**NGUYEN NGOC KHANH**

Phone: +65 9778 7635

Email: [khanh.nguyen.contact@gmail.com](mailto:khanh.nguyen.contact@gmail.com?subject=[Indst%20CV])

LinkedIn: [khanh-nguyen-446809140](https://www.linkedin.com/in/khanh-nguyen-446809140/)

Website: <https://khanhcsc.github.io/>

**EDUCATION**

**Nanyang Technological University, Singapore**

**Aug 2017 – Jul 2021 (Expected)**

* **Bachelor of Engineering (Computer Science)**
* Elective Tracks: High-Performance Computing, Artificial Intelligence, Data Science.
* Highest Distinction in: Compiler Techniques, Advanced Topics in Algorithms.

**PROJECTS**

**Academic Projects – Nanyang Technological University, Singapore**

**Final Year Project (CZ4079, individual): Cluster Analysis on Dynamic Graphs.**

* Reviewed literature in Graph Clustering:
  + Spectral Clustering
  + Node Embedding
  + Model-based Clustering.
* Extended and Analyzed Gibbs sampling algorithm on distant-dependence Chinese Restaurant Process for graph clustering problem.
* Introduced a novel cluster ensemble that is capable to respond the informative clustering evolution.
* Conducted intensive experiments to benchmark the performance of the new algorithms on both synthesis networks and real dynamic networks.

**Network Science (CZ4071, group-based): Survey and Implementation of NeurIPS'19 "Layer-Dependent Importance Sampling for Training Deep and Large Graph Convolutional Networks"**

* Generated networks based on stochastic block model.
* Summarized the mathematical formulation of the research using Tex.
* Implemented the correct version of code for the research using Pytorch.
* Performed experiments and analyzed the outputs and confirmed the superior results from the research.

**WORK EXPERIENCE / INTERNSHIP**

**Shopee, Singapore – Data Science Intern**

**Jan 2020 – Present**

Project 1 (Individual): Nebula 2 benchmark

* Generated LDBC SNB SF1000 dataset using Spark (~1.5TB)
* Transformed the dataset using Spark.
* Performed benchmark on Nebula 2 with different configurations.

Project 2 (Individual): Address NER

* Reviewed literature in Named-Entity Recognition and Classification.
* Implemented Lample et al ‘s work on two-level LSTM CRF using PyTorch.
* Trained the model that surpassed the previous approach (CRF) in F1 score for a dataset of size 1.6M addresses (achieve 0.98 of F1 score as compared to 0.92 previously).

**Nanyang Technological University, Singapore – SCALE Lab, Student Assistant for Research**

**Dec 2020 – Mar 2021**

Project 1 (Individual): Multi-Robot Patrolling Algorithms

* Reviewed literature for multiple travelling salesman problem (MTSP).
* Designed and proved the soundness and completeness of the MTSP solution for the real-world requirements.
* Designed and implemented relaxed algorithms to approximate MTSP solution based on spectral clustering method.
* Improved the baseline solution by designing and implementing a new heuristic for the multi-objective problem.
* Designed and implement a search algorithm based on Conflict-Based Search to convert the solution of MTSP to real-world requirements.

**Shopee, Singapore – Platform Engineering Intern**

**May 2020 – August 2020**

Project 1 (Group-based): Configuration Center Refactor

* Participated in DB Design, API Design
* Developed the Manager layer
* Tested and Fixed the functionalities of the (WIP) code.
* Updated the documentation for the API Design.

Task 1 (Individual): Extended HTTP Gateway for the RPC Agent

* Studied the design and Extended the functionality of the RPC Agent.
* Produced the Technical Requirement Design for review and Proceeded to develop the features according to the chosen design.

Task 2 (Individual): Investigation on a UI bug

* Studied the implementation of the component.
* Discussed with the authors of a third-party library (grpc-gateway)
* Produced the explanation and solution for the issue according to the inconsistency between two specifications. (swagger and proto).

Task 3 (Individual): Investigation on a timeout issue.

* Studied the implementation of the component and its features.
* Produced the explanation and a hotfix for the issue.
* Produced a longer-term solution that helped to improve the code performance

**AWARDS / ACHIEVEMENTS**

**Citi-Hackathon 2019**

**(Team-based)**

Best Application

**NSCC – APAC HPC-AI Competition 2019**

**(Team-based)**

Best HPC Performance

**National Data Science Challenge 2019**

**(Team-based)**

Champion Team

**IET Machine Learning Challenge 2018**

**(Team-based)**

Second Runner Up Team

**International Physics Olympiad 2015**

**(Individual)**

Silver Medal

**Asian Physics Olympiad 2015**

**(Individual)**

Gold Medal

Best Theoretical Result of Vietnamese Team

**Asian Physics Olympiad 2014**

**(Individual)**

Bronze Medal

**SKILLS / COMPETENCIES / INTERESTS**

**Technical Skills**

**Programming Skills**

Machine Learning

Python, MATLAB

Software Engineering

Others

Go, C++, Java, SQL

Latex