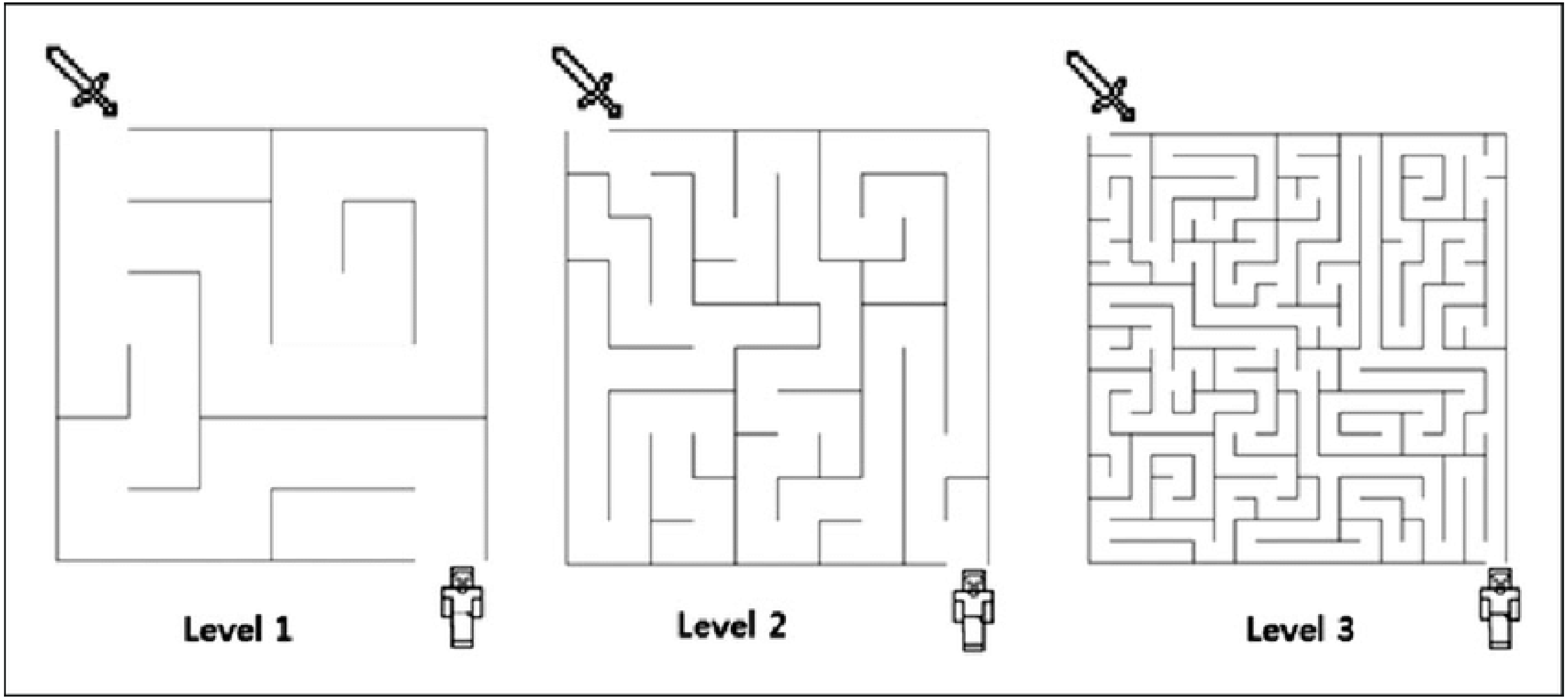
Towards Scale-Invariant Graph-related Problem Solving by Iterative Homogeneous GNNs

Hao TANG, Zhiao HUANG, Jiayuan GU, Bao-Liang LU, Hao SU

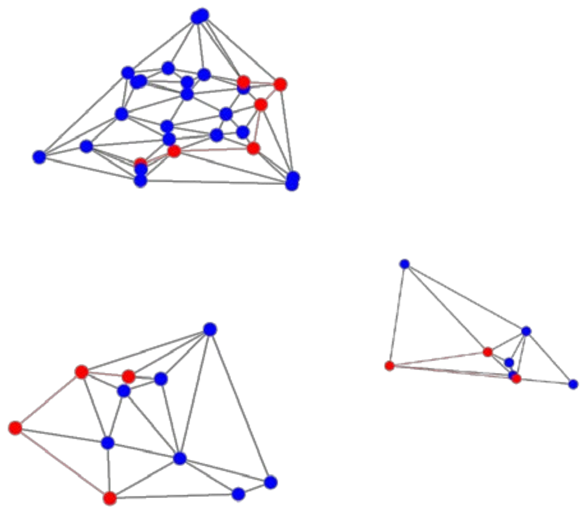
Scale Problems

人类是具有在小样本中学习知识，并应用到大样本中的能力的。

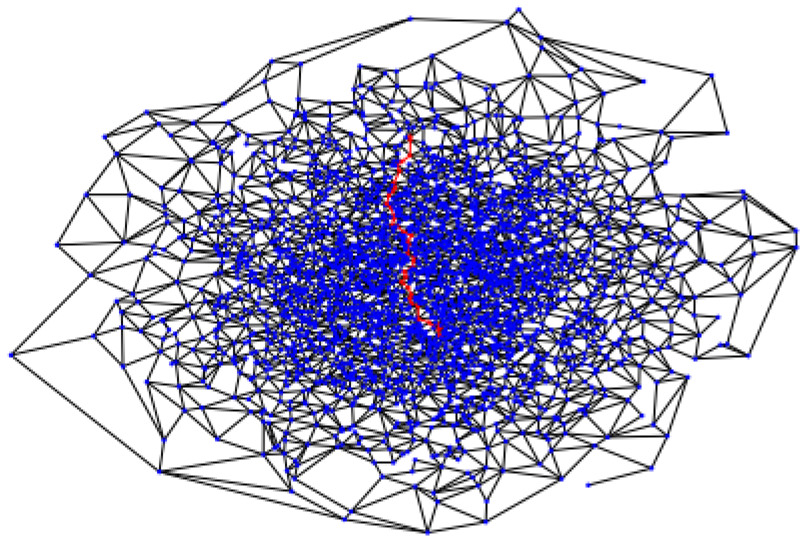


[1] Touati A, Baek Y. What Leads to Player’s Enjoyment and Achievement in a Mobile Learning Game?[J]. Journal of Educational Computing Research, 2018, 56(3): 344-368.
[2] Novaković, Predrag & Hornak, Milan & Zachar, Mgr. (2017). 3D Digital Recording of Archaeological, Architectural and Artistic Heritage. 10.4312/9789612378981.

Scale-Invariant Graph-related Problem Solving



Generalize w.r.t. Scales

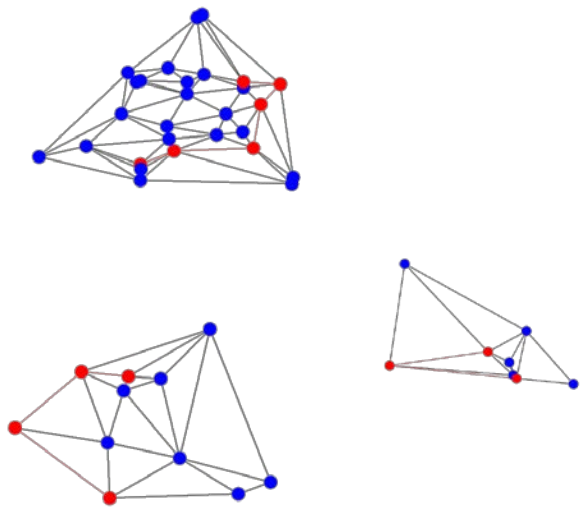


Trained on Smaller Graphs
(e.g. ≤ 40 nodes)

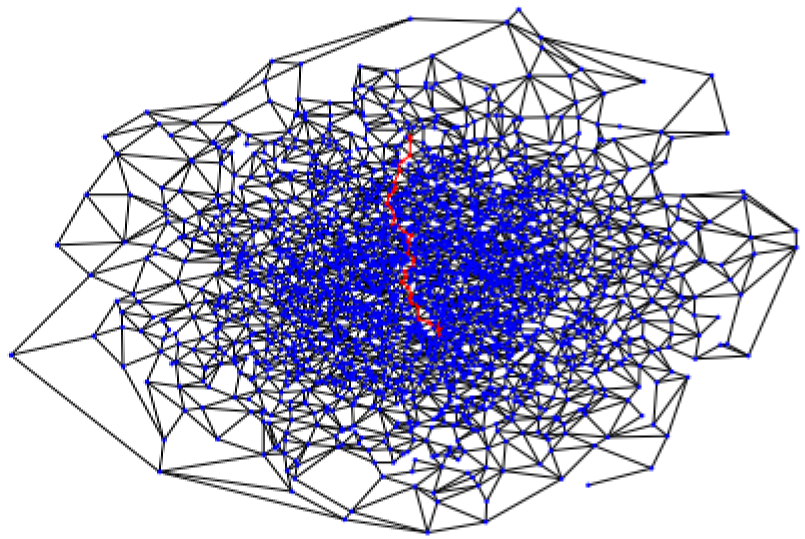
Tested on Larger Graphs
(e.g. 5000 nodes)

- 基于图强大的表征能力，很多问题可以转化为图相关问题。
- 我们想要实现在图尺度上的泛化，即在小图上训练，在大图上测试。

Scale-Invariant Graph-related Problem Solving



Generalize w.r.t. Scales



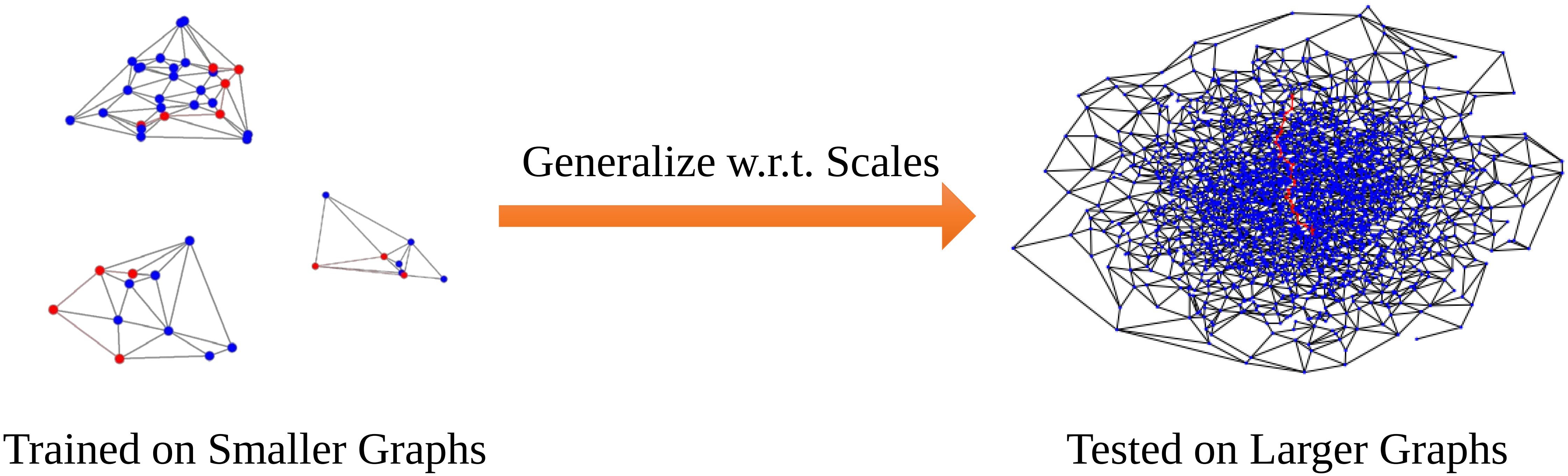
Trained on Smaller Graphs
(e.g. ≤ 40 nodes)

Tested on Larger Graphs
(e.g. 5000 nodes)

- 例如这里的最短路问题，我们实际实现了在小于40个节点的图上训练，并成功泛化到5000个节点的图上。

Problems of common GNNs

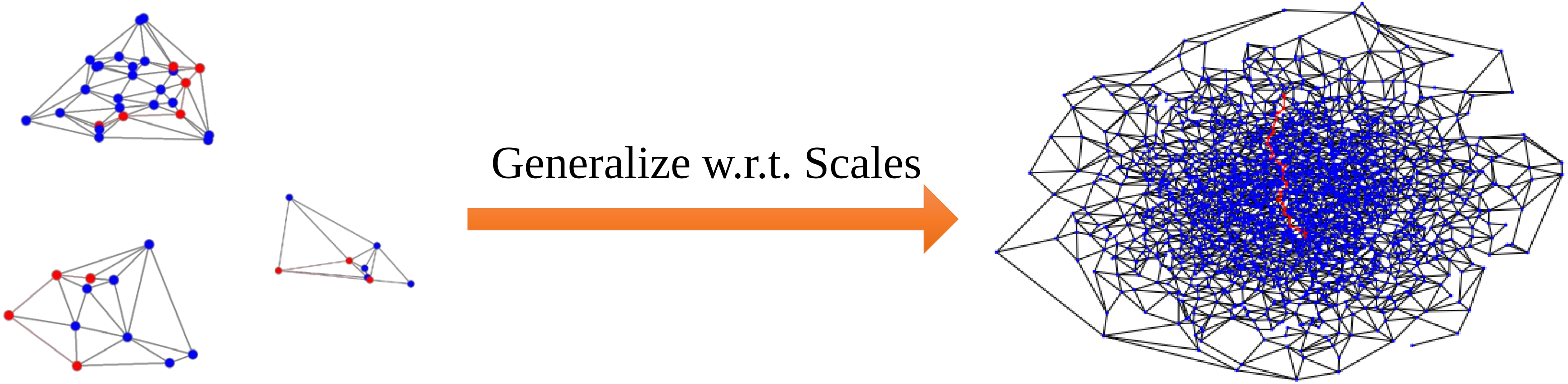
我们发现了传统图神经网络的两个问题，



- GNNs with bounded layer numbers cannot generalize to larger graphs.
- Out-of-range number encodings.

Problems and our Solutions

并提出相应的解决办法。



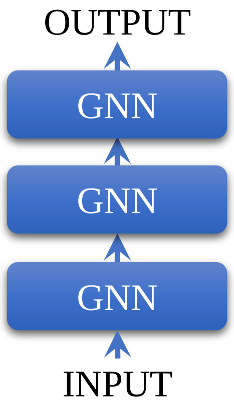
Trained on Smaller Graphs

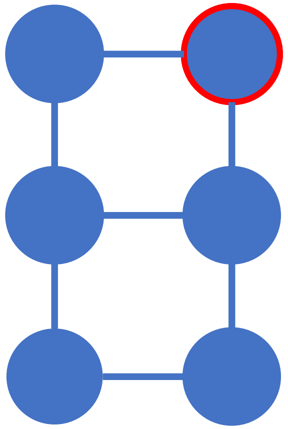
Tested on Larger Graphs

- GNNs with bounded layer numbers cannot generalize to larger graphs.
  ← **Iterative Graph Neural Networks.**
- Out-of-range number encodings. ← **Homogeneous Neural Networks**

GNNs with bounded layer numbers cannot generalize to larger graphs on many problems

Common GNNs



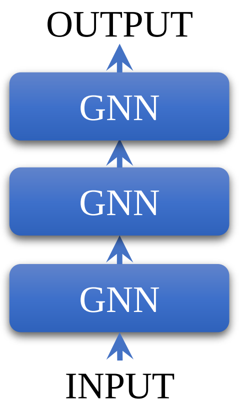


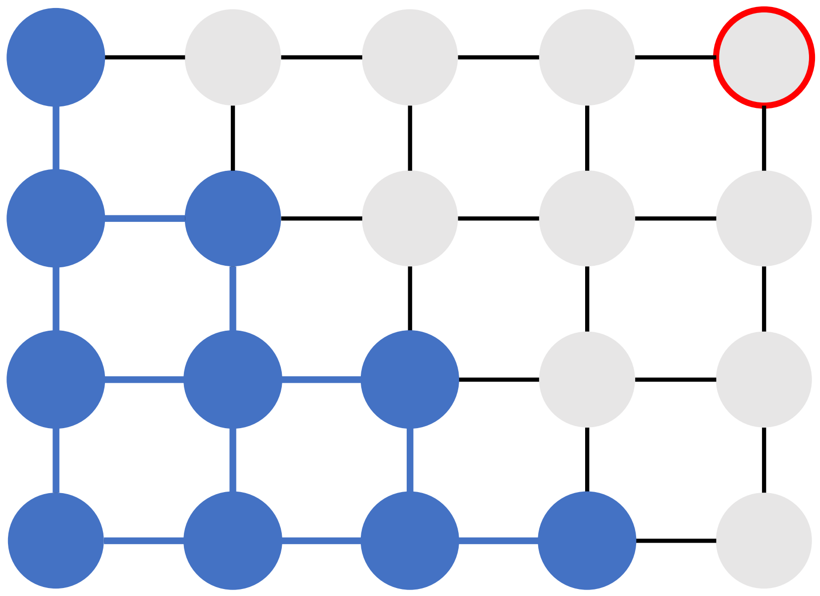
- GNNs with bounded layer numbers cannot generalize to larger graphs on many simple jobs, such as
  - send information from the bottom left node to the upper right node.
- Loukas [1] formally proves that
  - GNNs, which fall within the message passing framework, lose a significant portion of their power for solving many graph-related problems when their width and depth are restricted.

[1] Loukas, A. What graph neural networks cannot learn: depth vs width. In International Conference on Learning Representations. 2020.

GNNs with bounded layer numbers cannot generalize to larger graphs on many problems

Common GNNs



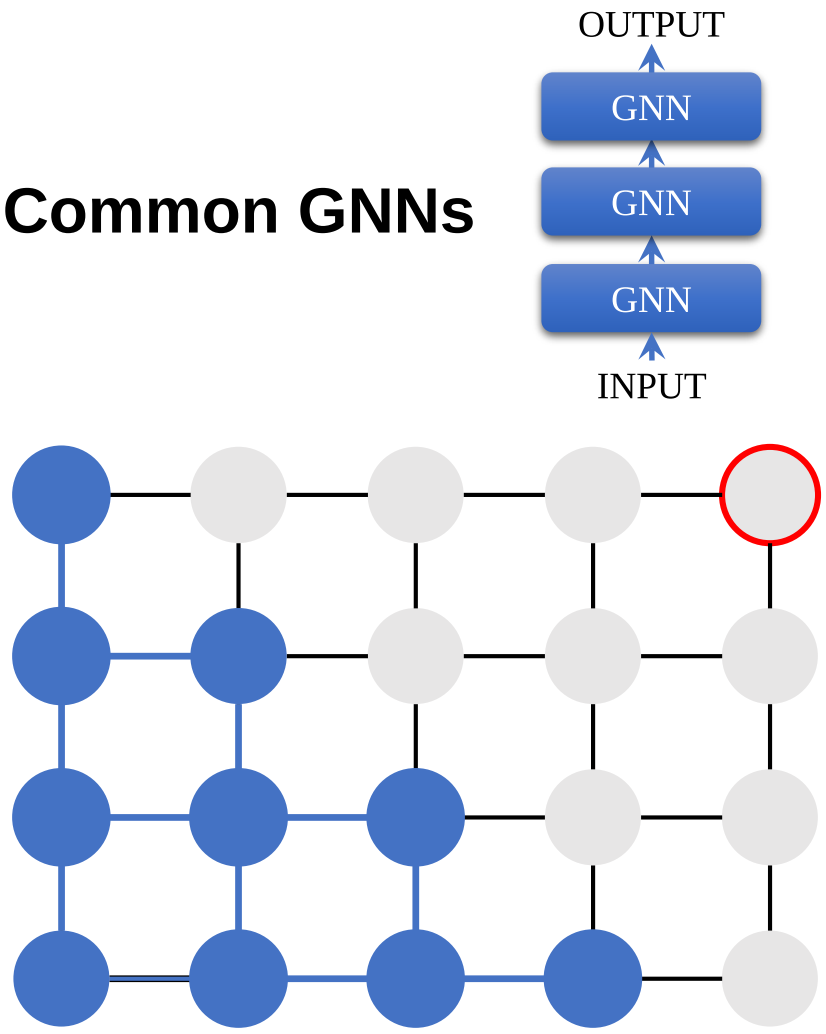


- GNNs with bounded layer numbers cannot generalize to larger graphs on many simple jobs, such as
  - send information from the bottom left node to the upper right node.
- Loukas [1] formally proves that
  - GNNs, which fall within the message passing framework, lose a significant portion of their power for solving many graph-related problems when their width and depth are restricted.

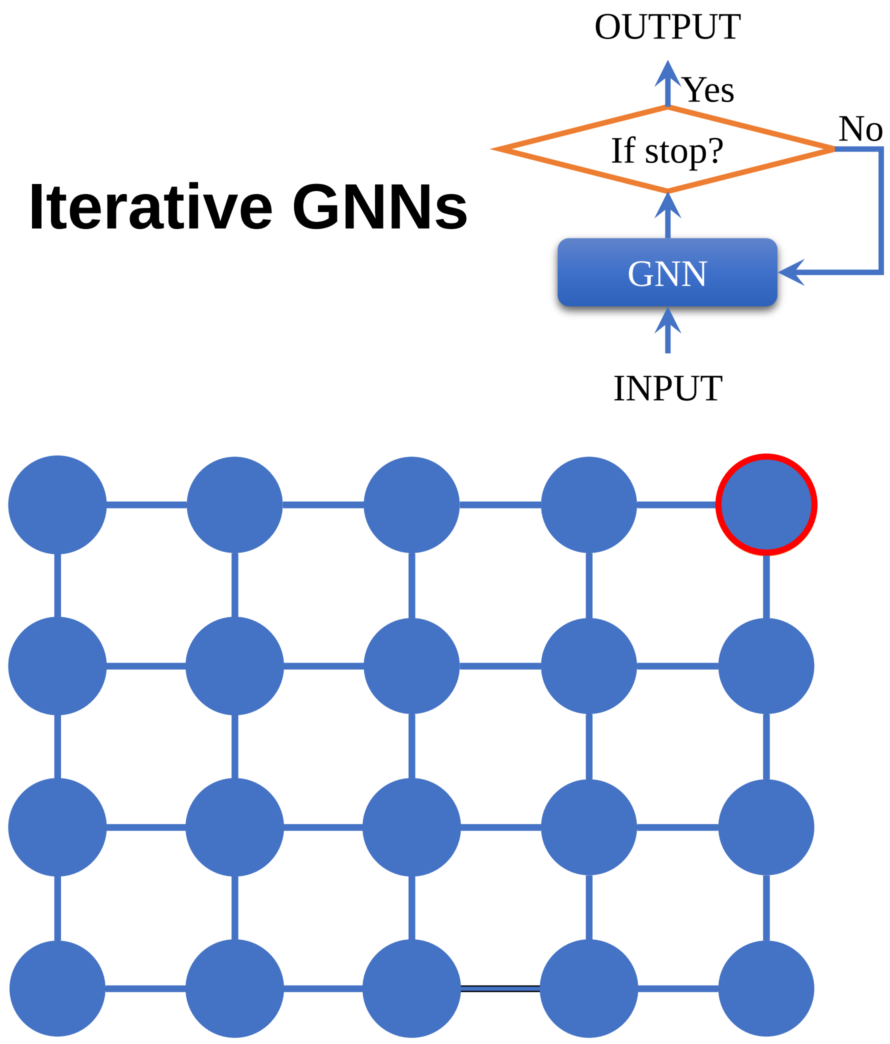
[1] Loukas, A. What graph neural networks cannot learn: depth vs width. In International Conference on Learning Representations. 2020.

Adaptive & unbounded iteration numbers by Iterative Graph Neural Networks (IterGNN)

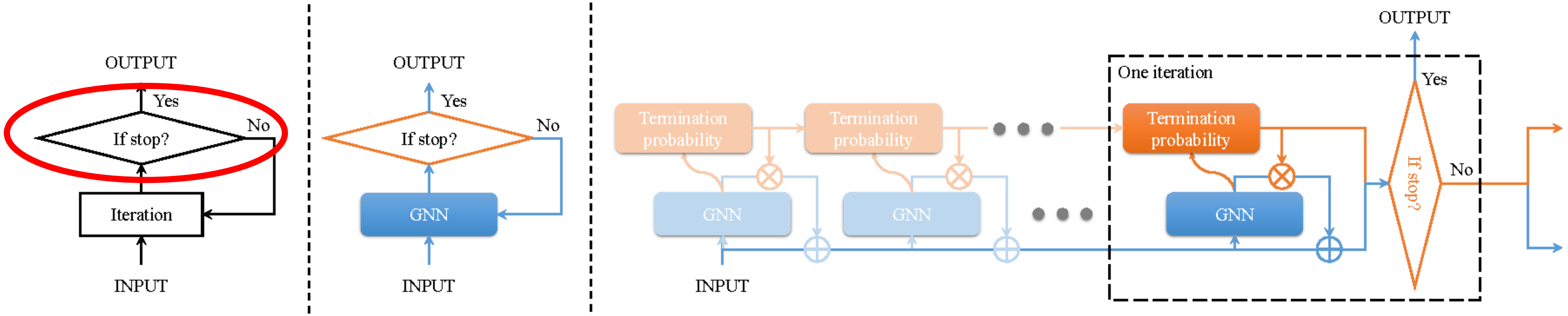
Common GNNs



Iterative GNNs



Iterative Graph Neural Networks (IterGNN)



(a) iterative algorithms

(b) general illustration of Iterative GNN

(c) concrete unfolded illustration of Iterative GNN at one iteration

Algorithm 1 Iterative module. g is the stopping criterion and f is the iteration body

input: initial feature x; stopping threshold ϵ
$k \leftarrow 1$
$h^0 \leftarrow x$
while $\prod_{i=1}^{k-1}(1-c^i) > \epsilon$ **do**
  $h^k \leftarrow f(h^{k-1})$
  $c^k \leftarrow g(h^k)$
  $k \leftarrow k+1$
end while
return $h = \sum_{k=1}^{\infty} c^k h^k \prod_{i=1}^{k-1}(1-c^i)$

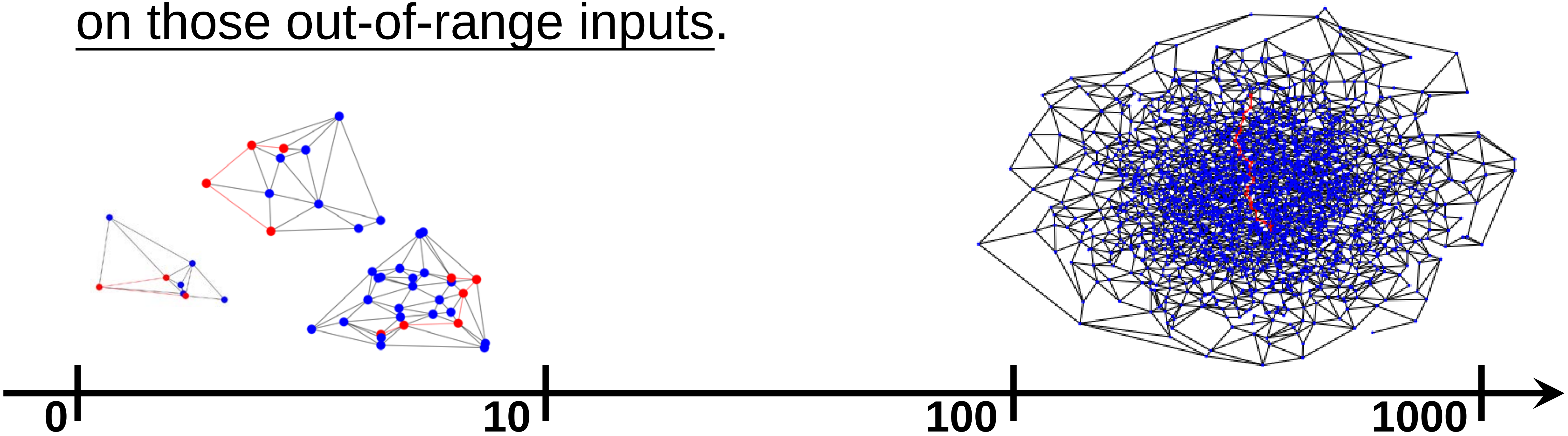
$$h^k = f(h^{k-1}) \in R^{N\times d}, k = 1,2,\ldots$$

$$c^k = g(h^k) \in [0,1), k = 1,2,\ldots$$

$$h = \sum_{k=1}^{\infty} \prod_{i=1}^{k-1}(1-c^i) c^k h^k$$

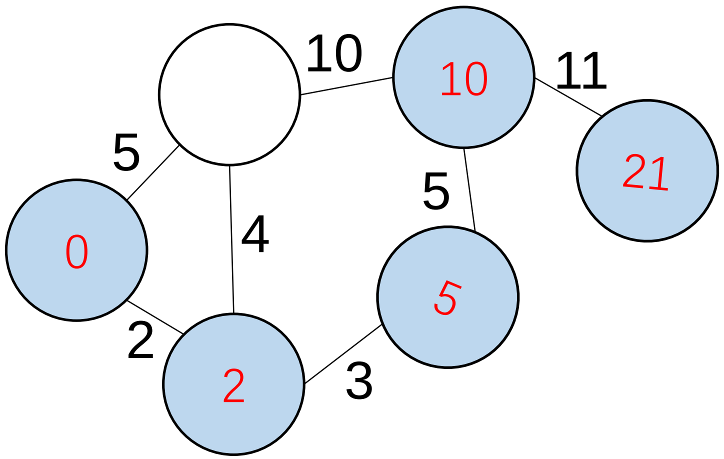
Out-of-range number encodings

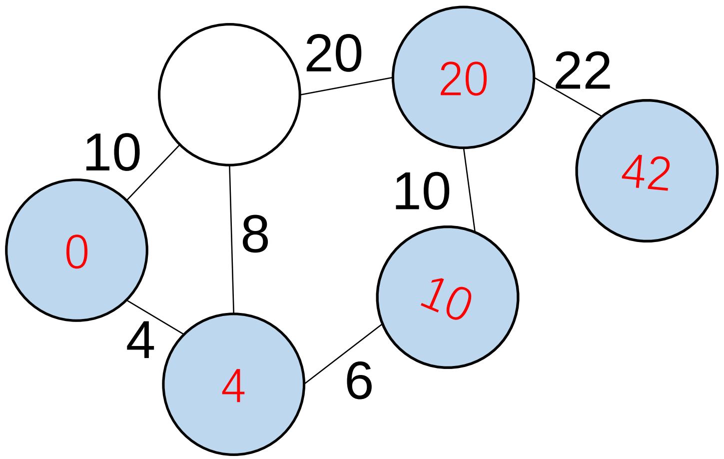
- The range of numbers to be encoded by the internal representation may deviate greatly for graphs of different scales.
- The performance of MLPs in GNNs are usually highly degraded on those out-of-range inputs.

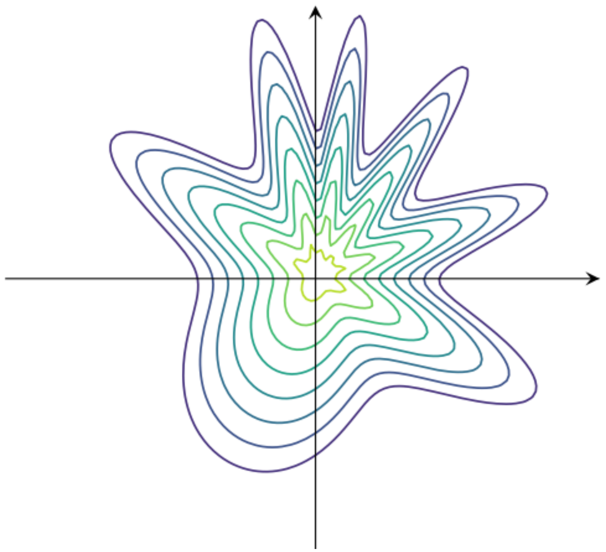


Graph properties to be encoded by internal representations (e.g. distances)

Solutions to many graph-related problems are homogeneous functions.





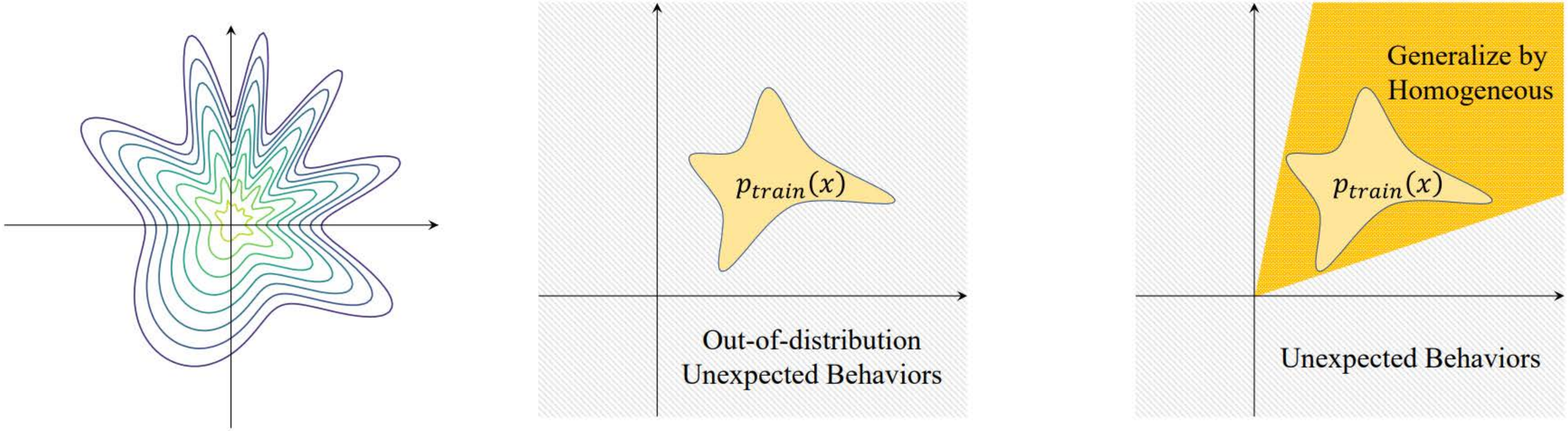


e.g. Solutions to shortest path are homogeneous.
(i.e. if we multiplies the edge weights by 2, the shortest path lengths are also multiplied by 2.)

$f(\lambda x) \equiv \lambda f(x)$,
$\forall \lambda > 0, x \in R^2$
(another example of homogeneous functions)

Homogeneous Neural Networks

- Generalization errors are bounded if both neural networks and the target function are homogeneous under proper conditions.
- Universal approximators of homogeneous functions are easy to build, e.g. $Relu(Wx + b) \Rightarrow Relu(Wx)$ in MLPs.

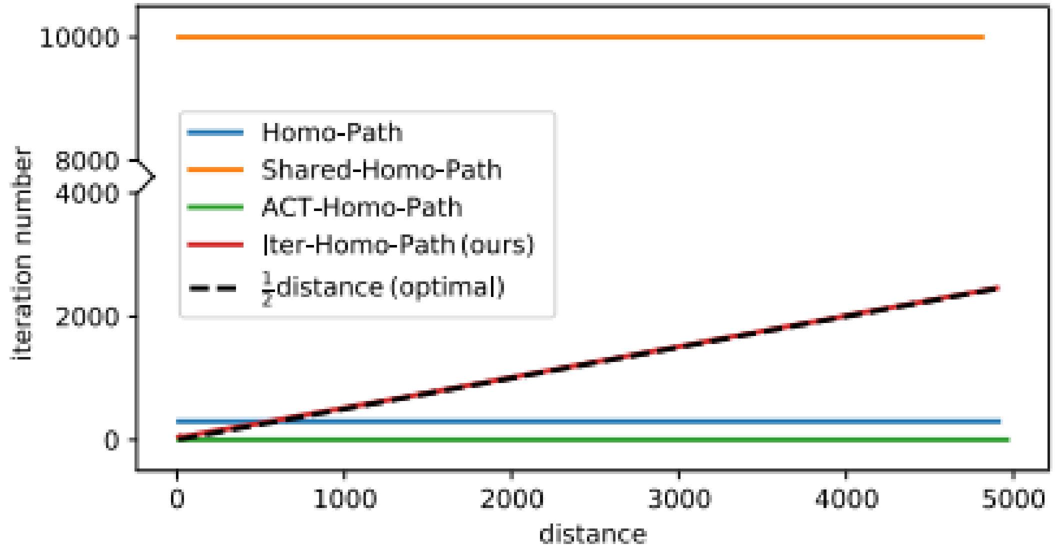


(a) $f(\lambda x) = \lambda f(x),\ \forall \lambda > 0, x \in R^2$ (b) Behaviors of MLP (c) Behaviors of Homogeneous MLP

Experimental Results

- Our model successfully generalized to much larger graphs on 3 graph-theory problems and 3 general reasoning tasks.
- Ablation studies to show the benefits of each component.
- Interpretable behaviors learnt by our iterative module.

Iter-Homo-Path	
100.0	
Homo-Path	Iter-Path
53.7	48.9
ACT-Homo-Path	Iter-Homo-GAT
52.7	2.9
Shared-Homo-Path	Iter-Homo-GCN
91.7	1.4



THANK YOU