Vol 5 Issue 3 April 2015

ISSN No : 2230-7850

International Multidisciplinary Research Journal

Indian Streams Research Journal

Executive Editor
Ashok Yakkaldevi

Editor-in-Chief
H.N.Jagtap

Welcome to ISRJ

RNI MAHMUL/2011/38595 **ISSN No.2230-7850**

Indian Streams Research Journal is a multidisciplinary research journal, published monthly in English, Hindi & Marathi Language. All research papers submitted to the journal will be double - blind peer reviewed referred by members of the editorial board.Readers will include investigator in universities, research institutes government and industry with research interest in the general subjects.

International Advisory Board

Editorial Board

Address:-Ashok Yakkaldevi 258/34, Raviwar Peth, Solapur - 413 005 Maharashtra, India
Cell : 9595 359 435, Ph No: 02172372010 Email: ayisrj@yahoo.in Website: www.isrj.org

Indian Streams Research Journal ISSN 2230-7850 Impact Factor : 3.1560(UIF)
Volume-5 | Issue-3 | April-2015 Available online at www.isrj.org

APPLICATIONS OF ANOMALY DETECTION TECHNIQUE IN RECENT TRENDS

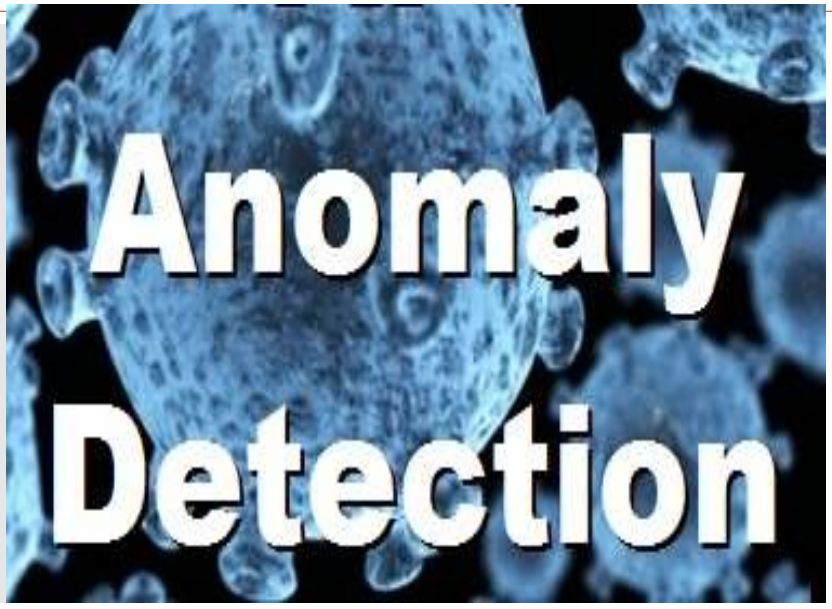
Jabez J.
Research Scholar, Sathyabama University, Chennai, India.

Short Profile

Jabez J. is a Research Scholar at Sathyabama University, Chennai, India. He has completed M.A. He has professional experience of 15 years.

Co-Author Details :

B. Muthu Kumar
Professor , Faculty of Computing, Sathyabama University,Chennai, India.



ABSTRACT:
Anomaly Detection Technique is one of the most valuable methods which can be applied in many fields in our day-to-day life. The technique can be used in numerous ways by developing several algorithms to find anomalies. To detect anomalies the main aspect is to identify accurate attributes. One has to pick a proper attribute based on the application requirement. Different application uses different attributes to predict anomalies. For e.g., anomalies in a network data can be predicted by attributes like arrival time of packets, the size of the packet, etc., Forest fire can be predicted by the anomalies that occur in the abiotic factors like temperature, sound , intensity of light etc.,. This paper describes a study on the general ideas of how the anomalies can be applied in various fields.

KEYWORDS
Anomaly Detection Technique , abiotic factors , anomalies.

1.INTRODUCTION:

In data mining, the recognition of objects is known as anomaly detection (or outlier detection), events or observations which do not match to an expected pattern or other items in a dataset. Generally, the anomalous objects can interpret with some types of difficulty such as bank fraud, a structural defect, medical problems or finding errors in text. Anomalies are also referred to as outliers, novelties, noise, deviations and exceptions. In particular, the perspective of maltreatment and network intrusion detection, the interesting objects are frequently not rare objects but unpredicted bursts in activity. This outline does not stick a general arithmetic description of an outlier as an uncommon object where many outlier discovery methods will be unsuccessful on such information, unless it has been combined properly. As an alternative, a cluster analysis algorithm may be capable to identify the micro clusters produced by these outlines.

At present three wide classes of anomaly discovery methods survive they are unsupervised . Unsupervised anomaly detection techniques discover anomalies in an unlabeled experimental data set under the hypothesis that can appear to fit smallest amount to the remnants of the data set. Supervised anomaly detection techniques necessitate a data set that is labeled as "normal" and "abnormal" and involves training a classifier. Semi-supervised anomaly recognition method build a replica instead of usual performance from a given regular training data set, and then testing the probability of a test example to be produced by the learnt representation.

Anomaly detection method is valid in a diversity of domains, such as intrusion detection, fraud detection, fault detection, system health monitoring, event detection in sensor networks, and detecting Eco-system disturbances. It is frequently used in preprocessing to take away anomalous data from the dataset. In supervised learning, eliminating the anomalous data from the dataset repeatedly results in a statistically important augment in correctness.

2.ARCHITECTURE OF ANOMALY DETECTION TECHNIQUE

Anomaly Detection can be applied in various fields based on the situation several algorithm can be developed. This algorithm decides the speed, accuracy of the anomaly pattern. But in general the following will be architecture of the anomaly detection technique.

Thus the input can be certain attributes depending on the application and situation where we are using anomaly detection. The efficiency of the detection rate depends on the efficiency of the algorithm. Output part is to justify whether to anomaly or not. This is shown in following Figure 1.

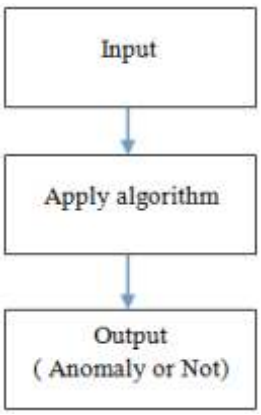


Figure 1: Architecture Diagram

3.DESCRIPTION ON APPLICATIONS OF ANOMALY DETECTION

3.1.Intrusion Detection

An Anomaly-Based Intrusion Detection System is a framework for detecting cyber intrusions and mistreatment by observing framework activity and categorizing it as either ordinary or abnormal. The categorization is based on heuristics or regulations, rather than outlines or marks, and tries to identify any kind of mishandling that cascades out of regular framework process. This is as contrasting to signature-based frameworks, which can only identify attacks for which a signature has been created beforehand.

In order to establish what attack transfer is, the framework must be qualified to distinguish typical framework actions. This can be done in numerous ways, most frequently with artificial intelligence sort of methods. Frameworks utilizing neural networks were been utilized to huge outcome. Another technique describes what typical usage of the framework includes utilizing a strict mathematical model, and flag any deviation from this as an attack. This is known as strict anomaly detection.

Anomaly-based Intrusion Detection does have a few limitations, specifically a high false-positive rate and the capability to be tricked by a properly transporting attack. Efforts were made to tackle these problems through methods utilized by PAYL and MCPAD.

Another researcher Jabez and Muthukumar [5] have proposed an intrusion detection framework which utilizes novel techniques like time probability based pattern detection method and Hyperbolic Hopfield Neural Network for detecting the intrusion from the real time network datasets. Their experimental results showed their proposed framework had obtained improved results than other frameworks.

Fraud detection is similar as that of intrusion detection except that the biometrics may also be involved in this kind of detection.

3.2.Anomaly detection in system health monitoring

In today's environment many types of diseases attack not only human even plants and animals. Hence we need an effective anomaly detection system to predict whether it is an anomaly or not.

The genes can be represented systematically by using dataset with several attributes. By using genetic algorithm and fuzzy logic we can predict the genetically disorder (if any). This is also one important anomaly detection application.

Many researchers presented numerous works in this field of system health monitoring area, one of which done by David L. Iverson [2] is stated below.

David has proposed the Inductive Monitoring System (IMS) software which was developed to offer a method to automatically create health checking acquaintance bases for frameworks that are either tricky to represent (simulate) with a computer or which necessitate computer replicas that are too composite to utilize for real time scrutinizing. IMS utilizes nominal data sets composed either directly from the framework or from simulations to make a knowledge base that can be utilized to identify anomalous actions in the framework. Machine learning and data mining methods are utilized to distinguish distinctive framework performance by removing common classes of insignificant data from archived data sets. He states that IMS is would be able to scrutinize the framework by evaluating real time operational data with these classes. He presents an explanation of learning and scrutinizing

techniques utilized by IMS and reviewing some current IMS outcomes.

3.3.Event Detection in Sensor Networks

Krasimira Kapitanova [7] has developed Event detection which is an essential part in many wireless sensor network (WSN) applications. Inspite of this, the area of event explanation has not been given sufficient concentration. The preponderance of existing event explanation approaches depends on utilizing exact values to identify event thresholds. However, they believed that crisp values cannot sufficiently hold the frequently in exact sensor analysis. They had demonstrated the utilization of fuzzy values in its place of crisp ones considering the advances of accurateness in event discovery. They also showed that their fuzzy logic approach offer superior recognition accuracy than a few of well established categorization algorithms. A drawback of utilizing fuzzy logic is the exponentially rising size of the rule-base. Sensor nodes have inadequate memory and accumulation of large rule-bases could be a challenge. To concentrate on this problem they developed a number of methods that help to decrease the size of the rule-base by more than 70% while conserving the level of event recognition accurateness respectively.

3.4.Detecting Eco-system Disturbances

Since Ecosystem being the most significant factor of life researchers has developed many techniques for the detection of disturbances in Ecosystem, a few of which are stated below.

Haibin Cheng [3] and his coworkers presented a case study on the application of data mining to the difficulty of identifying ecosystem conflicts from vegetation cover information gained through satellite observations. They described two anomaly detection approaches—moving average and random walk—for detecting such actions. They also illustrated how clustering can be utilized to position related occurrence of interruption events. Finally, they present a clustering-based system to help the visual investigation of ecosystem disturbances from high resolution data.

Another research work by Christopher Potter et al 2003 [1] on Ecosystem was proposed were scientists have yet to develop a verified methods to scrutinize and know key interruption actions and their past establishments at a worldwide scale. Their analysis was accomplished to estimate patterns in an 18-year documentation of worldwide satellite remarks of vegetation phenology from the Advanced Very High Resolution Radiometer (AVHRR) as a means to distinguish major ecosystem disturbance events and regimes. The Fraction Absorbed of Photosynthetically Active Radiation (FPAR) by vegetation canopies global was calculated at a monthly time gap from 1982 to 1999 and gridded at a spatial resolution of 0.5o latitude/longitude. Possible interruption actions of great level (>0.5 Mha) were recognized in the FPAR time sequence by finding anomalously low values (FPAR-LO) that lasted longer than 12 successive months at any 0.50 pixel. They discovered that almost 400 Mha of the worldwide land surface could be recognized with at least single FPAR-LO event over the 18-year time sequence. The majority of these possible interruption actions happen in tropical savanna and scrublands or in boreal forest ecosystem classes. Authentication of possible trouble actions from their FPAR-LO analysis was done utilizing documented accounts of the timing of large-scale wildfires at locations throughout the world. Interruptions regimes were further illustrated by association scrutiny with past weather anomalies. Assuming correctness of the FPAR satellite documentation to illustrate key ecosystem interruption actions, they estimated that almost 9 Pg of carbon has been missing from the terrestrial biosphere to the atmosphere as an outcome of large-scale ecosystem interruptions over this 18- year

time sequence.

3.5.Prediction of Forest Fire

Forest is one of the very important resources needed to human. Nowadays due to the overpopulation there is heavy loss of forest. In order make their shelter, sophisticated life, modern technology. There is many involve in illegal activities to destroy the resource of forest by fire and then using those land for some other purpose. Hence we need an effective system to predict the forest fire.

To detect this forest fire we can choose attributes intensity of light, Sound, temperature as attributes [6]. The following explains the method to detect forest fire.

Let 'X_i' be the temperature at any instant of time.
Let the average temperature at 'n' number of readings is denoted as 'P'
Therefore,

$$P = \sum_{i=0}^{n} X_i$$

Let 'Y_i' be the intensity of light at any instant of time.
Let the average temperature at 'n' number of readings is denoted as 'Q'.
Therefore,

$$Q = \sum_{j=1}^{n} Y_j$$

Let 'Z_k' be the sound at any instant of time
Let the average temperature at 'n' number of readings is denoted as 'R'.
Therefore,

$$R = \sum_{k=1}^{n} Z_k$$

In Normal Condition:

If (P < α or Q < ß or R < ?)

No Alarm

Where a ,ß,? are the threshold value range of temperature, Intensity of Light and sound respectively.

In Abnormal Condition: (i.e., during fire):

If (P > a && Q > ß && R > ?)

Alarm
Thus this is one of the simplest methods for finding anomaly in forest.

3.6.Anomaly Detection Machinery

"Anomaly detection can diagnose faults within transformers can greatly assist decision making on such issues as maintenance, performance and safety" [4]. In the power plant multi-agent system technology can be used as the underpinning platform for such condition monitoring systems can be used [8].

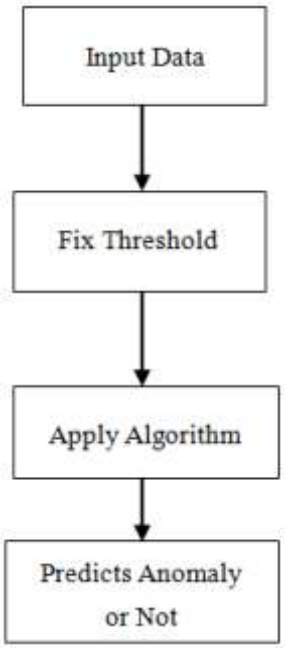


Figure 2: Flow Diagram

For any type of Anomaly the following steps can be applied:

Input data may be of any form it can be from any source like a sensor or real time scenarios. For all data a threshold is been fixed by us, this threshold may vary depending on the application of using, the efficiency on the output is depended on the algorithm what we use. Finally the anomalies are detected signinificantly.

4.CONCLUSION

Thus the above applications show the importance and accuracy of anomaly detection in several fields. This follows the generalized architecture for all fields. However keeping the necessity of anomalies in mind any user may develop any efficient algorithm by choosing the right attributes depends on the applications in highly effective.

5.REFERENCES

1.Christopher Potter, Pang-ning Tan, Michael Steinbach, Steven Klooster, Vipin Kumar, Ranga Myneni

and Vanessa Genovese. 2003. "Major disturbance events in terrestrial ecosystems detected using global satellite data sets." Global Change Biology.

2 .David L. Iverson. 2004. "Inductive System Health Monitoring", Proceedings of The International Conference on Artificial Intelligence (IC-AI'04), CSREA Press, Las Vegas, NV.

3.Haibin Cheng, Pang-Ning Tan, Christopher Potter and Steven Klooster. 2008. "Data mining for visual exploration and detection of ecosystem disturbances, Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems GIS'08.

4.IEEE. 2011. "Anomaly detection in power generation plants using similarity-based modeling and multivariate analysis", American Control Conference (ACC), ISSN : 0743-1619.

5.Jabez, J and Muthukumar, B. 2014. "Intrusion detection system: Time probability method and hyperbolic Hopfield neural network", Journal of Theoretical and Applied Information Technology, Volume 67: 1.

6 .Jabez , J, Anandha Mala.2012. G.S. " A Novel Fuzzy method for detecting anomalous events in Forest Ecosystem". NC4T'12, ISBN:978-81-922119-5-4, Sathyabama Univesity. 111–113.

7.Krasimira Kapitanova, Sang H. Son and Kyoung-Don Kang. 2010."Event Detection in Wireless Sensor Networks-Can Fuzzy Values Be Accurate?", Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Volume 49:168-184.

8.Power Systems. 2005. IEEE Transactions,. An agent-based anomaly detection architecture for condition monitoring, ISSN : 0885-8950. Volume:20 , Issue: 4,1675 – 1682

Publish Research Article
International Level Multidisciplinary Research Journal
For All Subjects

Dear Sir/Mam,

We invite unpublished Research Paper,Summary of Research Project,Theses,Books and Book Review for publication,you will be pleased to know that our journals are

Associated and Indexed,India

- International Scientific Journal Consortium
- OPEN J-GATE

Associated and Indexed,USA

- Google Scholar
- EBSCO
- DOAJ
- Index Copernicus
- Publication Index
- Academic Journal Database
- Contemporary Research Index
- Academic Paper Databse
- Digital Journals Database
- Current Index to Scholarly Journals
- Elite Scientific Journal Archive
- Directory Of Academic Resources
- Scholar Journal Index
- Recent Science Index
- Scientific Resources Database
- Directory Of Research Journal Indexing

Indian Streams Research Journal
258/34 Raviwar Peth Solapur-413005,Maharashtra
Contact-9595359435
E-Mail-ayisrj@yahoo.in/ayisrj2011@gmail.com
Website : www.isrj.org