Using and Evaluating Quantum Computing for Information Retrieval and Recommender Systems

Maurizio Ferrari Dacrema
Politecnico di Milano
Milano, Italy
maurizio.ferrari@polimi.it

Andrea Pasin*
Università degli Studi di Padova
Padova, Italy
andrea.pasin.1@phd.unipd.it

Paolo Cremonesi
Politecnico di Milano
Milano, Italy
paolo.cremonesi@polimi.it

Nicola Ferro
Università degli Studi di Padova
Padova, Italy
nicola.ferro@unipd.it

ABSTRACT

The field of *Quantum Computing (QC)* has gained significant popularity in recent years, due to its potential to provide benefits in terms of efficiency and effectiveness when employed to solve certain computationally intensive tasks. In both *Information Retrieval (IR)* and *Recommender Systems (RS)* we are required to build methods that apply complex processing on large and heterogeneous datasets, it is natural therefore to wonder whether QC could also be applied to boost their performance. The tutorial aims to provide first an introduction to QC for an audience that is not familiar with the technology, then to show how to apply the QC paradigm of *Quantum Annealing (QA)* to solve practical problems that are currently faced by IR and RS systems. During the tutorial, participants will be provided with the fundamentals required to understand QC and to apply it in practice by using a *real* D-Wave quantum annealer through APIs.

CCS CONCEPTS

• **Computer systems organization → Quantum computing**; • **Information systems → Information retrieval**; **Recommender systems**.

KEYWORDS

Quantum Computing, Quantum Annealing, Information Retrieval, Recommender Systems

ACM Reference Format:
Maurizio Ferrari Dacrema, Andrea Pasin, Paolo Cremonesi, and Nicola Ferro. 2024. Using and Evaluating Quantum Computing for Information Retrieval and Recommender Systems. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '24), July 14–18, 2024, Washington, DC, USA*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3626772.3661378

*Contact presenter.

SIGIR '24, July 14–18, 2024, Washington, DC, USA

ACM ISBN 979-8-4007-0431-4/24/07.
https://doi.org/10.1145/3626772.3661378

1 COVER SHEET

Title: Using and Evaluating Quantum Computing for Information Retrieval and Recommender Systems

Length: The duration of the tutorial will be half-day (3 hours) plus breaks and it will be subdivided into 4 parts.

Format: All tutorial presenters commit to be on-site.

1.1 Intended Audience

Targeted audience: The intended audience for this tutorial are researchers and practitioners coming from *Information Retrieval (IR)* and *Recommender Systems (RS)*, as well as from other fields, such as Natural Language Processing, Machine Learning, Big Data, Operations Research, and Optimization. In fact, QC and QA are powerful tools that can be applied to tackle problems in many different domains. Furthermore, while the practical part of the tutorial is focused on how to use QA for IR and RS, the problems we consider are quite general and have wide application to different research areas.

Prerequisite knowledge: This tutorial will be self-contained and has minimal prerequisite knowledge, mainly consisting of being familiar with the basic concepts of IR and RS. For those interested in the hands-on part, basic Python programming skills are required to interact with quantum annealers through the tools provided by D-Wave[1] but no specific knowledge of the D-Wave API is required ahead.

2 PRESENTERS

The tutorial has four presenters:

Maurizio Ferrari Dacrema is Assistant Professor at Politecnico di Milano, Italy. His main research focus is the application of Quantum Computing to machine learning tasks as well as the use of machine learning to improve the effectiveness of current generation quantum computers. He also has significant experience on reproducibility and evaluation of recommender systems. He won the ACM Best Paper Award at ACM RecSys 2019, has been teaching assistant at Politecnico di Milano since 2016, in particular for the MSc course on Recommender Systems since 2017, for the MSc course on Quantum Computing since 2023 and was Lecturer of the PhD course Applied Quantum Machine Learning in 2021.

[1]https://docs.ocean.dwavesys.com/en/stable/

He is Demo and Late-Breaking Results Co-Chair of RecSys 2024, presenter of the first edition of this proposed tutorial at ECIR 2024, and was co-organizer of the LERI workshop at RecSys 2023.

Andrea Pasin (main contact) is a PhD student at University of Padua, Italy, currently studying and investigating the possible applications of Quantum Annealing for Information Access systems to improve their performance. He was presenter of the first edition of this proposed tutorial at ECIR 2024.

Paolo Cremonesi is Full Professor of Recommender Systems and Quantum Computing at Politecnico di Milano, Italy. He has extensive experience on Quantum Computing, and he is the co-director of the Quantum Computing research activities within the Italian National Research Centre for High Performance Computing, Big Data and Quantum Computing (ICSC[2]). He has also extensive experience in the reproducibility and evaluation of recommender systems. He is member of the Steering Committee of ACM RecSys. He was General Chair of ACM RecSys 2017 and Program Chair of ACM TvX 2013. Paolo co-authored and presented several tutorials on: "Quantum Computing for Information Retrieval and Recommender Systems" at ECIR 2024, "Sequence-aware Recommender Systems" at WWW 2019, RecSys 2019 and UMAP 2019, "Cross-domain Recommender Systems" at RecSys 2014, "Evaluation of Recommender Systems" at UMAP 2013, "Recommender Systems" at TvX 2012.

Nicola Ferro is Full Professor in Computer Science at the Department of Information Engineering at the University of Padua, Italy. His main research interests are information retrieval, data management and representation, and their evaluation. He chairs the Steering Committee of CLEF, the European evaluation initiative on multimodal and multilingual information access systems, and the Steering Committee of ESSIR, the European Summer School on Information Retrieval. He is Senior PC Member in top-tier conferences, like ECIR, ACM SIGIR, ACM CIKM, WSDM. He was General Chair of ESSIR 2016 and Associate Editor for ACM TOIS. He was awarded the SIGIR Academy in 2023. He was presenter of the first edition of this proposed tutorial at ECIR 2024.

3 MOTIVATION

The interest for the *Quantum Computing (QC)* field has been rapidly growing in recent years [1], with many researchers and practitioners from different backgrounds building new methods using quantum computers to perform faster computations. Thanks to considerable technological advancements, it has been possible to use QC to tackle practical problems achieving good results in terms of efficiency and effectiveness [4, 7, 9, 17, 18, 25] and QC is moving more and more out from being a still exploratory research field into becoming a viable computing paradigm with more and more consolidated infrastructures and software engineering practices [20, 21].

IR and RS deal with large and heterogeneous datasets and rely on complex and computationally demanding algorithms and techniques. Therefore, both fields could substantially benefit from the efficiency improvements delivered by QC. Moreover, due to the intrinsic uncertainty and vagueness entailed by the retrieval and recommendation tasks, effectiveness is of primary concern for both IR and RS and this trade-off between efficiency and effectiveness is an interesting angle for QC itself.

Despite quantum mechanics concepts and formalism have been widely theorized and explored in the Quantum IR area [2, 11, 22, 24], the use of QC technologies themselves to formulate, implement, and execute IR and RS algorithms has been explored to a limited extent and it is much less familiar for researchers and practitioners in the area.

Finally, benchmarking and evaluation of QC is an open research area in itself [19] and almost unexplored when it comes to evaluation of QC for IR and RS.

For these reasons, we propose a tutorial covering the fundamental concepts of QC, with a focus on the practical application of the *Quantum Annealing (QA)* paradigm [25] through interactive coding sessions teaching participants how to use the cutting-edge quantum annealers to solve realistic problems. We have chosen QA, among the available paradigms, because it offers a good trade-off between computational power and a low access barrier. In fact, in order to perform a task on a quantum annealer one needs to describe it as an optimization problem using the relatively straightforward *Quadratic Unconstrained Binary Optimization (QUBO)* [6] formulation. This hides the underlying complexity of the QC system and makes it easily accessible for people without a background in quantum mechanics. Furthermore, QA has already been applied to IR and RS tasks [3, 4, 13, 14] showing that, although not always superior to classical methods, quantum annealers have matured enough to reliably tackle realistic problems. Finally, The tutorial will also present the lessons learned from QuantumCLEF[3] [15, 16], a new lab organized at CLEF 2024 that offers an infrastructure to develop and evaluate QA algorithms applied to IR and RS tasks amd that, to the best of our knowledge, is the first activity of this kind for IR and RS.

4 OBJECTIVES

The objective of this proposed tutorial is threefold. The first goal is to provide an introduction to QC and QA for an audience that is not familiar with it, with a particular focus on reducing what is often perceived as a high barrier of entry. The second goal is to present all necessary steps to formulate problems for QA in practice, with specific reference to algorithms for IR and RS. The third goal is to allow participants to run their own algorithms on a real QC with a hands-on session and to experiment with the technology.

Overall, we hope that this tutorial, and other related activities, sow the seeds of the growth of a vibrant IR and RS community, capable leveraging QC technologies to innovate the field.

5 RELEVANCE TO THE INFORMATION RETRIEVAL AND RELATED TUTORIALS

QC offers a computational paradigm that has the potential to significantly accelerate tasks that are difficult to tackle on traditional hardware. Both IR and RS face challenges in ensuring efficient processing and scalability. As such, successful application of QC would

[2]https://www.supercomputing-icsc.it

[3]https://qclef.dei.unipd.it/

be greatly impactful. Furthermore, QC may open new research directions allowing to perform efficiently tasks that are difficult to tackle on traditional hardware and that, due to this, have received less attention.

This proposal would be the second edition of our tutorial "Quantum Computing for Information Retrieval and Recommender Systems" presented at ECIR 2024[4] [5]. The tutorial was well received, with approximately 40 registered participants, and attendees said to be enthusiastic about what they learned from the tutorial.

This proposal builds upon the experience of the first edition, but it will also differentiate from it since we plan to expand the section concerning evaluation by including a discussion on the lessons learned from the QuantumCLEF lab, which currently has 18 registered participants[5]. Indeed, even if the actual CLEF conference happens in September, by July 2024 the activities of QuantumCLEF will be completed and the tutorial will thus allow us to accelerate knowledge transfer to the community.

To the best of our knowledge, are not aware of any other similar tutorial in the IR and RS fields.

6 FORMAT AND DETAILED SCHEDULE

The tutorial will cover theoretical and practical aspects underneath QC (and especially QA), by allowing participants to code and use real quantum annealers to solve optimization problems usually faced by many computer systems, including IR and RS systems.

6.1 Outline

The tutorial will be divided into 4 parts, starting from the theoretical aspects of QC, QA and the QUBO problem formulation. There will be a practical part where participants will be invited to code and solve a few selected problems using quantum annealers.

6.1.1 Part 1: QC Foundations (40 min). The first part consists of a gentle introduction to QC, showing its potential benefits as well as its limitations. We will also delve into the QA paradigm, trying to explain the difference between classical and quantum computation and dispelling some QC-related myths. It includes:

- an overview and basic understanding of QC and its potential benefits;
- a basic understanding of Quantum Circuit and Quantum Adiabatic models;
- relations between the classical optimizations meta-heuristics *Simulated Annealing* [23] and *Quantum Annealing* [8] with the Quantum Adiabatic model;
- the evolution of a quantum system to a state of minimal energy;
- how to represent the energy configuration of a quantum system (i.e., Hamiltonian) using the Ising model;
- the similarity between the Ising model and the QUBO representation for optimizing (NP-hard) problems.

6.1.2 Part 2: QUBO Formulation (50 min). This part will show how to represent classical binary optimization problems in QUBO formulation [10]. In particular, it will explain how to write NP-complete binary decision problems (e.g., number partitioning) and NP-hard binary optimization problems (e.g., quadratic assignment) in QUBO formulation, describing constraints and loss functions.

6.1.3 Part 3: QC for IR and RS and their Evaluation (30 min). This part will introduce IR and RS problems which can be solved by using QA , namely Feature Selection, Clustering and Model Boosting. Moreover, we will discuss how to evaluate such QA algorithms from both the efficiency and effectiveness point of view. Finally, it will introduce the QuantumCLEF lab and present the lessons learned from it.

6.1.4 Part 4: Hands-on (60 min). This part discusses how to use quantum annealers, which are available as a cloud service. It involves:

- the architecture and topology of a quantum annealer;
- how to use the QUBO formulation of a problem to program the quantum annealer via Minor Embedding;
- how the density of the QUBO problem impacts the number of variables required on the quantum annealer;
- how to program a quantum annealer and read the result (**hands-on**);
- Feature Selection and Clustering(**hands-on**);
- execution and evaluation of one of the above algorithms on the QuantumCLEF infrastructure (**hands-on**).

7 TYPE OF SUPPORT MATERIALS TO BE SUPPLIED TO ATTENDEES

The tutorial will include slides and Jupyter Notebooks. The materials will be available on GitHub prior to the tutorial and openly available afterwards, as was done for the first edition of the tutorial.

During the hands-on session, participants will also have access to the infrastructure made available by QuantumCLEF[6] that will provide virtual machines with all necessary tools already installed. This will enable participants play and experiment with D-Wave quantum annelears via Jupyter notebooks.

8 ACKNOWLEDGMENT

We acknowledge the financial support from ICSC - "National Research Centre in High Performance Computing, Big Data and Quantum Computing", funded by the European Union – NextGenerationEU.

We acknowledge the CINECA award under the ISCRA initiative, for the availability of high-performance computing resources and support.

REFERENCES

[1] J. Abhijith, A. Adedoyin, J. Ambrosiano, P. Anisimov, W. Casper, G. Chennupati, C. Coffrin, H. Djidjev, D. Gunter, S. Karra, N. Lemons, S. Lin, A. Malyzhenkov, D. Mascarenas, S. Mniszewski, B. Nadiga, D. O'malley, D. Oyen, S. Pakin, L. Prasad, R. Roberts, P. Romero, N. Santhi, N. Sinitsyn, P. J. Swart, J. G. Wendelberger, B. Yoon, R. Zamora, W. Zhu, S. Eidenbenz, A. Bärtschi, P. J. Coles, M. Vuffray, and A. Y. Lokhov. 2022. Quantum Algorithm Implementations for Beginners. *ACM Transactions on Quantum Computing (TQC)* 3, 4 (June 2022), 18:1–18:92.

[2] J. R. Busemeyer and P. D. Bruza. 2012. *Quantum Models of Cognition and Decision.* Cambridge University Press, Cambridge, UK.

[4] https://qclef.dei.unipd.it/ecir-2024-tutorial.html
[5] https://clef2024-labs-registration.dei.unipd.it/
[6] https://quantumclef.com/

[3] Maurizio Ferrari Dacrema, Nicolò Felicioni, and Paolo Cremonesi. 2021. Optimizing the Selection of Recommendation Carousels with Quantum Computing. In *RecSys '21: Fifteenth ACM Conference on Recommender Systems, Amsterdam, The Netherlands, 27 September 2021 - 1 October 2021*. ACM, 691–696. https://doi.org/10.1145/3460231.3478853

[4] Maurizio Ferrari Dacrema, Fabio Moroni, Riccardo Nembrini, Nicola Ferro, Guglielmo Faggioli, and Paolo Cremonesi. 2022. Towards Feature Selection for Ranking and Classification Exploiting Quantum Annealers. In *SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022*. ACM, 2814–2824. https://doi.org/10.1145/3477495.3531755

[5] M. Ferrari Dacrema, A. Pasin, P. Cremonesi, and N. Ferro. 2024. Quantum Computing for Information Retrieval and Recommender Systems, See [12], 358–362.

[6] F. Glover, G. Kochenberger, R. Hennig, and Y. Du. 2022. Quantum bridge analytics I: a tutorial on formulating and using QUBO models. *Annals of Operations Research* 314 (July 2022), 141–183.

[7] Tim Jaschek, Marko Bucyk, and Jaspreet S Oberoi. 2020. A quantum annealing-based approach to extreme clustering. In *Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC), Volume 2*. Springer, 169–189.

[8] Tadashi Kadowaki and Hidetoshi Nishimori. 1998. Quantum annealing in the transverse Ising model. *Phys. Rev. E* 58 (Nov 1998), 5355–5363. Issue 5. https://doi.org/10.1103/PhysRevE.58.5355

[9] Iordanis Kerenidis and Anupam Prakash. 2017. Quantum Recommendation Systems. In *8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9-11, 2017, Berkeley, CA, USA (LIPIcs, Vol. 67)*, Christos H. Papadimitriou (Ed.). Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 49:1–49:21. https://doi.org/10.4230/LIPICS.ITCS.2017.49

[10] Andrew Lucas. 2014. Ising formulations of many NP problems. *Frontiers in Physics* 2 (2014), 5.

[11] M. Melucci. 2015. *Introduction to Information Retrieval and Quantum Mechanics*. The Information Retrieval Series, Vol. 35. Springer.

[12] G. Nazli, N. Tonellotto, Y. He, A. Lipani, G. McDonald, C. Macdonald, and I. Ounis (Eds.). 2024. *Advances in Information Retrieval. Proc. 46th European Conference on IR Research (ECIR 2024) – Part V*. Lecture Notes in Computer Science (LNCS) 14612, Springer, Heidelberg, Germany.

[13] Riccardo Nembrini, Costantino Carugno, Maurizio Ferrari Dacrema, and Paolo Cremonesi. 2022. Towards Recommender Systems with Community Detection and Quantum Computing. In *RecSys '22: Sixteenth ACM Conference on Recommender Systems, Seattle, WA, USA, September 18 - 23, 2022*. ACM, 579–585. https://doi.org/10.1145/3523227.3551478

[14] Riccardo Nembrini, Maurizio Ferrari Dacrema, and Paolo Cremonesi. 2021. Feature Selection for Recommender Systems with Quantum Computing. *Entropy* 23, 8 (2021), 970. https://doi.org/10.3390/E23080970

[15] Andrea Pasin, Maurizio Ferrari Dacrema, Paolo Cremonesi, and Nicola Ferro. 2023. qCLEF: A Proposal to Evaluate Quantum Annealing for Information Retrieval and Recommender Systems. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction - 14th International Conference of the CLEF Association, CLEF 2023, Thessaloniki, Greece, September 18-21, 2023, Proceedings (Lecture Notes in Computer Science, Vol. 14163)*. Springer, 97–108. https://doi.org/10.1007/978-3-031-42448-9_9

[16] A. Pasin, M. Ferrari Dacrema, P. Cremonesi, and N. Ferro. 2024. QuantumCLEF – Quantum Computing at CLEF, See [12], 482–489.

[17] Giovanni Pilato and Filippo Vella. 2023. A Survey on Quantum Computing for Recommendation Systems. *Inf.* 14, 1 (2023), 20. https://doi.org/10.3390/info14010020

[18] Somayeh Bakhtiari Ramezani, Alexander Sommers, Harish Kumar Manchukonda, Shahram Rahimi, and Amin Amirlatifi. 2020. Machine Learning Algorithms in Quantum Computing: A Survey. In *2020 International Joint Conference on Neural Networks (IJCNN)*. 1–8. https://doi.org/10.1109/IJCNN48605.2020.9207714

[19] S. Resch Resch and U. R. Karpuzcu. 2022. Benchmarking Quantum Computers and the Impact of Quantum Noise. *ACM Computing Surveys (CSUR)* 54, 7 (September 2022), 142:1–142:35.

[20] M. A. Serrano, J. A. Cruz-Lemus, R. Pérez-Castillo, and M. Piattini. 2023. Quantum Software Components and Platforms: Overview and Quality Assessment. *ACM Computing Surveys (CSUR)* 55, 8 (August 2023), 164:1–164:31.

[21] M. A. Serrano, R. Pérez-Castillo, and M. Piattini (Eds.). 2022. *Quantum Software Engineering*. Springer International Publishing, Germany.

[22] S. Uprety, D. Gkoumas, and D. Song. 2021. A Survey of Quantum Theory Inspired Approaches to Information Retrieval. *ACM Computing Surveys (CSUR)* 53, 5 (September 2021), 98:1–98:39.

[23] Peter JM Van Laarhoven, Emile HL Aarts, Peter JM van Laarhoven, and Emile HL Aarts. 1987. *Simulated annealing*. Springer.

[24] C. J. van Rijsbergen. 2004. *The Geometry of Information Retrieval*. Cambridge University Press, Cambridge, UK.

[25] S. Yarkoni, E. Raponi, T. Bäck, and S. Schmitt. 2022. Quantum annealing for industry applications: introduction and review. *Reports on Progress in Physics* 85, 10 (October 2022), 104001:1–104001:27.