Stream Data Mining: A Survey

The paper discusses the survey done on “Stream data mining”. It explains the characteristics of a stream data. It also discussed the methodologies associated with solving the issues related to the stream data. The systems are replaced with special characteristics in order to mine the continuous high speed data streams. Which is not the case in traditional data mining. The paper discusses the algorithms that are used to mine data from a stream. Talks about which applications support the methodologies of this type of mining. It finally concludes to address the issues regarding the current research on data streams.

A data stream is a continuously flowing data, we cannot assume the order of the data and finally the length of the data stream is unbounded.

The data stream is analyzed using techniques such as sampling, sketching and synopsis methodologies. Some of the newer algorithms include approximation algorithm, sliding window. The newer algorithms addressed the computational challenges of the data stream.

The paper talks about the current state of research such as “the need of real world applications, including mobile devices” The survey is done based on a classification which is based on classification, clustering, aggregation, association, frequency counting and time series analysis.