Research Statement: Exploring Algorithmic Innovations in Graph and Hypergraph Analytics

As a seasoned researcher specialising in the algorithmic, applied, and topological aspects of non-relational data, particularly graphs and their higher-order counterpart, hypergraphs, my research endeavours have focused on advancing the frontiers of representation learning, topological data analysis, and algorithmic innovations.

Graph Representation Learning: My research in graph representation learning has introduced novel approaches that have significantly contributed to diverse domains. By pioneering the use of physics-driven Graph Neural Networks (GNNs), I have harnessed the power of low-resolution simulation data to efficiently and effectively predict flow fields on high-resolution meshes from complex aerodynamical systems [1]. I have also tackled critical challenges, such as enhancing the robustness of GNNs under adversarial attacks [9] and developing algorithms to measure and mitigate uncertainty associated with real-valued properties on uncertain graphs [10].

Higher-Order Graph Representations and Topological Data Analysis: In the realm of higher-order graph representations and topological data analysis, I have developed novel algorithms for identifying core nodes in hypergraphs [2], generating and estimating properties of random hypergraphs [3], visualising hypergraphs [5], and subsampling nodes to approximate and accelerate the computation of topological features [4]. My research in this domain has paved the way for enhanced understanding and analysis of complex relational structures.

Publication and Patent Portfolio: With a track record of publications in reputable peer-reviewed databases, data mining, and machine learning venues such as ICML [1] (with a spotlight), VLDB [2], DSAA [10], ECML-PKDD [6], DEXA [3, 4, 5], and high-impact journals like SEC [7], coupled with a patent granted at the UK IP office [8], my contributions have been recognised and disseminated widely in the research community. Notable publications include works on hypergraph core decomposition, random hypergraph generation, and topological data analysis.

Industry Collaborations and Impact: My research has had a tangible impact through collaborations with industry leaders such as Rolls-Royce@NTU Corporate Lab. Our joint efforts have led to significant breakthroughs in super-resolution problems for computational fluid dynamics (CFD), resulting in publications in prestigious conferences like ICML [1], IEEE CAI [11] and a patent granted at the UK Intellectual Property Office [8]. In addition, I have also collaborated with external collaborators and contributed to various topics, including hypergraph core decomposition [2], uncertain graph modelling [10], and adversarial robustness of GNNs [9], thereby advancing both real-world applications and theoretical understanding.

Future research Interests:

(1) **Inconsistencies and bias in LLMs.** Large language models have become pervasive in recent years. There are some methods that detect and mitigate bias in LLMs. However, these methods do not explain why such bias exists in existing LLMs. Precise formulation, exploration and mitigation of inconsistencies inherent in the response of LLMs are also exciting problems for future work.

(2) **Relational and tabular representation learning via higher-order primitives**. Learning representations of relational and tabular data via hypergraphs has not been done in the literature despite hypergraphs being a natural candidate for representing such data formats.

(3) Explainability and Robustness of Hypergraph- and other higher-order neural networks. Traditional representation learning for graphs is insufficient for hypergraphs because graphs cannot losslessly represent higher-order information. Recently, there has been an interest in designing neural networks that can learn representations for data structured as hypergraphs and simplicial complexes. However, their explainability and robustness are under-studied.

(4) **Explainability of GNNs in the physics applications.** Recently, data-driven GNN approaches for CFD have become a viable alternative to numerical simulation in terms of efficiency and precision. However, the explainability and reliability of these models have not been studied yet, which is important for their wider adoption in industrial applications.

(5) Exploration of aleatoric uncertainty of GNNs. Much effort has been devoted to understanding epistemic uncertainty in interpretable GNNs. However, the uncertainty stemming from the graph and its impact on the learned model is understudied.

Summary: My research journey has been characterised by a relentless pursuit of innovation and collaboration across diverse domains. With a focus on algorithmic advancements in graph and hypergraph analytics, I have collaborated with esteemed researchers from renowned institutions worldwide, including NUS (Singapore), NTU (Singapore), Aalborg University (Denmark), INRIA (France), Telecom Paris (France), University of Vienna (Austria), CENTAI Institute (Italy), and the University of Texas at Dallas (US). Together, we have tackled complex challenges spanning graphs, hypergraphs, uncertain graph modelling, and graph neural networks, contributing to the advancement of real-world applications and the enrichment of theoretical understanding in the field. I remain committed to driving impactful research that transcends boundaries and fosters innovation at the intersection of academia and industry.

References:

[1] *Loh Sher En Jessica, ***Naheed Anjum Arafat**, Wei Xian Lim, Wai Lee Chan, and Adams Wai Kin Kong. Finite volume features, global geometry representations, and residual training for deep learning-based CFD simulation. ICML 2024 (spotlight). * = equal contribution

[2] **Naheed Anjum Arafat**, Arijit Khan, Arpit Kumar Rai, and Bishwamittra Ghosh. Neighbourhood-based hypergraph core decomposition. pVLDB 2023.

[3] **Naheed Anjum Arafat,** Debabrota Basu, Laurent Decreusefond, and Stéphane Bressan. Construction and random generation of hypergraphs with prescribed degree and dimension sequences. DEXA 2020.

[4] **Naheed Anjum Arafat**, Debabrota Basu, and Stéphane Bressan. Topological data analysis with ε-net induced lazy witness complex. DEXA 2019.

[5] **Naheed Anjum Arafat** and Stéphane Bressan. Hypergraph drawing by force-directed placement. DEXA 2017.

[6] **Naheed Anjum Arafat**, Debabrota Basu, and Stéphane Bressan. ε-net induced lazy witness complexes on graphs. Workshop on Applications of Topological Data Analysis (ATDA), ECML-PKDD 2019.

[7] Siddhartha Shankar Das, Md Monirul Islam, and **Naheed Anjum Arafat**. Evolutionary algorithm using adaptive fuzzy dominance and reference point for many-objective optimization. Swarm and evolutionary computation, 2019.

[8] Fluid flow simulation. Inventors: Loh Sher En Jessica, **Naheed Anjum Arafat**, Adams Wai Kin Kong, Wai Lee Chan, Bryce D Conduit. Status (Sep 2024): Granted at the UK Intellectual Property Office (https://www.ipo.gov.uk/p-ipsum/Case/ApplicationNumber/GB2312389.6).

[9] **Naheed Anjum Arafat**, Debabrota Basu, Yulia Gel, and Yuzhou Chen. When witnesses defend: A witness graph topological layer for adversarial graph learning. AAAI 2025. (Under review)

[10] **Naheed Anjum Arafat**, Ehsan Bonabi Mobaraki, Arijit Khan, Yllka Velaj, and Francesco Bonchi. Estimate and Reduce Uncertainty in Uncertain Graphs, DSAA 2024.

[11] Wei Xian Lim, **Naheed Anjum Arafat**, Wai Lee Chan, and Adams Wai Kin Kong. Multi-Order Loss Functions For Accelerating Unsteady Flow Simulations with Physics-Based AI, IEEE CAI 2024.