

A Pragmatic Research Analytics on Real Time Cloud Platforms for Assorted Scenarios

Ravi Vasdev

M.Tech. Research Scholar

*Department of Computer Science and Engineering
Universal Group of Institutes
Lalru, Punjab, India*

Priyanka

Head of Department, Assistant Professor

*Department of Computer Science and Engineering
Universal Group of Institutes
Lalru, Punjab, India*

Abstract

Cloud Computing is one of the emerging domain now days with the assorted cloud service providers. Most of the effective cloud service providers are commercial and can be hired on per hour basis or per usage basis. In this research work, the effectual implementation of the performance evaluation of cloud services is done with the case scenarios of IBM Bluemix and RedHat OpenShift. These cloud service providers are considered as effectual and high performance, still there is need to evaluate the performance of these cloud platforms on real world code of social media extraction. In this research work, the analytics of PHP based code on the real clouds of IBM Bluemix and RedHat OpenShift is done and found the effectual and comparative results on multiple parameters in the real time environment. The RedHat OpenShift platform is found quite better in some specific applications and IBM Bluemix is effectual and performance aware in the particular domains.

Keywords : Cloud Computing, Cloud Environment, Internet of Things, Cloud of Things

Introduction

Cloud computing is a unique and effective way to use the computing infrastructure for effective communication and less overhead. The major technology behind grid as well as cloud computing is the virtualization technology. Using virtualization technology, the number of arbitrary and remote processors works in parallel for the execution of tasks in such a way that there is less complexity and higher integrity of the processes.

In cloud and grid architecture, there are number of remote computing devices and processors which work and execute the processes in parallel so that there is minimum load on the single node or machine. By this way, the overall load is balanced.

There are number of research areas in the domain of grid computing and cloud computing, still few areas are still in the preferences of the research community as these areas need updates and enhancements in regular basis.

The major areas of research in grid and cloud architecture are -

- Load Balancing
- Power Optimization
- Energy Optimization
- Secured Routing
- Task Scheduling
- Confidentiality and Integrity Management

Public and Private Key Cryptography and many others

Literature Review

Kune et al. [1] (2014) – proposed a Genetic Algorithm based scheduler for such Big Data Cloud where decoupled computational and information administrations are offered as administrations. The methodology is taking into account developmental systems focused on information conditions, computational assets and viable use of data transmission subsequently accomplishing higher throughputs

Mathew et al. [2] (2014) – presented a nitty gritty investigation of different undertaking booking

routines existing for the cloud environment. A brief investigation of different booking parameters considered in these systems is additionally examined in this paper.

Singh et al. [3] (2014) – proposed and implemented a hereditary calculation based planning methodology in which beginning populace is created with development rendition of Max Min by which the work can get more streamline results, regarding make compass. The execution of the proposed Modified Genetic Algorithm (MGA) and existing calculations have been assessed against the example information. Trial results demonstrate that proposed calculation shows adequate execution and beats the current calculations.

Verma et al. [4] (2014) – proposed the task planning and asset assignment are the key difficulties of distributed computing. Contrasted and framework environment, information exchange is a major overhead for cloud work processes. Along these lines, the expense emerging from information exchanges between assets and also execution costs should likewise be considered amid planning based upon client's Quality of Service (QoS) imperatives.

Javanmardi et al. [5] (2014) - presented and proposed the guide of hereditary calculation and fluffy hypothesis, the work presented a half genetic based job planning methodology, which considers the heap adjusting of the framework and decreases downright execution time and execution cost. The work attempted to adjust the standard Genetic calculation and to decrease the cycle of making populace with the guide of fluffy hypothesis. The principle objective of this exploration is to appoint the employments to the assets with considering the VM MIPS and length of jobs. The new calculation relegates the employments to the assets with considering the employment length and assets limits. The work assess the execution of our methodology with some renowned cloud booking models. The aftereffects of the investigations demonstrate the effectiveness of the proposed approach in term of execution time, execution cost and normal Degree of Imbalance (DI).

Rodriguez et al. [6] (2014) – proposed and presented an asset provisioning and planning system for investigative work processes on Infrastructure as a Service (IaaS) mists. This work exhibit a calculation taking into account the meta-heuristic improvement system, molecule swarm enhancement (PSO), which expects to minimize the general work process execution expense while meeting due date limitations. The heuristic is assessed utilizing CloudSim and different no doubt understood logical work processes of distinctive sizes. The outcomes demonstrated that the methodology performs better than the current best in class calculations.

Singh et al. [7] (2014) – proposed a hereditary calculation and recommended that timetables work process applications in questionable cloud environment and meet client characterized QoS limitations. A financial plan compelled time minimization hereditary calculation has been proposed which diminishes the disappointment rate and makespan of work process applications. It assigns those assets to work process application which are solid and expense of execution is under client spending plan. The execution of hereditary calculation has been contrasted and max-min and min-min planning calculations in inconsistent cloud environment.

Shojafar et al. [8] (2015) - scientifically demonstrated the streamlining issue which is arched with surely understood investigative conditions (particularly, Karush–Kuhn–Tucker conditions). The authors in this work contrasted the execution of our methodology with a few other cloud booking models.

Bollen et al. [9] (2010) – proposed the behavioral economics that emotions can profoundly affect individual behavior and decision-making. Does this also apply to societies at large, i.e. can societies experience mood states that affect their collective decision making? By extension is the public mood correlated or even predictive of economic indicators? Here we investigate whether measurements of collective mood states derived from large-scale Twitter feeds are correlated to the value of the Dow Jones Industrial Average (DJIA) over time. This work analyze the text content of

daily Twitter feeds by two mood tracking tools, namely OpinionFinder that measures positive vs. negative mood and Google-Profile of Mood States (GPOMS) that measures mood in terms of 6 dimensions (Calm, Alert, Sure, Vital, Kind, and Happy). This paper cross-validate the resulting mood time series by comparing their ability to detect the public's response to the presidential election and Thanksgiving day in 2008. A Granger causality analysis and a Self-Organizing Fuzzy Neural Network are then used to investigate the hypothesis that public mood states, as measured by the OpinionFinder and GPOMS mood time series, are predictive of changes in DJIA closing values. The results in this paper indicate that the accuracy of DJIA predictions can be significantly improved by the inclusion of specific public mood dimensions but not others. The authors find an accuracy of 87.6% in predicting the daily up and down changes in the closing values of the DJIA and a reduction of the Mean Average Percentage Error by more than 6%.

Bollen et al. [10] (2009) – proposed and presented that microblogging is a form of online communication by which users broadcast brief text updates, also known as tweets, to the public or a selected circle of contacts. A variegated mosaic of microblogging uses has emerged since the launch of Twitter in 2006: daily chatter, conversation, information sharing, and news commentary, among others. Regardless of their content and intended use, tweets often convey pertinent information about their author's mood status. As such, tweets can be regarded as temporally-authentic microscopic instantiations of public mood state. In this article, we perform a sentiment analysis of all public tweets broadcasted by Twitter users between August 1 and December 20, 2008. For every day in the timeline, we extract six dimensions of mood (tension, depression, anger, vigor, fatigue, confusion) using an extended version of the POMS, a well-established psychometric instrument. The authors in this paper compare the results to the values recorded by stock market and crude oil price indices and major events in media and popular culture, such as the U.S. Presidential Election of November 4, 2008 and Thanksgiving Day. This work finds that events in the social, political, cultural and economic sphere do have a significant,

immediate on the various dimensions of public mood. The authors speculate that large scale analyses of mood can provide a solid Platform to model collective emotive trends in terms of their predictive value with regards to existing social as well as economic indicators.

Davidov et al. (2010) [14] – proposed and presented that automated identification of diverse sentiment types can be beneficial for many NLP systems such as review summarization and public media analysis. In some of these systems there is an option of assigning a sentiment value to a single sentence or a very short text. In this paper the authors propose a supervised sentiment classification framework which is based on data from Twitter, a popular microblogging service. By utilizing 50 Twitter tags and 15 smileys as sentiment labels, this framework avoids the need for labor intensive manual annotation, allowing identification and classification of diverse sentiment types of short texts. The authors evaluate the contribution of different feature types for sentiment classification and show that the proposed framework successfully identifies sentiment types of untagged sentences. The quality of the sentiment identification was also confirmed by human judges. This paper also explores dependencies and overlap between different sentiment types represented by smileys and Twitter hashtags.

IBM Bluemix is one of the leading cloud service providers with number of services including Big Data Analytics, Mobile Applications, NoSQL Databases, Internet of Things (IoT), Natural Language Processing, Sentiment Analysis and a lot of related high performance computing services.

IBM Bluemix Platform can be accessed using freemium model and there is the feature to use of boilerplates. Using boilerplates, the creation and updation in the live code is very effective and easy so that the beginners can use the cloud services of IBM without any difficulty.

In IBM Bluemix, a detailed catalog is presented to the user having number of cloud services including the programming Platform for PHP, Java, Ruby, Python and many other languages. The cloud user

can use Bluemix as Platform as a service for hosting and running multiple applications.

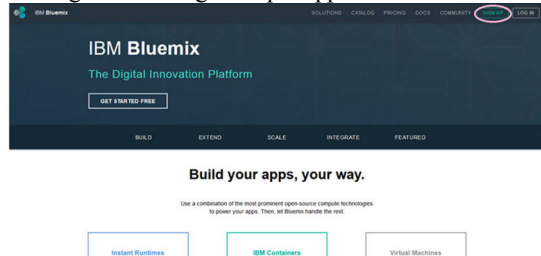


Fig. 1. IBM Bluemix

After logging in, Click on the 'Catalog' button to view and select different Platforms and services provided by IBM Bluemix.

In this example, the IoT Application will be created in IBM Bluemix. On Catalog page, click "Internet of Things Application Foundation Starter". This is simulation scenario of sensing the temperature.

As in figure 4, the App Name is to be specified while creating the live cloud application. Using this unique name, the application is identified and executed on the cloud servers.

The actual device id is mapped with the actual device. Suppose there is a particular machine and we want to sense its temperature using IoT. In this case, the unique machine id is mentioned in Node RED so that the application can sense the live temperature or any other required parameter.

IBM Bluemix provide PHP starter application template to the cloud users so that the own code can be added and pushed back to the cloud server. This starter application template is the boilerplate for PHP web application development in Bluemix. Initially, a new PHP starter application is activated and starter code is downloaded. After that the App name and Host is provided by the user. For further implementation, the cf (Cloud Foundry) command line tool is used installed. cf is the CLI (command line interface) that can be used to deploy and control the applications on Bluemix.

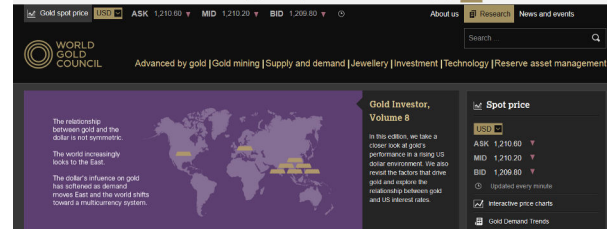


Fig. 2. Fetching Live Gold Data

OpenShift (<http://www.openshift.com>) is gained a lot of popularity with developers as it is having the enormous speed by which the quick spinning up a new development stack is possible.

To start using the OpenShift Platform, the very first step a developer requires is an account that provides access to create and deploy applications.

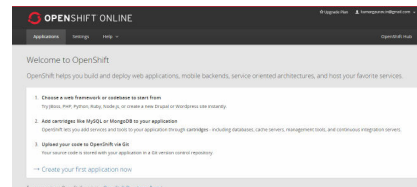


Fig. 3. Creating PHP Application Page in Red Hat OpenShift

After creating the new account, there is the option to create a new application. In the online Platform, the complete code can also be uploaded using Git so that the pre-built application can be launched on OpenShift Platform.

To work on a new application, the OpenShift Platform presents the number of options which are in the interest of developer to use. If a Python application is to be developed, the OpenShift gives the Platform to program and execute Python Code online.

After basic settings and selection of the public URL, a new application is created on which the code can be deployed and executed.

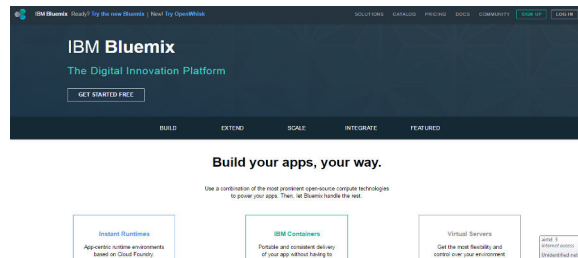


Fig. 4. Joining a IBM Bluemix Platform

The figure represents the panel for joining the IBM Bluemix Platform. Any user can create the account on IBM Bluemix and then experiments can be done.

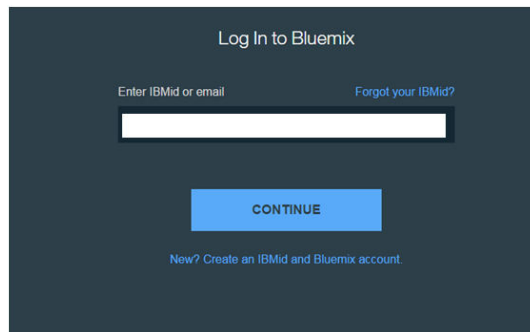


Fig. 5. Open personal account on IBM Platform

The personal account is opened on Bluemix and then the Bluemix id is provided. Once the IBM id is allocated, the implementation of cloud services can be done.

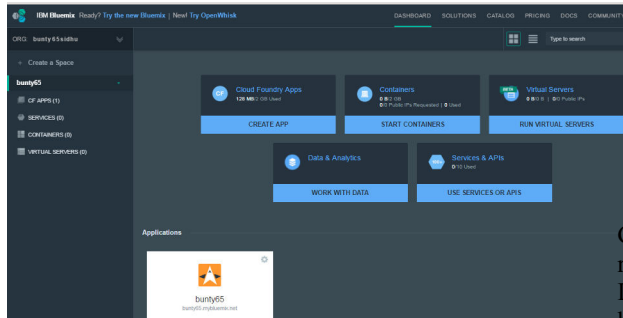


Fig. 6. Open personal coding IBM Platform

After creation and activation of account, the coding on any language can be done. IBM Bluemix provides the panel and Platform for coding or implementation of code so that the real experience of cloud can be done

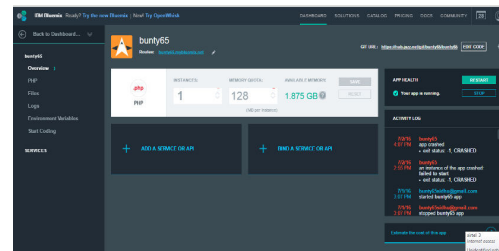


Fig. 7. Edit personal coding IBM Platform

The figure above depicts the customization of cloud environment adopted from IBM Bluemix. The Platform is having detailed catalogue and services provided at the Bluemix panel. Using this panel, the cloud user can adopt and customize the cloud infrastructure so that the execution on real cloud can be done. IBM Bluemix is a cloud Platform as a service (PaaS) developed by IBM. It supports several programming languages and services as well as integrated DevOps to build, run, deploy and manage applications on the cloud. Bluemix is based on Cloud Foundry open technology and runs on SoftLayer infrastructure.

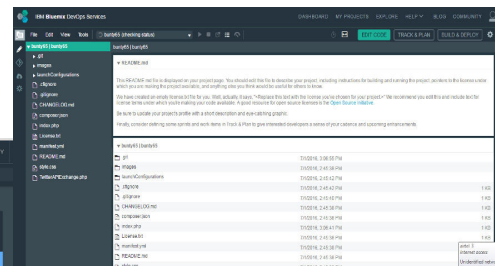


Fig. 8. Platform to add coding on live IBM record

Git is a commonly used decentralized source code repository. It was created by the Linux creator Linus Torvalds for the management of the Linux kernel source code. Whole services like GitHub are based around it. It is even used in IBM's DevOps Services alongside the IBM Rational Team Concert source code repository. So if you want to program

in the Linux world or use IBM's DevOps Services with Git, it helps to have a good understanding of Git.

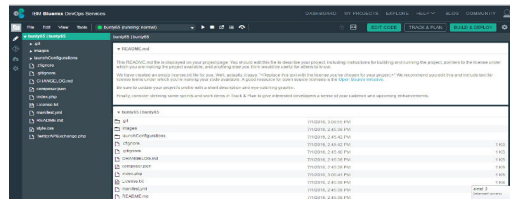


Fig. 9. Code And Deploy Live IBM Record

You can create a public or private Git repo that is hosted on Bluemix. You control access to the repo by adding members to your DevOps Services project.

If you want to store your source code in a hosted Git repo and you already have a local Git repo, create your DevOps Services project and then force push your local repository to the new hosted repo



Fig. 10. Live Data From IBM Platform

The Streaming APIs give developers low latency access to Twitter's global stream of Tweet data. A proper implementation of a streaming client will be pushed messages indicating Tweets and other events have occurred, without any of the overhead associated with polling a REST endpoint.

Connecting to the streaming API requires keeping a persistent HTTP connection open. In many cases this involves thinking about your application differently than if you were interacting with the REST API. For an example, consider a web application which accepts user requests, makes one or more requests to Twitter's API, then formats and

prints the result to the user, as a response to the user's initial request

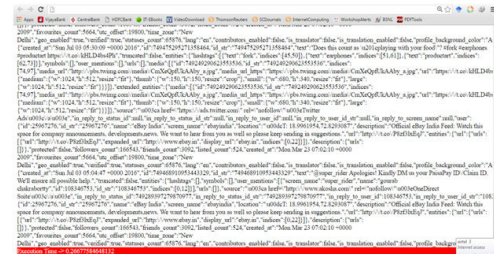


Fig. 11. To Note Time of Running Algorithm

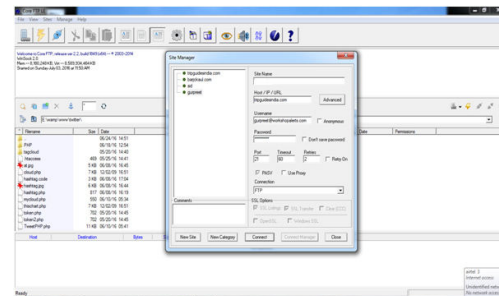


Fig. 12. Live Data From Redhat Platform

FTP Client makes use of FTP (File Transfer Protocol), FTPS, SFTP, WebDAV, S3, Google Drive, OneDrive, SSH, Terminal client. It allows you to transfer files between your local computer and a server on the Internet.

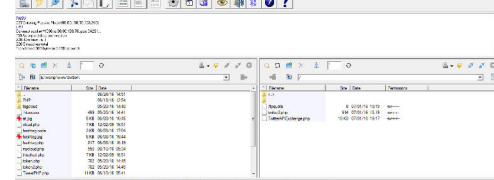


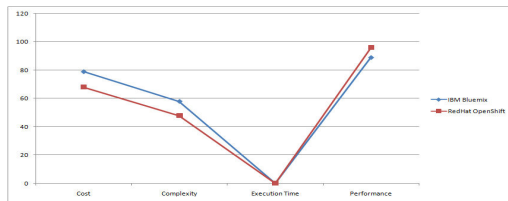
Fig. 13. Code From Redhat Platform



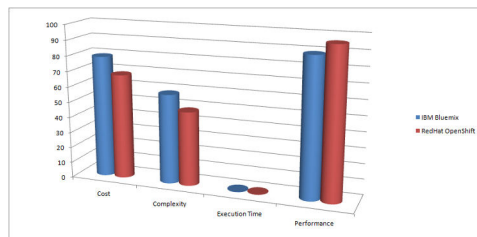
Fig. 14. Output From Redhat Platform

Table 1. Comparative Analysis on Assorted Parameters

	Co st	Com plexi ty	Executio n Time (ms)	Perfor mance
IBM Blue mix	78	59	0.271838	87
RedH at Open Shift	66	46	0.1297447	97

**Figure 15. Comparative Analysis**

A line graph is commonly used to display change over time as a series of data points connected by straight line segments on two axes. The line graph therefore helps to determine the relationship between two sets of values, with one data set always being dependent on the other set. From the graph it is clearly visible on the performance of IBM and RedHat in terms of assorted parameters.

**Figure 16. Bar Graph of Comparative Analysis**

Many social networks and apps have their own interface that programmers can work with. These interfaces are called APIs (short for Application Programming Interface).

Conclusion

There is huge scope of research and development using Python scripts and specialized APIs for

assorted applications including cyber security, data mining, Internet of Things, cloud simulation, grid implementation and many others. Python is one of the effective programming languages that can process and handle any type of data stream. In the later work, the live fetching of social media messages not integrated with assorted advance data mining tools for effective prediction and modeling. For this reason the data value are not real from the live data sources so we work on live data fetching. Later problem is that it is not clear which Platform of cloud provider's is best for working.

As a result of the advances in the innovation, the issues and further extension in science and building are turning out to be more convoluted than any time in recent memory. To tackle these muddled issues, network figuring turns into a well known apparatus. A framework domain gathers, coordinates, and uses heterogeneous or homogeneous assets scattered the world over by a rapid system. A lattice situation can be ordered into two sorts: processing networks and information frameworks. This exploration work concentrates on employment booking calculations and their execution on numerous parameters in the lattice environment. In processing matrix, occupation booking is a critical undertaking. A decent booking calculation can allocate occupations to assets productively and can adjust the framework load.

References

- [1] R. Kune and Konugurthi, Genetic Algorithm based Data-aware Group Scheduling for Big Data Clouds. In Proceedings of the 2014 IEEE/ACM International Symposium on Big Data Computing (pp. 96-104). IEEE Computer Society.
- [2] T. Mathew and Sekaran, Study and analysis of various task scheduling algorithms in the cloud computing environment. In Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on (pp. 658-664). IEEE.
- [3] S. Singh and M. Kalra, Scheduling of Independent Tasks in Cloud Computing Using Modified Genetic Algorithm. In Computational Intelligence and

- Communication Networks (CICN), 2014 International Conference on (pp. 565-569). IEEE.
- [4] A. Verma and S. Kaushal Deadline constraint heuristic-based genetic algorithm for workflow scheduling in cloud. *International Journal of Grid and Utility Computing*, 5(2), 96-106.
- [5] S. Javanmardi, M. Shojafar Hybrid job scheduling algorithm for cloud computing environment. In *Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014* (pp. 43-52). Springer International Publishing.
- [6] M. Rodriguez and R. Buyya Deadline based resource provisioning and scheduling algorithm for scientific workflows on clouds. *Cloud Computing, IEEE Transactions on*, 2(2), 222-235.
- [7] L. Singh and S. Singh, A Genetic Algorithm for Scheduling Workflow Applications in Unreliable Cloud Environment. In *Recent Trends in Computer Networks and Distributed Systems Security* (pp. 139-150). Springer Berlin Heidelberg.
- [8] M. Shojafar and S. Javanmardi, FUGE: A joint meta-heuristic approach to cloud job scheduling algorithm using fuzzy theory and a genetic method. *Cluster Computing*, 18(2), 829-844.
- [9] L. Yan, "Strengthen cloud computing security with federal identity management using hierarchical identity-based cryptography." In *Cloud Computing*, pp. 167-177. Springer Berlin Heidelberg, 2009
- [10] Goldwasser, "A digital signature scheme secure against adaptive chosen-message attacks." *SIAM Journal on Computing* 17, no. 2 (1988) : 281-308
- [11] S. Sanka and C. Chittaranjan, "Secure data access in cloud computing." In *Internet Multimedia Services Architecture and Application (IMSAA)*, 2010 IEEE 4th International Conference on, pp. 1-6. IEEE, 2010
- [12] M. Jensen and L. Iacono, "On technical security issues in cloud computing." In *Cloud Computing, 2009. CLOUD'09. IEEE International Conference on*, pp. 109-116. IEEE, 2009
- [13] R. Mariki and M. Eloff, "The management of security in cloud computing." In *Information Security for South Africa (ISSA)*, 2010, pp. 1-7. IEEE, 2010
- [14] W. Jansen, "Cloud hooks : Security and privacy issues in cloud computing." In *System Sciences (HICSS)*, 2011 44th Hawaii International Conference on, pp. 1-10. IEEE, 2011
- [15] L. Chang, "An authenticated key exchange scheme for efficient security-aware scheduling of scientific applications in cloud computing." In *Dependable, Autonomic and Secure Computing (DASC)*, 2011 IEEE Ninth International Conference on, pp. 372-379. IEEE, 2011
- [16] N. Santos, K. Nuno, "Towards trusted cloud computing." In *Proceedings of the 2009 conference on Hot topics in cloud computing*, pp. 3-3. 2009 –
- [17] A. Intisar, Evolution of grid computing architecture and grid adoption models. *IBM Systems Journal*, 43(4), 624-645, 2015
-