**Performance Improvement in Web caching along with prefetching**

**M.Tech. Thesis**

*by*

Bheri Srihari

**Roll Number: 14M505**



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY, HAMIRPUR

HIMACHAL PRADESH-177005 (INDIA)

July, 2015

**Performance Improvement in Web caching along with prefetching**

*A* Thesis s*ubmitted*

*in partial fulfilment of the requirements for the award for the Degree of*

**Master of Technology**

*by*

Bheri Srihari

**Roll Number: 14M505**

***Under the guidance of***

**Dr. T.P. SHARMA**



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY, HAMIRPUR

HIMACHAL PRADESH -177005 (INDIA)

June, 2016



**NATIONAL INSTITUTE OF TECHNOLOGY**

**HAMIRPUR (HP)**

CANDIDATE’S DECLARATION

I hereby certify that the work being presented in the Dissertation

Report entitled “**PERFORMANCE IMPROVEMENT IN WEB CACHING**

**ALONG WITH PREFETCHING**”, in partialfulfilment of the requirements for the awardof degree of the Master of Technology in Computer Science and Engineering submitted to the Department of Computer Science and Engineering at National Institute of Technology, Hamirpur (H.P.) is an authentic record of my dissertation carried out during a period from June 2015 to June 2016 under the supervision of **Dr. T. P. Sharma**, Associate Professor, Department of Computer Scienceand Engineering, NIT Hamirpur.

The matter presented in this dissertation has not been submitted by me or anyone else in any other university/institute for award of any degree.

(**BHERI SRIHARI)**

This is to certify that above statement made by the candidate is correct to the best of my knowledge.

Date:

**(Dr. T. P. Sharma)**

**Associate Professor**

The M.Tech.Viva-Voce Examinations of Bheri Srihari, Research

Scholar, has been held on………….

Signature of Supervisor Signature of External Examiner

i

**ACKNOWLEDGEMENTS**

First and Foremost I would like to express my deepest gratitude to my guide and mentor, **Dr. T. P. Sharma**, Assistant Professor, Computer Science and Engineering Department, National Institute of Technology, Hamirpur, under whose firm guidance, motivation and vigilant supervision I succeeded in completing my work. Without his guidance and persistent help this dissertation would not have been possible He infused into me the enthusiasm to work on this topic. His tolerant and supportive nature accepted my shortcomings and he synergized his impeccable knowledge with my curiosity to learn into this fruitful result.

I would like to offer my special thanks to **Dr. Narottam Chand,** Head, Computer Science and Engineering Department, who is always very helpful and constructive. Advice and comments given during seminar has been a great help to me. I express my gratitude to the Almighty for giving me strength and courage during the process of dissertation work and showering me with all the blessings.

I would sincerely thank **Dr. Naveen Chauhan**, for providing support and encouragement throughout my dissertation work.

I wish to express heartiest thanks to my friends Bheri Bala murali krishna and Satish Thulva for their helping and supportive nature and making the environment so comfortable.

Above all, I am feeling short of words to express my heartfelt gratitude to my affectionate parents whose firm belief and confidence in me gives me encouragement in every sphere of life and by whose efforts and blessings I have reached here. I extend thanks to my brothers whose love and affection supported me a lot.

(**BHERI SRIHARI**)

ii

**ABSTRACT**

In today’s day to day life, use of internet is growing exponentially, which causes more network traffic and delay. When the requests are served by nearby caches rather than distant servers, server loads and network traffic decrease and transactions are faster. Web caching along with prefetching is a powerful technique which is mainly used to reduce network traffic and increase performance of web. Web caching stores some of the web objects temporarily near to the client and prefetching fetches next user requests based on some criteria like past history and stores them in the cache near to the clients.

There are various issues while caching the objects like staled data, non-cacheable data and dynamically generated data. In this thesis I consider the problems like staled data, pre-fetch miss and optimal cache replacement mechanism. The main aim of the proposed algorithms is efficient use of web by reducing the band width usage and providing the data to the users with less amount of delay.

The work presents a web cache replacement algorithm if the cache size is full and a prediction algorithm to pre-fetch the next sub sequent requests based on the past user requests. The web caching and prefetching algorithms has been simulated based on discrete event simulation.

**TABLE OF CONTENTS**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Acknowledgements** | | **i** |
|  | **Abstract** | | **iii** |
|  | **Table of Contents** | | **iv** |
|  | **List of Figures** | | **vii** |
|  | **List of Tables** | | **viii** |
|  | **List of Abbreviations** | | **ix** |
| **1** | **INTRODUCTION** | | **1** |
|  | 1.1 | General Introduction | 1 |
|  | 1.2 | Web caching | 1 |
|  | 1.3 | Caching architectures | 2 |
|  | 1.4 | Desirable properties of web caching system | 3 |
|  | 1.5 | Dynamic data caching | 5 |
|  | 1.6 | Cache replacement mechanism | 5 |
|  | 1.7 | Metrics for Comparing web caching Techniques | 6 |
|  | 1.8 | Web prefetching | 8 |
|  | 1.9 | Integrating Web caching and prefetching | 10 |
|  | 1.10 | Challenges in web caching | 11 |
|  | 1.11 | Challenges in prefetching | 11 |
|  | 1.12 | Motivation of the work | 11 |
|  | 1.13 | Problem statement and objectives | 12 |
|  | 1.14 | Thesis outline | 12 |
| **2** | **LITERATURE REVIEW** | | **13** |
|  | 2.1 | Cache algorithms | 13 |
|  | 2.2 | Intelligent Web caching algorithms | 17 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | 2.3 | Prefetching algorithms | 18 |
|  | 2.4 | Integrating Web caching and prefetching | 21 |
| **3** | **WEB CACHING AND PREFETCHING ALGORITHMS** | | **23** |
|  | 3.1 | Proposed approach | 23 |
|  | 3.2 | Web caching | 24 |
|  | 3.3 | Web prefetching | 25 |
| **4** | **SIMULATION AND RESULTS** | | **26** |
|  | 4.1 | Introduction about Discrete event simulation | 26 |
|  | 4.2 | Simulation Assumptions | 27 |
|  | 4.3 | Events present in the simulation | 28 |
|  | 4.4 | Modules created in java |  |
|  | Results | |  |
| **5** | **CONCLUSION AND FUTURE SCOPE** | | **43** |
|  | 5.1 | Conclusion | 43 |
|  | 5.2 | Future Scope | 43 |
|  | **References** | | **44** |

v

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| Figure 1.1 | Types of web cache | 2 |
| Figure 1.2 | Web prefetching | 8 |
| Figure 3.1 | Proxy server cache along with prefetching | 23 |
| Figure 3.2 | Web sessions | 25 |
| Figure 3.3 | Prediction matrix | 25 |
| Figure 4.1 |  | 39 |
| Figure 4.2 |  | 40 |
| Figure 4.3 |  | 41 |
| Figure 4.4 |  | 42 |
|  |  |  |

vi

**LIST OF TABLES**

|  |  |  |
| --- | --- | --- |
| Table 2.1 | Comparison of different prefetching schemes at different locations | 9 |
|  |  |  |
| Table 2.2 | Comparison of conventional web cache algorithms | 16 |

vii

**LIST OF ABBREVIATIONS**

|  |  |  |
| --- | --- | --- |
| URL | - | Uniform Resource Locator |
| HR | - | Hit Ratio |
| BHR | - | Byte Hit Ratio |
| LSR | - | Latency Saving Ratio |
| LRU | - | Least Recently Used |
| LFU | - | Least frequently used |
| GDS | - | Greedy Dual Size |
| ANN | - | Artificial Neural Network |
| PSO | - | Particle Swam Optimization |
| GA | - | Generic Algorithm |

viii