Journal of Knowledge Information Technology and Systems

ISSN 1975-7700

http://www.kkits.or.kr

An Unsupervised Real-time Anomaly Detection of Furnace System based on HTM

Ky Vannroath, Jong-Min Bae, Hyun-Syug Kang*

Department of Computer Science, Gyeongsang National University

ABSTRACT

The immediate recognition of abnormal temperatures in the furnace system is very important for timely disaster control. Conventional abnormal temperature detection techniques typically use pre-programmed thresholds to identify abnormal patterns from a stream of temperature data in a statistical manner. Therefore, it is not suitable for use in a heating furnace system in which abnormal patterns must be detected in real time. Hierarchical Temporal Memory (HTM) is one of the machine learning models with excellent ability to analyze patterns of stream data. This model is a new machine intelligence technique that uses unsupervised continuous learning using Cortical Learning Algorithm (CLA), which is a time-based learning algorithm, and Spare Distributed Representation (SDR), which stores temporal and spatial patterns. This HTM, which is distinguished from existing deep learning or artificial neural network model, is capable of real-time continuous prediction from input of stream data and can be used to recognize abnormal states in various fields. HTM can be used as a robust model to detect abnormalities in successive temperature data in a heating furnace system because the furnace system continues to generate temperature stream data with a certain pattern over time. Here we developed an HTM-based, unsupervised, real-time anomaly detection system for heating furnaces. Performance evaluation shows that the suggested system can perform excellent abnormal real-time detection.

KEYWORDS: Anomaly detection, Real-time, Unsupervised, Furnace systems, HTM, Machine learning

ARTICLE INFO: Received 27 September 2018, Revised 9 November 2018, Accepted 7 December 2018.

*Corresponding author is with the Department of Computer Science, Gyeongsang National University, 501 JinjuDaero Jinju, 52828, KOREA.
E-mail address: hskang@gnu.ac.kr

1. Introduction

The automated monitoring system to check the condition of furnace system on a real-time basis plays the key role to eliminate system downtime or to mitigate potential service failures. Real-time anomaly detection engine, particularly autonomous temperature detection, is one of the most important factor in a furnace system.

The various attempts at using statistical methods or pre-programmed threshold have been made to detect the anomalous point in temperature streaming data, which send the alert when temperature data jump out of lower or upper limit temperature range[1-4]. However, these kinds of algorithms require many efforts in defining threshold values, and do not provide enough lead time to operations team dealing with the impending fluctuations of temperature incident. So the system having the ability identifying the unusual, unexpected, or surprising patterns in real-time is required for early incident detection in furnace system.

There are a lot of previous research trying to construct temperature anomaly detection systems to solve these problems. Mete CELIK using DBSCAN was suggested to discover anomalies in monthly temperature data[5]. DBSCAN using a density-based clustering algorithm defines anomalies in the data series. It shows its performance by comparing with the statistical anomaly detection method with two user-defined parameters. Even if Mete CELIK results in a good response from a monthly temperature dataset, this system is not practical in a steelmaking industry environment that requires a solution near to real-time detection. Similarly, Zhao proposes ensemble methods to improve the performance of density-based and rank-based algorithms using several datasets[6].

The neocortex is the complex memory system with hierarchical structures divided into regions[7]. It constantly predicts the future with the data it received and also learns new knowledge from past experiences. We call this neocortex theory as ‘memory prediction framework’[8].

HTM is a new machine intelligence model based on the neocortex theory[9]. It uses unsupervised continuous learning using Cortical Learning Algorithms (CLAs)[10-12], which are time-based learning algorithms, and Spare Distributed Representation (SDR)[13], which stores temporal and spatial patterns. It is distinguished from existing deep learning[14] or artificial neural network model[15].

In this paper, we use HTM machine learning algorithm for anomaly detection from temperature stream of furnace system.

This paper is structured as follows. Section 1 introduces a few challenging questions along with previous research to tackle the problems. Section 2 analyses some anomaly detection problems in the furnace system. Section 3 presents the system architecture and implementation based on HTM engine to perform the anomaly detection including experiment results. Section 4 evaluates our detection system, and section 5 describes conclusions.

2. Problems Analysis of Furnace System

2.1 Furnace system

Before diving into the logic behind our system we describe what is a furnace system. An industrial furnace system contains several sensors attached to each heater such as temperature sensor, vibration motor sensor, noise volume sensor and so on to watch over the whole operation of the system. The job of furnace system is to heat steel starting from pre-heater stage till uniform heater stage before performing metal bending procedure as <Figure 1>.

In the heating furnace system in the <Figure 1>, the steel plate of the yard is passed through the conveyor belt through several heaters in order to be machined, and is processed into a stable steel plate through the final step. In this case, if the temperature or the pressure intensity can not be adjusted properly at each step, a problem occurs in the uniformity of the steel sheet. Therefore, it is very important in the heating furnace system to determine the normal state by measuring the temperature of the heat applied by the heater in each stage in real time and to make a decision such as stopping the work. In order to efficiently process this problem, the heating furnace system arranges various related sensors in each stage, measures the sensing values, and then uses these values to control the heating furnace by sending them to the remote central computer.

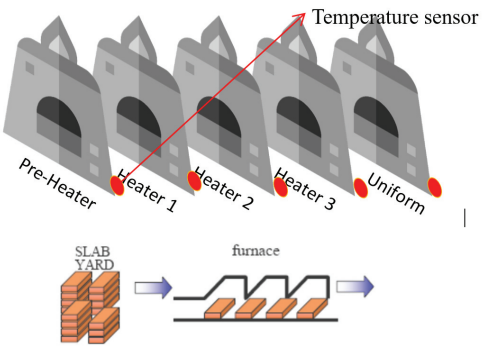

Figure 1. Furnace system

In this paper, we will cover only detecting anomaly behavior of temperature sensor inside furnace system as it is the most crucial factor among all others. The temperature inside the furnace system is controlled by input amount of gas and oxygen, and by the openness of valve of gas and oxygen.

2.2 Problems analysis

Real-time anomaly detection application promulgates unique constraints and challenges for machine learning models. It requires an astounding amount of continuous sequence data to be analyzed in real-time. With abundant amount of data to be properly learned, it can make prediction from the data set that would contain normal or diverse patterns. It is contrasted with anomaly detection using batch processing that the full dataset is available. Practical real-time anomaly detection system of furnace system imposes another constraint. Specifically, sensors inside furnace system keep sending data, in which those sensor streams are enormous in number at

high velocity and few seconds interval, leaving little opportunity for human intervention. Manual parameter tweaking and data labeling seem to be unviable for furnace system data. Therefore, operating in an unsupervised and automated fashion is necessary for anomaly detection of furnace system. HTM algorithms shows excellent performance to deal with detecting anomalies in streaming data. It meets our requirements to solve the above problems.

3. Anomaly Detection System for Furnace System

3.1 System architecture

<Figure 2> describes the architecture of unsupervised real-time anomaly detection system for furnace system. We construct our system using data server and HTM engine of NuPIC framework to detect anomaly of furnace system and generate anomaly score. NuPIC was developed by Numenta and derived from HTM machine intelligence algorithm[9]. In the <Figure 2>, preprocessing interface obtains raw metrics generated from each temperature sensor attached to heater directly and performs data format preprocessing. Detection system makes use of HTM engine to detect patterns, and models each heater automatically. Each individual metrics has a corresponding model which is used to generate anomaly score. An automated model creation also reduces or removes the pre-defined parameters for identifying what is normal and abnormal. It also eliminates the need to create new model manually if new heater is added.

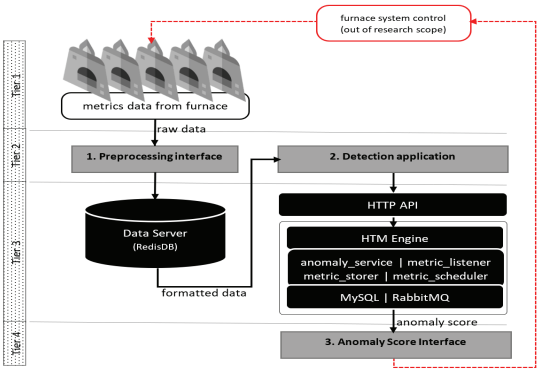

Figure 2. Architecture of anomaly detection system

Data server itself is constructed by combining html pages, JavaScript files, and many nodes of RedisDB to easily obtain temporal streaming data. Raw metrics generated from each heater contain various information such as airflow input, gas flow, and other chemical components that have the influence on the fluctuations of temperature value. That raw temperature value is scripted every minute and pushed into HTM engine which contain several services including anomaly service, metric listener service, metric storer service, and metric scheduler service. Then HTM engine processes those real-time raw metrics from furnace system, and produces an anomaly score for each metric.

3.2 Experimental results

We configured our user-define script to fetch data from furnace system in 1-minute time

interval. <Figure 3> shows the implementation result of our engine to generate abnormal scores in real time. We plotted this data as shown in <Figure 4>, <Figure 5>, and <Figure 6> for more visualization and intuitive control.

<Figure 4> shows temperature fluctuations for the first day resulting in high anomaly score around 0.5. It means that all data are treated as unexpected or high anomalous. That's what we expected because the engine has seen the temperature data for the first time and has no data previously learned. <Figure 5> shows the temperature fluctuations of the second day since the temperature data has been fetched into HTM engine. We can find that the system considered the temperature as normal, lowering its anomaly score to nearly zero in all the heaters, even though there are temporal fluctuations of temperature data.

Uid	Row id	Time stamp	Metric value	Raw anomaly score	Anomaly score	Is play value	Multi step best predictions
5441183c96a0497f86b35854 9f1f76ad	6022	2017-08-22 00:00:04	865.210661 148647	0	0.0445654 63	1000	null
dab122a1c53d49da9f593e80d c05b9af	6046	2017-08-22 00:00:07	820.623636 850596	0	0.0445654 63	1000	null
238a3aaa98f148758681b7c1e 7524d45	6001	2017-08-22 00:00:15	820.351505 501863	0.45	0.8159398 75	1000	null
5441183c96a0497f86b35854 9f1f76ad	6023	2017-08-22 00:00:17	869.909771 913561	0	0.0445654 63	1000	null
dab122a1c53d49da9f593e80d c05b9af	6047	2017-08-22 00:00:20	822.956055 236671	0	0.0445654 63	1000	null
238a3aaa98f148758681b7c1e 7524d45	6002	2017-08-22 00:00:28	827.181413 124897	0.9	0.9999999 99456085	10000 00001	null
5441183c96a0497f86b35854 9f1f76ad	60247	2017-08-22 00:00:30	865.781709 515306	0	0.0445654 63	1000	null
dab122a1c53d49da9f593e80d c05b9af	6048	2017-08-22 00:00:33	823.106836 594012	0	0.0445654 63	1000	null
238a3aaa98f148758681b7c1e 7524d45	6003	2017-08-22 00:00:41	826.419427 929976	0.8	0.999	1000	null
5441183c96a0497f86b35854 9f1f76ad	6025	2017-08-22 00:00:43	866.763260 268249	0	0.0445654 63	1000	null
dab122a1c53d49da9f593e80d c05b9af	6049	2017-08-22 00:00:46	821.493516 065957	0	0.0445654 63	1000	null

Figure 3. Original anomaly score

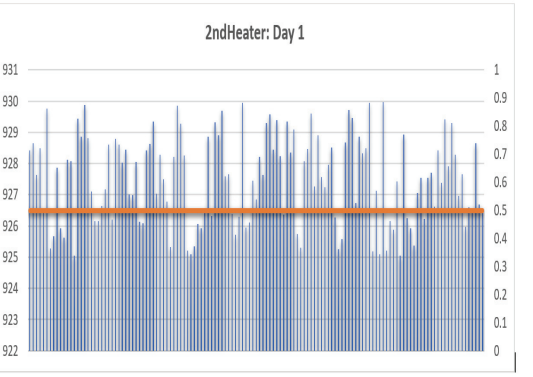

Figure 4. Day 1 of operation

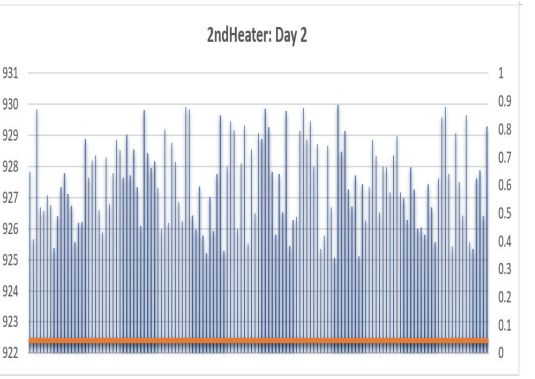

Figure 5. Day 2 of operation

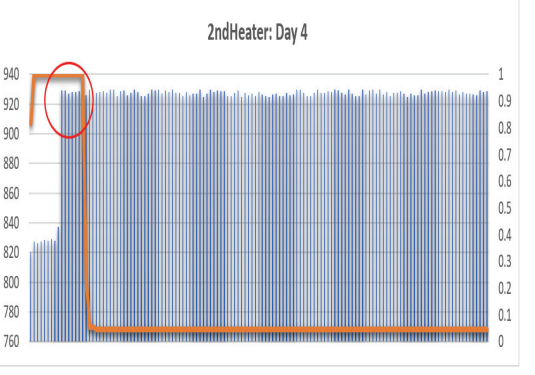

Figure 6. Day 4 of operation

This is the result of learning from the previous trend of temperature fluctuation. The system remains stable for four days. There was a technical error in the openness of valve to provide gas into furnace system when a heater starts. <Figure 6> shows a abnormal patterns of temperature fluctuations and its anomaly score. There was also human intervention after seeing high anomaly score. Again, this is what we expected as there was a technical mistake in furnace system. Refer [16] for more information.

4. Evaluation

We here evaluate our anomaly detection system for temperature of the furnace system. As mentioned, A stream of data is fed to our system continuously from temperature sensors of furnace system in a 1-minute interval. We found that our system has the ability to adapt very quickly to the random fluctuations of various heaters.

About 10,000 temperature records are generated per heater per day. Our system was able to judge normal and abnormal by using only input records for one day at the learning stage. Furnace system works normally on a daily operation. It generates high anomaly score and alerts when there was a technical error in the openness of valve providing gas into furnace system as <Figure 6>. It enables user to take action immediately toward specific abnormal heater found in the whole furnace system.

Our system provides a decent and compelling anomaly detection score. In our testing setup environment, all related servers are connected in a 100Mbps local area network having no bottleneck in data transmission over the network protocol. Our system can successfully generate anomaly score within 5 minutes for about 40,000 records, which is one-day records of a furnace system.

Qualitative investigations are still necessary for our system in order to compare system performance to others since our testing does not include any security related testing.

5. Conclusions

Sensors inside furnace system continuously send a stream of data. Immediate detection of anomaly from data generated by the temperature sensors in furnace system is very important for the timely disaster prevention. In this paper, we present an unsupervised real-time anomaly detection system for the temperature in the furnace system. It is implemented based on Hierarchical Temporal Memory (HTM) engine. HTM is an unsupervised machine learning model, which has special ability to analyze the pattern of data stream.

Though performance evaluation shows that the proposed system in this paper provides excellent anomaly detection, more extensive investigations are still necessary to apply this system to other areas. It is also necessary to find areas in which our solution is applicable.

References

[1] K. Korb, and A. Nicholson, *Bayesian artificial intelligence*, Chapman & Hall CRC Computer Science and Data Analysis, CRC

Press, 2011.

[2] A. Soule, K. Salamatian, and N. Taft, *Combining filtering and statistical methods for anomaly detection*, in Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement, p. 31, 2005.

[3] V. Chandola, V. Mithal, and V. Kumar, *Comparative evaluation of anomaly detection techniques for sequence data*, in Proceedings of the 2008 8th IEEE International Conference on Data Mining, pp. 743-748, 2008.

[4] H. Akouemo, and R. Povinelli, *Probabilistic anomaly detection in natural gas time series data*, International Journal of Forecast, pp. 948-956, 2015.

[5] M. Çelik, F. Dadaşer, and A. Şakir, *Anomaly detection in temperature data using DBSCAN algorithm*, International Symposium on Innovations in Intelligent Systems and Applications, pp. 91-95, 2011.

[6] Z. Zhao, K. Mehrotra, and C. Mohan, *Ensemble algorithms for unsupervised anomaly detection*, Lecture Notes in Computer Science, Springer, Vol. 9101, pp. 514-525. 2015.

[7] V. Mountcastle, *The columnar organization of the neocortex*, Brain, Oxford University Press, pp. 701-722, 1997.

[8] J. Hawkins, and S. Blakeslee, *On intelligence*, Times Books Henry Holt and Company, 2012.

[9] J. Hawkins, *Hierarchical temporal memory (HTM)*, in Numenta Whitepaper, http://www.numenta.org, 2011.

[10] S. Purdy *Encoding data for HTM systems*, in Numenta Whitepaper, http://www.numenta.org, 2015.

[11] Y. Cui, S. Ahmad, and J. Hawkins, *The HTM spatial pooler - A neocortical algorithm for online sparse distributed coding*, Frontiers in Computational Neuroscience, Vol. 110, pp. 1-15, 2017.

[12] Y. Cui, S. Ahmad, and J. Hawkins, *Continuous online sequence learning with an unsupervised neural network model*, Neural Computation, Vol. 28, pp. 2474-2504, 2016.

[13] S. Ahmad, and J. Hawkins. *Properties of sparse distributed representations and their application to hierarchical temporal memory*, in Numenta Whitepaper, http://www.numenta.org, 2015.

[14] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature 521, pp. 436-444, 2015.

[15] R. Douglas, and K. Martin, *Neuronal circuits of the neocortex*, Annual Review of Neuroscience 27, pp. 419-451, 2004.

[16] K. Vannroath, *An unsupervised real-time anomaly detection application of furnace system based on HTM,* Master Thesis, Department of Computer Science at Gyeongsang National University, Feb. 2018.

가열로 시스템의 HTM 엔진 기반 무감독 비정상 실시간 감지

끼완노앗[1], 배종민[2], 강현석[2]

[1] *경상대학교 컴퓨터과학과 석사*

[2] *경상대학교 컴퓨터과학과 교수*

요 약

가열로 시스템에서 비정상 온도의 즉각적인 인지는 시의적절한 재난 방제를 위해 매우 중요하다. 종래의 비정상 온도 탐지 기법들은 대개 통계학적인 방법으

로 일련의 온도 데이터들로부터 비정상적인 패턴을 식별하기 위해 사전에 프로그램된 임계치를 사용한다. 따라서 실시간으로 비정상적인 패턴을 알아내야 하는 가열로 시스템에 사용하기에는 적합하지 않다. 계층형 시간적 메모리(HTM: Hierarchical Temporal Memory)는 스트림 데이터의 패턴을 분석하는데 우수한 능력을 갖는 기계 학습 모델들 중의 하나이다. 이 모델은 시간 기반의 학습 알고리즘인 신피질 학습 알고리즘(CLA: Cortical Learning Algorithm)과 시공간적 패턴을 저장하는 희소 분산 표현(SDR: Spare Distributed Representation)을 이용해 무감독 연속 학습(Unsupervised continuous learning)을 하는 새로운 차세대 기계 지능 기법이다. 기존의 심층 학습(Deep learning)이나 인공 신경망 모델(Artificial Neural Networks Model)과는 구별되고 있는 이 HTM은 스트림 데이터의 입력으로부터 실시간 연속 예측이 가능하며 여러 분야의 비정상 상태를 인지하는데 사용이 가능하다. 이 HTM을 가열로 시스템에서 연속해서 발생되는 온도 데이터들에서 비정상 탐지를 할 수 있는 강력한 모델로 사용할 수 있다. 가열로 시스템으로 부터의 온도 데이터들은 시간에 따라 패턴을 갖는 스트림 데이터를 계속해서 발생시키기 때문이다. 우리는 여기에서 가열로 시스템을 위한 HTM 기반의 무감독 실시간 비정상 감지 시스템을 개발했다. 성능 평가를 통해 제안된 시스템이 우수한 비정상 실시간 감지를 할 수 있음을 보였다.

Ky Vannroath received the bachelor's degree in the Information Technology from University of Cambodia in 2012. He received the M.S. degree in the Department of Computer Science from Gyeongsang National University in 2018. His current research interests include database, artificial intelligence, and Health Care System.
E-mail address: vannroath.ky@gmail.com

Jong-Min Bae received the bachelor's degree in the Department of Mathematics Education from the Seoul National University in 1980. He received the M.S. degree and the Ph.D. degree in the Department of Computer Science and Statistics from Seoul National University in 1983 and 1995 respectively. He has been a professor in the Department of Computer Science at Gyeongsang National University since 1984. His current research interests include programming languages, compilers, and computer education.
E-mail address: jmbae@gnu.ac.kr

Hyun-Syug Kang received the bachelor's degree in the Department of Computer Science from the Dongguk University in 1981. He received the M.S. degree and the Ph.D. degree in the Department of Computer Science and Statistics from Seoul National University in 1983 and 1989 respectively. From 1981 to 1985, he was a researcher at ETRI. He was a professor in Chonbuk National University from 1985 to 1993. He has been a professor in the Department of Computer Science at Gyeongsang National University and researcherof the Research Institute of Natural Science since 1994. His current research interests include multimedia, embedded database, and intelligent system.
E-mail address: hskang@gnu.ac.kr