Dynamic Bundling with Large Language Models for Zero-Shot Inference on Text-Attributed Graphs

Presenter: Yusheng Zhao

Affiliation: Peking University

Background: Text-attributed Graphs (TAGs)

- A TAG contains text attributes associated with each node
- Very common in real-world applications:
  - Citation networks
  - Social networks
  - E-commerce networks
  - Webpage networks
- Our goal: zero-shot inference on TAGs
  - Input: TAG
  - Output: node labels
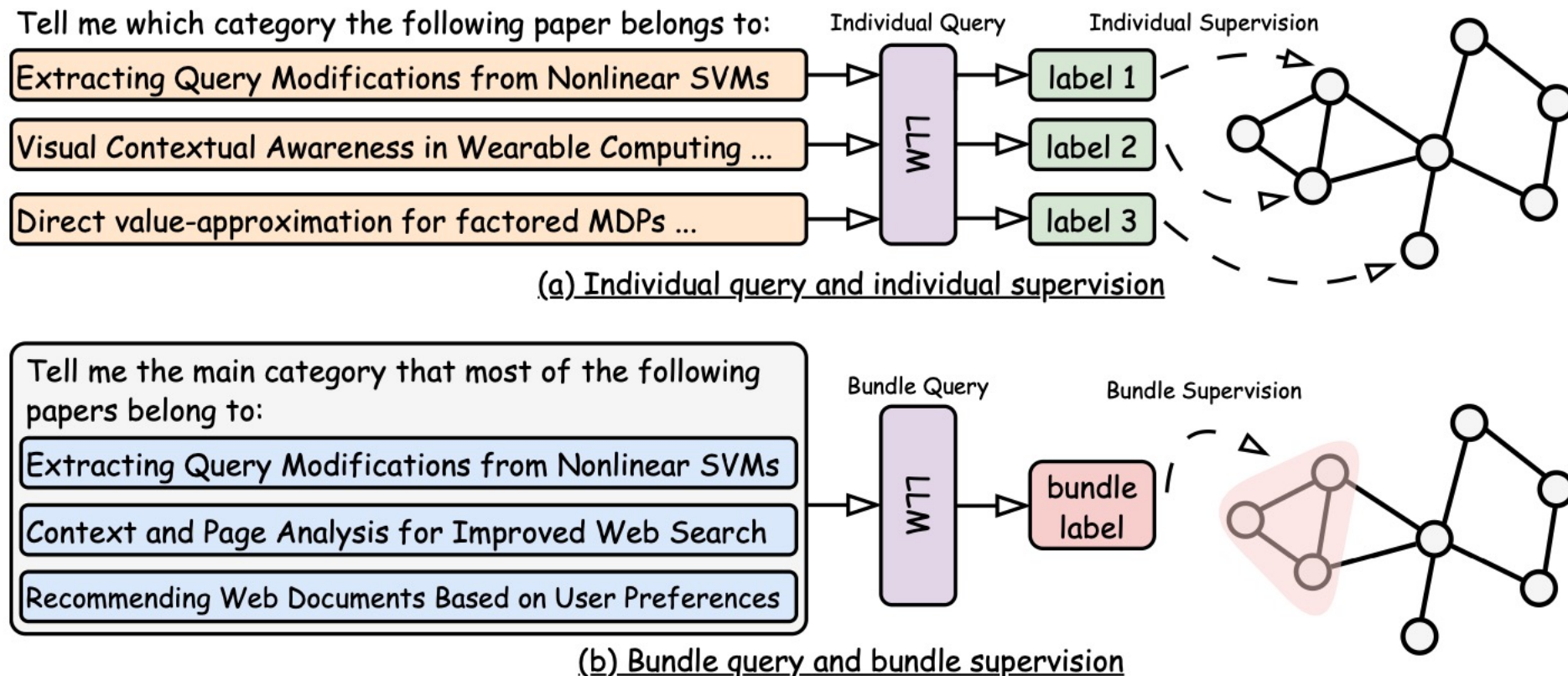
Core Idea: Bundling

Figure 1: (a) Querying LLMs with individual texts and supervising graph learning with individual labels. (b) By creating text bundles, we perform bundle queries to obtain bundle labels for supervision.

Methodology

- **Bundle Sampling**: sample the nodes/texts to form bundles
- **Bundle Query**: query the LLM with bundles of texts to obtain bundle labels
- **Bundle Supervision**: supervise a GNN with bundle labels
- **Bundle Refinement**: evict noisy nodes in the bundle

(a) Bundle Sampling (b) Bundle Query (c) Bundle Supervision (d) Bundle Refinement
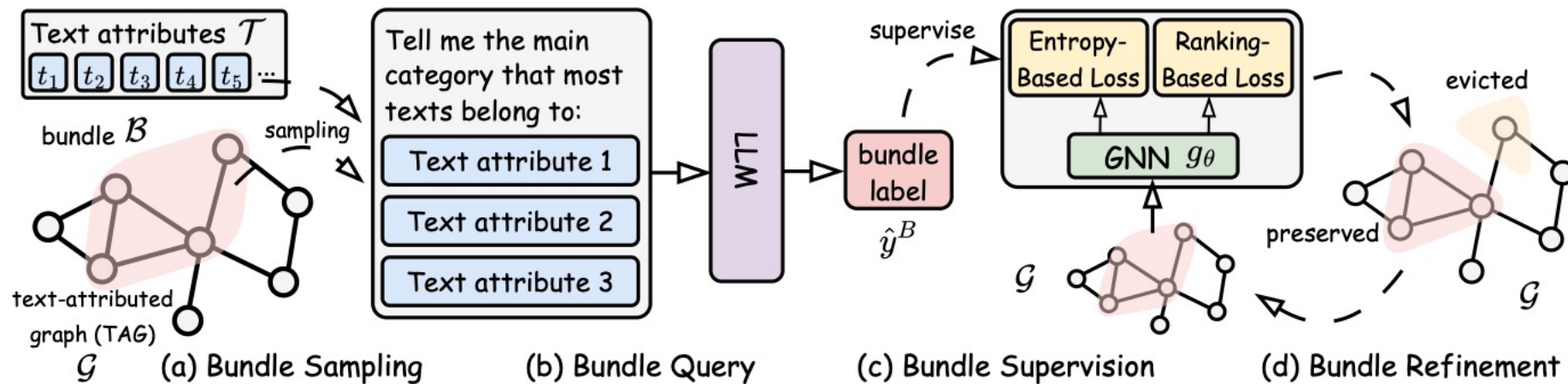
Experiments

Table 1: Prediction accuracies of our method compared to baselines across datasets. We mark the best results in **bold** and the second-best with underline.

Method	Cora	CiteSeer	WikiCS	History	Children	Sportsfit	Cornell	Texas	Wisc.	Wash.
SBERT	69.75	66.69	59.06	53.53	22.59	43.79	63.66	64.58	62.10	63.52
RoBERTa	70.71	66.95	59.08	55.39	24.25	41.51	61.68	62.25	60.33	60.60
TE-3-Large	71.90	66.24	61.78	50.15	24.68	58.39	81.50	75.42	73.14	66.35
LLM2Vec	67.34	67.13	62.34	53.14	25.56	57.00	81.26	76.68	73.36	65.92
GPT-3.5-turbo	70.11	66.83	65.53	55.07	29.73	<u>67.21</u>	45.54	56.14	58.86	51.09
GPT-4o	70.29	64.77	66.10	53.30	<u>30.76</u>	66.35	45.54	63.10	56.60	48.90
DGI	16.79	15.24	14.98	20.98	2.22	7.48	14.66	11.23	12.08	20.96
GraphMAE	15.13	8.11	8.91	36.36	7.24	30.50	23.04	17.65	23.02	24.89
OFA	20.36	41.31	30.77	8.25	3.05	15.18	29.84	11.77	4.80	6.04
GOFA	71.06	65.72	<u>68.62</u>	56.25	12.15	37.87	39.50	38.37	32.51	31.02
UniGLM	45.57	52.26	55.05	44.24	21.48	33.46	23.03	21.39	27.16	24.01
ZeroG	60.40	50.35	46.74	36.55	12.72	14.27	10.47	53.48	12.66	8.30
GraphGPT	17.48	13.93	33.59	12.31	9.94	4.53	10.18	18.48	12.35	20.64
LLAGA	11.62	19.52	10.98	7.95	10.09	1.84	12.57	15.51	15.09	10.48
LLM-BP	<u>72.59</u>	<u>69.51</u>	67.75	<u>59.86</u>	24.81	61.92	<u>83.28</u>	<u>81.66</u>	<u>77.75</u>	<u>73.14</u>
DENSE (ours)	**75.09**	**72.37**	**71.03**	**67.31**	**31.75**	**75.88**	**84.82**	**92.51**	**87.17**	**81.66**

Thank You

* If you are interested, feel free to read the original paper or contact the author.