B V RAJU COLLEGE

VISHNUPUR

BHIMAVARAM

FACULTY RESEARCH PUBLICATIONS

ACADAMIC YEAR 2018-19

INDEX

KNOWLEDGE REDUCTION IN MASSIVE PATIENT DATASETS USING ROUGH SET APPROACH

[1]K. B. V. Brahma Rao, [2]P. Suresh Varma, [3]R Krishnam Raju Indukuri, [4]M V Rama Sundari
[1]Associate Professor, [2]Professor, [3]Professor, [4]Professor
[1]Department of MCA,
[1]Department of MCA, B.V. Raju College, Vishnu Campus, Bhimavaram, A.P, India

Abstract : In order to eliminate redundancy of massive datasets, we developed parallel large-scale technique for knowledge reduction using rough set and MapReduce methods on patient massive datasets. Our technique will reduce the utilization of memory and processing time. The superfluous data is removed without significant accuracy loss using type of disease. In this paper we presented theoretical and experimental approach for knowledge reduction from large patient datasets using significance of attributes by organizing the data in discernibility and indiscernibility matrices. The experimental results demonstrate that the proposed parallel knowledge reduction method can efficiently process massive datasets on Hadoop platform, with highly speed up the grouping process and largely reduce the storage requirements. In all the experiments the introduced method based on significance of attributes is compared with the method based on positive region or information entropy. The comparison clearly shows that the former method outperforms the latter one.

Index Terms **- Big Data, MapReduce, Rough Set, Knowledge Reduction, HDFS.**

I. INTRODUCTION

The growing data in industry is like healthcare and scientific areas for the last eight years makes it difficult to store, manage and analyzing it either to make decisions or to retrieve the required data. In order to deal with the data explosion and knowledge reduction, we develop a parallel large-scale knowledge reduction method based on rough set method to acquire the knowledge using MapReduce technique.

We designed the parallel algorithm model for knowledge reduction using MapReduce which can be used to compute for the algorithms using indiscernibility matrices and functions. The proposed technique removes some superfluous data from a dataset by conserving its properties. The experimental results demonstrate the proposed technique that can efficiently process massive datasets with highly speedup the classification of data and largely reduce the storage requirements. In all the experiments the introduced method based on indiscernibility matrices is compared with the method based on positive region. This method clearly shows that the former method outperforms the latter one.

The information of a dataset attributes can classify into two classes called condition and decision (action) attributes. Each row of a decision table determines a decision rule, which specifies decisions (actions) that should be taken when conditions pointed out by condition attributes are satisfied. Objects in a decision table are used as labels of decision rules. Decision tables comprising inconsistent decision rules are called inconsistent (nondeterministic, conflicting); otherwise the table is consistent (deterministic, non-conflicting). The number of consistent rules in a decision table can be used as consistency factor of the decision table.

A set of decision rules is called a decision algorithm. Thus with each decision table we can associate a decision algorithm consisting of all decision rules occurring in the decision table. A decision table is a collection of data, whereas a decision algorithm is a collection of implications. To deal with data we use various mathematical methods, e.g., statistics but to analyze implications we must employ logical tools. Thus these two methods are not equivalent; however for simplicity we present here decision rules in the form of implications, without referring deeper to their logical nature.

An important issue in data analysis is discovering dependencies between attributes. We would need also a more general concept of dependency of attributes, called a partial dependency of attributes.

Knowledge finding has become a new challenge using big data. The Rough set theory has been successfully used in data mining. The MapReduce technique has been using for big data analysis in the recent times. Large amounts of data are collecting daily from various sources using sensors and devices in different formats by industries and scientific community. The size of the data may be zetta byte or yotta byte. The core data is processed by different applications and is used to convert the core data into the same format. To process this much of data, Google developed a software frame work is known as MapReduce.

The MapReduce technique supports large distributed datasets on clusters of computers which can analyze massive amounts of data. This has been a popular computing model for cloud computing platforms and is followed by Google's work, many implementations of MapReduce have emerged and lots of traditional methods combined with MapReduce have been presented here until now.

A parallel method is improving the performance of data mining for the effective computation of approximation. The parallel method makes our approach more ideal for executing large scale data using MapReduce technique. Mining the big data and knowledge discovery is a new challenge in the current days because the volume of data growing is at an unmanageable rate. The MapReduce technique has acknowledged much responsiveness from both scientific community and industry for its applicability in big data analysis.

Rough set model can be defined generally by means of topological operations, interior and closure, called approximations. This model processes incomplete data which is based on the lower and upper approximations and is defined as a pair of two crisp sets corresponding to approximations. The main advantage of rough set theory in data analysis is that it does not need any initial or supplementary information concerning data.

Rough set has also provided the necessary formalism and thoughts for the development of some propositional machine learning systems. Rough set theory has also been used for knowledge representation, dealing with imperfect data, reducing knowledge representation, data mining and for analyzing attribute dependencies. Rough set Theory has found many applications such as medical data, security analysis, power system, image processing, voice recognition and finance. This technique is one of the research areas that have successfully used for knowledge discovery or Data Mining in datasets.

To expand the application of rough sets in the field of big data mining and deal with massive data sets, the parallel computation of the rough set approximations is applied and this computation can be achieved by using MapReduce Technique.

1. Map-function: This function takes an input pair and produces a set of key, value pairs. The MapReduce groups together all values associated with the same key and sends them to the Reduce function.

2. Reduce-function: This function accepts key and a set of values for that key. It merges these values together to form a possibly smaller set of values by doing sorting and shuffling it produces reduced values get from Map function.

II. RELATED WORK

Weiping Cui [1] proposed method enables knowledge reduction algorithms to be applied over big data reduction problem without significant accuracy loss using information entropy. Jin Qian [2] discussed hierarchical attribute reduction algorithms in data and task parallel using MapReduce. Yan Zhao [3] discussed three different types of reducts can be constructed, keeping the indiscernibility, discernibility, and indiscernibility-and-discernibility relations, respectively. The existing methods for constructing the indiscernibility reducts also can be applied to construct the other two types of reducts.

J. Qian [4] proposed hybrid algorithms for attribute reduction. They first introduced a counting sort algorithm with time complexity for dealing with redundant and inconsistent data in a decision table and computing positive regions and core attributes. Then, hybrid attribute measures are constructed which reflect the significance of an attribute in positive regions and boundary regions. Yuhua Qian [5] discussed approximation reduct model to characterize the smallest attribute subset that preserves the lower approximation and upper approximation of all decision classes in this rough-set model. They used several key algorithms for finding an approximation.

Zdzislaw Pawlak [6] presented the basic concepts of rough set theory and point out some rough set based research directions and applications. Zdzistaw Pawlak [7] discussed approximate operations on sets, approximate equality of sets, and approximate inclusion of sets. The presented approach may be considered as an alternative to fuzzy sets theory and tolerance theory.

Jerzy Blaszczynski [8] presented a general rule induction algorithm based on sequential covering, suitable for variable consistency rough set approaches. This algorithm, called VC-DomLEM, can be used for both ordered and non-ordered data. Ching-Hsue Cheng [9] proposed four procedures in the hybrid model to provide efficient rules for forecasting, which are evolved from the extracted rules with high support value, by using the toolset based on the rough sets theory. The effectiveness of the proposed model is verified with two types of performance evaluations, accuracy and stock return.

Hongmei Chen [10] discussed updating approximations dynamically when attribute values are coarsened or refined. Jeffrey Dean [11] discussed MapReduce programming model and an associated implementation for processing and generating large datasets that is amenable to a broad variety of real-world tasks. Jeffrey Dean [12] designed a MapReduce programming model for more than ten thousand programs at Google, including algorithms for large-scale graph processing, text processing, machine learning, and statistical machine translation.

Chen Degang [13] presented a model to reduce the attributes of covering decision systems, which are databases characterized by covers. Zdzislaw Pawlak [14] discussed basic concepts of rough set theory and their granular structure. And also discussed the consequences of granularity of knowledge for reasoning about imprecise concepts. Guo-Fang Qiu [15] discussed characterizations of three important types of attribute sets in generalized approximation representation spaces, in which binary relations on the universe are reflexive.

Liangxiu Han [16] designed data mining and integration (DMI) model as a streaming data-flow graph: a directed acyclic graph (DAG) of Processing Elements (PEs). Isaac Triguero [17] presented a novel distributed partitioning methodology for prototype reduction techniques in nearest neighbor classification. Ashwin Srinivasan [18] examined the applicability to Inductive Logic Programming (ILP) of a popular distributed computing approach that provides a uniform way for performing data and task parallel computations in ILP.

Abhishek Verma [19] discussed how genetic algorithms (GAs) can be modeled into the MapReduce model. And also described the algorithm design and implementation of GAs on Hadoop, an open source implementation of MapReduce. Qinghua Hu [20] discussed a neighborhood rough set model to deal with the problem of heterogeneous feature subset selection. Daniel Zinn [21] presented a set of approaches for exploiting data parallelism in XML processing pipelines through novel compilation strategies to the MapReduce framework.

III. PROPOSED WORK

In the review of literature a very little work has been found towards knowledge reduction by removing superfluous data. No attempt is found to design data reduction using significance of attributes technique. This technique is used on discernibility and indiscernibility matrices. In this paper, we use Hadoop open source software. The Apache Software Foundation attempted for distributed storage and distributed processing of massive data on computer clusters. The commodity hardware is most sufficient in the clusters. The structure of Hadoop Distributed File System (HDFS) is described in [Figure-1].

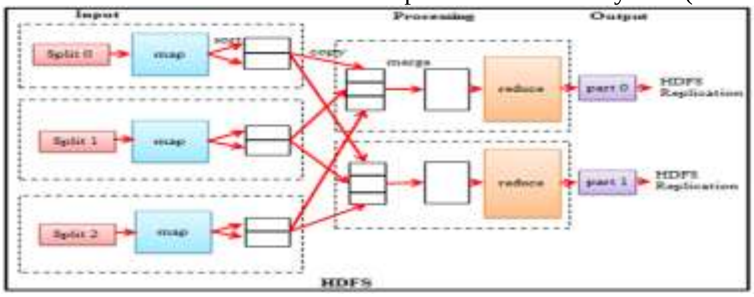

Figure 1 Structure of HDFS

Hadoop Distributed File System (HDFS) is the very large storage system for datasets used by Hadoop applications. HDFS creates multiple replicas of data blocks and assigns them on data nodes, to enable reliable tremendously rapid computations. Hadoop consist of two most important modules which are: File storage and Distributed processing system. The first module of file storage is known as "HDFS (Hadoop Distributed File System)". It is responsible for scalable, reliable, comparatively low cost storage. The files are stored across a collection of servers in HDFS and data availability is monitoring continually in a cluster servers. The second module of Hadoop is the parallel data processing system called "MapReduce". The Hadoop distributed file system and the MapReduce framework are running on the same set of nodes. The MapReduce programming allows the execution of Java code and also uses software written in other languages.

Basic Notions:

We here introduce about notions of the Pawlak rough set theory. The indiscernibility relation and equivalence class are important concepts in Pawlak's rough set theory. The indiscernibility relation expresses the fact that due to lack of information (or knowledge), we are unable to discern some objects by using available information. It determines a partition of U and is used to build the equivalence classes.

Definition 1: S = <U, C ∪ D, V, f> is a decision table, U = $\{x_1, x_2, \ldots, x_n\}$named domain is a finite non-empty set of objects. C = $\{c_1, c_2, \ldots, c_n\}$ is a set of conditional attributes describing the objects, and D is a set of decision attributes that indicates the classes of objects. C ∪ D = Ø. V = $\bigcup_{a\in C\cup D} V_a$, Va is a non-empty set of values of a ∈ C ∪ D. f: U x (C ∪ D) → V is an information function that maps an object in U to exactly one value in V_a, that means , f(x, a) = V means that the object *x* has the value v on attribute a.

Definition 2: An indiscernibility relation with respect to R ⊆ C ∪ D is defined as:

$$IND (R) = \{(x, y) \in U \times U \mid \forall a \in R, f(x, a) = f(y, a)\}$$

The partition generated by IND (R) is denoted as U│IND (R), U│R for short. Any elements in the U│R, [x] R = {y│∀ a ∈ R, f(x, a) = f(y, a)} is the equivalence classes.

Definition 3: For a decision table S = <U, C ∪ D, V, f>, for each subset of X ⊆ U and indiscernibility relation R ⊆ C ∪ D, the lower and upper approximations of X with respect to a partition R are defined as:

$$R (X)_{-} = \cup\{Y \in U \mid R : Y \subseteq X\}$$
$$R (X)^{-} = \cup\{Y \in U \mid R : Y \cap X\}$$

Definition 4: For a decision table S = <U, C ∪ D, V, f>, Let ∀ A ∈ C, then positive region POS(D│A) is given by:

$$POS(D \mid A) = \bigcup_{1\le i\le k} R(X)_{-}$$

Definition 5: Let U | A = $\{A_1, A_2, \ldots, A_r\}$, U | D = $\{d_1, d_2, \ldots, d_k\}$, then information entropy of A is given by :

$$H(A) = -\sum_{i=1}^{r} p(A_i)\log_2 p(A_i)$$

Conditional entropy of *D* conditioned on *A* is given by:

$$H(D|A) = -\sum_{i=1}^{r} p(A_i) \sum_{j=1}^{k} p(D_j \mid A_i)\log_2 p(D_j \mid A_i)$$

$$\text{Where} \quad p(D_j \mid A_i) = \frac{|D_j \cap A_i|}{|A_i|} \ (i = 1, 2, \ldots, k).$$

From the definition 4 and 5, the calculation form based positive region or information entropy both can be expressed as $\sum_{1\le i\le l} \Delta$ (Δ represents some sort of calculation of the same equivalence class), <equivalence class, some sort of calculation of the equivalence class> is like <k, v>. So the equivalence classes can be computed in parallel using MapReduce.

Definition 6: For a decision table *S*, let A ⊆ C, c ∈ C-A the information entropy attribute importance of c is given by:

$$Sig_{info} (c, A, D) = H(D \mid A) - H(D \mid A \cup c)$$

The positive region attribute importance of c is given by:

$$Sig_{pos} (c, A, D) = |POS(D \mid A \cup c| - |POS(D \mid A)|$$

Significance of attributes: The reduction of attributes can be generalized by introducing a concept of significance of attributes, which enables us evaluation of attributes not only by two-valued scale, dispensable and indispensable but also by assigning to an attribute a real number from the closed interval [0, 1], expressing how important is an attribute in the data table. Significance of an attribute can be evaluated by measuring effect of removing the attribute from an information table on classification defined by the table. The significance of the attribute 'a' calculated as

$$\sigma_{(C,D)}(a) = \frac{(\gamma(C,D) - \gamma(C-\{a\},D))}{\gamma(C,D)} = 1 - \frac{\gamma(C-\{a\},D)}{\gamma(C,D)}$$

Let C and D be sets of condition and decision attributes respectively and let 'a' be a condition attribute, i.e., $a \in C$. The significance of the attribute a can also be denoted by $\sigma(a)$ and obviously $0 \le \sigma(a) \le 1$. The more important is the attribute a is greater than $\sigma(a)$.

Knowledge reduction algorithm using MapReduce:

The emphasis of this paper is the calculation form based on the significance of attributes. In Parallel knowledge reduction algorithm, Map function is written to compute the significance of attributes by assigning a real number from the closed interval [0, 1]. Reduce function is written to reduce the attributes from the dataset by using interval value. The final summary of dataset consists only the required attribute without losing information contained in the dataset.

Knowledge reduction algorithm based on MapReduce is mainly including 3functions: the Map function (algorithm 1), the Reduce function (algorithm 2) and main function (algorithm 3).

Algorithm 1: Map (k, v)

Input: The selected condition attributes set C and decision attributes set D, 'a' is the conditions attribute a ∈ C.

Output: Cond_ Attrib, <g(a), v> // v is either zero or nonzero

1. For a ∈ C do
2. Compute sig(a) //significance of attribute 'a'
3. Emit key and value pairs (k, v) based on zero and one values as Cond_Attrib, <g(a), v>.

By algorithm 1, we can compute the significance of attributes of the given dataset.

Algorithm 2: Reduce (String Cond_Attrib, pairs [$<c_1, n_1>$, $<c_2, n_2>$, …])

Input: Conditional attributes set and its corresponding significance value list.

Output:

$sig\Delta^d$ is the importance of attributes d of the given set of attributes $<d, n> \in [<d_1, n_1>, <d_2, n_2>, \ldots]$

1. For do
   Classify the attributes into groups using the value received from algorithm 1.
2. Emit groups of attributes with same value.

By algorithm 2, we can produce different set of attributes using corresponding value.

Algorithm 3: Main function

Input: Set of decision attributes.

Output: Set of reduction attributes Red.

1. Red = Ø;
2. Compute conditional and decision attributes H(C│D).
3. Start a job,

{Execute algorithm 1 and algorithm 2, according to the result compare the significant value one attribute with remaining attributes and Sig_{info} (a ∈ C - Red) select the best attribute, Red = Red ∪ {a}}

5. Output Red.

By Algorithm 3, according to the significance attributes, we can determine an optimal set of attributes. In order to deal with the data explosion and knowledge scarcity, we have developed a parallel large-scale knowledge reduction method based on rough set for knowledge acquisition using MapReduce for massive patient datasets in this paper. It constructs the parallel algorithm framework model for knowledge reduction using MapReduce, which can be used to compute a reduction for the algorithms based on significance of attributes using discernibility and indiscernibility matrices. The proposed method enables knowledge reduction algorithm to be applied over massive datasets reduction problem without significant accuracy loss. The experimental results demonstrate that the proposed parallel knowledge reduction method can efficiently process massive datasets on Hadoop platform, which highly speed up the grouping process and largely reduce the storage requirements. In all the experiments the proposed method is compared with the method positive region or information entropy methods.

In order to eliminate redundancy of massive datasets, we have developed parallel large-scale technique for knowledge reduction using rough set and MapReduce methods on patient massive datasets. Our technique reduces the utilization of memory and processing time. The superfluous data is removed without significant accuracy loss using type of disease. In this paper we have presented theoretical and experimental approach for knowledge reduction from large patient datasets using significance of attributes by organizing the data in discernibility and indiscernibility matrices. The experimental results demonstrate that the proposed parallel knowledge reduction method can efficiently process massive datasets on Hadoop platform, with highly speed up the grouping process and largely reduce the storage requirements. In all the experiments the introduced method based on significance of attributes is compared with the method based on positive region or information entropy. The comparison clearly shows that the former method outperforms the latter one.

IV. RESULTS AND DISCUSSION

In this section, we propose to examine the efficiency of using MapReduce for big data parallel knowledge reduction, as embodied by computing attribute importance and performing parallel search. Section 4.1 describes the datasets used to evaluate the method. Section 4.2 shows the details of hardware and software used in these experiments. Section 4.3 presents and discusses the experimental results of three different algorithms achieved.

4.1. Data sets

We have been studying on the analysis of patient data for several years. Patient data is available to us and the experiment is very meaningful. So we have selected patient big data for this experiment. Then, we regard a patient big data set as a patient knowledge representation system and analysis the specific condition attributes of decision attribute. The condition attribute is influence factors of Heart Attack and the decision attribute is Heart Attack. This experiment can find out the significant influence attributes which affect heart attack. The patient data decision table contains 33 attributes and 1 decision attribute. The condition attributes are the influence factors of heart attack. The decision attribute is heart attack. The purpose is to remove the irrelevant attributes and confirm the attributes more important. Table 1 shows a decision table of patient heart attack.

4.2. Hardware and software used

The experiments have been carried out on six nodes in a cluster. The master node and five compute nodes. Each one of these computer nodes has the following features:

i. Processors: Intel Core i3 3rd generation
ii. Cores: 4 per processor (8 threads)
iii. Network: Gigabit Ethernet
iv. Hard drive: 1 TB
v. RAM: 4 GB
vi. The specific details of the software used are the following:
vii. MapReduce implementation: Hadoop 2.6.0. MapReduce 1 runtime (Classic).
viii. Cloudera's open-source Apache Hadoop distribution.
ix. Maximum maps tasks: 33.
x. Maximum reducer tasks: 1.
xi. Operating system: Ubuntu 15 Version OS 6.4.
xii. Java SE Development Kit: JDK1.7

Table 1 A Decision Table of Patient Disease Information

S. No.	**Condition Attributes**					**Decision Attribute**
	Patient Type	**Disease Type**	**Early Signs**	**Heart Cough**	**..**	**Heart Attack**
1	0	0	0	0	..	0
2	1	1	1	1	..	1
3	0	2	2	0	..	2
:	:	:	:	:	:	:

1. Patient Type: 0 – In Patient, 1 – Out Patient
2. Disease Type: 0 – Cardiomyopathies, 1 - Coronary Artery, 2 – Diabetes, 3 - Heart Valves, 4 - Heart Defects present at Birth, 5 - High Blood Pressure, 6 - Lung Disease such as Emphysema, 6 - Past Heart Attacks
3. Early Signs: 0 - Chest Discomfort. It's the most common sign of heart danger, 1 - Nausea, Indigestion, Heartburn, or Stomach Pain, 2 - Pain that Spreads to the Arm, 3 - Dizzy or Lightheaded, 4 - Throat or Jaw Pain, 5 - Get Exhausted Easily, 6 – Snoring, 7 – Sweating
4. Heart Cough: 0 – Yes, 1 – No
5. ….
6. Heart Attack: 0 - ST segment elevation myocardial infarction (STEMI), 1 - Non-ST segment elevation myocardial infarction (NSTEMI), 2 - coronary spasm, or unstable angina

Though there are many factors that affect the heart attack, we have selected certain available factors only. Different characteristic attributes have different dimensions. The unit and the order of magnitude are usually different. Considering the impact of the difference of dimension and magnitude on the results of the model evaluation, data normalization method should be used to convert the different characteristic values into dimensionless values. In this paper, we have quantified the attributes first, and then make the data dimensionless which produced Table 1 as the result.

4.3 Experimental Analysis:

To evaluate the performance the knowledge reduction algorithm we have considered measurements reduction of data size that effects utilization of memory. A series of experiments are conducted on the dataset and compared the results among knowledge reduction algorithm using significance of attributes, positive region and information entropy.

Table 2 Performance Metrics of Data Size and Reduction Data Size Using Different Knowledge Reduction Techniques

Data Size in MB	**Positive Region**	**Information Entropy**	**Significance of Attributes**	**Reduction in MB**	
				% of Difference w.r.t. Positive Region	**% of Difference w.r.t. Information entropy**
70	67	66	65	2.99	1.52
80	76	75	73	3.95	2.67
100	93	91	88	5.38	3.3
110	100	94	90	10	4.26
120	108	99	94	12.96	5.05
130	114	104	98	14.04	5.77
140	122	108	101	17.21	6.48
150	127	112	104	18.11	7.14
160	132	116	106	19.7	8.62
170	137	119	107	21.9	10.08
180	155	127	114	26.45	10.24
190	152	128	111	26.97	13.28
200	160	131	113	29.38	13.74

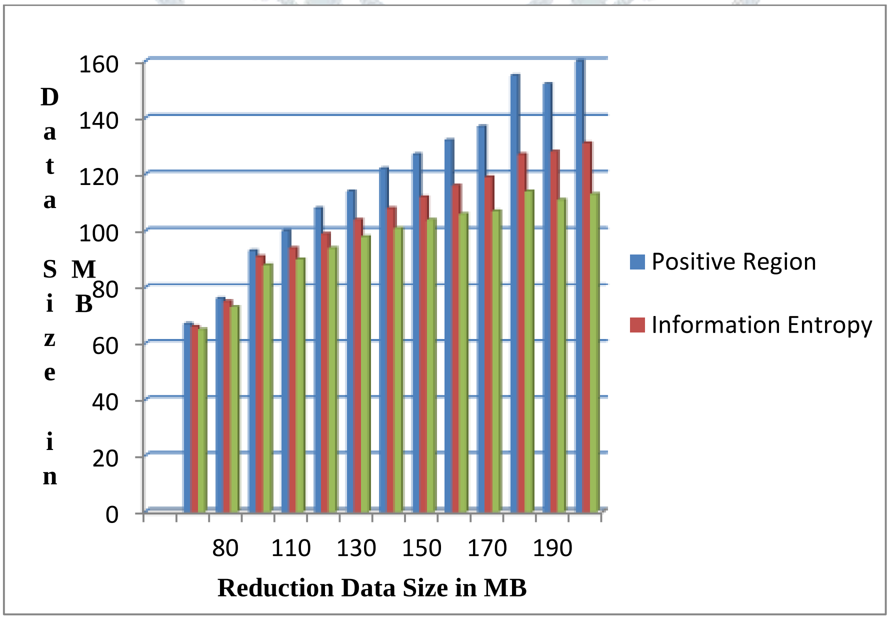

Figure 2 Comparison among different knowledge reduction algorithms

The performance metrics of core data size and reduction data size results are interesting when the starting core data size is 70 MB onwards, the corresponding reduction data size is 65 MB that affects utilization of memory. It shows that knowledge reduction data is giving better results than core data. When the data size is increasing, it shows that the knowledge reduction system using the significance of attributes technique is producing better results rather than positive region and information entropy techniques.

V. CONCLUSION

In this paper, we proposed knowledge reduction method using MapReduce that can handle big data. The MapReduce technique is an efficient computational model for distributed parallel processing with big data. The knowledge reduction algorithm based on significance of attributes using discernibility and indiscernibility matrices is successfully designed and is applied in the control experiment. The experimental results demonstrate that the knowledge reduction algorithm using MapReduce can scale well and efficiently process big data on Hadoop. The knowledge reduction algorithm based on significance of attributes can perform better than the knowledge reduction algorithm based on positive region or information entropy. Our future research work will focus on applications of the proposed parallel method in knowledge reduction using significance attributes based on rough sets.

VI. ACNOWLEDGEMENT

We thank Prof. Ch.S.V. Satyanarayana, Department of Mathematics, B.V. Raju College, Bhimavaram for assistance with Rough Set approach for comments that greatly improved the manuscript. We would also like to show our gratitude to the Dr. Ch.V. Srinivas, Principal, B.V. Raju College, Bhimavaram for sharing his pearls of wisdom with us during the course of this research.

Our sincere acknowledgements to the faculty of Vishnu Dental College, Bhimavaram for providing patients data for our analysis. This research was supported by B.V. Raju Foundation, Bhimavaram for providing the necessary facilities.

REFERENCES

[1] Weiping Cui, Lei Huang. A MapReduce Solution for Knowledge Reduction in Big Data. IJCSA; 2016 Jan. pp. 17-30.

[2] Jin Qian, Ping Lv, Xiaodong Yue, Caihui Liu, Zhengjun Jing. Hierarchical attribute reduction algorithms for big data using MapReduce. Qian; 2015 Sep. pp. 18-31.

[3] Yan Zhao, Yiyu Yao, Feng Luo. Data Analysis Based on Discernibility and Indiscernibility. Journal of Information Sciences, Elsevier; 2007 June, pp. 4959-4976.

[4] J. Qian, D.Q. Miao, Z.H. Zhang, W. Li. Hybrid approaches to attribute reduction based on indiscernibility and discernibility relation. International Journal of Approximate Reasoning, Elsevier; 2010 July, pp. 212-230.

[5] Yuhua Qian, Jiye Liang, and Chuangyin Dang. Incomplete multigranulation rough set. Systems Man & Cybernetics Part A Systems & Humans IEEE Transactions; 2010 Mar. pp. 420-431.

[6] Zdzislaw Pawlak, Andrzej Skowron. Rudiments of rough sets. Journal of Information Sciences, Elsevier; 2007 Sep. pp. 3-27.

[7] Zdzistaw Pawlak. Rough sets. International Journal of Computer & Information Sciences; 1982 Sep. pp. 341–356.

[8] Jerzy Blaszczynski, Roman Slowinski, Marcin Szelazg. Sequential covering rule induction algorithm for variable consistency rough set approaches. Journal of Information Sciences, Elsevier; 2010 Oct. pp. 987-1002.

[9] Ching-Hsue Cheng, Tai-Liang Chen, Liang-Ying Wei. A hybrid model based on rough sets theory and genetic algorithms for stock price forecasting. Journal of Information Sciences, Elsevier; 2010 Jan. pp. 1610-1629.

[10] Hongmei Chen, Tianrui Li, Shaojie Qiao, Da Ruan. A Rough Set Based Dynamic Maintenance Approach for Approximations in Coarsening and Refining Attribute Values. Jouranl of Inteligent Systems, Wiley Periodicals; 2010 Oct. pp. 1005-1026.

[11] Jeffrey Dean, Sanjay Ghemawat. MapReduce: Simplified Data Processing on Large Clusters. Communications of the ACM; 2008 Jan. pp. 107-113.

[12] Jeffrey Dean, Sanjay Ghemawat. MapReduce: A Flexible Data Processing Tool. communications of the ACM; 2010 Jan. pp. 72-77.

[13] Chen Degang, Wang Changzhong, Hu Qinghua. A new approach to attribute reduction of consistent and inconsistent covering decision systems with covering rough sets. Journal of Information Sciences, Elsevier; 2007 Feb. pp. 3500-3518.

[14] Zdzislaw Pawlak. Granularity of knowledge, indiscernibility and rough sets. Proceedings of the IEEE International Conference on Fuzzy Systems; 1998 Sep. 106-110.

[15] Guo-Fang Qiu, Wen-Xiu Zhang, Wei-Zhi Wu. Characterizations of attributes in generalized approximation representation spaces. Proceedings of the 10th International Conference on RSFDGrC; 2005 Aug. pp. 84-93.

[16] Liangxiu Han, Chee Sun Liew, Jano van Hemert, Malcolm Atkinson. A generic parallel processing model for facilitating data mining and integration. Jouranl of Parallel Computing, Elsevier; 2011 Feb. pp. 157-171.

[17] Isaac Triguero, Daniel Peralta, Jaume Bacardit, Salvador Garcia, Francisco Herrera. MRPR: A MapReduce solution for prototype reduction in big data classification. Neurocomputing; 2014 Oct. pp. 331-345.

[18] Ashwin Srinivasan, Tanveer A. Faruquie, Sachindra Joshi. Data and task parallelism in ILP using MapReduce. Jouranl of Machine Learning, Springer. 2011 May. pp. 141-168.

[19] Abhishek Verma, Xavier Llora, David E. Goldberg, Roy H. Campbell. Scaling genetic algorithms using MapReduce. Proceedings of the Ninth International Conference on Intelligent Systems Design and Applications, IEEE. 2009 Nov. pp. 13-18.

[20] Qinghua Hu, Daren Yu, Jinfu Liu, Congxin Wu. Neighborhood rough set based heterogeneous feature subset selection. Journal of Information Sciences, Elsevier. 2008 Sep. pp. 3577-3594.

[21] Daniel Zinn, Shawn Bowers, Sven Kohler, Bertram Ludascher. Parallelizing XML data-streaming workflows via MapReduce. Journal of Computer and System Sciences, Elsevier. 2010 Nov. pp. 447-463.

JOURNAL OF PURE AND APPLIED MICROBIOLOGY, Sept. 2018. Vol. **12**(3), p. 1667-1674

Isolation and Screening of Heavy Metal Resistant Microorganisms From Industrial Soil

V. Sridevi[1]*, M. Raghuram[1] and D. Ravisankar[2]

[1]Department of Microbiology, Acharya Nagarjuna University , Guntur - 522 508, India.
[2]Department of Biochemistry, Bhupathiraju Vissamraju College, Bhimavaram, India.

http://dx.doi.org/10.22207/JPAM.12.3.75

(Received: 12 August 2018; accepted: 11 September 2018)

Two industrial soils around west Godavari districts were analyzed for the concentration of nine heavy metals such as zinc, nickel, copper, manganese, cadmium, chromium, lead, iron and arsenic using atomic absorption spectrophotometer and simultaneously isolated nine morphologically distinct bacterial cultures from the same soil to check for the metal tolerance against copper, zinc and lead. The analyzed data revealed that iron and manganese metals were found to be the most abundant metals in these industrial soils and also noticed that the industrial soil 2 contained high amount of Cu, Zn and Pd. Control sample contained very low concentrations of the above mentioned heavy metals. That is therefore nine bacterial cultures (DPMC1-5 and TSC1-4) were checked for their metal resistance to Cu, Zn and Pb in nutrient broth at various metal concentrations. In in vitro application there was a drastically metal reduction observed in soils inoculated with nine bacterial strains. It can be concluded that this work will be the reference for the bioremediation of heavy metal polluted soils. As the industrialization in Bhimavaram creating the environmental pollution and hence treatment of industrial waste with metal resistant microorganisms called bioremediation is highly essential to clear the metal pollution instead of releasing industrial effluents into nearby water bodies or fields or choosing a costly chemical treatments.

Keywords: Industrial soils, heavy metal resistant bacteria, heavy metals, bioremediation.

West Godavari District has a richly cultivated land among all districts in Andhra Pradesh. West Godavari is popularly known as the Granary of India since about 50% of the state's rice production comes from the district. Cotton barrage was built on River Godavari at Dhavaleswaram channeling two canals, from which one canal passes through West Godavari, making the soil fertile. In the coastal belt of the district, a large portion of prawns and fish is exported to Japan, and the United States. Bhimavaram is a hub for Prawns export. It is the richest town in the State of Andhra Pradesh. Vendra paper mills in Bhimavaram, and Andhra Sugars, a sugar factory, in Tanuku are some of the more famous industries of the district. Due to the rapid industrialization and urbanization, many industries located in and around the district have released their untreated and semi treated effluents into the environment, more particularly to the nearby agriculture fields and water bodies. Even the pollutants such as heavy metals and other chemicals, which are present in the effluents moved through soil, surface water, sediments of the lake bed and percolated into ground water affecting the soil and groundwater quality. However, given significant industrialization and urbanization with ascendance of random dumping of untreated industrial wastewater and municipal sewage in to the environment, heavy metals have been reported to accumulate to unsafe or lethal concentrations[1-4]. High concentrations of heavy metals are dangerous to animals, plants and human cause many kinds of diseases. Indeed, even the purported essential components like Zn and Fe, if present in high concentration, are lethal in nature[5,6]. All heavy metal are not destructive to people,

* To whom all correspondence should be addressed.
Tel.: +91-8985144223;
E-mail: sridevivishnumolakala1@gmail.com

but rather at high focus levels they are poisonous in nature[7-10]. Studies have shown that toxic metals in soil influence plant growth[11]. Many studies show that soluble metal ions in the environment could be captured by microorganisms due to the negatively charged groups attached within their cell wall structure[12]. Bacteria, algae, and fungi or their separated components have been successfully used as bio-sorbents for heavy metal removal[13]. Heavy metal resistant bacteria represent a better alternative for heavy metal decontamination and have already been successfully applied for such purposes in the developed world[14-16]. In the present study, 9 different bacterial species were isolated from two different industrially polluted regions; they are Delta Paper Mill (DPMC1-5 colonies) and Tanuku Sugar Factory (TSC1-4 colonies), Bhimavaram. The metal removing ability of those isolates were investigated by in-vitro applications against copper, lead and zinc heavy metals. So in this investigation, an endeavor is made to set up the pathway of removal of heavy metals present in industrial polluted soils by metal resistant bacteria; thereby establishing the pathway of the bioremediation of metal polluted soils by microorganisms.

MATERIAL AND METHODS

Collection of Sample

Soil samples of Delta Paper Mill and Tanuku Sugar Factory from 2 different locations each; superficial and 6 inches depth were collected as polluted soil, Nearby agricultural soil was collected as cross contaminated soil and normal soil was collected as a control soil. Soil samples were collected with a sterile spatula in a polythene bags and were transported to the laboratory.

Preparation of Soil Sample

Under a sterile condition tenfold serial dilution was carried out in a laminar air flow. Stock suspension was prepared by adding 1 gm of the soil sample in 10 ml of distilled water in a first test tube for each soil sample and is vortexes. 1 ml of this solution is taken and transferred to the second test tube containing 10 ml of distilled water and is vortexes, labeled as $10^{-2}$and from second to third, it was labeled as $10^{-3}$and serial dilution was continued up to 10^{-8} . Hence each soil sample was serially diluted in a laminar air flow.

Isolation of Microorganisms

Nutrient agar medium have chosen to isolate bacteria from the soil samples. Media was prepared as per the composition and was sterilized in an autoclave conditions maintained was 121.5^0C temperature, 15 lbs pressure for 15minutes. After that media was poured in Petri plates and allowed to solidify and placed in an incubator at 37°C for 24 h in order to check its sterility.

Inoculation of Sample

An amount of 0.1 ml of each soil sample from selected dilutions (10^{-6}, 10^{-7} and 10^{-8}) was transferred on to previously labeled agar plates by spread plate method, each maintained in triplicate. Plates were incubated at 37^0C for 18-24hrs to obtain pure colonies.

Sub Culturing of Microorganisms

Isolated bacterial colonies with unique morphological characters were picked and sub cultured on fresh nutrient agar slants using sterile loop using streak plate method in laminar air flow to purify the isolates followed by incubation for 24 h at 37 °C and slants were stored at 4^0c for further investigation.

Analysis of Heavy metals in soil

Atomic absorption spectrometry was used for metal analysis**.** Soil samples were analyzed for heavy metal concentration for two times. First time metal analysis was done immediately after the collection of soil samples and second time metal analysis was done after the In vitro application (pot inoculation) to know whether the metals were reduced or not upon treating the metal inoculated soil with the isolated bacterial cultures.

Atomic Absorption spectrum

Model used was OPTIMA 8000, Category ICP-OES

Standard Preparation

Mixed elemental standard was prepared using 100 ppm is as follows:

0.5ppm-0.25ml of 100 ppm mixed standard transferred in to a 50 ml volumetric flask and diluted up to mark with milli Q water.

1ppm-05ml of 100 ppm mixed standard transferred in to a 50 ml volumetric flask and diluted up to mark with milli Q water.

1.5wg/L -0.75ml of 100 ppm mixed standard transferred in to a 50 ml volumetric flask and diluted up to marks with milli Q water.

Sample preparation

5 gm of soil sample was taken and transferred into a beaker to which 5 ml of concentrated HNO_3 and 20 ml of water was added. It was digested on a hot plate for about 30 minutes at 80^0 C; immediately it was cooled and diluted in 100 ml flask with water (Filtered the solution and aspirated it).

Determination of Metal resistance

To test for heavy metal resistance, the metals Cu, Zn and Pb, used in the form of $CuCl_2$, $ZnCl_2$, and $PbCl_2$, were added to sterile nutrient broth medium in varying concentrations viz: 5.0 ppm, 10.0, 20.0, 30.0, 40.0, 50.0, 60.0, 70.0, 80.0, 90.0, and 100.0 ppm/ml. Tubes were inoculated with an equal volume of selected bacterial cultures individually isolated from selected sites (Two from two different industries) and they were incubated at 37°C for 24 to 48 hours[17]. For comparing the growth response at different concentration, one set of nutrient broth (without the heavy metal) containing test tubes were inoculated with the selected organisms and used as control. Each dilution was maintained in triplicate for three times and Growth rate against metals were determined by taking O.D in turbidometer at 620nm.

In-vitro application (Pot inoculation)

For evaluation of bioremediation efficiency among selected bacterial cultures against copper, zinc and lead an attempt was made in the laboratory. Bioremediation of metal polluted soils was carried out in sterile clay pots. After exposure to bacterial cultures metal reduction were determined by atomic absorption spectrophotometer as described previously[18]. Briefly, 10 ml bacterial suspension which previously grown in 200 mL of sterile nutrient broth over night were inoculated in to sterile soil contaminated with 60ppm, 68ppm and 70ppm, 0f Cu, Zn and Pb respectively in sterile clay pots for 4 days. After incubation soil sample was dried in oven at 60°C for 10 hours. 10 mg dried soils were dissolved in 1 ml of concentrated nitric acid and diluted to 10 ml with DD/W. Blanks were treated in the same way and analyzed by atomic absorption spectrophotometer. Calibration curve of each heavy metal was drawn with working standard solution before testing. All the statistical analyses were carried out using Model no: Optima 8000 Spectrometer (Atomic): AES: ICP-OES.

RESULTS AND DISCUSSION

ASS model No: ST-EQ-084 was used for determination of the heavy metal contents of the contaminated soils. Soil samples were taken from 4 different locations; in the vicinity of a heavy metal industrial area in Vendra, Tanuku, Agricultural field near by industrial effluent out let and soil away

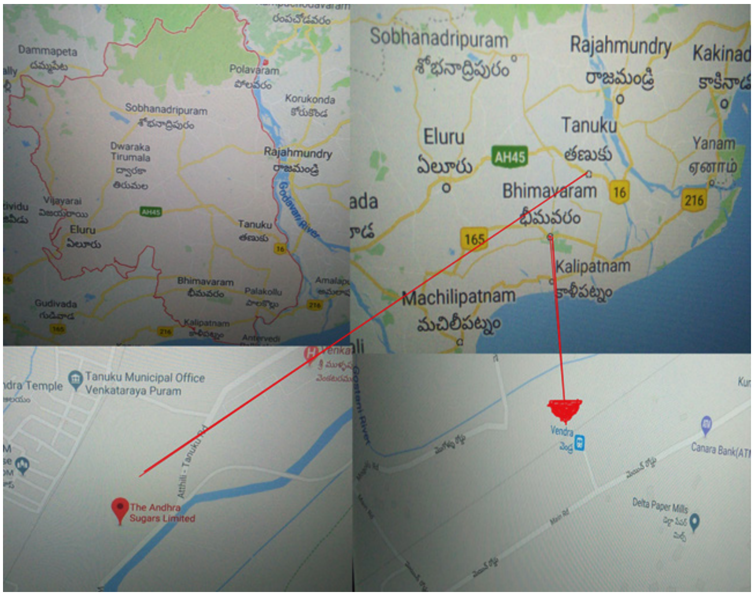

Fig. 1. Location

from pollutions to take as control; Bhimavaram. (Figure 1).

The heavy metal concentrations of these soils were included in the Table1. The concentrations of Zn, Pb,Cd,Ni,Cr,Cu,Fe and Mn in the industrial soils and agricultural soils were high as compared with the values of the control soil. However, the concentration of As and Cd in the soils was found to be BDL when compared with the Agricutrl soil and control soil. Fe was found to be more in agricultural soil, and it was also be obseved that there was a major difference in the concentration of Zn, Pb Cu and Mn between two industrial samples i.e 21ppm in DPM soil and 480 ppm in TSC soil, 50ppm in DPM soil and 100ppm in TSC soil, 320ppm in DPM soil and 570ppm in TSC soil and 820ppm in DPM soil and 270 in TSC soil respectively. The concentration of Zn, Cu and Pb was high in TSC soils than in DPM soils hence an attempt was made for bioremediation of these three metals in metal polluted TSC soils by the isolation of bacterrial cultures against Cu, Zn and Pb specifically taking them in its chloride forms as CuCl2, ZnCl2 and PbCl2.

Two different soil samples were innoculated on general purpose medium that is nutrient agar the temperature and pH of the collected samples were maintained at constant conditions that is 37°C and 7.2 respectively. The total bacterial count of each sample are given in Table 2. The total bacterial count ranged from 06 x 10^7 to 256 x 10^5(cfu/10 ml of sample).

During the period of study, a total number of 256 bacterial colonies were isolated from two industrial samples of two different locations each. Out of the 256 isolates, only 18 isolates were selected for further study on the basis of their morphological diversities. Among 18, 09 isolates, showed a different ranges of metal resistant against three different heavy metals namely $CuCl_2$, $ZnCl_2$ and $PbCl_2$ from a minium concentration of 10ppm to 100ppm. At the point when common habitats are discharged with metal pollutants, the indigenous microbial networks are probably going to contain microbial population of various taxonomic qualities, which are equipped for reducing the discharged chemicals [19]. Since industrial effluents contain extensive measure of substantial metals, the

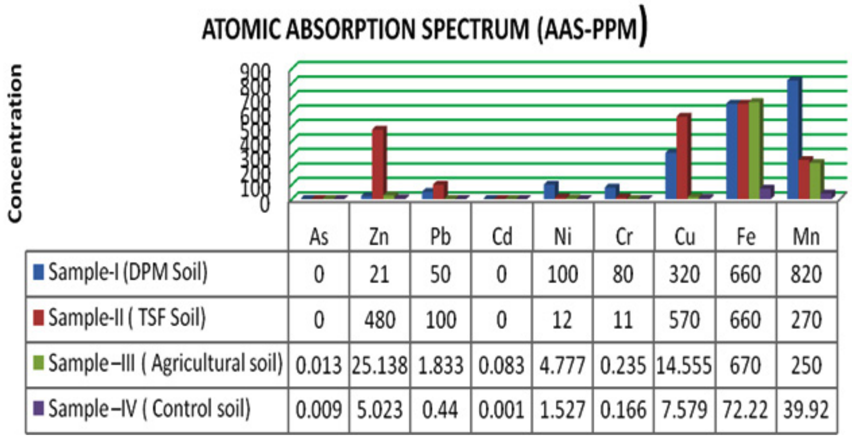

Fig. 2. Atomic Absorption Spectrum (AAS - PPM)

Table 1. Heavy metal analysis data (Atomic Absorption spectrum)

Soil sample type	As	Zn	Pb	Cd	Ni	Cr	Cu	Fe	Mn
Sample-I (DPM Soil)	BDL	21	50	BDL	10	80	320	660	820
Sample-II (TSC Soil)	BDL	480	100	BDL	12	11	570	660	270
Sample –III (Agricultural soil)	0.013	25.138	1.833	0.083	4.777	0.2350	14.555	670	250
Sample –IV (Control soil)	0.009	5.023	0.44	0.001	1.527	0.166	7.579	72.22	39.92

impact of Cu, Zn and Pb on bacterial development and the capacity to detoxify the toxins in vitro was the critical piece of the present examination. Metal resistance among microorganisms against Cu, Zn and Pb was considered evidently. To study metal resistance among selected colonies, initially bacterial cultures were allowed to grown on Nutrient medium contain Cu and Pb but the results got on the medium containing Cu was unsatisfactory since Cu can bound the agar components [20,21]. That is therefore in our investigation, the test was done in nutrient broth supplemented with Cu, Zn and Pb of various dilutions ranges from 10 ppm to 100 ppm. It was discovered that, the development of all the bacterial growth diminished with increasing metal concentration and the growth of each dilution were maintained in three replicates for three times and O.D was measured in turbidimetry at 620 nm and results were tabulated (Table 3, 4and 5)

Isolate DPMC1 showed highest resistant against $CuCl_2$ than $ZnCl_2$ and $PbCl_2$, where as TSC4 showed metal resistant against $CuCl_2$, $ZnCl_2$ and

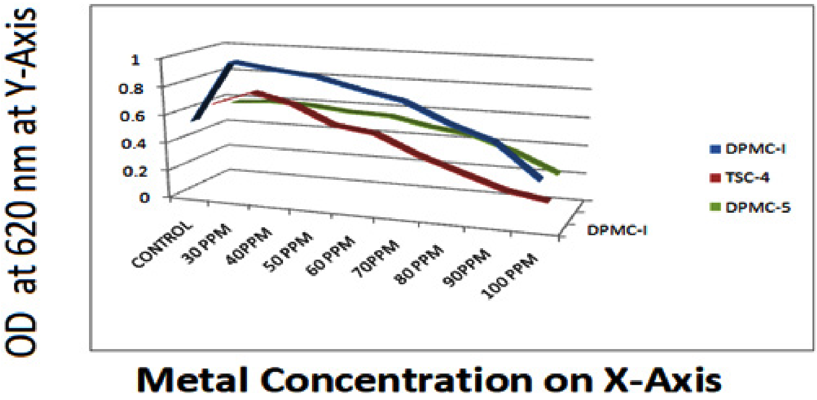

Fig. 3. Effect of copper on bacterial growth (OD at 620nm)

Table 2. Microbial count of the industrial soil samples

Sample types	Location	p H	Temperature	Colony forming units
DPM	Site -I	7.0	37^0C	$256x10^5$
	Site-II	7.0	37^0C	$83x10^7$
TSC	Site-I	7.0	37^0c	$10x10^6$
	Site-II	7.0	37^0c	$6x10^7$

Table 3. Effect of Copper on bacterial growth

S. No	Culture type	Control	$CuCl_2$ (30ppm/ml)	$CuCl_2$ (40ppm/ml)	$CuCl_2$ (50ppm/ml)	$CuCl_2$ (60ppm/ml)	$CuCl_2$ (70ppm/ml)	$CuCl_2$ (80ppm/ml)	$CuCl_2$ (90ppm/ml)	$CuCl_2$ (100ppm/ml)
01	DPMC-I	0.55	0.98	0.94	0.91	0.84	0.78	0.65	0.55	0.32
02	DPMC-2	0.49	0.57	0.58	0.48	0.38	0.38	0.22	0.10	0.06
03	DPMC-3	0.54	0.58	0.52	0.48	0.46	0.38	0.26	0.10	0.09
04	DPMC-4	0.56	0.60	0.55	0.55	0.44	0.32	0.22	0.10	0.07
05	DPMC-5	0.58	0.60	0.59	0.56	0.55	0.49	0.45	0.35	0.22
06	TSC-1	0.49	0.48	0.37	0.28	0.27	0.23	0.14	0.04	0.00
07	TSC-2	0.40	0.47	0.28	0.23	0.18	0.13	0.09	0.00	0.00
08	TSC-3	0.45	0.50	0.45	0.35	0.30	0.18	0.14	0.06	0.00
09	TSC-4	0.62	0.72	0.65	0.52	0.48	0.35	0.25	0.15	0.10

$PbCl_2$. DPMC2 was showed metal resistant agisnt $CuCl_2$ and $ZnCl_2$ and least resistant among all against $PbCl_2$.

DPMC3 and TSC4 have showed metal resistant more against $ZnCl_2$ than other cultures. DPMC4 and DPMC5 have showed metal resistant more agiainst $CuCl_2$ and $PbCl_2$ and least resistant against $ZnCl_2$ among all. TSC2 have showed least resistant among all for 3 heavy metals. TSC1 showed metal resistant agints Zncl2 and Pbcl2 and least against $CuCl_2$ like that of TSC2.

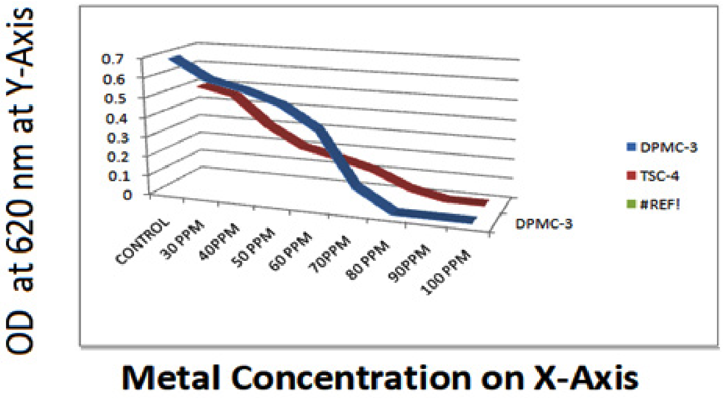

Fig. 4. Effect of Zinc on bacterial growth (OD at 620nm)

Table 4. Effect of Zinc on bacterial growth (OD at 620nm)

S. No	Culture type	$ZnCl_2$ (20ppm/ml)	$ZnCl_2$ (30ppm/ml)	$ZnCl_2$ (40ppm/ml)	$ZnCl_2$ (50ppm/ml)	$ZnCl_2$ (60ppm/ml)	$ZnCl_2$ (70ppm/ml)	$ZnCl_2$ (80ppm/ml)	$ZnCl_2$ (90ppm/ml)	$ZnCl_2$ (100ppm/ml)
01	DPMC-I	0.45	0.38	0.21	0.06	0.00	0.00	0.00	0.00	0.00
02	DPMC-2	0.43	0.34	0.20	0.06	0.00	0.00	0.00	0.00	0.00
03	DPMC-3	0.69	0.59	0.55	0.49	0.38	0.11	0.00	0.00	0.00
04	DPMC-4	0.48	0.43	0.27	0.09	0.00	0.00	0.00	0.00	0.00
05	DPMC-5	0.40	0.37	0.25	0.07	0.00	0.00	0.00	0.00	0.00
06	TSC-1	0.61	0.50	0.31	0.14	0.09	0.05	0.00	0.00	0.00
07	TSC-2	0.43	0.36	0.23	0.16	0.10	0.01	0.00	0.00	0.00
08	TSC-3	0.39	0.26	0.20	0.15	0.07	0.05	0.00	0.00	0.00
09	TSC-4	0.50	0.46	0.31	0.21	0.17	0.12	0.04	0.00	0.00

Table 5. Effect of Lead on bacterial growth (OD at 620nm)

S. No	Culture type	$PbCl_2$ (20ppm/ml)	$PbCl_2$ (30ppm/ml)	$PbCl_2$ (40ppm/ml)	$PbCl_2$ (50ppm/ml)	$PbCl_2$ (60ppm/ml)	$PbCl_2$ (70ppm/ml)	$PbCl_2$ (80ppm/ml)	$PbCl_2$ (90ppm/ml)	$ZnCl_2$ (100ppm/ml)
01	DPMC-I	0.33	0.22	0.20	0.15	0.14	0.00	0.00	0.00	0.00
02	DPMC-2	0.25	0.19	0.20	0.15	0.00	0.00	0.00	0.00	0.00
03	DPMC-3	0.31	0.31	0.18	0.00	0.00	0.00	0.00	0.00	0.00
04	DPMC-4	0.49	0.34	0.23	0.19	0.00	0.00	0.00	0.00	0.00
05	DPMC-5	0.50	0.28	0.15	0.10	0.00	0.00	0.00	0.00	0.00
06	TSC-1	0.42	0.33	0.25	0.10	0.06	0.00	0.00	0.00	0.00
07	TSC-2	0.48	0.30	0.20	0.10	0.02	0.00	0.00	0.00	0.00
08	TSC-3	0.55	0.47	0.28	0.17	0.01	0.00	0.00	0.00	0.00
09	TSC-4	0.64	0.55	0.46	0.39	0.25	0.16	0.09	0.00	0.00

The study was extended next to in vitro application for proving the cultures can be adopt for bioremediation of metal polluted soil. The data regarding in vitro application was tabulated in Table 6. There was a drastical metal reduction was observed in the atomic absorption report. And

Table 6. *In vitro* Application Report

S. No	Type of Bacterial culture	Copper	Zinc	Lead
1	DPMC1	5.102	8.082	10.160
2	DPMC2	7.290	7.748	10.582
3	DPMC3	8.102	5.780	8.020
4	DPMC4	7.956	10.918	7.810
5	DPMC5	7.028	10.748	7.688
6	TSC1	9.160	6.796	6.582
7	TSC2	9.580	10.510	8.802
8	TSC3	8.218	7.494	6.788
9	TSC4	5.782	5.536	5.780

it was obsreved and proved that the cultures which showed metal resistant in tube dilutions more, have also be reduced metal concentration in in vitro application (Fig. 6). Additionally investigation of the impacts of various supplements and conditions in their development is expected to distinguish their effectiveness as bioremediation specialists, where optimization of pH, temperature, and incubation time can impact metal obstruction limit[22] .

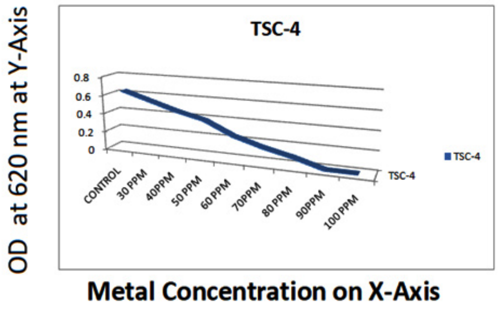

Fig. 5. Effect of Lead on bacterial growth (OD at 620nm)

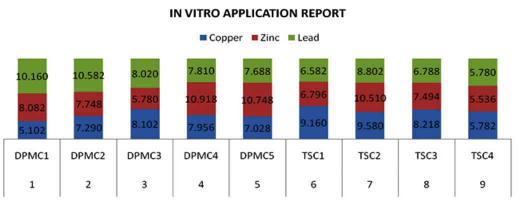

Fig. 6. *In vitro* Application report

CONCLUSION

The polluted soils are high medium to develop and spread of resistant microbial populations, which are resistant to various metals. The recognizable proof of opposition against various metals may give a valuable device to the synchronous observing of a few harmful contamination in the earth. It was demonstrated that industrial waste are the source of highly resistant forms. All the bacterial strains isolated in this investigation demonstrated against copper, zinc and lead have showed high levels of metal resistance. The future prospect lies in the use of these microorganisms for purposes like overwhelming metal remediation of metal polluted soils. All the results formulated in this study proved that DPMC1, DPMC5 and TSC4 cultures showed varing degree of metal resistant against Cu, Pb and Zn and that is therefore these cultures might be helpful for the detoxification of heavy metal pollution in the industrial effluents at industrial surroundings.

ACKNOWLEDGEMENTS

The authors are highly appreciative to the Department of Microbiology Acharya nagarjuna University Nagarjuna Nagar, ASN aqua labs Bhimavaram and B.V.Raju College Vishnupur Bhimavaram for their help and continues encouragement in the fruit full completion of this project work.

REFERENCES

1. Iwegbue, C. M. A., Nwajei, G. E., Ogala, J. E. and Overah, C. L. Determination of trace metal concentrations in soil profiles of municipal waste dumps in Nigeria. *Environmental Geochemistry and Health,* 2010; **32**: 415-430.
2. Jarup, L., Berglund, M., Elinder, C. G., Nordberg, G., Vahter, M., Health effects of cadmium exposure—a review of the literature and a risk estimate. *Scand. J. Work Environ. Health,* 1998; **24**: 1–52.
3. Kumar, A., Bisht, B. S. and Joshi, V. D. Bioremediation potential of three acclimated bacteria with reference to heavy metal removal from waste. *International Journal of Environmental Sciences,* 2011; **2**: 896-908.
4. Nies, D. H., Microbial heavy-metal resistance.

Applied Microbiology and Biotechnology, 1999; **51**: 730-750.
5. Aswathanarayana, U., Geoenvironment: an introduction, Rotterdam, the Netherlands, A.A. Balkema Publishers, 1995; pp: 270.
6. Govil, P.K., Reddy, G.L.N., and Krishna A.K., Soil contamination due to heavy metals in Patancheru industrial development area Environmental Geology, 2001; **41**: pp: 461-469.
7. Salem, H.M.; Eweida, E.A.; Farag, A. Heavy metals in drinking water and their environmental impact on human health. In ICEHM 2000; Cairo University: Giza, Egypt, 2000; pp. 542–556.
8. Wuana, R.A.; Okieimen, F.E. Heavy metals in contaminated soils: A review of sources, chemistry, risks and best available strategies for remediation. *ISRN Ecol*. 2011, Article 20
9. Padmavathiamma, P.K.; Li, L.Y. Phyto remediation technology: Hyper accumulation metals in plants. *Water Air Soil Poll.* 2007, **184**: 105–126.
10. Gadd, G.M. Bioremedial potential of microbial mechanisms of metal mobilization and immobilization. *Curr. Opin. Biotechnol.* 2000, **11**: 271–279.
11. Singh Jiwan;Kalamdhad Ajay S. Effects of Heavy Metals on Soil, Plants, Human Health and Aquatic Life: *Int. J. Res. Chem. Environ.* 2011; **1**(2); 15-21.
12. Doyle R, Matthews T, Streips U. Chemical basis for selectivity of metal ions by the Bacillus subtilis cell wall. *J Bacteriol,* 1980; : 471-480.
13. Wang L, Chua H, Zhou Q et al. Role of cell surface components on Cu2+ adsorption by Pseudomonas putida 5-x isolated from electroplating effluent. *Water Res,.* 2003; **37**: 561-568.
14. Nakamura K., Hagimine M., Sakai M. and Furukawa K. Removal of mercury from mercury-contaminated sediments using a combined method of chemical leaching and volatilization of mercury by bacteria. *Biodegradation,* 1999; **10**: 443-447.
15. Okino, S., Iwasaki, K., Yagi, O. and Tanaka, H. Removal of mercuric chloride by immobilized cells of genetically engineered mercury-volatilizing bacterium Pseudomonas putida Pp Y101/pS134. Bill. Environ. *Contam. Toxicol.* 2002; **68**: 712-719.
16. Sinha, K. R., Valani, D., Sinha, S., Singh, S. and Herat, S. Bioremediation of contaminated sites: A low-cost nature biotechnology for environmental clean-up by versatile microbes, plants and earthworms. In: Solid waste management and environmental remediation. *Timo Faerber and Johann Herzog*. 2009; ISBN: 978-1-60741-761-3.
17. Calomoris JJ, Armstrong TL & Seidler RJ. Association of metal- tolerance with multiple antibiotic resistances of bacteria isolated from drinking water. *Appl Environ Microbiol.,* 1984; **47**: 1238-1242.
18. Shakibaie, M. R., Dhakephallker, P. K., Kapadnis, B. P., Chopade, B. A., Removal of Silver from Phothographic waste Water effluents by Acinetobaeter baumnannii BL54: *Can. J. Microbiol,* 1999; **45**: 995-1000.
19. Bartha R & Atlas RM. The microbiology of aquatic oil spills. *Adv Appl Microbiol.* 1977; **22**: 225-266.
20. Ramamoorthy S & Kushnor DJ. Binding of mercuric and other metal ions by microbial growth media. *Microbial Ecol.* 1975; **2**: 162-175. 16.
21. Washington II JA, Snyder RJ, Kohner PC, Wiltse CG, Itstrup DM & McCall JT. Effect of cation content of agar on the activity of gentamycin, tobramycin and amikacin against Pseudomonas aeruginosa. *J. Infectious Diseases.* 1978; **137**: 103-111.
22. Shivakumar et al., C. K. Shivakumar, B. Thippeswamy, M. Krishnappa Optimization of heavy metals bioaccumulation in Aspergillus niger and Aspergillus flavus *Int. J. Environ. Biol.,* 2014 ; **4**(2): pp. 188-195.