International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 2, pp: 27-30 www.ijaem.net ISSN: 2395-5252

# Reflecting Water Pollution Using Data Mining: Survey Paper

S.Subashini [1], A.Geetha[2]
[1]*M.phil Research Scholar, Dept.of Computer Science,Chikkanna Government Arts College, Tirupur. Bharathiar University,Coimbatore,Tamilnadu,India.*
[2]*Head of the Department,Dept.of Computer Science, Chikkanna Government Arts College, Tirupur. Bharathiar University,Coimbatore,Tamilnadu,India.*

Date of Submission: 07-07-2020 Date of Acceptance: 21-07-2020

**ABSTRACT**
Water is important of human life. Fresh water 0.76%,Ground Water 1.69%,Oceans and Sces-96.54% and Ice and Permanent snow-1.74% in the world. The water occupies the world and also used the water in animals and plant. The water has pollution in many ways, such as Ground water pollution, Chemical pollution, Oxygen depletion pollution, Sewage lockage, Industrial waste.
This pollution affects the people and also animals. So many diseases will be attack the human. Such as Malaria, Dengue, Cholera, Typhoid and Dysentery. Waterborne Disease is World's Leading Killer. The World Health Organization says that every year more than 3.4 million people die as a result of water related diseases, making it the leading cause of disease and death around the world. In this paper discuses the water pollution on sewage lockage, used by a Fp-Growth algorithm and Python software. Then analysis the result based on the pH(potential Hydrogen). This paper prepares the three companies of sewage lockage, so collect the data for three companies. How much sewage to deliver the companies and it's problems. This project used by some parameter to analysis the result.
**KEYWORDS**: Pollution, Malaria, Dengue, Cholera, Fp-Growth.

## I. INTRODUCTION

Water is essential for all living organisms including, Such as human, animal, planet, food production and economic development.[1] Then the water mostly used for direct and indirect purpose. Direct purpose include bathing, drinking, and cooking, while examples of indirect purposes are use of water in processing wood to make paper and in producing steel for automobiles. The bulk of the world's water use is for agriculture, industry, and electricity. All living organisms need a good and quality water. Millions of women spend hours everyday collecting water, 2.6 billion people lack access to sanitation, and 1.8 million children die each year from diarrhea. Barriers to addressing water problems in developing nations include poverty, education, climate change, and poor governance. The Water can be many types of wasted, Such as,

- Washing machine. Washing machines are the biggest water users in the average home.
- Shower. Long showers and hot baths waste a ton of water.
- Faucet.
- Leaks.
- Other things to consider.

Tamil Nadu budget analysis 2019-2020 spending on rural development as a 2.6% of total budget.[2] This pollution affects the food chain also and damage the ecosystem. Fp-Growth Algorithm, Proposed by Han in, is an efficient and scalable method for mining the complete set of Frequent Patterns by Patterns Fragment Growth using an extended Prefix-Tree structure for storing compressed and crucial information about Frequent Patterns named Frequent-Pattern
Tree(FP-Tree). Olaniran 1995 to define water pollution to be the presence of excessive amount of a hazard(pollutants) in water.
Water pollution could be divided as
(a)Biological
(b) Chemical.

Understand the Water Quality and learn many type of water related problem, types, methods, and solutions.[3] Then divided the two category and then composed the two component. The Water Pollution related papers to be reviews and compare. Then summarized the all things and compare the strong & weak point existing research paper. Finally discuss the future research.

## II. LITERATURE REVIEW

Gaganjot Kaur Kang. Data-driven Water Quality Analysis and Prediction: A Survey. Water quality becomes one of the important quality factors for the quality life in smart cities. Recently, water quality has been degraded due to diverse forms of pollution caused by disposal of human wastes, industrial wastes, automobile wastes. The increasing pollution affects water quality and the quality of people's life. This paper is supported by Futurewei's support that is extended to the Center of Smart Technology, Computing, and Complex Systems at San Jose State University.

F. Karimipour1 (2005). Water Quality Management Using GIS Data Mining. Nowadays scientists, managers and decision makers have faced with ever increasing production of digital geospatial data acquired at various geometric, thematic and temporal characteristics. Geospatial information systems (GISs) have been widely considered to handle such a diverse range of geospatial data. The results have clearly identified the relationship between number and location of industrial pollutions and water quality indicators to be used in environmental protection and land use planning. This paper provided a brief review on geospatial data mining, the need and required analyses to address the importance of water quality management. Some applications of data mining technique have been tested.

P. B. Cheung1 (2014). Extension of Epanet for Pressure Driven Demand modeling in Water Distribution System. Water utilities for applications such as design, calibration, rehabilitation and operation. Currently, there are several hydraulic simulation models and most of them are based on demand driven analysis (DDA). In this paper, the authors describe an extension of the Epanet by OOTEN toolkit to directly

include pressure driven demand modeling. The data structures and algorithms within Epanet source code are modified to adequate PDA modeling.

Joshua Nizel Halder1 (2015). Water Pollution and its Impact on the Human

Health. River pollution has been one of the main topics in the environmental issue of urban Dhaka, the capital city of Bangladesh. This study was conducted to find out the pollution situation of Turag river and the health problem of the surrounding residents. The maximum recorded values of pH, color, turbidity, biochemical oxygen demand (BOD5), hardness, total

dissolved solids (TDS), chloride (Cl-), carbon-di-oxide (CO2) and chemical oxygen demand (COD) were7.1 mg/L, 625 ptcu, 97.2, 4.65 mg/L, 1816 mg/L, 676mg/L, 5 mg/L, 15.5, and 78 mg/L, respectively. The results of the sampling programmed clearly determine that the water quality of Turag river may not be in a position to sustain the aquatic life as well as not suitable for using for domestic purpose.

Anil K Dwivedi (2017). RESEARCHES IN WATER POLLUTION: Due to lack of time and resources, the sampling programmed was limited to four months duration, from April 2013 to July 2013.A REVIEW. More than 70% of the fresh water in liquid form of our country is converted into being unfit for consumption. Not only India, but other countries are also suffering from the same problem. This has been explained clearly by the help of considerable number of references in this paper. Various sources of pollution such as sewage discharge, industrial effluents and agricultural runoff and their potential has been studied in mass. In light of the above study we come to the conclusion that the level of water pollution have reached to the alarming stage. The quality of water in most part of the world has degraded, though the situation in India is more severe. Our body consists of about more than 10000 hormones and enzymes which are very specific in their requirement and kinetics.

Mehtab Haseena (2017) Water pollution and human health. Water covers about 70% Earth's surface. Safe drinking water is a basic need for all humans. The WHO reports that 80% diseases are waterborne. Industrialization, discharge of domestic waste, radioactive waste, population growth, excessive use of pesticides, fertilizers and leakage from water tanks are major sources of water pollution. These wastes have negative effects on human health. Water pollution is a global issue and world community is facing worst results of polluted water.

Major sources of water pollution are discharge of domestic and agriculture wastes, population growth, excessive use of pesticides and fertilizers and urbanization.

Mahipal Singh Sankhla (2018) Water Contamination through Pesticide & Their Toxic Effect on Human Health. Water is the most simple structure block of human life or Fresh water is essential for human health. The use of pesticides and nitrogen fertilizers in agriculture has grown dramatically over the past many years to contaminated water.

Environmental exposure of humans to pesticide is common effects in acute and chronic health effects, including acute and chronic neurotoxicity (insecticides, fungicides, fumigants), lung damage (parquet), chemical burns (anhydrous

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 2, pp: 27-30 www.ijaem.net ISSN: 2395-5252

ammonia), and infant methemoglobinemia (nitrate in groundwater). Pesticide are present at higher levels should be removed from drinking water for human safety. There is a need to maintain control on disposal of industrial waste or Agriculture waste in water bodies and to bio-monitor the trace elements in the water and other eatables. It is suggested that there must be proper waste disposal system and waste should be treated before entering in to water & soil.

## III. PROBLEM AND DIRECTION

**Data Mining:** Data mining is an approach for information extraction from huge amount of data stored in a database. In other words, data mining can be considered as an approach to determine the valid, novel, useful and ultimately understandable data patterns in a large database. It has to be pointed out that data mining analyses are valid while having huge amount of data; otherwise, the achieved results may not be extended further. One solution to this problem is to consider data warehousing (DW) concepts. The following fig 3.1.1.


**Fig:3.1.1 Data Mining Process**

**FP-Growth Algorithm:**

FP-growth algorithm find frequent item sets or pairs, sets of things that commonly occur together, by storing the dataset in a special structure called an FP-tree. The FP-growth algorithm scans the dataset only twice. The basic approach to finding frequent item sets using the FP-growth algorithm is as follows:

1. Build the FP-tree.
2. Mine frequent item sets from the FP-tree.
3. The FP stands for "frequent pattern." An FP-tree looks like other trees in computer science, but it has links connecting similar items. The linked items can be thought of as a linked list.
4. The FP tree is used to store the frequency of occurrence for sets of items. Sets are stored as paths in the tree. Sets with similar items will share part of the tree. Only when they differ will the tree split. Its defined figure3.2.1. A node identifies a single item from the set and the number of times it occurred in this sequence. A path will tell you how many times a sequence occurred. The links between similar items, known as node links, will be used to rapidly find the location of similar items.

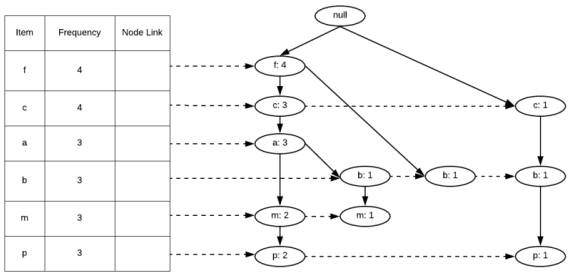


**Fig: 3.2.1 FP-Growth**

## IV. CONCLUSION:

In this paper defined by a various data mining techniques at different location. This paper researched by a sewage lockage water pollution in three companies [6]. The pollution will be growing day by day. Its affect many resource. FP-Growth Algorithm used by Python software. Then the companies are dispatch the water, in the water affect the human and animals also. These problem related dataset will be collect and some attribute used. The attribute of pH is based. That is calculate the pollution, Such as the pH of pure water is 7[7]. In general, water with a pH lower than 7 is considered acidic, and with a pH greeter then 7 is considered basic. The normal range for pH in surface water system is 6.5 to 8.5, and the pH range for ground water system is between 6 and 8.5 water with a pH > 8.5 could indicate that the water is hard.

## REFERENCE:

[1]. Gaganjot Kaur Kang. Data-driven Water Quality Analysis and Prediction: A Survey.
[2]. S. Pavithra Devi (2018) Data Mining Case Study for Water Quality Prediction using R Tool.
[3]. Mahipal Singh Sankhla1 (2018). Water Contamination through Pesticide & Their Toxic Effect on Human Health.
[4]. Anil K Dwivedi (2017). RESEARCHES IN WATER POLLUTION: A REVIEW.
[5]. Muhammad Faheem Malik (2017). Water pollution and human health.
[6]. SURVAY PAPER ON WATER POLLUTION (2015). Association Rule Mining for Ground water and Wastelands Using Apriori Algorithm: Case Study of Jodhpur District
[7]. Joshua Nizel Halder1 (2015). Water Pollution and its Impact on the Human Health.

[8]. P. B. Cheung1 (2014). Extension of Epanet for Pressure Driven Demand modeling in Water Distribution System.
[9]. Shukri I. Al-Hassen (2012). Levels of Heavy Metals Pollution in the Aquatic Environment of Basra City, Iraq.
[10]. F. Karimipour1 (2005). Water Quality Management Using GIS Data Mining.

International Journal of Advances in Engineering and Management
ISSN: 2395-5252

# IJAEM

Volume: 02 Issue: 01 DOI: 10.35629/5252

www.ijaem.net
Email id: ijaem.paper@gmail.com