

RESEARCH ARTICLE



OPEN ACCESS

Received: 02.12.2020

Accepted: 25.02.2021

Published: 09.04.2021

Citation: Sangeetha Devi A, Kalaivani A (2021) Knowledge graph embedding via the properties of mapping and dynamic matrix. Indian Journal of Science and Technology 14(11): 942-955. <https://doi.org/10.17485/IJST/v14i11.2172>

* **Corresponding author.**

sangeethadeviphd123@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2021 Sangeetha Devi & Kalaivani. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment ([iSee](#))

ISSN

Print: 0974-6846

Electronic: 0974-5645

Knowledge graph embedding via the properties of mapping and dynamic matrix

A Sangeetha Devi^{1*}, A Kalaivani²

¹ Associate Professor, Department of Science and Humanities (Mathematics), P. A. College of Engineering and Technology, Pollachi, Tamilnadu, India

² Assistant Professor, Department of Computer Technology, Nallamuthu Gounder Mahalingam College, Pollachi, Tamilnadu, India

Abstract

Objective: The main objective of this research is constructing a knowledge graph using a superior model that leverage the knowledge through relational mapping properties and the matrix constructed dynamically based on the changes in the relational mapping. **Methods:** This research work proposes symbolic objects in the knowledge graph that are signified by two vectors, the first vector denotes the entity/relation and the vector denote the matrix which is constructed with the help of mapping properties. The proposed methodology uses only minimum parameters and avoids multiplication processes which make the proposed scheme effective in the large scale graphs. The matrix representation with the assistance of mapping properties reduces the computational complication and the time requires to attain the knowledge is also minimized. **Findings:** The empirical research on knowledge graph mining gives significant insights and it is applied on benchmark datasets namely FB15K-237 and WN18RR. The performance are evaluated in terms of precision and recall in percentage.

Keywords: Knowledge embedding; Entity; Dynamic matrix; Relation and Edge

1 Introduction

Knowledge Graph (KG) is a kind of data structure and it is developed by incorporating the structural properties of graph. The semantic network of any relation is denoted by “node-edge-node”, which represent an “entity” or “concept”. The relationship among any entity is represented by the edges in the graph. The real world concepts are described via rich relations in the graph and it is extensively used in the fields of semantic search⁽¹⁾, medical⁽²⁾ and finance⁽³⁾. The insufficiency of information in the KG leads to the incompleteness of data that makes the data processing as incomplete.

The meagerness of the information makes the data processing incomplete and the process of completing the missing value is a tedious process. The completion of missing values and generating true relation in KG is a dynamic research area. The KG approaches available in the graph mining is established mainly by embedding of entity in graph and it ignores the impact on the diverse relations on the signification of the triple. The

knowledge graph is widely used in the applications of knowledge management, financial industry, data governance, automatic fraud detection and in insider trading.

The existing methods use the strategy of equal allocation of weights to various relations on the similar path of the graph⁽⁴⁾. Hence, the significance of the relation is treated correspondingly and it results in inaccuracy in prediction of links. The event prediction is significant and it is accomplished through event KG whereas the accuracy is important in event prediction⁽⁵⁾. If the critical path identification results with inaccurate value then the relation reasoning and the prediction process goes wrong, which gives entirely differing information. The process of decision analysis will mislead by the acquired inaccurate data. Therefore, enrichment of accuracy and reduction of computational complexity is necessary.

The similar studies on knowledge graph mining and two major graph mining category are summarized and two diverse aspects namely translational and compositional models are briefly explained here.

Translational Model

Translational models rely on the tail entity that is a result of a head entity through knowledge graph attained effectively. The formative work TransE scheme builds golden triplets namely HD, R and TL by considering the r relation and it is a translation from the head HD to tail TL⁽⁶⁾. Accordingly, head HD plus relation R is nearer to the tail TL and the relevant score for the function is estimated as:

$$f_s(HD, TL) = \|HD + R - TL\|_{l_1/l_2}^2 \quad (1)$$

The seminal TransE model is incorporated to the simple 1 to 1 relation and it faces the complexity for facing the relations such as N to N, 1 to N and N to 1. To rectify the complicated relation issue, a new TransH is developed that permit every entity with unique representation for every relation⁽⁷⁾. TransH architects the relation as a vector R on a hyperplane and the entity vectors such as HD and TL, where the entities are projected into a relation-specific hyperplanes (HD_{\perp} and TL_{\perp}). The score function is equated as:

$$f_s(HD, TL) = \|HD_{\perp} + R - TL_{\perp}\|_{l_1/l_2}^2 \quad (2)$$

TransDR⁽⁸⁾ is established with improvements and it incorporates the weight, which elucidate the eigenstate, relation specific space and mimesis. The entities hold a weight w_r for every relation that is used for the knowledge graph establishment model⁽⁹⁾. For every triplet in the graph, it contains specific triplet margin in the adaptive model⁽¹⁰⁾. In the dynamic model, relation vector is added with head entity and it is signified as orthogonal entity vector $(HD+R)^T TL \approx 0$ ⁽¹¹⁾. The representation ability is improved by the entity description in the knowledge graph⁽¹²⁾.

Compositional Model

In addition to the translation models, there are numerous knowledge graph approaches are available that are knowledge embedding models and it is reasonable for denoting the graphs as well as retrieving the knowledge.

Structured Embedding

Structure Embedding (SE) model⁽¹³⁾ considers the correspondence of head entity HD and tail entity TL. During the existence of triplet values (HD, TL, R), overlap in a definite relation space RS^n occur. It utilizes two mapping matrix namely $M_R HD$ and $M_R TL$, that extracts the needed features from the relevant entities of HD and TL. The scoring function of SE is equated by,

$$f_s(HD, TL) = \|M_{R-HD} HD + M_{r-TL} TL\|_{l_1/l_2}^2 \quad (3)$$

Unstructured Model

Unstructured Model (UM)⁽¹⁴⁾ deems the head entity HD and tail entity TL in every triplet holds similar semantic and the correlation amongst them is ignored. The scoring function utilizes the l_2 norm to restrain the embeddings, that is,

$$f_s(HD, TL) = \|HD + TL\|_{l_2}^2 \quad (4)$$

The diversification among the relations elucidates the features of minimum effectiveness in the knowledge graph.

Semantic Matching Energy

The main intent of semantic matching energy (SME)⁽¹⁵⁾ is to attain the accurate and correct triplet value where the feature vectors are orthogonal (head and tail).

After the process of mapping, every entity will be added with the bias of the relevant relation o15p. The feature extraction process is accomplished by non-linear and linear method. The scoring function is equated by,

$$f_s(HD, TL) = (M_H HD + M_{HD_R} R + B_{HD})^{TL} \cdot (M_{TL} TL + M_{TL_R} R + B_{TL}) \quad (5)$$

$$f_s(HD, TL) = (M_H HD \otimes M_{HD_R} R + B_{HD})^{TL} \cdot (M_{TL} TL \otimes M_{TL_R} R + B_{TL}) \quad (6)$$

For the non-linear case of feature extractors, the feature extractors are M_{HD} , M_{HD_R} , M_{TL} , M_{TL_R} and the hadamard product is \otimes and BR and BTL are consider as the biases for the values HD and TL. Various approaches have been developed that enhances the performance of SME.

Single Layer Model

In comparison, with the model SE, the single layer model (SLM) is effective in translated feature extraction and it incorporates the non-linear activation function. After the activation process of feature vector, they are orthogonal with the relation with the obtained feature vector and the bias values added to its relation⁽¹⁶⁾.

Neural Tensor Network

It is a more complex model and the tensor value is connected with the network that acts as an improved feature extractor.

Complex Embedding

The function of logistics inverse link is predicted using the complex embedding model (CompIEX)⁽¹⁷⁾. The function of logistic inverse link is,

$$P(K_{Rso} = 1) = \sigma(\varnothing(RL, s, o; \delta)) \quad (7)$$

Where the value of Y_{Rso} belongs to the value $\{-1, 1\}$ and it is the identified fact, RL is the relation, s and o denotes the subject value and object value respectively. The relevant parameters for relevant model is represented by δ . Generally, the scoring function σ is obtained from the factorization of relations observed in the graph.

The DisMult⁽¹⁸⁾ model utilizes bilinear scheme that denote the relations and entities. The logical rules in the knowledge graph are acquired by exploiting the learned embedding scheme. The HOLE⁽¹⁹⁾ model is formulated vector space with compositional model, which is based on the vector relation in the form of circular correlation.

Latent Factor Model

The latent factor model (LFM)^(20,21) adopts the structures of head entity is orthogonal to the tail entity, whereas the head is mapped in a definite relation space. The scoring function is determined by,

$$f_s(HD, TL) = HD^{TL} M_R TL \quad (8)$$

MvTransE and MSGE

The sub spaces are embedded with the representation learning model⁽⁶⁾ and word vectors are investigated from the unlabeled raw text. It is applicable to the scenario of zero-shot approach. The multiview translation learning (MvTransE)⁽⁷⁾ learns the fact that is relation from the perspective of local and global. The relation in the knowledge is learned from the correlation and it depicts the low rank structure over the embedded relation matrix⁽¹⁴⁾. A multi-step gated model (MSGE)⁽¹⁰⁾ is initiated for predicting the relation from the knowledge graph. Knowledge graph completion is a graph construction approach that is attained from the relation among the knowledge graph model. Knowledge graph approaches are widely used in the drug-drug interaction prediction⁽¹¹⁾ and the approach is designed to generate an effective prediction model. The knowledge graph

model is constructed to attain the effective relation prediction and every model has faced shortcomings which are rectified by introducing dynamic matrix with mapping properties.

In this paper, a novel scheme is introduced and it incorporates the dynamic matrix with mapping properties. The relation and entity is expressed by vectors where the denotation of a relation or an entity is by one vector and the embedding scheme in the entity is denoted by the other projection vector. The process of embedding entity into a vector space in the relation is will create the mapping matrices and each entity relation in the graph has distinctive mapping matrix. The matrix by vector and multiplication operations is replaced by simple vector operations. The proposed dynamic matrix is evaluated with the relation prediction and triplet classification. The proposed model is compared with previous KG models.

2 Proposed Model

The triplet values are signified as HD_i , R_i and TL_i , where $i=1,2,3,\dots,n_{TL}$. The HD_i signifies the head, R_i represents the relation and TL_i represents the tail. Their relevant embeddings are denoted by HD_i , R_i and TL_i , $i=1,2,3,\dots,n_{TL}$. The golden triplets are signified by \triangle and the negative triplets are signified by \triangle' . The relation set and entity set is represented by R_S and E_S , respectively. The identity matrix with the size $m \times n$ is signified by $IM^m \times n$.

3.1 Entities and Relations with Multiple Type

The consideration of diverse relations, CTransR partitions the triplets to a definite relation R into various groups and representation of vector is learned for every group. Conversely, every entity contains diversified types. In the knowledge graph approaches namely TransR/CTransR and TransH, all the varieties of entities share similar matrices and mapping vectors. Diversified nature of entities has diverse attribute and function. The process of sharing the similar parameters for a relation is insufficient and not applicable. The mapping properties must be similar for similar entities and dissimilar for other varieties. Figure 1 reveals numerous varieties of head and tail entities of relations location that are moderately acquired in the FB15k.

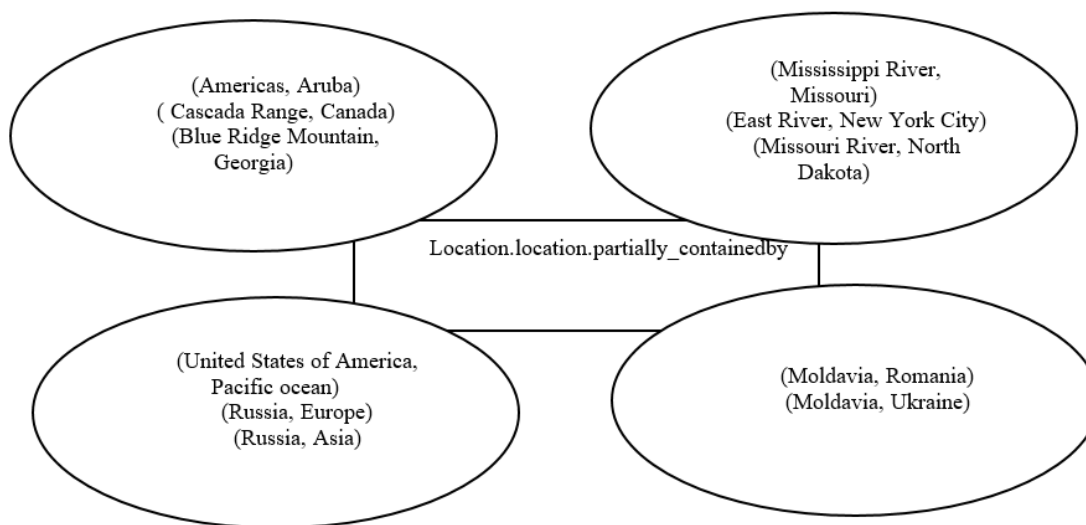


Fig 1. Various types of relation and entities

The process of mapping in a transaction among the relations and entities has numerous varieties. Hence, a well-organized scheme dynamic matrix with mapping properties is developed that considers diverse types of relation and knowledge to code the knowledge graph into the embedding vectors through matrices that utilizes mapping properties, which are generated by the projection vectors⁽²²⁾.

3.2 Dynamic Matrix and Mapping Property

In the dynamic matrix, every named symbol objects namely relations and entity are signified by vectors. Initially, the meaning of the entity is captured and subsequent process is carried to build the mapping matrix with relevant properties. For example, for the given triplet (HD, R, TL) and their vectors are $HD, HD_p, R, R_p, TL, TL_p$, where the projection vector is represented

by the subscript value p , where $HD, HD_p, TL, TL_p \in R^n$ and $R, R_p \in R^m$. The two matrixes with mapping properties are $M_{R_HD}, M_{R_TL} \in R^{n \times n}$ to venture the space of entity from the space of relation. They are equated as:

$$M_{R_HD} = R_p HD_p^\perp + IM^{m \times n} \quad (9)$$

$$M_{R_TL} = R_p TL_p^\perp + IM^{m \times n} \quad (10)$$

Therefore, relations and entities are determined by the mapping matrices and it offers sufficient interaction. Each mapping matrix is initialized with an identity matrix and it is added with the $IM^{m \times n}$ to M_{RHD} and M_{RTL} . With the mapping matrices, the projected vectors are defined as follows:

$$HD_\perp = M_{R_HD} HD, TL_\perp = M_{R_TL} TL \quad (11)$$

Every training triplet is assigned with the weight and the weight is estimated by,

$$w_R = \frac{1}{\log(HD_R p TL_R + TL_R p HD_R)} \quad (12)$$

The degree of mapping is simply estimated by the relationship to the count of average number of tail entities per every distinctive head entity and vice versa. The value $HD_R p TL_R$ represents the heads per tail and $TL_R p HD_R$ represents the tails per head in the relation R .

The estimated weight denotes the degree of mapping and the mapping property of a triplet relies on the relationship. The relation of husband to wife is considered as ONE-TO-ONE and the parents to children are MANY-TO-MANY. Hence, the value of weight is relation specific and Then the scoring function is equated as,

$$f_S(HD, TL) = -w_R \|HD_\perp + R - TL_\perp\|_2^2 \quad (13)$$

The investigation constraints used in this experiment is $\|HD\|_2 \leq 1, \|TL\|_2 \leq 1, \|R\|_2 \leq 1, \|HD_\perp\|_2 \leq 1$ and $\|TL_\perp\|_2 \leq 1$.

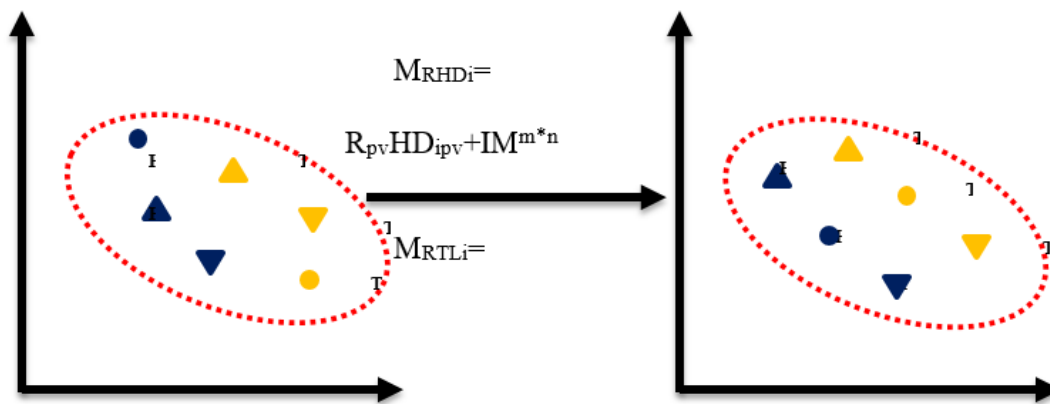


Fig 2. Simple notion of dynamic matrix representation

In Figure 2, every shape denotes every individual entity pair relies on the triplet that holds relation R . The mapping matrices are signified by M_{RHD} and M_{RTL} , where the head is HD and tail is TL in the matrix. The notion of projection vector and vector entities are R_{pv} and HD_{ip}, TL_{ip} . The vector projection of every entity is represented by $HD_{i\perp}$ and $TL_{i\perp}$ ($i=1,2,\dots,n$). The vector projected in the graph must fulfill the condition $HD_{i\perp} + R TL_{i\perp}$.

The dynamic matrix with mapping properties builds two matrices for every triplet and assigns a projection vector for every relation and entity. In addition, matrix multiplication is not applied on the dynamic matrix, which is replaced by the vector operation. Without dropping the generality, assume $m \geq n$, the expected vector can be estimated as,

$$HD_\perp = M_{R_HD} HD = HD_{pv}^\perp HDR_{pv} + [HD^\perp]^\perp \quad (14)$$

$$TL_{\perp} = M_{R_{TL}} TL = TL_{pv}^{\perp} TLR_{pv} = R_{pv} + [TL^{\perp}]^{\perp} \quad (15)$$

where HD is the head, M is the mapping matrix, R is the relation. TL is the tail, and pv represent the vector.

Therefore, the proposed method has minimum calculation and it makes training process as faster and it is applied to the large scale knowledge graphs. When compared with the existing approach, the graph construction and knowledge retrieval process in the proposed approach is simple in terms of computation.

3 Result and Discussion

To investigate the performance of proposed approach two benchmark data sets has been used, which are FB15k-237⁽¹⁹⁾ and WN18RR⁽¹⁹⁾. The identification information in the knowledge graph composed of identification of relation and classification of triplets. The performance of the proposed model is testified by the relation identification and also it can handle the inferences in the relation. The optimal results obtained by general inverse rule model and the effective prediction of link is attained when the proposed model is applied on the datasets FB15k-237 and WN18RR. The reverse relation in the dataset is achieved from the relevant sub-datasets of FB15k-237 and WN18RR²⁰. The information about the dataset is explained in the Table 1.

Table 1. Dataset description

Dataset	#relation	#entity
FB15K237	237	14541
WN18RR	11	40943

4.1 Prediction of Relation

The link prediction is used to calculate the missing values of golden triplets namely tail, head and relation. Every triplet in the dataset is interchanged with the entities in the dictionary values whereas tail or head is removed. The score value is estimated from the corrupted rank and triplet values which are sorted by the descending order. The entity with accurate entity is stored and the rank is emphasized as best entity. The proposed approach is investigated by two metrics namely average of every correct entity that is mean rank and the proportion of every correct entity ranked as Hits@10 that lies in the top 10. The best embedding model is accomplished by lower mean rank with higher Hits@10. The estimation setting of the proposed approach is Filter and Raw. In the filter, the corrupted triplets are removed by valid test before ranking. In Raw, corrupted triplets may available in graph that is also considered as correct triplet.

In the graph, the similarity among the triplets are estimated that is score value and the estimation of scoring function is given in section 3.2. The experiment is carried for two measures namely the average rank of all the correct entities that is Mean Rank and the proportion of entities in the top N that is Hits@1, Hits@3 and Hits@10. The evaluation setting Raw and Filter is used where the corrupted triplet also available in knowledge graphs as Raw and the corrupted are removed in the Filter data.

Table 2. Experimental results on relation prediction using the dataset FB15k-237

Dataset	FB15k-237							
Metric	Mean Rank		Hits@N					
			Hits@1		Hits@3		Hits@10	
	Filter	Raw	Filter	Raw	Filter	Raw	Filter	Raw
RESCAL	828	683	33.6	48.8	34.9	50.8	35.6	38.2
TransE	243	125	31.4	43.1	31.8	44.6	36.9	33.1
TransH	211	84	37.6	47.7	39.8	51.5	38.4	41.7
TransR	226	78	39.9	59.4	41.5	61.9	41.8	46.8
CtransR	233	82	41.6	59.1	43.4	64.1	43.0	51.2
Dy_Mat	199	75	49.8	64.7	51.8	69.1	47.1	54.4

The experimental outcomes on the dataset FB15k-237 is given in the Table 2. The proposed Dy_Mat is compared with the existing algorithms namely RESCAL, TransE, TransH, TransR and CTransR. For comparison multiple types of relations and entities are used from the available models. From the observation of results, it is concluded that the dynamic matrix with mapping properties outperforms the existing knowledge graph models, exclusively on the sparse dataset.

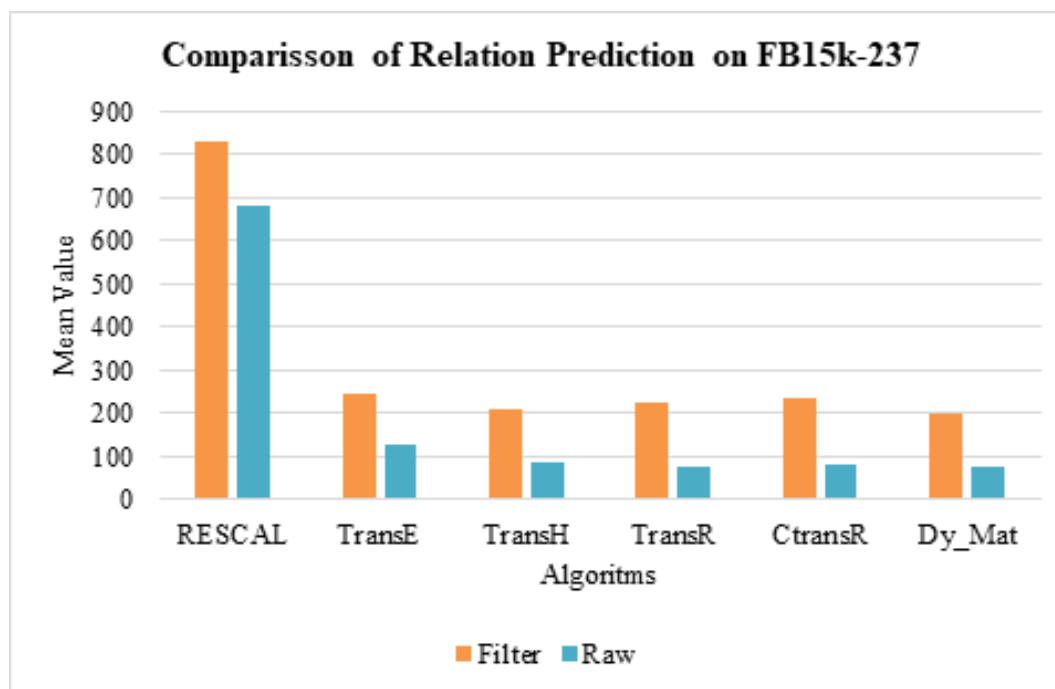


Fig 3. Comparison of relation prediction in FB15k-237 for the metric mean

In Figure 3, the Mean rank for Raw and Filtered relations in the knowledge model is depicted. The proposed approach attained minimum mean value, which shows the efficiency of the proposed Dy_Mat approach.

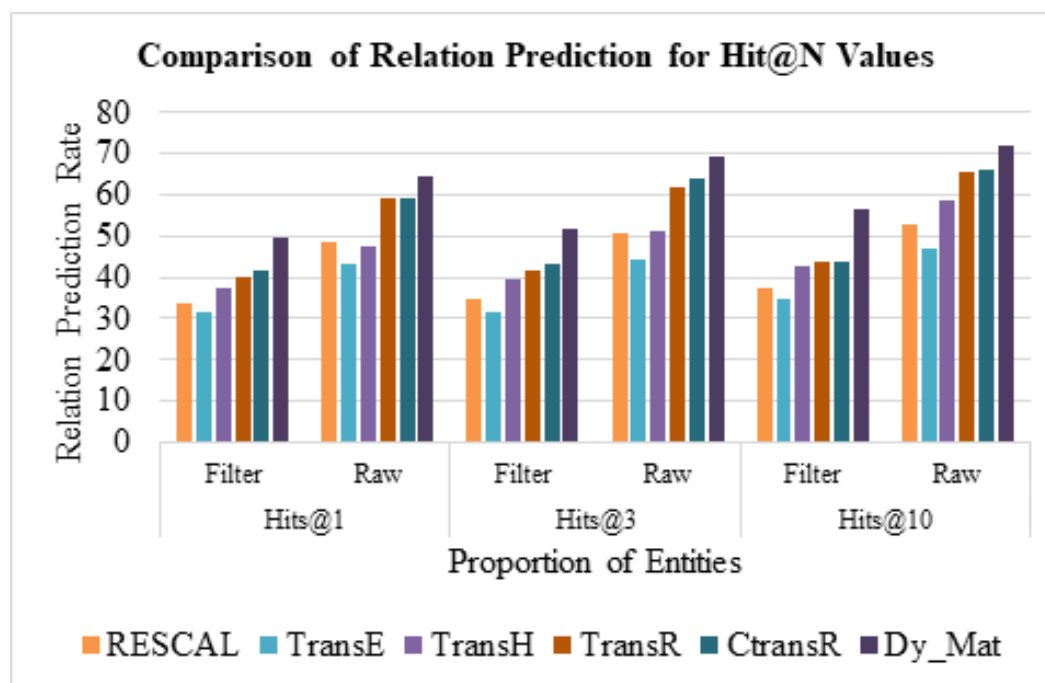


Fig 4. Comparison of relation prediction in FB15k-237 for Hit@N

Figure 4, depicts the illustration of relations among the graph at various Hit@N values in the knowledge model. The diverse relations used to compare the results that are 1-to-1, N-to-1, 1-to-N, and N-to-N. For the N-to-N relation in knowledge graph, the dynamic mapping matrix is enriched almost by 10.1% and the accuracy is enhanced in every relation.

Table 3. Experimental results on relation prediction using the dataset MN18RR

Dataset	MN18RR							
Metric	Mean Rank		Hits@N					
			Hits@1		Hits@3		Hits@10	
	Filter	Raw	Filter	Raw	Filter	Raw	Filter	Raw
RESCAL	1,163	1,180	33.6	48.8	34.9	50.8	37.2	52.8
TransE	251	263	31.4	43.1	31.8	44.6	34.9	47.1
TransH	303	318	37.6	47.7	39.8	51.5	42.5	58.5
TransR	219	232	39.9	59.4	41.5	61.9	43.8	65.5
CtransR	230	243	41.6	59.1	43.4	64.1	44.0	66.3
Dy_Mat	209	221	49.8	64.7	51.8	69.1	56.8	71.9

The experimental outcomes on the dataset MN18RR is given in the Table 3. The proposed Dy_Mat is compared with the existing algorithms namely RESCAL, TransE, TransH, TransR and CTransR. For comparison multiple types of relations and entities are used from the available models. The proposed approach achieved better result and it lever the intricate relation of relation and entity in the knowledge graph.

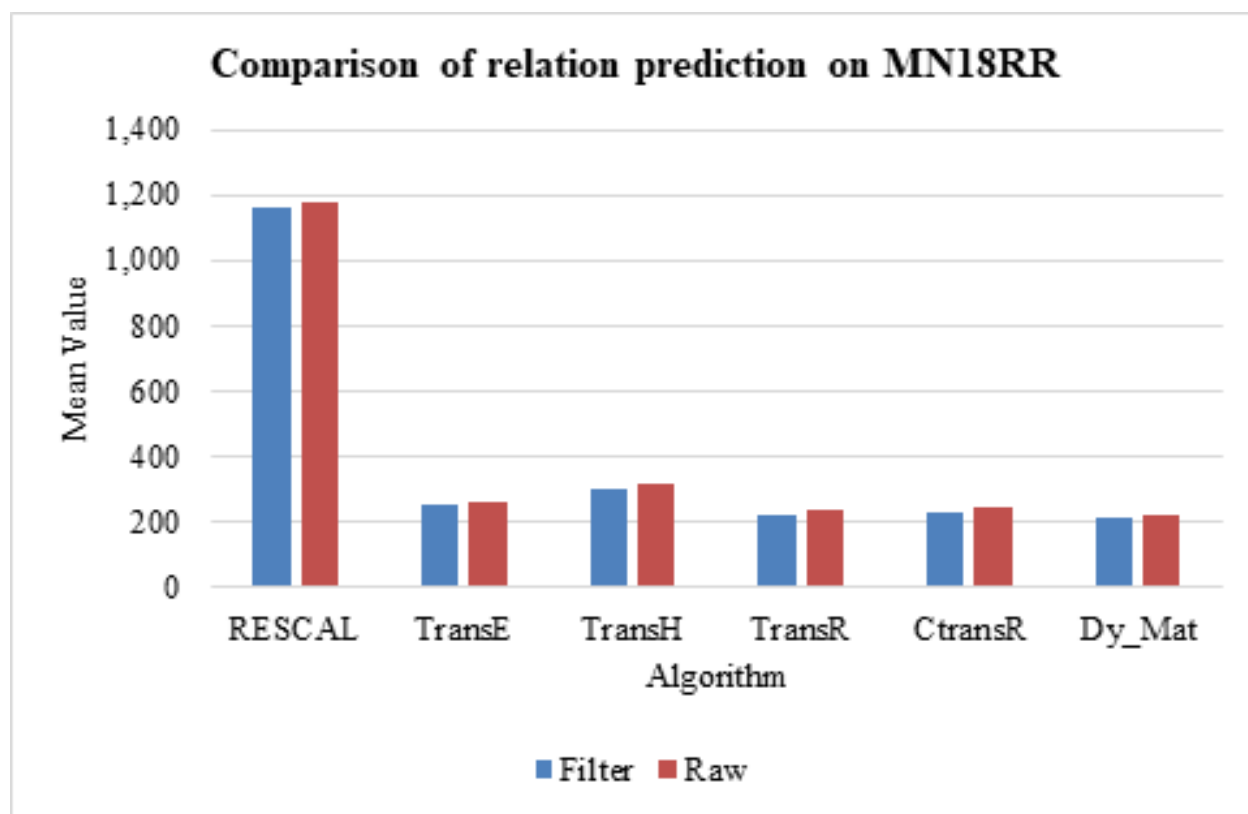


Fig 5. Comparison of relation prediction in MN18RR for the metric mean

In Figure 5, the Mean rank value for Raw and Filtered relations in the knowledge model is depicted. The proposed approach attained minimum mean value, which shows the efficiency of the proposed Dy_Mat approach.

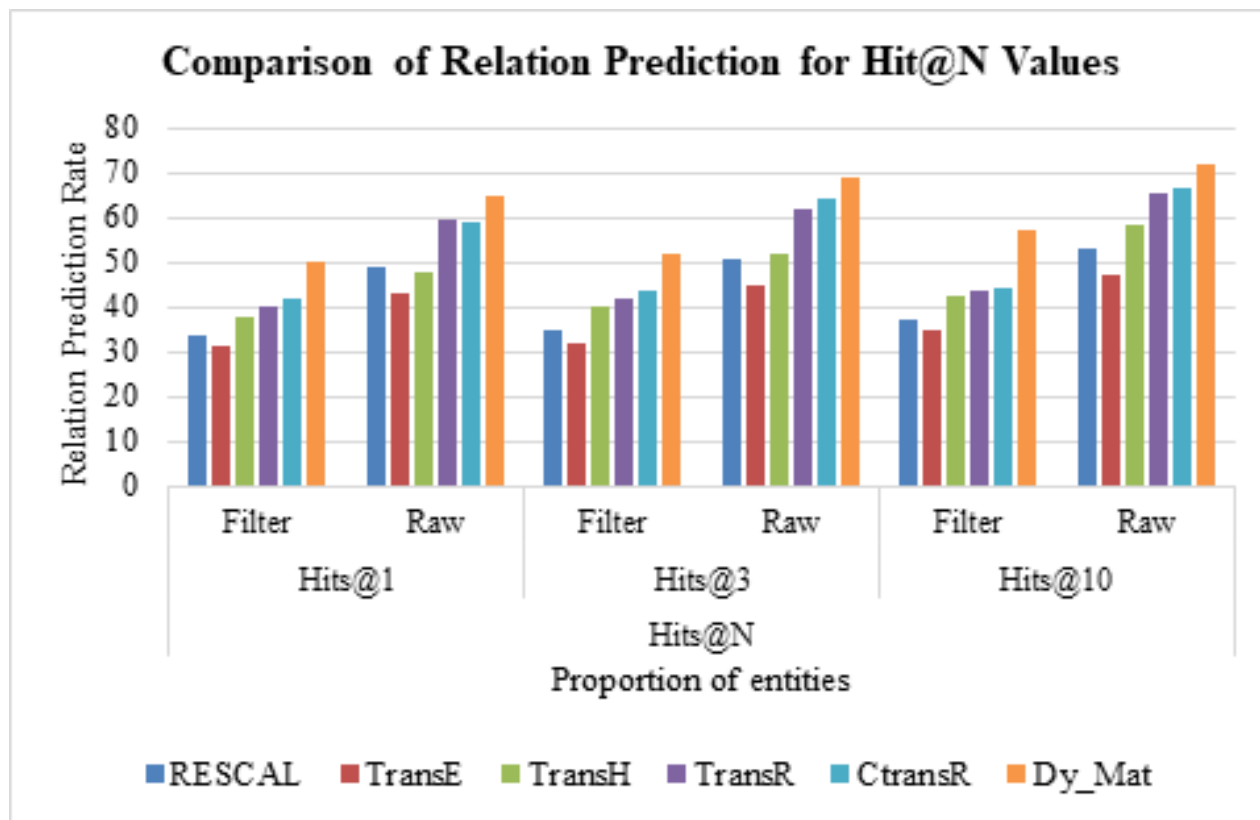


Fig 6. Comparison of relation prediction in MN18RR for Hit@N

Figure 6, depicts the illustration of relations among the graph at various Hit@N values in the knowledge model. The diverse relations used to compare the results that are 1-to-1, N-to-1, 1-to-N, and N-to-N. For the N-to-N relation in knowledge graph, the dynamic mapping matrix is enriched almost by 12.6% and the accuracy is enhanced in every relation. The diverse nature of graph and entities in the knowledge graph is a significant factor and the proposed dynamic matrix with mapping properties is appropriate for formulating the knowledge graphs.

Table 4. Experimental results on FB15k-237 by using the properties of mapping relations in percentage

Tasks	Prediction Tail (Hits@10)				Prediction Head (Hits@10)			
Relation Category	1-to-1	1-to-N	N-to-1	N-to-N	1-to-1	1-to-N	N-to-1	N-to-N
RESCAL	43.3	19.6	65.2	49.2	42.3	59.3	17.3	46.1
TransE	43.7	19.7	66.7	50.0	43.7	65.7	18.2	47.2
TransH	63.7	30.1	83.2	60.8	66.8	87.6	28.7	64.5
TransR	76.2	38.4	76.2	69.1	76.9	77.9	38.1	66.9
CtransR	77.4	37.8	37.8	70.3	78.6	77.8	36.4	68.0
Dy_Mat	78.6	36.5	42.1	72.7	79.2	78.8	37.1	69.3

The detailed outcomes of mapping properties and their relation on the dataset FB15k-237 are shown in the Table 4. This table shows the tail and head levels of the knowledge graph and it consists of diverse relations namely 1-to-1, 1-to-N, N-to-1 and N-to-N. The comparison is carried for existing RESCAL, TransE, TransH, TransR, CtransR and proposed Dy_Mat approach.

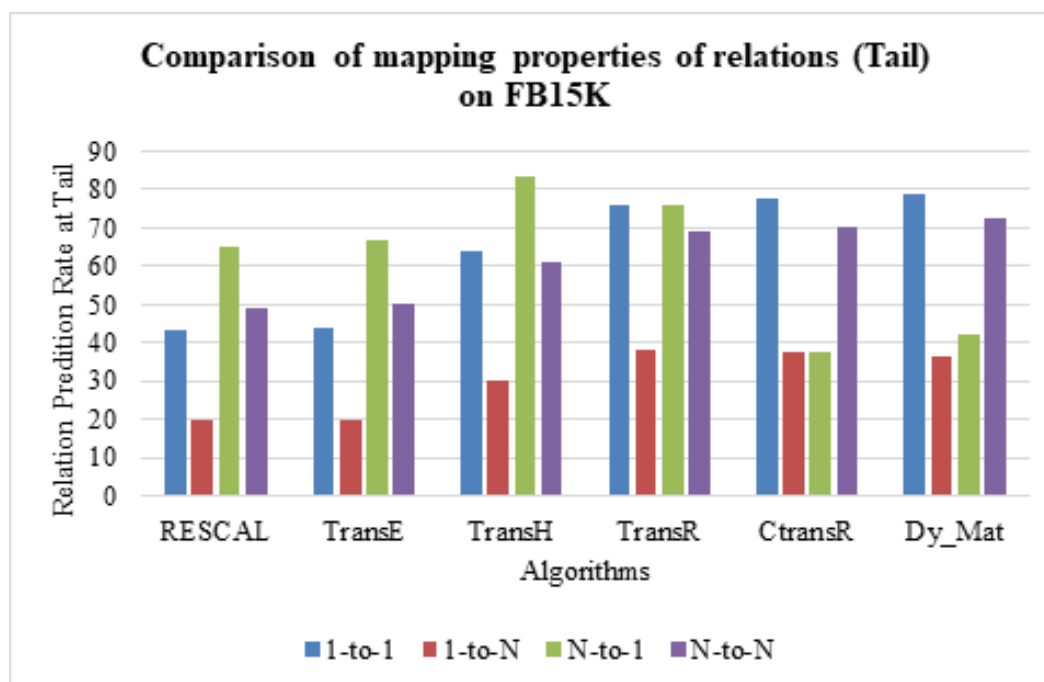


Fig 7. Comparison of mapping properties of relations (Tail) on FB15K

Figure 7, depicts the illustration of relations among the graph at various levels of relations in the Tail of the knowledge model on the FB15K dataset. The relation prediction rate at the tail of the knowledge graph is high for the Dy_Mat approach and the prediction rate is effective in the N-N and 1-1 relation when compared to the existing approaches.

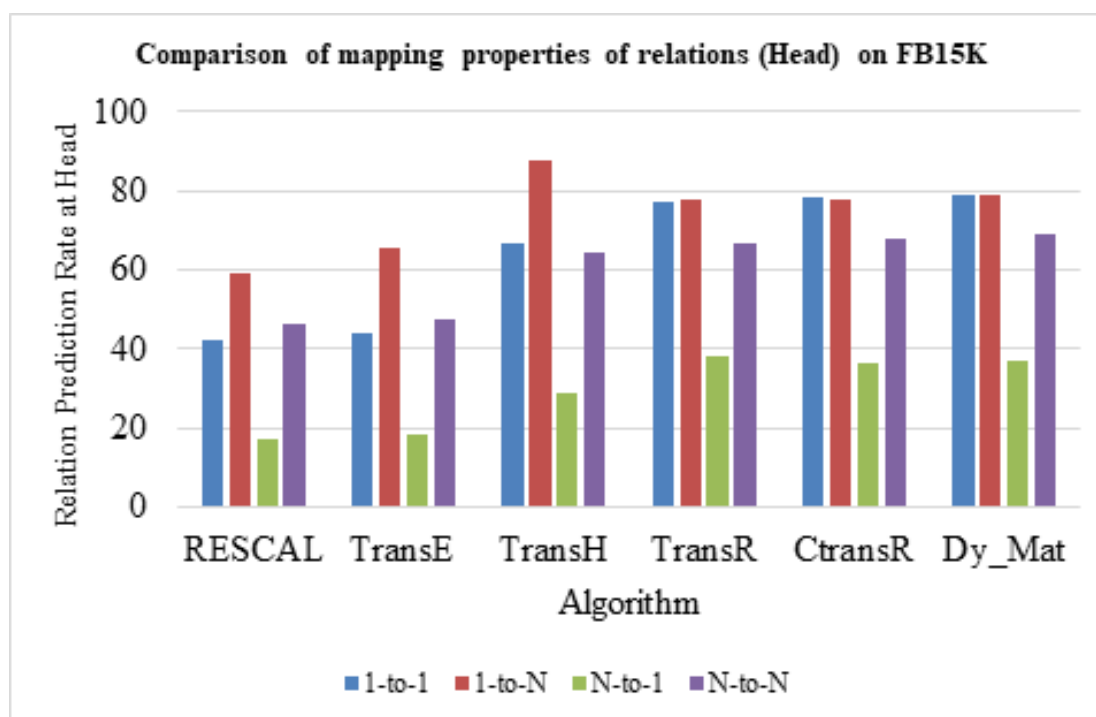


Fig 8. Comparison of mapping properties of relations (Head) on FB15K

Figure 8, depicts the illustration of relations among the graph at various levels of relations in the Head of the knowledge model on the FB15K dataset. The relation prediction rate at the head of the knowledge graph is high for the Dy_Mat approach when compared to the existing approaches.

Table 5. Experimental results on the dataset MN18RR by using the properties of mapping relations in percentage

Tasks	Prediction Tail (Hits@10)				Prediction Head (Hits@10)			
	1-to-1	1-to-N	N-to-1	N-to-N	1-to-1	1-to-N	N-to-1	N-to-N
RESCAL	45.1	20.3	75.2	59.2	52.3	59.3	27.1	56.1
TransE	44.3	21.7	76.4	56.6	53.6	65.7	28.5	57.2
TransH	69.1	34.5	85.2	66.1	69.9	87.6	38.7	67.4
TransR	86.1	40.5	84.2	72.1	79.5	77.9	39.1	68.1
CtransR	87.3	43.6	81.4	80.3	86.4	77.8	39.4	69.8
Dy_Mat	91.6	45.8	82.1	87.5	89.5	78.8	47.1	70.5

The detailed outcomes of mapping properties and their relation in the dataset MN18RR are shown in the Table 5. This table shows the tail and head levels of the knowledge graph and it consists of diverse relations namely 1-to-1, 1-to-N, N-to-1 and N-to-N. The comparison is carried for existing RESCAL, TransE, TransH, TransR, CtransR and proposed Dy_Mat approach.

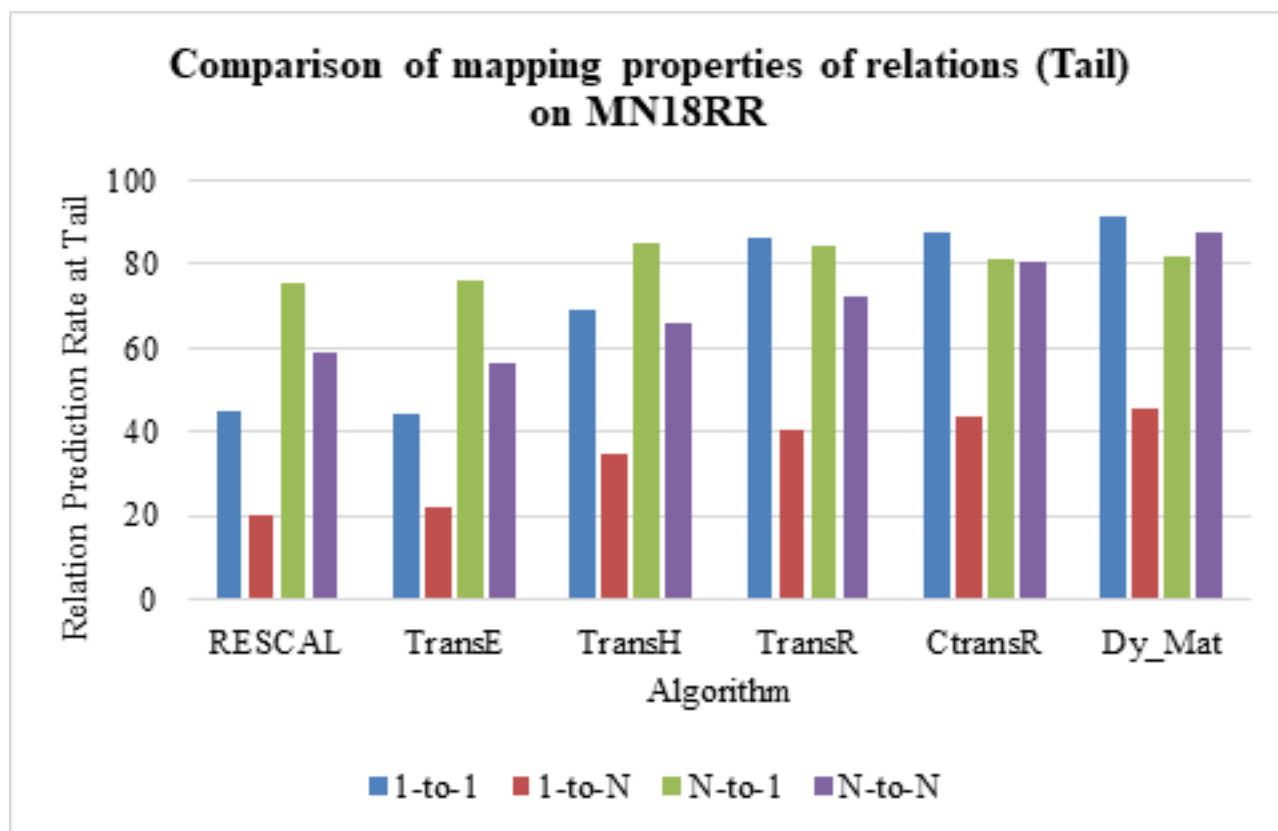


Fig 9. Comparison of mapping properties of relations (Tail) on MN18RR

Figure 9, depicts the illustration of relations among the graph at various levels of relations in the Tail of the knowledge model on the MN18RR dataset. The relation prediction rate at the tail of the knowledge graph is high for the Dy_Mat approach and the prediction rate is effective in the N-to-N, 1-to-N, and 1-to-1 relation when compared to the existing approaches.

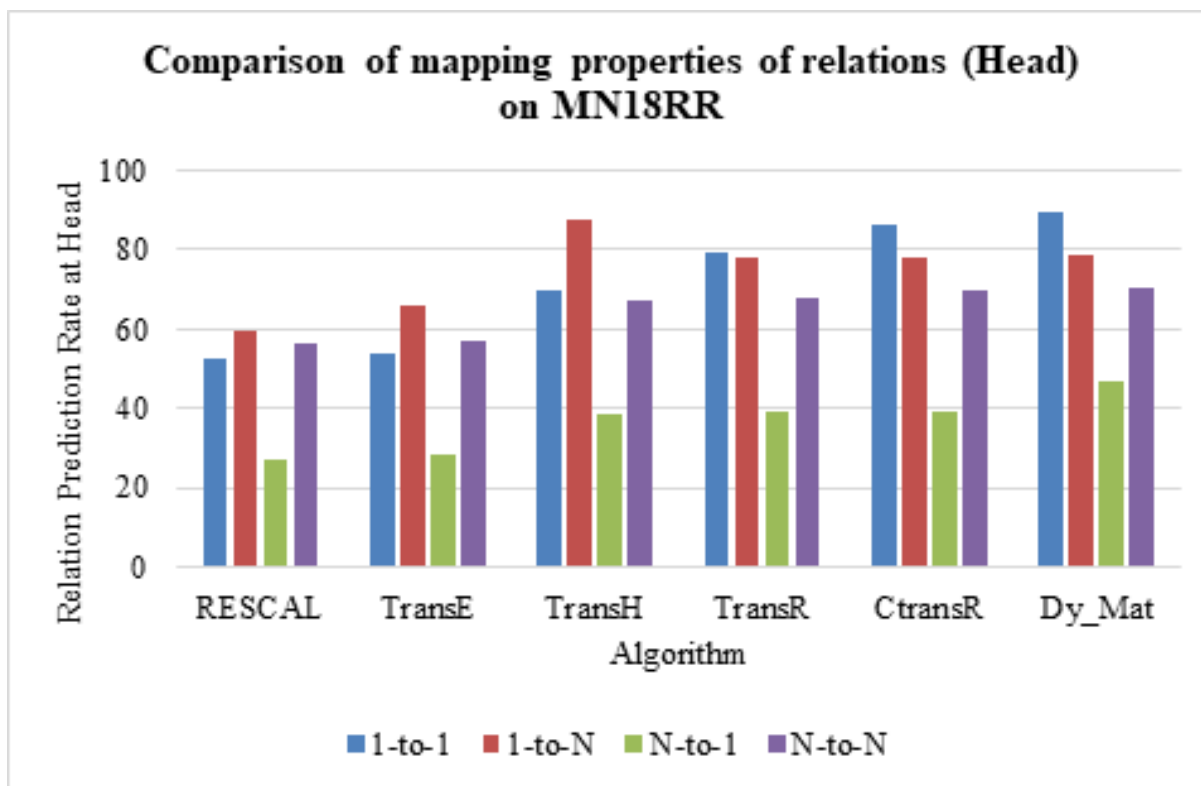


Fig 10. Comparison of mapping properties of relations (Head) on MN18RR

Figure 10, depicts the illustration of relations among the graph at various levels of relations in the Head of the knowledge model on the MN18RR dataset. The relation prediction rate at the tail of the knowledge graph is high for the Dy_Mat when compared to the existing approaches.

4.2 Triplet Classification

The process of classifying the triplet is to judge whether the triplets (HD, R, TL) is exact or not that is carried by the classification task. In this paper, two benchmark datasets are used namely FB15k-237 and MN18RR to investigate the approach. A threshold value is needed is assigned for every triplets and the threshold value is acquired by maximizing the accuracies of classification that is on the valid set.

Table 6. Triplet classification result

Dataset	FB15k-237	MN18RR
TransE	70.9	75.9
TransH	76.5	77.7
TransR	74.7	85.5
CtransR	81.9	85.7
RESCAL	83.1	87.1
Dy_Mat	86.4	88.7

In Table 6, outcome of the classification rate of existing and proposed approach is given. The marginal value γ elected among the values {1, 2, 5, 10}, the dimension of the relation vectors D_n and values of entity vectors D_m amongst the values {20, 50, 80, 100}. The enhanced configuration values acquired by valid set of values, they are $m, \gamma = 1, n = 100, B = 1000$ and taking L2 as variation on MN18RR; $m, \gamma = 2, n = 100, B = 4800$ and taking L1 as variation on FB15k-237. The proposed approach trains quicker than TransR (The available approach TransR necessitates 70 seconds and TransD simply expends 24 seconds on the dataset FB15k).

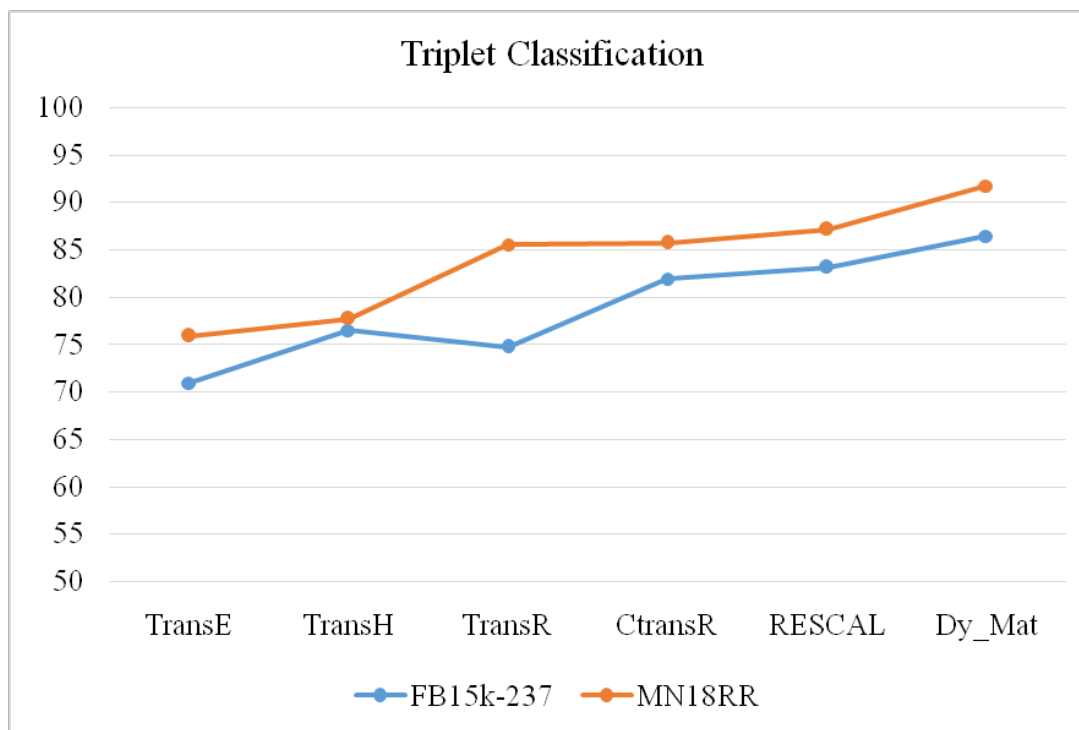


Fig 11. Comparison of triplet classification

In Figure 11, the classification range of the existing and proposed approach is given where the triplet classification rate is high in the proposed Dy_Mat approach. The classification is attained on two datasets namely FB15k-237 and MN18RR. The comparison is carried for existing RESCAL, TransE, TransH, TransR, CTransR and proposed Dy_Mat approach.

- The proposed novel method dynamic matrix using mapping properties will generate a mapping matrix dynamically for every pair of entity relation with the support of diverse nature of relation and entity. The representation of every vector to the space of relation vector is attained simply.
- The developed approach uses minimum parameters and the multiplication is not required, which makes the proposed scheme effective in the large scale graphs. The triple features in the graph will be explored deeper and the attainment of knowledge is effective.
- In the extensive experiment, dynamic matrix scheme outperforms the available knowledge graph approaches in the task of triplet classification and the relation prediction.

4 Conclusion

The dynamic matrix is introduced to mine the significant knowledge from the knowledge graph, the dynamic matrix implant the knowledge graphs into the vector space in a continuous form. The proposed dynamic matrix is less complex in processing and highly flexible than the existing knowledge graph approach RESCAL, TransE, TransH, TransR, CTransR and proposed Dy_Mat. The learning of the named symbols embedded in the relations or entity of a graph is complicated process. But, it is easily achieved by the dynamic matrix constructions. The experimental investigation shows that dynamic matrix using mapping properties outperforms the TransH, TransR/CTransR and TransE. Through the prediction of relation and classification of triplets, knowledge graph is evaluated. The classification results of triplet section shows that the new information are easily extracted from the available knowledge graph in the form of triplet.

References

- 1) Fu X, Ren X, Mengshoel OJ, Wu X. Stochastic optimization for market return prediction using financial knowledge graph. *IEEE International Conference on Big Knowledge*. 2018;p. 25–32.

- 2) Rotmensch M, Halpern Y, Tlimat A, Horng S, Sontag D. Learning a health knowledge graph from electronic medical records. *Scientific Reports*. 2017;7(1):1–11. Available from: <https://dx.doi.org/10.1038/s41598-017-05778-z>.
- 3) Xiong C, Power R, Callan J. Explicit semantic ranking for academic search via knowledge graph embedding. In: and others, editor. Proceedings of the 26th International Conference on World Wide Web. 2017;p. 1271–1279.
- 4) Nathani D, Chauhan J, Sharma C, Kaul M. Learning attention-based embeddings for relation prediction in knowledge graphs. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2019;p. 4710–4723.
- 5) Lyu X, Gu Y, Han X, Hou L, Li J, Liu Z. Adapting meta knowledge graph information for multi-hop reasoning over few-shot relations. In: and others, editor. Proceedings of the Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. 2019;p. 1–6.
- 6) Li C, Xian X, Ai X, Cui Z. Representation Learning of Knowledge Graphs with Embedding Subspaces. *Scientific Programming*. 2020;2020:1–10. Available from: <https://dx.doi.org/10.1155/2020/4741963>.
- 7) Bin C, Qin S, Rao G, Gu T, Chang L. Multiview Translation Learning for Knowledge Graph Embedding. *Scientific Programming*. 2020;2020:1–9. Available from: <https://dx.doi.org/10.1155/2020/7084958>. doi:10.1155/2020/7084958.
- 8) Tan Z, Zhao X, Fang Y, Xiao W, Tang J. Knowledge representation learning via dynamic relation spaces. 2016;p. 684–691.
- 9) Zhu JZ, Jia YT, Xu J. Modeling the Correlations of Relations for Knowledge Graph Embedding. *J Comput Sci Technol*. 2018;33:323–334.
- 10) Tan C, Yang K, Dai X, Huang S, Chen J. MSGE: A Multi-step Gated Model for Knowledge Graph Completion. vol. 2020. 2020;p. 424–435.
- 11) Chang L, Zhu M, Gu T, Bin C, Qian J, Zhang J. Knowledge Graph Embedding by Dynamic Translation. *IEEE Access*. 2017;5:20898–20907. Available from: <https://dx.doi.org/10.1109/access.2017.2759139>.
- 12) Xie R, Liu Z, Jia J, Luan H, Sun M. Representation learning of knowledge graphs with entity descriptions. 2016;p. 2659–2665.
- 13) Chen Z, Wang Y, Zhao B, Cheng J, Zhao X, &duan Z. Knowledge Graph Completion: A Review. *IEEE Access*. 2020;p. 1–1.
- 14) Celebi R, Uyar H, Yasar E, Gumus O, Dikenelli O, Dumontier M. Evaluation of knowledge graph embedding approaches for drug-drug interaction prediction in realistic settings. *BMC Bioinformatics*. 2019;20(1):726–726. Available from: <https://dx.doi.org/10.1186/s12859-019-3284-5>.
- 15) Zhao F, Sun H, Jin L, Jin H. Structure-augmented knowledge graph embedding for sparse data with rule learning. *Computer Communications*. 2020;159:271–278. Available from: <https://dx.doi.org/10.1016/j.comcom.2020.05.017>.
- 16) Sousa RT, Silva S, Pesquita C. Evolving knowledge graph similarity for supervised learning in complex biomedical domains. *BMC Bioinformatics*. 2020;21(1):6–6. Available from: <https://dx.doi.org/10.1186/s12859-019-3296-1>.
- 17) Trouillon T, Wealbl J, Riedel S, Gaussier E, Bouchard G. Complex embeddings for simple link prediction. In: and others, editor. Proceedings of the 33rd International Conference on Machine Learning;vol. 48. 2016;p. 2071–2080.
- 18) Yang B, Yih WT, He X, Gao J, Deng L. Embedding entities and relations for learning and inference in knowledge bases. In: and others, editor. Proceedings of the International Conference on Learning Representations. 2015.
- 19) Nickel M, Rosasco L, Poggio T. Holographic embeddings of knowledge graphs. In: and others, editor. Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence. 2016;p. 1955–1961.
- 20) Hu X, Duan J, Dang D. Scalable aggregate keyword query over knowledge graph. *Future Generation Computer Systems*. 2020;107:588–600. Available from: <https://dx.doi.org/10.1016/j.future.2020.02.011>.
- 21) Sutskever I, Tenenbaum JB, Salakhutdinov RR. Modelling relational data using bayesian clustered tensor factorization. In: and others, editor. InProceedings of the 22nd International Conference on Neural Information Processing Systems. 2019;p. 1821–1828.
- 22) Ji G, Liu K, He S, Zhao J. Knowledge graph completion with adaptive sparse transfer matrix. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. 2016;p. 985–991.