Quantum Computing for Information Retrieval and Recommender Systems

Maurizio Ferrari Dacrema[1], Andrea Pasin[2](✉), Paolo Cremonesi[1], and Nicola Ferro[2]

[1] Politecnico di Milano, Milan, Italy
{maurizio.ferrari,paolo.cremonesi}@polimi.it
[2] University of Padua, Padua, Italy
andrea.pasin.1@phd.unipd.it, nicola.ferro@unipd.it

Abstract. *Quantum Computing (QC)* is a research field that has been in the limelight in recent years. In fact, many researchers and practitioners believe that it can provide benefits in terms of efficiency and effectiveness when employed to solve certain computationally intensive tasks. In *Information Retrieval (IR)* and *Recommender Systems (RS)* we are required to process very large and heterogeneous datasets by means of complex operations, it is natural therefore to wonder whether QC could also be applied to boost their performance. The goal of this tutorial is to show how QC works to an audience that is not familiar with the technology, as well as 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.

Keywords: Quantum Computing · Quantum Annealing · Information Retrieval · Recommender Systems

1 Motivation

Quantum Computing (QC) is a rapidly growing field, involving an increasing number of researchers and practitioners from different backgrounds to develop new methods using quantum computers to perform faster computations. With QC it has been possible to tackle practical problems achieving good results in terms of efficiency and effectiveness [2,4,6,11,12]. Since *Information Retrieval (IR)* and *Recommender Systems (RS)* face big challenges due to the scalability of complex algorithms on top of very large datasets, QC seems very promising but its application to IR and RS has been explored only to a limited extent.

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 through interactive coding sessions teaching participants how to use the cutting-edge quantum annealers to solve realistic problems. The QA paradigm offers a good trade-off between computational power and a low

N. Goharian et al. (Eds.): ECIR 2024, LNCS 14612, pp. 358–362, 2024.
https://doi.org/10.1007/978-3-031-56069-9_47

access barrier since it requires to formulate the task one wishes to solve as an optimization problem using the *Quadratic Unconstrained Binary Optimization (QUBO)* [3] formulation, thus hiding the complexity of the underlying QC system and making it easily accessible even for people without a background in quantum physics. QA has already been applied to IR and RS tasks [1,2,8,9] showing that, although not always superior to classical methods, quantum annealers have matured enough to reliably tackle realistic problems.

The tutorial will also introduce QuantumCLEF[1] [10], a lab organized at CLEF 2024 that offers an infrastructure to develop and evaluate QA algorithms.

Overall, we aim at lowering the often perceived barrier-of-entry to QC and at providing participants the basic knowledge to develop and code QC algorithms.

2 Format

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

2.1 Materials

The tutorial will include slides and Jupyter Notebooks. The materials will be available on GitHub prior to the tutorial and openly available afterwards. A guide on how to create a free D-Wave account, set up the environment, and connect to quantum annealers will be provided before the tutorial. Participants will also have access to the infrastructure made available by QuantumCLEF.

2.2 Outline

The tutorial will be divided into 4 parts, starting from the theoretical aspects of QC, QA and the QUBO problem formulation. Then there will be a practical part where participants can code and solve some problems using quantum annealers.

Part 1: QC Foundations (30 min). The first part consists in a gentle introduction to QC, showing its potential benefits but also limitations. We will also delve into the QA paradigm. It comprises:

- overview and basic understanding of QC and its potential benefits;
- description of the paradigms of Quantum Circuit model and Adiabatic model.
- the relations between the classical optimizations meta-heuristics *Simulated Annealing* [13] and *Quantum-inspired Annealing* [5] to the Adiabatic model;
- how to represent the energy configuration of a quantum system (i.e., Hamiltonian) using the Ising model;

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

– the evolution of a quantum system to a state of minimal energy;
– the similarity between the Hamiltonian of a system and the QUBO representation for optimizing (NP-hard) problems.

Part 2: QUBO Formulation (30 min). This part will show how to represent classical binary optimization problems in QUBO formulation [7]. 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.

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 explain how to use its development and evaluation infrastructure.

Part 4: Hands-On (90 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**).

3 Audience

This tutorial is intended for people coming from IR and RS but also from other fields, such as Machine Learning, Big Data, Operations Research, and Optimization. In fact, QC and QA can be applied to solve problems in different domains and, even if the practical part is focused in using QA for IR and RS systems, the considered problems are very general and common to several research areas.

Targeted audience: Due to the topic's novelty, the target audience is of researchers and industry practitioners mainly belonging to IR and RS.

Prerequisite knowledge: This tutorial will be self-contained and has minimal prerequisite knowledge, mainly consisting on being familiar with the concept of decision and optimization problems. 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[2].

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

4 Tutorial History

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

5 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 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. He is Demo and Late-Breaking Results Co-Chair of RecSys 2024 and was co-organizer of the LERI workshop at RecSys 2023.

Andrea Pasin 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.

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-leader of the Quantum Computing research activities within the Italian National Research Centre for High Performance Computing, Big Data and Quantum Computing. He has also extensive experience in the reproducibility and evaluation of recommender systems. He has been member of the Steering Committee of ACM RecSys, among the most important conferences on Recommender Systems, since 2017. He was General Chair of ACM RecSys 2017 and Program Chair of ACM TvX 2013. Paolo co-authored and presented several tutorials: on "Recommender Systems" at TvX 2012, on "Evaluation of Recommender Systems" at UMAP 2013, on "Cross-domain Recommender Systems" at RecSys 2014, on "Sequence-aware Recommender Systems" at WWW 2019, RecSys 2019 and UMAP 2019. Paolo was also co-organizer of the LERI workshop at RecSys 2023.

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.

References

1. Ferrari Dacrema, M., Felicioni, N., Cremonesi, P.: Optimizing the selection of recommendation carousels with quantum computing. In: RecSys 2021: Fifteenth ACM Conference on Recommender Systems, Amsterdam, The Netherlands, 27 September 2021–1 October 2021, pp. 691–696. ACM (2021). https://doi.org/10.1145/3460231.3478853
2. Ferrari Dacrema, M., Moroni, F., Nembrini, R., Ferro, N., Faggioli, G., Cremonesi, P.: Towards feature selection for ranking and classification exploiting quantum annealers. In: SIGIR 2022: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, 11–15 July 2022, pp. 2814–2824. ACM (2022). https://doi.org/10.1145/3477495.3531755
3. Glover, F.W., Kochenberger, G.A., Du, Y.: Quantum bridge analytics I: a tutorial on formulating and using QUBO models. 4OR **17**(4), 335–371 (2019). https://doi.org/10.1007/s10288-019-00424-y
4. Jaschek, T., Bucyk, M., Oberoi, J.S.: A quantum annealing-based approach to extreme clustering. In: Arai, K., Kapoor, S., Bhatia, R. (eds.) FICC 2020. AISC, vol. 1130, pp. 169–189. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-39442-4_15
5. Kadowaki, T., Nishimori, H.: Quantum annealing in the transverse ising model. Phys. Rev. E **58**, 5355–5363 (1998). https://doi.org/10.1103/PhysRevE.58.5355
6. Kerenidis, I., Prakash, A.: Quantum recommendation systems. In: Papadimitriou, C.H. (ed.) 8th Innovations in Theoretical Computer Science Conference, ITCS 2017, Berkeley, CA, USA, 9–11 January 2017, LIPIcs, vol. 67, pp. 49:1–49:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2017). https://doi.org/10.4230/LIPICS.ITCS.2017.49
7. Lucas, A.: Ising formulations of many NP problems. Front. Phys. **2**, 5 (2014)
8. Nembrini, R., Carugno, C., Ferrari Dacrema, M., Cremonesi, P.: Towards recommender systems with community detection and quantum computing. In: RecSys 2022: Sixteenth ACM Conference on Recommender Systems, Seattle, WA, USA, 18–23 September 2022, pp. 579–585. ACM (2022). https://doi.org/10.1145/3523227.3551478
9. Nembrini, R., Ferrari Dacrema, M., Cremonesi, P.: Feature selection for recommender systems with quantum computing. Entropy **23**(8), 970 (2021). https://doi.org/10.3390/E23080970
10. Pasin, A., Ferrari Dacrema, M., Cremonesi, P., Ferro, N.: qCLEF: a proposal to evaluate quantum annealing for information retrieval and recommender systems. In: Arampatzis, A., et al. (eds.) CLEF 2023. LNCS, vol. 14163, pp. 97–108. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-42448-9_9
11. Pilato, G., Vella, F.: A survey on quantum computing for recommendation systems. Inf. **14**(1), 20 (2023). https://doi.org/10.3390/info14010020
12. Ramezani, S.B., Sommers, A., Manchukonda, H.K., Rahimi, S., Amirlatifi, A.: Machine learning algorithms in quantum computing: a survey. In: 2020 International Joint Conference on Neural Networks (IJCNN), pp. 1–8 (2020). https://doi.org/10.1109/IJCNN48605.2020.9207714
13. Van Laarhoven, P.J., Aarts, E.H., van Laarhoven, P.J., Aarts, E.H.: Simulated Annealing. Springer, Dordrecht (1987). https://doi.org/10.1007/978-94-015-7744-1_2