Latent Network Summarization: Bridging Network Embedding and Summarization

Di Jin, Ryan A. Rossi, Eunyee Koh, Sungchul Kim, Anup Rao, Danai Koutra

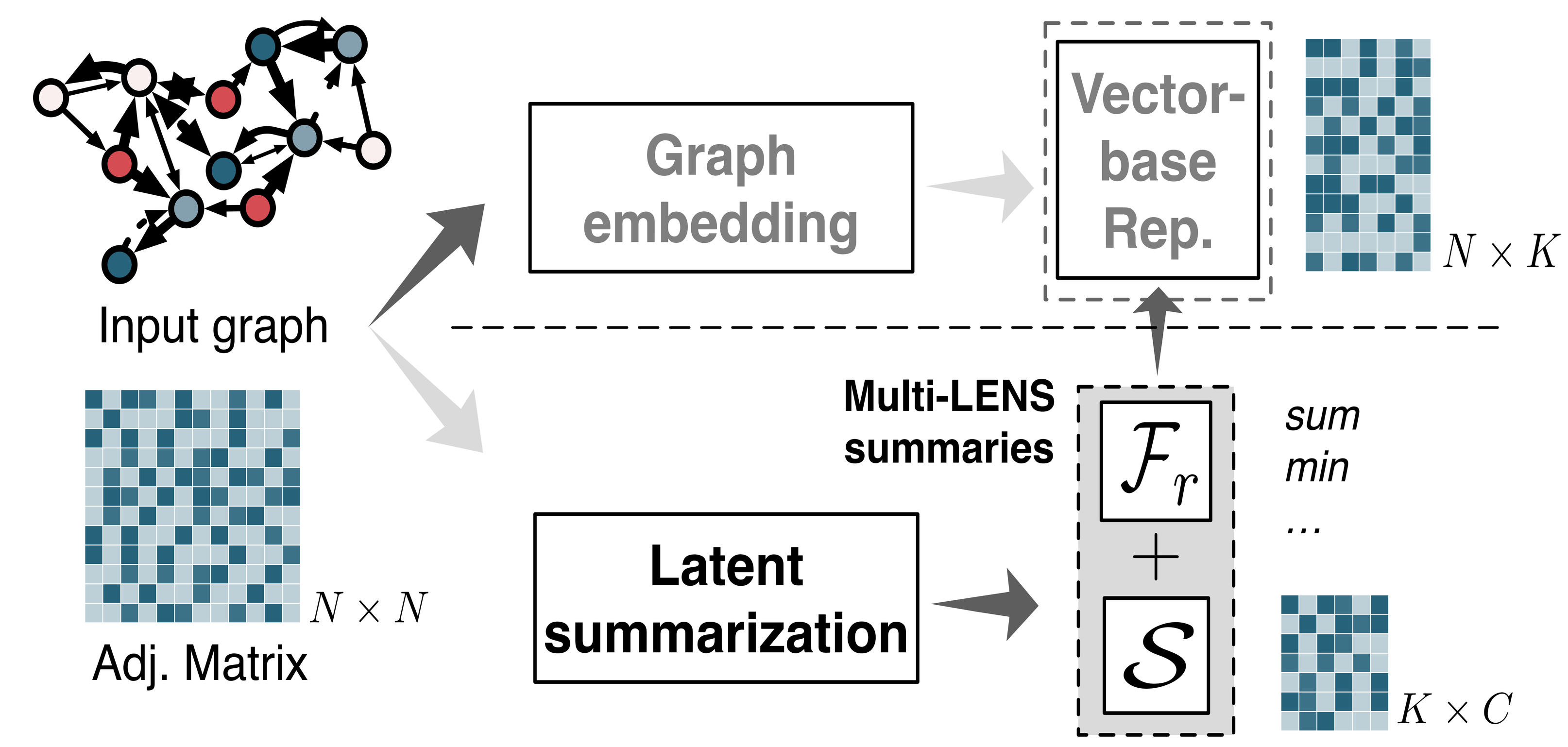
Motivation

- Node-wise embeddings are expensive in storage & computation.
- Graph summarization achieves efficiency, but does not retain info to derive node embeddings for ML tasks.

Problem

Given a heterogeneous graph G(N, M), latent graph summarization learns a compressed representation that captures the main graph structural information s.t.

- It is independent of graph size (N, M).
- It is capable of deriving node representations on the fly.

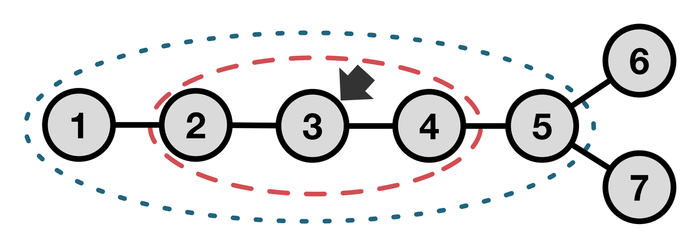


We propose the problem of *latent network summarization* that seeks to find a *compressed* graph representation that also retains the ability to derive node embeddings for ML tasks.

① Multi-level feature extraction

- **Relational operators** A general approach to collect structural features in a subgraph.
  - Denoted as $\phi(\mathbf{x}, \mathcal{S})$.
  - E.g., *sum* $\sum_{i \in \mathcal{S}} \mathbf{x}_i$

    max/min $\max/\min_{i \in \mathcal{S}} \mathbf{x}_i$

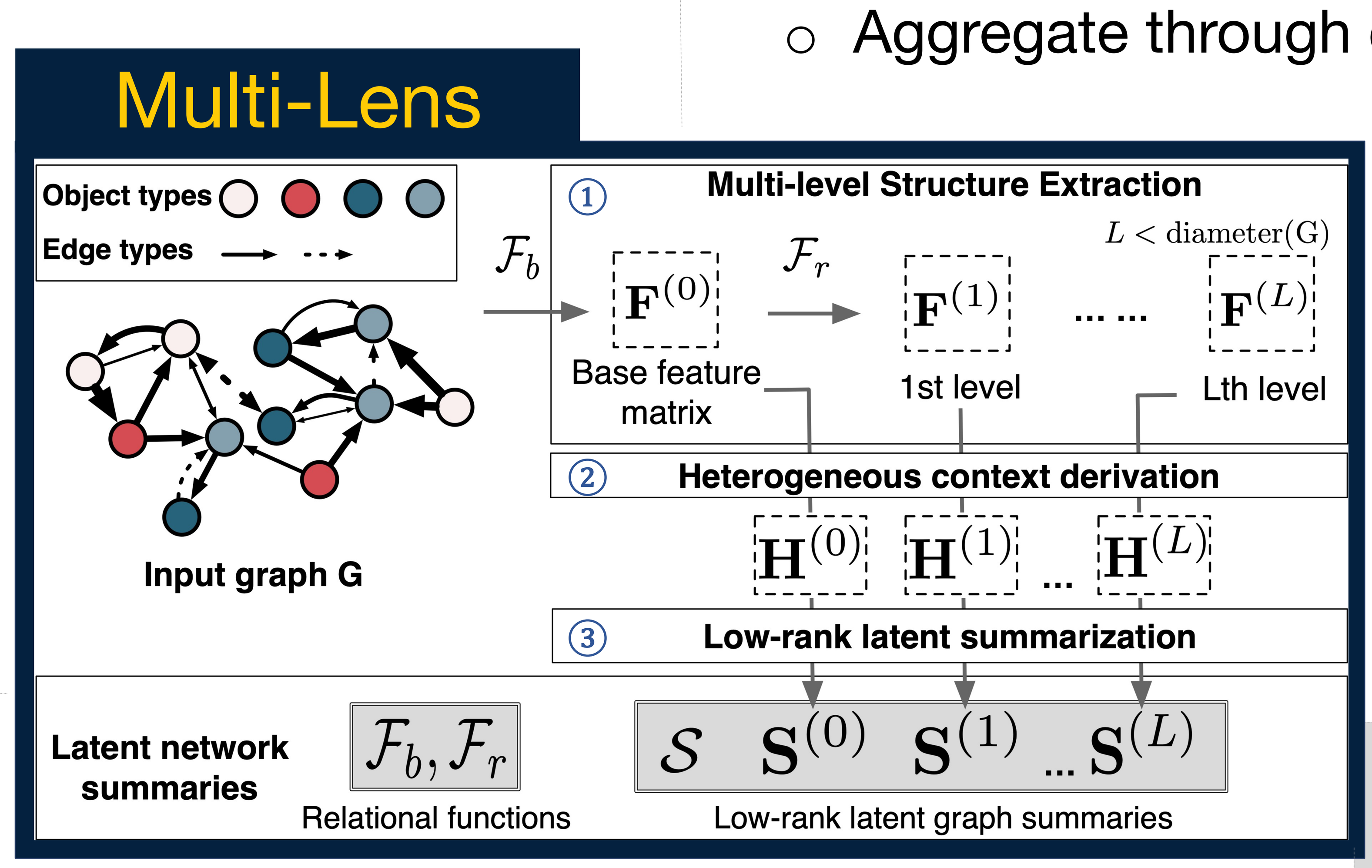


- **Relational functions** The composition of operators.

$$\underbrace{(\phi_k \circ \cdots \circ \phi_j \circ \phi_i)}_{h\ times}(\mathbf{x}, \mathcal{S})$$

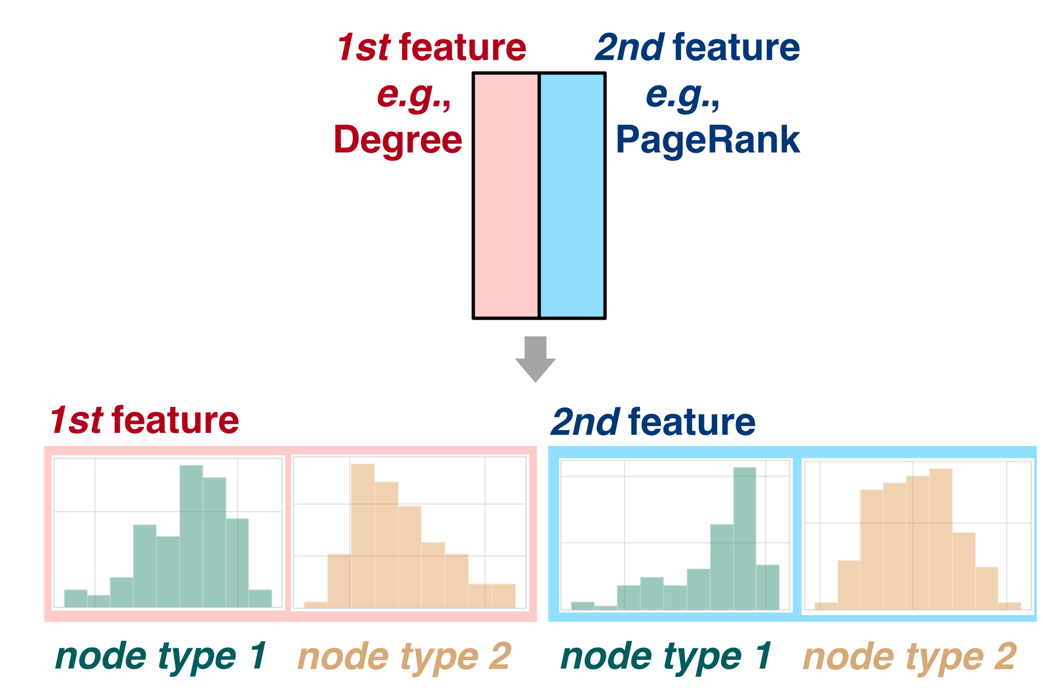
  - Produce (non)linear features in multi-level subgraphs
  - Stored as a part of summary

Multi-Lens



② Histogram-represented context

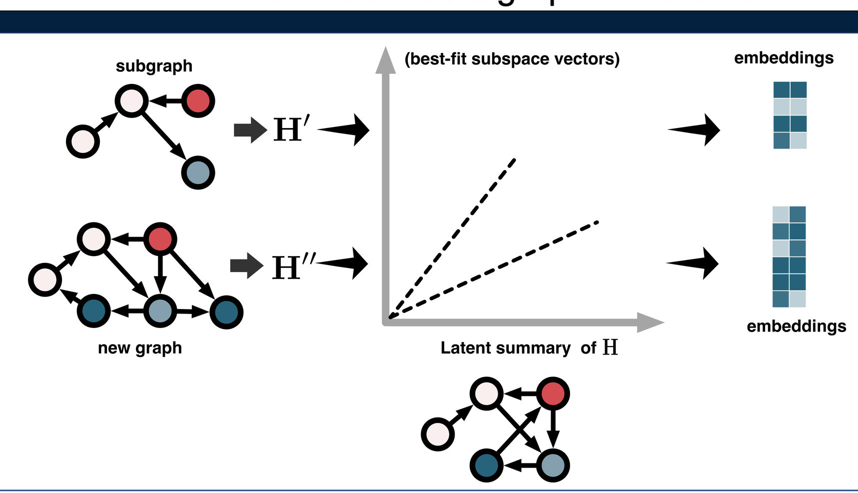
- Represent feature vectors as histograms with fixed lengths.
- Distinguish with respect to node type, edge type & edge directionality.
- Aggregate through concatenation.



③ Best-fit subspace vectors as latent summary

Low-rank approximation (e.g., SVD).

- Level-l embeddings (not stored). $\mathbf{Y}^{(\ell)} = \mathbf{U}^{(\ell)}\sqrt{\Sigma^{(\ell)}}$
- Level-l summary (stored) $\mathbf{S}^{(\ell)} = \sqrt{\Sigma^{(\ell)}}\mathbf{V}^{(\ell)T}$
- Stored to derive node embeddings.



Complexity

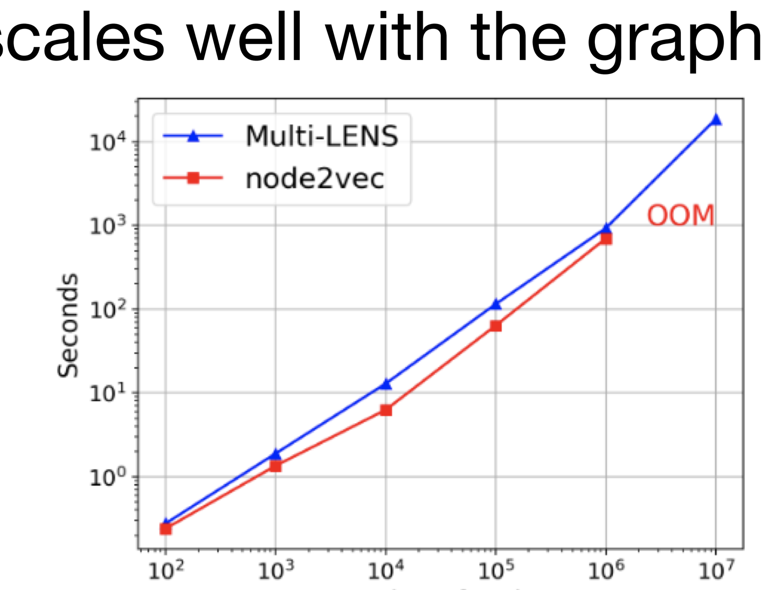
- Computational complexity linear on the # edges.
- Only ① and ③ are stored.
- Easy to parallelize.

Evaluation

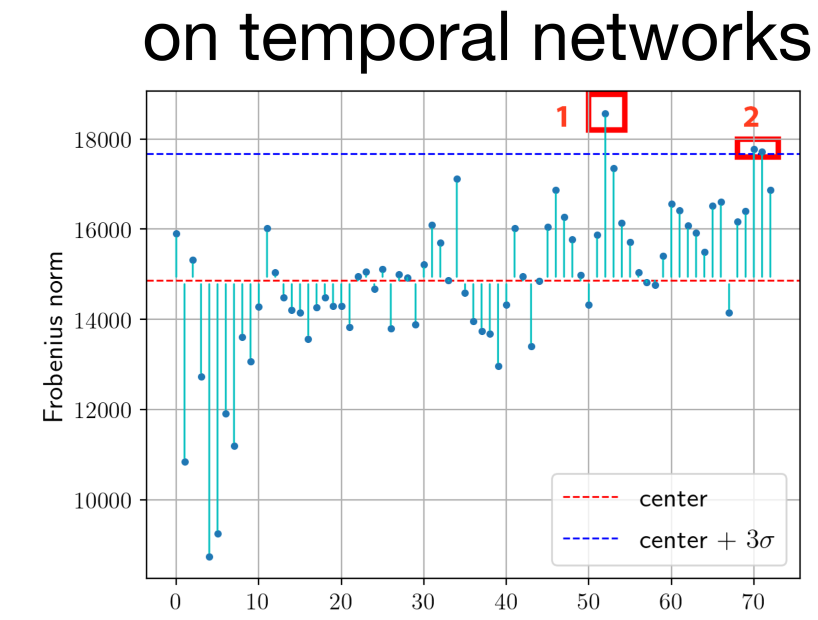
① Output storage. 3 – 2152 × less than baselines

Data	SE	LINE	n2vec	DW	m2vec	AspEm	G2G	ML (MB)
facebook	8.13x	8.48x	12.79x	12.84x	3.82x	8.50x	9.17x	0.58
yahoo	187.1x	180.0x	242.2x	231.0x	79.8x	197.4x	195.8x	0.62
dbpedia	710.0x	714.2x	996.4x	996.2x	-	749.2x	743.6x	0.81
digg	608.2x	612.8x	848.9x	830.3x	259.9x	641.7x	635.2x	0.54
bibson.	1512.1x	1523.0x	2152.5x	2152.5x	-	1595.8x	-	0.75

③ Scalability. Multi-Lens scales well with the graph size.



④ Inductive anomaly detection on temporal networks



② Link prediction. Multi-Lens outperforms all baselines by every metric.

Data	Metric	NA	SE	LINE	DW	n2vec	GR	s2vec	DNGR	m2vec	AspEm	G2G	ML($L = 1$)	ML($L = 2$)
facebook	**AUC**	0.6213	0.6717	0.7948	0.7396	0.7428	0.8157	0.8155	0.7894	0.7495	0.5886	0.7968	0.8703	**0.8709***
	ACC	0.5545	0.5995	0.7210	0.6460	0.6544	0.7368	0.7388	0.7062	0.7051	0.5628	0.7274	**0.7920***	0.7904
	F1 macro	0.5544	0.5716	0.7210	0.6296	0.6478	0.7367	0.7387	0.7060	0.7041	0.5628	0.7273	**0.7920***	0.7905
yahoo-msg	**AUC**	0.7189	0.5375	0.6745	0.7715	0.7830	0.7535	OOT	OOM	0.6708	0.5587	0.6988	0.8443	**0.8446***
	ACC	0.2811	0.5224	0.6269	0.6927	0.7036	0.6825			0.6164	0.5379	0.6564	**0.7587***	**0.7587***
	F1 macro	0.2343	0.5221	0.6265	0.6897	0.7016	0.6821			0.6145	0.5377	0.6562	**0.7577***	**0.7577***
dbpedia	**AUC**	0.6002	0.5211	0.9632	0.8739	0.8774	OOM	OOT	OOM	OOT	0.6364	0.7384	**0.9820***	0.9809
	ACC	0.3998	0.5399	0.9111	0.8436	0.8436					0.5869	0.6625	**0.9186**	0.9151
	F1 macro	0.2968	0.4539	0.9110	0.8402	0.8402					0.5860	0.6613	**0.9186**	0.9150
digg	**AUC**	0.7199	0.6625	0.9405	0.9664	0.9681	OOM	OOT	OOM	0.9552	0.5644	0.8978	**0.9894***	0.9893
	ACC	0.2801	0.6512	0.8709	0.9023	0.9049				0.8891	0.5459	0.8492	**0.9596***	0.9590
	F1 macro	0.2660	0.6223	0.8709	0.9019	0.9046				0.8890	0.5459	0.8492	**0.9595***	0.9590

References

- Rossi et al. *Deep inductive network representation learning*. WebConf'18 (companion)
- Hamilton, Ying, and Leskovec. *Inductive representation learning on large graphs*. NIPS'17
- Jin and Koutra. Exploratory analysis of graph data by leveraging domain knowledge. ICDM'17
- Koutra et al. Deltacon: A principled massive-graph similarity function. SDM'13