DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY

DECLARATION

I hereby declare that the project entitled "**APPLICATION OF ARTIFICIAL BEE COLONY ALGORITHM USING HADOOP**" 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.

SANJAY KUMAR
[2K13/SWE/20]

ACKNOWLEDGEMENT

In the sense of great pleasure and satisfaction I present this project entitled "**APPLICATION OF ARTIFICIAL BEE COLONY ALGORITHM USING HADOOP**".

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.

SANJAY KUMAR

[2K13/SWE/20]

ABSTRACT

Optimization problem is a very common problem of finding the best solution from all feasible solutions. It is also a very common problem both in theoretical and practical scenes. In most cases optimization problem is also been classified as a NP-Hard problem. In today's era, time and computing resources are most common limitations. People often choose such metaheuristic methods that evaluate the optimal value in a limited time and with a limited computing resource. Artificial Bee Colony (ABC) is newly proposed metaheuristic algorithm which is suggested very compromising in recent work. As compared with the other metaheuristic algorithms, ABC is very effective for high dimensional problems. ABC can also solve various optimization problems with complex nonlinearity because it is a fast and robust algorithm. In this work, we have proposed a parallel implementation of Artificial Bee Colony algorithm for analysis of big data by using MapReduce paradigm. During the implementation of this work, Hadoop has been used as the backend MapReduce platform. As the result shows our ABC algorithm is very effective and scalable.

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. Al-Madi, N., Aljarah, I., & Ludwig, S. A. (2014). Parallel Glowworm Swarm Optimization Clustering Algorithm based on MapReduce. IEEE.

4. Aridhi, S., Lacomme, P., Ren, L., & Vincent, B. (2015). A MapReduce-based approach for shortest path problem in large-scale networks. *Engineering Applications of Artificial Intelligence 41* (pp. 151-165). ELSEVIER.

5. Armour, F., Kaisler, S., & Espinosa, A. (2015). Introduction to Big Data Analytics: Concepts, Methods, Techniques and Applications. *48th Hawaii International Conference on System Sciences* (p. 886). IEEE.

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

7. Bing, L., & Chan, K. C. (2014). A Paralleled Big Data Algorithm with MapReduce Framework for Mining Twitter Data. *Fourth International Conference on Big Data and Cloud Computing* (pp. 121-128). IEEE.

8. Demchenko, Y., Laat, C. d., & Membrey, P. (2014). Defining Architecture Components of the Big Data Ecosystem. (pp. 104-112). IEEE.

9. 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.

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

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

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

13. Elagib, S. B., Najeeb, A. R., Hashim, A. H., & Olanrewaju, R. F. (2014). Big Data Analysis Solution Using MapReduce Framework. *5th International Conference on Computer & Communication Engineering* (pp. 127-130). IEEE.

14. Elanthiraiyan, N., & Arumugam, C. (2014). Parallelized ACO Algorithm for Regression Testing Prioritization in Hadoop Framework. *International Conference on Advanced Communication Control and Computing Technologies (lCACCCT* (pp. 1568-1571). IEEE.

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

16. Gohil, P., Garg, D., & Panchal, B. (2014). A Performance Analysis of MapReduce Applications on Big Data in Cloud based Hadoop. *S.A.Engineering College*. Chennai, Tamil Nadu, India: IEEE.

17. 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.

18. Gupta, A. (2015). Big Data Analysis Using Computational Intelligence and Hadoop: A Study. *2nd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1397-1401). IEEE.

19. *http://www.coreservlets.com/hadoop-tutorial/*. (n.d.). Retrieved 2015, from coreservlets.com: http://www.coreservlets.com

20. *http://www.tutorialspoint.com/hadoop/*. (n.d.). Retrieved 2015, from tutorialspoint: http://www.tutorialspoint.com

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

22. KARABOGA, D. AN IDEA BASED ON HONEY BEE SWARM FOR NUMERICAL OPTIMIZATION.

23. 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.

24. Kaswan, K. S., Choudhary, S., & Sharma, K. (2015). Applications Of Artificial Bee Colony optimization Technique : Survey. *2nd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 1660-1664). IEEE.

25. Ke, X., Jin, H., Xie, X., & Cao, J. (2015). A Distributed SVM Method based on the Iterative MapReduce. *9th International Conference on Semantic Computing (IEEE ICSC 2015)* (pp. 116-119). IEEE.

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

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

28. Kojima, M., Nakano, H., & Miyauchi, A. (2013). An Artificial Bee Colony Algorithm for Solving Dynamic Optimization Problems. *Congress on Evolutionary Computation* (pp. 2398-2405). Cancún, México: IEEE.

29. 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.

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

31. Ludwig, S. A. (2014). MapReduce-based Optimization of Overlay Networks using Particle Swarm Optimization. *GECCO'14, July 12–16, 2014*, (pp. 1031-1038). BC, Canada: Vancouver.

32. Maitrey, S., & Jha, C. K. (2015). Handling Big Data Efficiently by using Map Reduce Technique. *International Conference on Computational Intelligence & Communication Technology* (pp. 703-708). IEEE.

33. Manikandan, S. G., & Ravi, S. (2014). Big Data Analysis using Apache Hadoop. IEEE.

34. Marozzo, F., Talia, D., & Trunfio, P. (2011). A Framework for Managing MapReduce Applications in Dynamic Distributed Environments. *19th International Euromicro Conference on Parallel, Distributed and Network-Based Processing* (pp. 150-158). IEEE.

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

36. Nandimath, J., Patil, A., Banerjee, E., Kakade, P., & Vaidya, S. (2013). Big Data Analysis Using Apache Hadoop. *IRI 2013, August 14-16, 2013, San Francisco, California, USA* (pp. 700-703). IEEE.

37. 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).

38. 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

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

40. Singh, A., & Sundar, S. (2011). An artificial bee colony algorithm for the minimum routing cost spanning tree problem. (pp. 2489-2499). Springer.

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

42. SUBOTIC, M., TUBA, M., & STANAREVIC, N. Parallelization of the Artificial Bee Colony (ABC) Algorithm. *RECENT ADVANCES in NEURAL NETWORKS, FUZZY SYSTEMS & EVOLUTIONARY COMPUTING* (pp. 191-196). ISSN.

43. Tomar, R. S., Singh, S., Verma, S., & Tomar, G. S. (2013). A Novel ABC Optimization Algorithm for Graph Coloring Problem. *5th International Conference on Computational Intelligence and Communication Networks* (pp. 257-261). IEEE.

44. Tomar, R. S., Singh, S., Verma, S., & Tomar, G. S. (2013). A Novel ABC Optimization Algorithm for Graph Coloring Problem. *5th International Conference on Computational Intelligence and Communication Networks* (pp. 257-261). IEEE.

45. Verma, J. P., Patel, B., & Patel, A. l. (2015). Big Data Analysis: Recommendation System with Hadoop Framework. *International Conference on Computational Intelligence & Communication Technology* (pp. 92-97). IEEE.

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

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

48. Wu, B., Wu, G., & Yang, M. (2012). A MapReduce based Ant Colony Optimization Approach to Combinatorial Optimization Problems. *8th International Conference on Natural Computation (ICNC 2012)* (pp. 728-732). IEEE.

49. Xu, X., Ji, Z., Yuan, F., & Liu, X. (2014). A Novel Parallel Approach of Cuckoo Search using MapReduce. *International Conference on Computer, Communications and Information Technology (CCIT 2014)* (pp. 114-117). Atlantis Press.

50. Yang, J., & Li, X. (2013). MapReduce based Method for Big Data Semantic Clustring. *International Conference on Systems, Man, and Cybernetics* (pp. 2814-2819). IEEE.

51. Zhou, C. (2010). 1113-1114. *GECCO'10* (p. Fast Parallelization of Differential Evolution Algorithm Using MapReduce). ACM.