AUTOMATED RULE GENERATION
FOR
COMPLEX EVENT PROCESSING

DILUM CHANDIMA VITHANA

(158254P)

Degree of Master of Science

Department of Computer Science and Engineering

University of Moratuwa
Sri Lanka

February 2019

AUTOMATED RULE GENERATION
FOR
COMPLEX EVENT PROCESSING

DILUM CHANDIMA VITHANA

(158254P)

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree
Master of Science

Department of Computer Science and Engineering

University of Moratuwa
Sri Lanka

February 2019

DECLARATION

I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant the University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic or any other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:……………………...   **Date:** ………………………

Name: D.C Vithana

The above candidate has carried out research for the Masters/MPhil/Ph.D. thesis/dissertation under my supervision.

Signature of the supervisor:…………………...   **Date:** ………………………

Name: Dr. Surangika Ranathunga

Abstract

The key concept of Complex Event Processing (CEP) is setting up accurate queries to pick up the important events. Since all CEP engines support SQL like queries, this rule setup requires some technical skills plus domain knowledge. The best solution to address this issue is to automate the query generation of CEP. Existing automated query generation methodologies are computationally expensive and are not fully automated processes. This study addresses the above two issues by proposing a shapelet based approach. This new approach is not computationally expensive, and it is a fully automated process with zero manual user intervention required. The proposed method uses the computationally efficient algorithm called Fast Shapelet Selection (FSS) algorithm. This FSS algorithm is used to extract the shapelets from data set. Then extracted shaplets are used to generate CEP queries. This proposed method can be used analyze to multivariant time series and this is more efficient than previously proposed shapelet based approaches.

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Surangika Ranathunga for the valuable guidance given throughout to successfully complete this research. I would also like to thank Department of Computer Science and Engineering, Faculty of Engineering, University of Moratuwa Sri Lanka for their support.

Finally, I must express my very profound gratitude to my family for providing continuous support and encouragement.

TABLE OF CONTENTS

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AEP	Admissible Entropy Pruning
BF	Brute Force Algorithm
CEP	Complex Event Processing
DBMS	Database Management Systems
EDL	Event Description Language
FS	Fast Shapelet Discovery
FSS	Fast Shapelet Selection Algorithm
GRSF	Generalized Random Shapelet Forest
HMM	Hidden Markov Model (HMM)
IDP	Important Data Point
LS	Learning Shapelet Algorithm
nHMM	noise Hidden Markov Model
PLR	Piecewise Linear Representation (PLR)
PLR_IDP	Piecewise Linear Representation based on Important Data Points
RS	Random Shapelet (RS)
ST	Shapelet Transformation Algorithm (ST)
SDEA	Subsequence Distance Early Abandon (SDEA)
LFDP	Local Farthest Deviation Points (LFDP)