Joint Relational Co-evolution in Spatial-Temporal Knowledge Graph for SMEs Supply Chain Prediction

Presenter: Li Zimeng

Li, Youru, et al. Learning Joint Relational Co-evolution in Spatial-Temporal Knowledge Graph for SMEs Supply Chain Prediction. *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2023.

Background

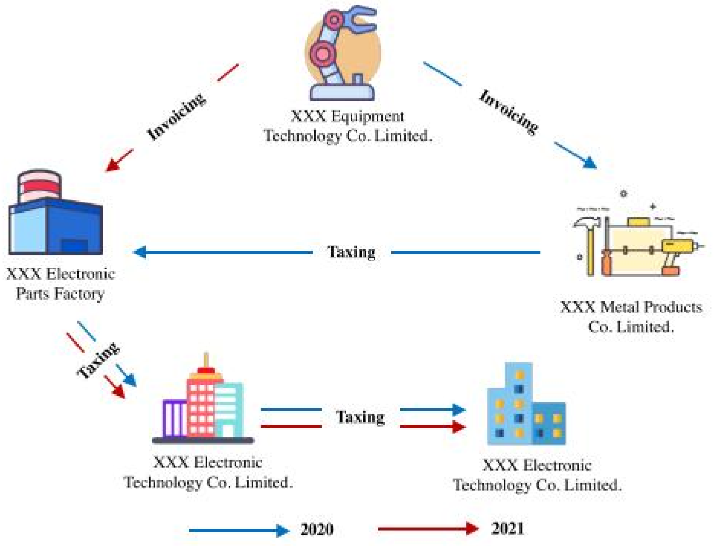
What is supply chain?

A supply chain, sometimes expressed as a "supply-chain", is a complex logistics system that consists of facilities that convert raw materials into finished products and distribute them to end consumers or end customers.

What can we benefit from supply chain prediction

With the help of supply chain networks,

- enterprises can improve their production and operation efficiency.
- banks can clarify the operational relationship between enterprises, so as to grant credits more accurately and provide financing assistance in a more targeted manner.



Motivation

- The real supply chain relationships among SMEs can change more profoundly than that, which reflects the necessity of modeling temporal dependency in supply chain relationships. However, existing models hardly consider the dynamics of SMEs relationships along the timeline.
- The cooperative correlation among the supply chain relationships should also be noted, which could be inferred from plural connection paths in KG.
- Plural geographical distribution of supply chains resulting from the uneven regional economic development and diverse industrial structures means that geographical information of SMEs should also be considered in supply chain prediction.

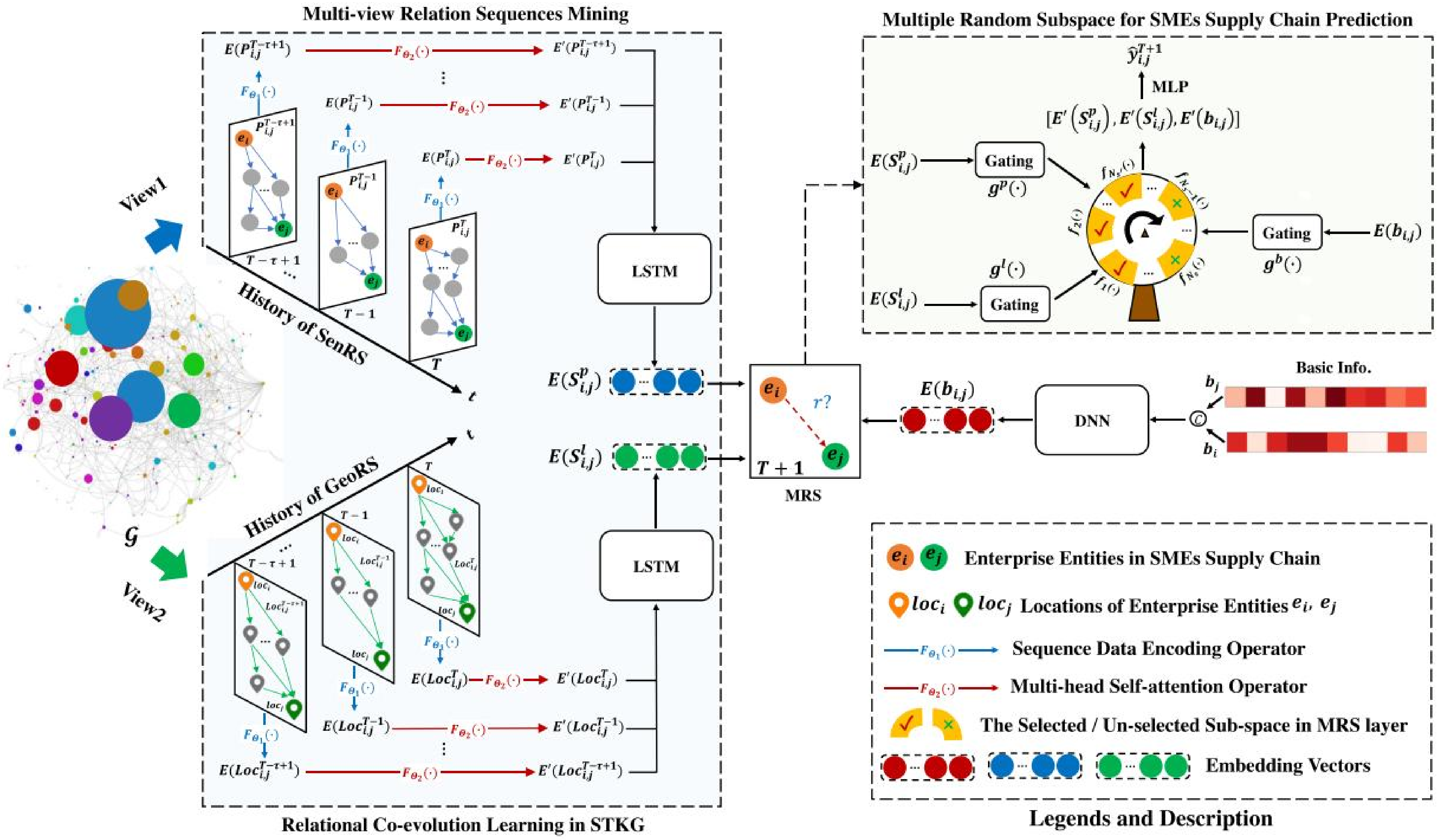
Formalization

- Key Notations
  - $E = \{e_1, e_2, \ldots, e_{|E|}\}$ denotes a set of SMEs.
  - $R=\{r_{i,j} \in \{0, 1, \emptyset\}\}$ denotes the set of supply chain relationships, where 1 or 0 denotes whether a supply chain relationship between them or not and Ø means an unknown relationship.
  - $B = \{b_1, b_2, \ldots, b_{|E|}\}$ denotes basic information of SMEs.

- Task Definition
  - To predict whether the enterprise e_i will have any supply chain relationships with e_j at future time:

$$\widehat{y}_{i,j}^{T+1} = \mathcal{F}_{\Theta}(< e_i, e_j > | R, B, \mathcal{G})$$

Model Design



Model Design

- Multi-view Relation Sequences Mining
  - Semantic-view Relation Sequence (SenRS)

  Sample K-shortest paths on KG at each time t, then embed K sequences as $E(P_{i,j}^{t}) \in R^{K\times d}$

$$p_{i,j}^{t}|_{k} = \{e_i, \psi(l_{e_i,\bar{e}_u}^{t}), \bar{e}_u^{t}, \ldots, \bar{e}_v^{t}, \psi(l_{\bar{e}_v,e_j}^{t}), e_j\}$$

  - Geographical-view Relation Sequence (GeoRS)

  For each p seq, the locations of SMEs constitute the loc seq. Embed the K loc seqs in the same way, getting $E(Loc_{i,j}^{t}) \in R^{K\times d}$

$$loc_{i,j}^{t}|_{k} = \{loc_i^{t}, loc_u^{t}, \ldots, loc_v^{t}, loc_j^{t}\}.$$

Model Design

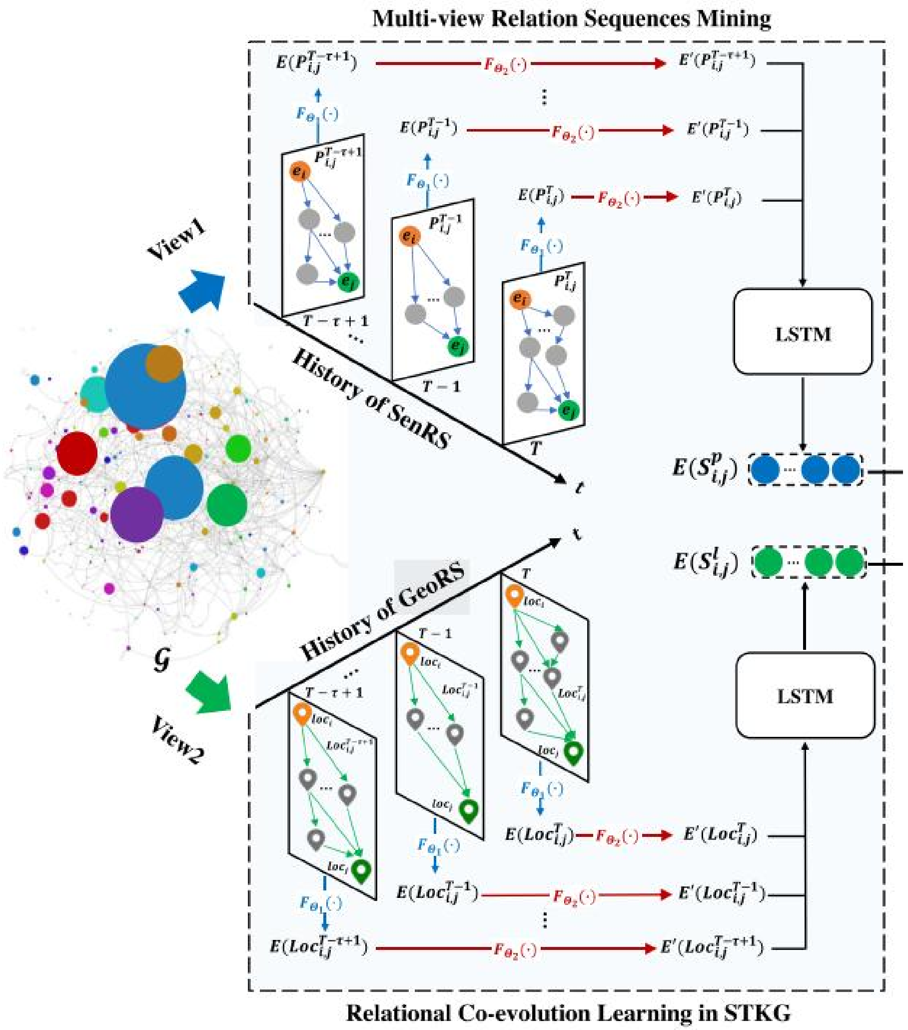
- Relational Co-evolution Learning
  - Cooperative Correlation Refining

$$head^{(h)}(E(P_{i,j}^{t})) = softmax(\frac{Q_{i,j}^{(h)} \cdot K_{i,j}^{(h)\mathrm{T}}}{\sqrt{d_1}}) \cdot V_{i,j}^{(h)}$$

$$\begin{aligned} E'(P_{i,j}^{t}) &= MultiHeadAttention(E(P_{i,j}^{t})) \\ &= [E'^{(1)}_{(P_{i,j}^{t})} \oplus E'^{(2)}_{(P_{i,j}^{t})} \ldots \oplus E'^{(N_H)}_{(P_{i,j}^{t})}] \cdot W^{H} \end{aligned}$$

  - Temporal Dependency Modeling

$$E(S_{i,j}^{P}) = LSTM^{(P)}(E'(P_{i,j}^{T-\tau+1}), \ldots, E'(P_{i,j}^{T-1}), E'(P_{i,j}^{T}))$$



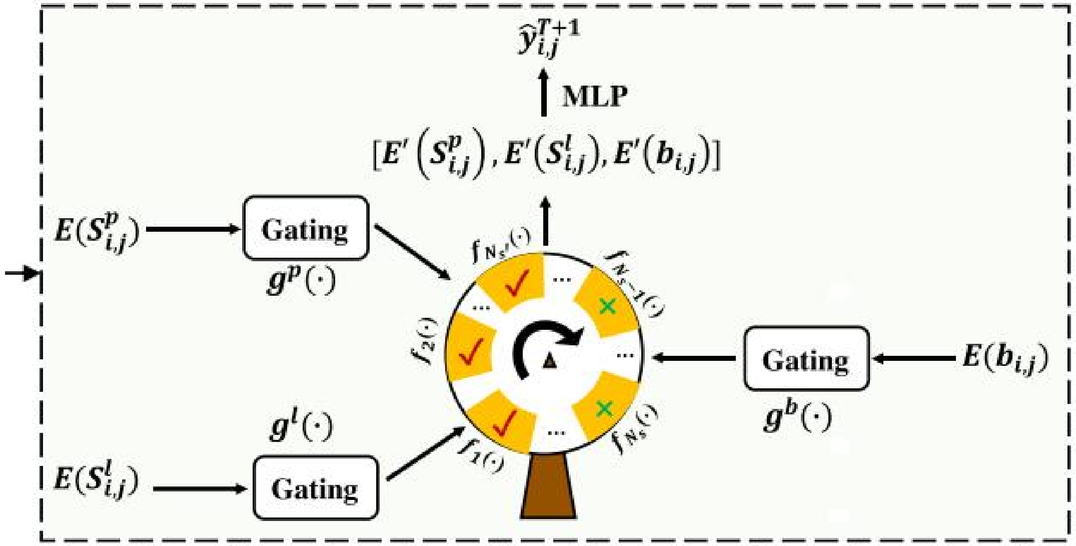
Model Design

- Multiple Random Subspaces
  - To disentangle the learned representations
  - Map learned representations into different N_s subspaces to capture connectivity patterns from different aspects
  - Randomly sample a subset of size N'_s from the N_s subspaces to alleviate winner-take-all issue and strengthen the uncertainty

$$E'(S^p_{i,j}) = \sum_{n=1}^{N'_s} g^p_n(E(S^p_{i,j})) \cdot f_n(E(S^p_{i,j}))$$

$$\hat{y}^{T+1}_{i,j} = MLP([E'(S^p_{i,j}), E'(S^l_{i,j}), E'(b_{i,j})])$$

Multiple Random Subspace for SMEs Supply Chain Prediction



Hint: this trick could be used to enhance representation learning and information fusion in a variety of tasks.

Optimization

The SMEs supply chain prediction task is defined as a binary classification problem, and the cross entropy is set as the loss function,

$$\mathcal{L}_{Class} = -\left(\sum_{y_{i,j}^{T+1} \in R^{+}} \log \widehat{y}_{i,j}^{T+1} + \sum_{y_{i,j}^{T+1} \in R^{-}} \log(1 - \widehat{y}_{i,j}^{T+1}) \right)$$

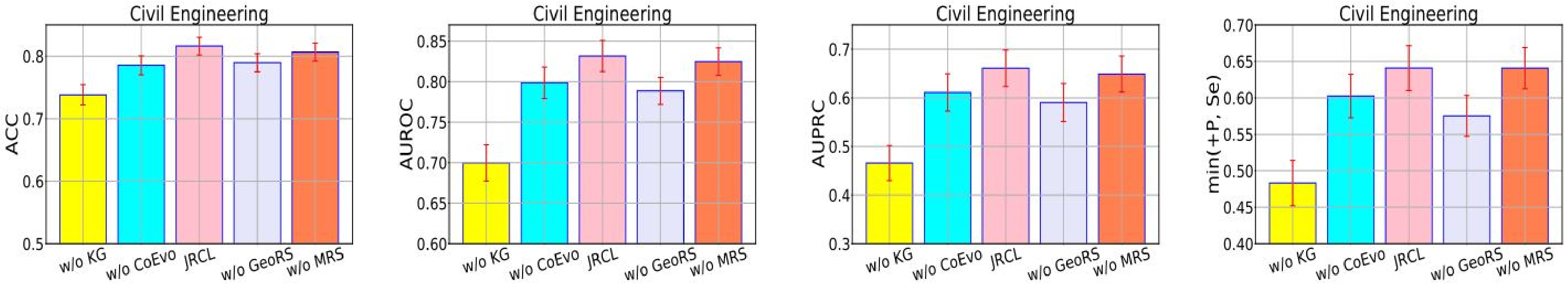
in which R+ and R− are the samples with positive and negative target values.

Performance Comparison

Dataset	Model	SMEs Supply Chain Prediction (Bootstrapping = 1000)			
		ACC	AUROC	AUPRC	min(Se, P+)
Civil Eng.	LR [3]	0.7219 ± 0.0161	0.6014 ± 0.0242	0.4027 ± 0.0355	0.4154 ± 0.0319
	GBDT [5]	0.7290 ± 0.0167	0.5769 ± 0.0231	0.3659 ± 0.0316	0.3509 ± 0.0288
	KGAT [35]	0.7272 ± 0.0163	0.6281 ± 0.0235	0.3920 ± 0.0324	0.4167 ± 0.0298
	KGIN [36]	0.7358 ± 0.0166	0.6302 ± 0.0238	0.4043 ± 0.0338	0.4132 ± 0.0318
	TiRGN [15]	**0.7437 ± 0.0167**	0.6403 ± 0.0240	0.4242 ± 0.0359	0.4593 ± 0.0307
	ST-GNN [40]	0.7402 ± 0.0162	**0.7307 ± 0.0204**	**0.4808 ± 0.0371**	**0.5028 ± 0.0310**
	JRCL	**0.8161 ± 0.0143**	**0.8317 ± 0.0169**	**0.6607 ± 0.0375**	**0.6407 ± 0.0283**

ACC: accuracy
AUROC: the area under the receiver operating characteristic curve
AUPRC: the area under the precision-recall curve
min (Se, P+) : the minimum precision and sensitivity

Ablation study



- JRCL (w/o KG): To demonstrate the effectiveness of introducing auxiliary information from **knowledge graphs**, we only applied a plain DNN to make a prediction with the basic information of SMEs.

- JRCL (w/o CoEvo): We also removed the cooperative **correlation refining module** to demonstrate its usefulness in capturing the inherent correlation among each individual sequence in multi-view relation sequences at time t.

- JRCL (w/o GeoRS): We only consider the SenRs to demonstrate the effectiveness of modeling the **geographical relationships** between enterprises in an explicit manner.

- JRCL (w/o MRS): We removed the **MRS module** to demonstrate the effectiveness of aggregating multiple representations by coordinating the coupling between temporal dependency and cooperative correlation adequately.

Thanks