Learning Causally Invariant Representations for Out-of-Distribution Generalization on Graphs

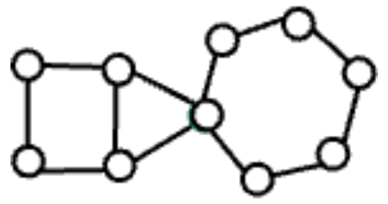
Yongqiang Chen

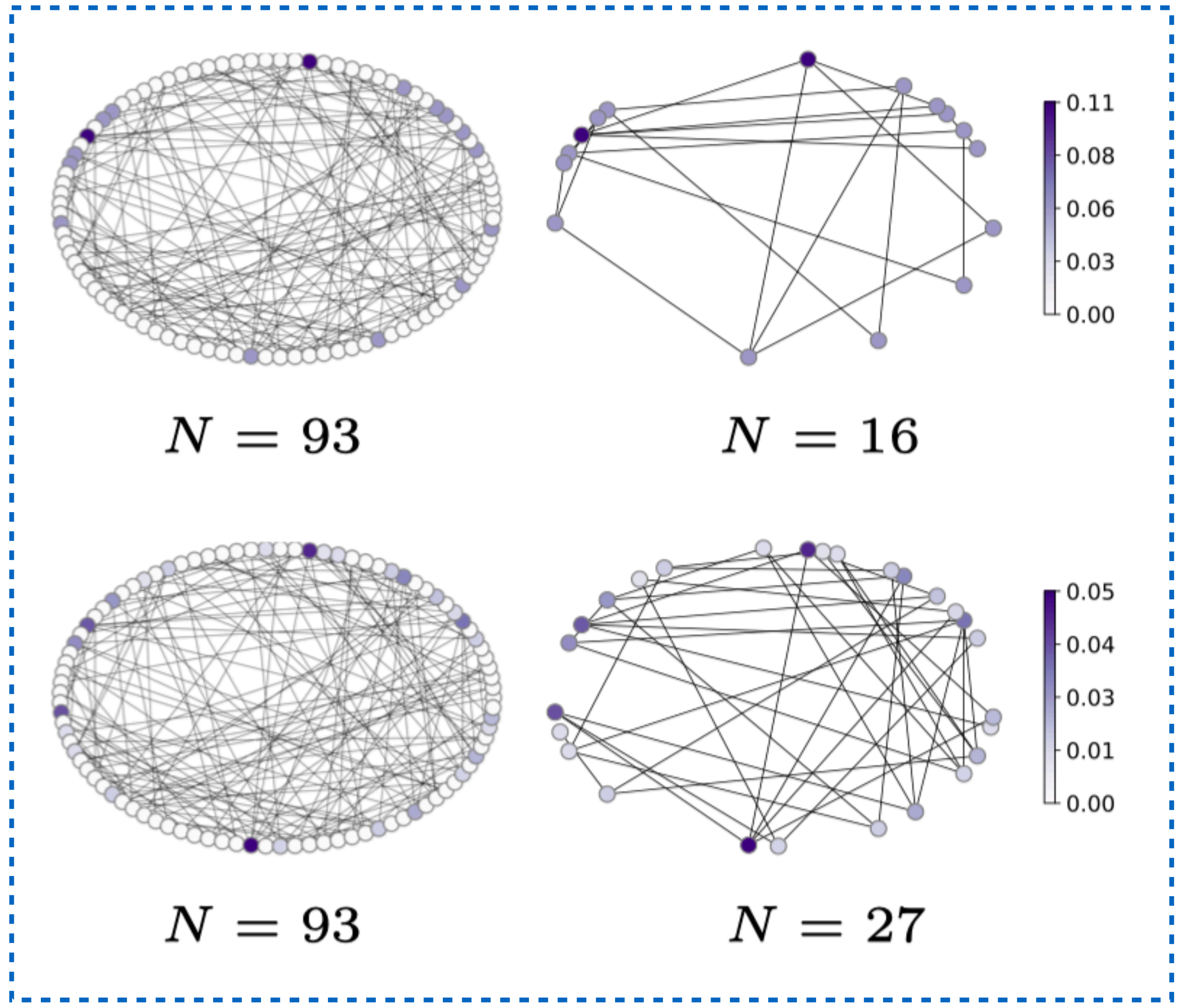
CUHK, Tencent AI Lab

with Yonggang Zhang, Yatao Bian, Han Yang, Binghui Xie, Kaili Ma, Tongliang Liu, Bo Han, and James Cheng

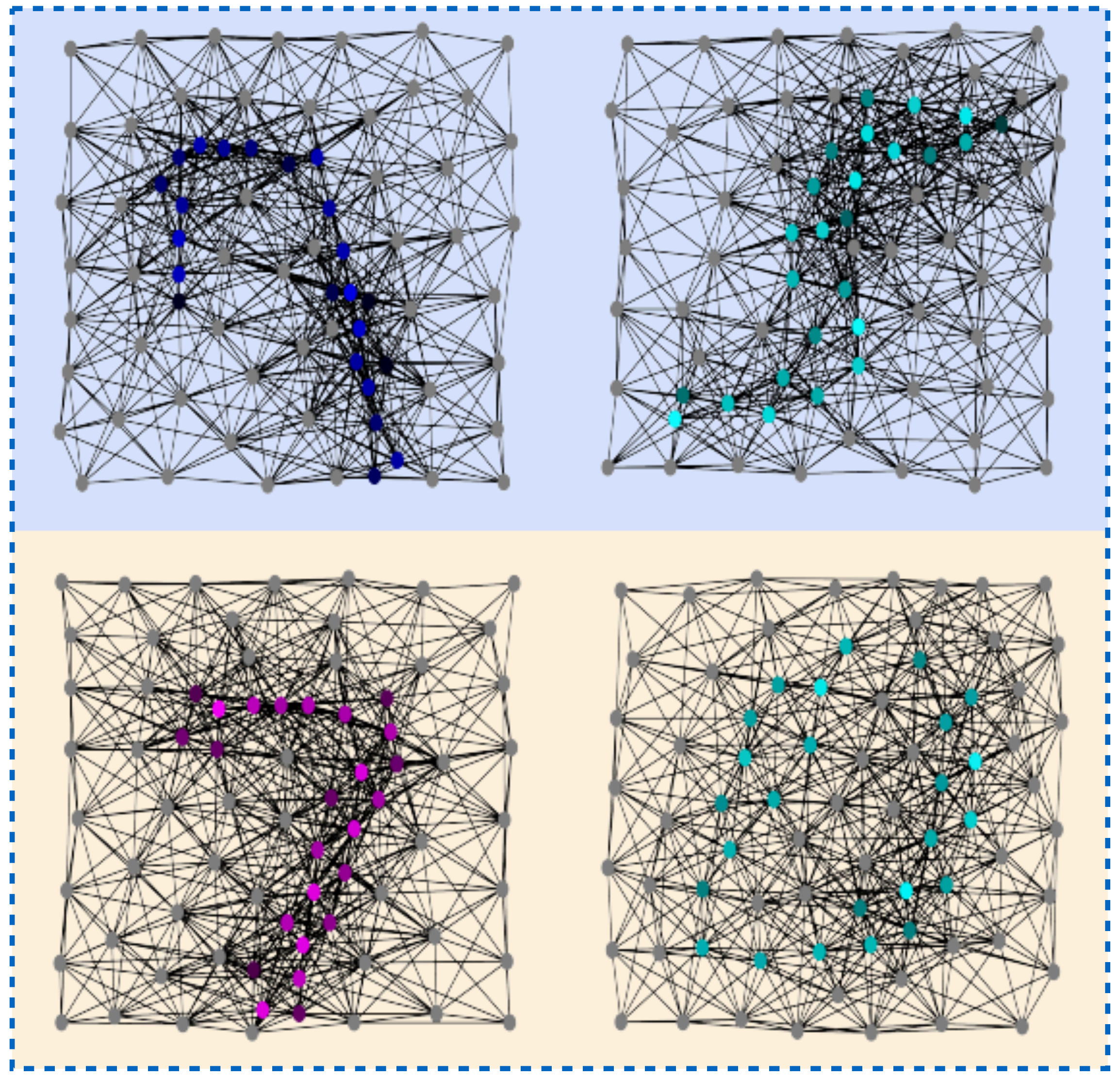
OOD generalization on graphs is more challenging

A Graph Neural Network (GNN) makes predictions taking both **structure-level** and node **attribute-level** features into account.

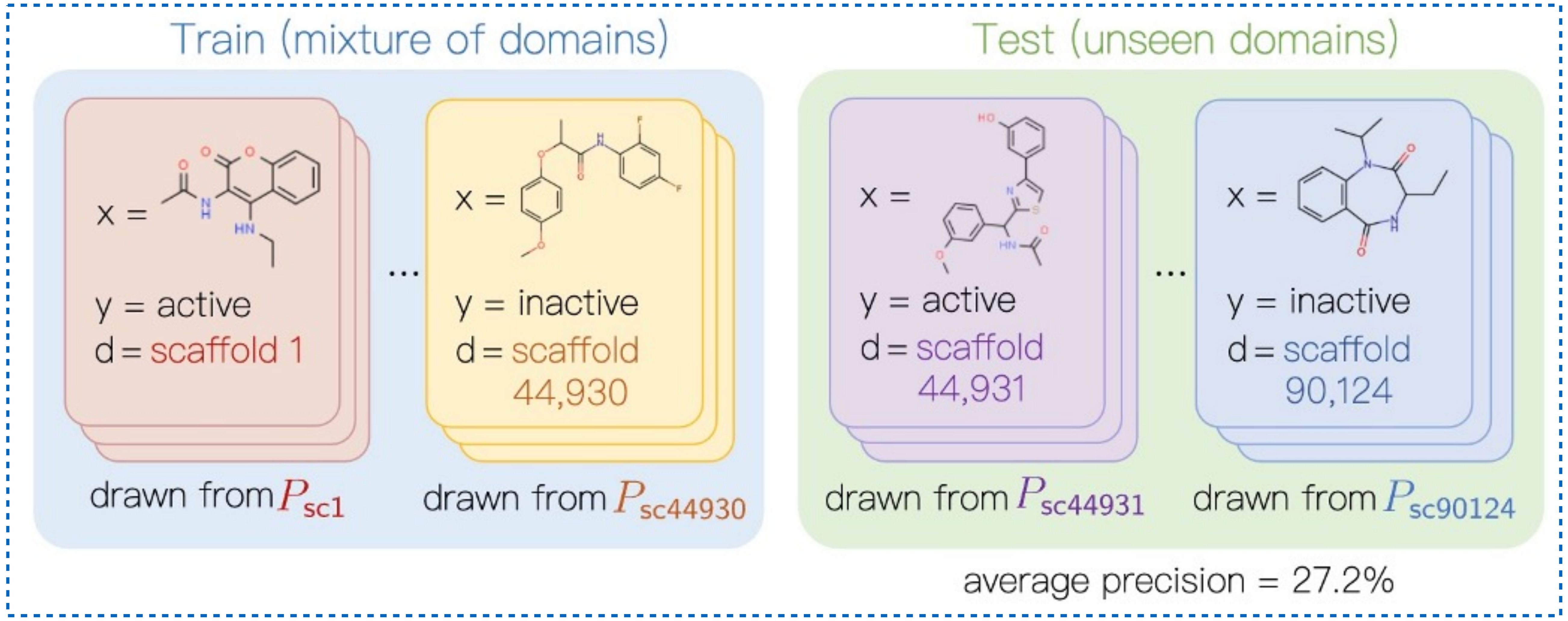
$f_{\mathrm{GNN}}($ {  } , {  } $)$ = **"House"**



Structure-level shifts



Attribute-level shifts

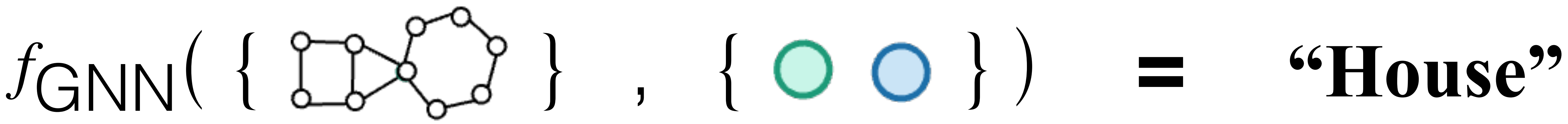


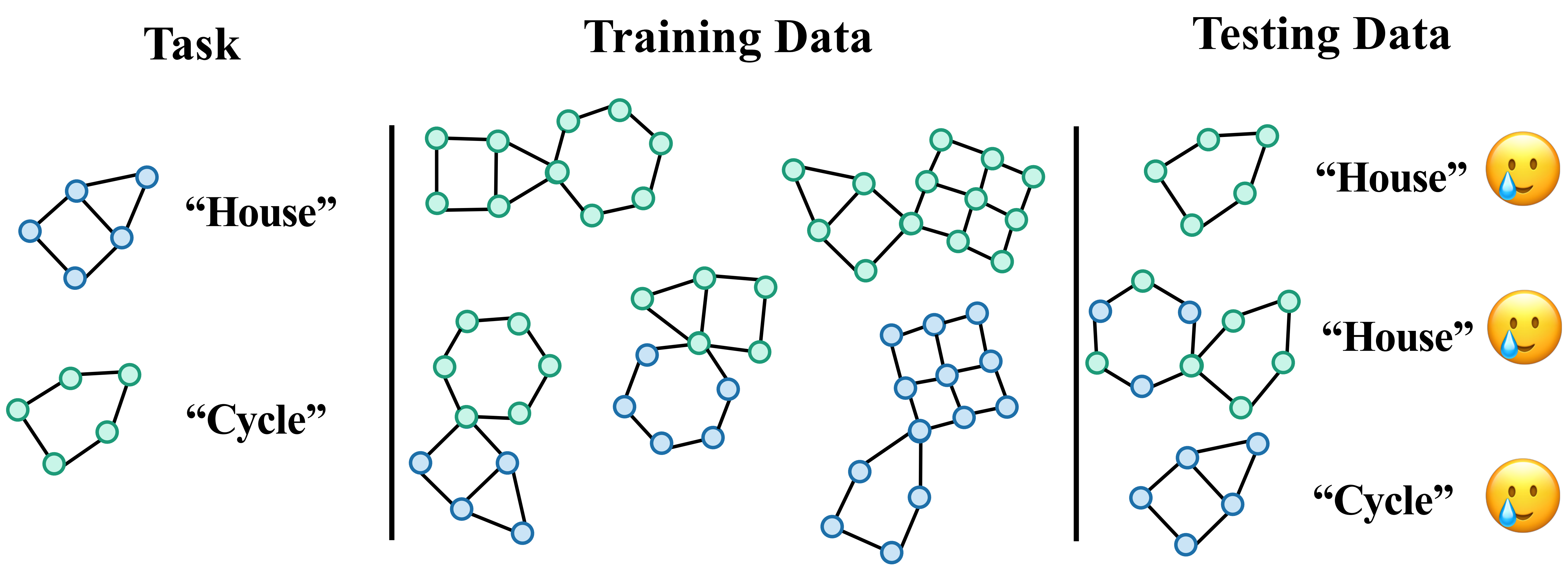
Mixture of structure-level and attribute-level shifts

(*Knyazev et al. 2019; Hu et al., 2020; Koh et al., 2021; Gui et al., 2022*)

OOD generalization on graphs is more challenging

A Graph Neural Network (GNN) makes predictions taking both **structure-level** and **attribute-level** features into account.





(Ying et al., 2019; Luo et al., 2020; Wu et al., 2022;)

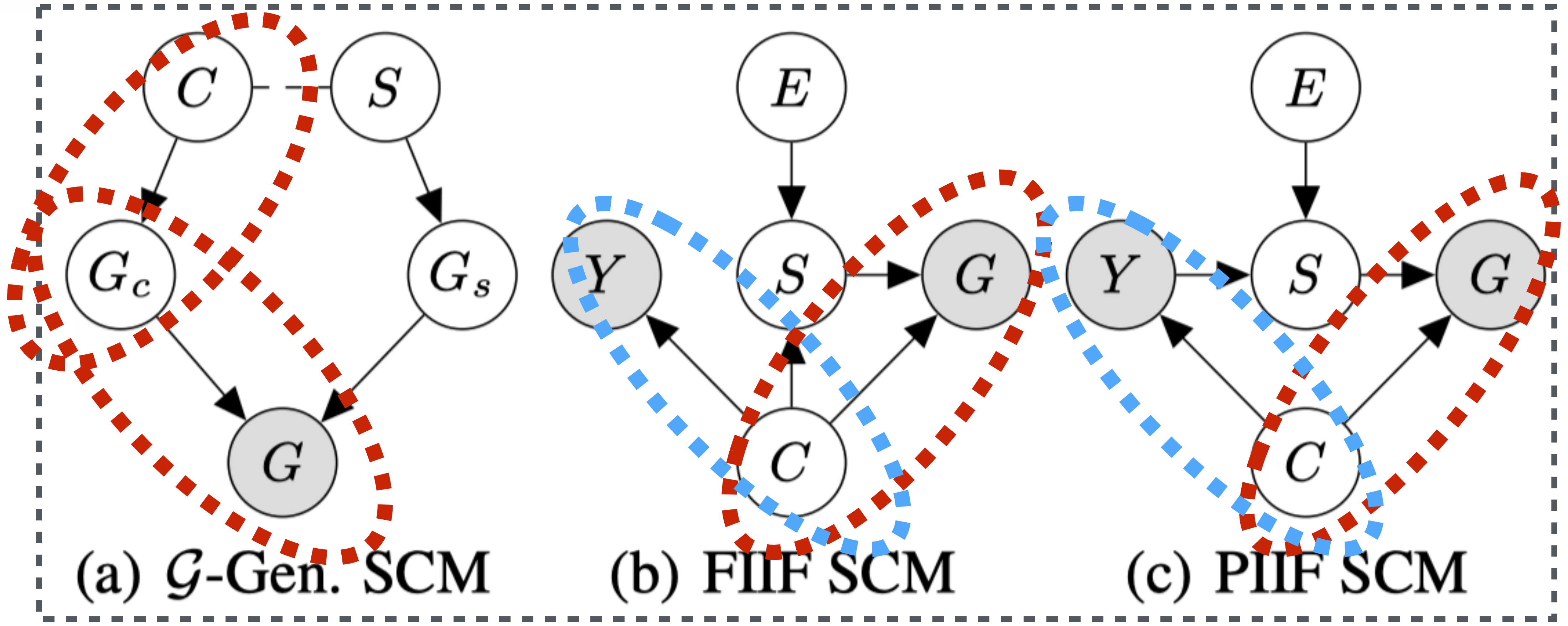
OOD generalization on graphs are **much more challenging!**

- Graphs are highly non-linear
- Attribute-level shifts
- Structure-level shifts
- Mixed shifts in different modes
- Expensive environment labels

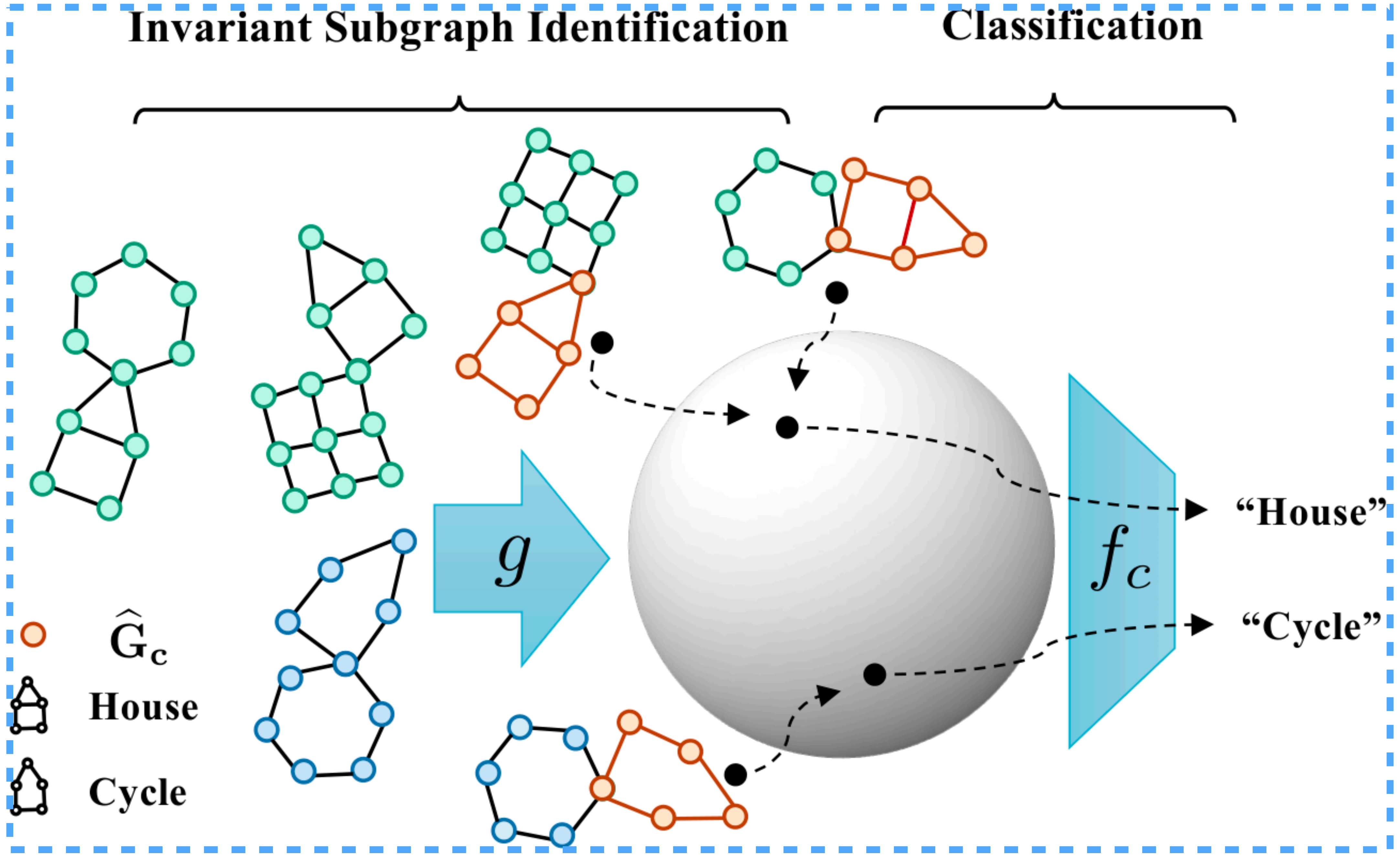
CIGA: Causality Inspired Invariant Graph LeArning

Step 1: Invariant subgraph identification

Featurizer GNN $g : \mathcal{G} \rightarrow \mathcal{G}_c$



Structural Causal Models



Step 2: Label prediction

Classifier GNN $f_c : \mathcal{G}_c \rightarrow \mathcal{Y}$

Overall objective

$$\max_{f_c,\, g} I(\widehat{G}_c; Y), \text{ s.t. } \widehat{G}_c \perp E,\ \widehat{G}_c = g(G),$$

Informative

Invariant

Summary

Through the lens of causality, we establish general SCMs to characterize the distribution shifts on graphs, and generalize the invariance principle to graphs.

We instantiate the invariance principle through a novel framework CIGA, where the prediction is decomposed into the subgraph identification and classification.

We show that the provable identification of the underlying invariant subgraph can be achieved using a contrastive strategy both theoretically and empirically.

Paper Code

Thank you!

Contact: yqchen@cse.cuhk.edu.hk