International Journal of

Digital Information and Wireless Communications

(IJDIWC)

SDIWC

www.sdiwc.net

Overview

The SDIWC International Journal of Digital Information and Wireless Communications is a refereed online journal designed to address the networking community from both academia and industry, to discuss recent advances in the broad and quickly-evolving fields of computer and communication networks, technology futures, national policies and standards and to highlight key issues, identify trends, and develop visions for the digital information domain.

In the field of Wireless communications; the topics include: Antenna systems and design, Channel Modeling and Propagation, Coding for Wireless Systems, Multiuser and Multiple Access Schemes, Optical Wireless Communications, Resource Allocation over Wireless Networks, Security, Authentication and Cryptography for Wireless Networks, Signal Processing Techniques and Tools, Software and Cognitive Radio, Wireless Traffic and Routing Ad-hoc networks, and Wireless system architectures and applications. As 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, IJDIWC 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 to the published materials.

Publisher

The Society of Digital Information and Wireless Communications
Miramar Tower, 132 Nathan Road, Tsim Sha Tsui, Kowloon, Hong Kong

Further Information

Website: http://sdiwc.net/ijdiwc, Email: ijdiwc@sdiwc.net,
Tel.: (202)-657-4603 - Inside USA; 001(202)-657-4603 - Outside USA.

Permissions

International Journal of Digital Information and Wireless Communications (IJDIWC) is an open access journal which means that all content is freely available without charge to the user or his/her institution. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles in this journal without asking prior permission from the publisher or the author. This is in accordance with the BOAI definition of open access.

Disclaimer

Statements of fact and opinion in the articles in the *International Journal of Digital Information and Wireless Communications (IJDIWC)* are those of the respective authors and contributors and not of the *International Journal of Digital Information and Wireless Communications (IJDIWC)* or *The Society of Digital Information and Wireless Communications (SDIWC)*. Neither *The Society of Digital Information and Wireless Communications* nor *International Journal of Digital Information and Wireless Communications (IJDIWC)* make any representation, express or implied, in respect of the accuracy of the material in this journal and cannot accept any legal responsibility or liability as to the errors or omissions that may be made. The reader should make his/her own evaluation as to the appropriateness or otherwise of any experimental technique described.

The issue date is October 2015.

TABLE OF CONTENTS
Original Articles

A Framework for Statistical Characterization of Indoor Data Traffic for Efficient Dynamic Spectrum Access in the 2.4 GHz ISM Band

Muhammad Khurram Ehsan and Dirk Dahlhaus
Communications Laboratory, University of Kassel, Germany
Email: khurram.ehsan@student.uni-kassel.de, dahlhaus@uni-kassel.de

ABSTRACT

The key for efficient dynamic spectrum access (DSA) is to model the spectral resources accurately. A large number of measurement campaigns have been performed to estimate the spectrum usage in outdoor and indoor scenarios. This spectrum usage estimation helps policy makers to optimize the spectrum management methodologies. The spectrum usage studies also assist researchers to constitute a way for efficient DSA using prior knowledge of the distribution of the observed data traffic in cognitive radio (CR) systems. In this paper we extend our previous work which statistically modeled the observed data traffic in the industrial, scientific and medical (ISM) band at 2.4 -GHz in two neighboring frequency subbands and time slots, respectively, to *three* neighboring frequency subbands and time slots, respectively. As before, the frequency and time correlation functions of the observed data traffic are modeled by an exponentially decaying function. The multivariate Gaussian mixture (MGM) is validated as a good candidate to model the joint distribution of measured data and also to estimate the correlation between the measured data in neighboring frequency subbands and as well as in time domain samples.

KEYWORDS

Wi-Spy Dual Band Spectrum Analyzer (WSDSA), Dynamic Spectrum Access (DSA), multivariate Gaussian mixture (MGM), k-means clustering, ISM data.

1 INTRODUCTION

It is a challenge for the industry as well as academia to optimize spectral resources in wireless communications. The numbers of customers are growing rapidly in the telecommunication industry due to innovative wireless technologies and high-speed data services. On the contrary, extensive measurement campaigns in different parts of the world [1] have claimed the under-utilization of the already allocated spectrum. The under-utilization has arisen due to the static allocation policies e.g. of the Federal Communications Commission (FCC). The advantages of static spectrum allocation policies are interference reduction and simple system hardware. On the other hand, the major disadvantage of these static policies is the inefficient utilization of radio spectral resources.

To address the issues between spectrum under-utilization and demand, cognitive radio (CR) technology stands as a promising candidate. CR provides the mechanism in which the available spatial and spectral resources are utilized by intelligently adapting parameters in the radio environment [2]. The switching between different operating frequencies, time duration to access specific frequency bands, variable operating frequencies and different spectrum ranges are representatives of intelligent adaptations and are collectively termed as dynamic spectrum access (DSA).

CR systems operate in three different ways, namely in so-called overlay, underlay and interweave modes. In an interweave system, secondary users (SUs) can access the spectrum in an opportunistic way only in the absence of a primary user (PU) and will stop transmission as soon as the PU resumes transmission. As a result, an efficient DSA with non-interfering SU signals arises from a correspondingly accurate statistical model for the existing data traffic observed in the radio environment. At the same time, the model should be simple enough to limit the complexity of resulting DSA approaches.

A lot of field measurement campaigns have been conducted in licensed (global system for mobile communication (GSM), universal mobile telecommunication system (UMTS)) and unlicensed (industrial, scientific and medical (ISM)) bands in both outdoor and indoor scenarios. The time-dependent spectrum occupancy using a four-state Markov model is discussed in [3]. The time-variant power spectrum at a CR receiver is measured in real-time and presented graphically in [4] and [5]. Interference temperature model based on real time measurements is discussed in [6]. Spectrum occupancy measurements for ultra high frequency (UHF) television (TV) bands are presented in [7]. The duty cycle models based on real time measurements are described for CR systems in [8] and [9]. In [10], real time measurement campaigns are conducted in different spectrum bands and also the average signal channel power in TV bands is modeled as a Gaussian random variable. The quantitative study of spectrum occupancy based on field measurements is presented in [11]. In [12], the average spectrum occupancies for UHF and very high frequency (VHF) bands are investigated. In [13], the authors assume that the signal power distribution of the PU in each subband is Gaussian. In [14], the autocorrelation function (ACF) of a wide-sense stationary (WSS) or short-range dependent process is modeled as a decaying exponential. The drawbacks in [3]-[11] are:

1. Simple occupancy matrices are used to estimate the data traffic in measured bands.
2. The data traffic is modeled using Markov chains following a Gaussian distribution under the assumption that the data traffic is independent and identically distributed (i.i.d) as well as unimodal, which clearly contradicts the properties of the data traffic observed in real-world scenarios.

In this paper, to model the data traffic measured in realistic scenarios more accurately, we have to deal with the correlation of the observed data traffic and the multimodality of the traffic data distribution. Here, we extend our previous work [15], which statistically modeled the observed data traffic in the ISM band at 2.4 GHz in two neighboring frequency subbands and time slots, respectively, to three neighboring frequency subbands and time slots, respectively. The observed data traffic in an indoor environment in the aforementioned ISM band is again modeled as a stochastic field in time and frequency. The process in time and frequency is assumed to be WSS with an autocorrelation function approximated by a decaying exponential. The multivariate Gaussian mixture (MGM) is considered as a suitable distribution model for the observed data traffic in this band. While there is no prioritization for the data traffic in the ISM band, we consider WLAN data traffic as a PU for simplicity.

The organization of the paper is as follows. The measurement scenarios and set-up are discussed in section 2. In section 3, the statistical characterization of ISM data traffic is detailed. The parametric modeling of the data traffic is described in section 4. In section 5, the MGM model is presented and validated using correlation functions. Conclusions are drawn in section 6.

2 MEASUREMET SCENARIOS AND SET-UP

2.1 Measurement Scenarios

Indoor measurements are conducted at the second floor of the engineering campus of the university of Kassel, Germany. The engineering campus is located in a residential area. We choose two different locations for conducting the measurement campaigns. At first, the measurements are taken in a computer laboratory. In order to analyze different user traffic profiles, measurements are also taken in an office room. At each location, measurements are taken for 8 hours per day for the duration of two weeks. In Figures 1 and 2, the snapshots of the measurements for an 8-hour day can be seen in an office room and a laboratory, respectively.
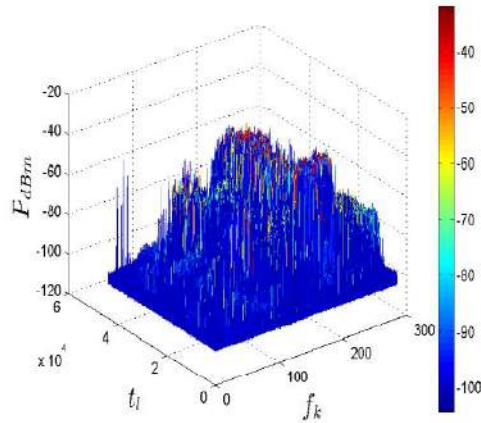
Figure 1. ISM data traffic measured in an office room.
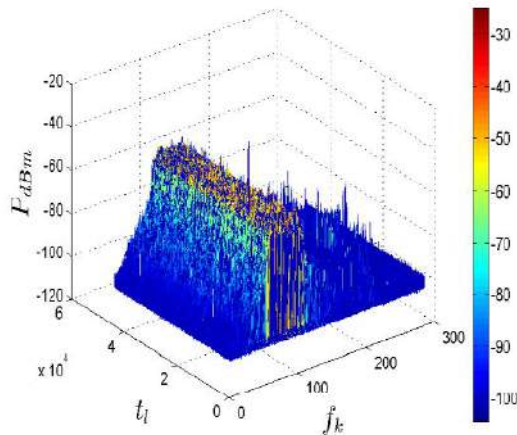
Figure 2. ISM data traffic measured in a computer lab.

Here, the traffic is characterized by the instantaneous power P in dBm (P_{dBm}).

2.2 The Testbed

A Wi-Spy Dual Band Spectrum Analyzer (WSDSA) with an omnidirectional antenna, as shown in Figure 3, is used to measure the ISM band activity.
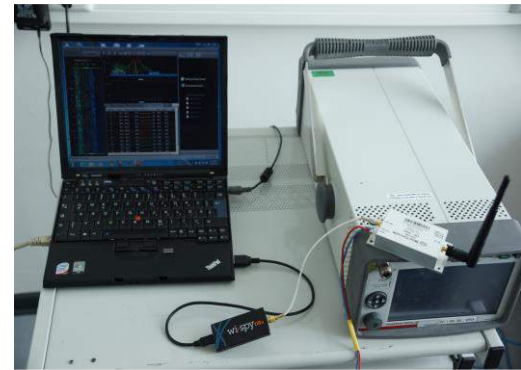
Figure 3. Testbed including WSDSA and omnidirectional antenna.

The term *activity* refers to the time-variant power spectral density (PSD) of the measured data, and it is this PSD as mentioned above that is considered to be traffic in the following. In this study, we only have the measurements at the 2.4 GHz band. Howeverm, the WSDSA can be used in both 2.4 GHz and 5 GHz bands to analyze the ISM data traffic.

The frequency span of the WSDSA is 83 MHz and it has a sweep time of 560 ms. A low noise amplifier (LNA) ZLR-3500+ is also introduced in the set-up, not only to have a better detection of weak signals, but also to improve the sensitivity of the measurement system. The frequency range to be scanned by the WSDSA is 2.4…2.483 GHz in our measurements. More of the measurement specifications are detailed in Table 1.

Table 1. Measurement Specifications

Parameters	Values
Frequency Range	2.4 GHz … 2.483 GHz
Frequency Span	83 MHz
Frequency Resolution	333 KHz
Measurement Duration	8 Hours
Sweep Time	560 ms
Preamplifier	21 dB

The testbed is interfaced with the Chanalyzer Pro® software that helps in the detailed visualization of the ISM data traffic observed by the WSDSA. The spectral view of the observed data given by Chanalyzer Pro® shows the spectrum usage over time which looks like a scrolling waterfall plot as shown in Figure 4.
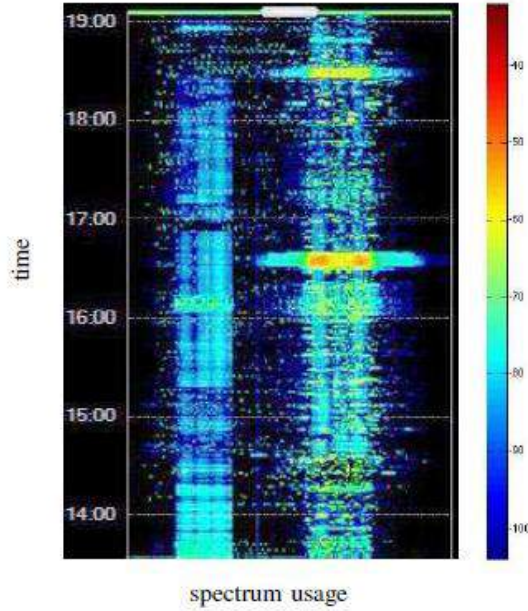
Figure 4. Waterfall plot showing spectrum usage over time.

2.3 Data Matrix

The real-time observed data traffic is saved in a comma-separated values (CSV) format by Chanalyzer Pro® to analyze and process it further. The observed data is then represented as a matrix where each row denotes the time instances having a resolution of $\Delta t = 560$ ms, while each column denotes the frequency subbands with a bandwidth of $\Delta f = $ 333 kHz. The obtained data matrix normalized to unit average power is represented as

$$\Xi = \begin{pmatrix} x(t_1, f_1) & \cdots & x(t_1, f_K) \\ \vdots & \ddots & \vdots \\ x(t_L, f_1) & \cdots & x(t_L, f_K) \end{pmatrix}, \qquad (1)$$

where $x(t_1, f_1)$ in the data matrix is recognized as an observation of the data traffic process random field $X(t_l, f_k)$ to denote the instantaneous power in dBm for each frequency subband $f_k = k\Delta f$, where $k = 1, 2, ..., K$ and the time instants are defined by $t_l = l\Delta t$ with $l = 1, ..., L$ and the upper limits $K = 285$ and $L = 1000$.

3 STATISTICAL CHARACTERIZATION OF ISM DATA TRAFFIC

It is assumed that the ISM data traffic $X(t_l, f_k)$ is a WSS process in both time and frequency. This is why the ACF of $X(t_l, f_k)$ is a function of the time and frequency differences, respectively, for a given value of the frequency and time, respectively.

3.1 Frequency Correlation Function

It is known that the frequency correlation function (FCF) for the ISM data traffic $X(t_l, f_k)$ which is dependent on Δk, is defined as

$$\phi_f(\Delta k) = E(X^*(t_l, f_k)X(t_l, f_{k+\Delta k})) \qquad (2)$$

where the estimate of $\phi_f(\Delta k)$ is defined by the empirical FCF $\hat{\phi}_f(\Delta k)$ using the measured ISM data traffic according to

$$\hat{\phi}_f(\Delta k) = \frac{1}{L(K - \Delta k)} \sum_{k=1}^{K-\Delta k} \sum_{l=1}^{L} x^*(t_l, f_k)x(t_l, f_{k+\Delta k}),$$
$$(3)$$

where $\Delta k = 0, ..., 284$. The FCF is modeled as a decaying exponential $\phi_f(\Delta k) = \rho_f^{\Delta k}$ due to the short-range dependency as described in [14]. Here, ρ_f represents the correlation between the frequency subbands with $\rho_f \in [0,1]$. A standard numerical method is employed to find the estimate $\hat{\phi}_f(\Delta k)$. The value of $\rho_f^{\Delta k}$ is computed for $r_f = 0, 0.01, 0.02, ..., 1$ and is chosen as the least squares estimate (LSE) between the modeled FCF $\phi_f(\Delta k)$ and the estimated FCF $\hat{\phi}_f(\Delta k)$. Both, the estimated FCF $\hat{\phi}_f(\Delta k)$ and the modeled FCF $\phi_f(\Delta k)$ are shown in Figure 5. The estimated value of $\hat{\rho}_{f,\text{LSE}}^{\Delta k}$ using the aforementioned approach is 0.56 when $\Delta k = 1$ and 0.31 for $\Delta k = 2$.
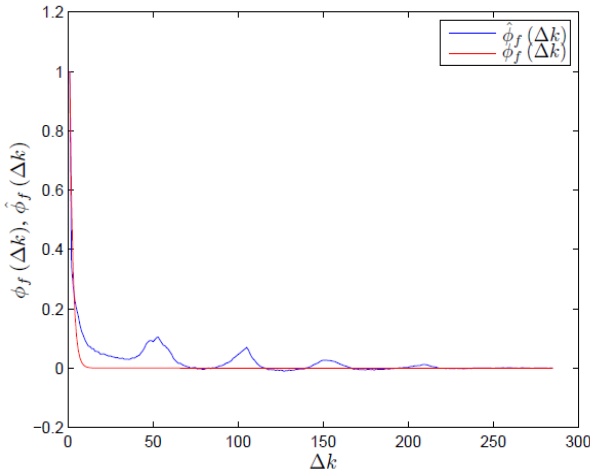
Figure 5. Empirical $\hat{\phi}_f(\Delta k)$ and modeled $\phi_f(\Delta k)$ Frequency Correlation Functions.

Figure 6. Empirical $\hat{\phi}_t(\Delta l)$ and modeled $\phi_t(\Delta l)$ Time Correlation Functions.

3.2 Time Correlation Function

The time correlation function (TCF) for the ISM data traffic $X(t_l, f_k)$, which is dependent on Δl, is expressed as

$$\phi_t(\Delta l) = E(X^*(t_l, f_k)X(t_{l+\Delta l}, f_k)) \qquad (4)$$

where the estimation of $\phi_t(\Delta l)$ is obtained by the empirical TCF $\hat{\phi}_t(\Delta l)$ using the measured ISM data traffic as

$$\hat{\phi}_t(\Delta l) = \frac{1}{K(L-\Delta l)} \sum_{l=1}^{L-\Delta l} \sum_{k=1}^{K} x^*(t_l, f_k)x(t_{l+\Delta l}, f_k)$$

$$(5)$$

with $\Delta l = 0,.....,999$. As in the case of FCF, the TCF is also modeled as $\phi_t(\Delta l) = \rho_t^{\Delta l}$ as done in [14], where ρ_t denotes the correlation between the ISM data traffic observed at different time instances and $\rho_t \in [0,1]$. The modeled TCF $\phi_t(\Delta l)$ and the empirical TCF $\hat{\phi}_t(\Delta l)$ according to the above model are shown in Figure 6.

By using the LSE approach, the estimated value of $\hat{\rho}_{t,\text{LSE}}^{\Delta l}$ is 0.23 for $\Delta l = 1$ and 0.06 for $\Delta l = 2$. The aforementioned characterization of ISM data traffic $X(t_l, f_k)$ is used to validate its model in Section 5.

4 PARAMETRIC MODELING

The second-order moments of the process $X(t_l, f_k)$ are estimated as described in Section 3, but the joint PDF is also required for the complete description of the process $X(t_l, f_k)$. The joint PDF is estimated by adapting the parametric modeling approach, which is recognized as a powerful statistical tool for estimating the parameters of the presumed distribution.

In [10], the authors accept that their work to model the average signal channel power in television bands as a Gaussian is not always in agreement with the observed data. The reason for the disagreement between the observed and modeled data is the missing consideration of the multimodality of the signal power distribution. The Gaussian mixture (GM) is a suitable option to address the multimodality of the distribution of the measured signal as described in [16].

In this work, we separate the estimation of the joint PDF of the process $X(t_l, f_k)$ into the estimation of the PDF of the neighboring process

samples in frequency and time. After the PDFs of the aforementioned vectors are estimated, the joint PDF of $X(t_l, f_k)$ can be obtained by considering the independence of the components within the GM. Consider $X = \left[X(t_{l_1}, f_{k_1}), X(t_{l_2}, f_{k_2}) \right]^T$. At first, we consider the samples being neighbored in frequency (NIF), where $k_1 = k = k_2 - 1$ and $l_1 = l_2 = \lambda l_0$. The obtained vector is $X_\lambda = X_{\text{NIF},\lambda}$ with $l_0 \geq 1$, where k is chosen arbitrarily within the set $k \in \{1, \ldots, K-1\}$. The vector of samples neighbored in time (NIT) is denoted by $X_\kappa = X_{\text{NIT},\kappa}$ where $k_1 = k_2 = \kappa k_0$ and $l_1 = l = l_2 - 1$. Similar to the case of NIF, the choice of l is given by $l \in \{1, \ldots, L-1\}$ and $k_0 \geq 1$.

It is assumed that the distribution of the data vector X is a multivariate Gaussian mixture (MGM) in both NIF and NIT scenarios. The sum of N Gaussian densities can be expressed as in [17] and [18] as

$$p(X \mid \Theta) = \sum_{n=1}^{N} \pi_n f(X_n \mid \mu_n, \Sigma_n), \qquad (6)$$

where $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ and each θ_n represents the parameter set $\theta_n = \{\pi_n, \mu_n, \Sigma_n\}$ and $f(X_n \mid \mu_n, \Sigma_n)$ denotes the Gaussian density of each component given as

$$f(X_n \mid \mu_n, \Sigma_n) = \frac{1}{2\pi(\det \Sigma_n)^{\frac{1}{2}}} \times$$
$$\exp\left\{ -\frac{1}{2}(X - \mu_n)^T \Sigma_n^{-1} (X - \mu_n) \right\}. \qquad (7)$$

In (7), μ_n denotes the mean vector, Σ_n represents the covariance matrix and π_n is the prior probability of the nth component.

4.1 Expectation-Maximization Algorithm for Estimating the Parameters of ISM Data Traffic

The Expectation-Maximization (EM) is considered as a suitable technique for estimating the parameters of the PDF of the observed data. This algorithm finds its applications in other fields including data clustering in machine learning, reconstruction of medical images and also in computer vision. The EM scheme is performed iteratively for different number of components of a MGM with $n = 1, 2, \ldots, N$. Initial prior probabilities π_n are assumed to be identical among the components. The covariance matrices Σ_n are initially identity matrices, while the choice of mean vectors μ_n is made by a k-means clustering, where k-means clustering chooses the mean vectors randomly [19].

The objective here is to estimate the parameter set Θ for a MGM having N components, which maximizes the log-likelihood function (LLF)

$$\Lambda_N = \log p(X_1, \ldots, X_M \mid \Theta) = \sum_{m=1}^{M} \log \sum_{n=1}^{N} \frac{\pi_n}{2\pi(\det \Sigma_n)^{\frac{1}{2}}}$$
$$\times \exp\left\{ -\frac{1}{2}(X_m - \mu_n)^T \Sigma_n^{-1}(X_m - \mu_n) \right\}, \qquad (8)$$

where x_m represents the observation of X_m for both NIF and NIT scenarios. The description of the conditional probability density of the mth observation $X_m = x_m$ given the nth component of the MGM distribution is given by

$$p(X_m \mid n, \Theta) = \frac{1}{2\pi(\det \Sigma_n)^{\frac{1}{2}}} \times$$
$$\exp\left\{ -\frac{1}{2}(X_m - \mu_n)^T \Sigma_n^{-1}(X_m - \mu_n) \right\}. \qquad (9)$$

The EM algorithm works iteratively where it alternates between the expectation step (E-step) to update the posterior probabilities and maximization step (M-step) to update the parameters of the components of the MGM.

E-Step: Assume that we have an estimate of $\hat{\Theta}^{(j-1)}$ as the $(j-1)$th iteration of EM completed. In the E-Step, the conditional distribution of the nth component of the MGM given the observation $X_m = x_m$ is determined using the conditional probabilities $p(X_m \mid n, \hat{\Theta}^{(j-1)})$ and $\hat{\Theta}^{(j-1)}$ according to

$$p(n \mid X_m, \hat{\Theta}^{(j-1)}) = \frac{\pi_n p(X_m \mid n, \hat{\Theta}^{(j-1)})}{\sum_{n=1}^{N} \pi_n p(X_m \mid n, \hat{\Theta}^{(j-1)})}. \quad (10)$$

M-Step: In this step, updated estimates of the parameter set Θ are obtained using the posterior probabilities $p(n \mid X_m, \hat{\Theta}^{(j-1)})$ determined in the E-step according to

$$\hat{\pi}_n^{(j)} = \frac{1}{M} \sum_{m=1}^{M} p(n \mid X_m, \hat{\Theta}^{(j-1)}) \quad (11)$$

$$\hat{\mu}_n^{(j)} = \frac{\sum_{m=1}^{M} X_m p(n \mid X_m, \hat{\Theta}^{(j-1)})}{\sum_{m=1}^{M} p(n \mid X_m, \hat{\Theta}^{(j-1)})} \quad (12)$$

$$\hat{\Sigma}_n^{(j)} = \frac{\sum_{m=1}^{M} (X_m - \hat{\mu}_n^{(j-1)})(X_m - \hat{\mu}_n^{(j-1)})^{\mathrm{T}} p(n \mid X_m, \hat{\Theta}^{(j-1)})}{\sum_{m=1}^{M} p(n \mid X_m, \hat{\Theta}^{(j-1)})}.$$

$$(13)$$

4.2 Model Selection Criteria

It is well-known that the LLF in (8) is an increasing function of the number of components. There are several information-theoretic approaches for the selection of a suitable value of N to minimize the corresponding cost function. In this work, we adopt a heuristic approach in which we choose the minimum value of N for which the relative increase in the LLF is below a given threshold. It can also be expressed alternatively as

$$N = \min\left\{ n \in \mathbb{N} : \left| \frac{L_n - L_{n-1}}{L_{n-1}} \right| \le t \right\} \quad (14)$$

where the threshold is chosen to be $t = 0.003$.

5 MGM VALIDATION

For efficient DSA, it is valuable to have the distribution of the observed data traffic. In perspective of CR, the FCF describes the frequency behavior of the ISM data traffic. The joint distribution of the neighboring frequency subbands assists secondary users to determine the temporal occupancy of the observed frequency subbands. The TCF illustrates the temporal behavior of the ISM data traffic. It is also helpful for secondary users to have information about the bandwidth occupation for the specific time duration if the joint distribution of neighboring time slots is already known.

In this section, the joint distributions of neighboring frequency subbands and neighboring time domain signals of ISM data traffic are modeled parametrically using MGMs as mentioned in Section 4. Moreover, MGM is also validated using FCF and TCF for neighboring frequency subbands and the neighboring time domain signals, respectively.

5.1 MGM Validation using Frequency Correlation

5.1.1 Two Neighboring Frequency Subbands

In order to validate the MGM using FCF, the estimated parameters of MGM are required. The analytical expression of the correlation ρ_f is obtained from the MGM and yields the MGM estimator

$$\hat{\rho}_{f,\mathrm{MGM}}^{\Delta k} = \rho_f(\hat{\Theta}) = \sum_{n=1}^{N} \hat{\pi}_n (\hat{\varepsilon}_n + \hat{\mu}_{1n}\hat{\mu}_{2n}) \quad (15)$$

where $\hat{\varepsilon}_n$ represents the estimated cross-covariance between the frequency subbands, $\hat{\mu}_n$ denotes the estimated mean vector of the components and $\hat{\pi}_n$ are the estimated prior probabilities of the MGM model. In the NIF case, the frequency subbands are denoted by the vector $X_\lambda = X_{\mathrm{NIF},\lambda}$ as mentioned in Section 4. For the

two neighboring frequency subbands where $l_0 = 1$, $M = L = 1000$ and $k = 99$, the number of components $N = 9$ is calculated by (14). The scatter plot of the measured data and modeled joint distribution of the data is shown in Figure 7.
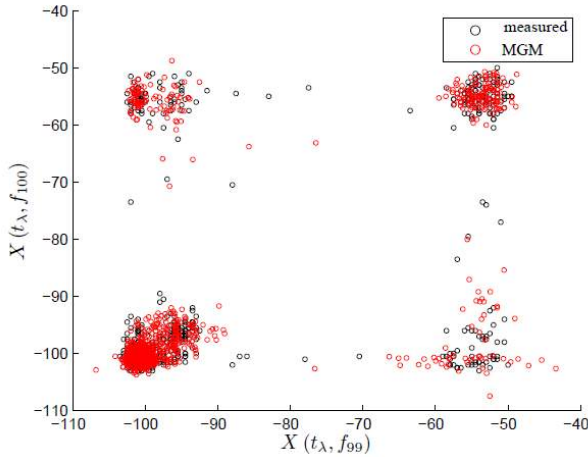
Figure 7. Joint distribution of $X_{\text{NIF},\lambda}$ with $l_0 = 1$, $M = L = 1000$, $k = 99$ and $N = 9$ modeled using MGM.

Figure 7 clearly shows the close agreement between the L measurements in neighboring frequency subbands and M samples from the MGM model.

The MGM estimator for the two neighboring frequency subbands when $\Delta k = 1$ is expressed as

$$\hat{\rho}^1_{f,\text{MGM}} = \rho_f(\hat{\Theta}) = \sum_{n=1}^{N} \hat{\pi}_n (\hat{\varepsilon}_n + \hat{\mu}_{1n} \hat{\mu}_{2n}). \quad (16)$$

The values of the components' mean vector are negligible. The estimated values of priors and cross-covariances are given in Table 2.

Table 2. Estimated parameters of MGM for two neighboring frequency subbands with $N = 9$

$\hat{\pi}_n$	0.0291	0.0303	0.1846	0.1160	0.0066	0.0413	0.4012	0.1631	0.0277
$\hat{\varepsilon}_n$	0.4673	-0.7824	0.4403	0.5928	52.553	0.0051	0.0162	0.0903	0.7995

We substitute these values in (16), which yields a value of $\hat{\rho}^1_{f,\text{MGM}} = 0.54$ and is in close agreement to $\hat{\rho}^1_{f,\text{LSE}} = 0.56$ calculated from the modeled FCF in Section 3.

5.1.2 Three Neighboring Frequency Subbands

Here we extend our validation of the MGM model to three neighboring frequency subbands where $l_0 = 1$, $M = L = 1000$ and $k = 106$. In this case, the number of components for $X_\lambda = X_{\text{NIF},\lambda}$ is estimated to be $N = 10$. The scatter plot of the measured data and modeled joint distribution of the data is shown in Figure 8.
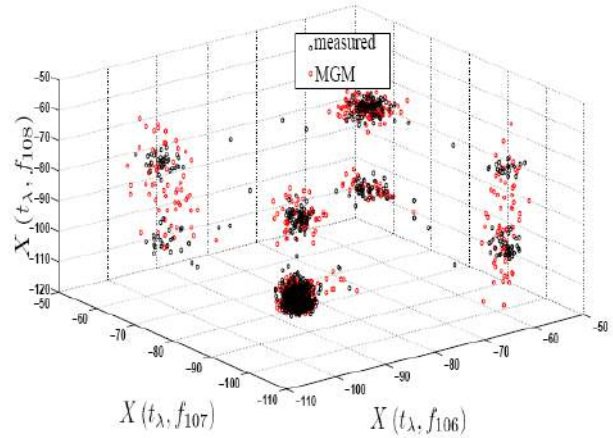
Figure 8. Joint distribution of $X_{\text{NIF},\lambda}$ with $l_0 = 1$, $M = L = 1000$, $k = 106$ and $N = 10$ modeled using MGM.

As was the case in Figure 7, the scatter plot in Figure 8 also shows a good fit between the measured and modeled data.

The MGM estimator for the three neighboring frequency subbands with $\Delta k = 1$ and $\Delta k = 2$ is given by (17) and (18) respectively,

$$\hat{\rho}^1_{f,\text{MGM}} = \rho_f(\hat{\Theta}) = \sum_{n=1}^{N} \hat{\pi}_n (\hat{\varepsilon}_n + \hat{\mu}_{1n} \hat{\mu}_{2n} \hat{\mu}_{3n}) \quad (17)$$

$$\hat{r}^2_{f,\text{MGM}} = r_f(\hat{Q}) = \sum_{n=1}^{N} \hat{p}_n (\hat{e}_n + \hat{m}_{1n} \hat{m}_{2n} \hat{m}_{3n}). \quad (18)$$

After substituting the estimated parameters from MGM in (17) we have the value of $\hat{\rho}^1_{f,\text{MGM}} = 0.55$ between the frequency subbands f_{106} and f_{107} whereas $\hat{\rho}^1_{f,\text{MGM}} = 0.51$ is obtained between frequency subbands f_{107} and f_{108}, both the estimated values of $\hat{\rho}^1_{f,\text{MGM}}$ are very close to $\hat{\rho}^1_{f,\text{LSE}} = 0.56$ estimated in Section 3. Moreover, $\hat{\rho}^2_{f,\text{MGM}} = 0.24$ is obtained after substituting the

estimated parameters in (18) also approximately in agreement with $\hat{\rho}_{f,\text{LSE}}^2 = 0.31$.

5.2 MGM validation using Time Correlation

5.2.1 In Two Neighboring Time Slots

To validate the MGM model of the ISM data traffic using TCF, the required MGM estimator is

$$\hat{r}_{t,\text{MGM}}^{\text{D}l} = r_t(\hat{\text{Q}}) = \overset{N}{\underset{n=1}{\mathring{\text{a}}}} \hat{\rho}_n(\hat{e}_n + \hat{m}_{1n}\hat{m}_{2n}). \quad (19)$$

In the NIT case, the vectors of neighboring time domain signals are denoted as $X_\kappa = X_{\text{NIT},\kappa}$.

For the case of two neighboring time domain signals where $k_0 = 1$, $M = K = 285$ and $l = 477$, the number of components estimated by (14) is to $N = 8$. A close agreement between the measured data and the modeled data based on the estimated parameter set $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ as shown in Figure 9 validates our assumption for the NIT case.

The modeled MGM is also validated for two neighboring time domain signals using the given estimator. The estimated values of priors and cross-covariances are given in Table 3.

Table 3. Estimated parameters of MGM for two neighboring time domain signals with $N = 8$

$\hat{\pi}_n$	0.2537	0.0175	0.0183	0.4502	0.0243	0.0206	0.1245	0.0908
$\hat{\varepsilon}_n$	0.3738	17.290	17.693	0.0043	2.7822	0.6155	0.0230	0.6122

After inserting the values of estimated parameters from Table 3 in the MGM estimator

$$\hat{r}_{t,\text{MGM}}^1 = r_t(\hat{\text{Q}}) = \overset{N}{\underset{n=1}{\mathring{\text{a}}}} \hat{\rho}_n(\hat{e}_n + \hat{m}_{1n}\hat{m}_{2n}). \quad (20)$$

We obtain $\hat{\rho}_{t,\text{MGM}}^1 = 0.2037$ with $\Delta l = 1$ which is clearly close to the estimated value of $\hat{\rho}_{f,\text{LSE}}^1 = 0.23$ using TCF in Section 3.
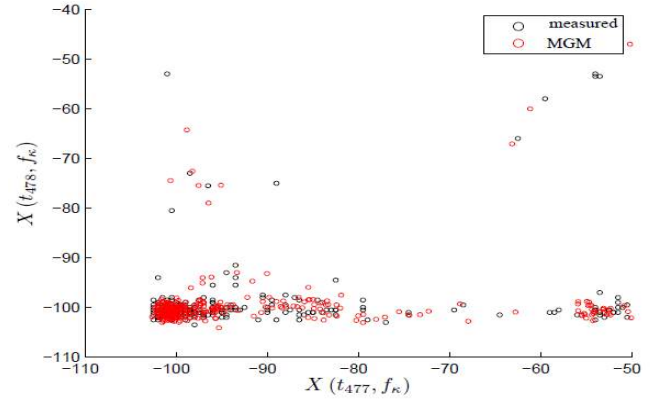
Figure 9. Joint distribution of $X_{\text{NIT},\kappa}$ with $k_0 = 1$, $M = K = 285$, $l = 477$ and $N = 8$ modeled using MGM.

5.2.2 Three Neighboring Time Slots

In case of the extension of MGM for three neighboring time domain signals where $k_0 = 1$, $M = K = 285$ and $l = 777$, the estimated number of components for three neighboring time slots using (14) is $N = 10$. In Figure 10, the scatter plots of the measured data and MGM samples show a good fit.
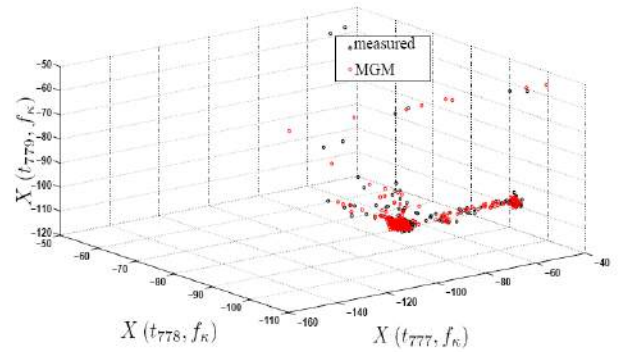
Figure 10. Joint distribution of $X_{\text{NIT},\kappa}$ with $k_0 = 1$, $M = K = 285$, $l = 777$ and $N = 10$ modeled using MGM.

The MGM estimator for three neighboring time domain signals with $\Delta l = 1$ and $\Delta l = 2$ is formulated as

$$\hat{\rho}_{t,\text{MGM}}^1 = \rho_t(\hat{\Theta}) = \sum_{n=1}^{N} \hat{\pi}_n(\hat{\varepsilon}_n + \hat{\mu}_{1n}\hat{\mu}_{2n}\hat{\mu}_{3n}) \quad (21)$$

$$\hat{\rho}^2_{t,\text{MGM}} = \rho_t(\hat{\Theta}) = \sum_{n=1}^{N} \hat{\pi}_n (\hat{\varepsilon}_n + \hat{\mu}_{1n} \hat{\mu}_{2n} \hat{\mu}_{3n}) . \quad (22)$$

After substituting the estimated parameters from MGM in (21) we find that the value of $\hat{\rho}^1_{t,\text{MGM}} = 0.19$ estimated between the signals at time slots t_{777} and t_{778} with $\hat{\rho}^1_{t,\text{MGM}} = 0.25$ estimated between the signals at time slots t_{778} and t_{779} are in close agreement with $\hat{\rho}^1_{t,\text{LSE}} = 0.23$. The value of $\hat{\rho}^2_{t,\text{MGM}} = 0.08$ obtained after substituting the estimated parameters in (22) is also nearly equal to the $\hat{\rho}^2_{t,\text{LSE}} = 0.06$ estimated using TCF in Section 3.

6 CONCLUSIONS

We conduct indoor measurement campaigns to analyze the RF activity. The frequency and time correlation functions of measured data traffic are modeled by decaying exponentials. MGM is considered to model the observed data traffic in neighboring frequency subbands and also at neighboring time instances by adopting a parametric approach. In order to estimate the parameters of MGM, the EM algorithm is used. The selection of a suitable initial mean vector is done using *k*-means clustering in a heuristic way. It is found that using real time measurements in CR, a suitable model for the observed data in the two neighboring frequency subbands is MGM with $N = 9$ and for the extended case of three neighboring frequency subbands with $N = 10$. For the observed data traffic in two neighboring time slots, MGM is selected with $N = 8$ and for the case of three neighboring time domain signals, the suitable choice is MGM with $N = 10$. It is also validated that the MGM estimator provides an accurate correlation between the frequency subbands and time domain signals.

REFERENCES

[1] J. Xue, Z. Feng, and K. Chen, "Beijing spectrum survey for cognitive radio applications," in Vehicular Technology Conference (VTC Fall)*, 2013 IEEE *78th*, Sept 2013, pp. 1–5.

[2] A. A. Ayeni, N. Faruk, N. T. Surajudeen-Bakinde, R. A. Okanlawon, and Y. A. Adediran, "Spatial spectrum utilization efficiency metric for spectrum sharing system," International Journal of Digital Information and Wireless Communications (IJDIWC), vol. 5, no. 1, pp. 44 –51, 2015.

[3] M. Biggs, A. Henley, and T. Clarkson, "Occupancy Analysis of the 2.4 GHz ISM Band," Communications, IEE Proceedings-, vol. 151, no. 5, pp. 481–488, Oct 2004.

[4] V. N. Q. Bao, L. Q. Cuong, L. Q. Phu, T. D. Thuan, N. T. Quy, and L. M. Trung, "Vietnam spectrum occupancy measurements and analysis for cognitive radio applications," in Advanced Technologies for Communications (ATC), 2011 International Conference on, Aug 2011, pp. 135–143.

[5] D. Roberson, C. Hood, J. L. LoCicero, and J. MacDonald, "Spectral occupancy and interference studies in support of cognitive radio technology deployment," in Networking Technologies for Software Defined Radio Networks, 2006. SDR '06. 1st IEEE Workshop on, Sept 2006, pp. 26–35.

[6] S. Pagadarai and A. M. Wyglinski, "A quantitative assessment of wireless spectrum measurements for dynamic spectrum access," in Cognitive Radio Oriented Wireless Networks and Communications, 2009. CROWNCOM '09. 4th International Conference on, June 2009, pp. 1–5.

[7] M. Dzulkifli, M. Kamarudin, and T. Rahman, "Spectrum occupancy at uhf tv band for cognitive radio applications," in *RF and Microwave Conference (RFM), 2011 IEEE International*, Dec 2011, pp. 111–114.

[8] K. Patil, K. Skouby, and R. Prasad, "Stochastic duty cycle model based on measurement for cognitive radio," in Wireless Personal MultimediaCommunications (WPMC), 2012, 15th International Symposium on, Sept 2012, pp. 128–132.

[9] M. Dzulkifli, M. Kamarudin, and T. Rahman, "Spectrum occupancy of malaysia radio environment for cognitive radio application," in Wireless Communications and Applications (ICWCA 2012), IET International Conference on, Oct 2012, pp. 1–6.

[10] F. C. Miguel Lopez-Benitez, "Spectrum occupancy in realistic scenarios and duty cycle model for cognitive radio," Advances in Electronics and Telecommunications, vol. 1, no. 1, pp. 26 –34, April 2010.

[11] L. Yin, S.-X. Yin, S. Wang, E.-Q. Zhang, W.-J. Ho, and S. fang LI, "Quantitative spectrum occupancy evaluation in china: based on a large scale concurrent spectrum measurement," The Journal of China Universities of Posts and Telecommunications, vol. 19, no. 3, pp. 122 – 128, 2012.

[12] A. Pintor, M. To, J. Salenga, G. Geslani, D. Agpawa, and M. Cabatuan, "Spectrum survey of vhf and uhf bands in the philippines," in TENCON2012 - 2012 IEEE Region 10 Conference, Nov 2012, pp. 1–6.

[13] C. Ghosh, S. Pagadarai, D. Agrawal, and A. M. Wyglinski, "A framework for statistical wireless spectrum occupancy modeling," Wireless Communications, IEEE Transactions on, vol. 9, no. 1, pp. 38–44, January 2010.

[14] K. K. Toshiyasu Kurasugi and Y. Takahashi, "A markovian model of coded video traffic which exhibits long-range dependence in statistical analysis," Journal of the Operations Research, vol. 42, no. 1, 1999.

[15] M. Ehsan and D. Dahlhaus, "Statistical modeling of ism data traffic in indoor environments for cognitive radio systems," in Digital Information, Networking, and Wireless Communications (DINWC), 2015 Third International Conference on, Feb 2015, pp. 88–93..

[16] M. F. Tommaso Costa, Giuseppe Boccignone, "Gaussian mixture model of heart rate variability," PLOS ONE, vol. 7, no. 5, May 2012.

[17] G. Mclachlan, Finite Mixture Models. Cambridge, MA: John Wiley & Sons, Inc., 2000.

[18] D. Choudhary and A. Robinson, "A new approach to wireless channel modeling using finite mixture models," International Journal of Digital Information and Wireless Communications (IJDIWC), vol. 4, no. 2, pp.44 –51, 2014.

 [19] Seber and G.A.F, Multivariate Observations. Hoboken, NJ: John Wiley and Sons, 1984.

Elucidating Digital Information and Wireless Communications Technology for an Electronic Working Logs and Applications

De-Ji Jang[1], Chih-Yung Chen[2] and Shieh-Shing Lin[3]
[1,3]Department of Electrical Engineering
[2]Department of Information Management
St. John's University
499, Sec. 4 Tam King Road, Tamsui, Taipei 25135, Taiwan
E-mail: [1]tedjang@gmail.com, [2]yung@mail.sju.edu.tw, [3]sslin@mail.sju.edu.tw

ABSTRACT

This work focuses on the troublesomeness of the traditional hardcopy working log sheets used in officers, proposing a Digital Information and Wireless Communications Technology based a Noble Design of Electronic Working Logs (EWL). This work also employs this dedicated tool to a real agent. There are some special features in this work: i) Easily operational interface, users can simply and instantly inquire the on-line request through the EWL; ii) Instantaneously monitor the situation, supervisors can clearly scrutinize the events; iii) Forever preservation, agent can conserve the electronic working logs endlessly due to the EWL. This work also implements the proposed EWL in a real officer agent and obtains some successful results.

KEYWORDS

Digital Information Technology, Wireless Communications Technology, Hardcopy working log sheet, Electronic Working Log, Supervisors.

1 INTRODUCTION

Digital Information and Wireless Communications Technology are widely spent in many filed listed [1]-[15]. Up to now, the Intelligent Information Systems and E-commerce Technologies are also broadly devoted in the literals listed in [16]-[42]. However, up to now, there are few documents exploring the application of this technology in officer working logs.

There are some drawbacks in conventional hardcopy working logs for officers: i) *too much time consumption*; officers should accomplish the corresponding data for the conventional hardcopy working logs while the duty is off and executing the duty transition in every mission. It is certainty too much time consumption for officers in every mission, especially, while the emergency case occurred. It could make cases worse to the delay for the next duty execution; ii) *hard to examine the hardcopy working logs*, the conventional hardcopy working logs are not easy being checked and Verification as to the history events for the executed duties. Moreover, supervisors cannot validate the corresponding records for the correlated events; iii) *difficult to preserve for the hardcopy working logs*, the contents of the conventional hardcopy working logs are consisted with traditional paper and they are not easily to kept in good condition owing to the wet in the air. In convention, five years is the period of keeping for conventional hardcopy working logs. Furthermore, they are obviously too much space consumption for every agent. Focusing on the inconvenience and drawback of the traditional hardcopy working logs sheet used in officers, this work proposed a Digital Information and Wireless Communications Technology based a Noble Design of Electronic Working Logs (EWL) and employed this dedicated tool to a real agent. There are some special features in this paper: i) Easily operational interface, users can simply and instantly inquire the on-line request of the EWL; ii) Instantaneously monitor the situation, supervisors can cleanly scrutinize the events; iii) Forever preservation, agent can conserve the electronic working logs forever due to the EWL. This work

also implemented the proposed EWL in a real officer agent and obtains some successful results.

Following of this work is organized in the subsequent manner. Section 2 presents the Texts Analysis for the conventional hardcopy working logs for routine duty in agent. A Digital Information and Wireless Communications Technology based a Noble Design of Electronic Working Logs (EWL) is proposed in Section 3. Section 4 implements the dedicated design EWL in a real Agent. Finally, this work makes a brief conclusion in Section 5.

2 TEXTS ANALYSIS

2.1 State of affairs of the Electronic Sheet in Agent

The applications of Digital Information and Wireless Communication Technology used in agent were raised from 1997. Up to now, few and/or not too many attempted to conduct these technologies to many agents. Nevertheless, there are very few applications corresponding to the Electronic Working Logs for officers. There are four stages of the project of Electronic/Network promotion in Agent from 1997 stated below:

2.1.1 Periods of Electronic government plan in Agent

Level 1. The mid-term plan of the Electronic/Network promotion in Agent was executed from 1998 to 2000.

Level 2. The project of the Electronic Promotion in agent is executed from 2001 to 2004 and the E-agent Project was executed from 2003 to 2007. Subsequently, the challenged Project and Digital Project in agent were executed from 2008.

Level 3. The Project of high quality Network Processing was executed from 2008 to 2111.

Level 4. The intelligent improvement Project in Agent was executed from 2011 to 2016.

2.2 The Original Investigation and On-Line Simulation of the Electronic Working Logs for duty in Agent

2.2.1 Novelty of the Electronic Working Logs

Owing to the following explanations, the Electronic Working Logs for duty was created;

a) There existed too much time and space consumption for agent and duty;

b) It is hard to examine the hardcopy working logs for supervisors and/or officers;

c) Additionally, it is difficult to preserve for the traditional paper hardcopy working logs.

Moreover, it is a trend for Agent to been Electronic/Network for Duty in this age.

2.3 Prototype of the System Development as to the Electronic Working Logs and On-Line Test

First of all, the prototype dedicated design of this Electronic Working Logs for duty is only executed in only one unit in a single agent to simulate the efficiency of this design. Furthermore, the following four steps tests are executed to justify and/or rectify the procedure of this EWL.

2.3.1 Process Stages

Step 1. System Analysis and Program Construction- Focusing on the inconvenience of the traditional paper hardcopy working logs sheet used in officers, this work rectify this drawback owing to the experience.

Step 2. System Test- Only one unit in a single agent of the Department of Information Management Office executes the system test and Hardware and Software combination.

Step 3. Traning for the users as to the officers in agent- One agent of the Department of Information Management Office executes the training.

Step 4. Spread this proposed Electronic Working Logs for duty to other units in this agent- Invite other units join this project.

Subsequently, the On-Line tests associated with other units in this agent are executed and spread this technology to all units in this simulated agent.

3 A DIGITAL INFORMATION AND WIRELESS COMMUNICATIONS TECHNOLOGY DESIGN OF EWL

3.1 Digital Information Technology-Prologue of the Web of the Electronic Working Logs for duty in Agent

*The Web of the EWL-*The details description corresponding to the Digital Information and Wireless Communications Technology based design Electronic Working Logs for duty was shown in the panel of the EWL. First of all, users should submit the personal identify of the username and password with the authorization passwords to get in the main system. Figure 1 shows the floating chat of Input and Platform of the Web of the EWL in main system. Figure 2 shows the real photograph of the system.
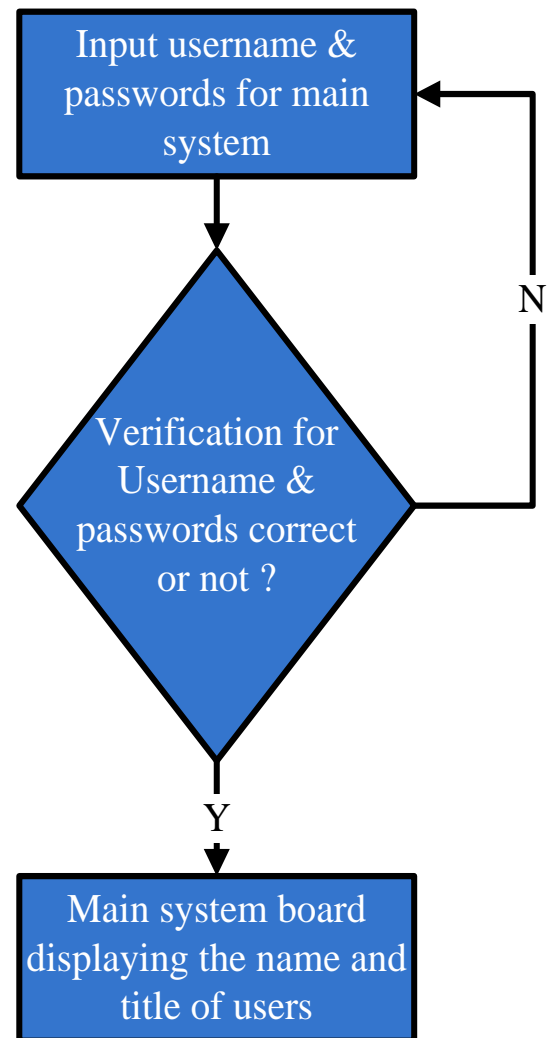
Figure 1. Floating chat of Input and Platform of the Web of the EWL in main system

Figure 2. The real photograph of the system

Information Management Catalog Introduction-
There are several icons shown in the Digital Information and Wireless Communications Technology based design EWL with different colors to make a difference, such as "Date", "Time" "Duty item", "Action signal of Y/N", "Officer name", "Status of T/C", "Edition", "Officer signature", "People Signature", "Acknowledge", "Delete", "Legal case description" and "Verification".

3.2 Wireless Communications Technology-the Operation of the Management System of Electronic Working Logs for duty in Agent

3.2.1 Wireless Communications Technology as to Operation Description-Acquirement Interface

There are some main items in the Digital Information and Wireless Communications Technology based design EWL as to the "Verification", *"Adding a new working log for duty", "Acquirement", "Output" and "Delete".*
Users selected the information shown in the Web through the EWL panel, pressing the button of "Acquirement" to execute the mission of checking and/or editing, then the following message with different colors icons corresponding to the message of "Date", "Time" "Duty item", "Action signal of Y/N", "Officer name", "Status of T/C", "Edition", "Officer signature", "People signature", "Acknowledge", "Delete", "Legal case description" and "Verification" will be responded immediately. Users can easily check and confirm those messages shown in the web through the EWL panel. In addition, those different colors will discriminate the different status of the processing Legal events. Figure 3 shows the real photograph of users to get in the system of EWL. Figure 4 shows the Process of users within Acquirement function. Figure 5 shows the Process of users to get in the system of EWL executing the function of Adding a new job. Figure 6 shows the real photograph of users to get in the system of EWL executing the function of Adding a new working log for duty in agent.

Figure 3. The real photographs of users to get in the system of EWL
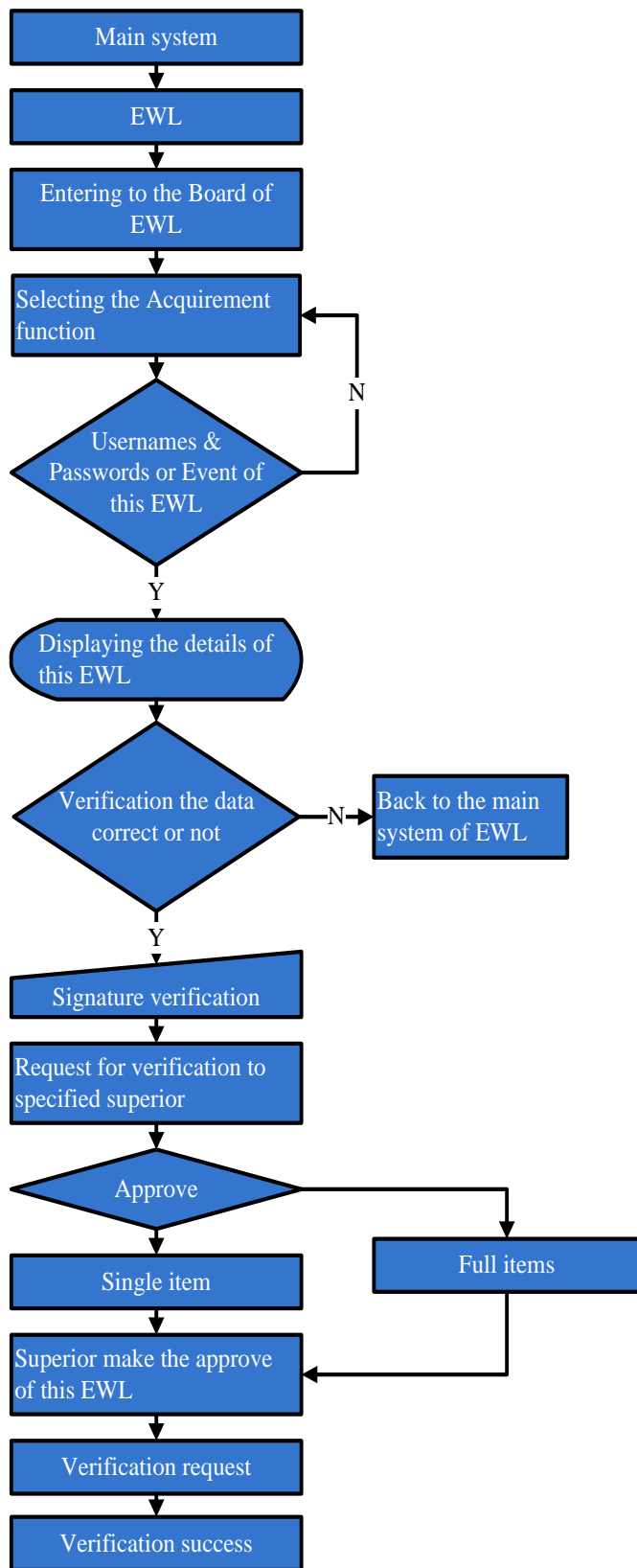
Figure 4. Process of users within Acquirement function
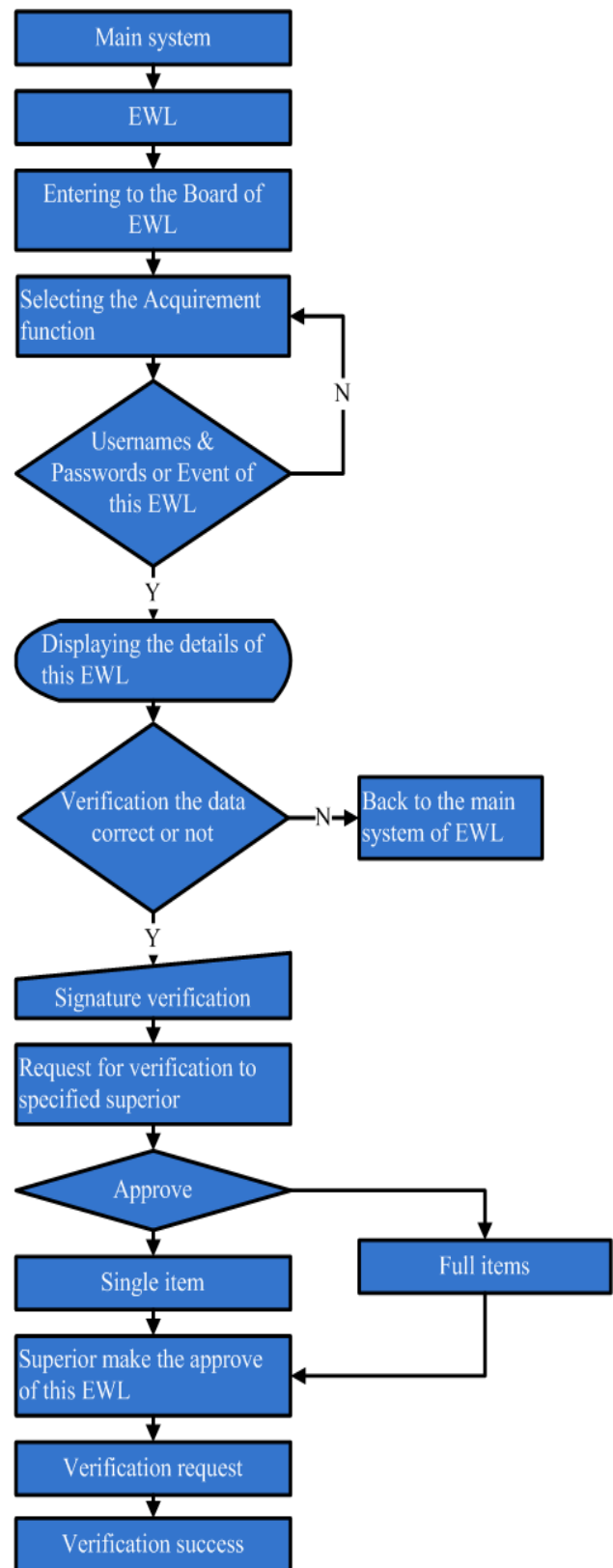
Figure 5. Process of EWL for users executing the function of Adding a new working log for duty

Figure 6. The real photograph of users get in the system of EWL executing the function of Adding a new working log for duty

"Date"- officer should fill in the currently data exactly.

"Time"-officer should fill in the currently time exactly.

"Duty item"- officer should fill in the duty items, such as the Traffic affair or others.

"Action signal of Y/N"- officer should fill in whether other Legal Affairs are existed and being proceeded.

"Officer name"-officers should fill in officer's name.

"Status of T/C"- officers should fill in the status of Temperately Record or Completely this event.

"Edition"- officers can modify the message of the execution legal event before "Output".

"Officer Signature"- officers should sign the officer's name to identify.

"People signature"- officers should acknowledge the people who involve the legal event to make a signature.

"Acknowledge"- officers should acknowledge the output command of this legal event to the Supervisor for Verification.

"Delete"- officers can delete this legal event with certainly reason descriptions and others.

"Legal case description"- officers should describe the details or make a brief description of the execution task.

"Verification"- officers should verify this event to output.

3.3 Digital Information and Wireless Communications Technology Process Description

Pressing the Panel command of the Digital Information and Wireless Communications Technology based design EWL "Adding a new working log for duty" to execute the Adding. The following steps should be completed.

i) Officers should fill in the exactly currently date and time. In addion, those data were beginning as the legal event occurred.

ii) Once other officers join this event, others should also fill in the corresponding names and ID number through the EWL panel. Finally, push the button of "Adding". Moreover, officers won't join this task due to urgent cases occurred beyond the officer controls. The officer can just push the button of "Delete".

iii) Other Legal Events item was employed to handle the corresponding legal events. The icon as to the "Y" represents the fact that there are existed other events should be handled. Officers should type the exactly the numbers of people and the details of the event. Otherwise, the icon presents "N".

iv) The process of the working logs is stated below. Users can also employ the corresponding phase words built-in the EWL memory for edition. Also,

users can type words directly through the EWL panel.

Finally, ***"Adding a new working log for duty"*** is completely executed. Users can also make some other "Adding", "Delete" or make a modification of the previous job. However, the proposed EWL will respond the executed results through the screen of the panel, such as the corresponding message of "Adding Success" or "Delete Success", "Verification Success", "Output Success" … *etc*. Once this action is not completed owing to some other typing errors or data losses, EWL will request the users to make a modification of the precious stage. It should be noticed that every action need be verified by the officers "Signature" as well as the People "Signature" before this Legal event was transferred to the Supervisor for Verification. However, if users are in temporary record for an executing duty of a EWL, she and/or he may employ the temporary register in the Digital Information and Wireless Communications Technology based design EWL. Figure 7 shows the process of the Status of users within Temporary Register.
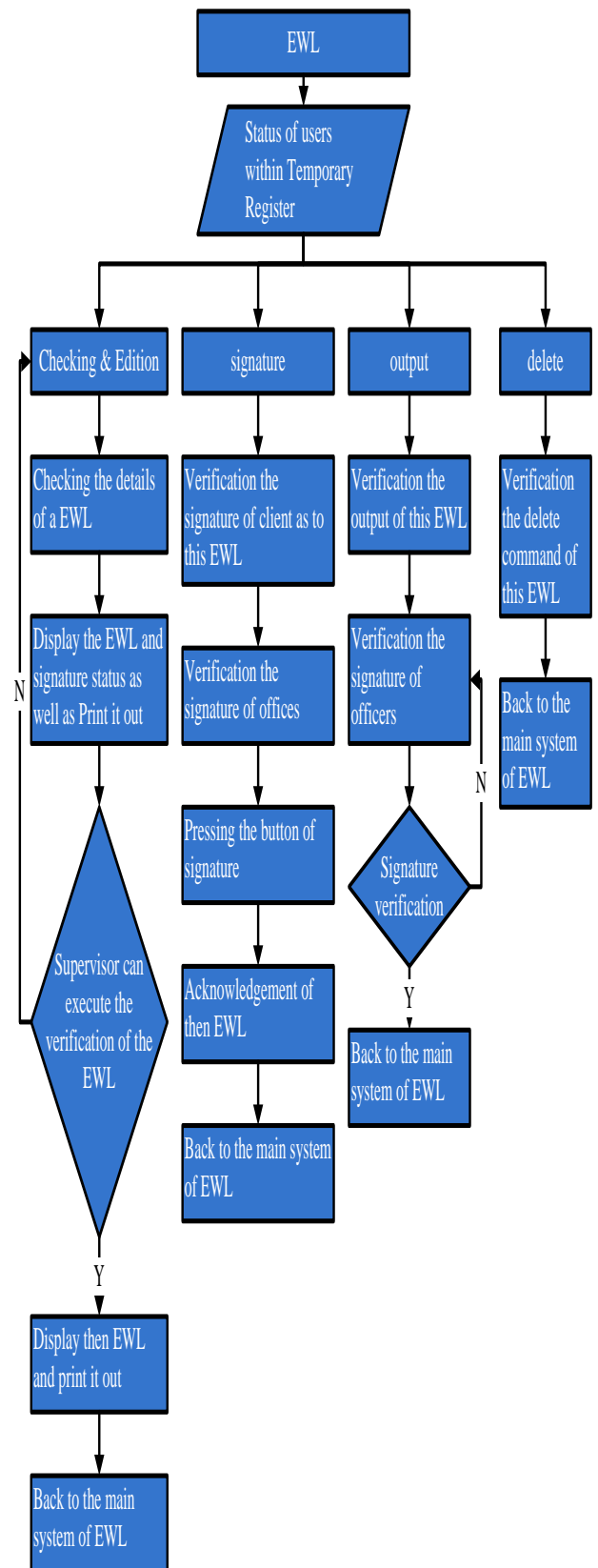
Figure 7. Process of the Status of users within Temporary Register

Figure 8 shows the real photographs of users to get in the system of EWL executing the function of Acquirement. Figure 9 shows the real photograph of users to get in the system of EWL executing the function of Signature. Figure 10 shows the Process of users within Output function. Figure 11 shows the real photograph of users to get in the system of EWL executing the function of Output. Figure 12 shows the real photograph of users gets in the system of EWL executing the function of Verification.

Figure 8. The real photograph of users to gets in the system of EWL executing the function of Acquirement

Figure 9. The real photograph of users to get in the system of EWL executing the function of Signature
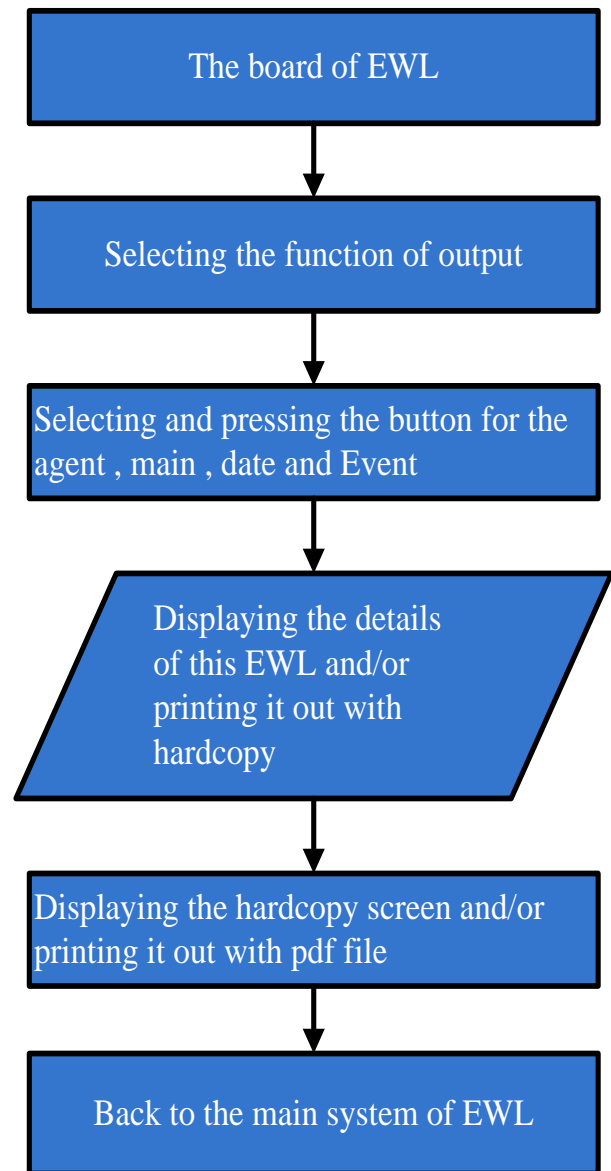
Figure 10. Process of users within Output function

Figure 11. The real photograph of users to get in the system of EWL executing the function of Output

Figure 12. The real photograph of users gets in the system of EWL executing the function of Verification

3.4 Operation Description- Edition, Delete, Officers and People Signature, Output, Verification, Export

The panel of EWL will present the corresponding phase of icons with different colors of "Date", "Time", "Items", "Other Events", "Edition", "Delete", "Temperately Records", "Officers and People Signature", "Output" and "Verification" guild users to operate.

Once this action is not completed owing to some other reasons of typing errors or data losses, EWL will request the users to make a modification of precious stage. It should be noticed that every action need be verified by the officers "Signature" before outputting this Legal event to the Supervisor for "Verification". Details are stated below.

"Temperately Record"- Users can execute the "Edition", "Signature" and "Delete". However, once this legal event was executed completely though the "Output Command". Users cannot make any other modification of this case. Once there existed some other reasons of typing errors or data losses, EWL will request the users to make a modification of this precious stage.

"Verification"-the supervisor with the authorization executed the Verification Command.

"Colors Discrimination"- Different Colors type of icons discriminated the different items job of EWL.

"Output Command"- officers can print it out of this handled Legal event through the printer and output this job to Supervisor for "Verification". However, every handled cased should be notified by Officer and People's Signature for recognition.

4 APPLICATIONS

4.1 A Real Agent is selected to simulate the practical operation of Digital Information and Wireless Communications Technology based design Electronic Working Logs for duty in Agent

There are three algorithmic stages to promote this Electronic Working Logs for duty.

4.1.1 Algorithmic Stages

A-Stage 1. First of all, 20 sets of facility as to these signature panels and the corresponding devices are gained to execute the Digital Information and Wireless Communications Technology based design the Electronic/Network EWL. There are seven units in this agent to execute this Electronic Working Logs for duty.

A-Stage 2. There are 222 sets of facility as to the signature panels and the corresponding devices are gained to carry out the Digital Information and Wireless Communications Technology based design the Electronic/Network EWL for duty. There increase *nine units* more with respect to the previous Stage of *seven units* in this agent to execute this Electronic Working Logs for duty.

A-Stage 3. There are 300 sets of facility as to the signature panels and the corresponding devices are gained to execute the Digital Information and Wireless Communications Technology based design the Electronic/Network EWL for duty. All units in this Agent are invited to join this Project to execute this Electronic Working Logs for duty.

4.2 Performance Evaluation of the proposed the Digital Information and Wireless Communications Technology based Electronic Working Logs for duty

The performance of this Digital Information and Wireless Communications Technology based

design Electronic Working Logs for duty in agent is stated below.

-*E-performance*, Electronic/Network EWL for duty concept substitutes the tradition hardcopy of paper working logs.

-*Time and Space saving,* this dedicated Digital Information and Wireless Communications Technology based design of Electronic Working Logs for duty is certainly superior with respect to the precious hardcopy working logs owing to time consumption and space consumption.

-*Economy efficiency*, the tradition working logs cost more than this dedicated design of Electronic Working Logs for duty.

-*Forever preservation*, agent can conserve the electronic working logs forever due to the EWL.

-*Justification*, officers can clearly and promptly investigate and/or check every past case of the corresponding message through the dedicated Digital Information and Wireless Communications Technology based design of Electronic Working Logs for duty.

5 CONCLUSIONS

This paper engrossed on the inconvenience of the traditional hardcopy working logs sheets used in officers in agent, proposing a Digital Information and Wireless Communications Technology based design of a Noble Design of Electronic Working Logs (EWL) for duty in agent. Furthermore, this paper also employed this dedicated tool to a real agent. There are some special features in this paper: i) Easily operational interface, users as well as officers and superiors can simply and instantly inquire the on-line request of the EWL; ii) Instantaneously monitor the situation, supervisors can clearly scrutinize the events; iii) Forever preservation, agent can conserve the electronic working logs forever due to the EWL. In additional, this paper also created a real Digital Information and Wireless Communications Technology based design of an EWL and implemented the proposed EWL in a real officer agent and obtains the following successful results: E-performance, Electronic/network EWL, Time and space saving, Economy efficiency, forever preservation, and Justification obtained for the implemented agent.

REFERENCES

[1] L.-T Liu, "The Research of dealing with the report of E-Platorm and service quality," S cinnection-To use as an example for Taipei County Government Police Bereau, vol. 1, no. 1, pp. 7-20, 2006.

[2] J.Y Kuo, "Topic Form System of Staff's Attendance in Taipei City Fire Department-Perspective From," Unified Theory of Acceptance and Use of Technolgy, vol. 1, pp. 1-2, 2011.

[3] B.-J Tai, "Empirical study on the Relationship between e-process Practices and Performance Improvement in the Software Development," Technology, vol. 1, pp. 6-14, 2004.

[4] U.-J Jeng, "The Study on the Degree of User Satisfaction for the e-Workflow System," Technology, vol. 1, pp.1-7, 2014.

[5] A. P. Balutis, "Monitoring the E-Government Revolution," Technology, vol. 27, Issue1, 1996.

[6] Adelina Tang, "Towards the Use of a Computing Ontology: A Curricula Management System Proposal," IJDIWC,The Society of Digital Information and Wireless Communications, pp. 100-101, 2011.

[7] Boonsit Yimwadsa na,Chalalai Chaihirunkarn,Apichaya Jai choom,Apichaya Thawornchak, "DocFlow: An Integrated Document Workflow for Business Process Management," IJDIWC, The Society of Digital Information and Wireless Communications, pp. 222, 2011.

[8] Juan Carlos López Pimentel,Víctor Fernando Ramos Fon Bon, Raúl Monroy, "A Web Service for Signing and Authenticating Digital Documents based on Symmetric Cryptography Protocol," IJDIWC,The Society of Digital Information and Wireless Communications, pp.4 40-442, 2011.

[9] Sattar J Aboud,Sufian Yousef, "A PRACTICAL PROXY SIGNATURE SCHEME," IJDIWC,The Society of Digital Information and Wireless Communications, pp. 297-298, 2012.

[10] AmrBadr-El-Din, "Object-Oriented in Organization Management: Organic Organization," IJDIWC,The Society of Digital Information and Wireless Communications, pp. 441-442, 2013.

[11] Niloofar Khanghahi,Ramin Nasiri,Mahsa Razavi, "A New Approach Towards Integrated Cloud Computing Architecture," IJDIWC,The Society of Digital

Information and Wireless Communications, pp. 24-30, 2014.

[12] Ahmad Mosallanejad, Rodziah Atan, "Masrah Azmi Murad, Rusli Abdullah," A Hierarchical Self-Healing SLA for Cloud Computing," IJDIWC,The Society of Digital Information and Wireless Communications, pp. 43, 2014.

[13] Thamilvaani Arvaree @ Alvar and Associate Prof. Dr. Rodziah Atan, "SERVICE AVAILABILITY AND ACCESSIBILITY OF REQUIREMENTS USING CLUSTERING IN CLOUD ENVIRONMENT," IJNCAA, The Society of Digital Information and Wireless Communications, pp. 457-460, 2011.

[14] Kenichi Takahashi , Takanori Matsuzaki , Tsunenori Mine , Takao Kawamura, Kazunori Sugahara, " Protection of Personal Information based on User Preference," IJNCAA,The Society of Digital Information and Wireless Communications,(ISSN:2220-9085),pp.800-801,2011.

[15] Ali A. Yassin , Hikmat Z. Neima , Haider Sh. Hashim, " Security and Integrity of Data in Cloud Computing Based on Feature Extraction of Handwriting Signature," IJCSDF,The Society of Digital Information and Wireless Communications,(ISSN:2305-0012),pp.93,2014.

[16] Information on: the Executive Yuan: Digital Learning National Technique Plan, Taiwan, pp. 3560-3470, 2013.

[17] C. S. Lin, "The Design and Implementation of Network Learning Environment," Information and Education Magazine, vol. 67, pp. 34-40, 1997.

[18] M. Reid: "Strength and conditioning in tennis Current research and practice," Journal of Science and Medicine in Sport, vol. 11, pp. 248-256, 2008.

[19] D. Karahoca, "Roles of teachers in e-learning How to engage students & how to get free e-learning and the future," Procedia Social and Behavioral Sciences, vol. 2, pp. 5775-5781, 2010.

[20] H. J. Chen, "Linking employees' e-learning system use to their overall job outcomes," An empirical study based on the IS success model, Computers & Education vol. 55, pp. 1628-1534, 2010.

[21] A. G. Dehkordi, "The effect of instructional-aid films on learning of table tennis techniques," Procedia Social and Behavioral Sciences vol. 15, pp. 1656-1662, 2011.

[22] T. F. Wu, S. L Fan and B. S. Chang: "The Policy to Construct the Middle-High School e-Learning Environment," A Competition Perspective. Education Data and Research, vol. 54, pp. 88-95, 2003.

[23] R. Hwung, "Development Strategy of Campus e-Learning," Teacher Word, vol. 179, pp. 26-33, 2004.

[24] S. B. Chang, N. H. Lyau, Y. J. Lee, W. T. Chen, H. W. Jyan and W. Y Wang, "A Research of the Promotion of e-Learning in Lying-Tung Technique University," Tech. Report, The Department of Commercial Technique Management of Lying-Tung Technique University, pp. 569-577, 2007.

[25] Information on: Encyclopedia American Online http://ea.grolier.com.

[26] C. Defta, "Information Security in E-learning Platforms," Social and Behavioral Sciences, vol. 15, pp. 2689-2697, 2011.

[27] K. Dilek, "Roles Of Teachers In E-Learning: How To Engage Students & How To Get Free E-Learning and the Future Procedia," Social and Behavioral Sciences, vol. 2, pp. 5775-5783, 2010.

[28] J. Iskra, J. Gasilervski, J. Hyjek, R. Zajac and M. P. Dyja, "Teaching Methods in Hurdle Races and Results in Special Field and Laboratory Tests," Academy of Physical Education in Katowice, pp. 41-48, 2012.

[29] S. Lee, G.-J. Nam, J. Chae, H. Kimn and A. J. Drake, "Two-dimensional position detection system with MEMS accelerometers, readout circuitry, and microprocessor for padless mouse applications," IEEE Transactions on Very Large Scale Integration (VLSI) Systems, pp. 347-355, 2005.

[30] M. Betke, J. Gips and P. Fleming, "The Camera Mouse: visual tracking of body features to provide computer access for people with severe disabilities," IEEE Transactions on Rehabilitation Engineering Neural Systems and Rehabilitation Engineering, pp. 236, 2002.

[31] T. Itou, Terao M., J. Nagata and M. Yoshida, "Mouse cursor control system using EMG," Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 598, 2001.

[32] C. Dagli and A. Kusiak, "Intellgent system in design and manufacturing," ASME press, pp. 986, 1994.

[33] C. W Siliva and R. Shoureshi, "Intelligent control system," ASME, pp. 8796, 1993.

[34] L. S. Chen, "Step Motor theory and Application," Jeng Shing Book Store, pp. 873, 1991.

[35] P. J Antsaklis and K. M. Passino, "An introduction to intelligent and autonomous control," Kluwer Academic Publishers, pp. 9843, 1993.

[36] F. L. Chen, T. Y. Ou., Grey, "Relation analysis and multilayer function link network sales forecasting model for perishable food in convenience store," Expert Systems with Application, vol. 36, pp. 7054-7063, 2009.

[37] O. G. Ali, S. Sayin, T. Woensel and J. Fransoo, "SKU demand forecasting in the presence of promotions," Expert Systems with Application, vol. 36, pp. 12340-12348, 2009.

[38] A. Allen, "Fashion retail forecasting by evolutionary neural networks," Journal of Production Economics, vol. 114, pp. 615-630, 2011.

[39] S. T. Huang, N. H. Chiu and L. W. Chen, "Integration of grey relational analysis with genetic algorithm for software effort estimation," European Journal of Operational Research, vol. 188, pp. 898-909, 2008.

[40] D. Karahoca, "Roles of teachers in e-learning: How to engage students & how to get free e-learning and the future," Procedia Social and Behavioral Sciences, vol. 2, pp. 5775-5783, 2010.

[41] H. J. Chen, "Linking employees' e-learning system use to their overall job outcomes: An empirical study based on the IS success model," Computers & Education, vol. 55, pp. 1628-1634, 2010.

[42] A. G. Dehkordi, "The effect of instructional-aid films on learning of table tennis techniques," Procedia Social and Behavioral Sciences, vol. 15, pp. 1656-1664, 2011.

Subroutine Entry Point Recognition Using Data Mining

Brian Knudson
Florida Institute of Technology
150 W. University Blvd Melbourne, FL 32901
bknudson2008@my.fit.edu

ABSTRACT

This paper introduces a novel approach to subroutine entry point recognition using data mining. The proposed method applies a Naïve Bayes classifier over features consisting of sequences of normalized disassembled instructions and sequences of preceding bytes. These features combined account for properties of compilers that introduce code at the start of subroutines and padding bytes before the start of subroutines. Experiments were conducted on a dataset consisting of Windows PE32 x86 binaries generated from a collection of small open-source applications for Windows using several compiler settings. Ten-fold cross-validation was applied for training and testing the classifier. The proposed method achieves an average true positive rate 98% with a false positive rate of 0.7% for certain features.

KEYWORDS

Subroutine entry point recognition; data mining; naïve bayes classifier; disassembled instructions.

1 INTRODUCTION

The security field of malware detection incorporates solutions that rely on static analysis of binary programs. Some approaches to static analysis consist of generating a disassembled representation of the program in order to recover instruction and control flow information used to describe its structural characteristics. One of the challenging tasks involved in static analysis on a binary level is the determination of subroutine entry points at which to begin the disassembly process. Compilers often generate similar code at the start of subroutines that can be recognized through data mining techniques.

Functions typically maintain storage and scope for local variables and parameters. Functions also have their own context, which allows them to be called recursively. At an assembly level, functions work within a local stack frame. The context of this stack frame is configured at the start of the function and released upon exit. Compilers often emit similar code that performs these types of actions. Additionally, compilers emit padding in between function boundaries for alignment purposes to improve performance. These properties maintain a degree of invariance that can be classified using data mining techniques.

This paper proposes a data mining method using a Naïve Bayes classifier to detect the entry point of subroutines in binary programs. The features selected for building the classifier consist of both sequences of normalized disassembled instructions at a given program point and the sequence of bytes preceding that program point.

This paper is organized as follows. Section 2 describes the background and relevant work of subroutine entry point recognition on binary programs. Section 3 introduces the proposed method of training a Naïve Bayes classifier over sequences of normalized assembly instructions and byte sequences immediately preceding the program point. Section 4 describes the design and implementation of the system developed to evaluate the proposed method. Section V describes the experiment methodology. Section 6 describes the experiment results. Section 7 discusses the experiment results. Section 8 offers conclusions and suggestions for further research.

2 BACKGROUND

In the domain of static analysis on binary programs, one of the major areas of research is disassembly, which is the "process of recovering a symbolic representation of a program's machine code instructions from its binary representation" [1]. Several techniques exist including flow insensitive, flow sensitive, hybrid, segmented, and heuristic-based.

The flow insensitive approach to disassembly does not follow the control flow of instructions in order to discover additional program locations to disassemble. The most common flow insensitive disassembly algorithm is linear sweep, in which the next program location to disassemble is the address that immediately follows the current disassembled instruction. GNU's *objdump* tool uses this algorithm [2]. Schwartz, Debray, and Andrews [3] improved this algorithm with their extended linear sweep technique that detects jump tables embedded within an instruction stream by leveraging relocation information. The main advantages of this approach are speed, coverage, and algorithmic simplicity. Since control flow is not followed as part of the disassembly process, the upper bound execution time is $O(n)$, where n is the size of the binary program. Since the entire binary is "swept" for instructions, this method produces high instruction coverage. The primary drawback to this method is accuracy. When code and data are interleaved within a binary, data addresses can be disassembled falsely as code, which can disrupt the flow of true instructions. Consequently, this technique provides an over-approximation of the program instructions with the potential for both false positive and false negative errors.

The flow sensitive disassembly approach follows the control flow of instructions in order to recover additional disassembly program locations. The most common algorithm within this approach is recursive traversal, in which all possible branch targets of a control flow instruction are used as additional disassembly points. The IDA Pro disassembler is an industry standard tool that employs this technique [4]. The key advantage of this approach is disassembly accuracy. Since only the instructions that actually flow from program entry points are disassembled, the identified instructions are accurate. However, binary programs may contain code that is only reachable through indirect control flow transfers whose target address is computed dynamically. In this case, flow sensitive disassembly approaches often fail to identify such code. Consequently, this approach yields an under-approximation of the program instructions and suffers from false negative errors. Additionally, the algorithmic complexity of this approach is greater since it constructs a control flow graph with nodes and edges. Care must be taken not to follow control flow loops.

The hybrid disassembly approach combines flow-sensitive and flow-insensitive techniques. Schwartz, Debray, and Andrews [3] developed a technique that performs an extended linear sweep disassembly and then verifies the disassembly results a function at a time using recursive traversal. Functions whose disassemblies do not match are then discarded. This approach combines the strength of coverage from flow-insensitive disassembly and the strength of accuracy from flow-sensitive disassembly. However, the limitation of this particular approach is that function boundaries are discovered primarily using the relocation information present in the evaluated sample binary programs. Not all binary programs contain relocation information, especially malware which can also inject unused false relocation information. Consequently, while this technique can verify the internal consistency of a control flow graph for a portion of the program, it does not necessarily advance the discovery of subroutine entry points.

The segmented disassembly approach conservatively creates disassembly segments, which are sequences of instructions starting at every offset within a program. Segments are then chained together based on control flow. Finally, segments are pruned that contain instructions leading to invalid instructions. This approach used by Kapoor [14] disassembles the entire program,

but it does attempt to identify subroutine boundaries.

Another approach relies on heuristics in order to discover disassembly points. Cifuentes and Fraboulet [5] depend on information stored in the file format of executables such as relocation data and debug symbols to help identify function boundaries. Other work by Jacobson et al. [6] detects and labels library functions within binaries by relying on the presence of export tables to inform the location of subroutines in a binary. The main limitation to these approaches is that malicious programs often omit this type of information in order to thwart analysis. Seminal work by Kruegel et al. [1] describes a methodology for function identification by searching for byte sequences that implement typical function prologues based on instructions that set up a new stack frame. The heuristics for this technique are applied at the disassembly level, which can be used on any program including malware. Once identified, they apply a novel speculative disassembly technique in which they create an over-approximate control flow graph from all possible branch instructions within a subroutine boundary. They apply several pruning techniques to remove nodes and edges that are likely invalid. Finally, any remaining gaps are filled in using disassembled instruction sequences that have a high statistical likelihood of being correct. The advantages of this approach include accuracy and expert knowledge in the form of disassembly heuristics. The main limitation of their technique is that it relies on manually produced expert knowledge of frequently used function prologues signatures.

Additional research has been conducted that leverages data mining and machine learning techniques for malware detection. Wang et al. [7] developed a virus detection system that applied support vector machines over DLL and API dependency features. Work by Christodorescu et al. [8] generated dependence graphs of system calls and applied data mining techniques to extract specifications of malicious behavior. Later, Schulz et al. [9] used Naïve Bayes classifiers to detect malware based on the presence of static features such as strings, byte sequences, and imported DLLs and APIs. Baldangombo et al. [10] extended their work using information gain, principal component analysis, and several classifiers (i.e. support vector machines, J48, and Naïve Bayes) over static features within portable executable (PE) files in order to detect malware.

More relevant work has examined the frequency of opcode-sequences and applied several data-mining algorithms to detect unknown malware [11]. They leveraged the WEKA tool suite in order to develop models based on decision trees, k-nearest neighbor, Bayesian networks, and support vector machines. Rosenblum et al. [12] applied machine learning techniques to detect the toolchain provenance of an executable such as the source language, compiler family, version, and optimization level. They developed features based on short sequences of instructions with wildcards and small, non-isomorphic subgraphs of the control flow graph, where each basic block within the subgraph is assigned a canonical label (color) based on its instruction content. Due to the large number of features, they apply feature selection to choose a subset of features most likely to be valuable for classification. They consider support vector machines and conditional random fields for classifiers.

The most recent work by Bao et al. [13] implements a system called BYTEWEIGHT that introduces an algorithm for automatically detecting function entry points in binaries. Their approach learns function entry point signatures by constructing a weighted prefix tree based on the instructions or bytes present for every point in a program. Each node in the tree is assigned a weight indicating the confidence that the sequence of bytes or instructions representing the path to that node in the tree actually belongs to a function entry point. Their technique achieves a 97% precision and 97% recall over a corpus of 2200 binaries from different operating systems, compilers, and optimization levels. BYTEWEIGHT does not attempt to identify function entry points using other surrounding information such as the bytes immediately preceding a function entry point. Since compilers

position subroutines in regular patterns and apply consistent alignment padding, the bytes preceding function entry points can improve recognition accuracy. Additionally, the training cost for constructing BYTEWEIGHT's weighted prefix trees (implemented as a *trie*) is on the order of 586 hours / 2064 binaries = 17 minutes per binary, which is higher than the proposed method using Naïve Bayes classification.

3 NAÏVE BAYES CLASSIFIER OVER NORMALIZED INSTRUCTION SEQUENCES AND PRECEDING BYTE SEQUENCES

This section describes the features over which a Naïve Bayes classifier is applied for the recognition of subroutine entry points.

3.1 Classification Features

The goal of detecting subroutine entry points is to determine if a given point in a binary program is the start of a subroutine based on the features present at that program point. The first type of feature included is the sequence of bytes immediately preceding a given program point. The second type of feature included is the sequence of normalized instructions starting at the given program point. The combination of these features yields a single feature vector that contributes to the description of a program point.

Let $\mathbf{b} = [b_{-n}, \ldots , b_{-2}, b_{-1}]$ be the sequence of bytes immediately preceding a given program point, where n is the number of preceding bytes and the subscript for each byte in the sequence represents the index of the byte relative to the program point. In some cases, the number of bytes preceding a program point that are characteristic of a subroutine entry point is variable. In order to account for this variance, an instance of $\mathbf{b}$ is generated for each possible sequence up to length n. For example, given the sequence of 3 bytes [0xff, 0x12, 0xab] preceding a given program point, the instances of $\mathbf{b}$ are:

    [none, none, 0xab]
    [none, 0x12, 0xab]

    [0xff, 0x12, 0xab]

Note that the value for bytes omitted for a length less than n are assigned a value of "none". The inclusion of all possible sequences allows for additional weight to be given to sequences that involve multiple characteristic bytes. In the case where a preceding byte does not exist in the binary, a value of "none" is assigned.

Let $\mathbf{i} = [i_0, i_1, \ldots, i_{k-1}]$ be the sequence of normalized instructions starting at a given program point, where k is the number of instructions. Each successive instruction in the sequence is determined by adding the length of the previous instruction to the address of the previous instruction and disassembling the instruction at the new location. A maximum of k sequential instructions are included in $\mathbf{i}$. When an unconditional branch instruction is encountered, no further instructions are added, and the sequence terminates. When an invalid instruction is encountered, a value of "invalid" is assigned to the instruction, no further instructions are added, and the sequence terminates. When the program point to be disassembled does not exist in the binary, a value of "missing" is assigned to the instruction, no further instructions are added, and the sequence terminates.

Variance in the length of characteristic instruction sequences are handled in the same manner as that for byte sequences. An instance of $\mathbf{i}$ is generated for each possible sequence up to length k. For example, given the sequence of 3 instructions ["mov edi, edi", "push ebp", "mov ebp, esp"], the instances of $\mathbf{i}$ are:

    ["mov edi, edi"]
    ["mov edi, edi", "push ebp"]
    ["mov edi, edi", "push ebp", "mov ebp, esp"]

Note that for instances with a length less than k, the other instructions in $\mathbf{i}$ are omitted.

Each instruction is normalized in order to remove references to immediate values or concrete memory addresses. For example, the instruction "add ecx, 5" is converted to "add ecx, *", and

"jmp 0x401000" is converted to "jmp *". Since these values vary significantly across programs, removing them provides a wide degree of signature matching.

Let **x** be a single feature that combines an instance of **b** and an instance of **i** such that $\mathbf{x} = [b_{-n}, \ldots, b_{-2}, b_{-1}, i_0, i_1, \ldots, i_{k-1}]$. The length of **x** will be at most $(n + k)$. For each program point, a set of features is generated for each combination of **b** and **i** at that point, such that the number of features generated is $(n * k)$. For example, given **b** with $n=2$ is [0x1, 0x2] and **i** with $k=3$ is ["inc", "dec", "retn"] for a program point, then the set of features at this program point are:

> [none, 0x2, "inc"]
> [none, 0x2, "inc", "dec"]
> [none, 0x2, "inc", "dec", "retn"]
> [0x1, 0x2, "inc"]
> [0x1, 0x2, "inc", "dec"]
> [0x1, 0x2, "inc", "dec", "retn"]

3.2 Naïve Bayes Classifier

Given a class variable C and a dependent feature vector x_1 through x_n, the Bayes' theorem states:

$$P(C \mid x_1, \ldots, x_n) = \frac{P(C)P(x_1, \ldots, x_n \mid C)}{P(x_1, \ldots, x_n)} \quad (1)$$

Naïve Bayes assumes that each feature vector is independent, which allows the theorem to be rewritten as:

$$P(C \mid x_1, \ldots, x_n) = \frac{P(C) \prod_{i=1}^{n} P(x_i \mid C)}{P(x_1, \ldots, x_n)} \quad (2)$$

Since $P(x_1, \ldots, x_n)$ is constant given the input, the classification rule can be simplified as:

$$P(C \mid x_1, \ldots, x_n) \propto P(C) \prod_{i=1}^{n} P(x_i \mid C) \quad (3)$$

In the case of subroutine entry point detection, C represents the class when a program point is a subroutine entry point. Each x_i is a single feature containing both a sequence of previous bytes and a sequence of normalized instructions.

$P(C)$ represents the probability that a feature instance is associated with a subroutine entry point. When training a Naïve Bayes classifier over a given dataset consisting of feature instances associated with program points taken from a collection of programs, this is the number of feature instances that are *a priori* known to be associated with subroutine entry points divided by the total number of feature instances in the dataset.

$P(x_i \mid C)$ represents the probability that the given feature x_i is a subroutine entry point. When training a Naïve Bayes classifier, this is the number of times this feature occurs in the dataset as part of a program point that is *a priori* known to be a subroutine entry point divided by the number of all feature instances associated with a program point that is known *a priori* to be a subroutine entry point.

Lastly, $P(C \mid x_1, \ldots, x_n)$ represents the probability that a program point is a subroutine entry point given the set of feature instances associated with that program point. The Naïve Bayes classifier is trained over a dataset in order to compute the values of $P(C)$ and $P(x_i \mid C)$ for each feature, which can then be used to compute $P(C \mid x_1, \ldots, x_n)$. This value is a proportional metric that can then be compared against a threshold in order to determine if the program point is classified as a function entry point.

Since the values of probabilities are fractional and multiplied as part of the Naïve Bayes theorem, it is convenient to convert to the log domain and perform calculations there. Consequently, the Naïve Bayes theorem in the log domain is re-written as follows:

$$\log(P(C \mid x_1, \ldots, x_n)) \propto \log(P(C)) + \sum_{i=1}^{n} \log(P(x_i \mid C))$$

$$(4)$$

4 SYSTEM DESIGN AND IMPLEMENTATION

The following subsections describe the system developed to evaluate the effectiveness of the proposed method.

4.1 Program Dataset Generation

The first step in the system is to generate a dataset. In order to provide ground truth for the program points known *a priori* to be subroutine entry points, a collection of open-source C++ applications was collected from the source code package provided by the "Windows Via C/C++" book [15]. This collection contains 26 small C++ application projects intended to teach the Windows Win32 API. Each project was compiled using 8 different optimization options under the Visual Studio 2012 C/C++ compiler. The output of this build process is 304 executable samples including executable images and dlls containing x86 code. For each sample, an associated Program Database (PDB) file was generated containing the symbol information that identifies the name and location of each function in the program. Next, the *TypeInfoDump* utility [16] was used to parse the PDB file and output the entry points of every function within the program for each program sample.

4.2 Feature Generation

The output of the program dataset generation step is an executable file and an associated ground truth file containing the program locations of subroutine entry points for every program sample. The next step in the process is generating feature instances for every program point in each program.

An IdaPython script for the IDA Pro disassembler was written to output feature instances for each binary sample. Configurable parameters are provided to the script, which specify the number of previous bytes n and the number of instructions k to use for features. For each section marked as CODE within the binary, the script performs the following at every offset:

- Generate the sequence of previous bytes **b** for every combination of sequences up to n length

- Decode and normalize the sequence of instructions **i** for every combination of sequences up to k length
- Generate feature instances for every combination of **b** and **i**
- Apply a label of "function" if the program point is a subroutine entry point based on the associated ground truth set of subroutine entry points, else "not_function"
- Assign a unique id composed of the sample program name and program point
- Output record of (unique id, feature instance, label) for every unique generated feature instance. (Note: the "unique" requirement eliminates redundant features, which cannot dominate the dataset.)

The output of the script is stored into a features file. A features file is produced for every program sample. The collection of features files for a given configuration of n and k serve as the feature dataset used during evaluation.

4.3 Classifier Training

The next phase in the process is to train a Naïve Bayes classifier over a collection of training features files. During training, each feature record (unique id, feature instance, label) is read for every training features file. For every unique feature within a record, a count c_f is maintained of all of the occurrences in which its label is "function". Additionally, the count of all records containing a label of "function" c_{f_total} is kept. Also, the total number of records read from the training features files is counted. Once all records have been read, then the probability of each feature with a label of "function" is computed by dividing c_f for that feature by c_{f_total}. After this is complete, the classifier is trained and ready for testing.

4.4 Classifier Testing

The last phase in the process is to test the Naïve Bayes classifier in detecting subroutine entry points over a collection of test feature files. During testing, each feature record is read for every test features file. The set of feature records

with the same unique id are collected together and serve as the feature instances that characterize a single program point. The probability that this program point is a subroutine entry point given its set feature instances is computed according to the Naïve Bayes theorem using the probabilities in the trained Naïve Bayes classifier. For each feature instance, the probability its feature is assigned the label "function" is retrieved from the trained Naïve Bayes classifier. This probability is converted into the log domain for convenience. The sum of this value for all feature instances is computed. This value is added to the log of the probability of finding a feature instance with the "function" label within the training dataset, which is just the number of training feature instances with the "function" label divided by the total number of training feature instances.

The result of this sum is a proportional measure of the probability that the given program point is a subroutine entry point. If this value exceeds a specified threshold, then the given program point is classified as a subroutine entry point. Since the ground truth label for the given program point is provided as part of its feature records, the classification can be evaluated as a true positive (TP), true negative (TN), false positive (FP), or false negative (FN).

For each program point in the test feature files, the count of true positives, true negatives, false positives, and false negatives are captured. From these counts, two evaluation metrics are computed:

True Positive Rate (TPR) = TP / (TP + FN) (5)
False Positive Rate (FPR) = FP / (FP + TN) (6)

These metrics are used to evaluate the performance of the classifier for a given threshold t and pair n, k. Values for TPR and FPR are computed for several values of t in order to generate a receiver operating characteristic (ROC) curve with TPR plotted along the vertical axis and FPR plotted along the horizontal axis.

5 EXPERIMENT METHODOLOGY

In order to evaluate the effectiveness of the proposed method, a 10-fold cross-validation method was applied over the set of 304 features files. In this method, the 304 feature files are separated into ten partitions. One partition is selected as the test set, and the other nine partitions are selected as the training set. The classifier is trained over the training set and evaluated over the test set to yield a pair of values for TPR and FPR. This process is repeated selecting a different partition for the test set until all partitions have been analyzed. The output pairs of TPR and FPR values are averaged in order to generate a fair evaluation of the classifier.

For a series of different threshold values t, the evaluation process above is repeated in order to generate a set of TPR and FPR pairs that can be plotted on an ROC curve. Lastly, the following combinations of n and k were evaluated to demonstrate the effect of changing the feature sizes on performance:

- n=1, k=5
- n=2, k=5
- n=3, k=5
- n=1, k=6
- n=1, k=7
- n=1, k=8

6 EXPERIMENT RESULTS

Figure 1 is a graph comparing the performance of each experiment combination in the form of a ROC curve.

Table 1 shows the execution performance for each experiment combination based on how long it took to generate the feature files and how long it took to complete a single cross-validation test.

All experiment combinations achieve an optimal TPR between 97% and 98% and FPR between 0.5% and 0.9%.

Table 1. Execution Performance

n,k	Feature Generation (min)	Cross-Validation (sec)
n=1,k=5	24	22
n=2,k=5	26	40
n=3,k=5	36	62
n=1,k=6	19	21
n=1,k=7	25	24
n=1,k=8	30	29

7 DISCUSSION

Based on the experimental results, subroutine entry point recognition is successful using variable length features of normalized instructions and preceding byte sequences, achieving an average true positive rate of 97% and false positive rate of 0.5%.

The time complexity of the algorithm is dependent on the number of feature instances generated per program point, which is directly computed from the values of n and k. While the time to generate a single feature is constant, the time complexity for generating all features for a single program point is $O(n*k)$. Given l program points, the overall processing time of the algorithm is $O(l*n*k)$. Cross-validation requires computing probabilities for each feature instance. Therefore, its time complexity is also $O(n*k)$. The execution performance measurements of the experiment results confirm the complexity estimation, showing a linear increase with respect to (n * k).

An interesting observation is that the accuracy of the classifier remains relatively constant regardless of the length of preceding bytes n or the number of instructions k. Consequently, selection of a lower value of ($n*k$) can significantly reduce overall execution time without a loss in recognition accuracy.

The longest experiment combination generates features and performs cross validation within 46 minutes (36 minutes + 10 * 62 seconds, where 10 is the number of cross validations) over 304 sample binaries, which yields a unit training cost

of 9 *seconds* per sample. When compared against the execution performance of BYTEWEIGHT (which has a training cost of 17 *minutes* per sample), the training cost of the proposed method is 99% less than that of BYTEWEIGHT. Since the accuracy of the proposed method is comparable to (and even slightly higher than) that of BYTEWEIGHT, training cost and execution time is the distinguishing merit of the proposed method.

8 CONCLUSION AND FUTURE WORK

This paper proposed a new subroutine entry point recognition algorithm using a Naïve Bayes classifier over features consisting of normalized assembly instructions and bytes preceding the program point. The experimental results demonstrate that the proposed method achieves a nearly perfect detection rate with a negligible false positive rate. Thus, it is concluded that this method can be used for certain subroutine entry point applications.

The proposed method would benefit from future work in extending the dataset to include samples generated from different compilers and operating systems. Additionally, the effect of increasing the value of n greater than 3 should be examined in combination values of k greater than 5. Another area of research is the inclusion of address alignment information, which is a property typically introduced by compilers.

REFERENCES

[1] Kruegel, C., Robertson, W., Valeur, F., & Vigna, G. (2004). Static disassembly of obfuscated binaries. USENIX security Symposium.

[2] Foundation, F. S. (2002, March). GNU Binary Utilities. Retrieved from http://www.gnu.org/software/binutils/manual

[3] Schwarz, B., Debray, S., & Andrews, G. (2002). Disassembly of executable code revisited. Proceedings of the 9th working conference on reverse engineering. IEEE.

[4] Hex-Rays. (2014). IDA. Retrieved from Hex-Rays: https://www.hex-rays.com/products/ida/

[5] Cifuentes, C., & Fraboulet, A. (1997). Intraprocedural static slicing of binary exectuables. Proceedings of the

International Conference on Software Maintenance (pp. 188-195). IEEE.

[6] Jacobson, E. R., Rosenblum, N., & Miller, B. (2011). Labeling library functions in stripped binaries. Proceedings of the 10th ACM SIGPLAN-SIGSOFT workshop on Program analysis for software tools (pp. 1-8). ACM.

[7] Wang, T. Y., Wu, C. H., & Hsieh, C. C. (2008). A virus prevention model based on static analysis and data mining methods. IEEE 8th International Conference on Computer and Information Technology Workshops (pp. 288-293). IEEE.

[8] Christodorescu, M., Jha, S., & Kruegel, C. (2008). Mining specifications of malicious behavior. Proceedings of the 1st India software engineering conference (pp. 5-14). ACM.

[9] Schultz, M., Eskin, E., Zadok, S., & Stolfo, S. (2011). Data mining methods for detection of new malicious executables. Proceedings of the 2001 IEEE Symposium on Security and Privacy (pp. 38-49). IEEE.

[10] Baldangombo, U., Jambaljav, N., & Horng, S.-J. (2013). A Static Malware Detection System Using Data Mining Methods. International Journal of Artificial Intelligence and Applications, 113-126.

[11] Santos, I., Brezor, F., Ugarte-Pedrero, X., & Bringas, P. G. (2013). Opcode sequences as a representation of executables for data-mining-based unknown malware detection. Information Sciences, 64-82.

[12] Rosenblum, N., Miller, B., & Zhu, X. (2011). Recovering the toolchain provenance of binary code. Proceedings of the 2011 International Symposium on Software Testing and Analysis (pp. 100-110). ACM.

[13] Bao, T., Burket, J., Woo, M., Turner, R., & Brumley, D. (2014). BYTEWEIGHT: Learning to recognize functions in binary code. SEC'14 Proceedings of the 23rd USENIX conference on Security Symposium, 845-860. ISBN 978-1-931971-15-7.

[14] Kapoor, A. (2004). An approach towards disassembly of malicious binary executables. Doctoral dissertation. University of Louisiana at Lafayette.

[15] Richter, J., & Nasarre, C. (2013). Windows Via C/C++. Retrieved from Wintellect: http://www.wintellect.com/DevCenter/wp-content/uploads/2014/10/Windows-via-CPlusPlus-Code.zip

[16] Starodumov, Olen. (2004). TypeInfoDump. Retrieved from http://www.debuginfo.com/download/typeinfodumpfull.zip

Figure 1. Experimental Performance Comparison ROC Curve
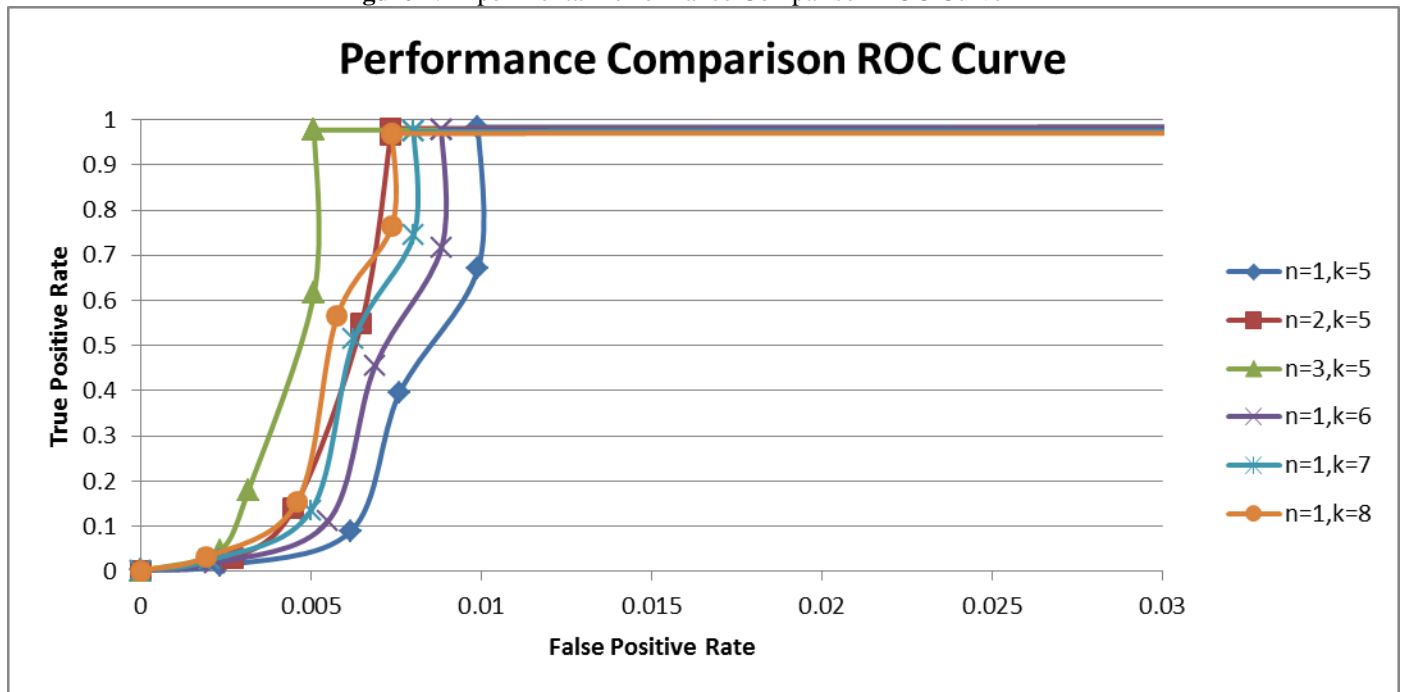
Lyapunov Stability Analysis for Consequential Message Path in Graded Cognitive Networks

Kavitha Sooda[1,2, †], **T. R. Gopalakrishnan Nair**[3,4]

[1]Advanced Networking Research Group (RIIC), DSI
and
[2]Associate Professor, Department of CSE
Nitte Meenakshi Institute of Technology
Bangalore-560064, India.
E-mail: kavithasooda@gmail.com

[3]Saudi ARAMCO Endowed Chair – Technology, PMU, KSA
and
[4] Advanced Networking Research Group
VP, Research, Research and Industry Incubation Center (RIIC), Dayananda Sagar Institutions
Bangalore-560078, India.
E-mail: trgnair@ieee.org, www.trgnair.org

ABSTRACT

In the domain of network routing, one of the key measure belongs to the stability analysis of the paths obtained by routing algorithms. Lyapunov analysis has been used widely for understanding the nature of the system output to know the stability intensity. Here in this article, we have compared the performance of two bio-inspired algorithms on a graded surface while determining the path. Also the study of the Lyapunov analysis is carried out on a restrcited search space. Here the Lyapunov function derived are best suited for the path derived on graded surface by Memetic and Artificial Bee Colony algorithms. Results show that the path obtained by the bio-inspired algorithms has the required stability as found by Lyapunov function for the transfer of data.

KEYWORDS

Artificial Bee Colony, Bio-inspired algorithms, Graded network, Intelligent routing, Lyapunov functions, Memetic, Reduced search space.

1 INTRODUCTION

The quality of the selection of path in current network depends on parameters like Bandwidth and Delay for data-forwarding purpose. This would involve hello packets to be transmitted to get the information of the required Quality of Service (QoS) parameters the user is looking for. The key necessary factor in today's Internet is to ensure that the nodes have knowledge of its environment. This has become an essential parameter as inter-networking is evolving to complex structures. This formation of complex structure is due to dynamic nature of the environment. Diagnosing failure of forward channel is difficult in such an environment. Thus, we need to ensure knowledge availability in the environment. This can be made possible if the topology of our interest can learn and reason out how to handle forwarding of packets. The advantage of this knowledge would be in the quality selection of the path. The literature shows that many progressive works has been carried out in this direction. One of the prominent approaches that had proved its proximity in achieving this is the nature-inspired algorithm that has facilitated to learn and adapt based on topological formation and emergent conditions [1]. Intelligent networks [2] require an approach, whereby each node cooperates with the data-distribution process and

makes use of an information-rich context. To get intelligent network working, there is a need to rethink on the architecture and protocols of the components in the global communication infrastructure. A prominent research direction orients its approach to mimic nature-like mechanisms to realize smarter communication networks that can make sense of the hidden communication patterns and do self-regulation within the topology for achieving intelligent routing. The routing approaches for intelligent routing are based on learning and reasoning techniques. This will require a database to store the learnt aspect, which can be referred to later in the future when necessary. The learnt aspect that will be utilized in future is based on the reasoning methods that will justify the result obtained.

This article presents the implementation of a novel idea of applying intelligence, which enables the nodes to make decisions and carry out the routing of packets. This intelligence at the node level makes the routing efficient and better [3]. A geographical search approximation is done using a cone, where we apply the routing algorithms on reduced search space and apply reasoning to derive the final path. Results show that the grading approach along with Artificial Bee Colony (ABC) and Memetic approaches has achieved the required routing path with promising speed [4-5]. The result thus obtained is verified by Lyapunov function to verify the paths obtained.

The rest of this article is organized as follows: in Section 2, related literature is described and in Section 3, the reasoning method is discussed in detail. Graded networks and Lyapunov analysis is dealt in Sections 4 and 5. Simulation results and analysis are presented in Section 6. Section 7 concludes.

2 RELATED WORK

The ability to be aware of network operations and subsequently adjust the operational parameters according to the needs of the scenario is referred to as cognitive network. It not only behaves based on a reference to an active network, but also includes an adaptation and learning technique [6-7], which makes the process different. Cognition was conceptualized by Mitola and later the idea of a feedback loop was derived from it. With the advancement in the field of cognition, the autonomic approach [8] has become more evident to solve a number of network-related issues. Some of the applications of cognitive approaches have been discussed below.

2.1 Autonomic network

An autonomic network is applicable in a situation where path determination is more predictable, and the search space can be handled well with an algorithm. When the environment is large, then cognitive approach will play a role in learning and reasoning. IBM is one of the organizations that proposed the autonomic networking concept. It described agents in three dimensions. The first dimension is the agency where it determines the degree of autonomy and authority assigned to the agent. The second dimension is about intelligence, describing the reasoning and learned behaviour. The third dimension is the mobility where the agents travel through the network. The current routing operates based on the packets received. This type of operation is asynchronous and is loosely coupled with a complex network. This satisfies the lowest degree of autonomy [9-10].

2.2 Motorola FOCALE

The autonomic management platform developed by Motorola is referred to as Foundation–Observation–Comparison–Action–Learn–rEason (FOCALE) for core networks [11]. This development was based on the human nervous system that performs unconscious actions. The system actually tries to figure out functions not known to the human. This system contains all the elements of cognitive approaches. It has two control loops, one for maintaining the current state and the other for reconfigurations. The idea behind this development was to consider the business-level requirement in simple language. It is later transferred to the autonomic configured network where it gets converted to a form that would sense and act according to the requirement of the end user.

2.3 Software-defined networking

The technique used in software defined networking (SDN) is based on a centralized controller that directly dictates where the packet should move [12-13]. There is support for multi-switch updates, which makes it strong without asking the user to program it with low-level commands. The user need not worry about the step-by-step updates. But programmers who use this must be able to distinguish among the multiple abstract updates that have been useful. This will definitely require a consistency class, mechanism and language to code for proper set-up of SDN.

3 IMPORTANCE OF REASONING

A cognitive network is a network with intelligent process that can perceive current network conditions, and then plan, decide and act on those conditions. The network can learn from these adaptations and use them to make future decisions, while taking into account end-to-end goals. The perception, decision, action and learning aspects of this definition are all elements of learning and reasoning. These capabilities are essential to ensure that the decisions made by a cognitive network will improve the network performance as measured against the end-to-end goals. Graded reasoning method is a promising approach for realizing the optimal path determination in the cognitive network. In this context, the estimation of grades at the nodes based on the reasoning technique has been a challenge for a long time.
This work deals with a novel estimation of reduced search space using a graded technique on cognitive network [14-15] using a reasoning mechanism. The transition from packet switching routing towards self-independent routing has been the research focus of the current decade. The realization of the self-independent nature of routing has been evolved through the cognitive approaches. Current routing schemes lack the ability to identify the relationship between the environment awareness and reasoning of the path determined by the technique. Thus, cognition plays a vital role in the process of remembering and formation of the path in Internet routing. The present article proposes and discusses in detail how an optimal path can be determined for graded cognitive network that has information about the environment [16-19].

3.1 Requirement for reasoning

Learning and reasoning play an important role in the cognitive network. Reasoning helps in the immediate decision process using the historical knowledge as well as the current state of the network. The primary goal of reasoning is to choose a set of actions. Learning is a long-term process that is based on the accumulation of knowledge on the perceived result of past actions. Cognitive network nodes learn by enriching the Information Base so that efficacy of reasoning improves in the future.

In general, any cognitive network structure will require some assistance to do a certain amount of job for accounting for an optimal solution. This is where agents will assist the network in doing the job. An agent [20-21] in general will work on a particular problem space, where it tries to achieve its goal and remember it for future purposes. To take proper action, it looks into what was learnt and also the current state to take a decision based on belief and goal.

The agent extracts the required parameters from Information Base and operates on the network topology as shown in Figure 1. The input to this network topology is the desired source and destination that comes as an instruction. The network topology carries out the reasoning to obtain the optimal path and to choose an action to be executed in the environment. The reasoning process is driven by the data structures in the Information Base.

In this article, once the awareness of the environment is learnt, and nodes have been graded and two algorithms have been used to determine the final path. First, ABC algorithm is used as developed by Karaboga [22], which was motivated by the intelligent behaviour of the bees.
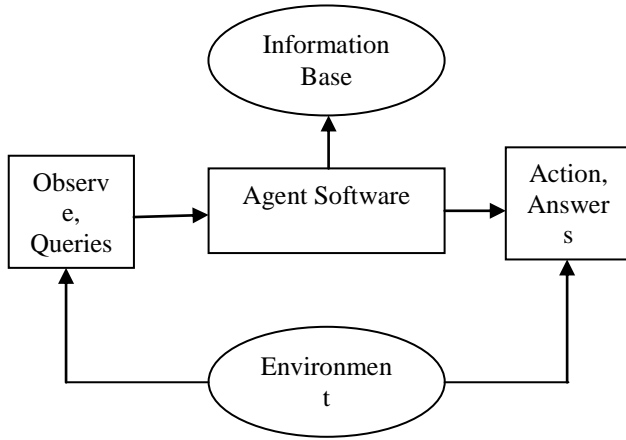
Figure 1. Standard agent model

It is an optimization tool, where search is performed in a distributed environment. Second, hybrid evolutionary algorithms that are commonly known as Memetic Algorithms (MAs) or genetic local search [23-25] are used. Such methods have been demonstrated to converge to high-quality solutions more efficiently than their conventional counterparts.

4 GRADING THE NETWORK

A grade is like an index, which represents the node QoS parameter-satisfied status. The status helps in estimating the selection of the node quickly once the network is realized. The grade technique has been applied segment wise. Here in each segment there exist thinking processes that run a performance function for determination of homogeneity among the nodes. This thinking process will have the awareness of the complete segment. Once the data have reached their end points in one segment, it will be handed over to the next segment for forwarding the data. Now the thinking process in this new segment will try to compare its evaluation of the nodes with the previous segment to see how the homogeneity property persists among the nodes, across the segments. This evaluation helps in the recognition of familiar pattern that would have been already encountered before. This qualifies for recognition ability from the learnt knowledge of the nodes. The parameters considered are based on bandwidth of the link, buffer resource allocation,

delay and packet loss. Based on each of these contributing factors, the grade estimation is calculated based on the lumped parameter equation, which will assign a value to the nodes associated with the metric chosen. The nodes participating in routing are studied under various parameters and evaluated based on Equation 1. The grade value (GV) obtained by this equation for a given node is thus stored in the memory of the node for further reference.

$GV = a_1 *$ contribution from B $+ a_2*$ contribution from De $+ a_3*$ contribution from RA $+ a_4 *$ contribution from PL, i.e.,

$$GV = 0.6* B + 0.2 *De + 0.1* RA + 0.1* PL \quad (1)$$

where,

B = bandwidth of a link; De = delay;
RA= resource availability and PL = packet loss.

Higher GV indicates better satisfaction of the parameters considered. Figure 2 depicts a pictorial representation of the GV assigned on the nodes in the search space. A GV of '5' is represented as red colour, GV of '4' is represented as green, GV of '3' is represented as blue, GV of '2' is represented as black and GV of '1' is represented as white colour.
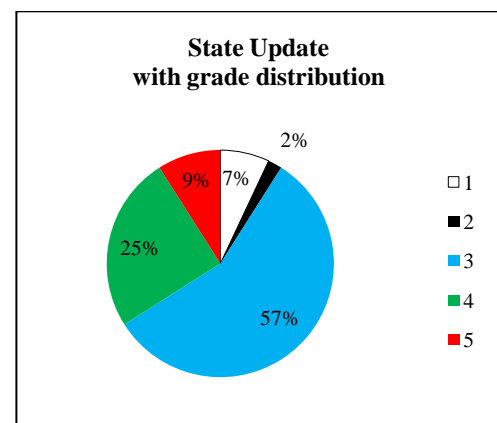
Figure 2. Nodes grade index distribution in set-up

The GV for a node is assigned based on the following criteria as depicted in Table 1.

5 LYAPUNOV STABILITY

Lyapunov stability analysis plays an important role in the stability analysis of control systems usually described by the state space equations for nonlinear systems. A Lyapunov function is that which allows one to deduce stability. There are two methods for stability analysis used by Lyapunov approach. First method is the Lyapunov Linearization that is applied for local stability in small range motion. This method fails to conclude about marginal stability. The second method overcomes this shortcoming and is used when it does not require the solutions of the differential equations for analysis. This method is called as the direct method.

Table 1. Grade Value (GV)

Levels	GV
All conditions satisfied	5
Only delay exists	4
Resource alone not satisfied	3
PL alone not satisfied	2
No bandwidth available/two or more QoS not satisfied	1
None of the parameters satisfied	0

Here, in this article, we develop the equations using the direct method. The design of the network topology is not explicitly dependent on time. Therefore, our system is a nonlinear system that is an autonomous system [26-30].

5.1 Stability of a system with inputs

Stability theory addresses the stability of solutions of differential equations and of trajectories of dynamical systems. If a small change happens in the hypothesis it leads to small variations in the conclusion of the theorem. We need to specify the metric used for determining the stability. In dynamical systems, an orbit is said to be in Lyapunov stable state if the forward orbit does not deviate more. Under favourable conditions the problem can be formulated using eigenvalues of matrices. A more general method involves Lyapunov functions. Stability of the system can be proven by various criteria. Asymptotic properties for differential equations and dynamical systems play an important role in determining what happens to the system after a long time. Stability in this article indicates the ability of the system to sustain quality message paths over a period of time ensuring lesser hop count on the path finalized for forwarding the data packets.

A system with parameters is written as
$$\dot{x} = f(x, p) , \qquad (2)$$

where p is the parameter that controls the function f's output. One needs to quantify the effect of inputs for considering the stability of the system. We have input-to-state stability mechanism for nonlinear systems for analysis purpose.

A nonlinear system does not satisfy the property of superposition and homogeneity and possess many equilibrium points. Lyapunov analysis is the best when we try to understand the system with respect to equilibrium points. Lyapunov function gives the information about the location of the limit set. It is nothing but the information about the asymptotic behaviour. Here in this article we have described three levels of asymptotic classification, namely low, medium and high. Lyapunov's method helps in determining the stability of the system without solving system equations. The functions in quadratic form give better understanding of Lyapunov equations.

The following are the properties of Lyapunov function, V:
- V is continuous.
- V has a unique minimum with respect to all other points in some neighbourhood of the equilibrium of interest.
- Along any trajectory of the system, the value of V never increases.

Here the stability of the working structure is classified as high, neutral or low (unstable) stability. These values help us in justifying the output as the output depends on the QoS metric values.

The input for the Lyapunov is an ordinary differential equation. Consider DE of the form

$$\frac{du}{dt} = f_L(u) \tag{3}$$

where $f: R \to R$ is a continuous differentiable function. A point u^* is called a fixed point of the DE if $f(u^*) = 0$. The variation equation of $du/dt = f(u)$ is given by

$$\frac{du}{dt} = (\frac{df_L}{du(u(t))})y \tag{4}$$

where it is assumed that f is continuous differentiable.

Here we have assumed that the destination nodes act as the equilibrium points. An equilibrium state x_e, of the system is said to be asymptotically stable if it is stable in the sense of Lyapunov and if every solution starting within $S(\delta)$ converges, without leaving $S(\epsilon)$, to x_e as t increases indefinitely.

5.2 Lyapunov equation for the system

Let $X^* = 0$ be an equilibrium point of V(X). Let V:D→R be a continuously differentiable function such that:
(i) V(0) = 0
(ii) V(X) > 0, in D−{0} means positive definite
(iii) $\dot{V}(X) < 0, in\ D - \{0\}$ means negative definite
Then $X = 0$ is asymptotically stable.
Consider the quadratic form of Lyapunov equation which is given as follows:
QF : Rn→R

$$QF(x) = x^T P x \tag{5}$$

where P is a symmetric matrix where $P = P^T = [P_{ij}]$.
It may be shown that a quadratic function QF is positive definite if all the eigenvalues of P are positive.
The stability model for the proposed work, which indicates the global asymptotical stability, is as follows:

$$\dot{x}_1 = ax_1 \tag{6}$$
$$\dot{x}_2 = -bx_2 \tag{7}$$

Here x_1 represents the bandwidth, which is a positive effect as it increases and x_2 represents the delay which is a negative effect as bandwidth decreases.

$$\frac{\partial V}{\partial X} = g(X) = \begin{bmatrix} K_1 & K \\ K & K_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{8}$$

$$\frac{\partial g_1}{\partial x_2} = \frac{\partial g_2}{\partial x_1} = K \tag{9}$$

Consider $K = 0$. Then,

$$g(X) = \frac{\partial V}{\partial X} = \begin{bmatrix} g_{1(X)} \\ g_{2(X)} \end{bmatrix} = \begin{bmatrix} K_1 x_1 \\ K_2 x_2 \end{bmatrix} \tag{10}$$

Here $K_1 = a$ and $K_2 = b$.

$$V(X) = \int_0^{x_1} g_1\ (\overline{x}_1,0)d\overline{x}_1 + \int_0^{x_2} g_2\ (x_1,\overline{x}_2)d\overline{x}_2 \tag{11}$$

$$= \frac{1}{2}(K_1 x_1^2 + K_2 x_2^2) \tag{12}$$

If $K_1 > 0$ and $K_2 < 0$, then V(X) is Lyapunov function. Therefore,

$$\dot{V}(X) = g^T(X)f(X)$$
$$= [K_1 x_1\ K_2 x_2].\begin{bmatrix} ax_1 \\ -bx_2 \end{bmatrix} \tag{13}$$
$$= a\ K_1 x_1^2 - bK_2 x_2^2 \tag{14}$$

If $K_1 < 0$ and $K_2 > 0$, then $\dot{V}(X)$ is negative definite.

Here we superimpose the f(destination node) = 0 as it is the point we are trying to reach from source. The following are the steps involved in obtaining the optimal path and determine its stability:
Step 1: sub_path = source node
Step 2: $f(x) = V(x_1, x_2) = Ax_1^2 + Bx_2^2$, $\tag{15}$
where A and B are the GV obtained for each node. For a node to be selected it should have satisfied the grade property parameters. Here we differentiate twice to get the GV. A is associated with Type-I check and B is associated with Type-II check.
Step 3: sub_path = sub_path + $f(x)$. Here + represents the concatenation of two nodes that have been selected after considering the GV.
Step 4: now choose the node x obtained from $f(x)$ and go to Step 2, if x is not the destination node, else go to Step 5.
Step 5: final path obtained, therefore we end.
We have considered $k_1 = 0.6$ and $k_2 = 0.2$. We have done Lyapunov calculation for only bandwidth

and delay parameters, and therefore the GV has been reassigned as shown in Table 2.

Table 2. Reassigned GV for Lyapunov calculation

B	De	SL
H	H	2
H	L	3
L	H	0
L	L	1

We substitute in Equations 6 and 7, for x_1 and x_2 depending on the criteria met by the nodes from Table 2. Consider the simulation run for 15 nodes as in Table 3. We have Lyapunov value defined for the same as in Table 4.

Table 3. Sample from simulation

Node number	RA	Neighbour node	B	De	PL
0	35	1	38	4	0
		6	21	2	1
		8	21	6	0
		10	38	6	0
1	38	0	38	4	0
		4	3	5	0
2	22	9	12	6	1
		10	11	4	1
3	22	10	11	4	1
4	36	1	3	5	0
		5	21	2	1
		7	18	2	0
		14	7	6	1
5	35	4	21	2	1
6	35	0	21	2	1
7	39	4	18	2	0
		11	16	4	0
8	39	0	21	6	0
		10	16	4	0
		13	33	3	0
9	24	2	12	6	1
		12	33	3	0
10	35	0	38	6	0
		2	11	4	1
		3	11	4	1

		8	16	4	0
11	35	7	16	4	0
12	36	9	33	3	0
13	36	8	33	3	0
14	19	4	7	6	1

Table 4. Lyapunov calculation for the sample of 15 nodes

Path	GV	Vx	$\dot{V}(X)$
0-8-10	4-4	0.5(0.6*2*2-0.2*2*2)+0.5(0.6*2*2-0.2*2*2)=1.6	(-0.6*0.6*2*2-0.2*0.2*2*2)+(-0.6*0.6*2*2-0.2*0.2*2*2)=−3.2
0-10	4	0.5(0.6*2*2-0.2*2*2)=0.8	(-0.6*0.6*2*2-0.2*0.2*2*2)=−1.6

In Vx calculation, we consider a as positive and b as negative and hence we have obtained the value. In $\dot{V}(X)$ calculation, a and k_1 are same and b and k_2 are of same value. Here k_1 is < 0 and hence has a negative value associated with it.

6 SIMULATION RESULTS

The simulation studies were carried out extensively using a Java-built simulation bed. In the simulation system developed for this total model of cognitive network, in any run, the user enters the number of nodes to be built for a random run. Depending on the number entered by the user, the nodes are spread across the region of view and links get generated and learning is performed as shown in Figure 3. The required QoS parameter is also assigned randomly.

We enter the source and destination node for determining the path. The learn button helps in remembering the path already executed on the run. This button also carries out grading of the node, which signifies the quality of the node. This information is useful for future, i.e. if such a pair of source and destination node is again looked for later, then the path is not computed by the
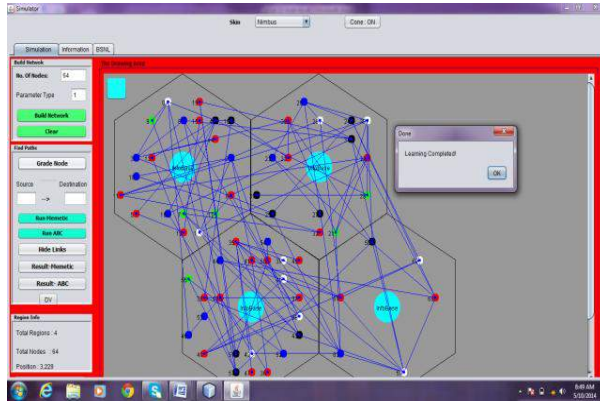
Figure 3. Distribution of nodes and links along with learning

algorithm, instead the path is determined from the Information Base. We choose the desired algorithm for viewing the output that uses the help of logical reasoning for the graded node.

The state updates were calculated depending on the programmable frequency factor like once in 30 min or once in 60 min. The whole network is analysed in abstract mode, hence there is no need for a real-time coupling of the system. Here the QoS parameters are learnt and updated based on the consumption of the resources during the simulation run. The initial state update is shown in Figure 1, and the final state update is shown in Figure 4.

Figure 4. Final state update

Here the data is assigned based on random function. Table 5 gives a detailed description of the experimental runs carried for node numbers

varying from 32 to 1024, which are all to the power of 2.

Table 5. Fitness value

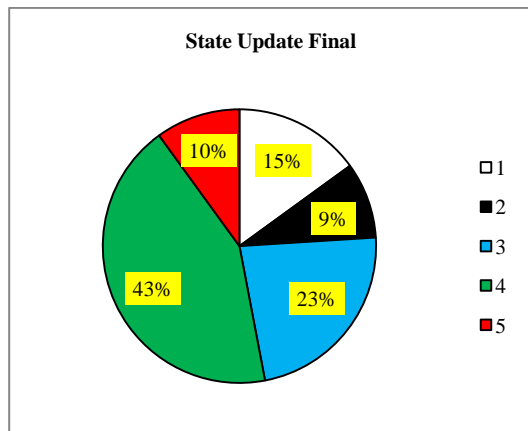|	Memetic		ABC	
No. of nodes	Fitness value	Bandwidth	Fitness value	Bandwidth
32	0.9	15	0.9	15
64	0.74	102	0.74	102
128	0.70	234	0.70	234
256	0.76	125	0.81	110
512	0.55	376	0.65	220
1024	0.54	717	0.72	534

Table 5 shows how the two algorithms perform on the graded network. The total throughput chosen by Memetic is equal or more than all six cases as compared to ABC.

Using the cone approach we reduce the search space. Each region has one agent that collects information about the nodes belonging to the region and also assigns the GV to each node. We observed that the simulation set-up time using cone is higher than without using the cone. This is because learning aspect takes place within the cone. The learnt knowledge is made available in the Information Base which is nothing but the agent. The stability of the output path obtained is justified by the GV as well as the Lyapunov value as shown in Table 6.

Table 6. Lyapunov value and grade equation value

$V(X)$	$\dot{V}(X)$	Average grade	GV
1.6	−3.2	4	0.8
5.2	−10.4	4	0.85
8.6	−17.2	3	0.828
6.2	−12.4	4	0.9
19.4	−38.8	4	0.938
26.2	−52.4	4	0.89

It was observed that longer paths had higher $V(X)$ than the shorter paths. The GV helps in classifying

the stability of the path. The graphs in Figures 5 and 6 show the hop count obtained when MA and ABC was carried out with cone and without cone. With respect to hop count, the cone actually helped in obtaining shorter path with required quality of the complete path.
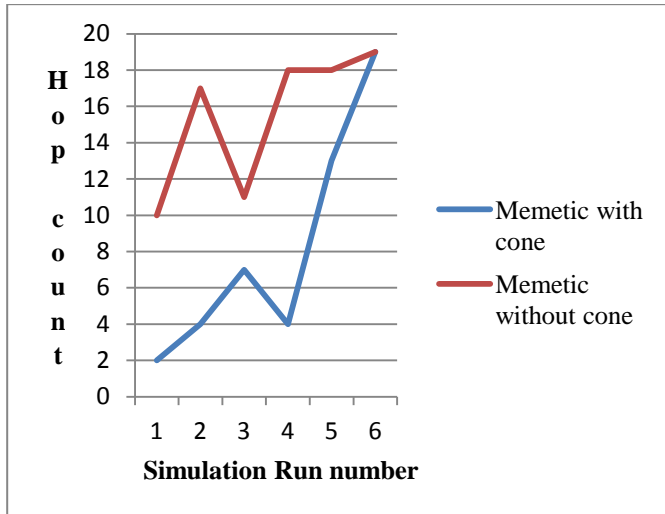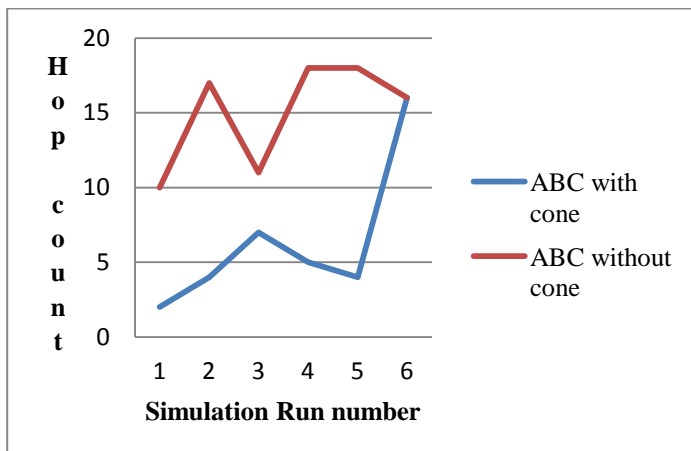

Figure 5. MA – hop count comparison

Figure 6. ABC algorithm – hop count comparison

A complete result of different run is provided in the Appendix as shown in Table A.1 and A.2. The initial assignment range of values assigned to the nodes which have been taken from the BSNL provider is described in Table A.3. For all runs the packet size is considered as 1Mb. The run for 32 nodes is captured and is as shown in Figure A.1. The values assigned for the QoS parameter in the simulation is provided in Table A.4. the Result

obtained for the 32 node run is as given in Table A.5 and A.6.

7 CONCLUSION

In this article, we have presented the successful implementation of optimal path determination by accounting the quality levels of the nodes and the way in which they are getting graded over the network, bringing in self-awareness concepts to network. This technique of obtaining an estimation of the quality of the nodes and its nature of distribution is realized through parametric grading. It presented the various modes of grading with state update process for maintaining apt cognition. Further on this set-up having cognitive ability, we reduce the search space by creating a geographical focus cone directed to the destination. The optimal path was obtained based on the nature-inspired algorithms. Further evaluation whether to calculate the path or not, is determined by the Information Base that remembers the path obtained earlier.

The results show that ABC algorithm outperformed the Memetic technique in terms of throughput and fitness for optimal path determination. Also it has shown the advantage of reduction in geographical search space to realize lesser hop count. The suggested algorithms are shown to be successful by the simulation. The results are justified by Lyapunov values which gives the proof that the obtained paths are of required quality.

8 REFERENCES

1. El-Abd, M.: Performance assessment of foraging algorithms vs. evolutionary algorithms. Information Sciences 182, 243–263 (2012).
2. III, M., Joseph.: Cognitive Radio – An Integrated Agent Architecture for Software Defined Radio (Ph.D. Dissertation), Kista, Sweden: Royal Institute of Technology (2000).
3. Hou, R., Lui, K., Leung, K., and Baker, F.: Performance analysis of quantization-based approximation algorithms for precomputing the supported QoS, Journal of Network and Computer Applications, Vol. 40, April , pp. 244–254 (2014).
4. Karaboga, D., Basturk, B.: On the performance of artificial bee colony (ABC) algorithm, Applied Soft Computing, vol. 8, pp. 687–697 (2008).

5. Tereshko, V.: Reaction–diffusion model of a honeybee colony's foraging behaviour, in: M. Schoenauer (ed.), Parallel Problem Solving from Nature VI, Lecture Notes in Computer Science, vol. 1917, Springer–Verlag, Berlin, pp. 807–816 (2000).

6. Thomas, R., W., DaSilva, L., A., Allen, B., Mackenzie.: Cognitive Networks. :Proceedings of the First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 352–360 (2005).

7. Fortuna, C., Mohorcic, M.: Trends in the development of communication networks: Cognitive networks. Journal of Computer Networks 53, 1354–1376 (2009).

8. Nair, T., R., G., Sooda, K.: An intelligent routing approach using genetic algorithms for quality graded network, International Journal of Intelligent Systems Technologies and Applications, vol. 11, 196–211., (2012).

9. Kephart, J., O., Chess, D., M.: The Vision of Autonomic Computing, IEEE Computer Society, 36(1), 41–50 (2003).

10. Raymer, D., Meer, S., Strassner, J.:From Autonomic Computing to Autonomic Networking: an Architectural Perspective, 5th IEEE Workshop on Engineering of Autonomic and Autonomous Systems, 174–183 (2008).

11. Konstantinou, A. V.: Towards Autonomic Networks, Ph.D. Thesis, Columbia University (2003).

12. Sooda, K., Nair, T., R., G.:. Optimal path selection in graded network using Artificial Bee Colony algorithm with agent enabled Information. 12th IEEE International Conference on Hybrid Intelligent Systems, Pune, 4–7 December (2012).

13. Reitblatt, M., *et al*.: Consistent Updates for Software-Defined Networks: Change You Can Believe In!, Hotnets '11, ACM, 2011.

14. Nair, T., R., G., Sooda, K.: Application of genetic algorithm on quality graded networks for intelligent routing. World Congress on Information and Communication Technologies, 11– 14 December 2011, Mumbai, India, pp. 558–563 (2011).

15. Lee, C., Wada, H., Suzuki, J.:Exploring biologically-inspired evolvable network applications with the BEYOND architecture: IEEE conference on Bio-Inspired Models of Network, Information and Computing Systems, 1-8 (2006).

16. Magnani, L.: Abduction, Reason, and Science. Processes of Discovery and Explanation. Kluwer Academic/Plenum Publishers, New York (2001).

17. Mohamed, D., T., Kotb, A., M., Ahmed, S., H.: A Medium Access Protocol for Cognitive Radio Networks Based on Packet's Collision and Channels' Usage, International Journal of Digital Information and Wireless Communications (IJDIWC) 4(3): 314-332 (2014).

18. Ong, Y., S.: Artificial intelligence technologies in complex engineering design, Ph.D. Thesis, School of Engineering Science, University of Southampton, UK (2002).

19. Choudhary, D., and Robinson, A.L.: A New Approach to Wireless Channel Modeling using Finite Mixture Models: International Journal of Digital Information and Wireless Communications (IJDIWC) 4(2): 169-183 (2014).

20. Nakano, T.: Biologically inspired network systems: A review and future prospects, IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, vol. 41, 630–643 (2011).

21. Hu, T., Lin, Z.: Properties of the composite quadratic Lyapunov functions, IEEE Transactions on Automatic Control, 49, 1162–1167 (2004).

22. Chen, X, Ong, Y., Lim, M., Tan, K.: A multi-facet survey on memetic computation. IEEE Transactions on Evolutionary Computation, 15(5), 591–607 (2011).

23. Sanusi, H., A., Zubair, A., Oladele, R., O.: Comparative assessment of genetic and memetic algorithms, Journal of Emerging Trends in Computing and Information Sciences, vol. 2, pp.498-508 (2011).

24. Krasnogor, N.: Studies on the Theory and Design Space of Memetic Algorithms, Ph.D. Thesis, Faculty of Computing, Mathematics and Engineering, University of the West of England, Bristol, UK (2002).

25. Karaboga, D., Ozturk, C.: A novel clustering approach: Artificial Bee Colony (ABC) algorithm. Applied Soft Computing vol 11, 652–657 (2011).

26. Naghshvar, M., Zhuang, H., Javidi, T.: A general class of throughput optimal routing policies in multi-hop wireless networks, IEEE Transactions on Information Theory, vol. 58, 2175–2193 (2012).

27. Marquez, H., J.: *Nonlinear Control Systems*, Wiley-Interscience, USA (2003).

28. Luenberger, D., G., Ye, Y.: *Linear and Nonlinear Programming*, 3rd edn, Springer, (2008).

29. Parks, P., C.: A. M. Lyapunov's stability theory—100 years on, IMA Journal of Mathematical Control & Information, vol. 9, 275–273 (1992).

30. Matousek, *R.,* IAENG, I.: Svarc, P., Pivoňka, P., Osmera, M., Seda.: Simple methods for stability analysis of nonlinear control systems, Proceedings of the World Congress on Engineering and Computer Science Vol II, WCECS, 20–22 October 2009, San Francisco, USA (2009).

APPENDIX

Table A1. QoS distribution

No. of nodes	Total links	QoS				Inside cone		
		B	De	RA	PL	Node	Links	Total paths
32	109	105	90	26	103	4	11	2
64	232	220	130	47	230	29	91	2
128	400	386	224	103	396	35	81	1
256	857	823	533	210	853	63	212	32
512	1648	1596	1096	425	1644	49	178	3
1024	3391	3269	2061	803	3387	364	1077	40

Table A2. Path and Grade Value (GV)

Path	GV
2-1-30	4-4
0-22-28-59-56	4-4-3-5
14-7-22-92-93-98-103-117	5-4-3-4-4-4-3
42-148-180-174-250	5-4-3-5
59-52-57-88-40-158-154-156-166-274-275-306-303-308	5-4-5-5-5-5-4-4-5-5-5-5-4
64-56-801-776-624-607-590-593-558-554-559-599-626-656-645-639-636-802-812-773	4-5-5-4-4-5-5-3-3-5-5-4-4-4-4-4-5-5-5

Table A3. Simulation set-up

Sl. No	Metric	Range	Threshold
1.	Bandwidth	2– 64 Mbps	4
2.	Delay	2–7 ms	4
3.	Packet loss	0–4	2
4.	Buffer space	15–40, each of 1 Mb length	20
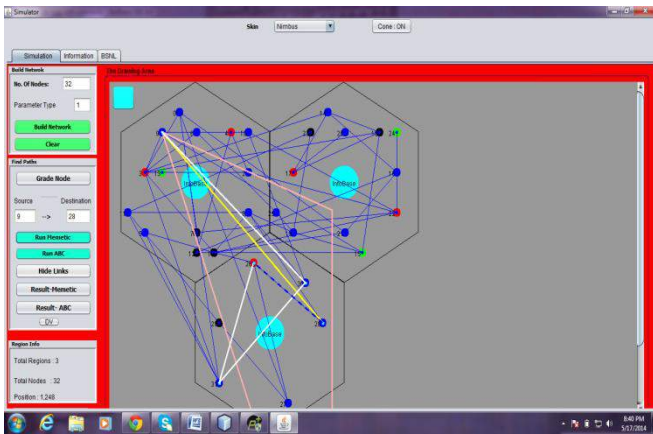
Figure A1. Run of 32 nodes

Table A4. QoS value assigned for 32 nodes

Node number	RA	Neighbour node	B	De	PL
0	21	3	57	6	0
		7	45	5	1
		21	46	5	1
		27	18	6	1
		31	6	4	1
1	19	3	45	6	1
		5	61	6	1
		8	62	6	0
		21	46	5	1
		22	12	5	1
		26	14	5	1
2	36	3	45	6	1
		4	29	2	0
		13	4	4	0
		22	10	4	0
3	34	0	57	6	0
		1	45	6	1
		2	45	6	1
		3	62	5	1
		6	33	4	1
		9	36	4	1
		14	51	6	1
		18	48	4	0
		31	53	2	0
4	33	2	29	2	0
		12	28	4	0
5	28	1	61	6	0
		31	53	2	0
6	36	3	33	4	1
		25	40	4	1
7	19	0	45	5	1
		22	10	4	0
8	38	1	62	6	0
		11	12	5	1
		19	58	2	0
		28	14	6	1
		30	2	2	1

		10	62	6	1
		15	45	4	0
		30	19	5	0
10	35	9	62	6	1
		22	43	3	1
		30	19	5	0
11	28	8	12	5	1
		23	63	6	1
		29	44	6	0
		31	36	3	0
12	18	4	28	4	0
		15	45	4	0
		16	56	5	0
13	18	2	4	4	0
14	26	3	51	6	1
		16	56	5	0
		17	46	4	1
		18	54	4	0
15	19	9	45	4	0
		12	45	4	0
		18	54	4	0
		19	49	3	1
16	20	12	56	5	0
		14	56	5	0
		21	12	5	1
		22	43	3	1
17	27	14	46	4	1
		18	58	2	0
		20	60	6	0
18	16	3	48	4	0

		14	54	4	0
		15	54	4	0
		17	58	2	0
19	26	8	58	2	0
		15	49	3	1
		23	8	2	1
		25	40	4	1
20	26	17	60	6	0
21	25	0	46	5	1
		1	46	5	1
		16	12	5	1
22	30	1	12	5	1
		2	10	4	0
		7	10	4	0
		10	43	3	1
		16	43	3	1
		24	8	2	1
23	18	11	63	6	1
		19	8	2	1
24	18	22	8	2	1
25	21	6	40	4	1
		19	40	4	1
26	26	1	14	5	1
		27	18	6	1
		28	44	6	0
		31	36	3	0
27	26	0	18	6	1
		26	18	6	1
		29	2	2	1
28	28	8	14	6	1
		26	44	6	0

29	15	11	44	6	0
		27	2	2	1
30	25	8	2	2	1
		9	19	5	0
		10	19	5	0
		31	19	5	1
31	26	0	6	4	1
		3	53	2	0
		5	53	2	0
		11	36	3	0
		26	36	3	0
		30	19	5	1

Table A5. Distribution analysis of 32 nodes

32 nodes						
B	De	RA	PL	Inside cone		
				Nodes	Links	Total path
107	58	23	111	10	26	36

Table A6. The algorithm output along with Bayesian and Lyapunov function

Output	Path no.	Path			Grade	
Memetic	35	9-30-31-26-28			4-4-5-4	
		GEV	Grade avg	Vx	Vx dot	
		0.8499	4	4.2	−8.4	

ABC	18	Path			Grade
		9-15-19-8-28			4-3-5-4
		GEV	Grade avg	Vx	Vx dot
		0.8499	4	4.2	−8.4
ABC-alternate best	7	Path			Grade
		9-10-22-2-3-0-27-26-28			4-5-4-4-4-4-4-4
		GEV	Grade avg	Vx	Vx dot
		0.825	4	7.4	−14.8
Best of Lyapunov	6	GEV	Vx		Vx dot
		0.8499	5.2		−10.4
Without cone Memetic	0	9-3-0-21-16-22-10-30-31-11-8-28			
Without cone ABC	1	9-3-0-21-16-22-10-30-31-11-23-19-8-28			
	31	9-3-14-16-22-10-30-31-11-23-19-8-28			

Determination of Performance Enhancement in IEEE 802.11 Multi-rate WLAN

Ibrahim Abubakar Alhaji

Department of Physics, Federal University, Kashere

P.M.B. 0182, Gombe, Nigeria

alhaji259@gmail.com

Bello Idrith Tijjani, Mutari Hajara Ali

Department of Physics, Bayero University, Kano

P.M.B. 3011, Kano, Nigeria

idrith@yahoo.com, mutariali@gmail.com

ABSTRACT

This work investigates performance anomaly of IEEE 802.11 multi-rate WLAN with respect to throughput and fairness. Even though in a given WLAN, hosts with different transmission rates exist, IEEE 802.11 allows them to only transmit equal-sized frames. This is unfair and reduces the overall throughput of the network. A new technique is proposed in this work to mitigate this problem by making the transmission time of the host equal instead, thereby adjusting their packet sizes in proportion to their individual transmission rates. Analytical and simulation analyses show that the proposed scheme eliminates the anomaly by improving the throughput and fairness index (FI) of a given WLAN.

KEYWORDS

Fairness Index, Multi-rate WLAN, Performance Anomaly, Throughput, Transmission Rate.

1 INTRODUCTION

The IEEE 802.11 [1] or Wi-Fi is the specification defined by Institute of Electrical and Electronics Engineers (IEEE) for wireless local area networks (WLANs). The standard defines two layers; physical layer (PHY), which specifies the modulation scheme used and the signaling characteristics for the transmission through radio frequencies. The second layer is medium access control (MAC) layer. This layer determines how the medium is used [2]. Different number of IEEE 802.11 standard exists (802.11, 802.11a, 802.11b, 802.11g, etc.).

IEEE 802.11 provides a way for allotting a part of the wireless channel bandwidth to some nodes using the Optional Point Coordination Function (PCF) or Hybrid Coordination Function (HCF) but the mandatory access protocol, Distributed Coordination Function (DCF), uses the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism to share the channel in a fair way. It has been observed, however, that in some scenarios in a wireless setting, the method (DCF) results in significant performance degradation. Some nodes, in WLANs, may be far away from their access point therefore the quality of their transmission, due to signal fading, is low. In this situation, IEEE 802.11 products degrade the bit rate from the nominal highest rate to lower or even lowest nominal rate [3]. That is, when a node detects a repeated transmission failure it reduces its bit rate. When there is at least a node with lower rate, IEEE 802.11 cell presents performance unfairness; the throughput of all

nodes transmitting at higher rate is degraded by that slow rate host [4]. The CSMA/CA is the basis for this anomaly as it guarantees that the probability of a long term channel access is equal for all nodes. Wireless devices adapt, normally, their transmission rate to the condition of the radio frequency (RF) channel. A device transmitting at a lower rate holds onto the RF channel longer than those with higher rates. This slows down the network and lowers the overall performance of the IEEE 802.11 WLANs. The paper [4] analyzed this anomaly of IEEE 802.11b and derived an important expression for the useful throughput but did not consider the fairness of the system. A time-based regulator algorithm that resides in the AP to provide time-based fairness by regulating packets is proposed in [5]. The limitation of the algorithm is that it only schedules transmission of packet from or to a device if it has not used up all of its available time. A maximum transfer unit (MTU) adjustment proportionate with the transmission rate is proposed in [6]. Their solution overcome the performance anomaly but did not address fairness problem in a multi-rate case. Authors in [7] proposed that the contention window (CW) be proportionate with the transmission rate on the expectation that the CW$_{min}$ will be a configurable parameter in IEEE 802.11e extension. The useful throughput derived by [4] is utilized by [8], where a transmission rate-based packet size adjustment (TRPSA) scheme is proposed. Their scheme improves the performance anomaly in IEEE 802.11a/b/g by increasing the throughput and the fairness of the system. However, the scheme is not tested against IEEE 802.11n which is the extension of the previous standards. The aims of this work are to derive the throughput equation of both transmission time-based scheme (TTBS) and IEEE 802.11 WLANs, use it to calculate the theoretical maximum value of their fairness indices and validate same through simulation using Optimized Network Engineering Tool (OPNET) modeler. The objective is to compare the performance metrics of TTBS and IEEE 802.11 WLANs in a multi-rate situation.

2 PERFOMANCE ANALYSIS

In this section we derived the equation of one of the performance metrics of IEEE 802.11 WLANs, i.e., the throughput. We then used the equation to calculate the theoretical maximum value of Jain's fairness index (Jain et al.1984; 1999); the other performance metric of IEEE 802.11 WLANs. Thereafter, we introduce our proposal called transmission time-based scheme (TTBS) to enhance and address the fairness issue in IEEE 802.11 WLANs.

According to [4] the time T_i for a single host or device to transmit a frame if propagation delay is neglected is given by:

$$T_i = t_i + t_{ov} \tag{1}$$

Where,

t_{ov}, t_i are the overhead time and the frame transmission time respectively given by:

$$t_{ov} = \text{DIFS} + t_{pr} + \text{SIFS} + t_{pr} + t_{ack} \quad \text{and}$$

$$t_i = \frac{S_i}{r_i} \tag{2}$$

S_i is the frame size while r_i is the transmission rate,

$DIFS$ and $SIFS$ are the distributed and short inter-frame space time given by:

$$DIFS = SIFS + 2 \times aSlotTime \tag{3}$$

The value of SIFS and aSlotTime is shown in table 1 for different PHYs and t$_{pr}$ is the physical layer convergence protocol (PLCP) preamble and header transmission time. S$_i$ is the frame size and r$_i$ the transmission rate.

When there is more than one host attempting to use or capture the wireless medium, the hosts have to contend for the medium. Therefore the overall transmission time has to account for the time spent in contention t_{cont} by the hosts. This implies that equation (1) has to be modified. The equation can now be written as:

$$T_i = t_i + t_{ov} + t_{cont} \tag{4}$$

Assuming that N hosts alternate transmission with all possible collisions and assuming further that multiple collisions are negligible [4], then, the total time during which all hosts transmit once is given by:

$$T_{tot} = \sum_{j=1}^{N} T_{r_j} + P_c(N) \times t_{jam} \times N \quad (5)$$

Where,

$\sum_{j=1}^{N} T_{r_j}$ = channel occupation time for all host with transmission rate r_j,

$P_c(N)$ = Probability of collision,

t_{jam} = average delay when packets form two node collision.

Therefore the channel utilization U_i by any node is given by:

$$U_i = \frac{T_i}{T_{tot}} \quad (6)$$

According to [4] the useful throughput P_i achievable by any host depends on the number of all hosts in the network and is given by:

$$P_i = \frac{t_i}{T_i} \quad (7)$$

Therefore, the throughput X_i at MAC layer for any host with transmission rate r_i is given by:

$$X_i = U_i \times P_i \times r_i \quad (8)$$

Putting equations (6) and (7) in (8) yields:

$$X_i = \frac{T_i}{T_{tot}} \times \frac{t_i}{T_i} \times r_i \quad (9)$$

but $t_i = \frac{S_i}{r_i}$ therefore equation (9) becomes:

$$X_i = \frac{S_i}{T_{tot}} \quad (10)$$

In IEEE 802.11 all hosts transmit equal-sized frames, therefore S_i is the same for all hosts. This causes the performance anomaly in a multi-rate multi-node situation as all nodes get the same throughput regardless of their transmission rates.

We now introduce the Jain's fairness index [9-10] which measures fair allocation of resources (in our case the wireless medium) in a distributed system. The fairness index (FI) is bounded between 0 and 1. If the index approaches 1 the system is very fair otherwise it is not. The FI in terms of transmission rate r_i and throughput X_i is given by:

$$FI = \frac{\left(\sum_{i=1}^{N} r_i/X_i\right)^2}{N \sum_{i=1}^{N} \left(r_i/X_i\right)^2} \quad (11)$$

We have shown in (10) that X_i is the same for all hosts in IEEE 802.11 WLANs. Therefore equation (11) becomes;

$$FI = \frac{\left(\sum_{i=1}^{N} r_i\right)^2}{N \sum_{i=1}^{N} (r_i)^2} \quad (12)$$

Let us take the example of IEEE 802.11b which specifies four different data rates of 1, 2, 5.5 and 11Mbps to calculate the value of the fairness index. The number of rates is four therefore N = 4. This gives us;

$$FI = 0.6084$$

This implies that the maximum theoretical value of FI achievable by IEEE 802.11b wireless system is 0.6084. That is the system is around 60% fair.

2.1 Transmission Time-Based Scheme (TTBS)

We now propose that instead of all hosts to have equal-sized frames, let the hosts have equal transmission time. This, we believe, will improve the network performance from time-based fairness point of view. That is, let t_i be the same for all host. But t_i is given by:

$$t_i = \frac{S_i}{r_i} \quad (13)$$

Let the frame size and transmission rate of the fastest node be s_{max} and r_{max} respectively. Further, let s_i and r_i be the frame size and

transmission rate of any other node in the network. Since transmission time is the same for all hosts in our proposal, therefore equation (13) can now be written as:

$$t_i = \frac{s_i}{r_i} = \frac{s_{max}}{r_{max}}$$

This implies that;

$$s_i = \frac{s_{max}}{r_{max}} \times r_i \qquad (14)$$

Equation (14) shows that if transmission time is equal then the frame size has to change. In other words for any given host its frame size is proportional to its transmission rate. Therefore according to equation (14), equation (10) is no longer constant. i.e., throughput is not the same for all hosts in our scheme and given s_{max}, r_{max} and r_i the frame size of any other node s_i can be calculated.

Now invoking the fairness index by substituting equation (10) in (11) yields:

$$FI = \frac{\left(\sum_{i=1}^{N} r_i/s_i\right)^2}{N \sum_{i=1}^{N} \left(r_i/s_i\right)^2} \qquad (15)$$

Now putting equation (13) in (15) gives:

$$FI = \frac{\left(\sum_{i=1}^{N} 1/t_i\right)^2}{N \sum_{i=1}^{N} \left(1/t_i\right)^2} \qquad (16)$$

In our scheme t_i is the same for all host, therefore equation (16) yields:

$$FI = \frac{(4/t)^2}{4 \times 4 \, (1/t)^2} = 1$$

$$\therefore \quad FI = 1 \qquad (17)$$

Equation (17) shows that the IEEE 802.11b wireless system can be absolutely fair using our scheme.

3 SIMULATIONS

We have designed and simulated twenty two (22) set of scenarios with simulation run-time of hundred seconds (100s) for each. The nodes communicate with an access point (AP) and transmit a constant bit rate (CBR) traffic to the AP. IEEE 802.11 standard specifies that all node transmit an equal-sized frame so we adopt a frame size of 1000byte. We further adopt IEEE 802.11b version of the standard for the scenarios as it specifies only four (4) different transmission rates of 1, 2, 5.5 and 11Mbps unlike IEEE 802.11a/g which specify up to eight (8) different transmission rates and only basic access method was used for the simulation. Optimized network engineering tool (OPNET) modeler academic edition version 17.5 was used for the simulations. In designing the scenarios in the modeler, packet inter-arrival time was set according to the transmission time of each node. Packets on-off generation time were set to 1 and 0 respectively while start time was set to 0.1. All other values used were the modeler's default unless stated otherwise.

3.1 Scenarios

In the first four scenarios nodes A, B, C and D, as shown in figure 1, transmit with the same transmission rate of 1, 2, 5.5 and 11Mbps. In the fifth scenario each node transmits with a different transmission rate. That is 1, 2, 5.5 and 11Mbps respectively. Designing and simulating our proposal, TTBS, requires that each node has its frame size proportional to its transmission rate i.e., 91, 182, 500 and 1000byte for 1, 2, 5.5 and 11Mbps in that order. Equation (14) was used to arrive at the given frame size values.

Now using the frame size values above we simulated another set of four scenarios where we varied the nodes number according to the transmission rates. That is in the first scenario we varied the number of 1Mbps node from 2-14 and kept the number of 2, 5.5 and 11Mbps nodes constant. We repeated this with 2, 5.5 and 11Mbps nodes. These set of scenarios were repeated with IEEE 802.11b for comparison.

Next we checked the influence of packet size variation by fixing the packet sizes of three nodes and allowing the other node to take values form 100-1500byte. We now compared TTBS against IEEE 802.11a, g and n in a multi-rate condition. The respective packet sizes for 802.11a/g/n are given in table 2. Lastly we compared TTBS and 802.11 in a dynamic multi-rate situation, where the four nodes (A, B, C and D) move towards the access point with a speed of 1, 1, 2 and 2m/s respectively.
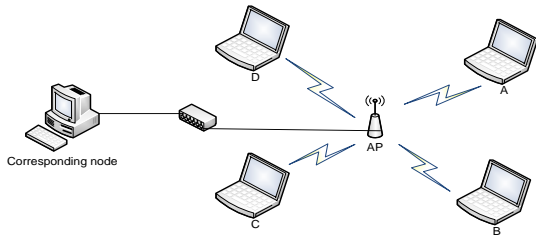
Table 1: Slot Time and SIFS Values for Each PHY

PHY	aSlotTime (s)	SIFS (s)
FHSS	50	28
DSSS	20	10
OFDM	9	16

Figure 1: Location of the Nodes

Table 2: 802.11a/g/n Rates with Their Proportional Packet

S/N	802.11a/g rate (Mbps)	802.11a/g packet size proportional to rate (byte)	802.11n rate (Mbps)	802.11n packet size proportional to rate (byte)
1	6	111	6.5(base)~60(max)	100
2	9	167	13(base)~120(max)	200
3	12	222	19.5(base)~180(max)	300
4	18	333	26(base)~240(max)	400
5	24	444	39(base)~360(max)	600
6	36	667	52(base)~480(max)	800
7	48	889	58.5(base)~540(max)	900
8	54	1000	65(base)~600(max)	1000

Sizes

4 RESULTS AND DISCUSSIONS

In this section the results of the simulation scenarios described in section 3.1 are presented and discussed.

### 4.1 Performance Anomaly when Multi-rate Nodes Coexist

Figure 2 shows that in a single-rate situation the wireless system is very fair and the throughput of the individual node is almost equal. We adopt

11Mbps single-rate, as figure 2 (a) reveals that the highest throughput is attained in 11Mbps single-rate scenario, for comparison with multi-rate where nodes A, B, C and D transmit with the rate of 1, 2, 5.5 and 11Mbps respectively. The average throughputs of A, B, C and D were found to be 1.95, 2.00, 1.97 and 1.97Mbps in that order and the system throughput was 7.9Mbps in figure 2. Figure 3 (a) shows the performance anomaly clearly as the throughputs of nodes A, B, C and D drop to 0.54, 0.47, 0.51 and 0.57Mbps and system throughput downs to just 2.10Mbps. Although the transmission rate of node D does not change, its throughput drops to 0.57Mbps about 71% loss and system throughput drops by 73%. The fairness index is around 1, most fair, when only single-rate multi-node is present, immediately multiple-rate coexist the index drops sharply to about 0.60 as predicted by equation (12). This clearly revealed the nature of the fairness issue in IEEE 802.11 multi-rate WLANs.

Comparing TTBS with 802.11 under multi-rate condition shows that the system throughput raises from 2.10Mbps in 802.11 to 3.26Mbps in TTBS about 36% appreciation (figure 4a). The system fairness rises to 1 unlike in 802.11 where it hovers around 0.6. This result indicates that the system performance under TTBS improves.
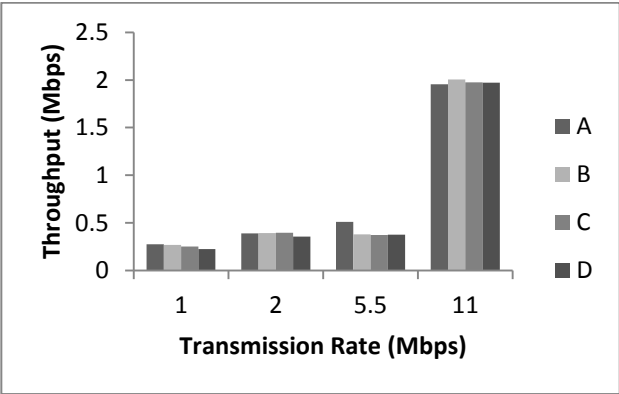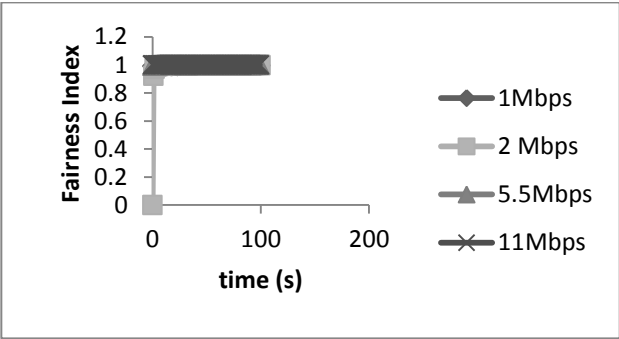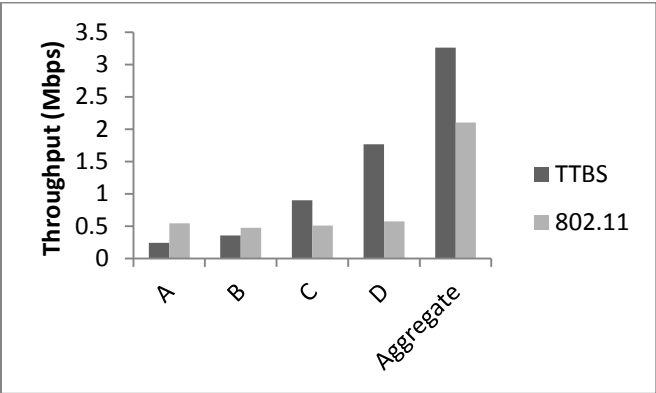
(a) Throughput
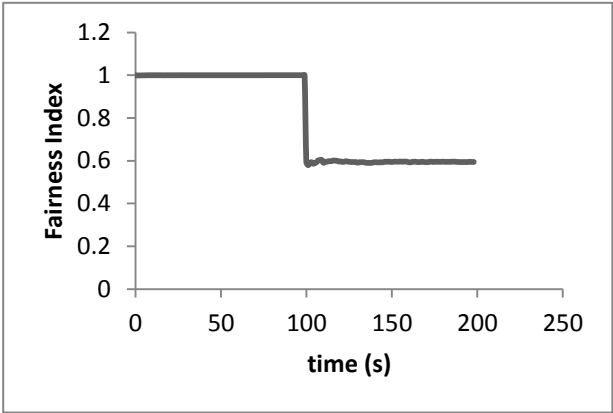
(c) Fairness Index

Figure 3: Comparison of Single and Multi-Rate Scenarios (a) Throughput (b) Fairness Index
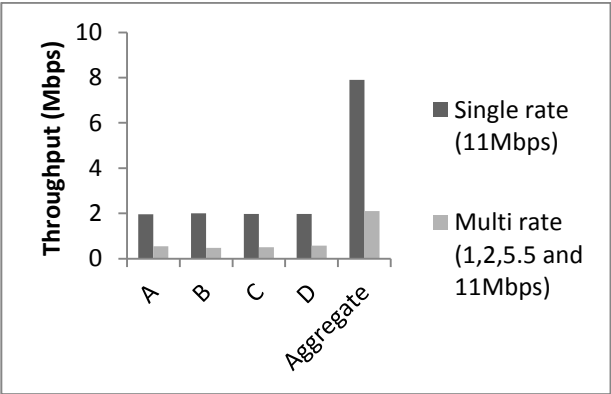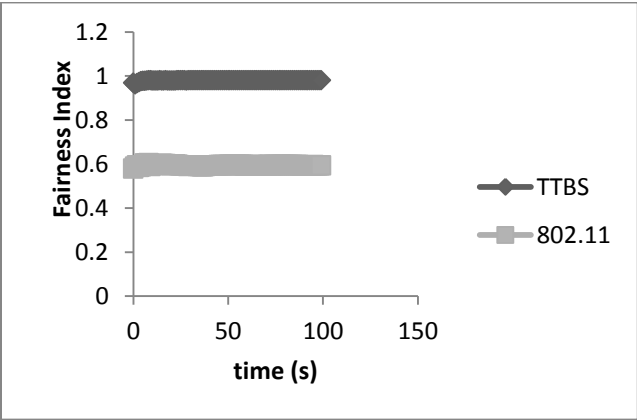

(b) Fairness Index

Figure 2: Single Rate Scenarios (a) Throughput (b) Fairness Index
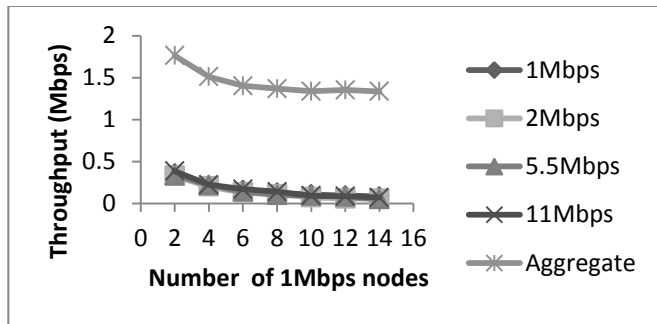

(a) Throughput

(a) Throughput

(b) Fairness Index

Figure 4: Comparison between Multi-rate TTBS and 802.11 (a) Throughput (b) Fairness Index

4.2 Influence of Node Number Variation on the System Performance

In figure 5 (a) and (b) we notice the influence of node number variation with respect to 1Mbps node (the node with the least transmission rate) on the system. It is as expected; the nodes in TTBS scenario have different throughput while maintaining almost equal throughput in 802.11. Even though the fairness index of TTBS decreases a bit, the system is still very fair unlike in 802.11. The system throughput decreases in both cases as only the number of nodes with the least transmission rate increases. When considering node number variation, the throughput and fairness index of the node under investigation are average values for any given number of the said node.

Figure 6(a) and (b) show that the system aggregate for 802.11 increases while that of TTBS decreases because all the nodes have nearly equal throughput in 802.11, so if the number of any node other than the node with the least transmission rate increases the system throughput appreciates unlike in TTBS where nodes have different throughput therefore the system throughput only appreciates when the higher nodes numbers increase. Figures 6, 7 and 8 support the argument we just stated. They also highlight the fact that the wireless system is most fair when the number of the node under variation is just two (2) beyond that the fairness index begins to fluctuate downward a bit as the wireless system allocates more resources to nodes with same transmission rate.
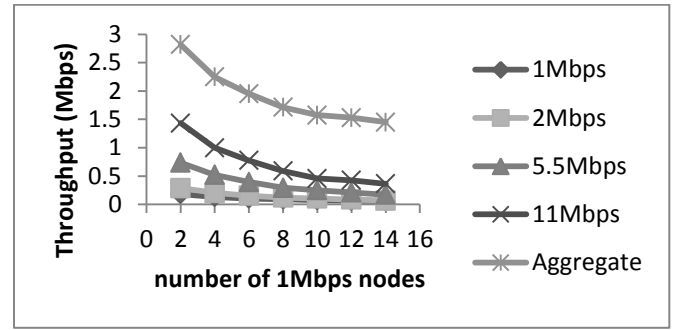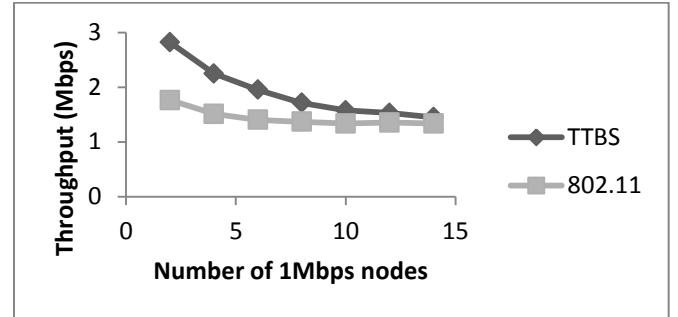
(b) TTBS

(c) Throughput

(d) Fairness Index

Figure 5: Comparison of the Performance of 802.11 and TTBS with Changing Number of **1Mbps** Node. (a) 802.11: Throughputs for Different Rates. (b) TTBS: Throughputs for Different Rates. (c) Total throughput. (d) Fairness Index

(a) 802.11

(a) 802.11

(b) TTBS

(a) TTBS
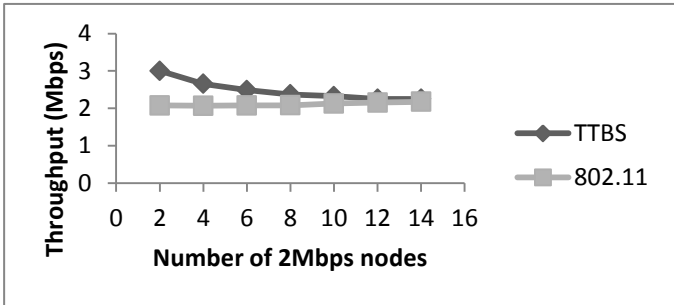
(c) Throughput
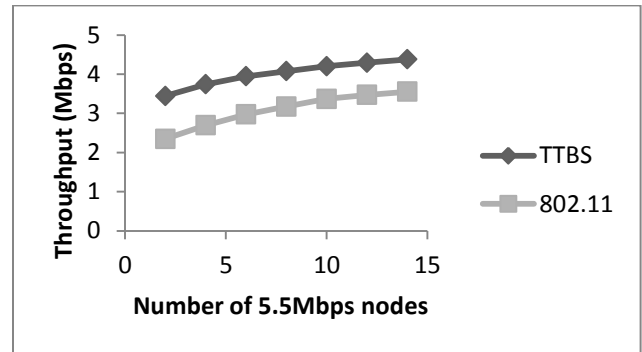
(b) Throughput
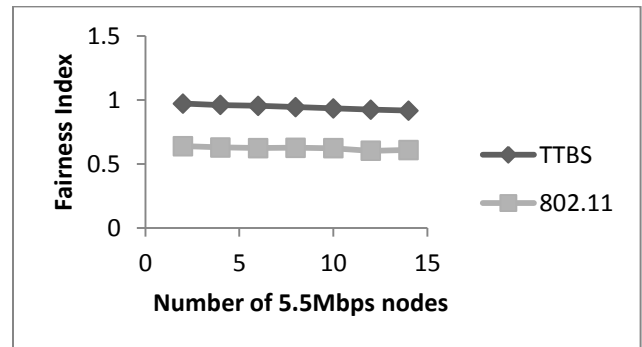
(d) Fairness Index

Figure 6: Comparison of the Performance of 802.11 and TTBS with Changing Number of **2Mbps** Node. (a) 802.11: Throughputs for Different Rates. (b) TTBS: Throughputs for Different Rates. (c) Total Throughput. (d) Fairness Index

(c) Fairness Index

Figure 7: Comparison of the Performance of 802.11 and TTBS with Changing Number of **5.5Mbps** Node. (a) 802.11: Throughputs for Different Rates. (b) TTBS: Throughputs for Different Rates. (c) Total Throughput. (d) Fairness Index
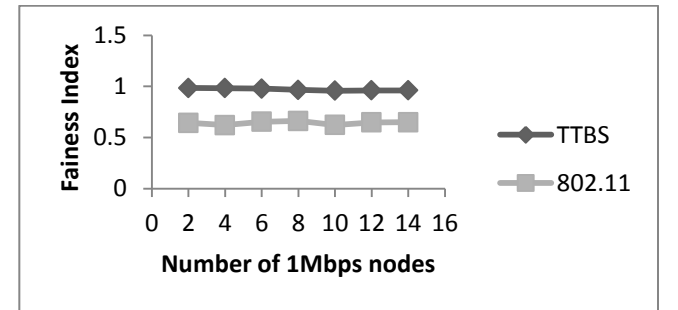

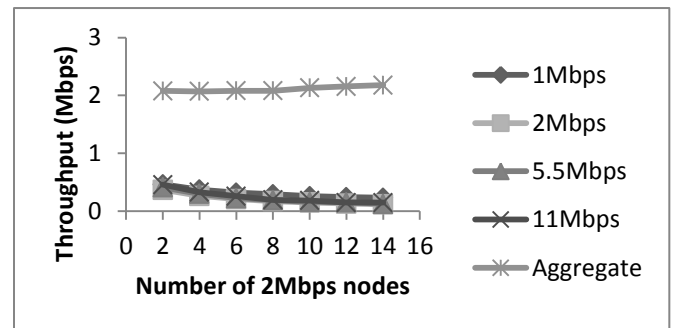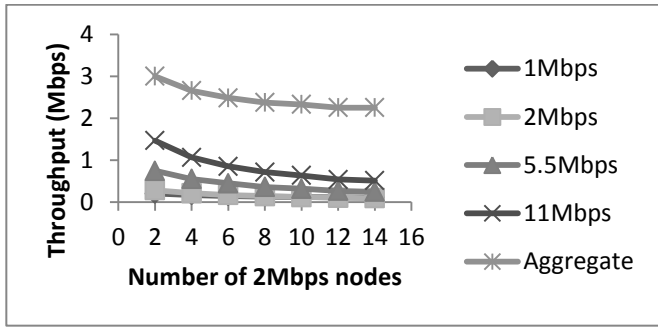(b) 802.11

(a) 802.11
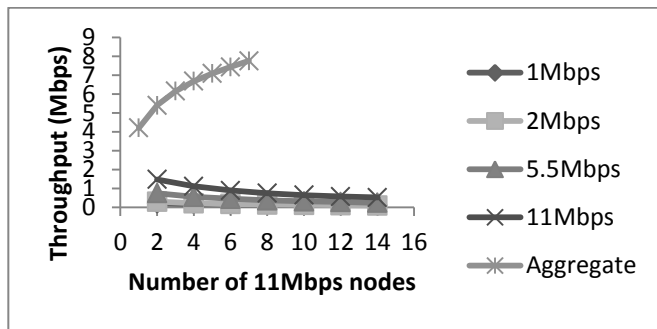
(b) TTBS
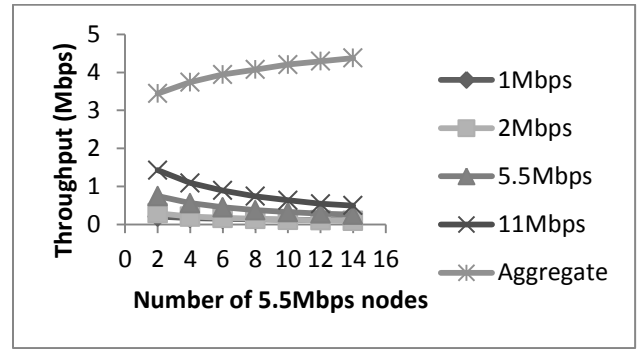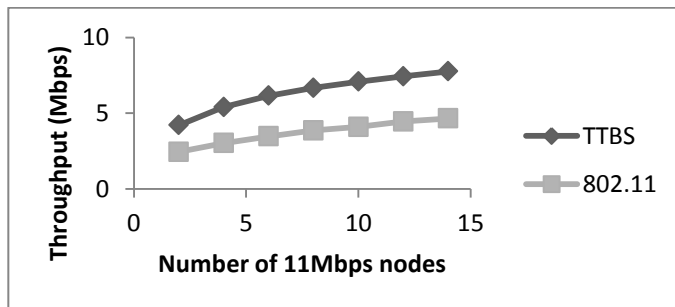

(c) Throughput
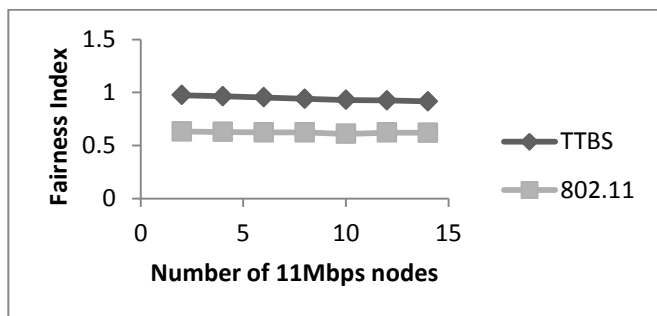
(d) Fairness Index

Figure 8: Comparison of the Performance of 802.11 and TTBS with Changing Number of **11Mbps** Node. (a) 802.11: Throughputs for Different Rates. (b) TTBS: Throughputs for Different Rates. (c) Total Throughput. (d) Fairness Index
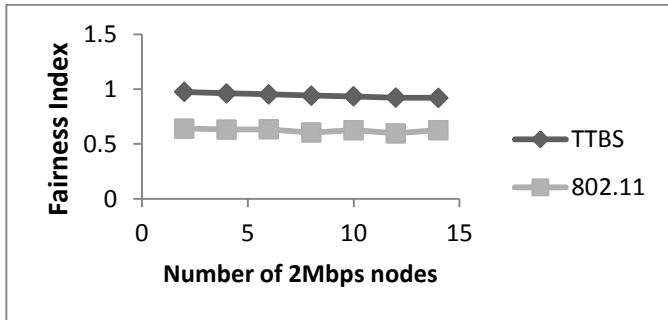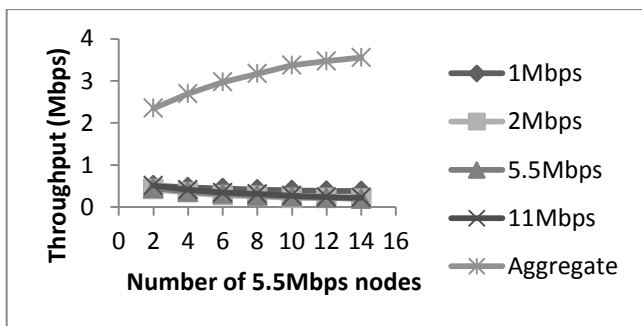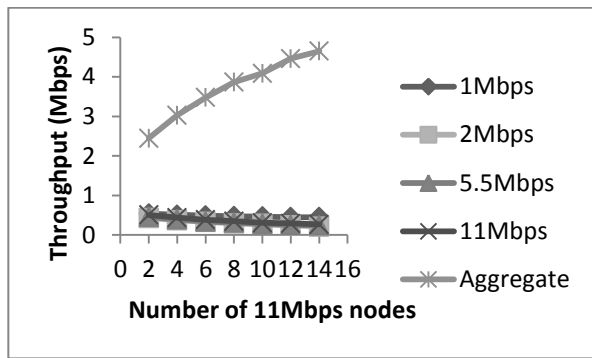
4.3 Influence of Packet Size Variation on System Performance

Now we study the influence of packet size variation on the network. Figure 9a reveals that whenever 1Mbps node has equal-sized packet as any other node it achieves the same throughput as that particular node. We further see that the throughput of the 1Mbps node appreciates as its packet size increases while the throughput of other nodes decreases. This is due to the fact that 1Mbps being the slowest of all captures the channel longer than the other nodes. The fairness index (b part of same figure) and system throughput are highest, 0.97 and 3.25Mbps, when the nodes packet sizes-transmission rates ratio is 91: 182: 500: 1000 = 1: 2: 5.5: 11. This confirms that TTBS can effectively improve the network performance.

We expect, in figure 10a, the 2Mbps to have same throughput as 1, 5.5 and 11Mbps nodes when its packet sizes are 91, 500 and 1000byte, our expectation was only met at 91byte. Still in the same figure, 2Mbps node gets same throughput with 5.5 and 11Mbps nodes at 600 and 1200bytes because 2Mbps node is twice faster than 1Mbps node therefore it cannot achieve same throughput with 5.5 and 11Mbps nodes at the same point the 1Mbps node did, it has to slow down to the speed of 1Mbps node by adjusting its packet size. The 2Mbps node coincides with 5.5Mbps node at 600byte but 11Mbps is twice faster than 5.5Mbps node, hence the coincidence of 2 and 11Mbps nodes at 1200byte.

Even though, figure 11a, the throughput of 5.5Mbps node drops when it assumes the packet sizes of lower rate nodes its throughput is still above theirs; it is 5.5 and 2.75 faster than 1 and 2Mbps nodes respectively. When the 5.5Mbps node assumes the packet size of 11Mbps node (1000byte) its throughput is less than that of 11Mbps node; it's twice slower than 11Mbps node. Since 11Mbps node is faster than 1, 2 and 5.5Mbps nodes, its throughput should be higher when it assumes their packet sizes. This is illustrated in figure 12a.

(a) Throughput

(b) Fairness Index

Figure 10: Packet size variation for **2Mbps** node (a) Throughput, (b) Fairness index
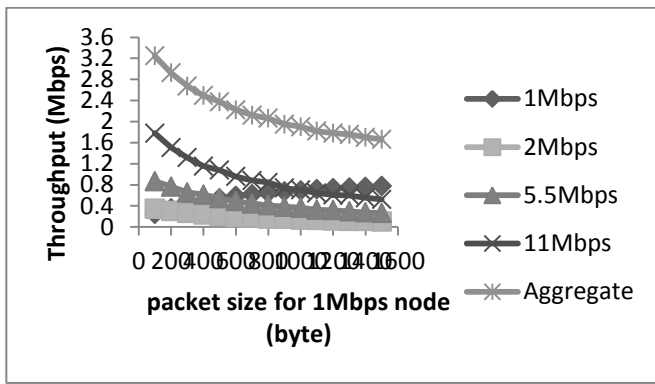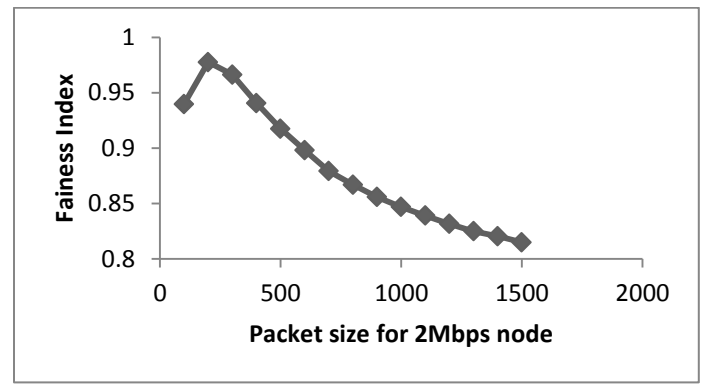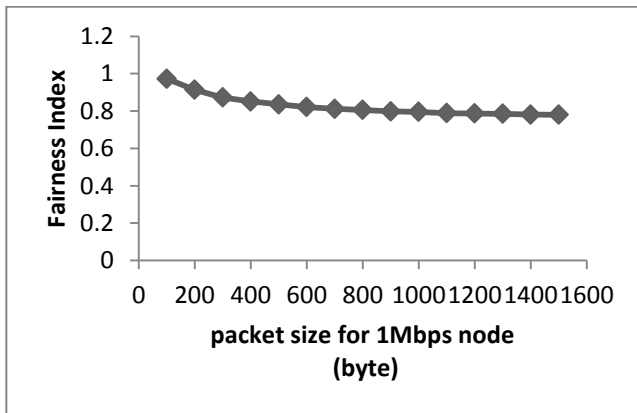
(b) Fairness Index

Figure 9: Packet Size Variation for **1Mbps** Node (a) Throughput, (b) Fairness Index

(a) Throughput
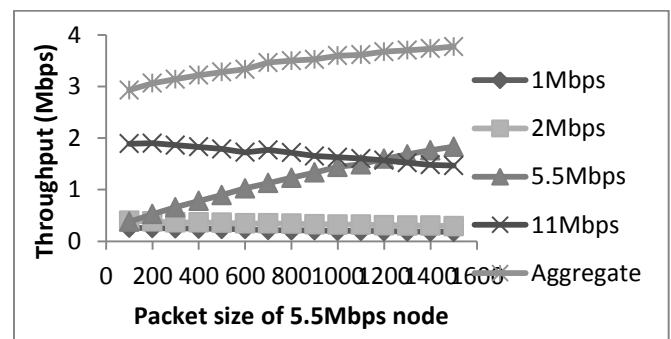
(a) Throughput

(b) Fairness Index

Figure 11: Packet Size Variation for **5.5Mbps** Node (a) Throughput, (b) Fairness Index
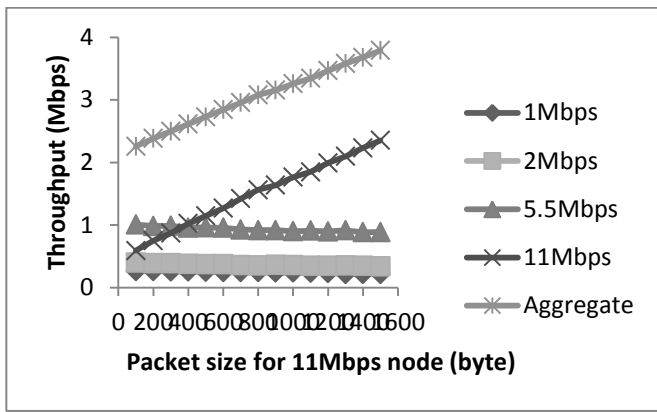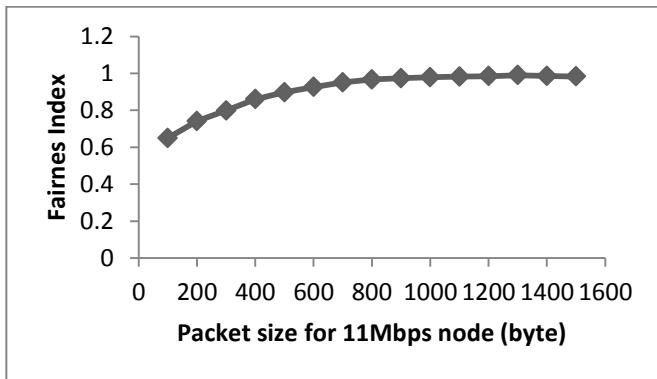
(a) Throughput

(b) Fairness Index

Figure 12: Packet Size Variation for **11Mbps** Node (a) Throughput, (b) Fairness Index

4.4 Performance Comparison of TTBS with IEEE 802.11a/g/n

When the throughputs of IEEE 802.11a and TTBS were compared in figure 13a their average values were found to be 14.25 and 18.08Mbps respectively. The TTBS outperforms 802.11a by about 21%. Their fairness indices stand at 0.96 and 0.68, a difference of 29% in favor of TTBS.

TTBS still improves the network performance when compared with 802.11g (figure 14). The average values of their throughputs are 15.60 and 13.19Mbps while their fairness indices stand at 0.94 and 0.70 respectively. TTBS increases the system throughput and fairness by 15 and 26% in that order thereby improving the network performance considerably.

Figure 15 reveals somehow a slightly different picture as both TTBS and 802.11n achieve similar

throughput of 16.59 and 16.63Mbps respectively. Actually the throughput of 802.11n is higher by just 0.003%. This is expected as 802.11n is the extension of 802.11a/b/g; there has been a lot of improvements in 802.11n as it utilizes multiple-inputs multiple-outputs (MIMO) antennas where only the signal with maximum strength is selected by the receiver. As expected TTBS achieves 0.64 fairness index increasing the system fairness by 63%.
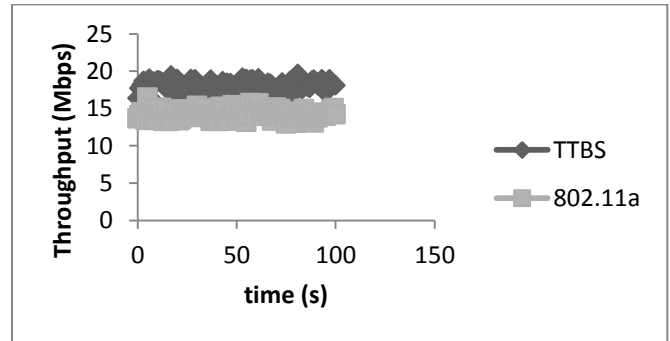
(a) Throughput
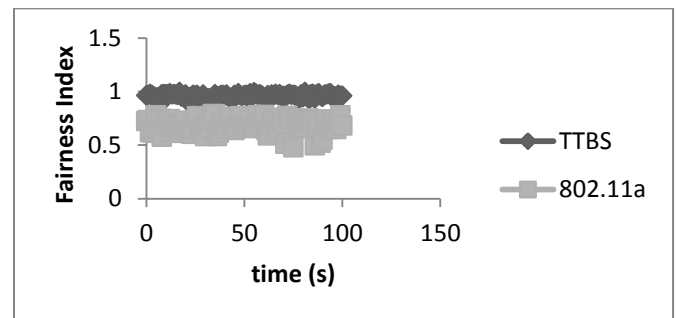
(b) Fairness Index

Figure 13: Comparison between TTBS and 802.11a. (a) Throughput, (b) Fairness Index
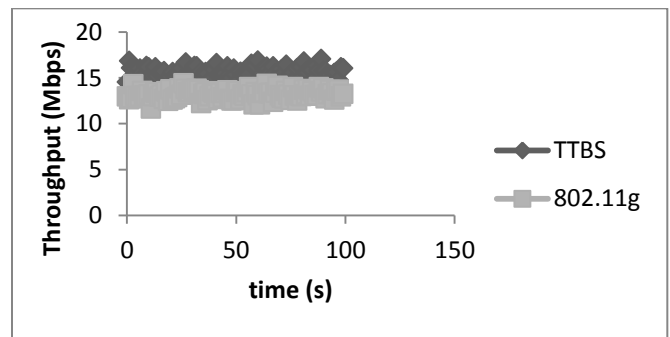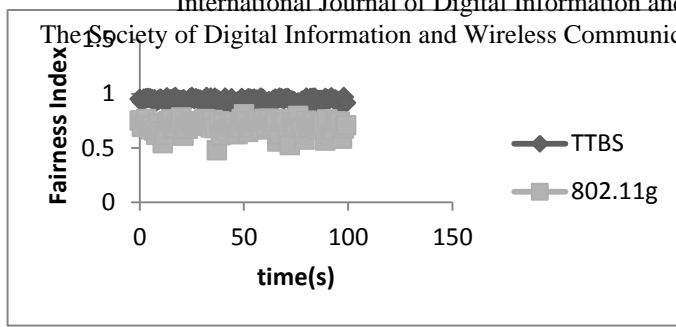
(a) Throughput

(b) Fairness Index

Figure 14: Comparison between TTBS and 802.11g. (a) Throughput (b) Fairness index
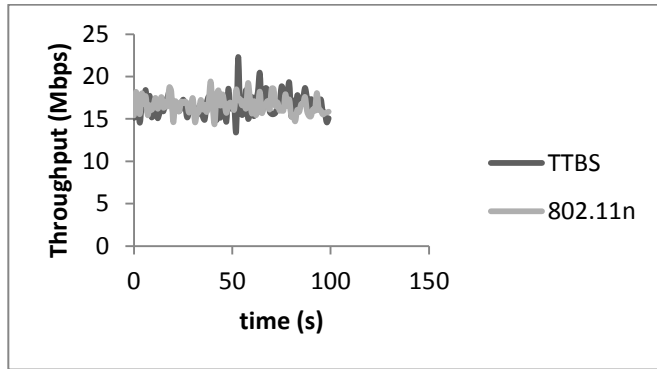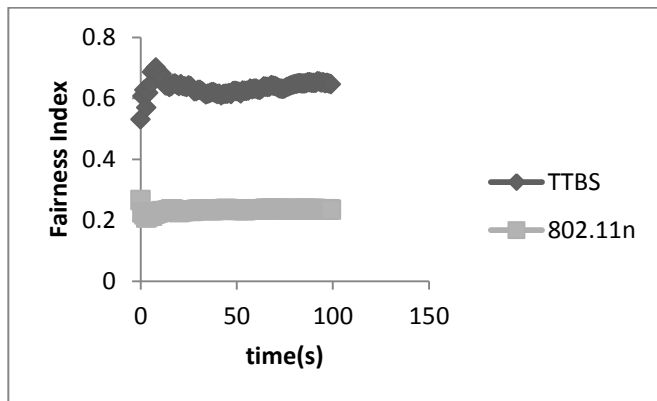
(a) Throughput

(b) Throughput

Figure 15: Comparison between TTBS and 802.11n. (a) Throughput, (b) Fairness Index

4.5 Performance Comparison between TTBS and 802.11b under Multi-rate Dynamic Condition

Figure 16 compared TTBS and 802.11b in a dynamic multi-rate condition. TTBS records a throughput of 2.62Mbps while 802.11b achieves just 2.09Mbps. The TTBS method improves the system throughput by 20%. In terms of system

fairness TTBS scores a fairness index of about 0.97 while 802.11b records just 0.62, 36% less than TTBS.
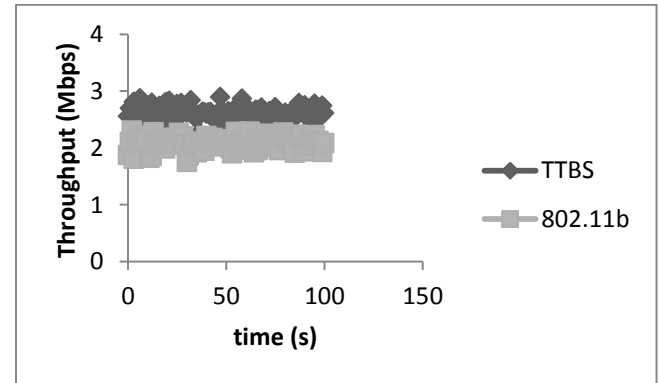
(a) Throughput
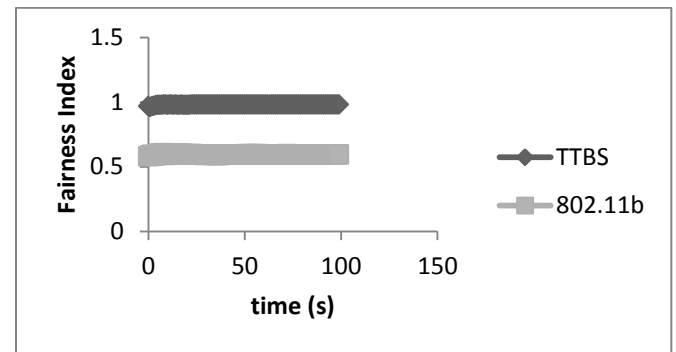
(b) Fairness Index

Figure 16: Comparison between TTBS and 802.11b in a Multi-rate Dynamic Situation. (a) Throughput, (b) Fairness Index

5 CONCLUSION

We have calculated the theoretical maximum values of IEEE 802.11b and TTBS fairness indices. The values of the indices were found to be 0.61 and 1 respectively. This result shows that IEEE 802.11b wireless system is just about sixty one percent (0.61%) fair in sharing the wireless medium among the participating nodes whereas TTBS, if adopted, makes wireless system hundred percent (100%) fair.

We then compared the performance metrics (throughput and fairness index) of transmission time-based scheme (TTBS) and IEEE 802.11 WLANs. We observed that, the percentage increase in throughput and fairness index of TTBS were found to be 21% & 29% when compared

268

with IEEE 802.11a, 35% & 39% in comparison with IEEE 802.11b and 15% & 26% when compared with IEEE 802.11g respectively. TTBS throughput percentage was lower by 0.003% in comparison with 802.11n but its fairness index is higher by 63%. In a dynamic multi-rate scenario TTBS throughput and fairness index percentages were also found to be higher by 20% and 36% in that order when compared with IEEE 802.11b.

The results obtained proved that the new scheme (TTBS) can actually enhance the network performance effectively.

6 ACKNOWLEDGEMENTS

REFERENCES

1. IEEE 802.11 Standard: Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. ANSI/IEEE 802.11 (1999).

2. Gupta, N., Kumar, P.R.: A Performance Analysis of the 802.11 Wireless LAN Medium Access Control. Communications in Information and systems, vol. 3, No. 4 (2004), pp. 279-304.

3. Holt, A., Huang, C.: 802.11 Wireless Networks: Security and Analysis. In A.J. Samees (ed.), Computer Communications and Networks. Cranfield University, Swindon, UK: Springer (2010).

4. Heusse, M., Rousseau, F., Berger-Sabbatel, G., Duda, A.: Performance Anomaly of 802.11b. Proceedings of IEEE INFOCOM, (2003), pp 836-843.

5. Tan, G., Guttag, J.: Time-based Fairness Improves Performance in Multi-rate WLANs. USENIX Annual Technical Conference, General Track (2004), pp. 269-282.

6. Yoo, S., Choi, J., Hwang, J., Chuck, Y.: Eliminating the Performance Anomaly of 802.11b. In P. Lorenz and P. Dini (Eds.): ICN 2005, LNCS 3421. Springer-Verlag Berlin Heidelberg, pp. 1055-1062.

7. Kim, H., Yun, S., Kang, I., Bahk, S.: Resolving 802.11 Performance Anomalies through QoS Differentiation. IEEE Communications Letters, vol. 9 (2005), No. 6, pp. 655-657.

8. Hu, S., Li, J.: Pan, G. (2014). Performance and Fairness Enhancement in IEEE 802.11 WLAN. International Journal of Electronics and Communications, vol. 68 (2014), pp. 667-675.

9. Jain, R. K, Chiu, D, W., Hawe, W. R.: A Quantitative Measure of Fairness and Discrimination for Resources Allocation System. Submitted for Publication to ACM Transaction on Computer System, (1984) Dec-TR-301.

10. Jain, R., Durresi, A., Babic, G.: Throughput fairness index: An Explanation. ATM Forum Contribution, No. 99-0045 (1999), pp. 1–13.

The Stakeholders of a User-Centred Design Process in Mobile Service Development

Ari Alamäki[1] and Amir Dirin[2]
HAAGA-HELIA University of Applied Sciences
Business Information Technology
Ratapihantie 13, 00520 Helsinki, Finland
{ari.alamaki, amir.dirin}@haaga-helia.fi

ABSTRACT

The use of agile methods in mobile service development has gained much attention in software design research. In addition, involving the potential stakeholders of the mobile service in the design and development process has become vital in achieving the best service experience. Many application development methodologies, such as user-centred design (UCD), ensure that stakeholders have direct involvement in the design and development. This paper describes experiences of designing a mobile concept where various stakeholders, such as tourism companies, target application users, business experts and the application developers, were involved in the design and development process. The main focus of this paper is the roles and contributions of stakeholders in mobile services for outdoor activities such as kayaking, hiking and biking. These outdoor activities are associated with different functional and non-functional requirements that are essential considerations during the design and development process. Therefore, we utilized the UCD principle as an appropriate method of involving all stakeholders in the design process, and we show that hearing the stakeholders' voice is vital in the design of outdoor-based mobile service development.

KEYWORDS

Mobile Service, User-centred Design, Software Development, Mobile Tourism, Project Management

1 INTRODUCTION

This paper describes the design process of a mobile web service for outdoor tourism activities where users regularly need navigation, guidance and nature- and route-specific information. The outdoor tourists, such as hikers, bikers or kayakers, are often operating in an unfamiliar environment. Therefore, geo-location is the important feature in the development of useful mobile services.

The ultimate goal of the research project was to design a geo-location service to support and improve customer satisfaction with small tourism companies by offering new types of digital guidance and navigation services. In developing this type of application, where different stakeholders are involved, it is essential to grasp the stakeholders' needs and requirements. In addition, the user's direct involvement in the application design and development can help in anticipating possible errors and failures. Therefore, we selected the user-centred design (UCD) [1] as the development approach. The mobile service design and development was accomplished as university research and a development project where partner companies delivered their insights through a steering group and acted as pilot companies to test the application.

The advancement in web technologies, such as cloud computing [2-6], Software as a Service (SaaS) [7-8], Service-oriented Architecture (SOA) and Web Services [9-12], as well as software development frameworks such as Vaadin and PhoneGap [13-15], have provided a unique opportunity to develop faster prototypes and robust mobile applications for various business segments. Therefore, more research is required in how to manage design and development projects in which various user groups and design teams could co-create new mobile service concepts. The methods of value co-creation are emphasized nowadays in marketing, sales and design literature, as the end users and customers are taking a more active role in service design projects [16-17]. Hence, they are not only consumers of products but also active partners in designing new services.

This research project applies mobile web technologies to design and develop a customized mobile guide service for several small and medium (SME) tourism companies to use simultaneously. The product designed in this study is based on one mobile web service where the target service is deployed to the mobile device's browser. Therefore, the designed mobile guide is not a native mobile application but is delivered as a SaaS model to the tourism companies who then deliver the URL link to their own customers, i.e. the end users. This study applied the Vaadin 6.0 Java framework [13-14], Apache TomCat and LAMP stack (Linux, Apache, MySQL and PHP) on the virtual server. The research and development goal was to design iteratively with all stakeholders' mobile web services to meet the needs of outdoors tourism business services. To achieve the research goals in the UCD process we defined the following sub-tasks.

1) Elicitation phase: Conducting user studies and gathering requirements in order to design the potential mobile web service concept. This phase is realized by collecting requirements through iterative concept and prototype designing, and gathering stakeholders' feedback. The main focus of the concept design is the end users' experiences of the provided services.

2) Evaluation and assessment phase: Piloting and testing the mobile guide prototypes with several stakeholders in the real usage environment to gain instant feedback on the concept and the proposed services.

2 MOBILE SERVICES and TOURISM

The tourism business was among the first to build electronic commerce and internet-based shopping solutions. According to a TripAdvisor survey [18], 38% of travellers have used their mobile devices to plan a trip, and 60% of respondents indicated that they have downloaded travel apps on their mobile devices in advance. The survey shows that the use of tourism-based mobile applications focuses particularly on researching restaurants (62%), checking flight statuses (51%) and researching attractions (46%). Google's study [19] provides

similar findings and emphasizes the importance of mobile services in the tourism business.

In outdoor activities, users often require navigation, guidance and other nature- and route-specific information. Tourists who are walking, hiking, biking or kayaking are frequently operating in an unfamiliar environment. Therefore, the aim is to design a tourism-based mobile system that supports most outdoor activities and is available from tourism companies [20, 21]. In addition to supporting the outdoor pursuits, the tourism companies could increase customer satisfaction by offering new types of digital guides and navigation aids as added-value services for their customers.

Tourism companies nowadays offer their customers hard-copy maps that contain the route details. However, it is easy to transfer the existing route and guide information, as highlighted on the printed map, to mobile devices. In addition to the static paper-based information, mobile services offer dynamic information for end users, such as speed, location, distances and social information related to specific locations. The existing tourism applications mainly provide information and support for city or urban navigation services [22, 23], in which the use of the mobile device does not significantly differ from typical mobile device usage, e.g. from checking SMS or answering a call.

3 USER-CENTRED DESIGN

Several studies [24-28] emphasize the involvement of stakeholders in the software development process as early as the first phases of the innovation and design stages. Developing an innovative mobile application and service is not a straightforward process. It requires several iterative experiments involving end users before any successful commercialization of the product. Blank [24] states that going backwards is considered a failure in the traditional linear product development model, whereas in the iterative development model going backwards is a natural and valuable part of learning and discovery. Ries [25] has also emphasized a cyclical development process where going backwards is an important element in continuous learning and a natural part of software development. Additionally, Blank [24] emphasizes that, unlike in the linear model, finding the right

customers and markets is unpredictable, and developers make several mistakes before they get it right. Blank also emphasizes the importance of the initial phases in his customer development model, namely determining customers' real problems and needs. Similarly, Ries [25] highlights the involvement of end users in the software innovation and development process, and claims you can begin the fruitful learning and discovery process with them by testing and measuring iteratively your minimum viable products and prototypes.

The paying customers do not necessary explicitly know what they need or want, but they can offer valuable comments and suggestions for your visualized drafts and prototypes [26, 27]. Those comments and feedback work as guidelines in searching the possibilities for successful product and service innovations. Moreover, usability and user experience considerations are increasingly important in contemporary mobile application developments [28, 29, 30]. Mobile applications' usability often mandates multi-level usability assessments. This complex and yet important process is accomplished by applying appropriate software development methods such as UCD or lean product development. Both of these methods consider users as key stakeholders at the various design and development stages.

Gould [29, 30] states that in a usable system we need to involve users continuously in the development process and based on their feedback refine the design concept. The term UCD was first used by Dan Norman during the 1980s after the publication of *User-Centered System Design; New Perspectives on Human-Computer Interaction* [31]. User-centred design and the development of interactive systems and devices have since increased in importance in product development as UCD both cuts costs [32] and improves usability. Additionally, it should also place a special focus on the business benefits, which are easier to identify when using rapid innovation methods where end users are involved in the same process as business owners.

The UCD process is divided into different phases of creating a usable mobile application [26-28]: 1) Concept phase: The users' needs and the opportunities are explored by applying different user study methods, such as interviews and questionnaires; 2) Requirements phase: Prepare a list of the requirements revealed during the previous phase. Applying various data analysis methods, such as interview transcripts, or task and environment analysis and affinity diagrams, assists in preparing such a list; 3) Prototype phase: The list of requirements is converted to a low-fidelity prototype and shared with the users. Based on users' feedback, the design is retuned; 4) Usability assessments phase: Users then assess the high-fidelity prototype through a usability test.

Unlike UCD, in which users are consulted at various stages of the mobile application concept development, the lean development principle is based on the values that the product provides to consumers. Lean principles originated from the lean manufacturing developed by Toyota. However, lean software development originated from a book written by Tom and Mary Poppendieck [33]; essentially, it is a software development model inspired by lean manufacturing and agile development principles. The lean model focuses on customer feedback and the reduction of waste. Based on lean software development principles, waste is defined as any part of the development process that does not create value for consumers.

Therefore, the first step in following the lean principle is to understand and identify activities that create value in a product. In the digital service business, this practically means that users will not be motivated to use a new digital service if they do not see or recognize added value or personal interest in the new digital solution [34, 35]. The second step in the lean principle emphasizes that quality ought not to be a separate phase, and instead requires consideration at all phases of the software development. Creating knowledge is the third step in the lean principle, and stresses sharing information among project workers and customers. Deferring commitment to the lean principle promotes the need for decision making at the last minute. The lean principle moreover recommends delivering smaller increments of the software product over shorter time intervals and promotes project workers as independent decision makers in their designated tasks, allowing them to achieve their goals more efficiently. Finally, the lean principle recommends optimizing the product

based on consumers' requests and hopes, just as UCD does [36].

Hence, in creating innovative and new applications in the software business, we are actually solving problems based on unknown proposals, as users cannot describe exactly what they need or want. Therefore, this requires iterative, lean and user-centred methods.

4 RESEARCH METHODOLOGY

The research approach is case-study [37] that applies the action research strategy [38], as the design team members were also actors in both the research and design work phases. Data was collected through semi-structured interviews and structured questionnaires from stakeholders during the design and development iterations. Table 1 shows the targets and phases of design iterations in designing the outdoor mobile guide service, the role and number of stakeholders involved and details of where the user encounters took place.

Table 1. Design iterations, stakeholders and context.

Design iterations of outdoor mobile guide service	Stake-holder	N=	Where the interview, interaction or evaluation took place
1. Business case and concept definition	Tourism company, ICT company and university-based tourism experts	6	Office
2. Validation of the technological solutions	Software developers	2	Street navigation, testing by kayaking on the Baltic Sea
3. Validation of the business viewpoint	Steering group members and two developers	6	Testing by kayaking on the Baltic Sea
4. Validation of the end users' viewpoint	Tourism university students	23	Testing by snowshoeing and hiking in the park
5. Validation of the expert users' viewpoint	Tourism experts	5	Testing by kayaking on the Baltic Sea
6. Validation of business scaling in the selected business segment	Customers and staff of tourism companies	50	Interaction at tourism centres and testing by kayaking on the sea and river

The study qualitatively analysed and applied the UCD principle to design a mobile web application for the outdoor tourism business. Outdoor tourism activities set special requirements for mobile application design processes. Therefore, in this paper we aim to reveal empirically the nature of the mobile application design process where several stakeholders are involved in the design and assessment actions throughout. The design process included tasks such as identifying requirements and creating mock-ups and prototypes with several iterative phases that utilize UCD principles.

5 DESIGNING and ASSESSING the MOBILE CONCEPT

5.1 Business Case and Concept Definition

The development team conducted a short study (n=6) to investigate the possibilities and the benefits of the mobile application for small tourism companies. The chief executive officer (CEO) of a tourism company, the CEO of a mobile software company, three tourism university lecturers and the director of a tourism education department participated in the survey and shared their insights on the possibilities for, and needs of an outdoor mobile guidance service. The university lecturers also work closely with several tourism companies, and therefore they have extensive knowledge of the possibilities and challenges for mobile services for tourism companies and their customers. The study mainly focused on revealing the possibilities for mobile services in this sector, and the tourism companies' expectations of, and business needs from, mobile-based services. By analysing the data the development team identified that location-based mobile services should include route information, a geo-location guide, maps suitable for outdoor

activities, points of interest, a link to the tourism company's e-commerce site, general and safety information and feedback functions. This was valuable information for our designers, and enabled them to propose an initial concept to the first partner companies. The designers themselves had prior experience of outdoor activities such as kayaking, camping and hiking, which helped greatly in coming up with additional features.

Figure 1. Example of the first conceptual plan and navigation structure.

Figure 2. Example of drafts of the mobile guide's graphical user interface design.

The proposed features were shared with companies, and their feedback was collected and analysed. On receiving the companies' confirmations based on the proposed features, we investigated the technological feasibility with programmers and software developers. Thus, the initial concept for a mobile web application to guide and inform kayakers and other outdoors tourists was designed and drafted (Figures 1 and 2).

5.2 Validation of the Technological Solutions

The first paper mock-ups and the proposed concept plans were assessed by the partner tourism company, which represented the voice of the customer at the beginning of the development period. After the partner companies confirmed the concept, the development team began implementation of the application prototype. The application prototype development was based on Vaadin [16], which is an open source rich internet application framework that consists of a server-side programming model and client-side development tools based on the Google Web Toolkit and HTML5. Moreover, it includes a server-side solution where the majority of the logic runs. Additionally, Ajax technology is used on the browser side to ensure interactivity and a better user experience, whereas on the client side Vaadin mainly utilizes the Google Web Toolkit, which is also used for rendering the resulting web pages. The framework is based on event-based programming and provides widgets and add-ons for helping developers to design and build faster and richer web applications and solutions.

Figure 3. Screenshots of the mobile guide application built on the Vaadin Java framework.

The Vaadin framework provides ready-made user interface elements, and therefore the visual outlook of the user interface was based on the

Vaadin framework's visual themes. Figure 3 shows screen shots of the mobile guide application prototype. The mobile application is compatible with iOS and Android devices, and the content is created, updated and managed by the external content editor, which was also built for this research project.

The software developers (n=2) themselves first assessed the functions of the prototype at the street level, as geo-location features require outdoor testing. They focused on assessing the software code related to the digital maps, navigation, digital compass and global positioning system (GPS). After the functionality tests on the street, the first field testing involved the designers using the mobile guide prototype in a target environment while kayaking on the Baltic Sea. This field test threw up new issues, which caused the development team to reconsider the design. First, the mobile device had to be in a plastic waterproof bag while kayaking, and this caused significant challenges for the user interface designers. Although this was recognized beforehand to some extent, the bright sunshine and glinting water caused considerable visibility challenges in the real usage environment. Therefore, the user interface designers were required to redesign, especially the visibility and size of the buttons and other navigation graphical user interface elements. In addition, the strength of the users' internet connection varied while on the sea, and while it was strong enough in places, in others it weakened significantly. Furthermore, the kayakers were often sitting near the sea's surface, and the bobbing of the water and breaking of waves also disturbed the internet connection. However, the GPS signal was excellent on the sea as unlike in cities there are no physical obstacles. The unstable or variable internet connection on the sea and archipelago surprised the development team, and caused further changes to the initial development plan. For example, the development team re-prioritized the requirements and dropped the features that required continual online access and made changes to the logic for the navigation software component. Before making these changes, the application prototype crashed every time during testing.

In the first plan, the development team aimed to include online route tracking and automatic saving in the service, but this feature had to be removed from the first release, despite it being flagged as useful by the tourism companies. It would have enabled the companies to record the kayaking routes of their customers and later allowed them to identify the most popular routes. It was decided to keep this feature in the product roadmap, with a view to reconsideration in the future.

5.3 Validation of the Business Viewpoint

The development proceeded to the second prototype, the so-called Alfa version, which implemented roughly 70% of the requirements and was ready for testing by a small group of potential customer companies. The steering group members of this project represented the potential customer tourism companies. Two group members were company representatives, one member worked as the CEO of an outdoor tourism company and another member was marketing director of a company specializing in location intelligence and digital map services. Therefore, they each had valuable viewpoints on the business perspective of the application. Overall, the testing group consisted of four steering group members and two developers (n=6).

The steering group members conducted the field testing on the archipelago of East Helsinki in Finland. They each had their own smartphones in plastic bags and the mobile web application was running in the background of their phones. The test group kayaked the *Kalliosaaren kierto* route with the help of the application guide. The test trip took five hours, during which we conducted semi-structured group interview to collect feedback on the users' experience and the development needs.

The overall analysis of the user feedback revealed that company representatives, i.e. business experts, were satisfied with the application. They expressed positive attitudes towards the concept and their experience of it as it helped them to navigate the archipelago and displayed the selected route information. The field test revealed that the application displayed the current kayaking speed and accurately calculated the distance to the next turning point and the final destination. If a kayaker turned away from the route, the navigation arrow turned red; when the direction was corrected, the

arrow returned to green, digitally guiding the kayaker back to the correct route.

The users also emphasized that they felt secure and safe by having the application at hand. Moreover, the field test demonstrated the potential business benefits for tourism companies. They stated that the mobile guide *"would work as a digital tourism guide that helps companies to scale business"*, which also indicates that better services could result in increased customer satisfaction. In general, the application received positive feedback, especially the user interface design that it was said supported use in challenging environments without any major technical or usability problems. The testers naturally suggested some new features, such as a warning mechanism, or at least better information about ferry routes, and a "home button" that would guide the kayaker directly back to the home harbour.

This testing round also raised some additional development requirements. For example, each route included several turning points, which can be defined on the editor site while creating a new route on the system. If a kayaker went 50 metres past a turning point the navigation compass began to point backwards, towards the turning point, unless the kayaker clicked the next-point button. Instructive information was required to do this, but smartphones' small screens do not allow the space to accommodate this extra information, and we found that only pop-up windows would be user friendly. The development team had already recognized this challenge, but no simple solutions were forthcoming; the software cannot know if you are lost and going in the wrong direction or whether you have passed the turning point but actually headed in the right direction. The solution requires more intelligence, and this need was added to the further development roadmap.

5.4 Validation of the End Users' Viewpoint

The tourism company Natura Viva arranged the field tests several months later, where (n= 17) tourism university students used the application while snowshoeing different routes on sea ice. The testers represented authentic end users as the tourism students were active in outdoor activities and many had specific experiences in the activities

for which the application was designed. Feedback was collected after testing using a structured questionnaire that included closed and open-ended questions. Thirteen of the seventeen testers returned the feedback forms; seven used an iPhone and six an Android device. Eleven of the thirteen used mobile applications daily or at least a few times a week. Similar tests were carried out two months later with a group of tourism students (n=6) that used the mobile guide application in navigating a route on the paths of a city park. Figure 4 presents a picture of the users' field test.

Figure 4. Tourism university students preparing to test the concept by snowshoeing a route guided by the prototype application.

The overall feedback we received was positive and optimistic. For example, typical replies to the question "Good things in this navigation application?" included: *"very clear with coloured arrows pointing to the right direction"; "it's simple and easy to follow"; "easy to see where you are, the direction and generally easy to use once it's started";* and *"it's really fun and easy and doesn't require any special skills to use it"*. The testers, however, raised development issues, such as: *"could add some sounds/feedback, and block the screensaver"; "the arrow was kind of restless at times, it was difficult to follow…"; "the current location should be in the centre…";* and *"sometimes the arrow was slow and showed green for every direction"*. These comments indicated that navigation was the key feature on which users focused, meaning it ought to work perfectly. The testers requested additional features, such as: *"a person telling you where to go, so you don't have to look at your phone constantly"; "voice feedback";*

"amount of steps taken and a voice telling you where to go"; and *"info about the things to see around you".* Some testers thought the application already had everything they would need, stating: *"Nothing. If I need a navigation app, navigation features are all I need";* and *"I can't think of any. Everything I needed was there".* The testers did not raise the requirement for radical changes or any compulsory need for new features. As the previous examples demonstrate, the users believed that the basic features would work well, and that any new features would improve and simplify the user experience, but only when the basic functions worked properly. The basic features are the foundation for the use of more advanced features, i.e. secondary features, for use only if the basic features of the application bring value. The basic features are essentially a doorway to the use of mobile applications.

5.5 Validation of the Expert Users' Viewpoint

The design team arranged a two-day testing trip where tourism experts (n=5) used the mobile application while kayaking three routes on the Helsinki archipelago of the Baltic Sea (Figure 5). The experts were a German navigation consultant and trainer with vast experience of outdoor tourism, wellness activities and associated navigation devices; two tourism experts (a lecturer and a project manager) from the tourism education department of a university; the CEO of an outdoor tourism company that offers services, e.g. kayaking, for tourists in the Helsinki area; and the project manager of this research and development project. Three of them had prior experiences about the application from the previous testing iterations. All participants used the mobile guide application while kayaking pre-defined routes during the trip.

Data related to the usage experiences and development proposals was collected from the semi-structured group interview conducted during the test trip while participants were using the mobile guide. The framework of interview themes concerned the benefits for users, development needs, technical problems and the business potential in European markets. Overall, the user experience was positive, although a bug in the iOS 6.0 operating system blocked the automatic geo-

location feature in the Safari browser. The automatic location information worked for a spell and then stopped, and the application had to be restarted.

Figure 5. A tourism expert testing the mobile guide while kayaking.

The test users raised significant benefits related to safety and entertainment aspects. The mobile guide provided users with the feeling of safety as it displayed their current location, speed and distance information related to the route. There was no similar application known to exist in European markets that allowed tourism companies to edit and manage content and with a user interface designed for use during outdoor activities, particularly kayaking and hiking.

5.6 Validation of Business Scaling in the Selected Business Segment

The design team arranged (n=4) company pilots after the application concept and functional prototype were readied for performance and usability assessments. Three company pilots were arranged with tourism companies that largely offer their customers kayaking, canoeing and hiking services on the sea or in river areas. The fourth pilot involved a hotel in Lapland, Finland, that targeted hikers.

All of the companies received their own application's internet address or uniform resource locator (URL) with their own content including points of interest, route content and instructions. The development team also printed a poster that invited customers to test the mobile guide application. The end users were personally met at the tourism premises, except the hotel guests were

planned to leave their written feedback in the hotel reception; due to its distance from Helsinki as face-to-face meetings were not feasible.

The design team conducted end user interactions with approximately 50 customers and staff members (n=50) in three different kayaking or paddling centres. Most end user interactions took place on land while users were preparing for their kayaking trip. In the first phase of end user interaction, the interview mainly focused on the interest of the new proposed mobile guide service, and on the second phase, if the potential end users interested to review or test the application, the user experience and technical functionalities.

In addition to the feedback documenting, data was also collected from the unstructured feedback function of the application. The users had possibility to give written feedback to the designers through the feedback functions of the application. Examples of the user experience and technical feedback received were: *"Does not work in HTC Sensation Z710e phone, application does not recognize the phone, e.g. Google Map works fine in the same phone and it has been tested over one day and the application started several times";* and *"The application basically works, but minor failures disturb usage".* The users also gave practical ideas for how to improve the application, such as: *"The navigation arrow could be larger…"; "…you could add a new feature that helps users: A touch and a user indicator (blue point) on the screen would change visible point instead of target point… …the blue point would appear "wave" element…";* and *"The approaching colours (green/red triangle) would change in the app. based on a 90-degree sector. Could you also do this for 45-degree sectors?"*

In this testing iteration, the design team also shared the mobile guide prototype with the staff of tourism companies and discussed with them their expectations and requirements. The staff of tourism companies were eager to know how their customers accepted the mobile guide application, and what kinds of benefits it could offer their business. All of them believed that mobile services would form an essential part of the tourism business in the future, with the "automated guide" and "digital tourism guide" seen as the mobile applications with most potential in outdoor activities, given that their

customers expect guide and route information. Investing in this would probably increase customer satisfaction and encourage fearful kayakers or hikers to purchase an outdoor trip or rent a kayak. However, the managers' open question remains: How will they earn a return on their investment if the mobile service is offered free of charge?

5.7 Limitations of the Pilot Testing

On starting prototype testing at the tourism company premises in the sixth iteration, we expected to encounter eager users because almost all of them were interested in the mobile guide application at the previous phases. However, the results of assessing the prototype indicated that only a few of the end users pro-actively downloaded or scanned the quick response (QR) code while participating in activities such as kayaking or hiking. Most of the end users who tested the application and provided feedback were asked to use the application, but few downloaded it voluntarily. The steps involved in testing consisted of loading or scanning the application to the browser, learning the key features, using it during the trip and reporting their experiences.

Kayaking is an activity that requires you to concentrate to maintain balance, which means if you are not careful there is a risk of damaging or wetting your own mobile phone while getting familiar with the functions of the mobile application. That is reasonable justification for why many kayakers did not try to use the new application while focusing on their kayaking. As a result, several users liked to hear and speak about the application, but were hesitant to try it in practice. The most active and enthusiastic users were those with some experience of mobile navigation devices; therefore, we call this group of users the early adapters and expert users. Often, the early adapters are interested in new technologies in their field of interest.

Despite the problems discussed above, during the prototype tests in the field environment the design team did collect enough feedback from the active potential users who voluntarily tested the application. In addition, conducting interviews with potential users provided us with substantial relevant feedback and increased our understanding of users'

expectations and priorities. However, it is important to mention that the hotel prototype testing was not as successful as the other company pilot tests because the design team failed to conduct face-to-face interviews with the target users. As a result, we gathered little relevant feedback from the hotel-based end users. It is difficult to estimate how many of them actually tested the application, but few returned the feedback form. Nevertheless, the post-interview with the hotel manager was beneficial and provided new insights regarding the application. The timing of the prototype testing was not optimal as most users hike in the autumn. In addition, most customer at the hotel were retirees and thus not the most eager testers of new mobile applications. Furthermore, the design team could not remotely support and motive potential end users in testing the mobile service in practice. The design team learned that visitors to hotels are not spontaneous or proactive testers, and instead need to be tempted or convinced to test through motivating or helping them in some way. Similar findings emerged from other piloting premises; the users needed to be motived to provide formal feedback.

6 DISCUSSION

6.1 The Key Roles of Stakeholders in Mobile Service Design

The UCD process should adopt a broad perspective and not focus only on the end users if a successful new service is to be produced. Instead, software and ICT services must satisfy other stakeholders, such as business managers, the software team, project management and, in many cases, the project financiers and providers [39, 40, 41]. It is essential to begin the iterative development process by identifying the most significant stakeholders and their impact on the overall application in UCD. This novel application design and development is a result of merging tourism business offerings with existing technological possibilities. Co-operation and interaction between software developers, project management, tourism experts, tourism companies and their potential customers provided useful resources for defining specification requirements

and realizing the expected user interface design. Conducting various prototyping phases, continuous testing events with target users of the application and several interview sessions with tourism companies' management have provided us with a great deal of valuable data. The analysis of the gathered data in this case has helped significantly in defining and prioritizing features and identifying the potential stakeholders' demands. Figure 6 demonstrates the three key stakeholders groups involved in the application concept design process.
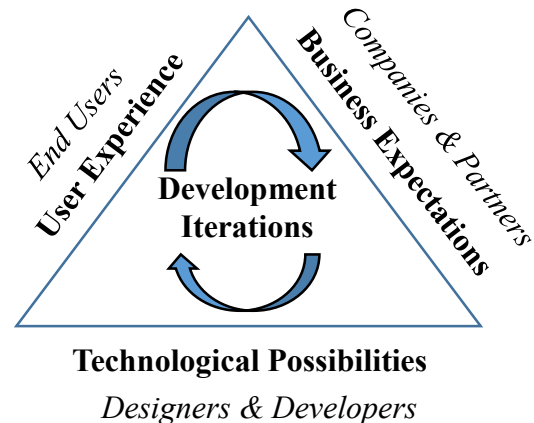
Figure 6. The key stakeholders and their contributions to the mobile service design process.

This paper emphasizes the importance of involving the most significant stakeholders in the design process, not only at the beginning of the project but throughout every phase of the design and development work. The diversified interaction and collaboration with end users and potential customer companies helped us to focus on the most important and requested features and to continually prioritize the product backlog, the development plan and the product roadmap.

6.2 The Main Contributions of Each Stakeholder

This paper emphasizes that working in a cross-disciplinary team ensures a broader viewpoint during the design process. The users' involvement may affect the specification requirements from the elicitation phase onwards, right up to the prototype evaluation phase. Therefore, as Figure 7 shows, using the mobile guide application can be categorized into three main phases: the end users

use it before their trip mainly for route planning; during their trip it is mainly for navigation and information about the route and points of interest; and after their trip they can share their experiences. The system administrator takes care of the configuration and the tourism companies' personnel manage the route and information about points of interest.
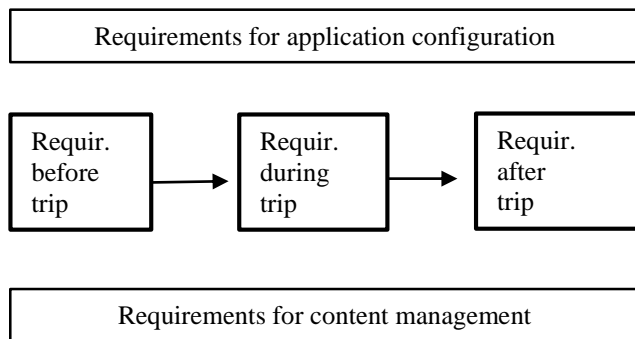
Figure 7. Requirement categories in designing a mobile guide concept for tourism businesses from the viewpoint of the end-user journey.

This design experiment showed how differently business managers and end users might focus on features that software developers view as self-evident. As an example, users were given the option to enable or disable the objects in the application's user interface, which meant they had the freedom to customize the application content. However, the end users did not speak about such possibilities, and it is probable that few took advantage of them. Customization features are usually advanced features, and most end users probably focused on the basic features while initially testing the application. This finding emphasizes the importance of road mapping in the research and development processes, especially from the viewpoint of how the users will prioritize features. By testing a set of features, we have been able to prioritize them and determine the key features, the so-called critical success factors, for new mobile services in the tourism market. It is also important to note that often there are conflicts in interest between the stakeholders involved in application development. As an example, the main concern of the tourism companies' management is

the business benefits, such as more revenue, increased customer satisfaction and differentiation from competitors, whereas their customers or end users are expecting direct benefits for their outdoor activities, such as automatic route guides and more relevant information.

6.3 User Involvement in Application Design

The management of pilot companies naturally considered the mobile services in a broader context than that considered by the end users. Therefore, there were several "layers" of interest within the development of the mobile service. The UCD principle ensures that all stakeholders' interests are taken into account, and several users have various roles in the digitalized service process. The prototype testing results indicated that the mobile service precipitated unanticipated and new added value for small tourism firms and their customers.

This project found that end users place significant value on the core features that offer them most benefits and satisfaction. However, these features have to be implemented in a proper way, one that results in an excellent user experience. The lack of proper functionalities or a poor design process results in a service failure from a usability point of view. Hence, the developers should recognize those features and their nuances as quickly as possible and prioritize their development effort by focusing on them. This can be achieved using rapid development methods and a mobile web development framework, and its add-ons, to shorten the lead-time of software product development cycles.

Mobile applications need to attract and attach themselves to the users emotionally to achieve sustainable usage [42, 43]. This is even more important in outdoor navigation applications, where the user not only has to emotionally attach to the application, they must also trust and feel secure with the application, a point raised during the field testing phase of this application. This is particularly important because the target users of the applications are often kayakers, joggers, cyclists or hikers who are alone or in a small group in rugged or rural areas. Moreover, the mobile service usage and functionalities should be easy to load and fast to navigate, and the functioning logics should

follow the user's mental models, i.e. their prior assumptions.

The field tests revealed that the end users and companies' staff members were satisfied with the overall concept and the application, and viewed it as a useful service that would provide added value for outdoor activities. Interestingly, many active kayakers used waterproof mobile devices and had downloaded some form of tracking application to their devices. The end users liked the information offered by the application, such as points of interest, guided route information, speeds and the distance to the destination. For instance, some users liked to see how fast they were able to kayak on a particular route or how far away the destination was, or to be able to locate the nearest point of interest.

6.4 Managing the User-centred Design Process

The involvement of several stakeholders in the design process requires careful project planning and management, as each stakeholder reviews the concept, requirements and user experience from a different perspective. As Ries [25] states, it is easier to build known software products for established markets than to innovate unknown product concepts for emerging early markets. Therefore, to optimize the use of resources and shorten the time-to-market, it is recommended to plan carefully which features could be tested using scenarios, screenshots, sketches and other mock-ups without the need for coding the application itself.

Table 2 shows six design iterations that involve different stakeholders and validation targets and content. The testing of usability and the researching of business benefits involve different development goals, although they can be tested using the same prototypes. This study shows that all six of these iterations bring value to the mobile service design process and have a unique role in the process. It is also recommended to involve all stakeholders in the design process as early as possible, as they will affect its success on launching the mobile service in any case.

Table 2. Design iterations, stakeholders and the targets of each design iteration.

Design iterations	Stakeholders and target of design iteration
1. Business case and concept definition	Business owners: Produce initial concept plan and visual drafts to create a common understanding of the goals
2. Validation of the technological solutions	Software developers: Validate technological solutions by testing functional prototypes in the real usage context
3. Validation of the business viewpoint	Company representatives: Validate the business viewpoint by assessing the concept using prototypes
4. Validation of the end users' viewpoint	Potential end users: Validate the end users' viewpoint by assessing the concept using prototypes
5. Validation of the expert users' viewpoint	Domain experts: Validate the marketing, sales and service viewpoints by assessing the concept using prototypes
6. Validation of the business scaling, delivery and service needs	Company pilots: Validate the scalability, delivery and service needs by piloting mobile services with real potential customer companies

This study points out that usability issues are only part of successful product design. For example, the managers of tourism companies were most appreciative of business benefits, while the end users focused on the guidance and navigation benefits of the application. To obtain relevant feedback in field testing situations, interviews and performance assessments in real usage environments should focus on emotion- and motivation-related issues, as they have a direct impact on the financial success and scaling of a new mobile service.

7 CONCLUSIONS and FUTURE WORKS

This case study, accomplished using an action research strategy, contributes to software development methods by describing an experiment in which users were involved in the mobile service design through six different design and assessment iterations. Each of the iterations had its own role and specific stakeholder group, and each helped the design team to gain rich and versatile user experience data. As a UCD method it worked

logically, and each phase synchronized with the others.

The study revealed that iterative UCD increases significantly the number of stakeholder touch points, and therefore helps to integrate business and user needs with new technological possibilities. We have also shown that the involvement of various stakeholders in new mobile service development ensures better integration of business expectations, mobile web technologies and user experiences. Hence, the UCD approach should adopt a broad perspective, not only the viewpoint of end users, as the successful new mobile service must satisfy business managers, end users, software developers, project management and, in many cases, the project financers.

The latest cloud computing, mobile technology advancements and UCD methods have created a unique opportunity to boost digital services, especially in the small business market segment. Hence, this requires fresh research on how such new digital services can be designed and developed more quickly by involving various stakeholder groups in an iterative design process.

ACKNOWLEDGEMENT

REFERENCES

1. ISO. Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems (ISO 9241-210:2010(E)). ISO. doi:10.1039/c0dt90114h, (2010)

2. Carlin, S., & Curran, K. Cloud Computing Technologies. International Journal of Cloud Computing and Services Science, 1, 59–65. doi:10.11591/closer.v1i2.486. (2012)

3. Tahir, A.O & Adeyinka, A. Autonomic Service Management in Mobile Cloud Infrastructures. In International Journal of New Computer Architectures and their Applications, 4(2): 91-107 (2014)

4. Kesavan, S. & Jayakumar, J. Network Performance Analysis of Cloud Based Multimedia Streaming Service. International Journal of New Computer Architectures and their Applications, 4(3): 156-166 (2014)

5. Makoto S., Masami M., Hiroshi D., Yoshiaki Y. Ciphertext Diverge-Merge Scheme of Identity-Based Encryption for Cloud-Based File Transmission Service. International Journal of Digital Information and Wireless Communications 5(1): 52-59 (2014)

6. Khanghahi, N., Nasiri R, Razavi M. A New Approach Towards Integrated Cloud Computing Architecture. International Journal of Digital Information and Wireless Communications 4(1): 24-34 (2014)

7. Benlian A., Hess T. & Buxmann P.. Drivers of SaaS-Adoption – An Empirical Study of Different Application Types. In Business & Information Systems Engineering 5 (2009)

8. Turner, M. and Budgen, D. and Brereton, P. (2003) 'Turning software into a service, Computer, 36 (10). pp. 38-44.

9. Liu, Z., Le Calvé, A., Cretton, F. & Glassey, N. Using Semantic Web Technologies in Heterogeneous Distributed Database System: A Case Study for Managing Energy Data on Mobile Devices. International Journal of New Computer Architectures and their Applications (IJNCAA) 4(2): 56-69 (2014)

10. Laplante, P., A., Zhang, J., & Voas, J. What's in a name? Distinguishing between SaaS and SOA. In IT Professionals vol 10, 46-50 (2008)

11. Josuttis, N. M. SOA in Practice. October (Vol. 253, p. 352). (2007)

12. Snell, J., Tidwell, D., & Kulchenko, P. Programming Web Services with SOAP. Language (p. xiii + 244). (2002)

13. Vaadin. Introducing Vaadin += GWT, Web Application Scalability : Vaadin goes Movies. Dock: World of Web Applications, 1–52. (2010)

14. Gröönroos, M. Book of Vaadin: Vaadin 6.4 https://vaadin.com/download (2010)

15. PhoneGap. http://phonegap.com/community/ [accessed 17.2.2015]

16. Alamäki, A., Kaski, T. Characteristics of Successful Sales Interaction in B2B Sales Meetings, International Journal of Social, Education,

Economics and Management Engineering, 9(4), 1037 – 1042 (2015)

17. Prahalad, C.K., and Ramaswamy, V. "Co-Creation experiences: The Next Practice in Value Creation," Journal of Interactive Marketing, 18 (3), pp. 5-13 (2004)

18. TripAdvisor. Mobile Device Travel Survey http://www.tripadvisor.com/PressCenter-i4720-c1-Press_Releases.html [accessed 27.10.2014]

19. Google. The 2014 Traveler's Road to Decision. June 2014, Google/Ipsos MediaCT. (2014)

20. Gavalas, D. & Kenteris, M. A web-based pervasive recommendation system for mobile tourist guides. Personal and Ubiquitous Computing 15:759–770, (2011)

21. Kennedy-Eden, H. & Gretzel, U. (2012) A taxonomy of mobile applications in tourism. E-review of Tourism Research, 10 (2), 47-50, (2012)

22. Gavalas, D., Konstantopoulos, C., Mastakas, K. & Pantziou, G. Mobile recommender systems in tourism. Journal of Network and Computer Applications 39, 319–333, (2014)

23. Schwinger, W., Grün, Ch. Ch. Pröll, Retschitzegger, W. & Schauerhuber, A. Context-awareness in Mobile Tourism Guides – A Comprehensive Survey. Handbook of Research on Mobile Multimedia, Second Edition, Vol 2, 298-314. (2007)

24. Blank, S.: The Four Steps to the Epiphany: Successful Strategies for Products That Win (2007)

25. Ries E.: The Lean Startup. How Constant Innovation Creates Radically Successful Businesses. London: Penguin Books (2010)

26. Dirin, A., Nieminen, M.: State-of-the-art of M-learnings ability and user experience. The Fourth International Conference on e-Learning (ICEL), Ostrava, Czech Republic, 201 (2013)

27. Dirin, A., Nieminen M.: Framework for Addressing Usability and User Experience in M-Learning, 2014. Journal of Computers (in press) (2014)

28. Nieminen, M.: Information Support for User-Oriented Development Organisation - Considerations Based on the Construction and Evaluation of Knowledge Storage, (2004)

29. Gould, J,D. & Lewis, C.: Designing for Usability: Key Principles and What Designers Think.

Communications of the ACM, 28(3), 300-311 (1985)

30. Gould, J.D., Boies, S,J. & Ukelson, J.: How to Design Usable Systems. In M.G.Helander, T.K.Landauer and P.V. Prasad eds. Handbook of Human-Computer Interaction. Amsterdam, Netherlands: Elsevier Science/North Holland (1997)

31. Norman, D., Draper, S.: User Centered System Design; New Perspectives on Human-Computer Interaction. L. Erlbaum Associates Inc. Hillsdale, NJ, USA (1986)

32. Bossert, J.L.: Quality Functional Deployment: A Practitioner's Approach. ASQC Quality Press. New York (1991)

33. Poppendieck. T. & M.: Lean Software Development: An Agile Toolkit. Addison-Wesley Professional (2003)

34. Kaski, T., Alamäki, A. & Moisio, A.: A Multi-Discipline Rapid Innovation Method. Interdisciplinary Studies Journal, Vol 3, Number 4, (2014)

35. Alamäki A. & Mäkinen K.: Competence Development Supported by Digital Means in a Knowledge-Intensive Company. In P. Nicholson, J. Thompson, etc. (Eds). E-Training Practices for Professional Organizations. Boston: Kluwer Academic Publishers (2005)

36. Dirin, A., Nieminen, M. "Managining m-learning application development: Roles and Responsibilities". ICAICTE 2014. (2014)

37. Eisenhardt, K., M.: Building Theories from Case Study Research. The Academy of Management Review, Vol. 14, No. 4 (Oct), pp. 532-55 (1989)

38. Brydon-Miller, M., Greenwood, D., & Maguire, P., Why action research? Action research, 1(1), 9-28. (2014)

39. Alamäki, A & Dirin, A. (2014). Designing mobile guide service for small tourism companies using user centered design principle. The International Conference of Computer Science, Computer Engineering and Social Media, Thessaloniki, Greece. December 12-14, (2014)

40. Mäki, M. & Alamäki, A. (2014). Developing and boosting online grocery shopping through the customer-oriented field study method. 7th International Conference of Education, Research and Innovation, Seville, Spain, November 17-19, (2014)

41. Hotti, V., Meriläinen, H., Salomaa H. & Lahtela, A.: ICT governance: Literature review and Survey in Northern Savonia. International Journal of Digital Information and Wireless Communications 4(3): 333-340, (2014)

42. Dirin, A., Nieminen, M., Kettunen, M.: Student capabilities to utilize m-learning service in new smart devices. International Conference on Advanced ICT for Education. ICAICTE 2013. September 20-22, (2013)

43. Hooper, M.A., Sant, P.: Understanding the Mobile User's Purchase-Decision Involvement. International Journal of Digital Information and Wireless Communications 5(3), 165-178, (2015)

Evaluating Cloud Computing Challenges for Non-Expert Decision-Makers

Karim Mualla *(Author)*

School of Energy, Geoscience, Infrastructure and Society
Heriot-Watt University
Edinburgh, United Kingdom
kjm33@hw.ac.uk

David Jenkins *(co-Author)*

School of Energy, Geoscience, Infrastructure and Society
Heriot-Watt University
Edinburgh, United Kingdom
d.p.jenkins@hw.ac.uk

ABSTRACT

Non-expert managers in various types of organizations usually find it challenging to decide whether to adopt cloud computing services, which are internet-based, instead of the conventional in-house technologies, which are physically owned and controlled on-premises. This paper evaluates cloud computing management challenges that were selected and ranked by a number of non-expert managers through a decision-making survey. The outcome argues that the Urgent Support Availability aspect is selected as the most worrying factor amongst the majority of non-expert managers taking into account the associated cost, and future demand changes. In addition, multiple decision-making considerations are identified through a cloud utilization framework and evaluated from a non-expert management perspective in relation to performance, network reliability, integration, quality of service, and actual business benefits.

KEYWORDS

Cloud Computing, Challenges, Decision Making, ICT Management, Non-expert.

1. INTRODUCTION

A basic definition of cloud computing for non-expert clients is the use of the Internet for the tasks performed on computers. The Cloud here represents the Internet. Virtualization of processing power, storage, and networking applications via cloud computing platforms allows organizations to operate heavy demand computing resources off-premises. While this approach reduces in-house costs and energy use, recent case-studies have highlighted complexities in the decision-making process associated with implementing various models of cloud computing. This complexity is due to the rapid evolvement of these technologies without standardization of approach by today's top providers. In addition, the difficulty of understanding and predicting ICT demand growth in organizations has caused managers not to take advantage appropriately of the cost-saving factor which accompanies cloud computing as marketed by top providers.

Non-expert managers look at cloud computing as the process of taking the services and tasks performed by computers and bringing them to the web. To a large extent this is correct. However, cloud computing technologies offer a wider range of processing, networking, and storage capabilities which assist organizations in performing many heavy or small ICT tasks at the cloud providers' datacenters. This is mostly achieved through on-demand, remotely-controlled, scalable, and pay-as-you-go approaches. In many cases, conventional in-house methods were observed not as cost-effective as current cloud-based services. Nevertheless, several challenges were observed in the long-run due to management aspects which can be summarized as follows: [1]

- Improper analysis of the organization's actual ICT requirements

- Unclear contracts with the cloud provider
- Unreliable internet performance which affects the entire cloud service-delivery process
- Security considerations and data integrity issues regarding the methods in which the cloud providers handle access to resources, store data, and secure the virtualized server infrastructure.

The above aspects can result in additional expenses and management complexities in the long-term as will be discussed in this paper. The different models and techniques of cloud computing deployments and services have a significant impact on the decision-making process in any organization's ICT environment. This paper discusses various management challenges of cloud computing, which were voted as most relevant by a number of non-expert managers through a risk-analysis survey. The outcome is analyzed against present decision-making considerations regarding the adoption of different types of cloud computing services. The study highlights non-expert clients as key users of the outcomes from this project given the diverse work objectives across today's organizations.

The structure of this paper is divided as follows: Section 2 will introduce briefly a background on cloud computing from the perspective of non-expert clients. In Section 3, a brief literature review will be discussed regarding cloud computing challenges and end-user potential risks. Section 4 will evaluate the client-cloud computing management challenges regarding the three cloud service delivery models (IaaS, PaaS, and SaaS). In Section 5, the selected client-cloud computing management challenges will be analyzed through a decision-making risk-analysis survey which targeted 54 non-expert management-level users. Following this, the paper will analyze data results and evaluate the outcome of the survey in relation to previously stated decision-making aspects. Section 6 will discuss a decision-making framework to potentially overcome the earlier challenges. At the end, conclusions and future works are listed.

2. BACKGROUND

Whether ICT clients realize it or not, cloud computing services are being used on a daily basis and for a long period of time. For example, internet email accounts, social networks, GPS locations, and numerous other forms of online data storage and sharing are constantly being accessed by millions of users worldwide [2]. These services are supplied by ICT providers that own virtualized datacenters for end-users to access through the Internet. In general terms, cloud-computing is a ubiquitous platform which provides on-demand ICT services through either the public Internet, or other privately-managed and secure tunneling networks like Virtual Private Networks (VPN) [3]. The Cloud concept came to life mainly because of the growing ICT requirements in almost each industry, which were not being fulfilled through previous models due to costly services and complex management procedures. However, multiple tradeoffs and challenges have risen as a result of the rapid evolvement of these technologies, while other challenges have remained from previous ICT models.

Several cloud computing scientists and organizations have identified different characteristics, service-delivery models, architectural types, and legal aspects of a system necessary to support cloud computing. According to the NIST definition of cloud computing concepts, five essential characteristics were necessary: On-Demand Self Service, Broad Network Access, Resource Pooling, Rapid Elasticity, and Measured Services [4]. In addition, experts from The Cloud Security Alliance have identified a sixth cloud characteristic and named it Multi Tenancy [5]. Furthermore, another client-cloud computing characteristic was widely discussed by many organizations is the Economy of Scale [6], which indicates the distributed manner of computing access and sharing of resources across the cloud. This characteristic is significant to this paper given the security considerations needed for non-expert clients to evaluate before signing contracts with the cloud provider.

Cloud computing hosting models were divided into four interrelated models as follows:

- Public: Cloud providers offer a full range of computing services via online means, which enables organizations to outsource the entire ICT infrastructure into the cloud.
- Private: Organizations operate either on-site, exclusively managed, or via a third-party outsourced cloud, or a combination of both.
- Community: Multiple organizations with similar operational goals and security policies, share the same virtual ICT services and platform, which can be managed by one of the above, a third-party, or a combination of all.
- Hybrid: Often the most preferable cloud deployment method for end-users, as it ensures additional management flexibilities regarding security, risk elimination, information systems portability, and better standardization. The hybrid solution offers a mixture of various sub-components from previous deployment approaches. In particular, this model irrespectively combines the technical and nontechnical aspects from Private, Public and Community models [7].

Moreover, client-cloud computing potential benefits extend beyond obtaining cost reductions and management flexibility. On this note, multiple energy saving characteristics were pointed out by academics and service providers given that ICT virtualization can have a significant potential for eliminating plugged-in equipment, thus minimizing associated electricity consumption, space and management. The Green characteristics of cloud computing are summarized as follows: [8] [9].

- Dynamic Provisioning: The ability to reduce unwanted cloud computing components through better matching of server capacity with actual clients' demand.
- Multi-Tenancy: The ability to normalize and flatten unmeasured peak loads by serving large numbers of clients on a shared hosting infrastructure.
- Server Utilization: The ability to operate servers at higher utilization rates via virtualization techniques.

- Data Center Efficiency: The ability to use advanced datacenter features which reduce the overall power loss through improved methods of power conditioning, air cooling, and other methods.

3. LITERATURE REVIEW

According to Carrenza and HP, upgrading an existing ICT system for three consecutive years is more costly than the system itself. This was studied on applying intensive cloud solutions for several large organizations across the United Kingdom [10]. As a result, these providers identified several vital security aspects related to the concept of virtualization in which non-expert users must thoroughly understand before outsourcing their critical business applications onto the cloud. On that ground, several reliability concerns were raised by clients after adopting software, platform, or infrastructure cloud services for at least a 1-year lifecycle.

One report argued that a slower pace of virtual ICT adoption is currently spreading across large organizations, simultaneously with the rapid evolution of cloud techniques [11]. These risks were argued to range from technical, management, all the way to legal aspects of ICT employment. Furthermore, numerous standards for specific industries were argued to be missing regarding optimizing the way in which cloud services are disparately purchased, supported, and governed.

One of the major issues in standardizing cloud computing is the large range of different purchase standards and technical definitions. Currently, these were estimated to reach nearly 160 different definitions around the world [12].These standards began developing in 1999 when Salesforce introduced the first online-based application [13].

A number of definitions and standards of cloud computing were published by top ICT providers such as Cisco, Microsoft and IBM. Many academics and papers stated that these cloud standards are developed inaccurately given that ICT providers usually tend to market their services in order to increase sales against other competitors

[14]. These actions cause decision-making challenges for non-expert clients as will be discussed in the next section.

Other general assumptions were arguing for the outsourcing of non-core ICT capacity into a third-party provider that owns the infrastructure. However, numerous growth-limiting barriers were explored concerning knowledge sharing and data breach risks [15]. Adopting a fully outsourced cloud computing solution is currently considered an unfavorable decision by most non-expert managers given the uncertainty of private data whereabouts and many other considerations related to less control over owned resources [16].

Other concerns regarding credibility and authenticity of cloud services were observed among managers from different organizations. According to a survey by the IDC Enterprise Panel in 2009, the following barriers were identified and rated depending on the level of concern in contrast to the acceptance percentages attained from purchasing on-demand cloud benefits (Figure 1) [17].

4. EVALUATION OF CLIENT-CLOUD MANAGEMENT CHALLENGES

With regard to the three primary service layers of cloud computing (SaaS, IaaS, and PaaS), the following table was constructed to evaluate the reliability and security challenges from the perspective of non-expert clients and in relation to each cloud layer separately (Table 1).

Many ICT providers are currently analyzing the adoption patterns in which their clients are turning towards cloud computing [18]. Their objective behind this analysis is essentially to identify the key concerns and challenges regarding why many clients are still reluctant to adopt cloud computing services for businesses and heavy ICT tasks. The following section presents a risk-analysis survey which addresses the point of view of 54 non-expert decision-makers from different specialties.

## 5. SURVEY-BASED EVALUATION FOR NON-EXPERT MANAGERS

This paper conducted a risk-analysis survey which targeted 54 non-expert management-level personnel form different organizations. The purpose was to collect data on cloud computing tradeoffs and management risks by following the viewpoint of non-expert decision makers across different types of industries. This survey includes a single rating-scale question which offers 5 multi-choices as available answers.

The Likert approach was selected for this survey given the nature of opposing opinions among different non-expert managers [19]. This was identified by this study from observing different ICT management aspects, such as the degree of concern towards the utilization of novel technologies among managers with medium-level technical background.

Table 1. Evaluation of Client-Cloud Computing Management Challenges regarding each Service-Layer

Cloud Model	Description	Example	Evaluations of Challenges
SaaS (Software as a Service)	Users access applications via network-hosted infrastructure like the Internet or VPN (e.g. Gmail).	Gmail, Blogger, Cisco WebEx, Flicker, Windows Live Meeting, Windows Office Live	Given that SaaS is mostly offered free of charge, or accompanied as an additional service with larger paid solution, Software is not installed on the users' servers or personal PCs. Therefore, access can occur strictly on-demand. As a result, only confined functionalities, selected configuration, service availability issues, and limited control of programs -to underlying ICT developments- are provided by the provider following the SaaS approach.
IaaS (Platform as a Service)	Clients develop software via a fully network-hosted platform	Force.com (development platform), GoGrid, Facebook Developers	Underlying cloud solutions, in addition to several dependencies like storage, network, servers and operating systems, are not controlled by the service-requester. However, more control is available than the SaaS model, as the main IT environment in the PaaS approach is considered *Closed* or *Contained* [20]. Nevertheless, availability restrictions are still considered a trade-off for managers against the traditional physically-owned ICT infrastructure.
PaaS (Infrastructure as a Service)	The Cloud provider rents out hardware, software, networking bandwidth, processing power, or data storage via virtual, on-demand accessing policies	Amazon EC2, IBM Cloud-works, Windows Azure	Even though non-expert managers have, to some degree, the ability to control, deploy, and run user-created programs such as operating systems, privately developed software, and networking components, however, the underlying cloud solution, is again, primarily managed by the cloud provider. Therefore, security access of information, user-group permissions and other administrative dependencies are all identified as management concerns.

(Scale: 1 = Not at all concerned 5 = Very concerned)
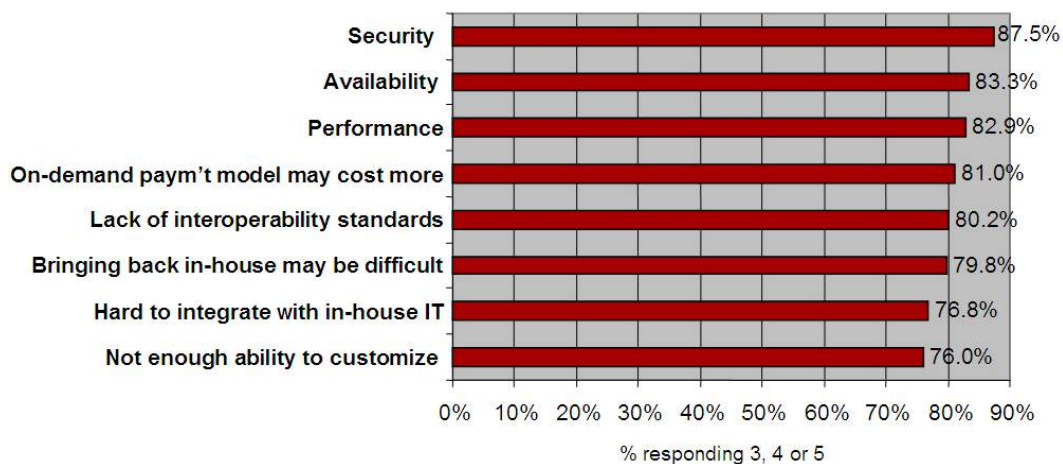
Figure 1. A 2009 Survey on Cloud Computing Management Concerns

The survey attempts to reflect the diverse attitude of these managers towards ICT budget acceptance, sustainability readiness, change management, and other organizational aspects [21]. The rating-scale survey was conducted via the popular online provider: Survey-Monkey. Furthermore, this survey was not structured to target a specific audience from a particular industry given that each relates to a different specialty and work nature, hence, is subject to dissimilar ICT requirements.

At first, we asked 54 decision-makers to select 12 statements of cloud computing adoption risks and challenges whereby in their opinion are most relevant to their organization. These were picked from a bigger pool of ICT risk statements and cloud adoption challenges which were picked from previous surveys and literature. We then asked the interviewees to rank those 12 categories according to their organization's priority by selecting one of five multi-choices which reflect five levels of concern. Table 2 presents the answers and the calculated percentages of the total ranking regarding each risk category. In addition, Figure 2 presents the completed survey findings via a bar chart, which was generated via Microsoft Excel from data inputs of Table 2.

It can be observed that the *Urgent Support Availability* aspect was classified as the most worrying factor among non-expert managers. This was demonstrated with a 4.13 average rate out of 54 participants in contrast to the rest of the statements. The *Government Hosting Regulations* came as the lowest concern and only received an average of 2.24. The two price-associated factors: *Unpredictable Costs in the Future* and *The 'on-demand' payment method of cloud computing might actually cost more than the traditional approach*; both came at positions 3 and 4 in order. In addition, the *security* risk category landed as the second most worrying aspect following *the rapid delivery of unpredictable maintenance*.

As mentioned earlier, given that the science of cloud computing is evolving at a faster pace than most of the other services provided by various industries, it is important to identify the patterns and changes in collecting data results when performing similar surveys across time. This risk-analysis survey was intended to illustrate a relatively different viewpoint of earlier cloud computing surveys. For instance, the IDC survey in 2009, which was discussed in the literature review section previously, has covered slightly different risk categories of cloud computing. The IDC survey results have shown obvious differences in comparison to this paper's survey. For example, both *Security* and *Availability* aspects have received the highest ranking in terms of end-users' concerns. On the other hand, this paper identified the *Support* and *Unpredictable Future Costs* aspects as the highest worrying factors among managers. Moreover, while most surveys addresses operational and administrative issues of cloud computing regarding the control and access of resources, this survey has restricted the range of audience to management-level users with only a medium or low technical background.

It can be concluded from the previous survey that most non-expert managers have similar concerns when it comes to unforeseeable long-term costs, contract management issues, performance difficulties, and integration with conventional systems. This was concluded as a result of selecting different companies with various ICT processes and applications as an audience to measure cloud computing decision-making challenges. In theory, each concern was addressed based on current ICT limitations observed by these managers within their organizations. On that account, a future research work can be suggested at this point to develop an automated filtering and comparison rule, which compares each of the previous cloud computing risk statements against the *Urgent Support Availability*. This can potentially support previous findings by highlighting the *Unpredictable Maintenance Delivery* as the most worrying aspect of different organizations' existing cloud solutions.

Table 2. Risk Analysis Survey Results

Level of Concern / Cloud Risk Category	Not worried at all	Slightly Worried	I don't mind	I am more worried	Extremely worried	Total Number of Participants	Average Rating
Government hosting regulations	46.30%	12.96%	18.52%	14.81	7.41%	54	2.24
Difficulties in going back to old hosting methods	28.85%	28.85%	26.92%	15.38%	0%	52	2.29
Unknown hosting locations	35.19%	16.67%	27.78%	16.67%	3.70%	54	2.37
Integration difficulties between the cloud and existing systems supplied by different vendors	15.38%	28.85%	26.92%	28.85%	0%	52	2.69
A complete service shutdown	3.92%	43.14%	17.65%	15.69%	19.61%	51	3.04
Contract management issues	7.55%	32.08%	15.09%	32.08%	13.21%	53	3.11
Performance issues	1.85%	33.33%	9.26%	50%	5.56%	54	3.24
Control over resources	1.89%	28.30%	9.43%	47.17%	13.21%	53	3.42
The 'on-demand' payment method of cloud computing might cost more than the traditional approach	5.66%	13.21%	3.77%	56.60%	20.75%	53	3.74
Unpredictable costs in the future	3.70%	14.81%	3.70%	55.56%	22.22%	54	3.78
Security (Data, access, permissions, sharing)	1.85%	16.67%	5.56%	46.30%	29.63%	54	3.85
Urgent support availability	0%	7.41%	7.41%	50%	35.19%	54	4.13
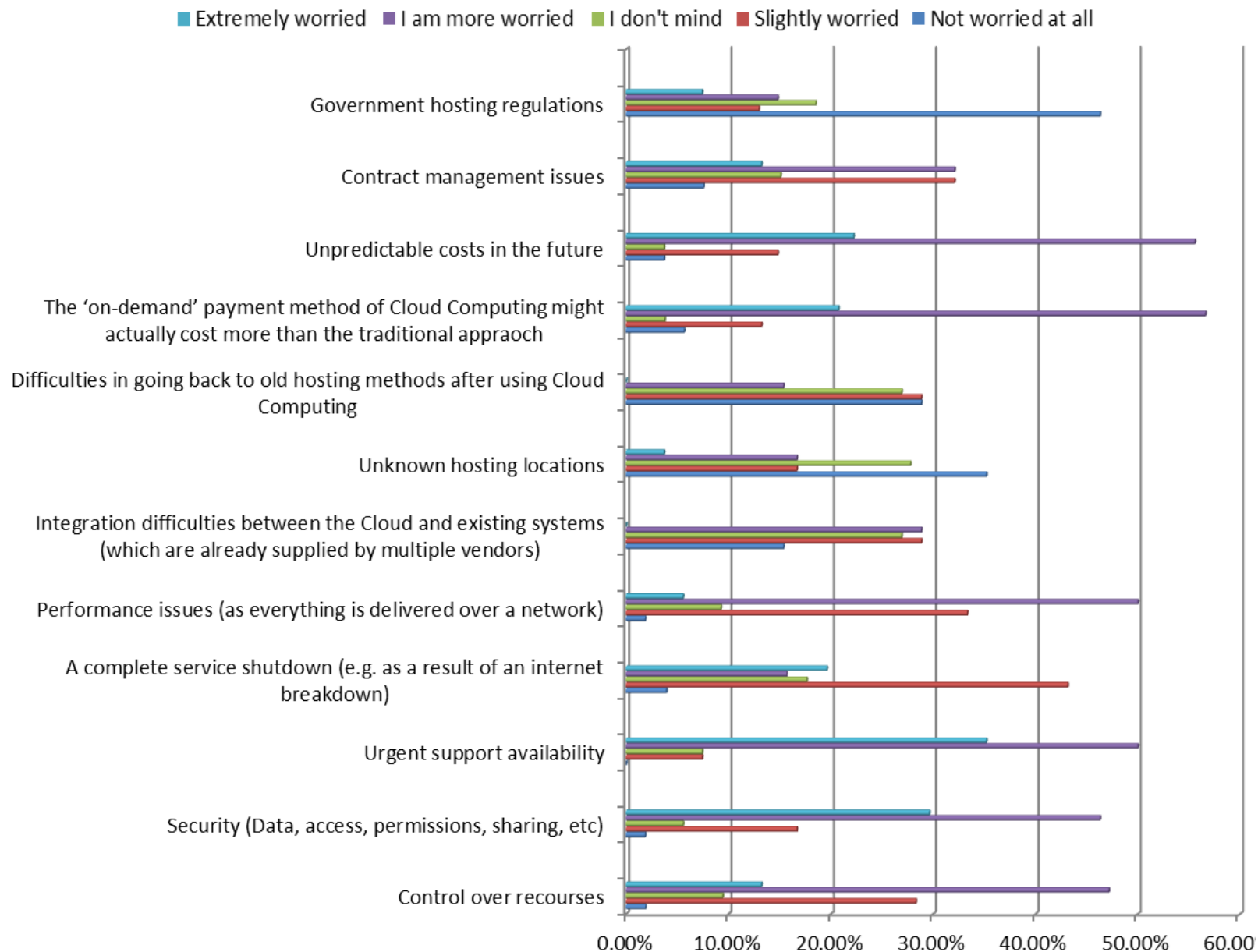
Figure 2. Survey Analysis: Microsoft Excel Representation of End-Users' Inputs

5.1 Summary of Results

The main objective of the earlier risk-analysis survey is to evaluate the level of concern of non-expert ICT clients towards cloud computing management and deployment. In conclusion, the previous collected data can be summarized in the following categories in relation to the highlighted cloud computing decision-making challenges:

Security and Privacy: According to IBM, the data security and privacy concerns rank top on almost all types of surveys [22]. On this account, cloud computing introduces another level of risk given that essential services are often outsourced to a third party, which makes the management process more challenging to demonstrate compliance, maintain data integrity, ensure privacy, support and service availability.

Actual Business Benefits: Most of today's non-expert managers are not convinced of the potential cost benefit of cloud computing. According to Netflex, in some heavy-scaling demand cases, cloud computing can be more costly than the conventional ICT approaches [23]. This can be determined in-house through a thorough identification of the organization's exact ICT requirements before adopting any models of cloud computing.

Furthermore, as noted from the earlier survey, one of the main concerns by managers is to realize the investment requisites to full potential, which adds value by making the cloud computing services part of their mainstream ICT portfolio. IBM argued that the return on investment (ROI) on utilizing cloud resources must be accomplished and verified by comparing certain management metrics of traditional ICT with cloud computing services [22]. As a result, this comparison will illustrate savings on future costs, which can lead to revenue, reduction in management effort and time, compliance, and better assessment of the organization's ICT workload and changes in demands.

Support and Service Quality: As can be viewed in the previous survey, Service quality is one of the biggest factors that non-expert managers highlighted as a challenge against outsourcing their ICT environments and business applications onto the cloud. On that ground, if the cloud providers' Service Level Agreements (SLAs) are not sufficient to guarantee the requirements for deploying applications on the cloud, then in most cases, these non-expert users need to ensure their contracts states that the provider will cover business loss for the amount of time the service was unavailable. These considerations are especially highlighted against the availability, performance and scalability sections in the contract with the cloud providers. Most of today's cloud computing contracts include a limited guarantee on service quality assurances. As a result, managers are reluctant to outsource their critical business infrastructure to the service providers' cloud datacenters.

Integration: Most organizations own legacy systems which require integration with specific types of cloud computing systems when outsourcing part of their applications onto the cloud. These applications usually have complex integration requirements to interact with other cloud or in-house systems. Non-expert mangers often sense that in terms of cost, effort, and time it is challenging from a technical and administrative perspectives to complete any needed integration with cloud-based systems. As a result, in many cases these managers would rather upgrade and invest more on existing in-house technologies. On this note, a proper evaluation of the cloud contract with the provider must be thoroughly examined given that most organizations have a major requirement to integrate cloud applications with the rest of the company's systems in a quick, easily-managed, and cost-efficient manner.

Performance: Most of today's cloud business applications require intensive bandwidth and a reliable internet connection whether delivered via software, platform or infrastructure cloud solutions. Cloud computing providers usually inform clients before signing any contracts that the performance of delivering complex services through the cloud is going to be unpredictable if the network bandwidth is not reliable and adequate. Therefore, as pointed out earlier, it has been observed that the majority of non-expert managers prefer to hold off any cloud outsourcing until an improved bandwidth with lower costs is made available by their organizations.

6. DECISION-MAKING FRAMEWORK

As discussed earlier, the delivery nature and characteristics of the current cloud computing services are constantly changing by top ICT providers without a unified standardization of their services. This was observed to result in gaps and unclear contracts between non-expert clients and the service providers regarding support, future costs, and other management concerns as discussed in the previous survey. This makes the management of cloud computing cost strategies, contracts, and resource control a difficult task across different sizes and types of organizations that employ ICT systems.

This paper has constructed a decision-making framework in order to assist non-expert clients in minimizing cloud computing potential threats to their in-house ICT management. The objective is to reduce deployment tradeoffs and support the decision-making process before outsourcing any core business applications onto a cloud computing platform. Another purpose of this framework is to assist non-expert managers to make effective decisions to minimize expenses and achieve a sustainable cloud computing environment with minimum management effort. The next section will

discuss each stage of the decision-making framework separately as follows (Figure 3).
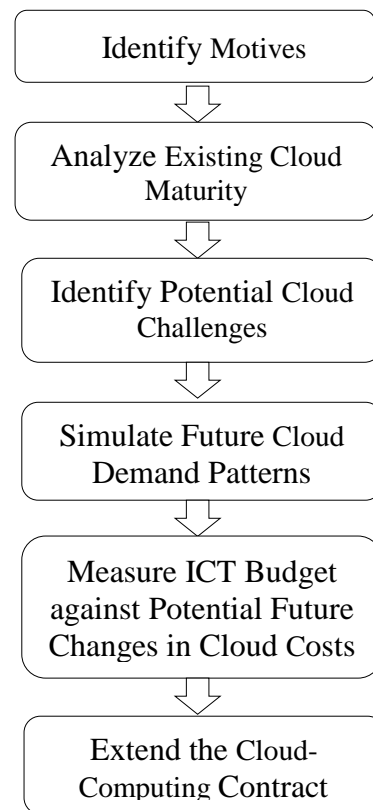
Figure 3. Cloud Computing Decision-Making Framework for Non-expert Clients

As illustrated in Figure 3, this paper has identified six key stages of the overall decision-making framework as follows:

- **Identify Motives:** It is recommended for any organization to thoroughly identify the main drivers of change and reasons behind moving certain applications onto the cloud before committing to any contracts with the service provider.

- **Analyze Existing Cloud Maturity:** As mentioned earlier, most organizations are already utilizing various types of cloud computing services. On that ground, a special management consideration is required before adopting new cloud services given that newer

features might include a duplicate of existing ones in some aspects, therefore, unnecessary costs can be added as a result of duplicating the same cloud services and purchasing unneeded resources.

- **Identify Potential Cloud Challenges:** Organizations adopt different work objectives and have various attributes such as size, ICT demand, and budget. As a result, particular cloud challenges can have more impact over the others as concluded in the previous survey. Therefore, it is recommended for organizations to identify the relevant areas of concern to their businesses and management processes. As a consequence, this can mitigate the level of concern by emphasizing on those aspects when signing a contract with the cloud provider. This can be achieved by requesting additional assurances and SLA guarantees from the cloud provider.

- **Simulate Future Cloud Demand Patterns:** As discusses earlier, one of the main cloud computing characteristics is the dynamic scalability which allow users to scale the capacity of their cloud resources up or down in a flexible, remote, and instant manner. This forms the fourth stage of this framework, which recommends non-expert managers to simulate their organizations' demand patterns across the off-peak and heavy demand periods prior to any actual cloud computing utilization.

- **Measure ICT Budget against Potential Future Changes in Cloud Cost:** Cloud computing providers such as Google and Amazon have changed their cloud pricing calculations and associated service features on several occasions in the last two years [24]. On that ground, this paper suggests that non-expert clients are recommended to measure results obtained from the previous stage with their allocated ICT budget for 3 to 5 years in advance. This stage is argued to help managers in predicting price changes in their

cloud services across time, which as a result would enable them to define and elaborate on these rules with the cloud provider at an earlier stage.

- **Extend the Cloud Computing Contract:** This forms the final stage of the decision-making framework after the non-expert cloud clients take into account all the previous stages. The main objective of this stage is to identify the potential threats and areas of ambiguity in the contract with the cloud provider, which can affect the organization's future ICT spending, management effort, and support.

7. CONCLUSION AND FUTURE WORK

Constructing long-term, sustainable, and cost-efficient strategies for any cloud deployment depends on the thorough identification of required services in-house and off-premises. This study points out that most of today's heavy-burdened organizations are outsourcing these services to costly independent suppliers which causes contract limitations, unnecessary management efforts, additional costs, and other decision-making complexities. These efforts are better employed by managers to enhance core competencies in their organizations, which potentially increases growth and attracts new business opportunities. On that ground, this paper has evaluated various management challenges of cloud computing, which were voted as most relevant by a number of non-expert managers through a risk-analysis survey. The outcome was analyzed against decision-making considerations for adopting different types of cloud computing services. The study highlighted non-expert clients as key users of the outcomes from this project given the diverse nature of the operational objectives in today's organizations.

Future work is structured to investigate and compare each cloud computing risk category

identified by the earlier survey with the actual cost, environmental, and management benefits obtained when cloud computing services are utilized. This analysis will highlight real-life examples of management issues and barriers experienced across those organizations, which as a result will allow decision-makers to measure the actual levels of management feasibility and efficiency from adopting cloud computing services against costs and other game changing factors in their ICT infrastructure.

REFERENCES

[1] M. Almorsy, J. Grundy, A. Ibrahim, "CollaborationBased Cloud Computing Security Management Framework", 2011.

[2] J. Rubner, "When the Sky's the limit". Collective Intelligence: Cloud Computing, 2011.

[3] R. Bernnat, W. Zink, N. Bieber, J. Strach, "Standardizing the Cloud: A Call to Action". Booz and Company Inc, 2012.

[4] P. Mell, T. Grance, "The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology", 2011.

[5] G. Brunette. R. Mogull, (2009). "Security Guidance for Critical Areas of Focus in Cloud Computing V2.1". CSA, Cloud Security Alliance.

[6] R. Buyya. J. Broberg. A. Goscinski, "Cloud Commuting: Principles and Paradigms". Wiley Press, New York, USA, 2011.

[7] BizCloud Corp. "Defining Cloud Deployment Models", http://bizcloudnetwork.com/defining-cloud-deployment-models, 2010

[8] C. Mines, "4 Reasons Why Cloud Computing is Also a Green Solution", GreenBiz, Web: http://www.greenbiz.com/blog/2011/07/27/4-reasons-why-cloud-computing-also-green-solution?page=0%2C1, 2011.

[9] A. Kofmehl, A. Levine, G. Falco, K. Schmidt, "Energy-Smart Buildings: Demonstrating how information technology can cut energy use and costs of real estate portfolios". Accenture Corporation, 2011.

[10] Carrenza Website, Enterprise Cloud Computing for Intelligent Buildings: Service Overview Manual for Tenants, 2015.

[11] R. Bernnat, W. Zink, N. Bieber, J. Strach, "Standardizing the Cloud: A Call to Action". Booz and Company Inc, 2012.

[12] A. Mohamed, "A History of Cloud Computing". Web Link: http://www.computerweekly.com, 2010.

[13] J. Brodkin, "Gartner: Seven Cloud-Computing Security Risks", InfoWorld Corporation, 2009.

[14] L. LaManna, "Top 9 Challenges in Cloud Computing". SAP - Business innovation. Web: http://blogs.sap.com/innovation/cloud-computing/top-9-challenges-in-cloud-computing-that-are-slowing-its-adoption-011918, 2012.

[15] S.O.Kuyoro, F. Ibikunle, O. Awodele, "Cloud Computing Security Issues and Challenges". International Journal of Computer Networks (IJCN), Volume (3), Issue (5), 2011.

[16] C. Gong, J. Liu, Q. Zhang, H. Chen, Z. Gong, "The Characteristics of Cloud Computing". Department of Computer Sciences, National University of Defense Technology, Changsha, China, 2010.

[17] F. Gens, "New IDC Cloud Services Survey: Top Benefits and Challenges". IDC Exchange, 2009.

[18] H. Lim, C. Babu, S. Shivnath, J. Chase, S. Parekh, "Automated Control in Cloud Computing: Challenges and Opportunities". ACDC, Barcelona, Spain, 2009.

[19] R. Johns, "Likert Items and Scales". University of Strathclyde. 2010.

[20] S. Yarwood, "Intelligent Buildings: Understanding and managing the security risks". IET: The Institution of Engineering and Technology, 2012.

[21] S. Reza, A. Adel, O. Justice, "Cloud Computing from SMES Perspective: a Survey-based Investigation", Journal of Information Technology Management Volume XXIV, Number 1, 2013.

[22] S. Sreekanth, "Top 5 Challenges to Cloud Computing", IBM, Web: https://www.ibm.com/developerworks/community/blogs/c2028fdc-41fe-4493-8257-33a59069fa04/entry/top_5_challenges_to_cloud_computing4?lang=en, 2011.

[23] G. Orzell, J. Becker, "Auto Scaling in the Amazon Cloud". Netflix Inc, the Netflix Tech Blog. 2012.

[24] U. Hölzle, "Google Cloud Platform Live: Blending IaaS and PaaS, Moore's Law for the cloud". Google Developers Blog: Senior Vice President, 2014.

The *International Journal of Digital Information and Wireless Communications* 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 *IJDIWC* publishes issues four (4) 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 IJDIWC. Papers should describe new contributions in the scope of IJDIWC 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 IJDIWC should not be previously published or be under review by any other publication. Original unpublished manuscripts are solicited in the following areas including but not limited to:

- Information Technology Strategies
- Information Technology Infrastructure
- Information Technology Human Resources
- System Development and Implementation
- Digital Communications
- Technology Developments
- Technology Futures
- National Policies and Standards
- Antenna Systems and Design
- Channel Modeling and Propagation
- Coding for Wireless Systems
- Multiuser and Multiple Access Schemes
- Optical Wireless Communications
- Resource Allocation over Wireless Networks
- Security; Authentication and Cryptography for Wireless Networks
- Signal Processing Techniques and Tools
- Software and Cognitive Radio
- Wireless Traffic and Routing Ad-hoc Networks
- Wireless System Architectures and Applications