

## THÔNG TIN VỀ LUẬN ÁN TIẾN SĨ

- Họ và tên nghiên cứu sinh: Hà Thị Kim Dung
- Giới tính: Nữ
- Ngày sinh: 03/10/1984
- Nơi sinh: Quảng Ninh
- Quyết định công nhận nghiên cứu sinh số: 1200/QĐ-CTSV, ngày 29 tháng 12 năm 2020 của Hiệu trưởng trường Đại học Công nghệ, Đại học Quốc gia Hà Nội.
- Các thay đổi trong quá trình đào tạo:
  - Quyết định số 167/QĐ-ĐHCN ngày 10 tháng 3 năm 2023 của Hiệu trưởng trường Đại học Công nghệ về việc bổ sung cán bộ hướng dẫn cho Nghiên cứu sinh Hà Thị Kim Dung, khóa QH-2020.
- Tên đề tài luận án: Một số thuật toán xấp xỉ cho bài toán tối ưu hàm dạng submodular với ràng buộc;  
(tên luận án chính thức đề nghị bảo vệ cấp Đại học Quốc gia)
- Chuyên ngành: Khoa học máy tính
- Mã số: 9480101
- Cán bộ hướng dẫn khoa học:
  - Hướng dẫn chính: PGS. TS. Hoàng Xuân Huân, Khoa CNTT, Đại học Công nghệ, ĐHQG Hà Nội;
  - Hướng dẫn phụ: TS Phạm Văn Cảnh, Khoa CNTT, Đại học Phenikaa.
- Tóm tắt các **kết quả mới** của luận án: Luận án nghiên cứu các thuật toán xấp xỉ cho bài toán tối ưu hàm submodular với dữ liệu lớn. Trong đó, luận án đưa ra các thuật toán xấp xỉ cạnh tranh cho 3 bài toán.

Đầu tiên, bài toán tối đa hàm mục tiêu  $k$ -submodular với ràng buộc chi phí (Bài toán  $k$ SMK). Luận án tập trung giải quyết bài toán khi hàm mục tiêu đơn điệu (monotone). Với trường hợp này, luận án đã đóng góp nhiều thuật toán cải tiến cho tỉ lệ xấp xỉ tốt hơn các thuật toán tốt nhất hiện nay nhưng giảm được độ phức tạp truy vấn tuyến tính thêm một hệ số. Ngoài ra, luận án cũng đề xuất thuật toán cải tiến để giải cho trường hợp hàm mục tiêu không đơn điệu (non-monotone). Trong trường hợp này, luận án trình bày một thuật toán

xấp xỉ cho tỉ lệ xấp xỉ tốt tương đương với thuật toán tốt nhất hiện nay nhưng cũng giảm được số truy vấn tuyến tính thêm một hệ số. Các thực nghiệm cho thấy các thuật toán của tác giả và cộng sự cho chất lượng cạnh tranh, giảm thời gian chạy và số truy vấn đáng kể, đặc biệt khi dữ liệu đầu vào và ngân sách tăng lên.

Tiếp theo, luận án nghiên cứu bài toán tối đa hàm submodular dưới ràng buộc chi phí với nhiễu (Bài toán SMKN). Qua đó, tác giả đề cập Thuật toán tham lam giải quyết bài toán. Sau đó, luận án đề xuất thuật toán Luồng (Streaming) để cải tiến thuật toán tham lam nhằm giảm thời gian chạy và tối ưu bộ nhớ. Tác giả cũng trình bày phân thực nghiệm cho thấy thuật toán cải tiến cho giá trị hàm mục tiêu tốt gần bằng Tham lam, trong khi thời gian chạy và bộ nhớ tốt hơn.

Cuối cùng, luận án nghiên cứu bài toán Submodular Cover trên lưới nguyên (Bài toán DRSC). Luận án đề xuất thuật toán tiêu chí kép dựa được thiết kế song song hoá cho chất lượng lời giải cạnh tranh song còn giảm được đáng kể số lượng truy vấn, số lượng vòng tuần tự của thuật toán, qua đó giảm được thời gian chạy.

12. Khả năng ứng dụng trong thực tiễn: Các thuật toán có thể ứng dụng trong các bài toán phổ biến trong Học máy và Trí tuệ nhân tạo như Tối đa ảnh hưởng, Tối đa doanh thu, Tóm tắt tài liệu, Tóm tắt hình ảnh, Đặt sensor cảm biến, Thiết lập cơ sở, Tối ưu cấp phát tài nguyên...

13. Những hướng nghiên cứu tiếp theo: Tiếp tục phát triển các thuật toán hiệu quả cho các bài toán; Mở rộng nghiên cứu sang các bài toán khác như Tối đa hàm submodular với ràng buộc d-knap; Tối đa hàm non-submodular; Bài toán kSMK trong môi trường nhiễu; Các bài toán SM giải trên lưới nguyên.

14. Các công trình đã công bố có liên quan đến luận án:

(liệt kê các công trình theo thứ tự thời gian)

- **Dung T. K Ha**, Canh V. Pham and Huan X. Hoang, *Submodular Maximization Subject to a Knapsack Constraint Under Noise Models*, Asia-Pacific Journal of Operational Research, Tập 39, Số 6, 2022 (**ISI/Q3**);

- Canh V Pham, **Dung KT Ha**, Huan X Hoang and Tan D Tran, *Fast Streaming Algorithms for k-Submodular Maximization under a Knapsack Constraint*, IEEE 9th International Conference on Data Science and Advanced Analytics (DSAA), 2022 (**ISI/RANK A**);

- Canh V. Pham and **Dung T.K. Ha**, *A note for approximating the submodular cover problem over integer lattice with low adaptive and query complexities*, Information Processing Letters, Volume 182, 2023, (ISI/Q3);
- **Dung T.K. Ha**, Canh V. Pham, Tan D. Tran, *Improved Approximation Algorithms for  $k$ -Submodular Maximization under a Knapsack Constraint*, Computers & Operations Research, 2023 (ISI/Q1);
- **Dung T. K Ha**, Canh V. Pham, Tan D. Tran and Huan X. Hoang, *Robust Approximation Algorithms for Non-monotone  $k$ -Submodular Maximization under a Knapsack Constraint*, The 15th IEEE International Conference on Knowledge And Systems Engineering (KSE 2023).

Ngày tháng năm 20  
**Xác nhận của cán bộ hướng dẫn 1**  
(Kí và ghi rõ họ tên)

Ngày tháng năm 20  
**Nghiên cứu sinh**  
(Kí và ghi rõ họ tên)

**Hoàng Xuân Huấn**  
**Xác nhận của cán bộ hướng dẫn 2**  
(Kí và ghi rõ họ tên)

**Phạm Văn Cảnh**

### INFORMATION ON DOCTORAL THESIS

1. Full name: Ha Thi Kim Dung
2. Sex: Female
3. Date of birth: October 3<sup>rd</sup>, 1984
4. Place of birth: Quang Ninh
5. Admission decision number: 1200/QĐ-CTSV Dated: Dec 29<sup>th</sup>, 2020.
6. Changes in academic process: (*List the forms of change and corresponding times*)
  - Decision No. 167/QĐ-DHCN dated March 10th, 2023 of the Principal of the University of Engineering and Technology on additional instructors for PhD student Ha Thi Kim Dung, course QH-2020;
7. Official thesis title: Approximation algorithms for constrained submodular optimization problems;
8. Major: Computer science
9. Code: 9480101
10. Supervisors:  
(*Full name, academic title and degree*)
  - 1<sup>st</sup> supervisor: Assoc. Prof. Hoàng Xuân Huân, Faculty of Information Technology, University of Engineering and Technology, Vietnam National University, Ha Noi;
  - 2<sup>nd</sup> supervisor: Ph.D. Phạm Văn Cảnh, Faculty of Information Technology, Phenikaa University.
11. Summary of the **new findings** of the thesis: The thesis researches approximation algorithms for submodular optimization problems with big data in which it proposes expensive approximation algorithms for 3 following problems.

First, the thesis proposes several approximation algorithms for the problem of  $k$ -submodular maximization under knapsack constraint (kSMK). It focuses on the monotone case. It contributes some algorithms that outperform state-of-the-art ones. It provides better approximation ratios but reduces query complexities with a factor than the best ones. Besides, some improved approximation algorithms for non-monotone are also devised in this thesis. In this case, it presents an approximation algorithm providing an equivalent

approximation ratio to the existing ones but reduces the query complexity with a factor. The experiment shows that the algorithms give competitive performance and considerably reduce running time and the number of queries, especially when the size of input data and the budget increase.

Second, the thesis works with the problem of submodular maximization subject to knapsack constraint under noises (SMKN). This work gives a Greedy version of SKMN and then proposes a Streaming one to improve the Greedy to reduce time consumption and optimize the memory. The experiment also illustrates the new one providing objective values close to the Greedy while providing better running time and memory.

The last one is to solve the problem of Submodular Cover over the Integer lattice (DRSC). The thesis devises a bi-criteria approximation algorithm designed parallelized by the Adaptivity concept to give competitive quality guarantees and significantly reduce the number of queries and the adaptive complexity, which are useful for reducing the running time.

12. Practical applicability, if any: The proposed algorithms are applicable for several public applications in Machine Learning and Artificial Intelligence, such as Influence Maximization, Revenue Maximization, Document Summarization, Image Summarization, Sensor Placement, Location Facility, Optimal Budget Allocation, etc.

13. Further research directions, if any: For future work, the research will continue to improve efficient approximation algorithms for submodular optimization problems. Extending to another problem such as Submodular Maximization under d-knapsack constraint, non-submodular maximization, or the kSMK problem under noises, etc.

14. Thesis-related publications:

*(List them in chronological order)*

- **Dung T. K Ha**, Canh V. Pham and Huan X. Hoang, *Submodular Maximization Subject to a Knapsack Constraint Under Noise Models*, Asia-Pacific Journal of Operational Research, Volume 39, Number 6, 2022 (**ISI/Q3**);

- Canh V Pham, Dung KT Ha, Huan X Hoang, and Tan D Tran, *Fast Streaming Algorithms for k-Submodular Maximization under a Knapsack Constraint*, IEEE 9th International Conference on Data Science and Advanced Analytics (DSAA), 2022 (**ISI/RANK A**);

- Canh V. Pham and **Dung T.K. Ha**, *a note for approximating the submodular cover problem over integer lattice with low adaptive and query complexities*, Information Processing Letters, Volume 182, 2023, (**ISI/Q3**);

- **Dung T.K. Ha**, Canh V. Pham, Tan D. Tran, *Improved Approximation Algorithms for  $k$ -Submodular Maximization under a Knapsack Constraint*, Computers & Operations Research, 2023 (ISI/Q1);

- **Dung T. K Ha**, Canh V. Pham, Tan D. Tran and Huan X. Hoang, *Robust Approximation Algorithms for Non-monotone  $k$ -Submodular Maximization under a Knapsack Constraint*, the 15th IEEE International Conference on Knowledge And Systems Engineering (KSE 2023).

Date: .....  
Signature: .....  
Full name: Hà Thị Kim Dung

Date: .....  
Signature: .....  
Full name: Hoàng Xuân Huân

Date: .....  
Signature: .....  
Full name: Phạm Văn Cảnh