INTERNATIONAL JOURNAL OF

New Computer Architectures and their Applications

SDIWC

www.sdiwc.net

Overview

The SDIWC International Journal of New Computer Architectures and Their Applications (IJNCAA) is a refereed online journal designed to address the following topics: new computer architectures, digital resources, and mobile devices, including cell phones. In our opinion, cell phones in their current state are really computers, and the gap between these devices and the capabilities of the computers will soon disappear. Original unpublished manuscripts are solicited in the areas such as computer architectures, parallel and distributed systems, microprocessors and microsystems, storage management, communications management, reliability, and VLSI.

One of the most important aims of this journal is to increase the usage and impact of knowledge as well as increasing the visibility and ease of use of scientific materials, IJNCAA does NOT CHARGE authors for any publication fee for online publishing of their materials in the journal and does NOT CHARGE readers or their institutions for accessing the published materials.

CONTENTS
ORIGINAL ARTICLES

A Performance Optimization Model of Task Scheduling towards Green Cloud Computing

Masnida Hussin, Raja Azlina Raja Mahmood, Nor Azura Husin, Noris Mohd. Norowi
Faculty of Computer Science & Information Technology
University Putra Malaysia, Selangor, Malaysia
masnida, raja_azlina, n_azura, noris @ upm.edu.my

Abstract: Cloud computing becomes a powerful trend in the development of ICT services. It allows dynamic resource scaling from infinite resource pool for supporting Cloud users. Such scenario leads to necessity of larger size of computing infrastructure and increases processing power. Demand on the cloud computing is continually growth that makes it changes to scope of green cloud computing. It aims to reduce energy consumption in Cloud computing while maintaining a better performance. However, there is lack of performance metric that analyzing trade-off between energy consumption and performance. Considering high volume of mixed users' requirements and diversity of services offered; an appropriate performance model for achieving better balance between Cloud performance and energy consumption is needed. In this work, we focus on green Cloud Computing through scheduling optimization model. Specifically, we investigate a relationship between performance metrics that chosen in scheduling approaches with energy consumption for energy efficiency. Through such relationship, we develop an energy-based performance model that provides a clear picture on parameter selection in scheduling for effective energy management. We believed that better understanding on how to model the scheduling performance will lead to green Cloud computing.

Keywords: green cloud; scheduling; energy efficiency; optimization model; energy management.

1. Introduction

Cloud computing is a state-of-the-art technology where it provides computing services i.e., IaaS, PaaS and SaaS to users [1, 2]. Basically, Cloud providers e.g., Amazon, Microsoft and Google provide their users with resource sharing model where resources (e.g., processors, storage) can be added and released easily either they are needed or otherwise. Particularly, large computing and storage infrastructure like Cloud needs more energy to generate sufficient electricity and cooling systems thus more expenses need to be invested. Furthermore, computer systems not only consuming vast amount of power also emit excessive heat; this often results in system unreliability and performance degradation. It has been reported in previous studies (e.g.,[3-5]) that system overheating causes system freeze and frequent system failures. According to the authors in [4] the highest energy cost of data centers are used to maintain the running servers. The servers need to be available and accessible throughout the year even though nobody is accessing them. Such situation incurs high cost in electric and cooling systems.

In order to sustain with good service reputation, the data centers needed to facilitate the processing and storage requirements for 24/7. The consequence of high resource availability in Cloud is not merely to financial expense but worse to environment through carbon footprint. There are more greenhouse gas (GHG) released in the atmosphere that leads to global warming, acid rain and smog [6]. Although there have been many research efforts to reduce the energy consumption in computing operation, there is still lack of a decision support system for choosing the right performance metric in dynamic computing environment. The large-scale distributed system like Cloud needs to support large number of users with diverse processing requirements. Meanwhile, the performance degradation at any stage of processing is unacceptable. It is a challenge to find the best trade-off between system performance and energy consumption in Cloud.

In task scheduling approach it mainly aims to maximize or at least sustain the system performance. Such scheduling approach able to promote better energy consumption by effectively scheduling users' requirements based on resource availability [7-10]. In the conventional scheduling optimization model, the performance parameters are usually specified as deterministic. However, in realistic energy management systems, many parameters are dealing with uncertain natures with multiple states and features. When it comes to green Cloud computing, several issues are need to

take into account, for example, security, accuracy and energy. In respect of energy saving, there are few aspects that needed to be highlighted such as power management, virtualization, and cooling. Energy efficiency for Cloud computing involved several series of interaction and interdependent between relevant system components, processes and metrics. Hence, it is needed a clear guideline for assessing performance and energy efficacy of Cloud especially in its scheduling and provisioning activities.

In this paper we focus on efficient energy consumption that influenced by parameter selection in the scheduling approach. We develop optimization model for energy-based scheduling that aims to provide a guideline for analyzing the better trade-off between system performance and energy consumption in Cloud. We then investigate the relationship of performance metrics that chosen in the scheduling approach with energy consumption to facilitate green computing. It is wise to effectively manage energy consumption in Cloud computing; this would in turn be very beneficial cost of computation.

The reminder of this paper is organized as follows. A review of related work is presented in Section 2. In Section 3 we describe the parameter categories that chosen for efficient energy management. Section 4 details our performance model for energy efficiency. Experimental setting and results are presented in Section 5. Finally, conclusions are made in Section 6.

2. Literature Review

The energy management has inspired many researchers [3, 4, 11-15] to focus on green Cloud computing. The scope of energy assessment for modeling its performance should be stretched further incorporating parameter selection and energy model that being used. In [16], the authors emphasized resource utilization in their energy model. They utilized the virtualization as a core technology to contribute green IT because it able to share limited resources with varies workloads. The quantity of physical machines (PMs) can be reduced where processing is performed by virtual machines (VMs). VMs are very flexible that act as independent servers that maximized resources utilization while achieving energy efficiency. However, the selection of parameters in their

performance model for energy efficiency is still an open issue. Some researchers focused on resource state in their energy models. According to the authors in [12], processors consumed approximately 32Watt when they are operated in idle mode that compared to storages merely used 6Watt. In peak processing mode the energy consumption of processors can be boosted up to more than 80Watt to 95Watt [17]. Energy efficient scheduling that proposed in [8] dynamically allocated users' tasks into processor to achieve better performance and minimize energy consumption. The system performance and energy consumption in their work is been measured throughout the task execution either during peak or idle state, then total energy consumption is recorded.

There are also some works that calculated the energy consumption through their modern scheduling approaches. Power-efficient scheduling in [7] assigned set of virtual machines (VMs) to physical machines (PMs) for data centers management. They used the consolidation fitness to determine the right VMs to replace PMs when existing PM is been switched off. However there is challenge to determine the most right VM due to unpredictable changes in the system workload. Power-aware mechanism in [18] is based on priority scheduling for efficient energy management. They adopted various heuristic for energy-aware scheduling algorithms that employed multi-objective function for diverse efficiency-performance tradeoffs. Their scheduling algorithms consist of three steps (i.e., job clustering, re-evaluating to give better scheduling alternatives and selecting the best schedule). They conducted the experiments in homogeneous environment that disguises Cloud computing environment. The existing researches that proposed energy efficient scheduling are able to reach appropriate balance between system performance and energy consumption. However, the dynamicity and heterogeneity on their computing environment are limited to some extent.

There are some researchers (e.g.,[4, 6, 19]) that highlighted the incorporation of low-energy computing nodes in heterogeneous distributed systems and able to achieve energy efficiency. Due to the scheduling approaches are subjected to system environment and scale, it required effective performance model for better evaluation

on both performance and energy usage. However, there is lack of exclusive parameter selection strategy for energy management in order to balance between system performance and energy consumption. In Table I, we summarized the selection of performance metrics and system behaviors that used in those existing scheduling approaches to design energy model.

Table 1. Relationship between Scheduling Rule and Energy Model

No.	Scheduling Rule	Energy Model
1	Heuristic approach used for pre-scheduling processing. Such predictable rule satisfies energy saving where it consumed less processing time.	It considers busy and idle states of processing elements.
2	Threshold for processor, memory, disk and communication link in resource allocation. If the threshold chosen is too low it might reduces resource utilization while setting too high threshold leads to communication bottleneck.	It considers resource utilization and overhead in the system.
3	Migrating or remapping strategy in scheduling using a set of virtual machines (VMs).	Virtualization leads to energy efficiency.

3. Classification of Scheduling Parameters

The existing scheduling approaches (e.g.,[3, 4, 7, 20]) analyzed the performance of their energy management system through various scheduling algorithms. They have chosen several types of parameters such as time-based perspective, utilization and overhead.

3.1. Time-based Metrics

Time-based metric is one of the most popular parameters that chosen for focusing green Cloud. Such parameter used to measure the effectiveness of scheduling decision for uncertain, dynamic and large-scale environment. The time-based metrics, for example, execution time, waiting time and response time is designed to monitor and manage queuing problems in scheduling. The major problem in scheduling is to allocate various users' jobs to be mapped and executed by the right resources. Due to the scheduler needs to get information from users and resources, the scheduling decisions are most of the time consumed more processing energy compared than storage. The resources need to be available to perform the job execution at most of the time. Therefore, in order to improve the execution time for minimizing energy consumption we need to take into account the optimization technique in designing scheduling approach.

Furthermore, the issue on suitable queue length comes into a picture when the total waiting time of jobs leads for high energy consumption. Note that the determination of queue size is significantly related to the scale of computing system. Specifically, the queue size contributes for better job waiting time that it relates to buffer management. The suitable size of queue needs to be identified to reduce data access time in the buffer. We do not want to power the entire memory module in lengthy time only for data accessing operation. Hence, the effective scheduling approach for dynamic environment required to define the (most) suitable queue length for reducing power consumption.

3.2 Utilization

There are scheduling approaches that calculate the energy consumption based on computing resources' busy and idle states. In particular, the resources that have high busy time means the system utilization is improved. Meanwhile the system is considered has low utilization when there are many resources that idle in a given observation time. However, from energy management perspective, high utilization may leads to large energy consumption. The resource that has high utilization certainly consumed high processing energy in order to complete the task execution. It is more crucial in the idle state of resources. There is huge percentage of power consumption (i.e., for electricity and cooling systems) to facilitate the running resources. The idle resources need to be available even though there is no processing happen in their space. Resource utilization only 53% of total energy

consumed in data centers [21]. Therefore, the best solution is to effectively manage resource utilization for energy efficiency. There is a big role for the scheduler to monitor resource utilization in dynamic environment.

3.3. Overhead

The storage and processing resources in Cloud must be highly available that it reflects Quality of Service (QoS). It included the ability of Cloud to adapt with unexpected failures, e.g., storage overloaded, traffic congestion and performance fluctuation. Such scenarios needed extra time for the Cloud providers to solve and fix without notify the users. Some strategies used replication of objects and services, and using redundant processing and communication mechanisms to solve the unexpected failure. In order to implement these strategies they need more than one communication paths that used for disseminating the same information, and several processing elements for processing the same action. In scheduling approach, we need to design extra procedure or policy to manage such unexpected failure and implement reserve strategy. It explicitly incurred extra communication and processing overheads in the systems. For the sake of energy efficiency, overhead should be minimized while maintaining system performance. Hence, tunable parameters in experimental setting are significant to thoroughly identify the system behavior/action in order to achieving the target results (better trade-off).

4. Optimization Stages for Energy Efficiency

There are various performance models that adopted in scheduling approach for energy efficiency (e.g., [3, 4, 7, 9, 15]). In this work, we specifically divide our optimization model for energy-based scheduling into three stages; (i) identifying, (ii) formulating, and (iii) modifying.

4.1 Identifying Stage

There are several energy issues such as energy waste, inaccurate energy measurement etc. that identified by existing researchers (e.g. [6, 14, 22]). The percentage of energy consumption in task scheduling is normally differs from one scheduler to another as there are other factors that contribute to such amount [16]. Some scheduling approaches

e.g., heuristic and game-theory might improve the system utilization but energy efficiency. It is because energy management not entirely dependent on the selected scheduling approaches. Furthermore, the system goal generally aims to increase the system performance while minimizing the energy consumption. These maximize and minimize objectives need to be carefully designed in order to avoid nastiest unbalance outcome.
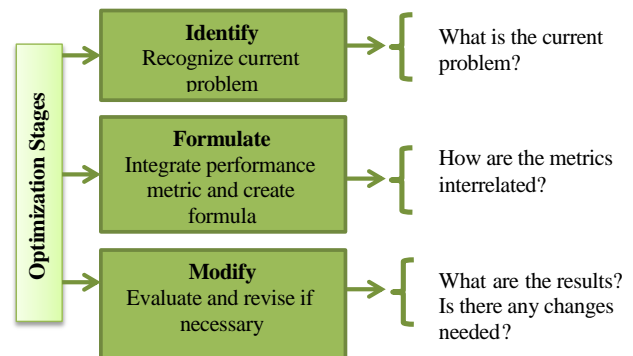
Figure 1. Energy-based Performance Model

In our optimization model, the identifying stage focuses on analyzing current energy-performance problem in the system. Initially, there is required to thoroughly understand the relationship between the system performance and energy consumption in the current computing system. Such review is important to determine a pattern on how the energy consumption crosses or touches the system performance during the execution process. The investigation on energy consumption in large-scale data center can be analyzed through its operational infrastructure. Such infrastructure can be classified based on power usage for physical equipment and processing condition. It leads to two different measurements of energy efficiency are; power usage effectiveness (PUE) and data center effectiveness (DCE) [16]. For PUE measurement, it concerns on total power used for IT equipment such as server, routers and cabling. Meanwhile DCE is calculated through the resource management strategy (i.e., scheduling, load balancing and security) that been applied in the server room or data center for complying the users' requirements. Both PUE and DCE are inter-related to each other for supporting computing operations.

Note that the high performance scheduler consumes high processing power at its peak processing time. This means to meet the users'

requirements for task scheduling purposes, hence increases the energy consumption exponentially. At other scenario, there is IT equipment in data center that used electricity to make them available 24/7 to facilitate processing purposes. Even though Cloud utilizes the virtualization technology, it still relies on the physical computing equipment at its end support. For monitoring energy consumption from IT equipment perspectives, the energy distribution must be accurately measured. It is due to the usage of those equipment contributes to electricity bill that basically relates to operational and maintenance costs. For example, the server rooms or data centers needed of mechanical and electrical (M&E) infrastructure, also ventilating or cooling infrastructure for supporting the operational in the room. Nowadays, the actual cost for managing the IT equipment either in server room or laboratory has become big issue in organization.

The performance model then should be focusing on how to utilize the usage of IT equipment hence the energy distribution can be optimally consumed. Several strategies such as scheduling, load balancing and authorization are needed to be highlighted in order to optimally manage and utilize the energy distribution. In this work, we proposed task scheduling strategy that acts as mediator in order to monitor system performance and processing power in a same time. The scheduler is designed to be capable for analyzing the effectiveness of the scheduling processes by identifying which scheduling pattern leads to energy waste. In priori, the scheduler should be expected to produce a scheduling pattern that gives better performance without enlarge energy consumed.

4.2 Formulating Stage

Cloud computing enables its services (i.e., IaaS, PaaS, SaaS) for various users at anywhere for anytime. It required to provide reliable in both computation and communication activities. In order to sustain the performance, thorough investigation of resource management specifically task scheduling is required. Note that the system performances are much influenced by the feature of system and characteristics of the users. Therefore, it is important to formulate the performance parameters in assessment method

that based on system environment and its scale. In this stage, we gather relative performance metrics to create formula for assessing energy efficiency. Such criteria can be defined according to several levels that represent their complexity in communication and processing activities.

In this work, we highlight three level of system complexities; (i) homogenous/heterogeneous, (ii) static/dynamic and (iii) local/centralized/distributed. High in complexity means huge amount of energy consumption that consumed for task scheduling decisions. For example, the system is considered consumed large amount of processing energy when there is a huge number of incoming tasks to be scheduled. Also, the data center is needed a large number of processing elements (PE) to offer high availability in processing. In some cases, the small number of PE is yet consumed large percentage of energy. It happens when the PE is operated for 24/7 and it needed very cold room to control the heat releases. Note that excessive heat emitted by these PE causes notable power consumption for cooling them. In addition, the energy consumption is proportional to the users' requirements. The resources operated in extensive processing time if there is huge number of workload in the system. Such situation leads to increase the energy usage in the whole system operation.

In response to the system and user criteria, selection of performance metric to achieve better trade-off between performance and energy consumption is a huge challenge. The users in large-scale computing system normally demand for varies of processing requirements. Hence, the system performance needed to fulfill the users' requirements in order to sustain the system's reputation. This is very important in Cloud because there are many Cloud providers that competing each other to provide better performance. Therefore, the performance metrics to evaluate energy consumption should be fragmented of a total/average of the system performance. The system performance might be reduced because its portion needs to embrace the energy consumed.

In response to the collection of performance metrics, we then raise issue on how to integrate each metric for energy efficiency. Basically, the metrics aim for better scheduling decisions that leads to improve the system performance. In response to energy efficiency, the design of

scheduling approach should able to monitor both performance and energy consumption for a given time duration. It is wised to capture energy consumption for a specific time duration that can be based on incoming workload while calculating the average of the execution time. It is because; by frequently measuring the energy consumption it implicitly increased a percentage of power usage. The right proportional of the system performance e.g., total execution time to measure the energy consumption must be significantly concerned.

For instances, the performance metric in Cloud data center is measured by processing overhead as follows.

$$\min_{\text{overhead}} = \left(\frac{\sum idle_{time} + \sum busy_{time}}{total\ number\ of\ task} \right)$$

(1)

The total energy consumption in the system can be calculated as given:

$$total\ energy\ consumption = \frac{\sum total\ overhead}{simulation\ time}$$

(2)

; assumed that the energy consumption is measured through simulation program. In such example, the total number of incoming task is recorded within the scheduling process that used to calculate the overhead. Meanwhile, the energy consumption is measured for entire simulation program. In some scheduling approaches (e.g.,[9, 21]), they used fix power consumption (in Watt) that identified at busy and idle states of processing element for calculating energy consumption.

Table 2. Suggestion Metrics

No.	Evaluation Scope	Metrics
1	Hardware measurement	Real-time voltage and current, processing state/frequency.
2	Data Center	Time-based performance, utilization, processing overhead.
3	Mobile Cloud Computing	Communication overhead (transaction delay and traffic congestion).

In this work, we also express some metric suggestion for energy efficiency that can be chosen based on a given evaluation scope (Table 2). From Table 2, it is merely that item number 2 and 3 are related to scheduling approach where it can be formulated to maintain the system performance while monitoring the energy consumed. Due to the large scale constitution of Cloud environments, such parameter selection is considered to be unbounded and still an open issue.

4.3 Modifying Stage

Energy-based scheduling approach aims to schedule users' tasks to be run on the computer systems that consumed low energy consumption. It means that the computer system can tune the processing states to adapt with system changes while reducing the system energy consumption. In response to this, the computer system needs to collect run-time information of applications, monitor the processing energy consumption states, notify about states changes and compute energy-based scheduling decisions. Through the monitoring process, the system administrator is able to identify the processing activities that consumed large amount of energy. Then, it can perform modification in the scheduling policy in order to balance between the system performance and energy consumption. Such modification procedure might be involves several alteration techniques (i.e., fading, chaining, looping etc.).

The most significant step in this modification stage is trial-and-error process. Due to task scheduling in dynamic computing system is known as nondeterministic polynomial time (NP) problem, in this work we suggest to use the concept of *generate-and-test* which is one of heuristic methods in solving NP issue. The *generate-and-test* process will be investigated on how the systems and applications respond to the processing state changes. We then adjust the performance metrics that fit to the target goal (better trade-off between performance and energy consumption).

In some cases, the modifying stage implicitly used to track any misidentify in the previous stages. The dynamic scheduling involved many challenge issues such as high complexity, huge overhead and performance degradation. The identification on the right system environment with the suitable performance metrics is needed to be thoroughly analyzed. The suitability in both criteria for optimizing the system can be solved

during modification process. For example, the waiting time might be not the right performance metric used for calculating energy consumption in the scheduler queue for public Cloud. It is due to the public Cloud is much expected by the large-scale number of workload that coming 24/7 that results extensive waiting time. Hence, in modification stage, the Cloud provider can change the performance metric by combining both waiting time and processing overhead in order to measure the energy consumption. The Cloud provider can design adaptive backup and maintenance activities in scheduling approach to reduce processing power that consumed for such activities.

5. Experimental Results

In this section, we experimentally evaluate several task scheduling approaches through our proposed optimization model for energy management. The scheduling approaches are chosen based on the common scheduling algorithms e.g., FCFS, Shortest-Job-First (preemptive), Shortest-Job-First (non-preemptive) and random. The main extension made to these algorithms is the incorporation of our optimization model. It means that the scheduler in each algorithm will complied with the all three stages in our optimization model. In order to evaluate the effectiveness of our performance optimization model, we compared with the scheduling approach that does not support the optimization stages (named *non-opt scheduler*).

In our experiments, there are two different Cloud components are storage and compute Clouds. In Cloud storage it involved data and information services. Each compute Cloud contains a varying number of compute resources ranging from 5 to 8 similar to that used in [8]. Both Cloud components are attached with a Cloud broker or so-called as a scheduler, in our study. There are several users that regularly request and get response from the Cloud. For this work, in every experiment there is considered 20 users with random task arrivals.

Results in Fig. 3 clearly demonstrate the competent capability of *opt-scheduler*. While *SJF-nonpreemp* showed appealing results, it is observed that average processing overhead in *opt-scheduler* is about 70% better compared to *non-opt scheduler*. Meanwhile, Fig. 4 shows a

reduction of more than 50% on average in energy consumption. Note that energy consumption with *opt-scheduler* exhibits better results in all scheduling algorithm and indicates that the differences is relatively small. In overall, the *opt-scheduler* able to deliver low energy consumption and better processing overhead in several scheduling policies. Therefore, it is significant in aggregating three optimization stages (Section 4) to strive for better performance and energy consumed.
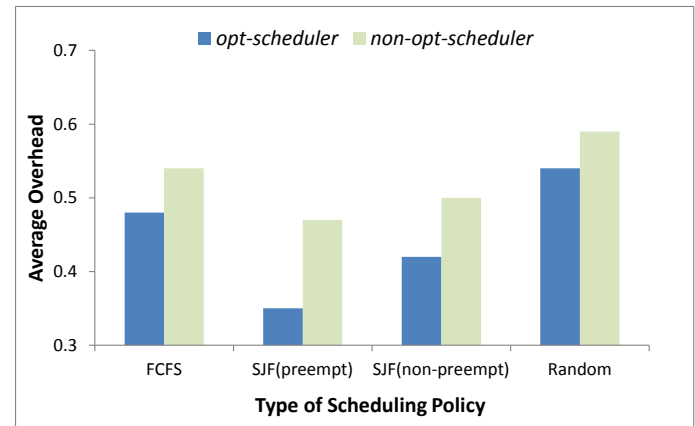
Figure 2: Average processing overhead for different scheduling policies
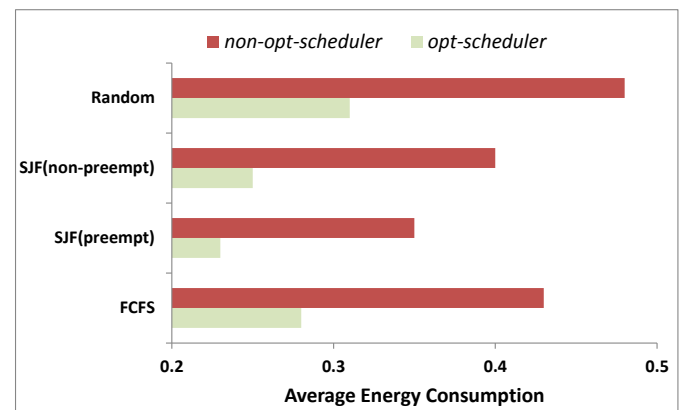
Figure 3: Average energy consumption for different scheduling policies

6. Conclusion

The concept of green computing has begun to spread in the past few years and still gaining its popularity. It is due to its significant performance, environmental and economic implications. In this paper, we analyze relationship between the parameter selection in scheduling approach and energy consumption that brings to energy efficiency. Specifically, we develop the

performance model for optimizing the task scheduling where the energy efficiency becomes the next goal. Note that, the (near) optimal scheduling for energy efficiency is still an open issue. Hence, there is a lot of potential for more research on its performance model and design. Optimistically, Cloud able to achieve better trade-off between performance and energy consumption when there is clear guideline for designing the energy-based performance model.

References

[1] C. Wu and R. Buyya, *Cloud Data Centers and Cost Modeling: A complete guide to planning, designing and building a cloud data center,*: Morgan Kaufmann, 2015.

[2] D. Rountree and I. Castrillo, *The Basics of Cloud Computing : Understanding the Fundamentals of Cloud Computing in Theory and Practise. .* MA USA: Elsevier, 2014.

[3] M. Hussin, Y. C. Lee, and A. Y. Zomaya, "Efficient Energy Management using Adaptive Reinforcement Learning-based Scheduling in Large-Scale Distributed Systems," presented at the 40th Int'l Conf. on Parallel Processing (ICPP2011), Taipei, Taiwan, 2011.

[4] Y. C. Lee and A. Y. Zomaya, "Energy efficient utilization of resources in cloud computing systems," *Journal of Supercomputing,* pp. 1-13, 2010.

[5] N. B. Rizvandi and A. Y. Zomaya, "A Primarily Survey on Energy Efficiency in Cloud and Distributed Computing Systems :," *ArXiv preprint arXiv:1104.5553,* 2012.

[6] Z. Abbasi, M. Jonas, A. Banerjee, S. Gupta, and G. Varsamopoulos, *Evolutionary Green Computing Solutions for Distributed Cyber Physical Systems* vol. 432: Springer, 2013.

[7] M. Sharifi, H. Salimi, and N. M., " Power-efficient distributed scheduling of virtual machines using workload-aware consolidation techniques.," *J. Supercomputing,* vol. 61, pp. 46-66, 2011.

[8] M. Hussin, Y. C. Lee, and A. Y. Zomaya, "Priority-Based Scheduling for Large-Scale Distribute Systems with Energy Awareness," in *Proc. of the 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing*, Sydney, Australia, 2011, pp. 503-509.

[9] K. Li, "Energy efficient scheduling of parallel tasks on multiprocessor computers," *Journal of Supercomputing,* pp. 1-25, 2010.

[10] N. Kim, J. Cho, and E. Seo, "Energy-credit scheduler: An energy-aware virtual machine scheduler for cloud systems," *Future Generation Computer Systems,* vol. 32, pp. 128-137, 2014.

[11] N. B. Rizvandi and A. Y. Zomaya, "A Primarily Survey on Energy Efficiency in Cloud and Distributed Computing Systems," *arXiv preprint* 2012.

[12] S. U. Khan and C. Ardil, "Energy Efficient Resource Allocation in Distributed Computing System," *World Academy of Science, Engineering and Technology,* pp. 667-673, 2009.

[13] W. X and W. Y, " Energy-efficient Multi-task Scheduling Based on MapReduce for Cloud Computing," *Seventh Int. Conf. Comput. Intell. Secur.,* 2011.

[14] D. Wei, L. Fangming, J. Hai, L. Bo, and L. Dan, "Harnessing renewable energy in cloud datacenters: opportunities and challenges," *Network, IEEE,* vol. 28, pp. 48-55, 2014.

[15] Kim, Nakku, J. Cho, and E. Seo, "Energy-credit scheduler: an energy-aware virtual machine scheduler for cloud systems.," *Future Generation Computer Systems,* vol. 32, pp. 128-137.

[16] A. Beloglazov, R. Buyya, Y. C. Lee, and A. Y. Zomaya, "A Taxonomy and Survey of Energy-Efficient Data Centers and Cloud Computing Systems," *Advances In Computers,* vol. 82, pp. 47-111, 2011.

[17] L. A. Barroso and U. Holzle, "The Case for Energy-Proportional Computing," *Journal Computer,* vol. 40, pp. 33-37, 2007.

[18] A. Aziz and H. El-Rewini, "Power efficient scheduling heuristics for energy conservation in computational grids.," *J. Supercomputing,* vol. 57, pp. 65-80, 2011.

[19] J. G. Koomey, *Estimating total power consumption by servers in the US and the world* vol. 15: February, 2007.

[20] C.-K. Tham and T. Luo, "Sensing-Driven Energy Purchasing in Smart Grid Cyber-Physical System," *IEEE Transaction on Systems, Man and Cybernetics: Systems,* vol. 43, pp. 773-784, 2013.

[21] L. Minas and B. Ellison, "InfoQ:The Problem of Power Consumption in Servers," in *Energy Efficiency for Information Technology*, ed: Intel PRESS, 2009.

[22] P. Schreier. (Dec 2008/Jan 2009) An Energy Crisis in HPC. *Scientific Computing World : HPC Projects*

Brain - Computer Interface for Communication and Estimation of Human Emotion from EEG and Video

Svetla Radeva and Dimitar Radev
University of Telecommunications and Post
1st. Acad. Stefan Mladenov Str., 1700 Sofia, Bulgaria
svetla_ktp@abv.bg, dradev@abv.bg

ABSTRACT

The brain-computer interface (BCI) aim to use Electroencephalography (EEG) or other measures of brain functions can be implemented for communication with smart devices for disabled persons. For connection with different smart devices was used recorded with experimental setup electrophysiological signals for execution of five different mental tasks. The recorded brain signals were processed for their transformation into commands to different devices. This signal processing aims to extract some specific features of brain signals and transform them into algorithms for connection with smart devices. Processed signals after noise filtering, clustering and classification with Bayesian Network classifier and pair-wise classifier was estimated and put into brain-computer interface for connection with smart devices. Recent advances in emotion recognition use a combination of two intrapersonal modalities - face and EEG to estimate emotion. In this research is made an attempt to combine received results on the base of record electrophysiological signals at execution of five different mental tasks with estimation of human emotion. This will help to provide a framework for reliable EEG emotional state estimation combined with facial emotion analysis in developed task-oriented BCI.

KEYWORDS

Brain-Computer Interface, Signal Processing, Neural Networks, Communication and Control.

1 INTRODUCTION

The brain-computer interface (BCI) aim to use different measures of brain functions for communication and control [1, 2]. Developing of such communication and control technology will improve the life of disabled persons. The brain-computer interfaces gives to these users possibility to switch on and off different devices [3, 4, 5], to run computer programs, to operate word-processing programs or neuro-prostheses [6, 7, 8].

The BCI gives possibility to control computer using brain directly, which is a very promising features for persons who cannot access traditional computer system due to their physical disabilities [9, 10, 11, 12].

The principle of BCI is based on measured brain activity while the user generates brain signals and to transform these signals into computer commands [13] for connection with smart device [14, 15, 16, 17].

Emotion analysis from human EEG signals is as well very attractive area for researchers. There already exist few types of approaches using EEG which are said to deliver promising emotion recognition results [18, 19].

The analysis of the existing research in the area shows that there is a large potential for improvement of the existing methods by means of using a novel approach to EEG emotion estimation combined with facial emotion analysis [20, 21].

Recent advances in emotion recognition use a combination of two intrapersonal modalities -

face and EEG to estimate emotion. This research consider implementation of BCI for communication and control, where estimation of emotional state supported by facial analysis is combined with estimation of different mental tasks execution.

2 SIGNAL PROCESSING AND CLUSTERING OF NEURONS

BCI system is realized on the following steps: when something has to be done a thought is developed into the brain which leads to development of a neuron potential pattern. Reading brain by register electrophysiological signals the developed potential pattern is transformed into an analyzable signal patterns. These signal patterns developed by BCI equipment and their spectrum is analyzed using various pattern analysis techniques. After recognition of human's intention about the task that brain wants to get from smart device or computer we can determine proper command (or sequence of commands) to execute the required task. Feedback to the user can be realized in various feedback-forms e.g. device switching on/off, emergency call, video, audio etc.

As is seen, for BCI communication with smart devices it is necessary to provide filtering of register brain signals and pattern analysis techniques for clustering of neurons and pattern recognition.

At any moment the human brain generates wave for a particular thought, but at the same time generates also some waves corresponding to other unnecessary thoughts. These additional waves appear because sometimes it is not possible to concentrate fully on a particular thought and they act as noise for the original waves. For handling with these noise-waves the user have to increase his concentration during the BCI process.

For solving this problem it is necessary to develop noise filtering mechanism that can detect the unrelated spectrum and filter them out from the useful spectrum [22].

The process of noise filters design requires some situation and application of specific knowledge from neurology. After some experiments for finding out the pattern of the signals was prepared and calibrated noise filter.

Another problem that has to be solved is connected with clustering of neurons, where it is necessary to divide 80-120 billion brain-neurons into few clusters. There is no exact answer on the question on what basis we should divide the neurons and the problem can be solved only experimentally. For solving these clustering problems is involved Artificial Intelligence and Artificial neural network [23]. There exists several alternatives for pattern analysis and recognition [24], but Artificial Intelligence and Artificial Neural Network provides provide very effective and useful algorithms for pattern recognition [25].

3 EXPERIMENTAL METHODS AND MEASUREMENTS

The experimental bran-computer interface setup for communication with smart devices was worked out at the University of Telecommunications and Post at Sofia.

The experimental BCI system includes 3D camera Panasonic HDC-Z0000, sender Spectrum DX9 DSMX, Sony GoPRO – GoPro HERO3, Nikon D902D smart TV Samsung UE-65HU8500 + LG60LA620S, ACER K11 Led projector, Linksys EA6900 AC1900 smart router, Pololu Zumo Shield, 8 core/32GB RAM/4TB HDD/3GB VGA computer for video processing that translate EEG signals into computer commands.

Two Electro-Caps (elastic electrode caps) was used to record each from positions C3, C4, P3, P4, O1, and O2, defined by the most popular "10-20 System of electrode placement" at experimental setup.

The implemented standard system of electrode placement defines a grid of most appropriate places for brain signals measurements on the human head.

The electrode locations are defined by either 10% or 20% increments between these specific positions. The possible placements of electrodes on the human scalp is presented on Figure 1.
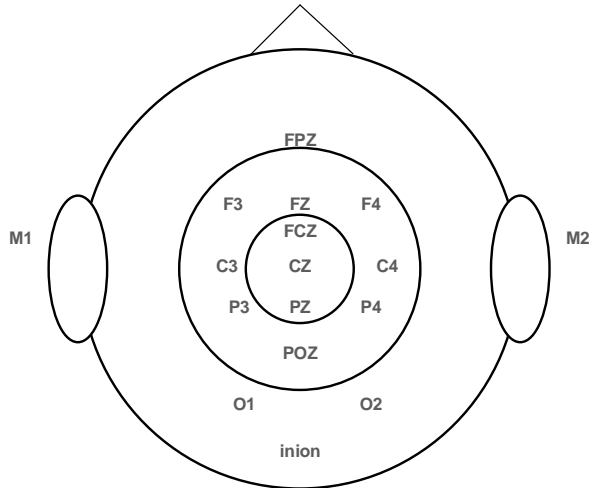

Figure 1. Electrodes placements model in experimental setup

Measurements were provided for the regions P3, P4 and Cz. The used device also includes a ground electrode connection that we attached to an ear lobe to provide electrical protection.

The electrodes were connected through a bank of Grass amplifiers. Data was recorded at a sampling rate of 250 Hz with a Lab Master 12 bit A/D converter mounted in computer.

Eye blinks were detected by means of a separate channel of data recorded from two electrodes placed above and below left eye. Data was recorded for 14 seconds during each task and each task was repeated 10 times per session.

After ensuring that participants are ready, the EEG electrodes are put on their scalp. The five cognitive tasks were explained and the participants have to perform them several times to ensure that they understood the tasks.

Participants at experiments performed the following tasks within experiments with their eyes closed:

- calm and relax task;

- letter - emergency call -subjects dial up 122;

- math task - imagined addition;

- counting task - count edges or planes of 3D graphics;

- imagine rotation of shown figure.

The recorded signals were passed through worked out filters realized by implementation of weights [26] and regression approach for modification of amplitude, which helps to remove noise from the original EEG signals.

During the measurements every task was repeated 20 times and register signal was divided into 200 sampling points. The duration of each record was 14 *sec*. and from each channel was received 4 000 samples.

4 DATA ANALYSIS AND CLASSIFICATION

Some basic signal processing of received time series was performed for classification of measured signals. The EEG signals were divided into small overlapping by 1 second intervals of 2 seconds.

For each of selected interval was computed the most significant features. These features was used for training a Bayesian Neural Network used for classification. After 10 000 training epochs the neural network is capable to provide classification of brain-neurons into clusters in real-time implementation.

The spectral power of the signal was investigated into a set of six standard frequency bands, according to different types of neural activity.

The frequency content of the processed signals was received after Fourier transform.

For comparison was provided classification of brain-neurons into clusters with pair-wise classifier.

Table 1 presents received results of the classification accuracies for each of 6 subjects, where N is subject number and M is mean value.

Table 1. Classification accuracies with Bayesian Network classifiers for five mental tasks

N	Mental Tasks				
	Base	Letter	Math	Count	Rotate
1.	91.3%	65.3%	73.7%	69.4%	72.5%
2.	92.4%	74.8%	76.5%	79.9%	69.2%
3.	87.8%	79.2%	62.6%	78.6%	81.1%
4.	90.2%	70.3%	69.2%	69.4%	75.3%
5.	93.7%	63.5%	75.3%	76.4%	66.1%
6.	89.3%	69.7%	73.7%	75.9%	79.5%
M	90.8%	70.46%	71.83%	74.93%	73.9%

In this investigation was used 18-fold cross validation. The classification accuracies with pair-wise classifier for the same five mental tasks is shown in Table 2.

Table 2. Classification accuracies with pair-wise classifier for five mental tasks

N	Mental Tasks				
	Base	Letter	Math	Count	Rotate
1	93.5%	67.1%	76.9%	69.3%	74.5%
2	92.4%	71.6%	78.3%	79.9%	63.2%
3	87.8%	70.8%	64.6%	78.6%	81.1%
4	90.2%	69.4%	63.8%	62.4%	79.3%
5	93.7%	67.9%	75.1%	72.5%	65.8%
6	89.3%	61.7%	82.3%	71.6%	73.3%
M	90.8%	68.08%	73.5%	72.38%	72.9%

The classification accuracies depending on the user fall into the interval between 62.6% and 81.1%.

5 EEG EMOTION RECOGNITION

Developed task-oriented BCI includes EEG emotion estimation combined with facial emotion analysis.
The technique for EEG analysis and feature extraction includes the following steps: first step is a preprocessing of EEG signals, the second step is built an augmented feature vector from statistical and time-frequency features, and the third step is classification, which consists of SVM core with probabilistic outputs.

For accuracy improvement in terms of facial muscle movement influence on system performance at EEG emotion recognition was made the following assumptions: the muscle movement is an additive noise and its influence can be reduced by filtering. The accuracy of the modality with better performing classifier can be used as a reliable output for the combined system output.

Prior to features extraction, the baseline drift in EEG and signals has to be removed. For this purpose was used mathematical morphology to get a filtered EEG signal x according to:

$$x = x' - \frac{(x' \circ SE) \bullet SE + (x' \bullet SE) \circ SE}{2} \quad (1)$$

where x' is the raw input signal, SE is a structuring element of type horizontal line with length $l = 500$, o denotes morphological opening and • denotes morphological closing. The length l meets the condition:

$$l = \frac{f_S}{2f_b}, \quad (2)$$

where $f_b = 0,1\ Hz$ is the lowest frequency in the spectrum of EEG signals and $f_s = 100\ Hz$ is the sampling frequency.

After decomposition of the signal into separate components $h_k[n]$ with defined instantaneous frequency:

$$x[n] = \sum_{k=1}^{K} h_k[n] + r[n], \quad (3)$$

where K is the number of intrinsic mode functions (IMF) and $r[n]$ is the residual. The decomposition is iterative procedure and it repeats until the residual becomes a monotonic function. Since the IMFs of an EEG signal represent the activity of the well-known waves (delta, theta, alpha and beta), we have

restricted the available IMFs up to 4 plus the residual. For each kth IMF was calculated its Hilbert transform $H_k[n]$ and the instantaneous frequency:

$$f_k[n] = \frac{f_S}{2\pi}\left(\theta_k[n] - \theta_k[n] - 1\right) \qquad (4)$$

The instantaneous frequency determined according to (4) appeared to be unstable, so for the final experiments was used Savitzky–Golay filter with polynomial order of 4 for $f_k[n]$ calculation. The window length for computation is chosen to be 41. The IMFs are not perfect analytic signals and we encountered the problem with negative values in $f_k[n]$. An IMF can be regarded as combination of amplitude and narrow-band frequency modulation. Suppressing the amplitude modulation part significantly reduces the appearance of negative frequencies. In a fixed number of iterations we normalize the IMF by its Hilbert envelope and finally the amount of negative frequencies is negligible. For a particular frequency band, we extract a signal $m[n]$ as follows:

$$m[n] = \begin{cases} \max\left\{\left|H_k[n]\right|\right\} & f_l \leq f_k[n] \leq f_h \\ 0 & \text{otherwise} \end{cases} \qquad (5)$$

where f_l and f_h are the lowest and the highest frequency in the band. The relative duration of a given EEG wave activity is found according to:

$$d = \frac{1}{N}\sum_{n=1}^{N} m'[n] \,, \qquad (6)$$

where $m'[n]$ is:

$$m'[n] = \begin{cases} 1 & V_l \leq m[n] \leq V_h \\ 0 & \text{otherwise} \end{cases} \qquad (7)$$

In (7) V_l and V_h are the lower and the higher voltages of a given wave activity. The upper threshold is used to avoid the possible artifacts in the signal.

The proposed model relies on a new approach for multi-view facial expression recognition, where was encoded appearance-based facial features using sparse codes and learn projections from non-frontal to frontal views using linear regression projection. We then reconstruct facial features from the projected sparse codes using a common global dictionary. Finally, the reconstructed features are used for facial expression recognition.

For detection of the frontal face and improved accuracy at face detection was used sub-window with the detected face through a Convolutional Neural Network.

At the process of face extraction it is necessary to detect and recognize a few face samples within the proposed detection interval. Afterwards was calculated an average feature vector which is estimated in the classifier for each window.

With the help of common spatial patterns (CSP) and linear SVM (support vector machines or support vector networks) was classified two emotional state – sad and happy. The received results show that Gamma band (30–100 Hz) is the most appropriate for EEG-based emotion classification.

After learning with supervised learning model and training with 6 000 epochs the classification process can be used in real-time for classification, where all samples are divided into two categories. The achieved classification accuracy is 70.78%.

6 CONCLUSIONS

An approach for human-computer interaction with classification of electrophysiological signals, recorded with brain-computer interface at different mental tasks is presented. With considered experimental setup of BCI were provided experiments with six subjects for execution of five mental tasks. The measured outputs after noise filtering were

classified with Bayesian Neural Network classifier and with pair-wise classifier.

Developed task-oriented BCI includes EEG emotion estimation combined with facial emotion analysis based on the following steps: preprocessing of EEG signals, built augmented feature vector from statistical and time-frequency features, and classification which consists of SVM core with probabilistic outputs. The received results are compared and shown.

7 ACKNOWLEDGMENTS

REFERENCES

[1] S. Coyle, T. Ward, and C Markham, "Brain – computer interfaces: A review," Interdisciplinary Science Reviews, vol. 28, February 2003, pp. 112 – 118.

[2] B. Z Allison, E. W. Wolpaw, and J. R. Wolpaw, "Brain – computer interface systems: Progress and prospects," Expert Rev. Med. Devices, vol. 4, Jul. 2007, pp. 463 – 474.

[3] F. Galán, M. Nuttin, E. Lew, P. Ferrez, G. Vanacker, J. Philips, and J. Millán, "A brain-actuated wheelchair: Asynchronous and non-invasive brain-computer interfaces for continuous control of robots," Clinical Neurophysiology, vol. 119, Sept. 2008, pp. 2159-2169.

[4] J. R. Wolpaw, "Brain–computer interfaces as new brain output pathways," Journal of Physiology, vol. 579, Mar. 2007, pp. 613-619.

[5] K. Forster, A. Biasiucci, R. Chavarriaga, J. Millán, D. Roggen, and G. Troster, "On the use of brain decoded signals for online user adaptive recognition systems," Proceedings of the Conference on Pervasive Computing, 2010, Online http://dx.doi.org/10.1016/j.clinph.2006.09.003.

[6] S. H. Fairclough, "Fundamentals of physiological computing," Interacting with computers, vol. 21, no. 1-2, 2009, pp. 133 –145.

[7] R. Palaniappan, "Brain – computer interface design using band powers extracted during mental tasks," Proc. IEEE EMBS Conference on Neural Engineering, 2005, pp.321–324.

[8] D. M. Taylor, S. I. Tillery, and A. B. Schwartz, "Direct cortical control of 3D neuroprosthetic devices," Science, vol. 296, Feb. 2002, pp. 829–832.

[9] J. Fogarty, A. J. Ko, H. H. Aung, E. Golden, K. P. Tang and S. E. Hudson, "Examining task engagement in sensor- based statistical models of human interrupt ability," Proc of SIGCHI Conference on Human Factors in Computing Systems, 2005, pp. 331 - 340.

[10] F. Azevedo, L. Carvalho, L. Grinberg, J. Farfel, R. Ferretti, R. Leite, W. Filho, and R. Lent, "Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain," Journal of Comparative Neurology vol. 513 May 2009, pp. 532–541.

[11] S. P. Fizgibbon, K. J. Pope, L. Mackenzie, C. R. Clark and J. O. Willoughby, "Cognitive tasks augment gamma EEG power," Clinical Neurophysiology, vol. 115, 2004, pp. 1802 - 1809..

[12] E. Iánez, J. M. Azorín, A. Úbeda, J. Ferrández, and E. Fernández, "Mental tasks-based brain-robot interface," Robotics and Autonomous Systems, vol. 58, Dec. 2010, pp.1238 - 1245.

[13] P. Sajda, E.Pohlmeyer, J. Wang, L. Parra, C. Christoforou, J. Dmochowski, B. Hanna, C. Bahlmann, M. Singh, and S. F. Chang, "In a Blink of an Eye and a Switch of a transistor: Cortically Coupled Computer Vision," Proceedings of the IEEE, vol. 98, Mar. 2010, pp. 462-478.

[14] T. Ros, M. A. Munneke, D. Ruge, J. H. Gruzelier, and J. C.Rothwell, "Endogenous control of waking brain rhythms induces neuroplasticity in humans," European Journal of Neuroscience vol. 31, Issue 4, Feb. 2010, pp. 770–778.

[15] A. Jaimes and N. Sebe, "Multimodal human-computer interaction: A survey," Computer Vision and Image Understanding, vol. 108, no. 1-2, 2007, pp. 116–134.

[16] L. Loke, A. T. Larssen, T. Robertson, and J. Edwards, "Understanding movement for interaction design: frameworks and approaches," Personal and Ubiquitous Computing, vol. 11, Aug. 2007, pp. 691–701.

[17] J. Millán, P. Ferrez, F. Galán, E. Lew, and R. Chavarriaga. 2007. "Non-invasive brain-machine interaction," in International Journal of Pattern

Recognition and Artificial Intelligence, vol. 20, Oct. pp. 1 - 13.

[18] Mu Li and Bao-Liang Lu, "Emotion Classification Based on Gamma-band EEG", Proceedings of the IEEE Conf. on Engineering in Medicine and Biology, Minneapolis, 3-6 Sept. 2009, pp. 1223-1226.

[19] Noppadon Jatupaiboon, Setha Panngum, and Pasin Israsena, "Real-Time EEG-Based Happiness Detection System", Scientific World Journal, 2013; 2013: Article ID 618649, 12 pages http://dx.doi.org/10.1155/2013/618649

[20] Wei-Long Zheng, Bo-Nan Dong and Bao-Liang Lu, "Multimodal Emotion Recognition using EEG and Eye Tracking Data", Proc. of the IEEE EMBC Conf. Chicago, 26-30 Aug., 2014, pp. 5040-5043.

[21] M.Schreuder, A. Riccio, M .Risetti, S. Dähne, A. Ramsay, J. Williamson, D. Mattia, and M. Tangermann. 2013. "User-centered design in brain-computer interfaces-a case study," in Artificial Intelligence in Medicine, vol. 59, pp. 71–80.

[22] J. J.Shih, D. J. Krusienski, and J. R. Wolpaw. 2012. "Brain-Computer Interfaces in Medicine," in Mayo Clinic Proceedings, vol. 87, pp. 268–279.

[23] J. Wolpaw and E. W. Wolpaw, "Brain-Computer Interfaces: Principles and Practice," Oxford University Press,. 2012.

[24] S. Kumar and M. Sharma, "BCI: Next Generation for HCI," Int. Jour. of Advanced Research in Computer Science and Software Engineering, vol. 2 Issue 3, Mar. 2012 pp.146 -151.

[25] J. Millán, "A local neural classifier for the recognition of EEG patterns associated to mental tasks," IEEE Transactions on Neural Networks, vol. 13, 2002, pp.222 – 226.

[26] P. Poolman, R. M. Frank, P. Luu, S. Pederson, and D. M. Tucker. 2008. "A single-trial analytic framework for EEG analysis and its application to target detection and classification," in NeuroImage, vol. 42, pp. 787-798.

Ultrasonic sensors gloves for blind people using Lilypad Arduino

Abdulrahman Alkandari
Computer department
College of Basic Education
(PAAET)
Kuwait
aam.alkandari@paaet.edu.kw

Areej Alajmi
Computer department
College of Basic
Education
(PAAET)
Kuwait
Areejal3jmii@gmail.com

Sarah Muhammad
Computer department
College of Basic
Education (PAAET)
Kuwait
Omjana1991@gmail.com

Altaf Alshammari
Om mabad middle school
Ministry of education
Kuwait
altaf.alshamri@hotmail.com

Abstract – A lots of blind people suffer in their own lives because of their vision loss. Vision is one of the five important senses in the human body. People with Vision loss have their own disability. Many countries around the world provide special assistance to these needed people to improve their life quality as good as possible. They provide them with special equipment for their disability to improve their daily life like voice message service, electronic stick that guide them while moving around and other specialized equipment. This paper present a project idea to establish and provide ultrasonic gloves to blind people for guiding them to their right roads without the need for other people assistant. This Can be done through Ultrasound waves that will be sent to the surrounding then it will be collected by detector in the gloves then to be sent as vibration or Audio signals to the blind's so they can be aware of their surroundings and they can choose their own road and way without other people assistant

Keywords: Ultrasonic sensor, Pit sensor, Global Positioning System (GPS) ,Programmable Interface controllers (PIC) and Lilypad Arduino .

1.   INTRODUCTION

Suffering From blindness is not temporary for certain time it's being blind the whole day the whole time every second and every minutes. Once the blind person wakes up from his bed in the morning his suffering start and his daily needs start. Blind people need more care to avoid risk of injuries and that affect people around them; people need to be near for them to avoid being injured. People around them will be exhausted from being attention to them and giving them all what they need. So Blind people must depend on themselves,

this paper propose gloves which helps Blind people to depend on themselves.

Blindness can be caused by physiological dysfunction, anatomical or neurological dysfunctions. Each blind person has his own blindness type. Many scores have been established to assess the extent of blindness [1]. Total blindness is the complete absence of vision in the person mainly neuronal blindness because of defect in the NLP.

Some researchers believe that blindness may positively affect the audio capability in the blind person, So their sense of hearing will be more accurate. So they can distinguish sounds and musical tones. Scientifics define people with small percentages of blindness is the person who have score of 200/20 at the best eye. Also the person who use eyeglasses or who have shortness of vision or have decrease in the visual fields less than 20 degree. Therefore, many countries are helping the development of researches and studies seeking to improve the lives of blind people with special needs to make them feel they are part of the surrounding community through associations and different courses that are giving to learn the mechanism of dealing with blindness and ignore the inability of weakness and inability to move around [1].

2.   PROBLEM STATEMENT

Around the world, Blind people face real problems. Their major problem lies in their sense of disability and sees what around them. Also they face a problem with other people hesitate in their service or assistance to them. In addition, many

people do not know how to deal with blind person. Blind people needs a boost of confidence in their self and their ability to achieve anything and rely on their self without the needs of others [3].

Many countries developed many mechanisms of services that contribute to make life easier for the blind, where they developed some rules and laws in the art of dealing with the blind without notifying disability and some companies and organizations are training and recruiting people with the ability to deal with blind. In addition, there are a lot of companies have developed special equipment used by the blind to help them [2].

There are a lot of emotions a blind person hides behind his eyelids that we cannot see them.
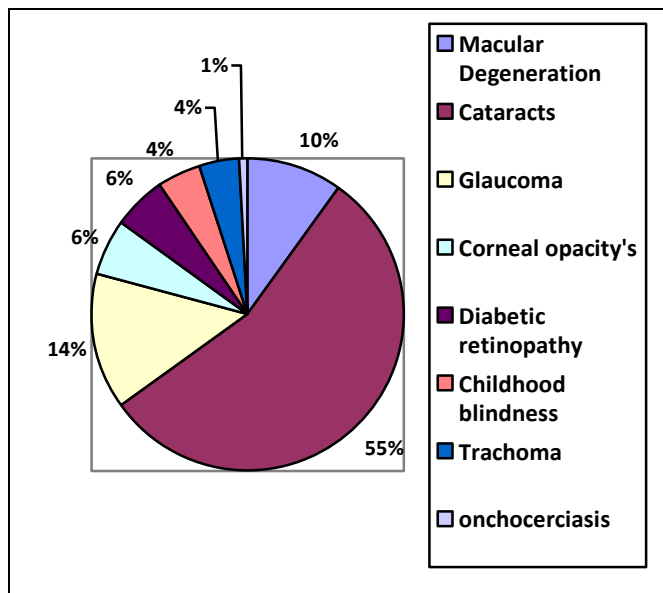

Figure 1. Chart of Percent of types of blindness

Blindness can be divided into several types [4]:
1. Partial blindness.
2. Blindness as a result of the incident.
3. Color blindness.
4. Psychological blindness.

Therefore, the prevalence of different types of blindness is difficult to assess because of the retinal differences between people this led to a lot of scientists to try to find different ways to solve this crisis. And this suffering as blindness varies

Many of them feel embarrassed to talk about their suffering in front of others or feel upset at their inability to see because of their illness or old age or the events that he suffered and caused him to lose sight. Some diseases that cause weakness in sight are listed (Fig.1), according to the World Health Organization in 2012 for the most common reasons of blindness (except for refractive errors):
1. Cataracts (47.9%),
2. Glaucoma (12.3%),
3. Macular Degeneration (8.7%),
4. Corneal opacities (5.1%),
5. Diabetic retinopathy (4.8%),
6. Childhood blindness (3.9%),
7. Trachoma (3.6%)
8. Onchocerciasis (0.8%) [4].

by categories as mentioned some of them is suffering from blindness only in old age. Some of them is infected since the births and other are blind in addition to deafness and inability to listen even for high voices. Since the Industrial Revolution in Europe still people trying to find and solve this problem. Some of blind people hires someone to be his assistant and assist him in his movement. In addition to using stick to navigate or use trained dogs for this purpose.

Thus, it is do not hide the extent of the suffering blind face behind their inability to see or the suffering of people around them. Many of those who help people blind feel bothers with the blind people and their special needs. Therefore, they should be find new style to help blind and rely on their self without resorting to other people [2].

3. HISTORY OF BLINDS AND STICK

Since ancient times, man used the stick and leaned on it. Also it used by a blind person in avoiding obstacles while walking and movement. The evolution of the idea of the white cane. In 1921 the process of James Biggs, who was living in Bristol County, England stick to become a new form of white form, after he was blinded in an accident wiped his eyes.

He has to live with his new environment, and he was horrified passers-by from tankers steam around his home.

Aftly white cane, he pay attention to the tanker drivers, especially at night and were keen to help him. And it became more pronounced stick to them, and felt the difference is clear. Here the White cane idea was started [1].

In America 1930, George Lyon developed the shape to make the lower end of the stick red color to be distinctive.

In 1931 d'Herbemont way of walking stick taught and learned in France. Then this event was spread in British newspapers, which helped to spread the idea.

In May 1931 BBC radio, broadcast White cane code for each world who are, and must be used by all blind. After that most of the blind rehabilitation centers embraced the idea and learn how to walk out of a lot of blind people in the world [1].

4.   RELATED WORKS

Chun-Ming Tsai explains Shopping Card interactive. The idea of interactive card love (Love Card) is to integrate slice of identification using radio waves (RFID) in the electronic card attached to the neck of the blind, the reaction after this card with existing products in the market and to pronounce the name of a blind product and its price [5].

Hyun Ju Lee, Mun Suk Choi and Chung Sei Rhee are considered that Braille credit card ATM cards whether credit or other are non-usable by blind category. The proposed solutions work credit card points Braille prominent with built-in audio system, so that the blind case to make a purchase and when you pass the card to pay your amount withdrawn and introduced braille pronunciation [6].

Wang Guolu, Qiu Kaijin, Xu hai and Chen Yao explained that Phone screen for the Blind Can blind interaction with mobile phones panel dedicated in Braille touch called (B-Touch). This phone contains the advanced capabilities, such as the existence of a system of direct and reader of books and a system to identify the objects, and the touch screen is formed in Braille by the content displayed [7].

Nguyen, Kennedy, and Cornwall demonstrate demonstrate that Color recognition audio voice is one of the problems facing the blind specifically in defining colors of objects. The touch alone is not enough to determine the color of the body, so the proposed technical solutions is a device (Bright-F), which specifies the color and depth and then translate them into spoken words [8].

Balachandran, Cecelja and Ptasinski were explained GPS for the blind. Perhaps the best technical developments in the field to help the blind was the launch of «GPS» navigation devices which you can see the blind site and notified acoustically close to the intersection and tell him about the restaurants or the nearby bus from sites. However, this technique cannot reach a high accuracy in determining the user's location, which is important for the Blind, where the closer the user to place about 3040 meters [9].

Chaurasia and Kavitha's paper was about Electronic stick. And it is an electronic rod designed as a long white stick for exports are really offering a blind frequency ultrasound feel them under his hand when he hit a snag in a certain way. They are able to detect obstacles in all directions at a distance of five meters. The made parties to this stick of lead material [10].

5.   HARDWARE REQUIREMENTS

Any system, application or device requires some of available equipment to work efficiently and properly, so that this system requires a selection equipment and taken into consideration that every part, which will provide to the system must be an occasion for the service to be useful and activity for the meant category [6].

System contains several tools , including Ultrasonic sensor which works as input data source ,vibration motor which work as output reaction ,microcontroller called Lilypad arduino board that programming by c++ language, fabric gloves worn by the user and superconductor fibers which connect all hardware parts together [6].
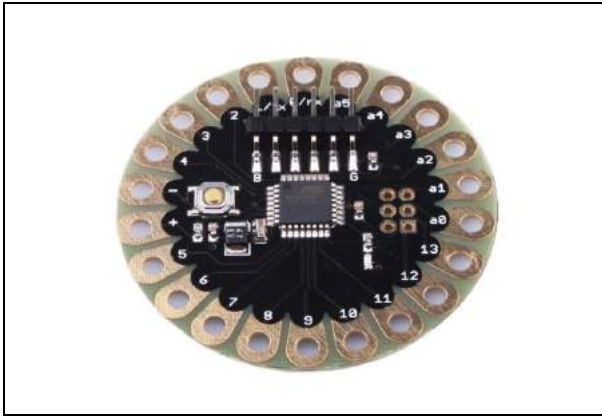
.Figure 2 LilyPad Arduino Board

Figure.2 shows LilyPad Arduino Board. Arduino is an intelligent microcontroller, which can sense the environment around it by receiving inputs from various sensors then affects surroundings by controlling motors, lights and actuators [5].

It is an open source board used to run electronic parts, it featuring with low cost as well as freedom Add electronic parts which connect with Arduino board through USB to connect it to a computer to program it by C++ programming language. Moreover, it can be linked with any other electronics part using electric Breadboard and has a port of power and needs a simple electrical energy can saved it by using normal batteries [7].

There are a lot of Arduino types that according vary to their functions and sizes , in blind gloves was selected Lilypad Arduino which featuring with small size, light weight and flat shape so that it can be stitched above the gloves.
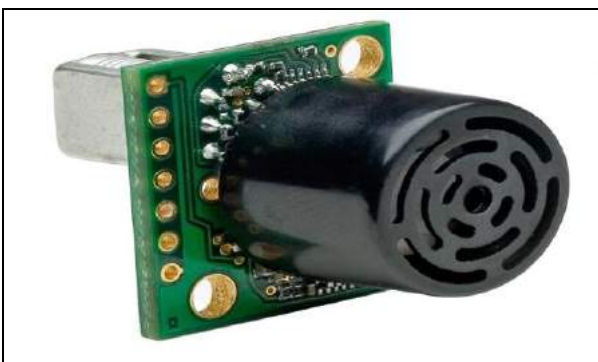

Figure 3 Ultrasonic sensor

Ultrasonic sensor work on a principle similar to radar systems, it convert high that intuit objects presence and it distance from the sensor via generate high-frequency waves around it. Then compute the time taken between sending waves and receiving its echo, that shown in figure.3.

Ultrasound waves longer such as IR that reflecting on the surface of the body out, but considered it performance higher and better because of the optical dispersion. Ultrasound has the best collection of readers where proved effective in a lot of robots and sensors industries, as well as a weak negative impact on the human body comparing with the infrared. In addition, ultrasound can measure the distance and angles where it is reliable and gives precise results and the margin of error is negligible, so they are suitable for the Blind and Visually Impaired [5].

The controller is a single chip containing a simple processor performs inputs processing and issued it in the form of output through (CPU). It contain non-volatile memory to process the results received in it work as an act (RAM). In addition, it contains a Read Only Memory (ROM) of the input and output (I / O) to control the time and distance, which is designed to perform and to control certain system preprogrammed.

In this system, the input, which is here the waves it is processed and analyzed estimate the time. It takes to issue Special Sensor Microwave ultrasound, the distance traveled, the time and the extent of return echo waveform them to issue the results of the analysis in the controller to be sent to the e-vibration installed on the gloves [7].

Figure 4 the vibrator device of the project

Figure.4 shows Vibration motor. It is a small electronic device gives a simple vibration to alert blind people to objects presence around him to avoid the collision, it is work with 5 voltage.

Figure 5 Electrically conductive wires

The connection between all hardware parts was be via simple electrically conductive wires (Fig.5) which sewing carefully above a fabric gloves.

Figure 6 fabric gloves

All hardware parts are stitching and installing above fabric gloves that wore by users, it shown in Figure 6.

6. IMPLEMENTATION

The project consists of right and left gloves worn by the user. To begin the gloves activity user need to direct his hand to know if there was handicap objects around through feeling vibration from gloves. Gloves alarm when any object is on its side. In addition, it alarm when any object is in front of the user.
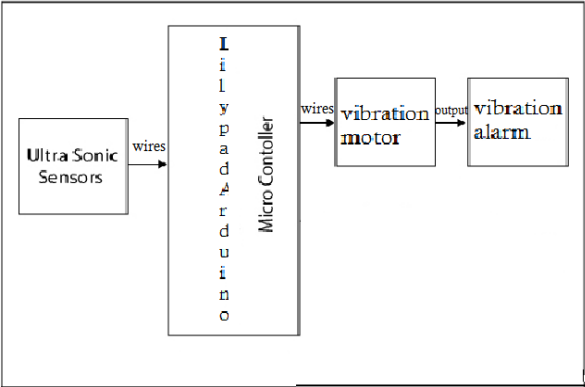

Figure 7. System Block diagram

The work of electronic parts distributed in figure.7. Ultrasonic sensor works as input device which sends data of objects around to Lilypad arduino microcontroller for processing and sentencing which respond should give Lilypad Arduino command vibration motor, which work as output device to respond through, gives vibration alarm.
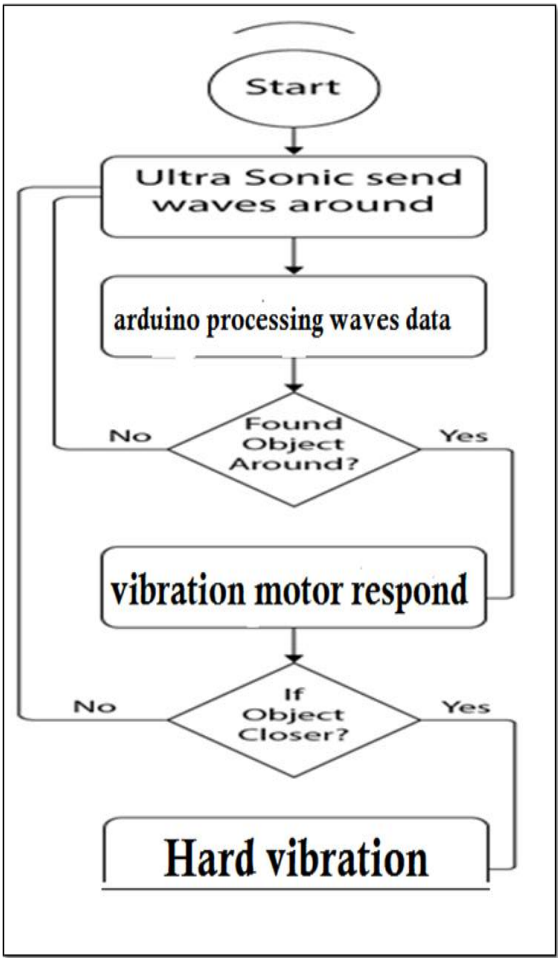

Figure 8 data flow diagram of the project

20

As shown in figure 8, the proposed device is divided into two main parts:

- Ultrasonic sensor (input device)
- Vibration motor (output device)

Ultrasonic sensor exports inputs data about objects distance around it to a Lilypad Arduino to process this data then commanding vibration motors to give output in form of vibration in accordance with programming of Lilypad Arduino.

As shown in Figure.8, Lilypad Arduino will give vibration motor order to alarm with continuous vibration when the object is close from the user. When the object keep away "few" from the user, the Lilypad Arduino will give vibration motor order to alarm with intermittent vibration. When the object keep away "more" from the user, the Lilypad Arduino will order a vibration motor to give not vibration. Ultrasonic will back to send waves to find objects. Then repeat the process [9].
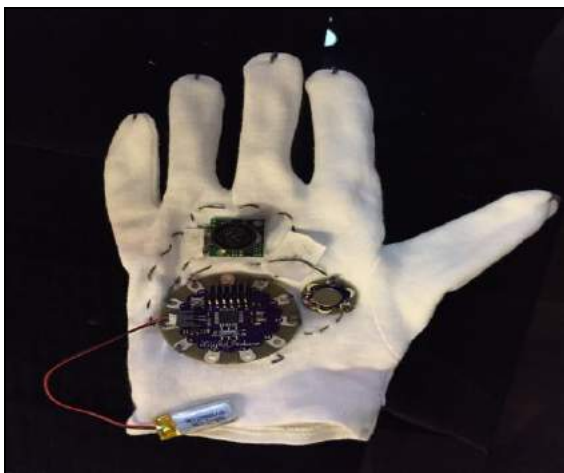
7. RESULTS

Figure 9. The gloves with the parts

Figure 9 shows the hardware parts that works out this proposed system implemented above the glove with the finally shape.

Many experiments dedicated to blind gloves was conducted to know the quality of performance expected of them. It have been tested on more than one way, to make the visually impaired more capable of dealing with the surrounding environment and negotiate with them

properly. As mentioned earlier, the circles sensor gives information about objects surroundings and the environment around the user and alerting through vibration, where it made about the accuracy of 1 meter, from the user's real surrounding areas.

The overall result of the experiment gloves, came as follows:

i. The highest response was whenever it is closer to the object.
ii. Middle respond was found in middle distance from the object.
iii. No respond where found on far distance.

Table.1 contains numerical minutes of the experience on the gloves and it shows the results in cm unity:

Table 1 Test Results of the Gloves

Obstacle	Test
From 0 to 30 cm	Hard vibration
From 30 to 60 cm	Intermittently vibration
From 60 cm to 100 cm	Intermittently vibration
From 100 cm and up	Vibration not work

Befits of using the blinds gloves

Wearing blind glove can make the user feel more secure from of the difficult things facing him. The blind person walks without knowing what are the hidden risks in the next step? It may be a hole or ladder or wall in front of him, the gloves here have a significant role in securing the next step for him.

Also to feel blind person that he independent; he can go anywhere without waiting for someone facilities. That making him feel a degree of independence and self-reliance. On the other hand, it is easy to use and that because of the simplicity in its size.

In addition, it will be develop more models, which comes with the modern models.

8. CONCLUSION

This paper proposed Ultrasonic sensors gloves for blind people using Lilypad Arduino. That it is a solution for the visually impaired and those who do not wish to carry on his stick give them the

instructions of places and surrounding objects. The blind will be able to move from side to side and from one place to another without the need to help others to learn about the highs and surrounding objects.

This system would be more setups sophisticated and assistant for the visually impaired around the world, and so high for his efficiency. In addition to the low-cost of manufacture it, and according to the accuracy of the sensor.

In spite of this, it is recommended to restructure the pieces added or commercial quality to its requirements to lead properly. In the future, some improvements will be improved to the system in order to meet approbation of users such as:

- Increasing other type of sensors as pulse and temperature sensors.
- Increasing navigation system and an internal memory to store all places and coordinates experienced by the user
- Adding automatic synchronization between the gloves and mobile phones to be identified on the site by a person other relatives or theirs.
- Adding voice guidance property to make it easier to stay away from the process of risk if the vibration were unavailing.

REFERENCES

[1] Mohd Helmy Abd Wahab, Amirul A. Talib, HerdawatieA.Kadir, Ayob Johari, A.Noraziah, Roslina M. Sidek, Ariffin A. "Smart cane: assistive cane for visuallyimpaired people", IJCSI, Vol.8 Issue 4, July 2011.

[2] Yuan, D.; Manduchi, R.; "Dynamic environmentexploration using a virtual white cane", in ComputerVision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference.

[3] F. van der Heijden, P.P.L. Regtien, "Wearable navigation assistance - a tool for the blind" MEASUREMENTSCIENCE REVIEW, Volume 5, Section 2, 2005

[4] Sylvain Cardin, Daniel Thalmann and Frederic Vexo, "Wearable Obstacle Detection System for visually impaired People" In VR workshop on haptic and tactile perception of deformable objects (No. VRLAB-CONF-2005-019, pp. p-50).

[5] Chun-Ming Tsai Machine Learning and Cybernetics (ICMLC), 2012 International Conference on Volume: 5, DOI:10.1109/ICMLC.2012. 6359666, Publication, Year: 2012, Page(s): 1901- 1906.

[6] Ju Lee, Hyun, Mun Suk Choi, and Chung Sei Rhee. "Traceability of double spending in secure electronic cash system." In Computer Networks and Mobile Computing, 2003. ICCNMC 2003. 2003 International Conference on, pp. 330-333. IEEE, 2003.

[7] Guolu, Wang, Qiu Kaijin, and Chen Yao. "The design and implementation of a gravity sensor-based mobile phone for the blind." In Software Engineering and Service Science (ICSESS), 2013 4th IEEE International Conference on, pp. 570-574. IEEE, 2013.

[8] THONG NGUYEN, D., Hamish Kennedy, and Jeff Cornwall. "A vocalized color recognition system for the blind." IEEE transactions on instrumentation and measurement 33, no. 2 (1984): 122-126.

[9] Balachandran, Wamadeva, Franjo Cecelja, and Piotr Ptasinski. "A GPS based navigation aid for the blind." In Applied Electromagnetics and Communications, 2003. ICECom 2003. 17th International Conference on, pp. 34-36. IEEE, 2003.

[10] Chaurasia, Shashank, and K. V. N. Kavitha. "An electronic walking stick for blinds." In Information Communication and Embedded Systems (ICICES), 2014 International Conference on, pp. 1-5. IEEE, 2014.

Evaluation of Performance of Smart Devices in Closed System Models

Cecilia Twinamatsiko[1],   Kyanda Swaib Kaawaase[2],   Michael Okopa[3]
Makerere University[1,2], International University of East Africa[3]
Kampala, Uganda
hopececilia1@gmail.com, kswaibk@cis.mak.ac.ug, okopamike@gmail.com

ABSTRACT

Smartphones and tablets are becoming dominant devices in the present market. Apart from offering the same features as traditional mobile phones, smart devices provide several other features as well. Designing Smartphones to balance the tradeoff between minimum power dissipation and increased performance becomes a challenge. This is due to the fact that users try to optimize for mobility and lightness rather than for maximum computing performance and throughput. In an effort to study the performance of smart devices, we modeled the components of smart devices using an $M/M/1//N$ closed queue system. In this model, the number of requests in the system is fixed. The performance of requests under the closed and open system models are compared in terms of average response time. We observe that average response time generally increases with increase in arrival rate, and increase in load values. It is further observed that requests experience lower average response time under the open system model at low arrival rates, low service times and low load compared to closed system model. However, at high arrival rates, high service rates and high load values, requests experience lower response time under the closed system model than under the open system model. The analytical results indicate that closed system models offer better performance at high arrival rates and high load values, whereas open system models offer better performance at low arrival rates and low load values.

KEYWORDS

Closed system model, open system model, smart devices, embedded systems, workload generators

1   INTRODUCTION

Despite the market's heterogeneity, smart devices, wireless broadband, and network-based cloud computing constitutes a perfect storm of opportunity for application developers, luring their attention towards the new platforms [1]. Smart devices are becoming dominant devices in the present market. Apart from offering the same features as traditional devices, smart devices provide several other features. Some of these features include amongst others high speed access to Internet using Wi-Fi and mobile data network, running several applications & games, capturing & sharing pictures to social platforms, audio-video capturing and sharing [2].

The performance limitation of smart devices originate mainly from the fact that they have embedded systems [3]. An embedded system is a computer system made for specific control functions inside a larger system frequently having real-time computing constraints.

Normally, embedded systems have different design constraints than desktop computing applications. Instead embedded systems try to optimize for mobility and lightness rather than for maximum computing performance and throughput. Since many users want their smartphones to be more powerful, have high throughput, and at the same time, small and light, it presents a constraint to designers [4]. This constraint acts as a recipe to analyze the factors that affect system performance, for example, number of processor cores, clock frequency, processor types, load, service rate, arrival rate and so on.

System researchers are aware of the importance of representing a system accurately and this involves accurately representing the scheduling of requests, service distributions, arrival of requests into systems, etc. Representing systems accurately involves many things, including accurately representing the scheduling of requests, service distributions, correlation between requests, etc. One factor that researchers pay little attention to is whether the job arrivals obey a closed or an open system model.

In a closed system model, new job arrivals are only triggered by job completions [3]. In addition, in a closed system model, it is assumed that there is some fixed number of users, who use the system forever. These users are called the multiprogramming level (MPL) [3, 5, 6]. Examples of closed systems include TPC-W [8] used as a database benchmark for e-commerce, RUBiS [9] which is an Auction website benchmark, AuthMark [10] which is a Web authentication and authorization benchmark, etc.

On the other hand, in an open system model, new jobs arrive independent of job completions [11, 12]. In an open system model there is a stream of arriving users and each user is assumed to submit one job to the system, wait to receive the response, and then leave. The differentiating feature of an open system is that a request completion does not trigger new requests, instead a new request is only triggered by a new user arriving.

The differences between open and closed system models motivate the need for researchers to investigate the performance of devices in either open or closed systems. This background therefore provides a simple recipe for investigating the performance of smart devices in closed system models.

2 RELATED WORK

Schroeder *et al.* [3] studied the performance of closed versus open workload generators under many applications, like static and dynamic web servers, database servers, auctioning sites, and supercomputing centers. The study noted that there was a huge difference between the two systems. For example, under a fixed load, the mean response time for an open system model can exceed that for a closed system model by a high magnitude. In addition, while scheduling to favour short jobs is extremely effective in reducing response times in an open system, it has very little effect in a closed system model. The deduction is that variability in the job sizes has a higher effect in an open system than a closed one.

Bondi et al [14] studied a general network of FCFS queues and concluded that service variability is more dominant in open systems and less pronounced in closed systems (provided the MPL is not too large). However, this study was primarily restricted to FCFS queues.

In a similar study, Schatte et al. [15, 16] studied a single FCFS queue in a closed loop with think time. The study revealed that, as the MPL grows to infinity, the closed system converges monotonically to an open system. This result provides a fundamental understanding of the effect of the MPL parameter.

M. Choi *et al.* [1] developed a queuing model to analyze the performance of the components of Smart devices in terms of average waiting time and mean queue length using an open system model. The components analyzed include; Central processing Unit (CPU), RAM, GSM, LCD, and Graphics. The authors assumed an open queuing system model. Although many applications can be modeled using an open system model, open system models do not capture interactivity in systems which is well captured by closed system models. For example, interactive behaviors like human users interacting with a system, threads contending for a lock, and networked servers waiting for a response message are so common. In a closed network, it is possible to model a set of users submitting requests to a system, waiting for results, and

then submitting more requests.

The expressions for average response time and number of requests in the system respectively are derived in [1] and given below:

Average response time, $E[S]$ is given as:

$$E[S] = \frac{1}{\mu(1-\rho)} = \frac{\rho}{\lambda(1-\rho)} \qquad (1)$$

Similarly, the mean number of requests, $E[X]$ is given as:

$$E[S] = \frac{\rho}{(1-\rho)} = \frac{\lambda}{(\mu-\lambda)} \qquad (2)$$

The differences between open and closed system models shown across a range of applications motivates the need for system designers to be able to determine how to choose if an open or closed model is more appropriate for the system they are targeting. Although many applications can be modeled using an open system model, open system models do not capture interactivity in systems which is well captured by closed system models.

In this paper, we model the performance of Smart devices in a closed system in terms of average response time and mean queue length so as to capture the interactivity.

The contributions of this paper are two fold. Firstly, we developed closed system model using queuing theory to analyze the performance of Smart devices in terms of average response time and mean queue length. Secondly, this research evaluated the performance of the closed system models while comparing it to the open system models.

The rest of the paper is organized as follows: in the next section, we present the system models. We evaluate the proposed models in section 4 and finally conclude the paper in section 5.

3 SYSTEM MODELS

We consider a closed system model, where new job arrivals are only triggered by job completions as shown in figure 1. In a closed

system model, new job arrivals are only triggered by job completions.
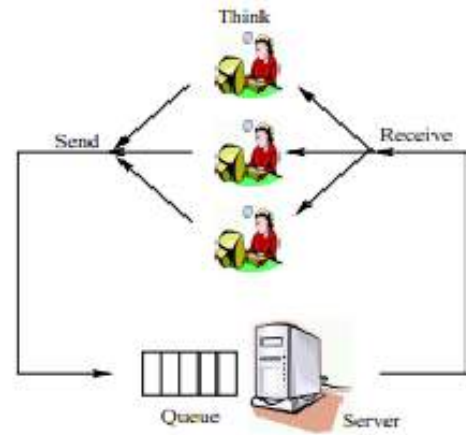

Figure 1. Closed system. Adopted from [3]

In a closed network, we can model a set of users submitting requests to a system, waiting for results, then submitting more requests. Examples of closed systems include TPC-W [8] used as a database benchmark for e-commerce, RUBiS [9] which is an Auction website benchmark, AuthMark [10] which is a Web authentication and authorization benchmark, etc. In addition, we consider the following;

(a) a smart phone with the following components; CPU, RAM, graphics, GSM, LCD, Backlight, and Rest.

(b) we assume that the throughput is known (to be equal to the arrival rate), and also that there is no probability of incomplete transfer in this system, so there is no retrial path to go back.

(c) Although the CPU components of recent smartphones can have more than one CPU, known as dual-core or quad-core, however we assume a smartphone with single-core in this study.

In the next section, we derive the expression for average response time and average queue

length.

3.1 Expression for average response time and average queue length

Specifically, we use an $M/M/1//N$ queue system, where the first M denotes Poisson distribution which best models random arrivals into systems. The *pdf* (probability density function) of a Poisson distribution is given by [13]:

$$P(x = k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad (3)$$

The second M represents an exponentially distributed service time. The *pdf* (probability density function) of an exponential distribution is given as [13]: $f(x) = \mu e^{-\mu x}, \quad x \geq 0, \ \mu \geq 0$, where μ is the mean service rate. One server and FCFS service discipline are assumed for the system model, while at the same time the population of the system is finite.

We used the tagged job approach where we track the experience a job undergoes, i.e, when a new customer arrives to the queue, there will be some number of other customers already present and waiting, including a customer that is in service.
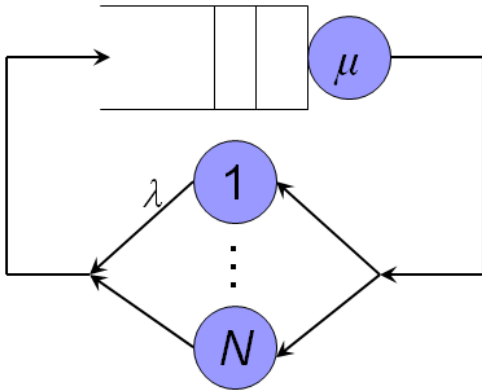

Figure 2. $M/M/1//N$ closed queue system

Figure 2 shows an $M/M/1//N$ closed queue system. The corresponding state transition diagram is given in figure 3.
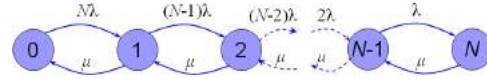

Figure 3. State transition diagram for $M/M/1//N$ closed queue system

Using the birth-death result we can obtain individual probabilities as:

$$\mu\pi_1 = N\lambda\pi_o \qquad (4)$$

$$\mu\pi_2 = (N-1)\lambda\pi_1 \qquad (5)$$

$$\mu\pi_3 = (N-2)\lambda\pi_2 \qquad (6)$$

Combining equations 4, 5, and 6 we obtain;

$$\pi_3 = N(N-1)(N-2)(\frac{\lambda}{\mu})^3 \qquad (7)$$

Equation 7 can be generalized to any probability state i as:

$$\pi_j = \frac{N!}{(N-j)!} \cdot \rho^j \cdot \pi_o, j = 1, 2, ..., N. \qquad (8)$$

Using the fact that total probability is one we obtain;

$$\sum_{j=1}^{N} \frac{N!}{(N-j)!} \cdot \rho^j \pi_o = 1 \qquad (9)$$

and hence,

$$\pi_o = \left[\sum_{j=0}^{N} \frac{N!}{(N-j)!} \cdot \rho^j \right]^{-1} \qquad (10)$$

Using little's law [13], the average queue length, $E[X]$ is given by;

$$E[X] = \mu(1 - \pi_o)E[S] \qquad (11)$$

where $E[S]$ is the average response time. For the thinking part of the closed system,

26

$E[N - X] = \mu(1 - \pi_o).\frac{1}{\lambda}.$

$E[N] - E[X] = \mu(1 - \pi_o).\frac{1}{\lambda}$, from which we obtain $E[X] = N - \mu(1 - \pi_o).\frac{1}{\lambda}$.

Using equation 11 we obtain;

$$E[S] = \frac{E[X]}{\mu(1 - \pi_o)} \qquad (12)$$

Therefore, the average response time is given as;

$$E[S] = \frac{N - \mu(1 - \pi_o).\frac{1}{\lambda}}{\mu(1 - \pi_o)} = \frac{N}{\mu(1 - \pi_o)} - \frac{1}{\lambda} \qquad (13)$$

where π_o is as given in equation 10.

Similarly, the average number of requests in the system can be deduced from little's theorem as follows: $E[X] = \lambda E[S]$. Dividing equation 13 by λ we obtain the expression for the average queue length of requests in the system is given as:

$$E[X] = \left(\frac{N.\rho}{(1 - \pi_o)} - 1\right), \qquad (14)$$

where π_o is as given in equation 10.

In the next section, we present the performance evaluation of closed against open system models.

4 PERFORMANCE EVALUATION

In this section, we use the derived models to evaluate its performance. In particular, we analyze the variation of average response time and average queue length with arrival rate of requests, load in the system, and service rate. In each case, we compare the performance under closed and open system models.

For numerical evaluation we considered the following hypothetical data, these data is the same as the configurations used in [1]. Average arrival rate used is 1000 to 1500 requests/second, average service rate is 1500 to 1800 trequests/second, load used is in the range 0 to 1.

4.1 Comparison of open and closed systems in terms of average response time

In this section, we compare the performance of requests under open and closed system models in terms of average response time. We used equations 1 and 13 to plot the following graphs.
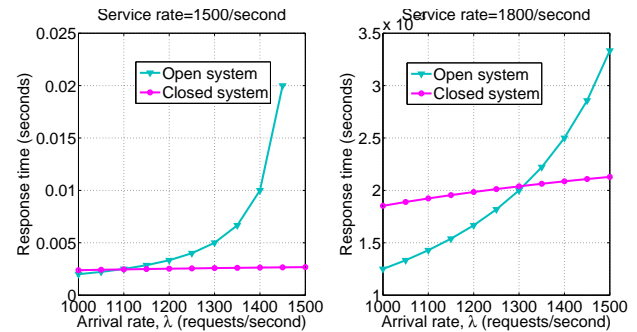

Figure 4. Response time against arrival rate.

Figure 4 shows the variation of average response time against arrival rate of requests into the system. We observe that average response time increases with increase in arrival rate regardless of the service rate. We further observe that requests experience lower response time under the open system at low arrival rates, however at high arrival rates, requests experience lower response time under closed systems. The difference in performance of requests between open and closed systems is more pronounced at high service rate compared to low service rate. This is shown when service rate of 1500 requests/second is used compared to 1800 requests/second. Whereas the performance of requests under open and closed systems are the same at arrival rate of 1100 requests/second when the service rate is 1500 requests/second, it is the same at arrival rate of 1300 requests/second when the service rate is 1800 requests/second.

Figure 5 shows the graph of average response time against load for different service rates. We observe that average response time increases with increase in load in the system regardless of the service rate. We also note that requests experience lower average response time under the open system at low load, however when the
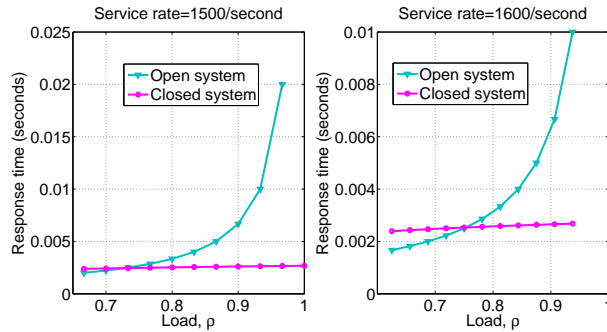
Figure 5. Response time against load.

load increases, requests experience lower average response time under the closed system model. The difference in performance between open and closed system is more pronounced at high service rate compared to low service rate.
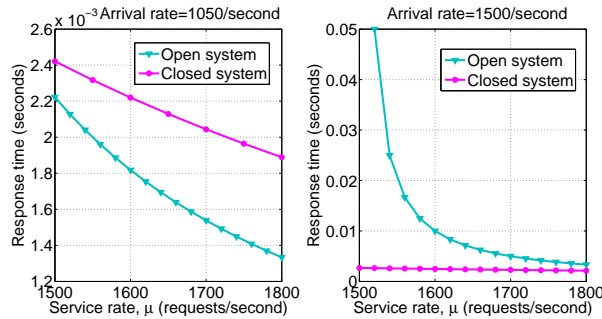

Figure 6. Response time gainst service rate.

Figure 6 shows the graph of response time against service rate for different arrival rates. We observe from the graph that response time decreases with increase in service time. In other words, as the service rate increases, the response time decreases. Increase in service rate implies increase in the number of requests served per unit time, hence decrease in average response time. We also observe that requests experience lower response time under the open system at low service rate, however at high service rate, requests experience lower response

time under the closed system compared to under the open system.

4.2 Comparison of open and closed system models in terms of average queue length

In this section, we compare the performance of requests under open and closed system models in terms of average queue length. We used equations 2 and 14 to plot the following graphs.
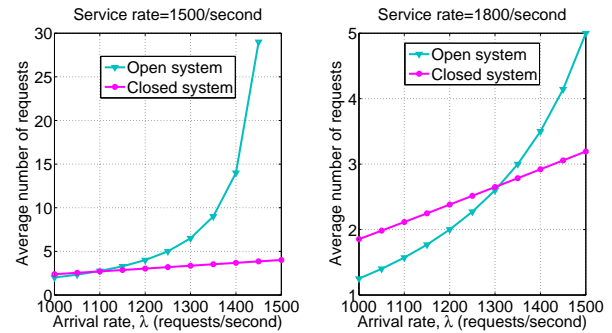

Figure 7. Average queue length against arrival rate.

Figure 7 shows the variation of average queue length against arrival rate of requests into the system. We observe that average queue length increases with increase in arrival rate regardless of the service rate. We further observe that the number of requests in the system under the open system model is lower than under the closed system model at lower arrival rates. However, at higher arrival rates, the number of requests in the system under the open system model is higher than under the closed system model. Whereas the performance of requests under open and closed systems are the same at arrival rate of 1100 requests/second when the service rate is 1500 requests/second, it is the same at arrival rate of 1300 requests/second when the service rate is 1800 requests/second. Figure 8 shows the graph of average queue length in the system against the load. We observe that the number of requests in
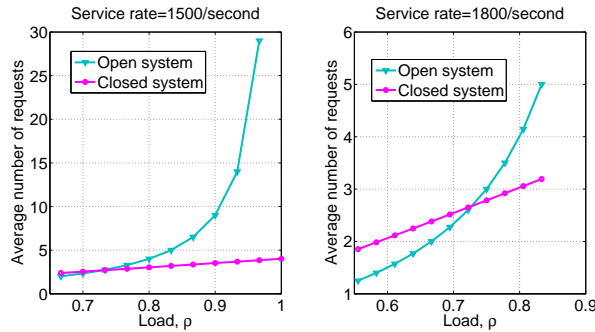
Figure 8. Average queue length against load.

the system increases with increase in load regardless of the service rate. We further observe that the number of requests under the open system model is lower than under the closed system model at lower load values. On the other hand, the number of requests under the open system model is higher than under the closed system model at higher load values. Figure
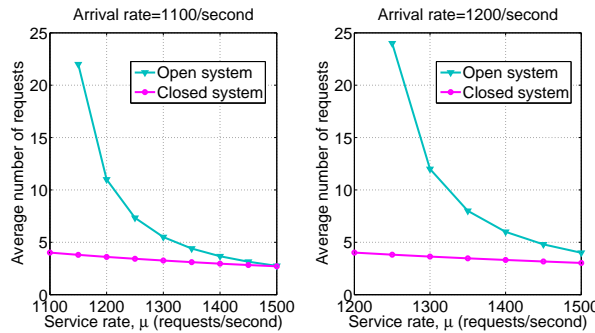

Figure 9. Average queue length against service rate.

9 shows the graph of average queue length in the system against the service rate for different arrival rates. We observe from the graph that the average queue length decreases with increase in service time. We also observe that the average queue length in the system under the open system is generally higher than under the closed system model.

4.3 Performance under the closed system model

In this section, we investigate the effect of

varying parameters on the performance of requests under the closed system model. We used equation 13 and equation 14 to plot a graph of average response time and average queue length against arrival rate, service rate, and load while varying the number of requests in the system.
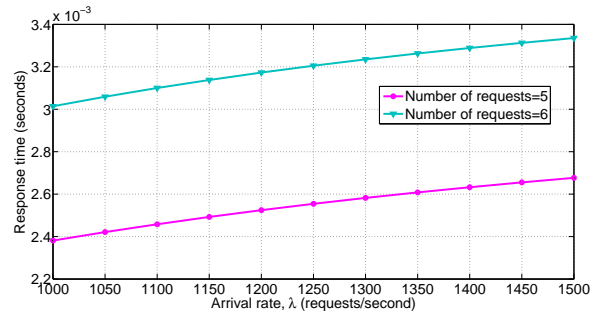

Figure 10. Response time against arrival rate.

We observe from figure 10 that average response time increases with increase in arrival rate. We further observe that the average response time is higher when the number of requests in the system is higher as compared to when the number of requests in the system is lower. For example, when the arrival rate of requests into the system is 1300 requests/second, the average response time when the number of requests in the system is 5 is approximately $2.6 * 10^{-3}$ seconds while when the number of requests in the system is 6, its approximately $3.2 * 10^{-3}$ seconds. The increase in average response time as arrival rate increases is due to the fact that increase in arrival rate leads to increase in the number of requests in the system and therefore an increase in average response time.

We observe from figure 11 that the average response time of requests generally decreases with increase in service time. We further observe that average response time is lower when the number of requests in the system is lower compared to when it is high. When the arrival rate is high, the average response time is lower. On the other hand, when the average arrival rate is low, the average response time is lower.
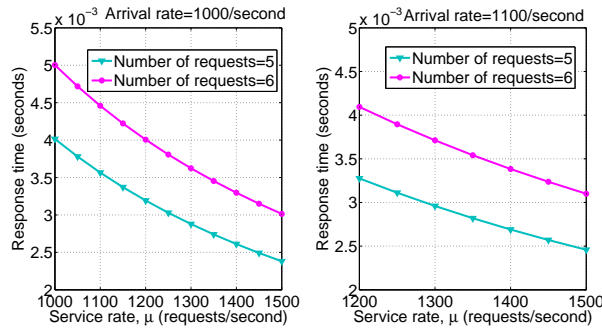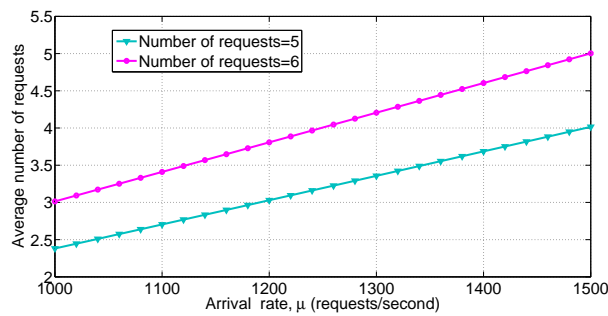
Figure 11. Response time gainst service rate.

Figure 12. Average queue length against arrival rate.

We observe from figure 12 that the average queue length increases as the arrival rate increases. In addition, the average queue length is higher when the number of requests in the system is high compared to when it is low. For example, when the arrival rate is 1200 requests/second, and the number of requests in the system is 5, the average queue length is observed to be 3. On the other hand, when the number of requests in the system is 6, the average queue length is 3.8 requests.
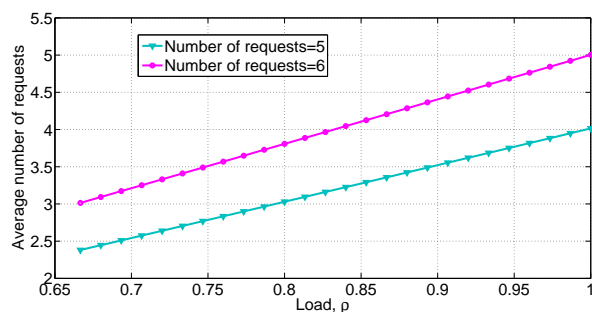

Figure 13. Average queue length against load.

We observe from figure 13 that average queue length generally increases with increase in load. We further observe that the average queue length is higher when the number of requests in the system is high compared to when it is low. For example when the load is 0.9, and the number of requests in the system is 5, the average queue length is 3.5 requests. On the other hand, when the load is 0.9, and the number of requests in the system is 6, the average queue length is 4.4 requests.

We observe from figure 14 that the average queue length decreases with increase in service rate regardless of the number of requests in the system. We also observe that the average queue length is lower when the number of requests in the system is low compared to when it is high. This is noted when the number of requests in the system is 5 compared to when it is 6. For example, when the service rate is 1200 requests/second, the average queue length when the number of requests in the system is 5 is 3.6 requests whereas when the number of requests in the system is 6, the average queue length is 4.5 requests.
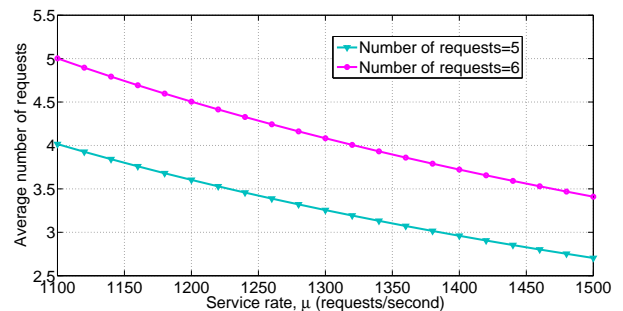

Figure 14. Average number of requests against service rate

5 CASE STUDY FROM FRANCE FOOTBALL WORLD CUP 1998 TRACE

In this section, we conducted performance test on smartphones.

We used trace data from Internet sites serving the 1998 World Cup. The data used are available from the Internet Traffic Archive (see [17] and http://ita.ee.lbl.gov/html/traces.html). This repository provides detailed information about the 1.3 billion requests received by World Cup sites over 92 daysfrom April 26, 1998, to July 26, 1998. We assume that all requests submitted must first pass through the Internet sites for providing HTTP service before moving on to the smart devices.

The data trace indicated in [17] exhibits arrival-rate fluctuations corresponding to light, medium and heavy loadings in this temporal order. More specifically, the trace allows us to study performance under various load conditions, as follows:

- Light load. In the time interval [0, 1200], the arrival rate is relatively low (below 1200 requests/s), and the resultant utilization is also low (40%).

- Medium load.In the time interval (1200, 2400], the arrival rate is between 1200 and 1800 requests/s, and the resultant utilization is intermediate (50%).

- Heavy load. In the time interval (2400, 3600], the arrival rate often exceeds 2000 requests/s, and the corresponding utilization is high (75%).

We use two performance metric to test the performance of smart devices, that is, average response time and mean queue length.

5.1 Comparison of open and closed systems in terms of average response time

In this section, we compare the performance of open and closed systems in terms of average response time using trace data from Internet sites serving the 1998 World Cup.

5.1.1 Performance under light load

Under light load, we consider request average arrival rate of below 1200 requests/s and

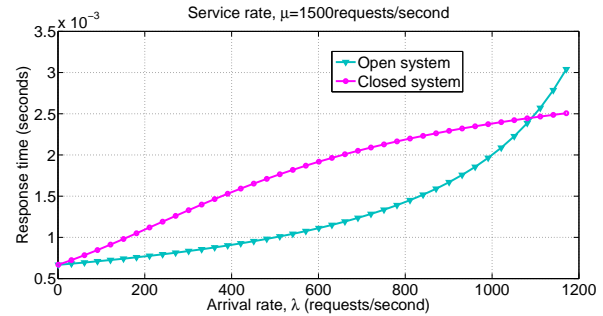service rate of 1500 requests/s. Figure 15

Figure 15. Response time against arrival rate.

shows the variation of average response time against arrival rate of requests into the system when the load is light. We observe that average response time generally increases with increase in arrival rate. We note that requests experience lower average response time under open systems at lower arrival rates, however at higher arrival rates, requests experience lower response time under closed systems. The difference in performance of requests between open and closed systems is more pronounced at arrival rate of between 400 requests/second and 900 requests/s. On the other hand, the performance of requests under open and closed systems are the same when the arrival rate is approximately 1100 requests/s.

5.2 Performance under medium load

Under medium load, we consider request average arrival rate of between 1200 and 1800 requests/s and service rate of 1800 requests/s to ensure the arrival rate is less than service rate ($\lambda < \mu$), otherwise the queue grows infinitely long. Figure 16 shows the variation of average response time against arrival rate of requests into the system under medium load. We observe that average response time generally increases with increase in arrival rate. We further observe that requests experience lower average response time under open systems at lower arrival rates, however at higher arrival rates, requests experience lower response time under
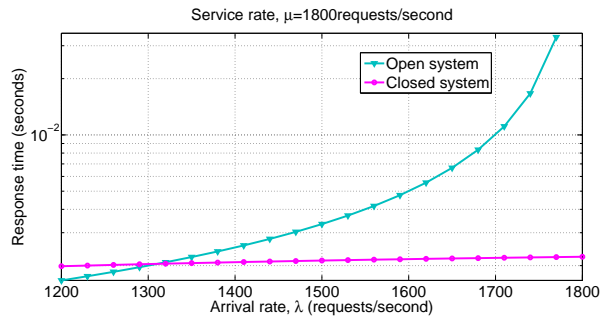
Figure 16. Response time against arrival rate.

closed systems. The difference in performance of requests between open and closed systems is more pronounced at higher arrival rates. It is also observed that the performance of requests under open and closed systems are the same when the arrival rate is approximately 1300 requests/s, after the arrival rate of 1300 requests/s, requests perform better under closed systems.

5.3 Performance under heavy load

Under heavy load, we consider request average arrival rate of between 2000 and 9000 requests/s and service rate of 9000 requests/s to ensure the arrival rate is less than service rate ($\lambda < \mu$), otherwise the queue grows infinitely long. Figure 17 shows the variation of average
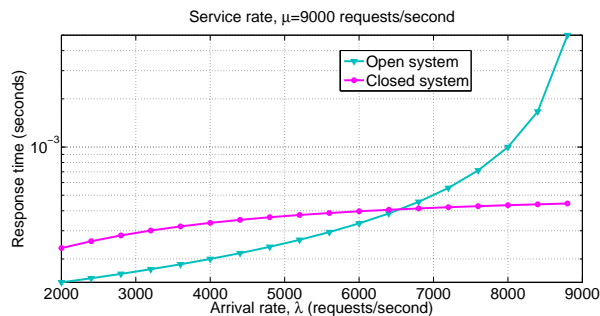

Figure 17. Response time against arrival rate.

response time against arrival rate of requests into the system under heavy load. We observe that average response time generally increases with increase in arrival rate.

We further observe that requests experience lower average response time under open systems at lower arrival rates, however at higher arrival rates, requests experience lower response time under closed systems. It is further observed that the performance of requests under open and closed systems are the same when the arrival rate is approximately 6500 requests/s, after the arrival rate of 6500 requests/s, requests perform better under closed systems.

6 CONCLUSION

In an effort to study the performance of smart devices, we modeled the components of smart devices using an $M/M/1//N$ closed queue system. In this model, the number of requests in the system is fixed. The performance of requests under the open and closed system models are compared.

We observe that average response time generally increases with increase in arrival rate, and increase in load values. The average response time however, decreases with increase in the service rate. It is further observed that requests experience lower average response time under the open system model at low arrival rates, low service times and low load. However, at high arrival rates, high service rates and high load values, requests experience lower response time under the closed system model.

A similar trend is observed with the average queue length which generally increases with increase in arrival rate, and increase with increase in load values. The average queue length however decreases with increase in the service rate. It is further observed that requests experience lower average queue length under the open system model at low arrival rate, low service rate and low load. However, at high arrival rate, high service rate and high load values, requests experience lower average queue length. Therefore, closed system models is observed to offer better performance at high arrival rates and high load values than open

system models.

These results are also confirmed when trace data from Internet sites serving the 1998 World Cup was used. The analytical results indicate that closed system models offer better performance at high arrival rates and high load values, whereas open system models offer better performance at low arrival rates and low load values.

ACKNOWLEDGMENT

REFERENCES

[1] M. Choi, "Power and Performance Analysis of Smart Devices," in Proc. International Journal of Smart Home, vol. 7, no. 3, May 2013.

[2] "http://www.gartner.com/newsroom/id/1924314 ," Feb. 2012, press release.

[3] B. Schroeder, A. Wierman, and M. Harchol-Balter, "Open Versus Closed: A Cautionary Tale," in Proc. the 3rd conference on Networked Systems Design & Implementation, vol. 3, 2006, pp. 8-18.

[4] J. Kro, A. Brunnert, C. Prehofer, T.A. Runkler, H. Krcmar, "Model-based Performance Evaluation of Large-Scale Smart Metering Architectures," ACM 2015, PP. 9-12.

[5] K. Kant, V. Tewari, and R. Iyer, "GEIST: Generator of e-commerce and internet server traffic," in Proc.International Symposium on Performance Analysis of Systems and Software, 2001.

[6] P. Barford and M. Crovella, "The surge traffic generator: Generating representative web workloads for network and server performance evaluation," in Proc.the ACM SIGMETRICS, 1998.

[7] E. Nahum, M. Rosu, S. Seshan, and J. Almeida, "The effects of wide-area conditions on www server performance," in Proc. ACM SIGMETRICS, 2001.

[8] "Transaction Processing Performance Council, TPC benchmark W(webcommerce)," Feb. 2002, number Revision 1.8.

[9] C.Amza, E.Cecchet, and W. Zwaenepoel, "Specification and implementation of dynamic website benchmarks," 2002.

[10] Mindcraft, "The authmark benchmark http://www.mindcraft.com/authmark/," Feb. 2002.

[11] M. Yuksel, B. Sikdar, K. S. Vastola, and B. Szymanski, "Workload generation for ns simulations of wide area networks and the internet," in Proc.Comm. Net. and Dist. Sys. Mod. and Sim, 2000.

[12] Z. Liu, N. Niclausse, and C. Jalpa-Villanueva., "Traffic model and per-formance evaluation of web servers," in Proc.Performance Evaluation, 2001, pp. 77-100.

[13] L. Kleinrock, Queueing Systems, Volume I& II. Computer Applications. John Wiley & Sons, 1976.

[14] A. B. Bondi and W. Whitt, "The influence of service-time variability in a closed network of queues," Proc. Performance Evaluation, 1986, PP. 219-234.

[15] P. Schatte, "On conditional busy periods in n/M/GI/1 queue," Proc. Mathematics Operations for Optimization, 1983.

[16] P. Schatte, "The M/GI/1 queue as limit of closed queueing systems," Proc. Mathematics Operations for Optimization, 1984.

[17] A. Martin and J. Tai, "Workload Characterization of the 1998 World Cup Web Site", IEEE, 2000. pp.30-37.

International Journal of
NEW COMPUTER ARCHITECTURES AND THEIR APPLICATIONS

The *International Journal of New Computer Architectures and Their Applications* aims to provide a forum for scientists, engineers, and practitioners to present their latest research results, ideas, developments and applications in the field of computer architectures, information technology, and mobile technologies. The IJNCAA is published four times a year and accepts three types of papers as follows:

1. **Research papers**: that are presenting and discussing the latest, and the most profound research results in the scope of IJNCAA. Papers should describe new contributions in the scope of IJNCAA and support claims of novelty with citations to the relevant literature.
2. **Technical papers**: that are establishing meaningful forum between practitioners and researchers with useful solutions in various fields of digital security and forensics. It includes all kinds of practical applications, which covers principles, projects, missions, techniques, tools, methods, processes etc.
3. **Review papers**: that are critically analyzing past and current research trends in the field.

Manuscripts submitted to IJNCAA **should not be previously published or be under review** by any other publication. Plagiarism is a serious academic offense and will not be tolerated in any sort! Any case of plagiarism would lead to life-time abundance of all authors for publishing in any of our journals or conferences.

Original unpublished manuscripts are solicited in the following areas including but not limited to:

- Computer Architectures
- Parallel and Distributed Systems
- Storage Management
- Microprocessors and Microsystems
- Communications Management
- Reliability
- VLSI