



## Regional electricity consumption analysis for consumers using data mining techniques and consumer meter reading data



Ravindra R. Rathod\*, Rahul Dev Garg

Indian Institute of Technology Roorkee, Roorkee, India

### ARTICLE INFO

#### Article history:

Received 29 May 2015

Received in revised form 2 June 2015

Accepted 26 November 2015

Available online 22 December 2015

#### Keywords:

Data mining

Clustering

Association rules

Electricity consumption

Geographical features

### ABSTRACT

Data Mining (DM) techniques are employed to discover electricity consumption pattern at regional level in a city and used to extract knowledge concerning to the electricity consumption with respect to atmospheric temperature and physical distance from geographic features like river, farm, ground and highway. In order to form the different clusters of temperature and consumers based on the basis of electricity consumption *K*-means clustering algorithm is applied. Association rule analysis is carried out to form association rules on electricity consumption to describe the result of physical distance between natural geographic objects and various regions. The work includes pre-processing of data, application of DM algorithms and the interpretation of the discovered knowledge. To validate the proposed work, real databases of around twenty thousand consumers from Sangli city are used.

© 2015 Elsevier Ltd. All rights reserved.

### Introduction

Data mining is used for extracting hidden predictive information from large databases and is used in almost every discipline of science and engineering applications. The principal data mining techniques are regression analysis, decision tree, classification and prediction, clustering, association rule analysis and combination of these techniques. Recently data mining techniques are used in electricity distribution system for electricity supply forecasting [1], designing tariff plans [2,3] and to carry out consumer classification [4–7]. The growth in population, improved living standards and dependency on electronic and electrical appliances in day to day life attracts data miners to analyse consumers and regional analysis on the basis of electricity consumption, revenue collection, theft and fraud detection and other dependent factors contributing on fluctuations in electricity demand and consumption [8].

The electricity demand is increasing every year in almost every sector [9]. Installing and starting new electricity generation plants is difficult due to environmental preservation awareness and pollution control policies of government. To overcome this problem, proper electricity consumption profiling can be used to find alternate ways for managing electricity loads in near future with existing electricity generation capacity. Electricity profiling can be

carried out for individual consumer, on small area and at regional levels.

Different regions in a city have a different electricity consumption pattern with respect to locality and nearest man-made geographical features. Similarly, regions at different geographical locations may have similar or matching electricity consumption pattern in association with similar geographical features such as cities on the bank of rivers or at same altitude (hilly areas). The consumption pattern of region is highly influenced by geographical features and environmental conditions like atmospheric temperature, humidity and rainfall [10].

### Related work

Data mining uses different approaches and builds different models depending upon the type of data and objectives. Broadly, data mining methods are classified as predictive and descriptive. Predictive methods are classification techniques and descriptive method covers association rules and clustering [11]. Depending on data types used for analysis, data mining methods are classified as web and text mining (internet application), multimedia mining (audio/video applications), spatial data mining (remote sensing and GIS application) and time series data mining (business and financial applications) [12].

Data mining techniques are applied to extract useful information from electricity consumption data of consumers. Data mining methods are used at different stages to generate useful information for preparing regional electricity consumption profile. Each region

\* Corresponding author.

E-mail addresses: [rrr.wce@gmail.com](mailto:rrr.wce@gmail.com) (R.R. Rathod), [garg\\_fce@iitr.ac.in](mailto:garg_fce@iitr.ac.in) (R.D. Garg).

has few thousands of consumers where electricity is supplied in a hierarchical manner. Generated electricity is transmitted through transmission lines to electricity distribution network, substations, regional or area transformers, various electricity poles and at the end to consumer.

Data mining is used for short and long term forecasting of electricity demand, planning, fraud, theft and fault detection [13–16]. Many studies are carried out using consumers load profile analysis [3–5,10,17,18] for tariff planning and consumer characterization using house size and economic condition of consumer [19]. A seasonal electricity consumption analysis is carried out to identify environmental effect on consumers electricity consumption [2,7]. The load profile of individual consumer varies differently depending on income level, residence locality, residential building type, total number of persons living together and environmental factors [19,20]. Some studies were also carried out using sample data from city to study load profile of few consumers representing entire class or cluster they belong to [1,2,5,21,22].

In present study, the effect of various geographical objects nearer to the region and atmospheric temperature of the city on electricity consumption is studied using clustering and association rule analysis. To perform the analyses data of about 20,000 consumers from electricity distribution region (feeder) of Sangli city has been used. The consumers are divided into regions by government agency MahaDiscom (Maharashtra state electricity distribution company limited), which supplies electricity to all consumers. After grouping each consumer in individual regions, data cleaning is carried out to remove the consumers having very low annual electricity consumption of 150 units per annum (average monthly consumption less than 15 units; 1 unit of electricity equals to 1 kW-h consumption for 1 hr).

#### *Geographic information system (GIS)*

The term geographic information system (GIS) is defined by Chrisman [23] as “Geographic information system – Organized activity by which people measure and represent geographic phenomena then transform these representations into other forms while interacting with social structure. GIS system helps to manage spatially or geographically references (from earth) data by storing, transforming, accessing, editing and displaying for studying environmental and man-made (artificial) processes, detecting changes and trends for the possible planning and analysis purpose.

Commonly GIS is used for applications involving planning transportation network (Road, train) [24,25], municipal tax/revenue and sewer, water supply management [26], flood assessment modeling [27], urban growth and sprawl assessment [28], land and natural resources management system, planning information system [23].

Wang et al. [29] presented a model using GIS and spatial database using city type (major), highway (close) and population (>million). The presented model represents spatial features as attribute data to deduce rules for logistic industry in china. Such attribute reduction policy is very useful to minimize spatial data storage and processing cost while preserving spatial feature data in attribute form.

#### *Data mining and GIS*

Development and advances in remote sensing (RS), global positioning system (GPS) and GIS has generated massive amount of spatial and non-spatial. Such data is very helpful for various planning purposes. Data mining plays an important role in assessing massive GIS data for proper planning using spatial and non-spatial data [30,31].

Many researchers applied data mining techniques (spatial and non-spatial) for applications such as; car crash test in transportation using road data (width, roughness) [13] through GPS and GIS. An approach to build spatial database to perform spatial association rule mining is presented by Bogorny et al. [32] to show spatial relationship between spatial features (gas station, hospital, water bodies and street) with non-spatial feature population.

Some GIS applications includes electricity distribution network information system [33] used to manage entire electrical network, mapping electrical power distribution network [34,35], electricity loss, theft and fraud detection in power distribution network [34], managing electricity feeders for effective distribution of electricity to consumers [36], electricity distribution network load analysis and load pattern classification of consumers using clustering technique [17]. In electricity distribution system; spatial data mining has wide range of applications. For example, a GIS study gives us the high electricity consumption, high electricity losses or high revenue generating areas from a city or state. Such relationships between spatial features (area or zone) and non-spatial features (electricity consumption or revenue collection) are easy to discover with the help of GIS and data mining. Common spatial data mining tasks [31] are; spatial classification and prediction, spatial association rule mining, spatial cluster analysis and geo-visualization.

#### **Case study on a consumer's electricity consumption data from Sangli city**

The study has been carried out on consumer's data from Sangli city (16.86° N, 74.57° E) located on the banks of river Krishna in the western part of Maharashtra state, India. The valley of the river Krishna offers many irrigation and agricultural advantages to Sangli city and district. Due to large scale irrigation, agriculture farming is the main business around Sangli district and nearby region. The physical location of Sangli city comes under Deccan plateau, nearly 120 km away from Arabian Sea and at 549 m height from mean sea level. Fig. 1 shows location map of Sangli city used for study.

#### *Data selection, cleaning and pre-processing*

The dynamic expansion of cities due to development in infrastructure, industries and residential growth increases electricity consumers which produce huge electricity consumption data. Study and analysis of consumers electricity consumption data is helpful for detecting consumption pattern at various levels such as individual consumer, zone and at regional level in electricity distribution network. To analyse huge data, data pre-processing operations are carried out in first step before applying data mining algorithm [11,12]. Data pre-processing includes preparation of data in desired form to work, which is clean and free from any noise. It is also used for reduction of actual huge data into summative workable data to avoid unnecessary processing of unwanted, meaningless data. Consumers monthly electricity consumption values are checked and inconsistent consumers are removed from study data.

#### *Clustering*

Clustering is an important technique applied to form groups or clusters of data, which represents a common property of entire elements within group [37–40]. Each element within cluster represents a class or common property among them. Clustering algorithms are used for various data types including numerical, categorical and multimedia [15,16]. In this proposed study work

Download English Version:

<https://daneshyari.com/en/article/398440>

Download Persian Version:

<https://daneshyari.com/article/398440>

[Daneshyari.com](https://daneshyari.com)