

# **Informatics Institute of Technology**

In collaboration with

**University of Westminster, UK.**

## **AutoDbIndexer - Expert System for Automated Database Indexing**

A dissertation by

**Deepadarshan Gnanajothi (2013164)**

Supervised By

**Mr. Saman Hettiarachchi**

Submitted in partial fulfillment of the requirements for the

BSc (Hons) Software Engineering degree

Department of Computing

© The copyright for this project and all its associated products resides with  
Informatics Institute of Technology.

**ABSTRACT**

The collection, processing and analysis capacity of large amounts of data is crucial for the purpose of extrapolating new learning in business, logical, and medicinal applications. In the era of Big Data, Database Management Systems are a vital element which acts as a repository where the data is aggregated and stored. For such valuable databases, database tuning is paramount. Tuning a database to improve performance is historically a difficult task in database administration. A research article by ACM SIGMOD states that tuning expense is evaluated to be right around half of the aggregate cost of a database, and numerous DBAs spend about a quarter of their time on database tuning activities. Database Tuning is not a one-time task, majority of the real-world applications' workloads are highly dynamic, unpredictable and subject to change based on various factors such as time, weather, trends, politics, business processes etc. Therefore, tuning configurations such as indexes soon become invalid and cause performance issues. It is DBAs' job to continuously watch the database for performance issues and tune it frequently to maintain the performance level. The process is very time consuming and includes several steps. It is impossible for a human operator to study and analyze thousands of performance statistics, historical tuning data and workload patterns in order to come up with optimal tuning recommendations. A technique called online-indexing was introduced to solve this problem by autonomously applying indexes based on the workloads. But the existing online-indexing systems has significant drawbacks in index selection and overhead management. Therefore, a continuous workload monitoring and automated tuning system called AutoDbIndexer is proposed, which unremittingly monitors the changing workloads of databases and applies indexes based on database access trends. The solution improves the query execution time by 28%, creating negligible overhead to the DBMS.

Subject Descriptors:

• **Information systems~Database utilities and tools** • **Information systems~Autonomous database administration** • **Information systems~Indexed file organization** • *Information systems~Nearest-neighbor search* • *Information systems~Expert systems* • **Computing methodologies~Supervised learning by regression** • *Computing methodologies~Vagueness and fuzzy logic* • *Theory of computation~Oracles and decision trees* • *Theory of computation~Pattern matching*

Key Words:

**Automated Database Tuning, Online Indexing, Index Recommendation Engine, Artificial Intelligence, Expert System**