

## INDUSTRY WORKSHOP

### DIGITAL TRANSFORMATION – PRACTICAL VIEWS

In Conjunction with  
Symposium on Computer Science & Engineering 2021 (SCSE 2021)



#### **Digital Transformation: Trends, Challenges/Opportunities and Practices**

**Mr. Phan Thanh Son**

*Director, Strategic Business, FPT Corporation*

*Chief Business Development Officer, FPT Information Systems*



#### **Digital Transformation in Business Administration – Practical Applications**

**Dr. Quach Ngoc Long**

*Founder and CEO of WORKIT*

*Founder and CEO of EPACIFIC*



#### **Towards a Modern Supply Chain Management Platform: Seamless interactivity across the entire supply chain with standardized master data**

**Mr. Lai Duc Nhuan**

*Chairman, CEO of ATALINK*

*Chairman of LARION Computing*

*Co-founder of MEPERIA*



#### **Managing Enterprise's Content in Digital Transformation**

**Mr. Ho Thai Truong Giang**

*Co-founder of EyeQ Tech*

*Co-founder of INTELIN TECHNOLOGY CORP*

*Used to be Manager of IT and Operations Departments of ACB*

## KEYNOTES

### MCU PROCESS VITAL SIGN ACQUISITION USING CONTACTLESS DOPPLER RADAR

**Prof. Koichiro Ishibashi**

The University of Electro-Communications, Tokyo, Japan

#### Abstract

Recently a lot of researchers have investigated Vital Sign Acquisition by Medical Doppler Radar. The vital sign can be acquired contactless manner so that we can easily acquire the vital signs rather than contactless system such as ECC or PPG. This paper demonstrates the signal processing of the vital sign signals on time domain and peak detection. The process is so light that low performance MCU can process them to acquire the vital signs at fast time.

#### Biography of Prof. Koichiro Ishibashi



Dr. Koichiro Ishibashi has been a professor of The University of Electro-Communications (UEC), Tokyo, Japan since 2011. He has been the Director of UEC ASEAN Research and Education Center (UAREC) since 2017. He has been serving a guest professor at Ho Chi Minh City University of Technology and Ho Chi Minh City University of Science since 2012.

After receiving PH. D degree from Tokyo Institute of Technology in 1985, he worked at Central Research Laboratory, Hitachi Ltd., at Semiconductor Technology Academic Research Center (STARC) and at Renesas Electronics, where he had investigated low power technologies for high density SRAMs and MCUs. Since 2005, he has been a Fellow of IEEE for the technical contributions to developments of low-power SRAMs and MCUs.

He has presented more than 200 academic papers at international conferences and journals including more than 20 invited presentations. He was awarded R&D 100 for the development of SH4 Series Microprocessor in 1999.

His current interests are IoT technologies including Ultra low power LSI design technology, Technologies for energy harvesting sensor networks and applications, and Biomedical electronics using contactless sensors and data processing by AI.

## KEYNOTES

### AI-EMPOWERED EDGE COMPUTING WITH EFFICIENT AND SECURE DISTRIBUTED LEARNING

**Dr. Diep Nguyen**

University of Technology Sydney, NSW, Australia

#### Abstract

By “moving” computing resources closer to users and infrastructure, edge computing has become an enabler of new emerging services and smart infrastructures like intelligent transportations, industry automation, real-time gaming, virtual/augmented reality. This talk explores how recent advances in distributed learning (DL) can empower edge computing to turn existing infrastructure into smart one with augmented intelligence. Specifically, we first discuss hurdles like low-latency requirement, straggling effect, and users’ security-related risks of DL in wireless edge networks. We then present potential solutions by leveraging coded computing, network economics to address the straggling problem, protect users' privacy, and incentivize users to efficiently contribute to the learning process. We will also present an in-network computation framework to enable large-scale distributed learning systems in wireless edge networks.

#### Biography of Dr. DIEP NGUYEN



Diep N. Nguyen (Senior Member, IEEE) received the M.E. degree in electrical and computer engineering from the University of California at San Diego (UCSD) and the Ph.D. degree in electrical and computer engineering from The University of Arizona (UA). He is currently a Faculty Member of the Faculty of Engineering and Information Technology, University of Technology Sydney (UTS).

Before joining the UTS, he was a DECRA Research Fellow at Macquarie University, and a member of Technical Staff at Broadcom, CA, USA, and ARCON Corporation, Boston, consulting the Federal Administration of Aviation on turning detection of UAVs and aircraft, U.S. Air Force Research Lab on antijamming.

His current research interests include computer networking, wireless communications, and machine learning applications, with an emphasis on systems performance and security/privacy. He has received several awards from LG Electronics, UCSD, UA, U.S. National Science Foundation, and Australian Research Council. He is an Editor, Associate Editor of the IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE ACCESS, IEEE SENSORS JOURNAL, and IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY (OJ-COMS).

# AN EFFICIENT APPROACH TO MEASURE THE DIFFICULTY DEGREE OF PRACTICAL PROGRAMMING EXERCISES BASED ON STUDENT PERFORMANCES

**Huy Tran<sup>(1), (2)</sup>, Tien Vu Van<sup>(1), (2)</sup>, Hoang Nguyen Viet<sup>(1), (2)</sup>, Duy Tran Ngoc Bao<sup>(1), (2)</sup>,  
Thinh Tien Nguyen<sup>(1), (2)</sup>, and Thanh Van Le<sup>(1), (2)</sup>**

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT),

Ho Chi Minh City, Vietnam

<sup>(2)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM),

Ho Chi Minh City, Vietnam

Email: {huy.tran14; tien.vuvan3499; hoang.nguyen.k2017; duytnb; ntthinh;  
ltvan}@hcmut.edu.vn

## ABSTRACT

This study examines the generality of easy to hard practice questions in programming subjects. One of the most important contributions is to propose four new formulas for determining the difficulty degree of questions. These formulas aim to describe different aspects of difficulty degree from the learner's perspective instead of the instructor's subjective opinions. Then, we used clustering technique to group the questions into three easy, medium and difficult degrees. The results will be the baseline to consider the generality of the exercise sets according to each topic. The proposed solution is then tested on the data set that includes the results of the two subjects: Programming Fundamentals, Data Structures and Algorithms from Ho Chi Minh City University of Technology. The most important result is to suggest the instructors complete various degrees according to each topic for better evaluating student's performance.

**Keywords:** *e-learning, difficulty degree, automatic question classification, student's perception, effective coverage exercise.*

# INFECTIOUS DISEASE SCREENING SYSTEM USING MEDICAL RADAR AND DATA QUALITY ASSESSMENT BY EFFICIENT NEURAL NETWORK HARDWARE

**Duc-Tho Mai<sup>(1)</sup>, Koki Kumagai<sup>(1)</sup>, Guanghao Sun<sup>(1)</sup>, *Senior Member, IEEE* and Koichiro Ishibashi<sup>(1)</sup>, *Fellow, IEEE***

<sup>(1)</sup> Graduate School of Informatics and Engineering,  
The University of Electro-Communications, Tokyo 182-8585, Japan  
Email: thomaiduc@uec.ac.jp

## ABSTRACT

In recent years, various state-of-the-art machine learning algorithms enable improving medical diagnoses and developing efficient tools to process complex biomedical signals. Neural network models demonstrate tremendous efficiency in numerous applications requiring biomedical signal processing due to their pattern-matching capabilities and human brain-like features. Vital signs mixed with noise acquired via the non-contact medical radar method are enormous challenges for researchers who want to analyze and process these signals. This paper presents a hardware architecture of  $3 \times 6 \times 5 \times 4$  Neural Networks to take advantage of the hardware resource, enable computation acceleration, and ensure the reliability of classified signals for infectious disease screening systems. The proposed optimized architecture uses a 16-bit floating-point data representation, is synthesized on Xilinx Artix-7 XC7A200T board at a maximum operating frequency of 119.02 MHz, and low hardware resource utilization. The experimental results show that the design can reach about 93% classification accuracy in test results.

**Keywords:** *Infectious Disease Screening, Neural Networks, Hardware Implementation, Vital Signs, Non-contact Medical Radar.*

# GENERATING WEBSITE CODES FROM IMAGES

**Lai Nguyen Ha My<sup>(1),(2)</sup>, Huynh Ngoc Thien<sup>(1),(2)</sup>, and Nguyen Duc Dung<sup>(1),(2)</sup>**

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT),

Ho Chi Minh City, Vietnam

<sup>(2)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM),

Ho Chi Minh City, Vietnam

Email: nddung@hcmut.edu.vn

## ABSTRACT

Brainstorm website layout ideas usually start with everyone giving out their mockups, and all the team members will discuss to finalize the layout of the user interface. Once a vision of that mockup is accepted, it is given to the designer to sketch it digitally on computer software (i.e., Photoshop, Figma, Sketch). When the designer completes, the developer based on the final design to code the UI/UX of the website. As we can see, the process requires three stages, which can be time-consuming. Therefore, if anyone has an idea for the professional website layout, they can visualize it by drawing on sketches. However, it can be impossible for them to make a usable website without designers and website developers. Due to that reason, our primary goal in this paper is to help individuals transform their hand-drawn sketch images into a website that can be deployed. To achieve that goal, we present two approaches: classical computer vision techniques and the other using a deep learning model to detect the sketch and execute the conversion. Furthermore, our evaluation shows that deep learning is the most promising direction. Still, classical techniques also improve the model's input data by applying it in the pre-processing image.

**Keywords:** *Computer vision, Hand-drawn sketch, UI/UX, Deep learning.*

# AN EFFICIENT HARDWARE IMPLEMENTATION OF CONVOLUTIONAL NEURAL NETWORK IN DETECT BREAST CANCER HISTOPATHOLOGY IMAGE

**Vo Tan Phat** <sup>(1), (2)</sup>, and **Hoang Trang** <sup>(1), (2)</sup>

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT),

Ho Chi Minh City, Vietnam

<sup>(2)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM),

Ho Chi Minh City, Vietnam

Email: hoangtrang@hcmut.edu.vn

## ABSTRACT

This paper presents our work on evaluating the effectiveness of a novel deep convolutional neural network architecture (CNN) for classifying breast histology images for cancer risk factors as negative or positive. Also, the hardware structure of the proposed model was successfully synthesized and verified. The results indicate that a CNN trained on a small dataset achieved an overall AUC (Area under the receiver operating characteristic curve - ROC Curve) value of 0.922 on a set of 55505 test images. In addition, the time it takes to classify each image is within 3.8 milliseconds instead of a task that even trained pathologists take hours to complete.

**Keywords:** *Convolutional Neural Network, FPGA, Breast Cancer, Hardware Implementation.*

# A MULTITASK DATA-DRIVEN MODEL FOR BATTERY REMAINING USEFUL LIFE PREDICTION

**Thien Pham<sup>(1), (4)</sup>, Loi Truong<sup>(1), (4)</sup>, Mao Nguyen<sup>(2), (4)</sup>, Anh Bui<sup>(3), (4)</sup>,  
Dang Nguyen<sup>(1), (4)</sup>, Dat Nguyen<sup>(1), (4)</sup>, Akhil Garg<sup>(5)</sup>, Liang Gao<sup>(5)</sup>, and Tho Quan<sup>(1), (4)</sup>**

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam

<sup>(2)</sup> Ho Chi Minh City University of Science (HCMUS), Ho Chi Minh City, Vietnam

<sup>(3)</sup> Ho Chi Minh City University of Technology and Education (HCMUTE),  
Ho Chi Minh City, Vietnam

<sup>(4)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM), Ho Chi Minh City, Vietnam

<sup>(5)</sup> School of Mechanical Science and Engineering Huazhong University of Science and  
Technology Wuhan, China

Email: pcthien.sdh20@hcmut.edu.vn

## ABSTRACT

Lithium-ion batteries (LIBs) have recently been used widely in moving devices. Understand status of the batteries can help to predict the failure and improve the effectiveness of using them. There are some lithium-ion information that define the battery health over time. These are state-of-charge (SOC), state-of-health (SOH), and remaining-useful-life (RUL). Normally, a LIB is working under charging and discharging cycles continuously. In this paper, we will focus on the data dependency of different time-slots in a cycle and in a sequence of cycles to retrieve RUL. We leverage multi-channel inputs such as temperature, voltage, current and the nature of peaks cross the cycles to improve our prediction. Comparing to existing methods, the experiments show that we can improve from 0.040 to 0.033 (reduce 17.5%) in RMSE loss, which is significant.

**Keywords:** *Lithium-ion battery, time-series, Conv2DLSTM, RUL.*



# NEIGHBORHOOD SEARCH FOR SOLVING PERSONAL SCHEDULING PROBLEM IN AVAILABLE TIME WINDOWS WITH SPLIT-MIN AND DEADLINE CONSTRAINTS

**Trang Hong Son** <sup>(1), (2), (3)</sup>, **Tran Van Lang** <sup>(4)</sup>, and **Nguyen Huynh-Tuong** <sup>(1), (2)</sup>

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam

<sup>(2)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM), Ho Chi Minh City, Vietnam

<sup>(3)</sup> Hoa Sen University, Vietnam

<sup>(4)</sup> HCMC University of Foreign Languages - Information Technology, Vietnam

Email: {8140009@hcmut.edu.vn, son.tranghong@hoasen.edu.vn}, langtv@huflit.edu.vn, htnguyen@hcmut.edu.vn

## ABSTRACT

The scheduling of individual jobs with certain constraints so that efficiency is a matter of concern. Jobs have deadlines to complete, can be broken down but not too small, and will be scheduled into some available time windows. The goal of the problem is to find a solution so that all jobs are completed as soon as possible. This problem is proved to be a strongly NP-hard problem. The implementation of the proposed MILP model using a CPLEX solver was also conducted to find optimal solution for the small-size dataset. For large-size dataset, heuristic algorithms are recommended such as First Come First Served (FCFS), Earliest Deadline (EDL), and neighborhood search including Stochastic Hill Climbing (SHC), Random Restart Hill Climbing (RRHC), Simulated Annealing (SA) to find a good solution in an acceptable time. Experimental results will present in detail the performance among the groups of exact, heuristic, and neighborhood search methods.

**Keywords:** *splitting-job, available time-window, deadline constraint, FCFS rule, EDL rule, neighborhood search, hill climbing algorithm, simulated annealing algorithm.*

# AN FPGA-BASED CONVOLUTION IP CORE FOR DEEP NEURAL NETWORKS ACCELERATION

**Xuan-Quang Nguyen** <sup>(1), (2)</sup>, and **Cuong Pham-Quoc** <sup>(1), (2)</sup>

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT),  
Ho Chi Minh City, Vietnam

<sup>(2)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM),  
Ho Chi Minh City, Vietnam

Email: cuongpham@hcmut.edu.vn

## ABSTRACT

The development of machine learning has made a revolution in various applications such as object detection, image/video recognition, and semantic segmentation. Neural networks, a class of machine learning, play a crucial role in this process because of their remarkable improvement over traditional algorithms. However, neural networks are now going deeper and cost a significant amount of computation operations. Therefore, they usually work ineffectively in edge devices that have limited resources and low performance. In this paper, we research a solution to accelerate the neural network inference phase using FPGA-based platforms. We analyze neural network models, their mathematical operations, and the inference phase in various platforms. We also profile the characteristics that affect the performance of neural network inference. Based on the analysis, we propose an architecture to accelerate the convolution operation used in most neural networks and takes up most of the computations in networks in terms of parallelism, data reuse, and memory management. We conduct different experiments to validate the FPGA-based convolution core architecture as well as to compare performance. Experimental results show that the core is platform-independent. The core outperforms a quad-core ARM processor functioning at 1.2GHz and a 6-core Intel CPU with speed-ups of up to 15.69× and 2.78×, respectively.

**Keywords:** *Convolution neural network; Reconfigurable hardware; FPGA design.*

# PGN-LM MODEL AND FORCING-SEQ2SEQ MODEL: MULTIPLE AUTOMATIC MODELS OF TITLE GENERATION FOR NATURAL TEXT USING DEEP LEARNING

Nhan To Thuan<sup>(1), (2)</sup>, Nguyen Thi Hiep<sup>(1), (2)</sup>, and Tho Quan Thanh<sup>(1), (2)</sup>

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT),

Ho Chi Minh City, Vietnam

<sup>(2)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM),

Ho Chi Minh City, Vietnam

Email: {1970022; 1870365; qttho}@hcmut.edu.vn

## ABSTRACT

In the current era, the amount of information from the Internet in general and the electronic press in particular has increased rapidly and has extremely useful information value in all aspects of life, many popular users have posted several high-quality writings as casual blogs, notes or reviews. Some of them are even selected by editors to be published in professional venues. However, the original posts often come without titles, which are needed to be manually added by the editing teams. This task would be done automatically, with the recent advancement of AI techniques, especially deep learning. Even though auto-title can be considered as a specific case of text summarization, this job poses some major different requirements. Basically, a title is generally short but it needs to capture major content while still maintaining the writing style of the original document. To fulfill those constraints, we introduce PGN- LM Model, an architecture evolved from the Pointer Generator Network, with the ability to solve Out-of-Vocabulary problems that traditional Seq2Seq models cannot handle, and at the same time combined with language modeling techniques. In addition, we also introduce a model called Forcing-Seq2Seq Model, an enhanced Seq2Seq architecture, in which the classical TF-IDF scores are incorporated with Named Entity Recognition method to identify the major keywords of the original texts. To enforce the appearance of those keywords in the generated titles, the specific Teacher Forcing mechanism combined with the language model technique are employed. We have tested our approaches with real datasets and obtained promising initial results, on both metrics of machine and human perspectives.

**Keywords:** *Sequence to sequence, Attention Mechanism, Named Entity Recognition, TF-IDF, Language Model, Teacher Forcing, Pointer Generator Network.*

# ENHANCING SECURITY AND ROBUSTNESS FOR SDN-ENABLED CLOUD NETWORKS

**Long Tan Le<sup>(1),(2)</sup>, and Tran Ngoc Thinh<sup>(1),(2)</sup>**

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT),

Ho Chi Minh City, Vietnam

<sup>(2)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM),

Ho Chi Minh City, Vietnam

Email: {1870385; tnthinh}@hcmut.edu.vn

## ABSTRACT

Software-Defined Networking is an emerging network architecture which promises to solve the limitations associated with current cloud computing systems based on traditional network. The main idea behind SDN is to separate control plane from networking devices, thereby providing a centralized control layer integrable to cloud-based infrastructure. The integration of SDN and Cloud Computing brings immense benefits to network deployment and management, however, this model still faces many critical challenges with regards to availability, scalability and security. In this study, we present a security and robustness SDN-Enabled Cloud model using OpenStack and OpenDaylight. In particular, we design and implement a security clustering-based SDN Controller for monitoring and managing cloud networking, and a hardware platform to accelerate packet processing in virtual switches. We evaluate our proposed model on a practical cloud testbed consisting of several physical and virtual nodes. The experiment results show that the SDN controller cluster significantly improve robustness for the network even in case of being attacked by abnormal network traffic; while the hardware-accelerated switches can be operated in high-performance and well-adapted to the cloud environment.

**Keywords:** *Cloud Computing, SDN, OpenStack, OpenDaylight, Network Security.*

# UNBIASED PAIRWISE APPROACH TOWARD LEARNING-TO-RANK: AN EMPIRICAL STUDY

**Son Thanh Le** <sup>(1), (2)</sup>, **Ha Manh Tran** <sup>(3)</sup>, **Quang Duy Nguyen** <sup>(1), (2)</sup>,  
**and Sinh Van Nguyen** <sup>(1), (2)</sup>

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam

<sup>(2)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM), Ho Chi Minh City, Vietnam

<sup>(3)</sup> Department of Information Technology,

HCMC University of Foreign Languages - Information Technology, Vietnam

Email: hatm@hufliit.edu.vn

## ABSTRACT

With the bloom of information technology in recent decades, people are constantly being exposed to a huge amount of information. Learning-to-rank comes out as one of the solutions to ease out the mentioned obstacle by trying to rearrange objects according to their degrees of importance or relevance. This solution usually applies machine learning techniques to construct ranking models in information retrieval systems. The aim of this study is to explore and experiment the existing learning-to-rank approaches with real-life logs data. The study also includes estimating and minimizing the bias noise found in the click-through data of the recorded logs. Evaluation results have presented the advantage and disadvantage of the experimented approaches in realistic settings.

**Keywords:** *Unbiased Pairwise, Learning-to-Rank, Machine Learning, Click Logs.*

# REAL-TIME FACE DETECTION AND HUMAN TRACKING SYSTEM ON FPGA CYCLONE-V

Huu Luong Nguyen<sup>(1), (2)</sup>, Minh Son Nguyen<sup>(1), (2)</sup>, and Tri Nhut Do<sup>(1), (2)</sup>

<sup>(1)</sup> Ho Chi Minh City University of Technology (HCMUT),

Ho Chi Minh City, Vietnam

<sup>(2)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM),

Ho Chi Minh City, Vietnam

Email: {luongnh; sonnm; trinhutdo}@uit.edu.vn

## ABSTRACT

Face detection in image sequence (real-time video stream) has been an active research area in the computer vision field in recent years due to its potential applications such as surveillance cameras, human computer interfaces, smart rooms, intelligent robots and biomedical image analysis. Face detection is a process that determines whether an image has a face or not. In this paper, an embedded system for detecting and tracking human faces in real-time video stream implemented on FPGA DE10-NANO is proposed. The system can be divided into two parts: data streaming, data processing. Experimental results show that the system is capable of accurately detecting faces of up to 5 different people at a distance of up to 1.5 meters from the camera, coexisting in the same frame in resolution of 320x240 pixels with a detection speed of only several hundred milliseconds prove the feasibility of the system. A comparison with similar existing projects will be discussed for evaluation and conclusion as well.

**Keywords:** *Embedded system, human tracking system, face detection system, FPGA.*

# TWO-PHASE DEFECT DETECTION USING CLUSTERING AND CLASSIFICATION METHODS

**Ha Manh Tran<sup>(1)</sup>, Tuan Anh Nguyen<sup>(1)</sup>, Son Thanh Le<sup>(2),(3)</sup>,  
Giang Vu Truong Huynh<sup>(4)</sup>, and Tuan Bao Lam<sup>(4)</sup>**

<sup>(1)</sup> Department of Information Technology,

HCMC University of Foreign Languages - Information Technology, Vietnam

<sup>(2)</sup> School of Computer Science and Engineering, International University,

Ho Chi Minh City, Vietnam

<sup>(3)</sup> Vietnam National University Ho Chi Minh City (VNU-HCM), Ho Chi Minh City, Vietnam

<sup>(4)</sup> Center of Information Management, Vietnam Posts and Telecommunications Group,

Tay Ninh Province, Vietnam

Email: hatm@huflit.edu.vn

## ABSTRACT

Autonomous fault management of network and distributed systems is a challenging research problem and attracts many research activities. Solving this problem heavily depends on expertise knowledge and supporting tools for monitoring and detecting defects automatically. Recent research activities have focused on machine learning techniques that scrutinize system output data for mining abnormal events and detecting defects. This paper proposes a two-phase defect detection for network and distributed systems using log messages clustering and classification. The approach takes advantage of K-means clustering method to obtain abnormal messages and random forest method to detect the relationship of the abnormal messages and the existing defects. Several experiments have evaluated the performance of this approach using the log message data of Hadoop Distributed File System (HDFS) and the bug report data of Bug Tracking System (BTS). Evaluation results have disclosed some remarks with lessons learned.

**Keywords:** *Defect Detection, Clustering, Classification, HDFS Logs, BTS Bugs, Network Fault Management.*