

Personal Statement

Rodrigo Caetano de Oliveira Rocha

Since the very beginning of my undergraduate course I have been focused on following an academic career. Ever since I first discovered the pleasure of exploring the frontiers of science, I have been really dedicated and passionate with scientific research. Regardless of the field or whether it is a more theoretical or practical research, I truly believe that the scientific method is the best way we humans can develop groundbreaking ideas, theories and technologies capable of changing the world for better.

During my Bachelor's degree at Pontifícia Universidade Católica de Minas Gerais (PUC Minas), I worked as a Teaching Assistant in Differential and Integral Calculus I and also Algorithms and Data Structures III. Those experiences helped me develop teaching skills, which have been very important for me as a professor. I also had many opportunities to work on interesting programming endeavours; for example, some of the projects I developed during the undergraduate course include a fully functional compiler for a subset of the Oberon programming language, a parallel and distributed factorization of big integers, besides working in computer vision and image processing during an undergraduate research scholarship. For those and other reasons, I concluded my Bachelor's degree (*summa cum laude*) in Computer Science, for which I had the privilege of being awarded with the Merit Medal of Honour from PUC Minas. For graduating with honours, and first in my class, I also received the certificate of Outstanding Student from the Brazilian Society of Computer Science (Sociedade Brasileira de Computação - SBC).

After receiving my Bachelor's degree, I applied for a Master's degree at Universidade Federal de Minas Gerais (UFMG), which is considered to be among the best Brazilian universities in Computer Science. My initial research project was on Computer Networks. After the first semester, having published a paper [4] related to Network Security with Software Defined Networks, I changed my research focus to scalable parallel and distributed frameworks for data-intensive (a.k.a. Big-Data) applications. In particular, I focused my Master's research on the development of the Watershed framework, which is a distributed system for processing large-scale data streams, inspired on the dataflow model. I designed a modular abstraction for the stream communication channels, allowing for reusable and programmable streams. I also integrated Watershed with the Hadoop framework, including YARN for job scheduling and cluster resource management, HDFS as a distributed file system, and Zookeeper for distributed coordination and synchronization. This work resulted in a paper published in an international workshop [2]. That paper was selected as one of the best papers on the presented event, so we have been invited to submit an extended version for a Special Issue of the journal *Concurrency and Computation Practice and Experience* (accepted; to appear in the journal) [3].

I also worked in a secondary project with a professor from the Department of Mathematics at UFMG, where we developed a distributed algorithm for cycle detection in large-scale graphs. We were able to derive a correctness proof of this distributed algorithm in addition to also providing an analysis of its theoretical performance. The algorithm was also implemented using the GraphChi framework, which is a disk-based execution framework for large-scale graph processing. This work resulted in a published paper [5] that was mentioned with honours, for being among the five best papers in the presented symposium (for which I earned a certificate).

After finishing my Master's degree, I have been working at PUC Minas as an Assistant Professor, where I have been teaching several Computer Science and Information Systems courses, including Parallel Programming and Theory and Algorithms in Graphs. I have also been collaborating as a co-adviser of students working as undergraduate researches, which resulted in two papers [1, 6] that was published in an undergraduate research workshop, one of which [1] was also mentioned with honours (earning a certificate for being among the best papers in the workshop). Besides those papers, I have also submitted to the journal

Concurrency and Computation Practice and Experience another paper about an automatic tiling technique, using genetic algorithms, applied to stencil computations, allowing for processing large inputs on GPUs (under revision). In addition, I have been in three examining boards of undergraduate final projects, which was an extremely valuable experience for me.

With the purpose of building a solid academic career, I intend to pursue my postgraduate study at the Lancaster University. My interest in School of Computing and Communications of the Lancaster University is due to its prestigious and worldwide recognized excellence in research. Specifically, the Ph.D. position offered by Dr. Zheng Wang at Lancaster University is clearly related to the postgraduate research that I intend to explore. A Ph.D. program in the area of automatic parallelisation would provide me with an excellent opportunity to develop myself as a researcher. If accepted, I will certainly dedicate myself to make the best of such an opportunity of being part of an internationally leading group in computing systems.

References

- [1] Alyson Deives Pereira, Sérgio Vitarelli Silva, Rodrigo Caetano de Oliveira Rocha, Luís Fabrício Wanderley Góes, and Márcio Castro. Stencilbench: Um benchmark sintético para avaliação de frameworks do padrão estêncil. In *XVI Workshop de Iniciação Científica em Arquitetura de Computadores e Computação de Alto Desempenho (WSCAD-WIC)*, Santa Catarina - Brazil, Oct. 2015. SBC.
- [2] Rodrigo Caetano Rocha, Renato Ferreira, Wagner Meira Jr., and Dorgival Guedes. Watershed reengineering: Making streams programmable. In *26th IEEE International Symposium on Computer Architecture and High Performance Computing Workshop, SBAC-PAD Workshop 2014*, pages 120–125, Paris - France, 2014. IEEE.
- [3] Rodrigo Caetano Rocha, Bruno Hott, Vinícius Dias, Renato Ferreira, Wagner Meira Jr., and Dorgival Guedes. Watershed-ng: an extensible distributed stream processing framework. (*To appear in*) *Concurrency and Computation: Practice and Experience*.
- [4] Rodrigo Caetano Rocha, Humberto Marques-Neto, and Dorgival Guedes. SMTD: Uma aplicação de redes definidas por software no contexto de redes domésticas. In *SBRC 2013 - Workshop de Pesquisa Experimental da Internet do Futuro (WPEIF)*, Brasília, May 2013. SBC.
- [5] Rodrigo Caetano Rocha and Bhalchandra Digambar Thatte. Distributed cycle detection in large-scale sparse graphs. In *XLVII SBPO 2015 - Simpósio Brasileiro de Pesquisa Operacional*, Pernambuco - Brazil, Aug. 2015. SBPO.
- [6] Joao Saffran, Rodrigo Caetano de Oliveira Rocha, and Luís Fabrício Wanderley Góes. Algoritmo de regras de associação paralelo para arquiteturas multicore e manycore. In *XVI Workshop de Iniciação Científica em Arquitetura de Computadores e Computação de Alto Desempenho (WSCAD-WIC)*, Santa Catarina - Brazil, Oct. 2015. SBC.