DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY

DECLARATION

I hereby declare that the project entitled "**IMPLEMENTING PARALLEL PSO ALGORITHM USING MAPREDUCE ARCHITECTURE**" submitted by me in the partial fulfillment of the requirements for the award of the degree of Master of Technology (Software Engineering) of Delhi Technological University is record of my own work carried under the supervision and guidance of **Dr. Kapil Sharma.**

To the best of my knowledge this project has not been submitted to Delhi Technological University or any other University or Institute for the award of any degree.

PRATEEK KUMAR VARSHNEY
[2K13/SWE/11]

ACKNOWLEDGEMENT

In the sense of great pleasure and satisfaction I present this project entitled "**IMPLEMENTING PARALLEL PSO ALGORITHM USING MAPREDUCE ARCHITECTURE**".

The completion of this project is no doubt a product of invaluable support and contribution of number of people.

I would like to express my sincere thanks to my guide **Dr. Kapil Sharma** (Associate Professor, Department of Computer Science and Engineering) for his continuous help and valuable suggestions and also providing encouraging environment, without which my project and its documentation would not have been possible.

The completion of any task is not only the reward to the person activity involved in accomplishing it, but also the person involved in inspiring and guiding. I am grateful to my friends and family for their constant motivation and comments that has helped me to complete this report.

PRATEEK KUMAR VARSHNEY

[2K13/SWE/11]

ABSTRACT

Optimization is the problem of finding minimum or maximum of a given objective function relative to some set, often representing a range of choices available in a certain situation. Particle Swarm Optimization (PSO) is a simple and effective evolutionary algorithm, but it may take a reasonable time to optimize complex objective functions which are deceptive or expensive. To avoid being trapped in local optima, Particle Swarm Optimization requires extensive exploration for multimodal and multidimensional functions. Expensive functions whose computational complexity may arise from dependence on detailed simulations or large datasets, takes a long time to evaluate. For such functions PSO must be parallelized to use multiprocessor systems and clusters efficiently. Parallelization of PSO can lead to scalable speedup in performance. PSO can be naturally expressed in Google's MapReduce framework to develop a simple and robust parallel implementation.

To improve optimization of difficult objective functions and to improve parallel performance, modifications could be made to this flexible implementation of the algorithm. In the proposed work the classic Particle Swarm Optimization Algorithm has been implemented on Big Data platform Hadoop using MapReduce Architecture. The algorithm has been applied to optimize parameters of basic COCOMO Model need to calculate effort of the project.

The experiments show that the Hadoop could carry out big data calculations which normal serial PSO could not. The proposed model would have better efficiency for intensive computational functions.

TABLE OF CONTENTS

APPENDIX A

LIST OF FIGURES

APPENDIX B

LIST OF TABLES

REFERENCES

1. Ahmed, Alam, A., & Jamil. (2014). Hadoop Architecture and Its Issues. *2014 International Conference on Computational Science and Computational Intelligence* (pp. 288-291). IEEE.

2. Ahmed, H., Ismail, M. A., & Hyder, M. F. (2014). Performance Optimization of Hadoop Cluster Using Linux Services. *IEEE* (pp. 167-172). IEEE.

3. Belal, M., & Ghazawi, T. E. (2004). PARALLEL MODELS FOR PARTICLE SWARM OPTIMIZERS. *The International Journal of Intelligent Computing and Information Sciences* .

4. Dubey, Dwivedi, K., & Kumar, S. (2014). Analytical Review on Hadoop Distributed File System. *5th International Conference- Confluence The Next Generation Information Technology Summit* (pp. 174-181). IEEE.

5. Eberhart, Kennedy, J., & Russell. (1995). Particle Swarm Optimization. *IEEE*, (pp. 1942-1948).

6. Eberhart, R. C., & Shi, Y. (2001). Particle Swarm Optimization: Developments, Applications and Resources. *IEEE* (pp. 81-86). IEEE.

7. Eiben, A. E., & Smith, J. E. (2003). *Introduction to Evolutionary Computing.*

8. Ghemawat, Dean, J., & Sanjay. (2008, January). MapReduce: Simplified Data Processing on Large Clusters. *COMMUNICATIONS OF THE ACM* , pp. 107-113.

9. Grolinger, K., Hayes, M., Higashino, W. A., L'Heureux, A., Allison, D. S., & Capretz, M. A. (2014). Challenges for MapReduce in Big Data. *2014 IEEE 10th World Congress on Services* (pp. 182-189). IEEE.

10. *http://www.coreservlets.com/hadoop-tutorial/*. Retrieved 2015, from coreservlets.com: http://www.coreservlets.com

11. *http://www.tutorialspoint.com/hadoop/*. Retrieved 2015, from tutorialspoint: http://www.tutorialspoint.com

12. Jagtap, & D., D. (2014). Big Data using Hadoop. *International Journal of Engineering Research and General Science Volume 2, Issue 6,* .

13. Karun, A. K., & Chitharanjan, K. (2013). A Review on Hadoop – HDFS Infrastructure Extensions. *2013 IEEE Conference on Information and Communication Technologies (ICT 2013)* (pp. 132-137). IEEE.

14. Kečo, D., & Subasi, A. (2012). Parallelization of genetic algorithms using Hadoop Map/Reduce. *SOUTHEAST EUROPE JOURNAL OF SOFT COMPUTING* , 56-59.

15. Kennedy, M. C., & James. (2002, February). The Particle Swarm—Explosion, Stability, and Convergence in a Multidimensional Complex Space. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION* .

16. Lin, C. -Y., Pai, Y. -M., Tsai, K. -H., & Wen, C. H.-P. (2013). Parallelizing Modified Cuckoo Search on MapReduce Architecture. *JOURNAL OF ELECTRONIC SCIENCE AND TECHNOLOGY, VOL. 11, NO. 2* , 115-123.

17. Ludwig, S. A. (2014). MapReduce-based Optimization of Overlay Networks using Particle Swarm Optimization. *GECCO'14, ACM*, (pp. 1031-1038).

18. McNabb, A. W., Monson, C. K., & D., K. (2012). Parallel PSO Using MapReduce. *IEEE*.

19. Patel, A. B., Birla, M., & Nair, U. (2012). Addressing Big Data Problem Using Hadoop and Map Reduce. *2012 NIRMA UNIVERSITY INTERNATIONAL CONFERENCE ON ENGINEERING, NUiCONE-2012*, (pp. 1-5).

20. Shelley, P. (n.d.). *big-data-spectrum.pdf*. Retrieved from http://www.infosys.com/: http://www.infosys.com/cloud/resource-center/Documents/big-data-spectrum.pdf

21. SINANC, SAGIROGLU, S., & Duygu. (2013). Big Data: A Review. *IEEE* (pp. 42-47). IEEE.

22. Singh, K. K., & Ravinder. (2014). Hadoop: Addressing Challenges of Big Data. *IEEE International Advance Computing Conference (IACC)* (pp. 686-689). IEEE.

23. Wang, J., Yuan, D., & Jiang, M. (2012). Parallel K-PSO Based on MapReduce. *IEEE*, (pp. 1203-1208).