Anomaly Detection for Cognitive Intelligence

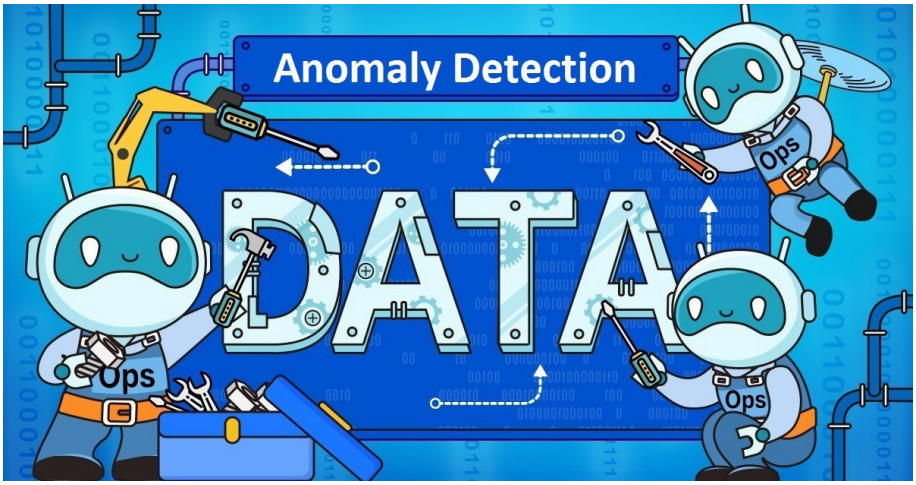
– **Anil Lamba**
Practice Lead - Cyber Security, EXL Service Inc., NJ, USA
https://orcid.org/0000-0002-7793-785X dranillamba@outlook.com

Article History
Paper Nomenclature:
Experiential Research Paper (ERP)
Paper Code: CYBNMV1N4SEP2019ERP1
Submission Online: 09-Sep-2019
Manuscript Acknowledged: 10-Sep-2019
Originality Check: 14-Sep-2019
Originality Test Ratio: 6%
Peer Reviewers Comment: 24-Sep-2019
Blind Reviewers Remarks: 26-Sep-2019
Author Revert: 27-Sep-2019
Camera-Ready-Copy: 28-Sep-2019
Editorial Board Citation: 30-Sep-2019
Published Online First: 2-Dec-2019

Today's industrial organizations are often tasked with objectives that are seemingly at odds with each other. They need to increase productivity while reducing machine failure or enhance product quality while speeding up time to market. Achieving these goals simultaneously can be incredibly challenging — if not impossible.

How does one work around this dilemma? The trick lies in cognitive anomaly detection and prediction, which is a process that leverages unsupervised learning (cognitive computing) and pattern recognition to quickly and accurately identify the anomalies hidden in your Industrial Internet of Things data. The use of machine learning algorithms minimizes the appearances of false alarms.

Keywords : Anomaly detection | Internet of things (IoT) | Cognitive Intelligence | Data Security



Introduction

Anomaly detection is a method used to ascertainstrange patterns that do not conform to expected behavior, known-as outliers. It can be used for many purposes from intrusion-detection to system-health monitoring, and from false-charges detection in credit-card transactions to Issues\fault detection in operational environments.

What Are Anomalies?

Anomalies can be broadly categorized as:

Point anomalies: A single-occurrence of data which is too far off from the rest. For-example: Identifyinga credit-card fake-transaction basis "amount-spent."

Contextual anomalies: This relates when occurrences are any context-specific and usually it is common in time-bound data. For Example: Paying 50$ on groceries every-day during the festival-season is usual, but may-be unusual otherwise.

Collective anomalies: This related to a pattern-of-data and its occurrences in a group helps in identifying anomalies. Business use case: Copying-data from a remotely-connected machine to your local machine unpredictably, an anomaly that would be flagged as a potential cyber-attack.

Anomaly Detection Techniques -

- **Simple Statistical Methods**
- **Machine Learning-Based Approaches**
- **Density-Based Anomaly Detection**
- **Clustering-Based Anomaly Detection**
- **Machine learning for anomaly detection (MLAD):-**
  - **Supervised MLAD**
  - **Un-Supervised MLAD**

Machine-learning techniques are now receiving significant attention among the anomaly-detection researchers.

Anomaly Detection in IoT

Internet of things in simple terms is the connection between everyday usage devices with internet. So, once the device is connected with internet access to the data of that device is obtained. Essential tools that require anomaly detection techniques are the ones used in industries and business organizations which has sensors. Data has been received continuously with every passing second through these sensors. Notably, in the maintenance of the systems, the sensors have to be monitored to predict the anomaly. These predictions have high economic value because in IoT as multiple things are interlinked.

While working using immense data, the following things are required?

- Feature selection
- Feature Transformation
- Pattern recognition

What are the features that are responsible for the anomaly?

What transformations should be made on these features to detect patterns?

What patterns signify anomaly?

To explain with a simple example, 'let's consider a sensor gives temperature values of special equipment in an industry. Change in sensor values is used to know if the equipment is stable or about to fail. Tracking is done on the statistical measures mean and standard deviation of the temperatures over some time. If there are changes, that is a shift in a mean or considerable fluctuations in standard deviation values, there is something wrong with the equipment, and immediate action is required.

This notification is sent as an alert. With the advances in technology, multiple machine learning and statistical techniques are used to identify and predict anomalies accurately. The significant advantage is once the complete system is automated one need not always keep track of equipment to know if everything is okay or not.

The need for Anomaly Detection?

To detect the unknown. Earlier, anomalies seldom occur. If the anomaly is not detected and rightful actions are not taken, soon the consequences may prove to be costly in situations like Network intrusion, change in log patterns, data leak, fraud transactions, Insider trading and many more. Just imagine the loss that could incur if any of the lists mentioned above occurs.

In a high velocity business, many things occur simultaneously, and different people/roles are charged with monitoring those activities. For example, at the level of the underlying infrastructure, a technical IT group carefully monitors the operation and performance of the network, the servers, the communication links, and so on. At the business application level, an entirely different group monitors factors such as web page load times, database response time, and user experience. At the business level, analysts watch shopping cart conversions by geography and by user profile, conversions per advertising campaign, or whatever KPIs are important to the business.

Anomalies in one area can affect other areas, but the association might never be made if the metrics are not analysed on a holistic level. This is what a large-scale anomaly detection system should do. Still, the question arises, why care about anomalies, especially if they simply seem to be just "blips" in the business that appear from time to time? Those blips might represent significant opportunities to save money (or prevent losing it) and to potentially create new business opportunities. Consider these real-life incidents:

- An e-commerce company sells gift cards and sees an unexpected increase in the number of cards purchased. While that sounds like a great thing, there is also a corresponding drop in the revenue expected for the gift cards. Something strange is going on, and it turns out to be a price glitch— something quite common for e-commerce companies. Without looking at these two metrics together, it is hard to understand that there is a business incident that could cost the company a lot of money if not caught and addressed quickly.
- A mobile game company notices a decrease in installations of one of its games, but may discover it several weeks in and not know why. With an anomaly detection

system, it is easy to determine that a problem with the cross-promotion mechanism serving up the wrong ads or ads with the wrong link is what led to the decline.

As businesses grow, more incidents go undetected unless an anomaly detection system is directed to make sense of the massive volume of metrics available to every online business. Of course, not every metric is directly tied to money— but most metrics are tied to revenue in some way. Say an online news site counts visitor to its website.

By itself, the visitor count doesn't lead to revenue, but the more visitors the news site gets, the more opportunity there is to generate revenue from ads on the pages, or to convert people who read the news site from free to paid subscribers. Most companies today tend to do manual detection of anomalous incidents.

Cognitive Analytics – In the contemporary world, the analytics mainly target at predictions with impeccably high accuracy trying to be as close to human understanding as possible, basically trying to mimic machines with the Cognitive intelligence that humans have. This analysis should be accurate, fast and with constant Learning. The output is expected in real time and also predict future events.

Layers of Cognitive Intelligence

Cognitive Computing helps in accelerating human intelligence by human Learning, thinking and adaptivity. By moving towards the machine capacity, it not only helps in augmenting the human potential; instead, it will increase the creativity of the individual and create new waves of innovations. The key areas of capability are –

- Sensory Perception - Machines are enabled in such a way that they can stimulate the senses of humans such as smell, touch, taste, and hearing. Therefore, they are developed in terms of machine simulation such as visual and auditory perception.

- Deduction, Reasoning, and Learning - In this case, the machines are simulated with human thinking for decision making. Therefore various technologies such as machine learning, Deep Learning, and neural networks are deployed as a system to intelligence to extract meaningful and useful information and apply the judgment.

- Data processing - In this larger dataset is accessed to facilitate the decision-making process and provides practical suggestions. Therefore,hyperscale-computing, knowledge representation, and natural language processing togetherly provide the required processing power to enable the system for engaging in real time.

Features of Cognitive Computing Solutions

The purpose of cognitive Computing is to create the frame for computing such that complex problems are solved easily without human intervention. Features are listed below –

- Adaptive - It is one of the first steps in developing the machine learning based cognitive system. The solution imitates to adapt the human ability with the Learning from the surroundings. It is dynamic for data gathering, understanding goals, and requirements.

- Interactive - The cognitive solution should dynamically interact bidirectionally in nature with each element in the system such as processes, devices, users and cloud services. The system can understand human input and provides the results using natural language processing and deep learning models.

- Iterative and Stateful - The system should be able to learn from previous iterations and ready to return the information which is specifically crucial at that time. The system must follow data quality and visualization methodologies so that it provides enough information and the data sources can operate the reliable and updated data.

- Contextual - The system should be able to understand, identify and even extract the contextual elements from the data such as meaning, syntax, time, location, task, goal and many more. The system removes the multiple sources of information like structured, sensor inputs, unstructured and semi-structured data.

Working of Cognitive Computing

Cognitive applications use deep Learning and neural network algorithms to control technological applications such as data mining, pattern recognition, and natural language processing.

The system gathers a variety of information and processes it with the previous report it already knows. After the completion of data analysis, it integrates with the adaptive page displays to visualize the content for specific audiences at specific situations.

How to find anomalies?

- **Supervised Anomaly detection – In this approach, historical data is used which says data points and class defining if each position is abnormal or not. It is similar to the classification problem. So, the class variable as the anomaly column is taken and applied with the models such as "Random Forest," "XGB," "SVM" or regression algorithms to train the data. This model is used to the new data point to know if it is an anomaly or not. One should be careful about the ratio of an anomaly to no-anomaly in the dataset. It 'shouldn't be too high – for example, more than 1:10 since it becomes a class imbalance problem.**
- **Unsupervised Anomaly detection –** Clustering techniques are applied in this as it is not known before if a data point is anomaly or not. So, clustering algorithms are used to detect anomalies.
- **K-Means clustering –** This algorithm is required to be given with the number of clusters to be formed as an initial input, based on that value the algorithm provides the same amount of clusters as an output. At present, by this process, it will consider all the data points and forms clusters. Restrict the distance of the boundary from the centroid from each cluster. The restriction can be 95 percentile or 99 percentile based on the requirement. The points outside this range after the clustering process is done are the anomalies.
- **DBSCAN –** This is also a clustering algorithm which is different from k-means. In DBSCAN minimum points and maximum distance as parameters are chosen. The algorithm initially starts with a random point and select the locations in the neighborhood and links them. Each of those will continue the same process. With this, the patterns are, and all the possible clusters are formed. The leftover points are considered as anomalies.

In the case of Supervised Algorithm, the data should be trained already, and the model can detect only that kind of anomaly which is learned by it before. Therefore, this algorithm is not feasible for detecting all types of anomalies. So, Unsupervised Algorithms can be used instead of supervised algorithms. This can identify any anomalies from

Dr Anil Lamba is a notable industry speaker, researcher, an innovator, and an influencer with proven success in spearheading Strategic Information Security Initiativesand Large-scale IT Infrastructure projects across industry verticals. He is Ph.D. Cyber Security, CISA ® and hold various other impressive industry credentials*. He has helped bring about a profound shift in cybersecurity defense. Throughout his career, Anil Lamba has parlayed his extensive background in security and a deep knowledge to help organizations build and implement strategic cybersecurity solutions.

Industry Credentials*: Ph.D. Cyber Security, M.B.A. – Strategic Project Management, CISA ®, CISSP, CDCP, CPD, CFE, PMP, Amazon Web Services (AWS) Certified Architect, AZURE Certified, Prince2, ITIL Expert, ISO 27001 Lead, Auditor, MCSE, 6σ Sigma Green Belt, CEH and CCNA.

He has performed various IT Security & Governance initiatives viz. Cloud Security Audits, Secure SDLC Audits, Penetration tests, Tabletop exercises, Data Warehouse (i.e. Enterprise Data Lake & Grid) Audits, Regulatory as well as Industry standard assessments and Pre-Mergers & Acquisitions assessment projects across industry verticals. Anil has gained his expert reputation by staying at the leading edge of security research and mentoring security teams, developers and audit experts across various industries.

Anil has strived to change the way organizations approach security, placing an emphasis on making informed decisions regarding products and services to best protect the organization, employees, and customers. He has spent much of his career meeting with CISOs and CIOs to advise and educate them on threats, required policies, processes, and expertise.

Anil has given more than 29 researches and 6 conference papers to information & cyber security world. His career has focused on high impact research in Cyber Security as well as its applied practice in real world and bringing it to everyone's use and education for free. His researches has educated the industry and the general public on the evolving threats to our interconnected world. He serves on the board member, reviewer& editor of several computer science related research journals.

Per Anil, "Effective cyber security requires science, engineering, business, policy and people skills. My goal has been and is to instill this culture in the discipline and provide leadership in all elements."

Professionally, He is Practice Lead – Cyber Security for a Consulting company named EXL. As a security professional, Anil has spent more than 15 years in cybersecurity operations leadership and influencing policy level decisions in multiple client organizations and helping clients understand cyber security challenges,. Some of these organizations included some of the largest pharmaceutical, Fortune 500 banking & financial and telecommunications companies in the U.S.

Anil leverages his skills and pervasive industry experience to help customers understand risks in their systems and develops programs to mitigate those risks. He has automated various cyber security, privacy and data security audit programs which has in-turn saved a lot of time for our clients as well as improved the practice. He has been a trusted advisor to the organization's lawful team to ensure that the required security clauses are built into supplier and customer contracts. He has also enabled engineering development teams of our clients with secure code practices and have performed reviews and quality assurance tests for functional and security verification.

He has always volunteered himself to act as a cybersecurity adviser for undergraduate college students conducting independent researches. He has presented research and other topics at many conferences over the years. His published researches and conference papers since 2014 has led to many thought provoking examples for augmenting better security. His Ph.D. thesis on Cyber Crime became a famous literature for inspiring cyber security and law students.

dranillamba@outlook.com

data, i.e. the anomalies that are never seen before.

Design Principles Of Anomaly Detection

Based on our experience, there are five main design considerations when building an automated anomaly detection system:

1. Timeliness – How rapidly does the organization need an answer to determine if something is an abnormality or not? In real-time, after a day, week, month or an year?

2. Scale – Will the data-sets be on a large-scale with millions of metrics or a relatively small-scale with hundreds of metrics?

3. Rate of change – Does the data tend to change-rapidly, or is being analyzed relatively-static?

4. Conciseness – Do we need an answer that tells the whole-picture, or would it be enough to detect irregularities at each metric-level by itself?

5. Definition of incidents – Are the expected incidents well-defined? Is anything known about them in advance in terms of what types of things can be anomalous in the data? Can incidents be categorized over time?

Importance of Real-Time Anomaly Detection

In specific use cases anomaly detection has to work in real-time. As soon as the anomaly is detected several measures can be taken to mitigate the loss. The techniques used in real-time anomaly detection have to evolve with time. Static methods based on an existing training data which was formed taking an actual sample may not serve the purpose of fundamental discoveries. 'That's because data changes are fast with immense volume and accordingly the models have to learn from data for rightful predictions.

The actions that are carried out for solving the problem of anomaly can be delayed but the detection of an anomaly in real time cannot be missed. This is because the data containing the anomaly can consist of information that can further lead to loss or gain in business.

For building the real-time anomaly detection platform following are the requirements

- Data-Collection
- Data-Aggregation
- Ability to Visualize
- Various-Alerts/ Received-Notifications
- Analytics

Annexure I

Submission Date	Submission Id	Word Count	Character Count
14-Sep-2019	1173716172 (turnitin)	2973	16678

ORIGINALITY REPORT

6%	3%	2%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	proceedings.mlr.press Internet Source	2%
2	Submitted to Middle East Technical University Student Paper	2%
3	pythonbytes.fm Internet Source	1%
4	Shijoe Jose, D. Malathi, Bharath Reddy, Dorathi Jayaseeli. "A Survey on Anomaly Based Host Intrusion Detection System", Journal of Physics: Conference Series, 2018 Publication	1%
5	Robert J. Freund. "Chapter 19 Cognitive Computing and Managing Complexity in Open Innovation Model", Springer Nature, 2017 Publication	<1%
6	Submitted to Bellevue University Student Paper	<1%
7	www.profsandhu.com Internet Source	<1%

Note: www.Cybernomics.in Uses a "Turnitin" https://www.turnitin.com which is an American commercial, Internet-based plagiarism detection service, a subsidiary of Advance and also offers plagiarism-detection service for newspaper editors and book and magazine publishers called iThenticate..

Reviewers Comment

Review 1: This article properly defines the use of technology in today's world.

Review 2: That's all focuses on importance of cyber security

Review 3 Data mining, anomaly detection is the identification of rare items, this article have proper information on anomaly system.

Editorial Excerpt

The article has 6% of plagiarism which is an accepted percentage for publication. The finding related to this particular manuscript seems to be noteworthy and defend a topic "**cognitive intelligence**" in the present industrial organization which often deal with odd objective, the manuscript is of great importance and use as referral for a learner. hence is has been earmarked and decided under **"Experiential Research Paper"** category.

Citation

Anil Lamba
"Anomaly Detection for Cognitive Intelligence"
Volume-1, Issue-4, Sep 2019. (www.cybernomics.in)

Frequency: Monthly, Published: 2019
Conflict of Interest: Author of a Paper had no conflict neither financially nor academically.

Scholastic Seed Inc.
www.scholasticseed.in