International Journal of Innovative
Computing, Information and Control
Volume **5**, Number **5**, May **2009**

 ISSN 1349-4198
pp. **1265–1273**

LOAD SHEDDING FOR WINDOWED NON-EQUIJOIN OVER SENSOR DATA STREAMS

Jiadong Ren and Cong Huo

College of Information Science and Engineering
Yanshan University
Qinhuangdao City, Hebei Province, P. R. China
jdren@ysu.edu.cn; hcedu@163.com

Received December 2007; revised May 2008

Abstract. *Many load shedding approaches for window equijoin over data streams have been proposed. However, load shedding method for non-equijoin over sensor data streams is not mentioned at all. Since non-equijoin queries are popular in wireless sensor network applications, it should be processed effectively when CPU or memory resources are limited to keep up with input streams. In this paper, we propose a novel load shedding algorithm for sliding window non-equijoin over sensor data streams under limited CPU resources. A clustering-based Range-Count AVL tree index (CRCA-index) is constructed for each sensor stream by clustering technique. And the statistical information of non-equijoin attribute can be obtained accurately. With CRCA-index, the contribution towards the join results of each tuple can be estimated exactly. A part of new tuples are selected without performing the join operation, since these tuples will produce less non-equijoin results. Experimental results show that the CRCA-index is low storage cost and fast search time, and our approach is more efficient than other existing approaches for sensor streams non-equijoin query.*

Keywords: Wireless sensor network, Data streams, Non-equijoin query, Load shedding algorithm, Window index

1. **Introduction.** Data streams have been emerged in many applications, such as network traffic monitoring, sensor data analysis and financial fraud. Queries over data streams are required to be processed in real time. Since data streams and streaming query is continuous, the management of data streams has become a challenging problem for current processing systems [1]. An innovative broadcasting scheme is developed, which aimed at higher reliability in broadcasting for ad hoc wireless networks [2]. A promising solution for this problem is to design data stream management systems, exemplified by STREAM [3] and Aurora [4].

Streaming join query is a complex and important operation in the context of data stream systems. For example, users can obtain the temperature information with the same humidity from two sensor streams, which are generated by two sensor nodes at the given location in the sensor network. Unbounded memory, communications and CPU resources will be required for processing the join over data streams. Punctuated predication has been introduced to process streaming join, but only a few streams in applications have this predication [5]. Another promising solution is to use window predication, including sliding, landmark and snapshot window. A new tuple from one stream will be joined with tuples in the window of another stream in the windowed join algorithms. A unit-time-basis cost model is introduced to analyze the performance of streams join algorithms, and four algorithms are investigated for evaluating moving window joins over pairs of streams