational Linguistics, 2022, pp. 6160–6172. [Online]. Available: https://doi.org/10.18653/v1/2022.findings-emnlp.458

[10] K. Chen, Y. Wang, Y. Li, and A. Li, "Rotateqvs: representing temporal information as rotations in quaternion vector space for temporal knowledge graph completion," *arXiv preprint arXiv:2203.07993*, 2022.

[11] 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 (IJCNN)*. IEEE, 2023, pp. 01–08.

[12] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," *Advances in neural information processing systems*, vol. 26, 2013.

[13] Z. Wang, J. Zhang, J. Feng, and Z. Chen, "Knowledge graph embedding by translating on hyperplanes," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 28, no. 1, 2014.

[14] Y. Lin, Z. Liu, M. Sun, Y. Liu, and X. Zhu, "Learning entity and relation embeddings for knowledge graph completion," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 29, no. 1, 2015.

[15] 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*, vol. 34, no. 03, 2020, pp. 3065–3072.

[16] M. Nickel, V. Tresp, H.-P. Kriegel *et al.*, "A three-way model for collective learning on multi-relational data." in *Icml*, vol. 11, no. 10.5555, 2011, pp. 3 104 482–3 104 584.

[17] 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*, E. Riloff, D. Chiang, J. Hockenmaier, and J. Tsujii, Eds. Brussels, Belgium: Association for Computational Linguistics, Oct.-Nov. 2018, pp. 4816–4821. [Online]. Available: https://aclanthology.org/D18-1516

[18] T. Trouillon, J. Welbl, S. Riedel, É. Gaussier, and G. Bouchard, "Complex embeddings for simple link prediction," in *International conference on machine learning*. PMLR, 2016, pp. 2071–2080.

[19] 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*, vol. 35, no. 8, 2021, pp. 6894–6902.

[20] J. Leblay and M. W. Chekol, "Deriving validity time in knowledge graph," in *Companion Proceedings of the The Web Conference 2018*, 2018, pp. 1771–1776.

[21] S. S. Dasgupta, S. N. Ray, and P. Talukdar, "Hyte: Hyperplane-based temporally aware knowledge graph embedding," in *Proceedings of the 2018 conference on empirical methods in natural language processing*, 2018, pp. 2001–2011.

[22] C. Xu, M. Nayyeri, F. Alkhoury, H. S. Yazdi, and J. Lehmann, "Tero: A time-aware knowledge graph embedding via temporal rotation," in *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*, D. Scott, N. Bel, and C. Zong, Eds. International Committee on Computational Linguistics, 2020, pp. 1583–1593. [Online]. Available: https://doi.org/10.18653/v1/2020.coling-main.139

[23] T. Lacroix, G. Obozinski, and N. Usunier, "Tensor decompositions for temporal knowledge base completion," in *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net, 2020. [Online]. Available: https://openreview.net/forum?id=rke2P1BFwS

[24] A. Sadeghian, M. Armandpour, A. Colas, and D. Z. Wang, "Chronor: Rotation based temporal knowledge graph embedding," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 7, 2021, pp. 6471–6479.

[25] 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*, 2021, pp. 2569–2578.

[26] 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*, 2022, pp. 2529–2538.

[27] W. R. Hamilton, "Lxxviii. on quaternions; or on a new system of imaginaries in algebra: To the editors of the philosophical magazine and journal," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 25, no. 169, pp. 489–495, 1844.

[28] Clifford, "Preliminary sketch of biquaternions," *Proceedings of the London Mathematical Society*, vol. 1, no. 1, pp. 381–395, 1871.

[29] Z. Luo, W. Xu, W. Liu, J. Bian, J. Yin, and T.-Y. Liu, "KGE-CL: Contrastive learning of tensor decomposition based knowledge graph embeddings," 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. Gyeongju, Republic of Korea: International Committee on Computational Linguistics, Oct. 2022, pp. 2598–2607. [Online]. Available: https://aclanthology.org/2022.coling-1.229

[30] A. García-Durán, S. Dumančić, and M. Niepert, "Learning sequence encoders for temporal knowledge graph completion," *arXiv preprint arXiv:1809.03202*, 2018.

[31] R. Trivedi, H. Dai, Y. Wang, and L. Song, "Know-evolve: Deep temporal reasoning for dynamic knowledge graphs," in *international conference on machine learning*. PMLR, 2017, pp. 3462–3471.

[32] B. Yang, W.-t. Yih, X. He, J. Gao, and L. Deng, "Embedding entities and relations for learning and inference in knowledge bases," *arXiv preprint arXiv:1412.6575*, 2014.

[33] Y. Li, X. Zhang, B. Zhang, F. Huang, X. Chen, M. Lu, and S. Ma, "Sane: Space adaptation network for temporal knowledge graph completion," *Information Sciences*, vol. 667, p. 120430, 2024.