

# “Complex Event Processing Application for Smart City”

<sup>[1]</sup> Poonam Hajare, <sup>[2]</sup> Rajeshri Kachole, <sup>[3]</sup> Ankush Deshmukh, <sup>[4]</sup> V. S. Mahalle  
<sup>[1][2][3]</sup> Third Year B. E., <sup>[4]</sup> Assistant Professor

<sup>[1][2][3][4]</sup> Department of CSE, Shri Sant Gahanna Maharaj College of Engineering  
(SSGMCE) Shegaon-444203

<sup>[1]</sup> hajarepv123@gmail.com, <sup>[2]</sup> rajeshrikachole2012@gmail.com, <sup>[3]</sup> deshmukhank@gmail.com,  
<sup>[4]</sup> vsmahalle@gmail.com

**Abstract:**-- Smart city is defined as the ability to incorporate multiple technological solutions in secure fashion to manage the city assets. Emerging needs to make cities smarter, as proposed by IMB in Smarter planet program which mention the collaboration between the different city agencies such as health care agencies, transport agencies, various govt. agencies etc. This collaboration in smart cities will generate a huge data set. By applying complex event processing on these data set we can solve the various real time problems related to above mentioned agencies. Event processing is nothing but processing of the past or real time dataset to generate the new conclusion. These conclusions are helpful to find the opportunities or threats about any particular event. Complex event processing uses data mining to process the given dataset and give the essential event pattern as output. There is very less research work in India for calculating required conclusion from multiple dataset. This proposed system will help to propagate the optimum conclusion.

**Index Terms:** Event Processing, data mining, smart city, ID3 Algorithm.

## I. INTRODUCTION

Data Mining is process of extracting an event pattern from large database. Event is nothing but changing from one state to another. Suppose there is a car in showroom at that time the state of that car is 'SELL'. After purchasing that car by customer, the state of that car will be 'SOLD'. The main goal of Data Mining is to extract relevant event pattern for performing various operations. In data mining term event processing refers to any type of meaningful computation on event for getting efficient output. The incoming source data can be come from any type of data streaming sources. To store such a large amount of data with various type is not possible by DBMS(Data Base Management System). For that there are two type of system:

- 1) Data Stream Management Systems (DSMS)
- 2) Complex Event Processing (CEP) systems.

In this paper we are focusing on Complex Event Processing as an Application for 'Smart City'. To get quick response Complex Event Processing is new paradigm to process on data stream. CEP is used to analyze event pattern and it will also help in business side for better communication with IT and service department.

## II. CONSEQUENCES FOR SMART CITY:

The goal of smart city building is to improve quality of life by using various technologies to improve efficiency of services. Smart city uses information and communication technology [ICT] to enhance quality, performance and interactivity between citizen and government. Smart city criteria depend on six different aspects, and these are as follows.

- ❖ Economy
- ❖ People
- ❖ Governance
- ❖ Mobility
- ❖ Environment
- ❖ Living.

Out of above mentioned six criteria, for complex event processing we are focusing on mobility i.e. referring to accessibility of modern and fast transport system.

## III. LONDON CASE STUDY:

In London, the case study has been done to examine and analyze the data coming from different data sources like traffic data, data about event happening in city, weather condition etc. And to find the reasons and

solutions about traffic disturbances by applying complex event processing mechanism. As per the case study held in London it is seen that we can correlate data between traffic and various events happening in city (like cultural event). The step in above case study was to identify incoming data sources.

**Traffic data source:** Bing map, Traffic for London.

**Weather conditions. Sources:** Weather Underground, Yahoo! Weather, AccuWeather, etc.

**Events happening in the city:** Eventful.com, upcoming.org, last.fm, zvents.com, socialevnts.com.

All the data coming from these resources is stored in single database. The dataset obtained by these resources are dynamic in nature. That means they are updated after specified time interval. By analyzing the dataset obtained from the above data sources various problematic decisions like arranging the events, time and place about event and decision about rout of travelling etc.

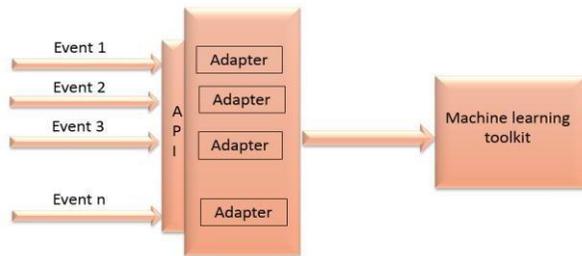


FIGURE (1) Data Sources

#### IV. PROPOSED METHODOLOGY:

Road traffic conditions in India are getting worse day by day. The average number of vehicles in India is growing at the rate 10.16 percent annually, since last five years. Spending hours in traffic jam has become part and parcel of metropolitan life style, leading to health and environmental hazards. The vehicle penetration in metropolitan cities like Mumbai is suffering from about 590 vehicles per Km of road stretch and Bangalore with around 5 million of vehicle ply over a network that extends barely up to 3000kms. There could be an approaches to solve this problem which is nothing but our 'Smart City' project will be based on these 3 scenario.

- ❖ Traffic Data
- ❖ Weather Report
- ❖ Cultural Event

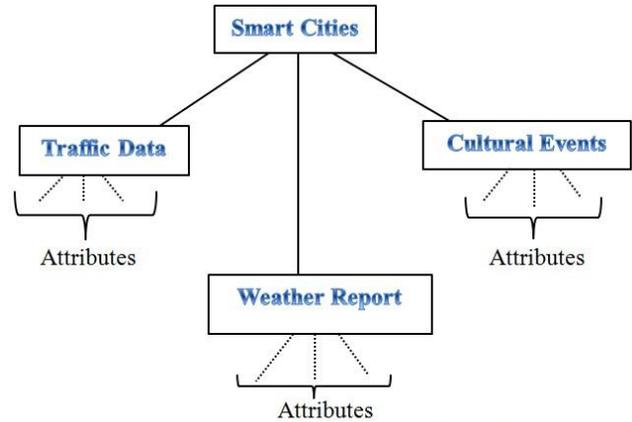


Figure. Decision Tree Of Complex Event Processing.

FIGURE (2) Decision tree for CEP

In case of traffic data, let's take an example, when we are driving a car then, we may need information regarding speed breaker, turning, traffic on rout etc. In case of weather report, we may need information about today's weather, forecast of next two days weather etc. In case of cultural event, we need information about functional hall. By using complex event processing we can combine all events from dataset obtain from given resources and generate conclusion for current and future incident. Our proposed methodology has four basic steps. And they are as follows.

- ❖ Dataset Collection
- ❖ Data Processing
- ❖ Data Mining
- ❖ Result Set Generation

#### V. DATASET COLLECTION:

In dataset collection we need to collect datasets about traffic, which we can collect from various sources like Google map, Bing map, various traffic related app databases etc.

Datasets related weather can be collected from meteorology department forecast, news, and various mobile apps AccuWeather etc. Dataset related to cultural events can be obtained from many cultural websites, some events are fix events, and information related events can be obtained from social media and newspapers. All datasets collected from various data sources is stored in a single database to maintain correlation between them and to make final conclusion from them.

#### VI. DATA PROCESSING:

The data collected from various sources is not structured data. To make that data structured we need to apply various mechanisms like data selection, data

clearing, data transformation etc. As all the data collected in data collection step is not fully useful so we need to delete useless data. Data cleaning involve error removing from dataset. Errors like missing values or inconsistent data etc. We need to apply various techniques to remove these anomalies. Only data cleaning will not allow us to apply data mining techniques on datasets. We need to transform our dataset. We have to apply smoothing aggregation and normalization to transform our dataset.

## VII. DATA MINING:

After removing all the anomalies from the collected dataset it's time to apply data mining techniques on it. Data mining can be done on any type of data i.e. text or media, only algorithm or approach varies. There are many algorithms like c4.5, k-means, Apriori and many more. In 1 clustering and association analysis are done for data mining. We perform data mining to discover the interested pattern. Only discovering interested Pattern is not sufficient, we need to remove the redundant patterns from the extracted data.

### A. Result set generation:

Result set generation is the last phase of data mining. In this phase the virtual representation of discovered result is done. Data set which is extracted by using interested pattern is called as result set. This result set is used to generate the final conclusion.

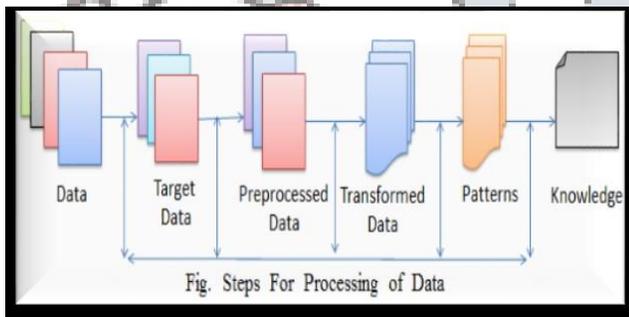


FIGURE (3) Steps For Processing of Data

## VIII. ALGORITHM:

- Step 1:** Create a root node for the tree
- Step 2:** If all examples are positive, Return the single-node tree Root, with label = +.
- Step 3:** If all examples are negative, Return the single-node tree Root, with label = -.
- Step 4:** If number of predicting attributes is empty, then Return the single node tree Root, with label = most common value of the target attribute in the examples.
- Else
- A = The Attribute that best classifies examples.
- Decision Tree attribute for Root = A.
- For each possible value,  $vi$ , of A,

Add a new tree branch below Root, corresponding to the test  $A = vi$ .

Let Examples( $vi$ ), be the subset of examples that have the value  $vi$  for A

If Examples ( $vi$ ) is empty

Then below this new branch add a leaf node with label = most common target value in the examples

Else below this new branch add the sub tree ID3 (Examples ( $vi$ ), Target Attribute, Attributes – {A})

End

**Step 5:** Return Root.

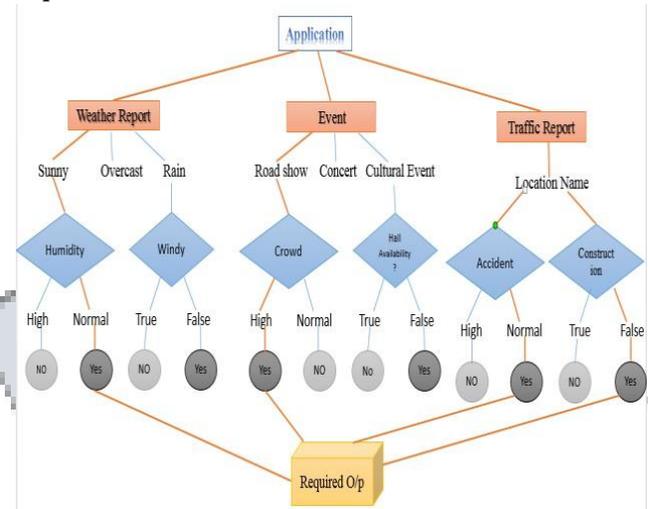


FIGURE (4) Decision Tree with Example.

## IX. CONCLUSION:

By analyzing the today's various problems related to day to day life, like traffic problems, weather problems and problems like finding the place to arrange different events need to solve using today's technology. Today's need to upgrade our metropolitan cities to smart city for renovation of our standard of living can be solved by using latest technology. This paper emphasis a solution to our different problems mentioned above. The proposed technology help us to find real time and optimum solution to our problems. And also to enhance the standard of living.

## FUTURE WORK:

We can increase our data sources to large number to focus on many more problems like mentioned in this paper. We can extend our work to various government departments, educational departments, public sector and many more to improve the standard of city and make it "SMARTER CITY".

Consider an example of Mumbai as a 'Smart City', we can collect traffic data of Mumbai from Google and Bing map etc. and weather report from meteorology department, AccuWeather app etc. and data of cultural

event from the Social Networking Sites and other website. By collecting above data we can bind it into single unit. After that we can extract target data that is useful event pattern then apply some pre-processing on that data. For storing multiple event patterns we can use event cloud. For getting appropriate event pattern, we can calculate entropy and information gain by using some formula. There may be common attributes in above mentioned events like location, time etc. by calculating entropy we can combine those attributes and store complete data in less memory. Suppose, there is a medical emergency and there is more than one route to reach the hospital then ambulance driver can search the shortest route with less traffic. Suppose user want to search functional hall with less traffic and know about weather of next day then also this proposed methodology can be used.

### REFERENCE

- [1] Stonebraker, M., Çetintemel, U., Zdonik, S. "The 8 requirements of real-time stream processing" ACM SIGMOD Record. 34, 42-47 (2005).
- [2] Cugola, G., Margara, A. "Processing Flows of Information" From Data Stream to Complex Event Processing. ACM Computing Surveys.
- [3] Niblett, P.: Event Processing In Action. (2010).
- [4] Wang, Q., Meegan, J., Freund, T., Li, F.T., Cosgrove, M.: Smarter City: The Event Driven Realization of City-Wide Collaboration. 2010 International Conference on Management of e-Commerce and eGovernment. 195-199 (2010).
- [5] Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanovic, N., Meijers, E.: Smart cities Ranking of European medium-sized cities. Vienna, Austria (2007).
- [6] Transport for London, Live Traffic Disruptions – Data Dictionary, (last accessed 1st September 2012), <http://www.tfl.gov.uk/assets/downloads/businessandpartners/data-dictionary-live-traffic-disruptions.pdf>
- [7] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten. The WEKA Data Mining Software: An Update. SIGKDD Explorations, Volume