

## **ANALYZING HIGH DIMENSIONAL DATA MINING IN TIME SERIES DATABASE WITH REDUCED DIMENSIONALITY**

**SANDEEP KADAM<sup>a</sup> AND DEVENDRA SINGH THAKORE<sup>b</sup>**

<sup>a</sup>M. E. (Computer Engineering),  
Bharati Vidyapeeth College of Engineering, Pune – 411041, India.  
<sup>b</sup>Professor, Computer Engineering Department,  
Bharati Vidyapeeth College of Engineering, Pune – 411041, India.

### **Abstract**

With rapid growth in ecommerce applications or any e-business application use of high dimensional data becomes very common. Thus mining this high dimensional data is the major problem faced by any business application. Time series data is also high dimensional data where information is collected with sequence of well defined data points. There are new difficulties and aspects in mining time series data due to their high dimensionality nature. There is need for the approach to reduce the dimensionality of time series data. Numbers of representations are proposed for time series data to reduce the dimensionality. Symbolic representation is one of the representations which are helpful in reducing dimensionality of high dimensional database. In this approach of representation we represent data with some alphabetic strings. To achieve symbolic representation of high dimensional time series data we first convert it to Piecewise Aggregate Approximation and then converted to symbolic strings. This representation is having advantage of lower bound along with dimensionality reduction and numerosity reduction. Most important fact of this approach is that traditional data mining techniques with few changes can be used for this representation of time series data.

-----  
**Keywords:** Data Mining, High Dimensionality, Lower bound, Time Series, Dimensionality reduction.