UNIVERSITY TEKNOLOGI MARA

ANT COLONY ALGORITHM FOR TEXT CLASSIFICATION IN MULTICORE-MULTITHREAD ENVIRONMENT

AHMAD NAZMI BIN FADZAL

Thesis submitted in fulfilment
of the requirements for the degree of
Master of Science

Faculty of Computer and Mathematical Sciences

January 2017

CONFIRMATION BY PANEL OF EXAMINERS

I certify that a Panel of Examiners has met on 10th August 2016 to conduct the final examination of Ahmad Nazmi Bin Fadzal on his Master of Computer Science thesis entitled "Ant Colony Algorithm on Text Classification in Multicore-Multithread Environment" in accordance with Universiti Teknologi MARA Act 1976 (Akta 173). The Panel of Examiners recommends that the student be awarded the relevant degree. The panel of Examiners was as follows:

Seripah Awang Kechil, PhD
Associate Professor
Fakulti Sains Komputer dan Matematik
University Teknologi MARA
(Chairman)

Shuzlina Binti Abdul Rahman, PhD
Associate Professor
Faculty of Computer and Mathematical Sciences
Universiti Teknologi MARA
(Internal Examiner)

Mohd Zakree Bin Ahmad Nazri, PhD
Associate Professor
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
(External Examiner)

DR MOHAMMAD NAWAWI
DATO' HAJI SEROJI, PhD
Dean
Institute of Graduate Studies
Universiti Teknologi MARA
Date : 19th January 2017

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the results of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student   :   Ahmad Nazmi Bin Fadzal

Student I.D. No.   :   2012308179

Programme   :   Master of Computer Science - CS750

Faculty   :   Computer and Mathematical Science

Thesis Title   :   Ant Colony Algorithm for Text Classification in Multicore-Multithread Environment

Signature of Student   :

Date   :   January 2017

ABSTRACT

In the age of wide digital usage, text classification is one of the significant prominent attribute required in order to automatically arrange emails, articles, and other textual data in an organization. Unclassified data can lead to slower data retrieval thus a reliable method is required to effectively retrieve data efficiently and in systematic manner. Ant Colony Optimization (ACO) is a bio-inspired technique that was introduced to solve Non-Polynomial hard problem of high text data dimension that is similar to Traveling Salesman Problem (TSP) using probabilistic way. Pheromone concept is the main criterion that distinguish ACO to other algorithms. Based on the concept, pheromone saturation is used to combine stackable solution pattern that is discovered while straying to different term node to build a path. ACO classification accuracy is compared to Genetic Algorithm classifier which also a wrapper method. On integration of the technique, ACO is proposed to work in a multicore-multithread environment to gain additional execution time advantage. In multicore-multithread environment, the adjustment aims to make artificial ants communicate across the physical core of processor. As a trade to the investment for more computing power, the execution time reduction is expected to show an improvement without compromising the original classification accuracy. The unthreaded and multicore-multithreaded version of ACO was experimented and compared in term of accuracy and execution time. It was found that the result show a positive improvement.

TABLE OF CONTENTS

Page