Sri Lanka Institute of Information Technology

IT3071- Machine Learning and Optimization Methods

Lecturer in Charge: Mr. Samadhi Chathuranga Rathnayake

Message Passing Neural Networks

Selected problem: Bridging the Efficiency

Identified unsolved research gap : Accuracy Trade-off in Dynamic (MPNNs) through Sparse Pseudo-node Message Passing

Student Registration Number	Student Name
IT23246096	D.D Devanmith
IT23426658	Kalpitha W.K.V
IT23269866	K.A.U Gimhani
IT23377844	Dinethra M.M.T

Table of Content

1.Introduction and Background

Graph-structured data play a crucial role in modern computing, as they effectively represent complex relationships among entities such as users, molecules, or sensors. Applications in areas like social networks, molecular chemistry, and recommendation systems rely heavily on learning from such structured data. To handle these relationships, **Graph Neural Networks (GNNs)** have emerged as a powerful learning framework capable of propagating and aggregating information across connected nodes.

Within this broader class, **Message Passing Neural Networks (MPNNs)** provide a unified approach where node representations are updated by exchanging messages with their neighbours. This iterative process enables models to capture both structural and feature-based dependencies within a graph. Despite their success on small or moderately sized graphs, MPNNs face significant challenges in handling **large-scale or dynamic graphs**, where both node features and edges evolve over time.

A key issue lies in the **trade-off between accuracy and efficiency**. As networks grow deeper to capture long-range dependencies, computational cost increases drastically, and phenomena such as **over-squashing** and **over-smoothing** emerge—limiting the model's ability to preserve meaningful information. The study **"Towards Dynamic Message Passing on Graphs (N^2)" (Sun et al., 2024)** highlights this limitation, emphasizing that dynamic models often struggle to balance adaptability with efficiency.

While dynamic MPNNs attempt to update their connections during training to improve flexibility, they typically require dense connectivity or global updates that remain computationally heavy. This makes real-time or large-scale deployment difficult. Hence, there is an open research need to design **lightweight and adaptive message-passing mechanisms** that maintain high accuracy while minimizing redundant communication. Addressing this challenge provides the foundation for the present research direction — **bridging the efficiency–accuracy trade-off in dynamic MPNNs through sparse pseudo-node message passing.**

2. Identification of Research problems

Research problem/gap :

Improving the Efficiency of Message Passing Neural Networks (MPNNs) : Accuracy Trade-off in Dynamic MPNNs through Sparse Pseudo-node Message Passing.

Long-range data transfer and accuracy of models have been improved by recent developments in Message Passing Neural Networks (MPNNs), which include dynamic rewiring, virtual nodes, and hierarchy representations. For vast and dynamic graphs, these methods' scalability and real-time usability are constrained by the rapid introduction of new efficiency bottlenecks.

DRew: Dynamically Rewired Message Passing (Di Giovanni et al., 2023), Probabilistic Graph Rewiring via Virtual Nodes (Qian et al., 2024), and Efficient and Effective Implicit Dynamic GNN (Zhong et al., 2024) are a few important studies that address efficiency in MPNNs, according to a review of recent literature. For better message flow and reduce overload, these works use dynamic rewiring, virtual nodes, or pseudo-nodes. Most still use dense message-passing structures which result in significant memory and computational overhead, despite their efficiency.

Sun et al.'s paper "Towards Dynamic Message Passing on Graphs (N^2)" (NeurIPS 2024) received special attention since it applies adaptive paths and pseudo-nodes to enhance connectivity between distant graph sections. The model clearly shows a trade-off between accuracy and efficiency, even with its success. Each pseudo-node is connected to several graph regions by the N2 framework, which significantly enhances information flow but results in nearly quadruple message volume growth. These intricate linkages result in duplicate calculations and significant memory usage. Additionally, all connections are seen as equally crucial even when many of them have minimal impact on prediction accuracy since pseudo-node interconnections are equally dense rather than adaptive.

A clear research gap was identified as the outcome of this analysis: Accuracy Trade-off in Sparse Pseudo-node Message Passing in Dynamic MPNNs.

Thus, the creation of a Sparse Pseudo-node Message Passing (SPMP) mechanism is the focus of this work.

reduces over-squashing by maintaining long-range information flow. Uses controlled sparsity (such as top-k focused selection) to reduce redundant computation and effectively adjusts to changes in the graph instead of requiring a complete recalculation. By finding an approach to this challenge, MPNNs will be able to balance accuracy and efficiency, which will make them more useful for massive, real-time graph-learning applications.

3. Motivation and Justification

3.1 Overview

The increasing need for graph learning methods capable of efficiently handling large, evolving datasets motivates this work. Dynamic graphs, which continually change as nodes and edges are added or removed, are fundamental to applications such as traffic forecasting, social network analysis, financial fraud detection, and recommendation systems. Standard message passing neural networks (MPNNs) are ill-suited for these environments because any change in the graph requires a full recompilation of the message-passing steps, making real-time operation impractical.

Recent approaches, like IDGNN (Zhong et al., 2024), improve computational efficiency through implicit programming techniques rather than exploiting structural sparsity. Similarly, DRew (Di Giovanni et al., 2023) modifies edge connections dynamically but still depends on dense message-passing, limiting scalability.

An alternative solution involves introducing pseudo-nodes to facilitate long-distance information propagation without traversing every node in the graph. Existing designs, such as N2 (Sun et al., 2024), connect pseudo-nodes densely across the network, resulting in excessive message passing and increased computational cost. This limitation motivates a sparse approach, where only the most relevant connections are maintained.

Controlling connectivity in this manner allows the network to focus on critical relationships while ignoring less significant ones. As noted by Hwang et al. (2022), excessive virtual node connections can lead to over-smoothing and negatively impact model performance. By applying sparsity guided by attention or relevance scores, it is possible to maintain predictive accuracy while reducing computational overhead.

In this study, we propose Sparse Pseudo-node Message Passing (SPMP), a framework designed to:

- Limit computation by retaining only the top k most informative connections.
- Adapt efficiently to dynamic changes in the graph without recomputing all messages.
- Employ attention-based selection mechanisms to preserve model accuracy.By balancing efficiency with prediction quality, SPMP offers both theoretical insights and practical benefits for deploying dynamic MPNNs on large, real-time graphs.

4. Literature Review

4.1 Research paper 1:
"Probabilistic Graph Rewiring via Virtual Nodes" *(Qian et al., 2024)*

1)Background

Graph Neural Networks (GNNs) pass information between connected nodes to learn graph patterns. However, they often suffer from **over-squashing**, where data from distant nodes becomes compressed, weakening the model's ability to capture global relationships. This limits accuracy and increases computational cost in large or changing graphs, making the balance between efficiency and accuracy a major challenge.

2) Research focus

Qian et al. (2024) focus on **structural limitations in message passing** that cause over-squashing and propose a method called **Probabilistic Graph Rewiring via Virtual Nodes**. Their approach introduces **virtual (pseudo) nodes** that act as intermediate hubs, facilitating long-range communication between otherwise distant nodes. Unlike full rewiring, which adds many edges and increases computational cost, their method uses **probabilistic rewiring** to selectively connect nodes through these virtual intermediates. This reduces message bottlenecks, improves information flow, and helps maintain global dependencies without excessively increasing the graph's density or computational burden.

3) Key Findings

- Over-squashing arises when vast neighborhoods are compressed into fixed-size node representations.
- Graph bottlenecks limit the ability of messages to travel across remote subgraphs.
- The proposed **probabilistic rewiring** mechanism creates virtual connections that enhance message propagation while keeping the graph structure manageable.
- **Virtual nodes** effectively serve as shortcuts for long-range dependencies, improving performance in large graphs without requiring deep or computationally expensive GNN layers.

4) Relevance to our research problem

- **Targets the bottleneck issue:** The paper's use of virtual nodes demonstrates how adding selective connections can overcome over-squashing — a key idea your research extends through **sparse pseudo-node message passing**.
- **Supports efficiency–accuracy balance:** While Qian et al. improved propagation through probabilistic rewiring, they did not fully explore the **trade-off between model accuracy and computational cost**. Your research builds on this by explicitly modeling and optimizing that trade-off.
- **Dynamic adaptability:** Their static rewiring method doesn't adapt to changing graph structures, leaving space for your work to propose **dynamic sparse pseudo-node integration** that scales with evolving graphs.

5) Indirect Support

- **Theoretical validation:** Their findings confirm that adding pseudo-nodes effectively mitigates over-squashing, providing theoretical backing for your approach.
- **Justification for sparsity:** While their method probabilistically rewires, your work goes further by **sparsely selecting pseudo-node connections**, ensuring reduced computation while maintaining message fidelity.
- **Scalability insight:** The paper recognizes over-squashing as both an **efficiency and scalability issue**, reinforcing your focus on **dynamic and real-time MPNNs**, where the cost of dense rewiring would be prohibitive.

6) Summary

Qian et al. (2024) show that virtual nodes and probabilistic rewiring reduce over-squashing by improving information flow across graphs. Yet, their method remains static and doesn't balance accuracy with computational cost. Your research advances this idea by using **sparse and adaptive pseudo-nodes**, achieving efficient message passing while maintaining accuracy in dynamic graph environments.

Link: http://arxiv.org/pdf/2405.17311

4.2 Research paper 2:

“On the Bottleneck of Graph Neural Networks” *(Alon & Yahav, 2020)*

1) Background

Graph Neural Networks (GNNs) use message transmission for gathering data from nodes that are nearby. However, over-squashing, an occurrence where data obtained from distant nodes has been reduced into limited-size connections, leads them to struggle with long-range relationships and lose important global context. GNN performance is hampered by this, particularly in big or rapidly changing graphs.

2) Research focus

In this study, structural problems in GNNs that result in over-squashing—the compression and loss of data from distant nodes during message passing—are examined. It illustrates how the diagram's topology—such as the nodes that link several remote subgraphs—directly influences how well messages spread. Standard message-passing layers are unable to record long-range relationships in graphs with such bottlenecks because the fixed-size node embeddings are unable to process all incoming information. This explains why traditional GNNs perform poorly on tasks requiring broad context since their algorithms are ineffective at collecting data from remote regions of the graph.

3) Key Findings

- Compressing messages from vast neighbourhoods into fixed-size node representations leads to over-squashing.
- Information loss is made worse by graph bottlenecks, such as nodes that link distant subgraphs.
- Signals cannot be propagated over these bottlenecks efficiently by standard GNNs.
- By offering shortcuts for message flow, approaches like pseudo-nodes or network remodelling help reduce over-squashing.

4) Relevance to our research problem

- Targets the bottleneck: By enhancing graph connectedness via sparse pseudo-node message passing, long-range dependencies can be maintained without the need for expensive, deep message-passing layers.
- Justifies the precision of pseudo-nodes: Pseudo-nodes mitigate over-squashing and ensure message fidelity by preserving data collected by remote nodes.
- Trade-off between accuracy and efficiency: Introducing pseudo-nodes improves information flow but adds computation, highlighting the necessity of striking the right balance among the two in dynamic MPNNs.

5) Indirect Support

- The work provides a solid theoretical justification for the introduction of pseudo-nodes by describing why long-range dependencies in traditional GNNs fail. This reinforces the reasons for your research decisions.
- Separation maintains accuracy and efficiency: It validates your argument that sparse pseudo-node interconnections can accomplish the same advantages with fewer processes by demonstrating that deliberately adding links (as opposed to completely connecting nodes) mitigates over-squashing.
- For graph learning to be scaled and real-time, bottlenecks are essential:
  The findings highlight that dealing with over-squashing is a scalability and efficiency problem as well as a quality issue, which is crucial for your work on dynamic MPNNs because the graph topology changes over time.

6) Summary

Bottlenecks in GNNs that result in over-squashing, when information about faraway networks is lost during communication transit, are identified by Alon & Yahav (2020). They demonstrate how communication is affected by node topology and how regular GNNs have difficulty with long-range connections. This can be mitigated by methods like pseudo-nodes or rewiring, which strike a balance between efficiency and accuracy.

Link: https://arxiv.org/abs/2006.05205

4.3 Research paper 3:

"DRew: Dynamically Rewired Message Passing with Delay" (*Di Giovanni et al., 2023*)

1) Background

Important long-range dependencies are lost when information from distant nodes in a graph is compressed over numerous hops into a fixed-size representation, a phenomenon referred to as over-squashing that affects message-passing neural network models (MPNNs). Furthermore, the inductive bias of graph distance—that is, the belief that nodes that are closer would normally interact earlier—is frequently abandoned in the interest of making the graph better connected, which is the goal of many current rewiring techniques. This is highlighted in the research as a disadvantage of earlier rewiring techniques.

2) Research focus

To combat over-squashing, the study suggests a system known as DRew ("Dynamically Rewired Message-Passing with Delay"), which does so by:

To prevent over-squashing, the article proposes DRew (Dynamically Rewired Message Passing using Delay) which uses layer-dependent rewiring and delay mechanism to progressively enable connections to skip based on distance. This method enhances efficiency on long-range jobs by preserving the inductive leaning of graph structure and may be used with any MPNN.

3) Key Findings

DRew performs multi-hop MPNNs and graph Transformers in a number of long-range contact tasks.

- Over-squashing is lessened by the waiting mechanism and gradual densification (rewiring over layers) than by static rewiring.
- Research on Machine Learning Proceedings According to the paper's theoretical analysis, the suggested structure may reduce the information bottlenecks present in conventional message-passing channels.
- They demonstrate that it is better to maintain the inductive distance bias (i.e., closer nodes interact faster) as opposed to just instantly joining distant nodes.

4) Relevance to our research problem

The DRew study offers a strong theoretical foundation for your research by solving a similar basic problem of long-range dependency failure brought on by over-squashing. Its delay-assisted layer-dependent rewiring supports your objective of maintaining accuracy while managing computational price and sparsity. Your sparse pseudo-node method, which selectively links nodes for effective long-range communication, is like Drew's controlled rewiring. Similar to how your dynamic pseudo-nodes change over time, the paper's dynamic rewiring increases accuracy without having a lot of work, which further represents the accuracy-efficiency trade-off you examine.

5) Indirect Support

- Provides theoretical support showing how rewiring and delay reduce over-squashing, backing the effectiveness of sparse pseudo-node methods.
- Demonstrates that structured (not full) connectivity preserves inductive bias and accuracy, supporting scarification with pseudo-nodes.
- Encourages dynamic, layer-wise connectivity — aligning with your dynamic MPNN framework.
- Emphasizes long-range dependency handling, reinforcing your focus on improving information flow through pseudo-nodes.

6) Summary

By combining layer-wise rewiring and delayed connection skips, Drew offers an efficient way for solving over-squashing in MPNNs while maintaining distance inductive bias. This allows distant nodes to communicate, but not instantly within each layer. On long-range graph tasks, tests indicate this method works more efficiently than conventional graph Processors and multi-hop MPNNs, and research backs up the bottleneck reduction. This study provides methodological inspiration (dynamic rewiring/delay) and conceptual justification (addressing over-squashing) for your research on sparse pseudo-node message passing.

Link : https://arXiv.org/abs/2305.08018

4.4 Research paper 4:

"Efficient and Effective Implicit Dynamic Graph Neural Network (IDGNN)." *(Yongjian Zhong, Hieu Vu, Tianbao Yang, Bijaya Adhikari – 2024)*

1) Background

Implicit Graph Neural Networks (GNNs), which consider the node attachments as fixed-points of a transmission process, have become effective methods for identifying long-range dependencies in static graphs. However, there are more challenges when applied to dynamic networks (where nodes and edges change over time) because features are aggregated over time and also across graph neighbourhoods, which exacerbates problems like long-range dependence breakdown and over-smoothing. Based to the IDGNN article, no implicit GNN model had ever been specifically developed for dynamic graphs.

2)Research Focus

To efficiently imitate changing graph structures, the paper presents a unique architecture called IDGNN (Implicit Dynamic Graph Neural Network). By providing convergence to a stable fixed-point formulation even in dynamic situations, it guarantees well-posedness. IDGNN avoids the high computational costs of conventional iteration implicit modelling by enabling a single-loop training procedure via a bilevel optimisation framework. The findings show that IDGNN successfully captures long-range dependencies across time in dynamic graphs, resulting in improved efficiency and accuracy in both classification and regression tasks.

3) Key Findings

- On real-world datasets, IDGNN outperforms state-of-the-art dynamic GNN baselines in both classification and regression tasks.
- According to conventional implicit methods for training, the recommended bilevel optimisation and single-loop training algorithm provide a notable speedup (improvement to around 1600× in some situations). Because of the fixed-point structure, the model retains stable connections over time, reducing oversmoothing and maintaining long-range relationship identification.

4) Relevance to Current Research

The paper presents a novel framework called IDGNN (Implicit Dynamic Graph Neural Network), which guarantees continuous resolution to a fixed-point description even in dynamic graph environments. The computational cost of conventional implicit models is greatly decreased by using a bilevel optimisation technique, which permits a more effective single-loop training process. IDGNN exhibits exceptional performance and efficiency for both classification and regression tasks on dynamic datasets by successfully capturing long-range dependencies over designing networks.

5) Indirect Support

- provides a strong mathematical foundation for the use of pseudo-nodes by demonstrating how implicit models preserves long-range dependency.
- Illustrates how learning and structured propagation can strike a compromise between efficiency and accuracy, supporting your sparse connectivity strategy.
- Highlights the significance of effective training and dynamic connectivity in relation to your dynamic MPNN with sparse pseudo-nodes.
- Suggests that optimising the accuracy-efficiency trade-offs in dynamic MPNNs is a significant and current area of research.

6) Summary

The first-of-its-kind implicit dynamic graph neural network architecture is presented in the IDGNN paper. It conceptually assures a fixed-point solution and provides an effective learning method through bilevel optimisation. It significantly increases training efficiency while excelling at capturing long-range dependencies over time in dynamical graphs. In your study of sparse pseudo-node message forwarding in dynamic MPNNs, it provides both theoretical support as well as helpful suggestions for finding an agreement among efficiency and accuracy in changing graph architectures.

Link: https://arxiv.org/abs/2406.17894

4.5 Research paper 5:
"An Analysis of Virtual Nodes in GNNs for Link Prediction"*(Hwang et al., 2022)*

1) Background

Graph Neural Networks (GNNs) aggregate information from neighboring nodes to understand graph structures. However, they often struggle to capture relationships between distant nodes because of **over-squashing**, where information is compressed within limited message paths. This reduces their ability to represent global context efficiently, especially in large or dynamic graphs where maintaining all connections becomes computationally expensive.

2) Research focus

Hwang et al. (2022) explored the use of **virtual nodes** in GNN architectures to improve information flow across distant regions of a graph. Their study analyzed how the **connectivity patterns and density** of these virtual nodes affect both model performance and generalization. Instead of focusing only on accuracy, they also examined whether adding too many virtual-node connections could introduce redundancy and unnecessary computation, reducing overall efficiency.

3) Key Findings

- Adding virtual nodes can enhance long-range message passing and reduce over-squashing.
- However, **dense virtual-node connections** often lead to redundant communication and slower training.
- Sparse or selectively connected virtual nodes maintain performance while lowering computation.
- The optimal structure of virtual-node links strongly influences both **accuracy and efficiency** in GNNs.

4) Relevance to our research problem

- **Supports sparse message passing:** Their results show that excessive virtual-node connectivity can be wasteful, directly supporting your idea of **sparse pseudo-node message passing**.
- **Addresses efficiency–accuracy balance:** Hwang et al. demonstrate that while virtual nodes help information flow, a trade-off exists—too many links harm efficiency without improving accuracy.
- **Reinforces adaptive design:** Your work extends their findings by proposing **adaptive sparse pseudo-nodes** that dynamically adjust to changing graph structures, optimizing both computation and performance.

5) Indirect Support

- Provides **empirical evidence** that sparse pseudo-node connections are more effective than dense ones.
- Strengthens your argument that efficient GNN design depends on balancing **connectivity and cost**.
- Highlights that over-squashing and redundancy are not only accuracy issues but also **scalability challenges**—a critical point for dynamic MPNNs that evolve over time.

6) Summary

Hwang et al. (2022) analyzed how virtual nodes influence message passing in GNNs and found that while they help capture long-range dependencies, dense connections often reduce efficiency. Their findings emphasize the need for **sparse and adaptive linking**, forming a strong foundation for your research on **sparse pseudo-node message passing**, which aims to optimize the trade-off between efficiency and accuracy in dynamic MPNNs.

Link: https://openreview.net/pdf?id=dI6KBKNRp7

5. Comparative Analysis of Reviewed Studies

Paper	Model / Approach	Main Contribution / Findings	Limitations
Probabilistic Graph Rewiring via Virtual Nodes (Qian et al., 2024)	Virtual node–based probabilistic graph rewiring	Improves long-range message passing and reduces over-squashing through probabilistic rewiring.	Does not explore sparse or adaptive pseudo-node links; lacks explicit efficiency–accuracy trade-off.
On the Bottleneck of Graph Neural Networks (Alon & Yahav, 2020)	Theoretical analysis of GNN message-passing limits	Identifies over-squashing as the main issue in capturing long-range dependencies.	Provides no practical solution; only theoretical justification.
DRew: Dynamically Rewired Message Passing with Delay (Di Giovanni et al., 2023)	Dynamic rewiring mechanism	Adjusts connections across layers to improve message propagation.	Ignores pseudo-nodes and sparsity; efficiency optimization not addressed.
Efficient and Effective Implicit Dynamic GNN (IDGNN, Zhong et al., 2024)	Implicit update method for dynamic graphs	Enhances efficiency in dynamic GNNs with lower computational cost.	Does not use pseudo-nodes or sparse communication strategies.
An Analysis of Virtual Nodes in GNNs (Hwang et al., 2022)	Empirical study on virtual node effects	Shows that virtual-node links can help long-range propagation but may cause redundancy when dense.	Focuses on static graphs; lacks adaptive or sparse connectivity exploration.

6.Proposed Future Directions

- **Sparse and Adaptive Pseudo-node Message Passing:** Introduce only selected pseudo-nodes and adaptively link them to capture long-range dependencies efficiently, balancing accuracy and computation for dynamic graphs.

- **Dynamic Edge Rewiring for Sparse Graphs:** Apply adaptive rewiring with sparsity constraints to reduce unnecessary connections while maintaining effective message flow.

- **Efficiency-focused Implicit Updates:** Integrate implicit update techniques to lower computational cost, updating only relevant nodes or pseudo-nodes in evolving graphs.

- **Adaptive Trade-off Modelling:** Explicitly model the efficiency–accuracy trade-off so the network dynamically decides the number and placement of pseudo-node connections.

- **Task-specific Sparse Optimization:** Optimize pseudo-node placement for particular tasks (e.g., link prediction, node classification) using methods like attention or reinforcement learning.

7. Potential Contributions

- Sparse Pseudo-node Message Passing: proposes an original sparse pseudo-node architecture to improve dynamic MPNNs' long-range communication.

- Improved Graph Connectivity: Presents an adaptive scarification technique that builds pseudo-node linkages in a chosen manner according to message significance and graph topology.

- Precision-Efficacy Trade-off Modelling: Provides a quantitative framework for evaluating and weighing computing cost and model correctness.

- Pseudo-nodes can vary as time passes with dynamic graph adaptability, preserving accuracy and efficiency as the graph structure shifts.

- Reduction of Over-squashing: Presents experimental support and a theoretical justification for how sparse pseudo-nodes reduce message bottlenecks.

- Layer-wise Information Control: increases the flow of information depth-wise without needing undue computation by implementing layer-dependent pseudo-node updates.

- Integration of Probabilistic and Dynamic Concepts: Integrates concepts from dynamic message forwarding (Di Giovanni et al., 2023) and probabilistic graph rewiring (Qian et al., 2024).

- By guaranteeing compatibility with present GNN architectures, structure generalisation makes it flexible enough to handle an array of dynamic graph jobs.

- By extending fundamental theories (Alon & Yahav, 2020) into a realisable, efficiency-conscious approach, it crosses the gap between theoretical and practical insights.

- Guidelines for Graph Scarification: Provides useful information on whether and how model performance is improved by sparse pseudo-node connections.

- The work provides a link across dynamic efficiency-focused GNN studies and static virtual-node research, making it a contribution to the dynamic MPNN literature.

8. Solutions

Overview of the Concept

The ASPI Framework introduces a programmable and flexible pseudo-node mechanism that dynamically modifies the number and location of pseudo-nodes in response to the intensity of the flow of messages and real-time graph evolution. To achieve both accuracy and efficiency in dynamic MPNNs, it integrates topology-aware learning, flexibility in time, and attention-based scarification.

Why It's Novel

Learning-driven sparse pseudo-node adaptation, which ASPI presents compared with different research that use static or compact pseudo-nodes, enables the network to develop its own connection for the greatest accuracy–efficiency balance.This makes it especially well-suited for real-time dynamic MPNNs, where links' structure and importance change over time.

Key Innovations

- In order to improve long-range communication in bottleneck areas, adaptive pseudo-node allocation dynamically adds or removes pseudo-nodes during training based on node centrality or message entropy.
- Attention-guided Sparse Connectivity: Reduces redundant message forwarding by establishing only high-impact pseudo-node links using an attention-based controller.
- The Time Evolution Method maintains pseudo-nodes' temporal memory intact by trimming or updating them as the graph changes without requiring retraining.
- Dual Optimisation Goal: Using a dynamic trade-off coefficient (λ) to balance efficiency (lowering connection density) and accuracy (minimising over-squashing). Topology-aware RL Agent: Uses reinforcement education to maximise the density and location of pseudo-nodes according to cost and performance.
- In order to encourage effective local and global message propagation with the least amount of work, hierarchical sparse aggregation groups pseudo-nodes into clusters.

9. Conclusion & References

8.1 Conclusion

This study proposes a sparse pseudo-node message passing framework to investigate the accuracy–efficiency trade-off in Dynamic Message Passing Neural Networks (MPNNs). Focussing on findings of recent research in implicit graph neural networks, dynamic rewiring, and virtual nodes, the suggested approach improves long-range transfer of information while reducing computational costs. The article shows how sparse pseudo-nodes can successfully reduce over-squashing while improving scalability in dynamic graph environments through theoretical reasoning and comparative analysis. All things examined, this work promotes the creation of MPNN architectures that are more accurate, efficient, and scalable for dynamic graph applications in the real world.

8.2 References

- Stack exchange network – Anology between MPNNs and GCNs: https://stats.stackexchange.com/questions/619757/analogy-between-mpnns-and-gcns
- We got the unsolved research gap from this paper – "Towards Dynamic Message Passing on Graphs (N^2)" : https://arxiv.org/abs/2410.23686
- W. Hamilton, Z. Ying, and J. Leskovec, "Inductive Representation Learning on Large Graphs," *Advances in Neural Information Processing Systems (NeurIPS)*, 2017: https://arxiv.org/abs/1706.02216
- OpenReview.net - On dimensionality of feature vectors in MPNNs: https://openreview.net/forum?id=UjDp4Wkq2V