http://www.icact.org/journal.asp

ICACT-TACT JOURNAL

Transactions on Advanced Communications Technology

Volume 3 Issue 3, May 2014, ISSN: 2288-0003

Editor-in-Chief
Prof. Thomas Byeongnam YOON, PhD.

GIRI Global IT Research Institute

Volume. 3 Issue. 3

Automatic Music Genre Classification Using Timbral Texture and Rhythmic Content Features

Babu Kaji Baniya*, Deepak Ghimire, Joonwhoan Lee

Division of Computer Science and Engineering
Chonbuk National University, Jeonju 761-756, South Korea
everwith_7,deep,chlee@jbnu.ac.kr

***Abstract*— Music genre classification is a vital component for the music information retrieval system. There are two important components to be considered for better genre classification, which are audio feature extraction and classifier. This paper incorporates two different kinds of features for genre classification, timbral texture and rhythmic content features. Timbral texture contains the Mel-frequency Cepstral Coefficient (MFCC) with other several spectral features. Before choosing a timbral feature we explore which feature contributes a less significant role on genre discrimination. This facilitates the reduction of feature dimension. For the timbral features up to the 4-th order central moments and the covariance components of mutual features are considered to improve the overall classification result. For the rhythmic content the features extracted from beat histogram are selected. In the paper Extreme Learning Machine (ELM) with bagging is used as the classifier for classifying the genres. Based on the proposed feature sets and classifier, experiments are performed with two well-known datasets: GTZAN and the ISMIR2004 databases with ten and six different music genres, respectively. The proposed method acquires better and competitive classification accuracy compared to the existing approaches for both data sets.**

***Keyword*— Classification, music genres, ELM (Extreme Learning Machine) with bagging, covariance matrix, timbral texture, rhythmic contents**

I. INTRODUCTION

TUTOMATIC music genre classification is an important for the information retrieval task since it can be applied for practical purposes such as efficient organization of data collections in the digital music industry. There have been several well-known distinct approaches put forward on this. Still, efficient and accurate automatic music information processing remains as the key issue, and it has been consistently attracting the attention of a growing number of researchers, musicians, and composers. A current challenging topic in automatic music information retrieval is the problem of organizing, describing, and categorizing music contents on the internet [1]. Although music genre classification is done manually, sometimes it is difficult to precisely define the genre of music content. The reason for such difficulties is due to fact that music is a state of art that evolves, where composers and musicians have been influenced by the music of other genres. Despite these difficulties, there are still some possibilities that remain for genre classification. The audio signals of music belonging to the same genre mean they share the certain common characteristics, because they are composed of similar types of instruments, having similar rhythmic patterns, and similar pitch distributions [2]. The extracted features must be comprehensive (representing music very well), compact, and effective.

The overview of our music genre classification is shown in Fig.1. It depicts the backbone of genre categorization. There are two associated problems that need to be addressed in genre classification, i.e., feature extraction and classification. The first stage is to extract the meaningful and relevant features from audio that could sufficiently discriminate the music genre. The next stage is to classify the genre based on the extracted features. In our method the extreme learning machine (ELM) combined with bagging is used as a classifier. Several bags of the dataset are constructed and each bag is trained using individual ELMs. The final decision is made based on the majority voting score. ELM is an unstable classifier, therefore ELM combined with bagging increases the stability, as well as generalization performance of the classifier.

For constructing a robust music genre classifier, extracting features that allows direct access to the relevant genre-specific information is crucial. Most musical genre classification systems utilize the low-level spectral features of the short time audio signal in the range of 10ms to 100ms, such as pitch extraction, mel-frequency cepstral coefficients (MFCCs), and other timbral texture features [3]. Then the short-time low-level spectral features are integrated on long duration. The most widely used integrating method is mean and standard deviation of the short time feature [4, 5].

In this paper, we attempt to implement timbral texture features which represent short-time spectral information, and rhythmic content features like beat histogram which represent

Manuscript received May 9, 2014. This work was partially supported by a National Research Foundation of Korea grant funded by the Korean government (2011-0022152) and PK21PLUS project.

Babu Kaji Baniya is with the Computer Engineering Department of Chonbuk National University, Jeonju, South Korea (e-mail: everwith_7@jbnu.ac.kr).

Deepak Ghimire is with the Computer Science and Engineering Department of Chonbuk National University, Jeonju, South Korea (e-mail: deep@jbnu.ac.kr).

Corresponding author: Joonwhoan Lee is with the Computer Engineering Department of Chonbuk National University, Jeonju, South Korea (phone: +82-10-9855-2406; Fax: +82-63-270-2406; e-mail: chlee@jbnu.ac.kr).

the long-term properties. Timbral texture features include spectral centroid, flux, rolloff, flatness, energy, zero crossing, and MFCCs, respectively. We divide the timbral texture features into two groups for convenience; the first group (FG1) does not include MFCCs and the second group (SG2) includes only MFCCs. After the frame-wise extraction of each timbral texture feature among FG1 from all genres of music, the next stage is to calculate the standard deviation for all genres of music. The aim of calculating the standard deviation for each feature in whole genres is to find out which feature is insignificant for genre discrimination. The feature which has a small value of standard deviation contributes the insignificant impact on genre discrimination. Based on the standard deviation value, we considered a limited number of timbral features.

A Similar procedure has been preceded for the SG2 of MFCC features as well. Out of thirteen, twelve coefficients give meaningful standard deviation values. This shows that twelve MFCC coefficients are meaningful for genre classification. For our experiment, we consider both the first seven and twelve coefficients separately for genre classification.

Timbral texture features are based on short time low-level spectral components that are integrated on long duration. The integration method is mean and standard deviation. Beside this, high order moments such as skewness and kurtosis are also implemented for integrating the frame-wise features. The aim of considering the high order moment is that even if there are the same values of mean and standard deviation, the position of location (shape of skewness and kurtosis position) could be different because each feature cannot be modelled by the Gaussian distribution.

Ultimately, the high order moment increases the classification accuracy when it is combined with other low level spectral features. It generally provides the supplementary statistical information for the audio signal. Skewness is a measure of the asymmetry of the data distribution regarding the sample mean, which represents the relative disposition of the tonal and non-tonal components of the audio signal. Kurtosis is the measure for the degree of peakedness or flatness of a distribution [6]. Therefore we have considered 4n components for the n texture features.

In addition we propose to use the covariance components of mutual timbral texture features. Each of them gives the statistical property of mutual random variables associated features. For each song the covariance values of selected features from FG1 and SG2 are calculated, respectively. Therefore, additional $n(n\text{-}1)/2$ components are included for n timbral texture features.

Note that we can have $4n+ n(n\text{-}1)/2$ for n features, which increases rapidly as the number of features increases. This is the reason why we remove the relatively less important features by checking the corresponding variances.

Rhythm is a property of an audio signal that represents a changing pattern of timbral and energy over time. Rhythmic features characterize the movement of music signals over time and contain such information as the regularity of the rhythm, beat, tempo, and time signature. The feature set for representing the rhythmic structure is based on detecting the most salient periodicities of the signal and it is usually extracted from the beat histogram. Rhythmic content features contain relative amplitude of the first and second histogram peaks, period of first and second peaks, ratio of the amplitude of the second peak divided by the amplitude of first peak, and overall sum of the histogram.

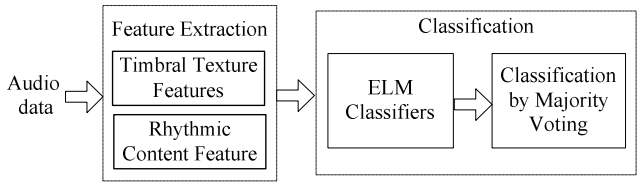


Fig.1. Overview of music genre classification

There are different types of classifiers which have been proposed for genre classification. We prefer the distinct classifier than the previously applied one. Extreme Learning Machine (ELM) is a recently proposed classifier which has high generalizing capability and takes minimum time for training. The reason for selecting ELM is that it does not require a tuning parameter, has the smallest training error, and is free from the local maxima problem. However, ELM is unstable because the weights connected with hidden units are randomly determined. Therefore, we combine ELM with bagging in order to increase the stability. Bagging is almost always more accurate than a single classifier. Other classifiers like K-Nearest Neighbour (K-NN), Neural Networks (NN) have some drawbacks. In case of neural network, when learning rate is too small, the algorithm converges very slowly. It also requires a tuning parameter and probably faces the local maxima problem. K-Nearest Neighbour is a simple nonparametric classifier. It is proven that the error of K-NN's is twice large than Bayesian error rate.

This paper is organized as follows. A review of related work is provided in section II. Feature extraction is the critical portion of genre classification; and is describes in section III. Section IV deals with the classifier, similarly section V explains the experimental setup and data preparation, and section VI explains the result and analysis. Finally, section VII describes the conclusion of the proposed method and future work of the genre.

II. RELATED WORK

Many different features have been introduced for music genre classification. The primary aim of feature extraction is to acquire a meaningful representative part of music in the reduced form. The acoustic features include tonality, pitch, beat, and symbolic features extracted from the scores, and text-based features can be obtained from the song lyrics. In this paper, we only focus on timbral texture and rhythmic content which are sub-groups of content-based features.

The content-based acoustic features are divided into timbral texture features, rhythmic content features, and pitch content features [7]. Timbral features are often calculated for every short-time frame of sound based on the Short Time Fourier Transform (STFT) [8]. Timbral texture features contain MFCCs, spectral centroid, spectral flatness, spectral flux, spectral rolloff, zero crossing, energy, and Linear Prediction Coefficients (LPCs) [7, 8]. These features are widely used in different applications based on the requirement of applications. MFCCs have been

extensively used in speech recognition [8]. Later, MFCC features are used for discriminating the music and speech as well. Rhythmic content features possess information about continuity of rhythm, beat and tempo. Tempo and beat tracking are excessively used in music search and retrieval systems. The tempo value is a number which represents the speed of music or music measured by beats per minute (bmp) [9, 10]. The pitch content feature deals with frequency information of music.

Bergstra et al. [11] extracts the several timbral texture features like MFCCs, spectral centroid, spectral flux, spectral rolloff, zero crossing, energy, and Linear Prediction Coefficients. These features are almost similar with the features used in [3, 5]. AdaBoost is used as a classifier.

C.-H. Lee et al. [12] considers the Octave-Based Spectral Contrast (OSC) and MFCC for feature extraction. There is a range of nine different frequencies in octave-based spectral contrast. Music genres are classified by using Linear Discriminant Analysis (LDA). Recently, Seo et al. [13] also implemented the Octave-Based Spectral Contrast (OSC) for feature extraction. Beside this, he consider the high order moment for improving the performance of classification accuracy. The genre classification is performed by using Support Vector Machines (SVM).

Li et al. [1] mention several audio feature extraction methodologies. Later, he proposed a new approach for feature extraction, i.e. Daubechies Wavelet Coefficients Histograms (DWCHs). The effectiveness of this new feature is compared using various machine learning algorithms, SVMs, Gaussian Mixture Models (GMMs), K-NNs, and LDAs.

The spectral similarity of the timbral texture feature is described by Pampalk et al. [14]. The audio signal is chopped into thousands of very short frames and their order in time is ignored. Each frame is described by MFCCs. The large set of frames is summarized by a model obtained by clustering the frames. The distance between two pieces is computed by comparing their cluster models. Later, GMM is considered for genre classification.

Tzanetakis and Cook [7] proposed a comprehensive set of features for direct modelling of music signals and explore the different applications of those features for musical genre classification using K-Nearest Neighbor and GMM. Other researchers like Lambrou et al. [15] use statistical features in the temporal domain as well as three different wavelet transform domains to classify music into rock, piano, and jazz.Soltau et al. [16] propose an approach of representing temporal structures of input signals. He shows that this new set of abstract features can be learned via artificial neural networks and can be used for music genre identification. Deshpande et al. [17] use Gaussian Mixtures, SVM, and K-Nearest Neighbor to classify the music into rock, piano, and jazz based on timbral texture features.

III. FEATURE EXTRACTION

Feature extraction encompasses the analysis and extraction of meaningful information from audio in order to obtain a compact and concise description that could be machine readable. Features are usually selected in the context of a specific task and domain. The features that are used in our research are divided into two categories, the timbral texture feature and rhythmic content feature.

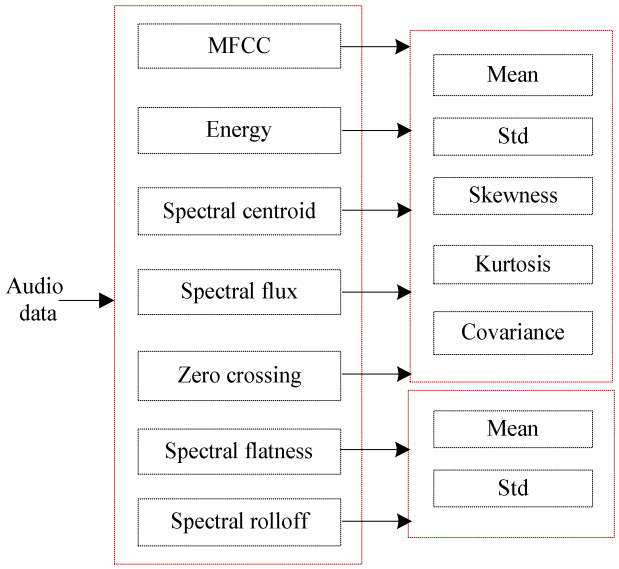


Fig. 2. Overview of Timbral texture features extraction of audio.

A. *Timbral Texture features*

These features are used to differentiate mixture of sounds that possibly have similar pitch and rhythm [8]. The features used to represent timbral texture are based on standard features proposed for music-speech discrimination [18]. To extract the timbral features, audio signals are first divided into frames by applying a windowing function at fixed intervals. The window function of this research is hamming window which helps to remove the edge effects. Timbral texture features in Fig.2 have been computed and later we calculated different statistical values like mean, standard deviation, skewness, kurtosis, and covariance matrix from feature values. The mean (μ) and standard deviation (σ) for frame-wise feature values (X_n) in a N -frame song are given by

$$Mean(\mu) = \frac{1}{N}\sum_{n=1}^{N} X_n \tag{1}$$

$$Std(\sigma) = \frac{1}{N}\sum_{n=1}^{N} (X_n - \mu)^2 \tag{2}$$

The skewness is a measure of asymmetry of the distribution, which can represent the relative disposition of the tonal and non-tonal components of each band. If the tonal components occur frequently in a band, the distribution of its spectrum will be left-skewed otherwise it will be right-skewed. Mathematically, the skewness in a song can be defined as

$$Skewness = \frac{\sum_{n=1}^{N} (X_n - \mu)^3}{(N-1)\sigma^3} \tag{3}$$

Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. That is, data sets with high kurtosis tend to have a distinct peak near the mean. It is difficult to specify the exact contribution of kurtosis in music genre classification [13]. However, the kurtosis measure can sketch the effective dynamic range of the audio spectrum. Mathematically it can be defined as

$$Kurtosis = \frac{\sum_{n=1}^{N}(X_n-\mu)^4}{(N-1)\sigma^4} - 3 \quad (4)$$

Covariance is measured between two random variables or features. The aim of considering the covariance is usually to see if there is any relationship between the random variables. It is useful to measure the polarity and the degree of the correlation between two features. The covariance of two features X_n and Y_n in a song is given as

$$Cov(X_n, Y_n) = \frac{1}{N}\sum_{n=1}^{N}(X_n-\mu_X)(Y_n-\mu_Y), \quad (5)$$

where μ_X and μ_Y are corresponding means of X_n and Y_n, respectively. For n timbral texture features we acquired $n(n-1)/2$ mutual covariance values.

We consider two groups of timbral texture features FG1 and SG2 described as

1) FG1 features

Spectral flux: It is defined as the variation value of the spectrum between the adjacent two frames in a short-time analyze window. It measures how quickly the power spectrum changes and is used to determine the timbral of an audio signal.

$$F_t = \sum_{n=1}^{N}(N_t[n]-N_{t-1}[n])^2 \quad (6)$$

where $N_t[n]$ and $N_{t-1}[n\text{-}1]$ are normalized magnitudes of the Fourier transform at the present frame *t*, and previous frame *t*-1 respectively.

Spectral centroid: The spectral centroid is described as the gravity centre of the spectral energy. It determines the point in the spectrum where most of the energy is concentrated and is correlated with the dominant frequency of the signal. It is closely related to the brightness of a single tone.

$$C_t = \frac{\sum_{n=1}^{N} M_t[n]*n}{\sum_{n=1}^{N} M_t[n]} \quad (7)$$

where $M_t[n]$ is the magnitude of the Fourier transform at frame *t* and frequency bin *n*.

Short Time Energy: The short time energy measurement of an audio signal can be used to determine voiced and unvoiced speech. It can also be used to detect the transition from unvoiced to voice and vice versa [19]. The energy of voiced speech is much greater than the energy of unvoiced speech. Short-time energy can be defined as

$$E_n = \sum_{m=1}^{N}[x(m)w(n-m)]^2 \quad (8)$$

where, $x(m)$ is discrete time audio signal, n is time index of short-time energy, and $w(m)$ is window of length N.

Zero Crossing: It is a process of measuring the number of times in a given time interval that the amplitude of speech signals crosses through a value of zero. It is random in nature. Moreover, the zero crossing rate for unvoiced speech is greater than that of voice speech. Moreover, it is often used as a crucial parameter for voiced/unvoiced classification and end point detection.

$$Z_t = \frac{1}{2}\sum_{n=1}^{N}|\mathrm{sgn}(x[n]-\mathrm{sgn}(x[n-1])| \quad (9)$$

where sgn is a short notation of sign function. The sgn function is 1 for positive arguments and 0 for negative arguments and x[n] is the time domain for signal for frame t.

Spectral Rolloff: It is a measure of the bandwidth of the audio signal. It is the fraction of bins in the power spectrum in which 85% of the power is at lower frequencies.

$$\sum_{n=1}^{R_t} M_t[n] = 0.85\sum_{n-1}^{N} M_t[n] \quad (10)$$

where $M_t[n]$ is the magnitude of the Fourier transform at frame *t* and frequency bin *n*.

Spectral flatness: It is used to characterize an audio spectrum. Spectral flatness is typically measured in decibels, and provides a way to quantify how tone like a sound is, as opposed to being noise-like.

$$F = \frac{\exp\left(\frac{1}{N}\sum_{m=0}^{N-1}\ln x(m)\right)}{\frac{1}{N}\sum_{m=0}^{N-1}x(m)} \quad (11)$$

where $x(m)$ represents the magnitude of bin number *m*.

From the above mentioned features in FG1, the normalized standard deviation of all the data has been calculated. Since the standard deviation generally depends on the mean value in general, the standard deviation is divided by corresponding mean to find out the less important features. Note that a smaller value of the standard deviation means a smaller change in the values of the frame-wise timbral texture feature. This means any derived central moments from the feature and the covariance with the feature is not significant for the discrimination of music genres. Therefore we removed such features to reduce the feature dimension.

Spectral centroid, flux, short time energy, and zero crossing possess large normalized standard deviations compared to the rolloff and flatness as shown in Table IV. We only consider four features (Spectral centroid, flux, short time energy, and zero crossing) and their mean, std, skewness, kurtosis and $n(n\text{-}1)/2$ covariance components, respectively. The feature dimension is given in Table I.

2) SG2 Features: Mel-Frequency Cepstral Coefficients

TABLE I
FEATURE DIMENSION OF FOUR DIFFERENT TIMBRAL TEXTURE FEATURES

Mean	Std. dev.	Skew.	Kurt.	Cov.	Total features
4	4	4	4	6	22

Earlier MFCCs widely used in automatic speech recognition later on evolved into one of the prominent techniques in every domain of audio retrieval. They represent most distinctive information of signal. MFCCs have been successfully implemented to timbral measurements by H. Terasawa [20].

We took the MFCC feature based on the paper that mentioned the mel frequency cepstral coefficients for music modelling [21]. Fig. 3 shows the process of creating MFCC features. The first step is to divide the audio signal into frames, by applying a window function at fixed intervals. The aim is to

model small (having 10ms) sections of the signal that are statistically stationary. The window function is hamming window. We generate the cepstral feature vector for each frame. The next step is to take the Discrete Fourier Transform (DFT). The phase information has been discarded because perceptual studies have shown that the amplitude of the spectrum is much more important than the phase. The logarithm of the amplitude spectrum has been taken because the perceived loudness of a signal has been estimated to be approximately logarithmic. The next stage is to smooth the spectrum and emphasize perceptually meaningful frequencies. This is achieved by collecting the spectral components into frequency bins. As we know, lower frequencies are perceptually more important than the higher frequencies. Therefore, the bin spacing follows the so-called 'Mel' frequency scale [22]. The components of the Mel-spectral vectors calculated for each frame are highly correlated. In order to reduce the number of parameters in the MFCC, we need to apply a transform to the Mel-spectral vectors which decorrelates their components. The cepstral features of each frame are obtained by using DCT.

by applying a window function at fixed intervals. The aim is to model small (having 10ms) sections of the signal that are statistically stationary. The window function is hamming window. We generate the cepstral feature vector for each frame. The next step is to take the Discrete Fourier Transform (DFT). The phase information has been discarded because perceptual studies have shown that the amplitude of the spectrum is much more important than the phase. The logarithm of the amplitude spectrum has been taken because the perceived loudness of a signal has been estimated to be approximately logarithmic. The next stage is to smooth the spectrum and emphasize perceptually meaningful frequencies. This is achieved by collecting the spectral components into frequency bins. As we know, lower frequencies are perceptually more important than the higher frequencies. Therefore, the bin spacing follows the so-called 'Mel' frequency scale [22]. The components of the Mel-spectral vectors calculated for each frame are highly correlated. In order to reduce the number of parameters in the MFCC, we need to apply a transform to the Mel-spectral vectors which decorrelates their components. The cepstral features of each frame are obtained by using DCT.

There are thirteen coefficients in the mel-frequency cepstral

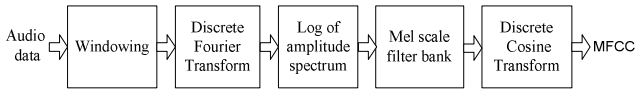


Fig. 3. Mel frequency cepstral coefficients feature extraction of audio.

coefficient. After analysis of the normalized variance we selected 12 out of 13 coefficients. The last coefficient has a very small value of the variance as shown in Table V, so that it could be removed. We try to implement the early seven and twelve coefficients separately. The first choice is just for reducing the dimension. The different feature dimension of MFCC while considering seven and twelve coefficients are given in table II and III.

TABLE II
FEATURE DIMENSION OF MFCC CONSIDERING FIRST SEVEN COEFFICIENTS

Mean	Std. dev.	Skew.	Kurt.	Cov.	Total features
7	7	7	7	21	49

TABLE III
FEATURE DIMENSION OF MFCC CONSIDERING TWELVE COEFFICIENTS

Mean	Std. dev.	Skew.	Kurt.	Cov.	Total features
12	12	12	12	66	114

B. Rhythmic Content Features

Rhythmic content features characterize the movement of music signals over time and contain such information as the regularity of the rhythm, beat, and tempo. For the rhythmic feature, beat histogram has been taken. It is a compact global representation of the rhythmic content of audio music. The beat histogram [5] can be obtained by the wavelet decomposition of a signal and can be interpreted as successive high-pass and low-pass filtering of the time domain signal. The decomposition is defined by

$$y_{high}[k] = \sum_n x[n]g[2k-n] \quad (12)$$

$$y_{low}[k] = \sum_n x[n]h[2k-n] \quad (13)$$

where $y_{high}[k]$ and $y_{low}[k]$ are the output of high-pass and low-pass filters respectively, and $g[n]$ and $h[n]$ are the filter coefficients for the high-pass and low-pass filters associated to the wavelet function for fourth order Daubechies wavelets (DW) [22]. Wavelet Transform deals with the similarity of the decomposed signal to the octave filter band. Once the signal is decomposed, the additional signal processing operation is required. The building blocks as shown in Fig. 4 are used for the beat analysis feature extraction.

1) Full Wave Rectification:

$$y[n] = abs(x[n]) \quad (14)$$

where x[n] is the output of the wavelet decomposition at that specific scales.

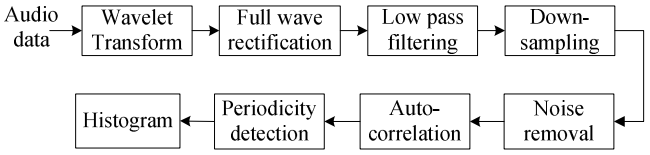


Fig.4. The block diagram of beat histogram for feature extraction

2) Low-Pass Filtering:

$$a[n] = (1-\alpha)y[n] + \alpha a[n-1] \quad (15)$$

For one-pole filter with an alpha value of 0.99 which is used to smooth the envelope.

3) Downsampling:

$$b[n] = a[kn] \quad (16)$$

Downsampling the signal reduces computation for the autocorrelation calculation without affecting the performance of the algorithm. The value of k is 16.

4) Normalization (mean removal)

Mean removal is applied in order to make the signal centered to zero for the autocorrelation stage.

$$c[n] = b[n] - E[b[n]] \quad (17)$$

5) Autocorrelation

$$d[k] = \frac{1}{N}\sum_{n} c[n]c[n-k] \quad (18)$$

where c[n] is periodic signal with period N.

6) Periodicity detection and beat histogram calculation:

There are six different features extracted from the beat histogram. They are relative amplitude of the first and second histogram peak, period of the first and second histogram peak measure in beat per minute (bpm), ratio of the amplitude of the second peak divided by the amplitude of the first peak, and overall sum of the histogram.

IV. CLASSIFIER

Traditionally, all the parameters of the feed-forward networks need to be tuned and thus there exists the dependency between different layers of parameters (weights and biases). In particular the gradient descent-based methods have been used in various learning algorithms of feed-forward neural networks [23]. However, the weakness of this kind of learning method is that it is generally very slow due to diverse learning steps and may easily converge to local minima. They also require many iterative learning steps in order to obtain better learning performance.

ELM [24] resolves the problem associated with the gradient-based algorithm by analytically calculating the optimal weights of single-hidden layer feed-forward neural networks (SLFNs). Where the weights between input layers and the hidden layer biases are arbitrarily selected and then the optimal values for the weights between the hidden layer and output layer are determined by calculating the linear matrix equations.

For N distinct samples and $\tilde{N}$ hidden nodes, the activation function g(x) of the SLFN neural network is defined as

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i.x_j + b_i) = o_j, \quad j = 1, ..., N \quad (19)$$

where $w_i = [w_{i1}, w_{i2}, .., w_{in}]^T$ is the weight vector connecting the ith hidden node and the input nodes, $\beta_i = [\beta_{i1}, \beta_{i2}, .., \beta_{im}]^T$ is the weight vector connecting the i-th hidden nodes and output nodes, and b_i is the threshold of the i-th hidden node. $w_i.x_j$ denotes the inner product of w_i and x_i.

The standard SLFNs with $\tilde{N}$ hidden nodes with the activation function $g(x)$ can approximate these N samples with zero error means that $\sum_{j=1}^{\tilde{N}} || o_j - t_j || = 0$ i.e., there exist β_i, w_i, and b_i such that

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i.x_j + b_i) = t_j, \quad j = 1, ..., N \quad (20)$$

where t_j is the target vector of the j-th input data. Equation (19) can be written as a matrix equation to form a new equation by using the output matrix of the hidden layer H.

$$H\beta = T \quad (21)$$

where

$$H = \begin{bmatrix} g(w_1.x_1 + b_1) & \cdots & g(w_{\tilde{N}}.x_1 + b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ g(w_1.x_N + b_1) & \cdots & g(w_{\tilde{N}}.x_N + b_{\tilde{N}}) \end{bmatrix}_{N \times \tilde{N}} \quad (22)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \quad and \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \quad (23)$$

From the above equation (21), the target vector T and the output matrix of the hidden layer H can comprise a linear system. Thus, the learning procedure of the network helps to find the optimal weight matrix β between the output layer and the hidden layer β can be determined by using the Moore-Penrose generalized inverse of H.

$$\hat{\beta} = H^{\dagger}T \quad (24)$$

From the above equation (24) we can draw the following important properties. The first one is that we can take minimum training error, because the solution $\hat{\beta} = H^{\dagger}T$ is one of the least-square solutions of the general linear system $H\beta = T$. In addition, the optimal $\tilde{\beta}$ is also minimum norm among these solutions. Thus, ELM has the best generalization performance compared to the typical back propagation network. In summary the ELM algorithm can be summarized as follows.

Algorithm ELM: For the given training set $\aleph = \{(x_i, t_i) \mid x_i \in R^n, t_i \in R^m, i = 1, ..., N\}$, activation function $g(x)$, and hidden neuron number $\tilde{N}$,

1) Assign random input weight w_i and bias b_i, i=1,…, $\tilde{N}$.
2) Calculate the hidden layer output matrix H.
3) Calculate the output weight β:

$$\hat{\beta} = H^{\dagger}T$$

Where $H^{\dagger}$ is the Moore-Penrose generalized inverse of hidden the layer output matrix H.

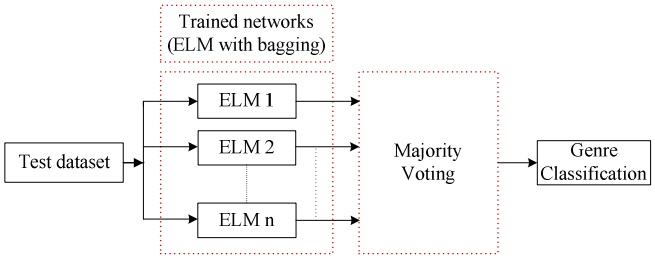


Fig. 5: Block diagram of music genre classification

A. Bagging Algorithm

Bagging [25] is a well-known ensemble learning algorithm that has been shown to be very effective in improving generalization performance compared to the individual base models. Breiman indicated that bagging is a smoothing operation which turns out

to be advantageous when aiming to improve the predicative performance of regression or classification. It is a "bootstrap" ensemble method that creates bags for its ensemble by training each classifier on a random redistribution of the training set. Each classifier's training set is generated by randomly drawing, with replacement; many of the original samples may be repeated in the resulting training set while others may be left out. Each bag classifier in the ensemble is generated with a different random sample of the training set. The algorithm then applies a base classifier to classify each bag. Bagging is almost always more accurate than a single classifier. Finally the decision is taken by majority voting of all the base classifier results. Fig. 5 shows the overview of genre classification. Our base classifier is ELM.

The Bagging Algorithm

Inputs: Training set *S*, based classifier *L*, integer *T* (number of bootstrap sample)

for $i = 1$ to T {

S_i=bootstrap sample from *S* (i.i.d. sample with replacement)

$E_i = L(S_i)$ }

$$E^*(x) = \arg\max_{y \in Y} \sum_{i:E_i(x)=y} 1 \quad \text{(the most often predicated label } y\text{)}$$

Output: Compound classifier E^*

V. Experimental setup and Data Preparation

Different datasets widely used for music genre classification are employed for performance comparison. The first dataset (GTZAN) consists of 1000 songs over ten different genres: Classical, Blues, Hiphop, Pop, Rock, Gazz, Reggae, Metal, Disco, and Country. Each class consists of 100 songs having duration of 30s. The dataset was collected by Geroge Tzanetakis [26]. Each song in the database was stored as a 22050Hz, 16bits, and mono audio file. The second dataset is ISMIR2004 [27] which were used in the Music Genre Classification Contest 2004. This dataset has an unequal number of distributions of music tracks in each class. It consists of six different classes: Classical, Pop and Rock, Metal and Punk, Electronic, World, and Jazz and Blues respectively. This dataset consists of 1458 music tracks in which 729 music tracks are used for training and the other 729 tracks for testing. The audio files are stored in MP3 format having a sampling frequency of 44.1 kHz, 128-kbps, 16 bit, and stereo files. For our research, each stereo MP3 file was first converted into a 44.1 kHz, 16 bit, mono audio file before feature extraction. In summary, the music tracks used for training/testing include 320/320 tracks of Classical, 115/114 tracks of Electronic, 26/26 tracks of Jazz/Blues, 45/45 tracks of Metal/Punk, 101/102 tracks of Rock/Pop, and 122/122 tracks of World music genre.

A five-fold cross validation scheme is used to evaluate the performance of the proposed system in the GTZEN dataset whereas in order to compare our proposed method with the results from the ISMIR2004 Music Genre Classification Content, our experiment on the ISMIR2004 genre dataset used the same training and testing set as in the contest. In the contest, the classification performance is evaluated based on 50:50 training and testing set instead of five-fold cross validation.

VI. Result and Analysis

At first, in order to reduce the dimensionality of the extracted feature set, the normalized standard deviation of each timbral texture feature is calculated in both the GTZAN and ISMIR2004 dataset. As the number of timbral texture feature increases the dimensionality of the extracted feature set increases rapidly, therefore we removed the relatively less important features by checking the corresponding normalized standard deviations. Table IV and V contain the average normalized standard deviation of the GTZAN and ISMIR2004 datasets. The data seen in table IV indicates that flatness and rolloff are less significant for genre classification than the other four in FG1. A similar approach is also applied for SG2 of MFCC coefficients. There is only one MFCC coefficient which has a relatively small value of normalized standard deviation. Aside from that coefficient, the remaining twelve coefficients are useful for genre classification.

TABLE IV
NORMALIZED STANDARD DEVIATION OF TIMBRAL TEXTURE FEATURES (EXCLUDING MFCC)

	Energy	Centroid	Flux	Zerocrossing	Flatness	Rolloff
Nor.Std	1.085	0.690	0.886	0.671	0.239	0.274

TABLE V
NORMALIZED STANDARD DEVIATION OF MFCC

MFCC Coefficients	1	2	3	4	5	6	7
Normalized Standard Deviation	2.03	2.31	2.36	2.43	2.49	2.38	2.43
MFCC Coefficients (contd.)	8	9	10	11	12	13	
Normalized Standard Deviation	2.24	2.28	2.48	2.29	1.60	0.02	

The genre classification result of the GTZAN and ISMIR2004 datasets is shown in table VI and VII respectively. Several experiments have been conducted among the different feature sets. The first experiment was conducted within timbral texture features in FG1 like spectral centroid, flux, energy, and zero crossing for genre classification (excluding MFCC). The second and third experiments were only conducted for SG2 with seven and twelve mel-frequency cepstral coefficients (feature dimension shown in table II and III respectively). The fourth and fifth experiments only considered mean, standard deviation, skewness, and kurtosis of timbral texture feature including seven and twelve MFCC coefficients separately with the rhythmic content feature. The final experiment was conducted taking the covariance matrix and all other features. The experiment was performed into two different steps to find the classification accuracy in regard to minimum (seven MFCC coefficients) and maximum (twelve MFCC coefficients) feature dimensions considering mean, standard deviation, skewness, and kurtosis.

The combination of different feature sets gives different classification accuracy. The feature extracted from FG1 of energy, centroid, flux, and zero crossing gives 68.33 % of accuracy. Similarly, SG2 (MFCC feature sets) with seven and

twelve coefficients give 64.62% and 66.26% accuracy respectively. This experimental result shows that seven or eleven coefficients of MFCC do not make a big difference in genre classification. We also tried to find out the overall impact of classification accuracy with covariance components. The classification accuracy of GTZAN dataset without covariance components comes around 78.26% (7-MFCCs) and 80.21% (12-MFCCs) respectively. Among them, maximum accuracy is obtained while combining the covariance components with other feature sets. The classification accuracy of the GTZAN dataset increases from 80.21% to 85.58% while including covariance components. Similarly, ISMIR2004 classification accuracy also increases from 81.53% to 86.46%. Hence, covariance components had significant impact in improving the genre classification.

TABLE VI
CLASSIFICATION ACCURACY (CA) OF GTZAN DATASET IN DIFFERENT FEATURE SETS

Feature set	ELM(CA)
Energy+Centroid+Flux+Zerocrossing	68.33%
MFCC (7 coff)	64.62%
MFCC (12 coff)	66.26%
[ECFZ+MFCC, 7 coff.] [without covariance]+beat histogram	78.26%
[ECFZ+MFCC, 12 coff.] [without covariance]+beat histogram	80.21%
[ECFZ+MFCC , 7 coff.] [with covariance]+beat histogram	84.52%
[ECFZ+MFCC, 12 coff.][with covariance]+beat histogram	85.15%

TABLE VII
CLASSIFICATION ACCURACY (CA) OF ISMIR2004 DATASET IN DIFFERENT FEATURE SETS

Feature set	ELM(CA)
Energy+Centroid+Flux+Zerocrossing (ECFZ)	73.62%
MFCC (7 coff)	66.58%
MFCC (12 coff)	68.78%
(ECFZ+MFCC (7 coff.)) (without covariance)+beat histogram	79.65%
(ECFZ+MFCC (12 coff.)) (without covariance)+beat histogram	81.53%
(ECFZ+MFCC (7 coff.)) (with covariance)+beat histogram	85.15%
(ECFZ+MFCC (12 coff.)) (with covariance)+beat histogram	86.46%

In our approach, the extreme learning machine combined with bagging algorithm is used for the classification of the music genre. Before we applied the ELM with bagging as a classifier, we attempted to find out how many bags were needed to obtain maximum classification accuracy. There were twenty three bags combined in case of the GTZAN dataset to get maximum classification as shown in Fig.6. In the ISMIR2004 dataset maximum classification accuracy was achieved when twenty five bags were used as shown in Fig.7.

Table VIII compares our proposed method with other approaches in terms of average classification accuracy in the GTZAN dataset. It is clear that our proposed method achieves the classification accuracy of 85.15% which is better than other approaches. Similarly, Table IX shows the comparison results with previous different approaches as well as the ISMIR2004 Music Genre Classification Contest. The classification accuracy is 86.46%. It is also comparatively competitive with Chang-Hsing Lee's method [26] and better than all other approaches including the ISMIR Music Genre Classification Contest (classification accuracy 84.07%) shown in table IX.

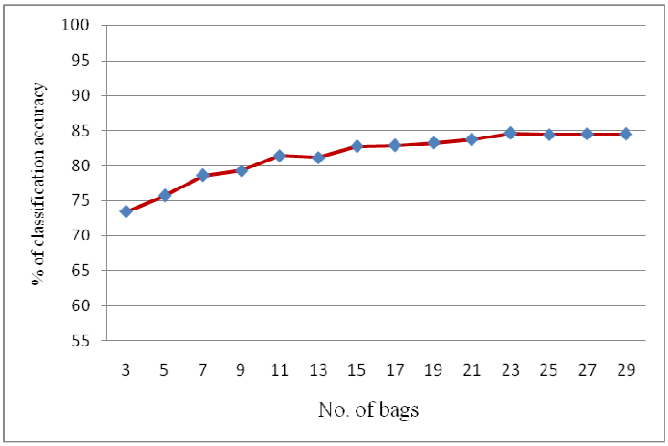


Fig. 6: Number of bags Vs classification accuracy of ZIGEN dataset.

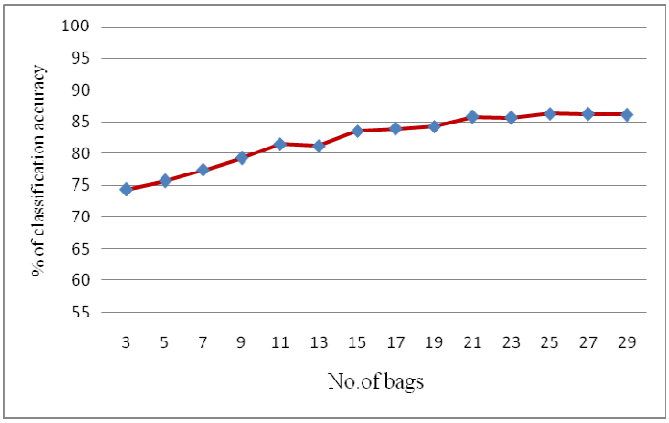


Fig.7: Number of bags Vs classification accuracy of ISMIR2004 dataset.

TABLE VIII
COMPARISON OF CLASSIFICATION ACCURACY WITH OTHER APPROACH OF GTZAN DATASETS (OUR APPROACH BASED ON FIVE-FOLD CROSS VALIDATION

Reference	CA
Our approach	85.15%
Jin S. Seo [11]	84.09%
Bergstra et al [9]	82.50%
Li et al.[1]	78.50%
Tzanetakis [5]	61.00%

TABLE IX
COMPARISON OF CLASSIFICATION ACCURACY WITH OTHER APPROACH OF ISMIR2004 DATASETS

Reference	CA
Our approach (ECFZ+MFCC+beat histogram)	86.46%
Chang-Hsing Lee[10]	86.83%
Jin S. Seo [11]	84.90%
Pampalk et al. [12]	84.07%
Bergstra et al [9]	82.34%
Our approach (ECFZ+MFCC+beat histogram)	86.46%

To get a better picture of the classification accuracy of an individual music genre, the confusion matrices are given. The confusion matrix is *n* x *n* matrix, at which each column of the matrix represents the instances in a predicted class, while each row represents the instances in an actual class. The diagonal entries of the confusion matrix are the rates of music genre classification that are correctly classified, while the off-diagonal entries correspond to misclassification rates.

Table VIII shows the confusion matrix of the GTZAN dataset. The genres are arranged in the order of Classical (Cl), Blues (Bl), Hiphop (Hi), Pop (Po), Rock (Ro), Jazz (Ja), Reggae (Re), Metal (Me), Disco (Di), and Country (Co) respectively. Similarly, Table IX shows the confusion matrix of genre classification of the ISMIR2004 dataset. The genres are arranged in the order of Classical (Cl), Pop and Rock (P&R), Metal and Punk (M&P), Electronic (Ele), World (Wo), and Jazz and Blues (J&B) respectively.

Form the confusion matrix of the GTZAN dataset; we can see that some music genres are classified with significant accuracy like Classical, Pop, Metal, and Reggae. Except for Rock, other music genre classification rates are also competitive. Rock music has a minimum classification rate. It is confused with Metal. Beside this, it is diverse in nature as compared to other genre and also overlaps its characteristic with other genres. Music genre Disco and Country are also confused with Reggae and Metal respectively.

Similarly, Table XI shows the confusion matrix of the ISMIR2004. The classification rate of Classical, Pop and Rock, and Electronic are significant. Genres like Jazz and Blues, and World are also relatively better than Metal and Punk. The World music is diverse in nature, so it is confused with Classical and Pop and Rock. Genre like Jazz and Blues are also confused with Classical. Among the six genres, Electronic has minimum classification rate as compared to others.

TABLE X
CONFUSION MATRIX OF GTZAN DATASETS CLASSIFICATION ACCURACY WITH TIMBRAL TEXTURE AND RHYTHMIC CONTENT FEATURES

	Cl	Bl	Hi	Po	Ro	Ja	Re	Me	Di	Co
Cl	**95.0**	3.67	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.33
Bl	0.0	**85.25**	5.25	0.0	0.0	0.0	2.48	4.52	2.50	0.0
Hi	0.0	3.24	**81.26**	2.06	7.39	0.0	0.0	0.0	8.03	0.0
Po	0.0	0.0	0.0	**96.82**	1.14	0.0	0.0	0.0	2.04	0.0
Ro	2.18	1.15	0.0	1.39	**74.92**	0.0	3.55	7.09	5.03	2.69
Ja	5.21	4.0	0.0	0.0	2.52	**83.12**	0.0	0.0	0.0	5.15
Re	0.0	0.0	0.0	8.64	2.23	3.68	**85.45**	0.0	0.0	0.0
Me	0.0	0.0	0.0	3.89	0.0	5.13	0.0	**89.57**	2.28	1.15
Di	0.0	0.0	2.35	0.0	4.95	0.0	6.25	2.35	**83.90**	0.0
Co	0.0	6.12	0.0	4.19	0.0	0.0	0.0	9.18	0.0	**80.51**

TABLE XI
CONFUSION MATRIX OF ISMIR2004 DATASET CLASSIFICATION ACCURACY OF TIMBRAL TEXTURE AND RHYTHMIC CONTENT FEATURES

	Cl	P&R	M&P	Ele	Wo	J&B
Cl	**95.75**	0.0	0.0	3.14	1.13	0.0
P&R	1.45	**90.62**	0.0	4.58	3.35	0.0
M&P	7.18	10.19	**78.63**	0.0	2.52	1.48
Ele	0.0	3.28	4.46	**87.02**	3.85	1.39
Wo	6.47	7.82	0.0	2.86	**80.85**	0.0
J&B	8.14	2.62	0.0	3.32	0.0	**85.92**

VII. CONCLUSIONS

In this paper, first we analysed the validity of timbral texture features. The validity criterion is determined by the normalized standard deviation of each feature. In the second stage, the frame-wise features have been integrated by using central moments including mean, standard deviation, skewness and kurtosis. Also, we propose the covariance components between timbral texture frame-wise features to be included for improving the classification performance. By considering these feature values, several experiments have been performed separately to analyse classification accuracy among the different feature sets. The ELMs combined with bagging is used to build the classifier. The ELM is an unstable classifier; therefore ELMs with bagging improved the classification accuracy as well as the generalization performance.

The classification accuracy of both datasets (GTZAN and ISMIR2004) is shown in table IV and V, respectively. According to our proposed method, the classification accuracy of 85.15% is achieved in GTZAN datasets. The experimental results on the ISMIR2004 genre datasets have also shown that our proposed approach achieves higher classification accuracy (86.46%) than the ISMIR Music Genre Classification Contest with classification accuracy (84.07%) competitive with Chang-Hsing Lee (86.83%).

Experimental results show that there are no significant differences seen while considering seven and twelve MFCC coefficients for genre classification. It concludes that the minimum feature dimension is also sufficient for music genre discrimination. In addition, our experiment shows that covariance components have a significant impact in improving the genre classification. By adding the components we could improve approximately 5% of the overall accuracy. We expect that a more accurate classifier can be constructed with more features added such as segment-based ones after partitioning audio data into pieces, even though it increases the complexity of the classifier.

Acknowledgement

This work was partially supported by the National Research Foundation of Korea grant funded by the Korean government 2011-0022152 and BK21PLUS project.

REFERENCES

[1] Tao Li, Mitsunori Ogihara, and Qi Li, "A comparative study on content-based music genre classification", Proceedings of the 26th Annual International *ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 282-289, Toronto, Canada, 2003

[2] Xu, N. C. Maddage, and X. Shao, "Automatic music classification and summarization," *IEEE Trans.Speech Aud. Processing*, vol. 13,no. 3, pp. 441–450, May 2005.

[3] A. Meng, P. Ahrendt, J. Larsen, L. Hansen, Temporal feature integration for music genre classification, *IEEE Transactions on Audio, Speech, and Language Processing* 15 (2007).

[4] Babu Kaji Baniya, Deepak Ghimire, and Joonwhoan Lee, "Evaluation of different audio features for musical genre classification" *in proc. IEEE workshop on Signal Processing Systems*, Taipei, Taiwan, 2013

[5] Babu Kaji Baniya, Deepak Ghimire, and Joonwhoan Lee, "A Novel Approach of Automatic Music Genre Classification Based on Timbral Texture and Rhythmic Content Features", *Int. conference on Advance Communications Technology* (ICACT), pp.96-102, 2014

[6] R. Groeneveld, G. Meeden, Measuring skewness and kurtosis, The Statistician 33 (1984) 391-399
[7] Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Trans Speech Audio Process*., vol. 10, no.3, pp. 293-302, Jul. 2002
[8] L. Rabiner and B. Juang. Fundamentals of Speech Recognition. Prentice-Hall, NJ, 1993.
[9] B. Logan. Mel frequency cepstral coefficients for music modeling. *In Proc. Int. Symposium on Music Information Retrieval ISMIR*, 2000.
[10] Gouyon, A. Klapuri, S. Dixon, M. Alonso, G. Tzanetakis, C. Uhle , P. Cano, An Experimental Comparison of Audio Tempo Induction Algorithms, *IEEE Transactions on Speech and Audio Processing*, vol.14, page 1832-44, 2006.
[11] Bergstra, J., Casagrande, N., Erhan, D., Eck, D. and Kegl B. "Aggregate features and AdaBoost for music classification", *Machine Learning*, Vol. 65, No. 2-3, pp. 473-484, 2006.
[12] Chang-Hsing Lee, Jau-Ling Shih, Kun-Ming Yu, and Hwai-San Lin, "Automatic Music Genre Classification Based on Modulation Spectral Analysis of Spectral and Cepstral Features", *IEEE Trans. of Multimedia*, Vol.11, no. 4, pp. 670-82, June 2009.
[13] Jin S. Seo, Seungjae Lee, Higher-order moments for musical genre classification, *Signal Processing*, vol. 91, Issue 8, pp. 2154-57, 2011
[14] Pampalk, E., Flexer, A. and Widmer, G. "Improvements of audio-based music similarity and genre classification", *Proceedings of the Sixth International Symposium on Music Information Retrieval*, London, UK, 2005.
[15] T. Lambrou, P. Kudumakis, R. Speller, M. Sandler, and A. Linney. Classification of audio signals using statistical features on time and wavelet transform domains. *In Proc. Int. Conf. Acoustic, Speech, and Signal Processing (ICASSP-98)*, volume 6, pages 3621–3624, 1998.
[16] H. Soltau, T. Schultz, and M. Westphal. Recognition of music types. *In Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing*, 1998.
[17] H. Deshpande, R. Singh, and U. Nam. Classification of music signals in the visual domain. In Proceedings of the COST-G6 *Conference on Digital Audio Effects*, 2001.
[18] E. Scheirer and M. Slaney, "Construction and evaluation of a robust multifeature speech/music discriminator," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing (ICASSP), 1997, pp.1331-1334.*
[19] T.Zhang and C.J.Kuo, "Audio Content Analysis for Online Audiovisual Data Segmentation and Classification," *IEEE Trans. Speech and Audio Processing*, Vo1.9, No.4, pp. 441-457, May 2000.
[20] H. Terasawa, M. Slaney, and J. Berger. "Perceptual distance in timbrals space". *In Proceedings of Eleventh Meeting of the International Conference on Auditory Display*, pages 61-68 Limerick, Ireland, July 2005.
[21] Beth Logan, Mel Frequency Cepstral Coefficients for Music Modeling, Int. *Symposium on Music Information Retrieval* (2000)
[22] I. Daubechies, "Orthonormal bases of compactly supported wavelets," Commun. Pure Appl. Math, vol. 41, pp. 909-996, 1988
[23] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. 4, pp. 251—257, 1991.
[24] Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme Learning Machine: Theory and Applications. *Neurocomputing* 2006, 70, 489-501.
[25] Beriman, L. Bagging Predictors. *Machine Learning* 1996, 24, 123-140.
[26] Marasys, "Data sets" http://marsysas.info/download/data
[27] [Online]. Available: http://ismir2004.ismir.net/ISMIR_Contest.html

BIOGRAPHIES

Babu Kaji Baniya received the B.E. degree in Computer Engineering from Pokhara University, Nepal in 2005 and M.E. degree in Electronic Engineering from Chonbuk National University, Rep. of Korea in 2010. Currently he is pursuing his Ph.D. degree in Computer Science and Engineering at Chonbuk National University, Rep. of Korea from 2011. His main research interest includes audio signal processing, music information retrieval, source separation, pattern recognition etc.

Deepak Ghimire received the B.E. degree in Computer Engineering from Pokhara University, Nepal in 2007 and M.S. degree in Computer Science and Engineering from Chonbuk National University, Rep. of Korea in 2011. Currently he is pursuing his Ph.D. degree in Computer Science and Engineering at Chonbuk National University, Rep. of Korea from 2011. His main research interest includes image processing, computer vision, pattern recognition, facial emotion analysis etc.

Joonwhoan Lee received his BS degree in Electronic Engineering from the University of Hanyang, Rep. of Korea in 1980. He received his MS degree in Electrical and Electronics Engineering from KAIST University, Rep. of Korea in 1982 and the Ph.D. degree in Electrical and Computer Engineering from University of Missouri, USA, in 1990. He is currently a Professor in Department of Computer Engineering, Chonbuk National University, Rep. of Korea. His research interests include image processing, computer vision, emotion engineering etc.

Performance Analysis of Power Allocation and Relay Location in a Cooperative Relay Network

Muhammad H.D. Khan, Mohammed S. Elmusrati

Communication and System Engineering Group, University of Vaasa, Finland
hazandanish@yahoo.com, moel@uwasa.fi

***Abstract*— Transmission using cooperative relays is a new paradigm in wireless communication. The cooperating relays facilitate the process of communication by performing the operations like data transmission and data processing in a distributed manner. In this wireless system every node is an active element and can act as a relay node. So when a group of these cooperating nodes are involved in a communication stream, a virtual multiple input multiple output (MIMO) system is formed which provides the networks with additional benefits of spectral efficiency and error reduction. Since this system is based on the traditional wireless sensor network (WSN) in which each node has a limited power and computational resource. Therefore, energy efficiency achieved by employing cooperating relays is not sufficient enough. Some extra measures need to be taken to decrease the power consumption of the network. This paper is an effort in this direction, as a power efficient allocation algorithm has been proposed which allocates transmission power optimally to the source node and the involved relay nodes. In the first part of the paper mathematical expressions have been derived for various phases in a cooperative relay transmission. The performance efficiency of the system has been presented using average bit error rate (ABER) as a performance criteria. In the second part, a power allocation algorithm has been derived and employed in a multi-hop cooperative relay network having 4-nodes, with amplify and forward (AF) protocol as its relaying technique. The efficiency of the power allocation algorithm (OPA) has been further investigated with respect to relay location in a network. Simulation results validate the performance efficiency of OPA in different transmission scenarios.**

***Keyword*— Power Allocation, Cooperative Relay Network, Multi-hop Transmission, Relaying Technique, Multiple Input Multiple Output and Diversity Combining.**

I. INTRODUCTION

For the past few years there has been a substantial increase in wireless communication services and applications. This advancement has brought us to the critical issue of efficiently utilizing the network resources. Communication through cooperating relay nodes is one such technology that uses the network resources dynamically rather than the traditional fixed access approach. These cooperating relays share their network resources to enhance the network productivity by providing distributed transmission [1]. This new wireless network seems quite promising, as it incorporates extra capabilities to a traditional WSN and at the same time increases spectrum efficiency in a significant way [2]. In a nutshell, a cooperative relay network (CRN) can provide its end users with high bandwidth efficiency, extended service coverage and ubiquitous connectivity.

In this technique spatial diversity gain is achieved by sharing the antenna resource of cooperating relay nodes to form a virtual MIMO system, as proposed in [3][4]. The source node and the involved relay nodes simultaneously transmit data over independent fading paths. At the receiver node these multiple streams of same correlated data is used to achieve a spatial diversity gain. The gain achieved due to diversity offers advantages like reduction in both error rate and the required transmission power. The relaying protocols being run on the nodes can be categorized into different types [5]. Among these types, amplify and forward (AF) protocol is a widely adopted one. This protocol simply amplifies the received signal and then forwards it to the next node without any complex processing, making it an efficient and less complex technique [6]. So, the system analysis has been carried out using AF relaying protocol.

In order to practically integrate this promising technique in the future wireless networks there are few issues, which need to be addressed. Among them transmission power management is a crucial one. Although extensive research has been carried out for relay networks but very few are focusing the issue of power efficiency. Optimal power allocation (OPA), is one of the techniques that allocates the transmission power optimally between the source and the involved relay nodes. Therefore, a substantial amount of node power is saved while maintaining the link quality [7]. The power allocation of a two-hop AF cooperation network was examined in [8] and [9] for Rayleigh fading, but they only considered a single relay network. In [10], the power allocation has been discussed for a multi-node network but decode-and-forward (DF) relaying protocol has been employed over Rician fading channel.

Manuscript received May 10, 2014. This work was supported in part by the Communication and System Engineering Group at University of Vaasa, 65200 Vaasa, Finland. Part of the results in this paper was presented at the International Conference on Advanced Communications Technology, Pyeongchang, South Korea, 2014.

Muhammad H. D. Khan is with the University of Vaasa, 65200 Vaasa, Finland. (phone: +92-346-5280626; e-mail: hazandanish@yahoo.com).

Mohammed S. Elmusrati is with the University of Vaasa, 65200 Vaasa, Finland. (phone: +358-504-003763; e-mail: moel@uwasa.fi).

Moreover in [11] [12] and [13], OPA has been derived for AF dual-hop network with arbitrary number of relays over Rayleigh fading. In [14], an OPA technique has been proposed for a multi-hop cooperative relay network but the network scenarios based on the location of relay in the network have been discussed very briefly.

The remainder of this paper is organized as follows. Section II, presents the network model and parameters. In Sections III, the mathematical expressions of the received signals have been derived for different phases of transmission. ABER has been taken as criteria to show the performance efficiency of power allocation algorithm. In Sections IV, the expressions for the required transmission power of the source and the relay nodes have been derived for the corresponding 4-node network configuration. The performance efficiency gained by employing the proposed technique is presented in form of performance curves, in Section V. Section VI, concludes the paper.

II. Multi-Hop Cooperative Network

A. Network Model

For this scenario the network model is shown below in Fig 1. This network has four communication nodes; a source node *S*, two relay nodes namely *R1* and *R2* and lastly a destination node *D*. The channel gains for the source to destination, source to relay1, relay1 to relay2 and relay2 to destination are represented by terms $G_{SD,}$ $G_{SR1,}$ G_{R1R2} and G_{R2D} respectively. The network model is designed based on few assumptions. The channel is assumed to be a Rayleigh fading channel and is normalized so that the fading coefficient matrix is complex Gaussian with zero-mean and variance σ^2. Every communicating node in the network will obey rules of half duplex transmission that means it will either transmit or listen at any given instant. The multiple signals received by the receiver node will be combined using Maximal Ratio Combining (MRC) technique. Finally, the power constraints are applied for the whole communication link instead of intermediate hops and AF protocol has been employed as a relaying strategy.

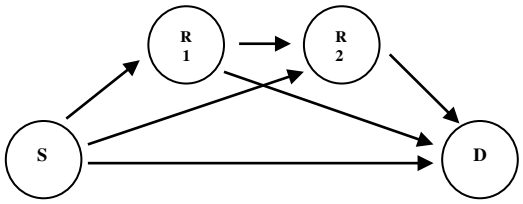


Figure 1. Cooperative Relay Network (2-Relays)

In the first phase the source node *S* broadcasts the data *X* with its transmission power P_S to both relays and destination node, as shown in the next figure. The channel gains for the source-relay1 and from source- relay2 links are G_{SR1} and G_{SR2} respectively. The AWGN present at the relay 1 and relay 2 nodes will be n_{SR1} and n_{SR2} respectively. The equation of received signals in the phase-I at the relay 1 and relay 2 will be expressed as Y_{R1} and Y'_{R2} respectively

$$\mathrm{Y_{R1}}[\mathrm{n}] = \sqrt{P_S}\ \ G_{SR1}.\,\mathrm{X}[\mathrm{n}] + \mathrm{n_{SR1}}[\mathrm{n}] \tag{1}$$

$$\mathrm{Y'_{R2}}[\mathrm{n}] = \sqrt{P_S}\ \ G_{SR2}.\,\mathrm{X}[\mathrm{n}] + \mathrm{n_{SR2}}[\mathrm{n}] \tag{2}$$

After Phase-II, the relay2 node has received two signals Y'_{R2} and Y_{R2}, first one was in phase-I (broadcasting) and the second one was forwarded by relay1 with its amplification factor β_{R1} in phase-II. G_{R1R2} is the channel gain for relay1-relay2 link and n_{R1R2} is the noise at the relay node2 in phase-II. But since the link between source-relay2 is weak as compared to the link between relay1 and relay2 so the node relay2 discards the signal Y'_{R2} and the received signal Y_{R2} is defined as:

$$\mathrm{Y_{R2}}[\mathrm{n}] = \beta_{\mathrm{R1}}\,.\,\mathrm{Y_{R1}}[\mathrm{n}]\,.\ G_{R1R2} + \mathrm{n_{R1R2}}[\mathrm{n}] \tag{3}$$

Relay gains β_{R1} and β_{R2} are used to properly fine-tune the powers at the corresponding relays to reduce variations in the source-relay and relay–destination links. The relaying node's amplifier can provide a maximum gain defined by the following expressions:

$$\beta_{\mathrm{R1}} = \sqrt{\frac{P_{R1}}{P_S|G_{SR1}|^2 + N_0}} \tag{4}$$

$$\beta_{\mathrm{R2}} = \sqrt{\frac{P_{R2}}{P_{R1}|G_{R1R2}|^2 + N_0}} \tag{5}$$

Since, amplify-and-forward protocol may induce some noise amplification but the MRC detector employed is quite competitive therefore with the help of its weights it can compensate effect of induced noise [15]. The output of the MRC detector at the destination node will be:

$$\mathrm{Y_{MRC}}[\mathrm{n}] = \alpha_0(\sqrt{P_S}\ .G_{SD}.\,\mathrm{X}[\mathrm{n}] + \mathrm{n_{SD}}[\mathrm{n}]) + \alpha_1(\beta_{\mathrm{R2}}\,.\,\mathrm{Y_{R2}}[\mathrm{n}].\,G_{R2D} + \mathrm{n_{R2D}}[\mathrm{n}]) \tag{6}$$

Substituting equation (3) into (6):

$$\mathrm{Y_{MRC}}[\mathrm{n}] = \alpha_0(\sqrt{P_S}\ .G_{SD}.\,\mathrm{X}[\mathrm{n}] + \mathrm{n_{SD}}[\mathrm{n}]) + \alpha_1(\beta_{\mathrm{R2}}\,.\,(\beta_{\mathrm{R1}}\,.\,\mathrm{Y_{R1}}[\mathrm{n}]\,.\,G_{R1R2} + \mathrm{n_{R1R2}}[\mathrm{n}])\,.\ G_{R2D} + \mathrm{n_{R2D}}[\mathrm{n}]) \tag{7}$$

Substituting equation (1) into (7):

$$\mathrm{Y_{MRC}}[\mathrm{n}] = \alpha_0(\sqrt{P_S}\ .G_{SD}.\,\mathrm{X}[\mathrm{n}] + \mathrm{n_{SD}}[\mathrm{n}]) + \alpha_1(\beta_{\mathrm{R2}}\,.\,(\beta_{\mathrm{R1}}\,.\,(\sqrt{P_S}\ \ G_{SR1}.\,\mathrm{X}[\mathrm{n}] + \mathrm{n_{SR1}}[\mathrm{n}])\,.\,G_{R1R2} + \mathrm{n_{R1R2}}[\mathrm{n}]).\,G_{R2D} + \mathrm{n_{R2D}}[\mathrm{n}]) \tag{8}$$

In the above equation (6-8), α_0 and α_1 are the weights of the maximal-ratio-combiner. These combining weights compensate for the effects of likely incorrect decisions and are therefore chosen accordingly. Here $\alpha_{1,}$ is the weight of the signals which is being forwarded by the relays placed in a linear topology.

$$\alpha_0 = \frac{\sqrt{P_S}\,G_{SD}{}^*}{N_0}$$

$$\alpha_1 = \frac{\sqrt{P_S}\ .\ \beta_{R1}.\ \beta_{R2}\,.\,G_{SR1}{}^*\,.\ G_{R1R2}{}^*\,.\ G_{R2D}{}^*}{N_0}$$

III. Optimal Power Allocation

Now, the probability of bit error for the 4-node configuration will be calculated using Moment Generating Function (MGF) approach. With the help of moment generating function approach the error probability for the QPSK modulation scheme will be expressed as an exponential function of γ, defined below.

$$P_e = \frac{1}{\pi}\int_0^{\left(\frac{N-1}{N}\right)\pi}\prod_{n=0}^{1} N_{\gamma_n}\left(\frac{g_{QPSK}}{sin^2\theta}\right)d\theta \quad (9)$$

$$N_{\gamma_n} = \int_0^{\infty} P_{\gamma_n}(\gamma)e^{s\gamma}d\gamma \quad (10)$$

$$N_{\gamma_n}\left(\frac{1}{sin\theta^2}\right) = \left(1+\frac{g_{QPSK}}{sin^2\theta}\gamma_n\right)^{-1} \quad (11)$$

$$N_{\gamma_n}\left(\frac{1}{sin\theta^2}\right) = \left(1+\frac{\gamma_n}{sin^2\theta}\right)^{-1} \quad (12)$$

Since ($\gamma_n >> 1$) equation (12) becomes:

$$N_{\gamma_n}\left(\frac{\gamma_n}{sin\theta^2}\right) \cong N_{\gamma_n}\left(\frac{1}{sin\theta^2}\right)$$

$$P_e = \frac{3}{8}(\gamma_0.\gamma_1)^{-1} \quad (13)$$

$$P_e = \frac{3{N_0}^2}{8}\left[\frac{1}{P_S.\ \sigma^2_{SD}\ .\ \sigma^2_{SR1}} + \frac{1}{P_S.P_1.\ \sigma^2_{SD}\ .\ \sigma^2_{R1R2}} + \frac{1}{P_S.P_2.\ \sigma^2_{SD}\ .\ \sigma^2_{R2D}}\right] \quad (14)$$

As the outage probability has been formulated for this 4-node configuration, now this system will be expressed as a constrained optimization problem with transmission power as a constraint entity.

$$\text{Minimize}\quad P_e = \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\prod_{n=0}^{2} N_{\gamma_n}\left(\frac{1}{sin^2\theta}\right)d\theta$$

$$\text{Subject to}\quad P_s + \sum_1^N P_R \le P_T$$

Now as we have established the constrained optimization problem, a closed form expression of this optimization problem will be found using Lagrange Method. The Lagrange cost function of this problem is defined as:

$$J = P_e + \lambda\,(P_s + \sum_1^2 P_R - P_T) \quad (15)$$

$$J = \frac{3{N_0}^2}{8}\left(\frac{1}{{P_S}^2\,\sigma^2_{SD}\ \sigma^2_{SR1}} + \frac{1}{P_S P_1\,\sigma^2_{SD}\ \sigma^2_{R1R2}} + \frac{1}{P_S P_2\,\sigma^2_{SD}\ \sigma^2_{R2D}}\right) + \lambda(P_S + P_1 + P_2 - P_T) \quad (16)$$

Now taking the partial derivatives of the Lagrange cost function J (P, λ), with respect to P_s, P_1, P_2 and λ. Afterwards these derivatives will be equated to zero, like defined below

$$\frac{\partial J}{\partial P_S} = \frac{3N_0^2}{8}\left[\frac{-2P_S.\sigma^2_{SD}\ .\sigma^2_{SR1}}{(P_S{}^2.\sigma^2_{SD}\ .\sigma^2_{SR1})^2} - \frac{P_1.\sigma^2_{SD}\ .\sigma^2_{R1R2}}{(P_S.P_1.\sigma^2_{SD}\ .\sigma^2_{R1R2})^2} - \frac{P_2.\sigma^2_{SD}\ .\sigma^2_{R2D}}{(P_S.P_2.\sigma^2_{SD}\ .\sigma^2_{R2D})^2}\right] + \lambda = 0 \quad (17)$$

$$\frac{\partial J}{\partial P_1} = \frac{3N_0^2}{8}\left[\frac{-P_S.\sigma^2_{SD}\ .\sigma^2_{R1R2}}{(P_S.P_1.\sigma^2_{SD}\ .\sigma^2_{R1R2})^2}\right] + \lambda = 0 \quad (18)$$

$$\frac{\partial J}{\partial P_2} = \frac{3N_0^2}{8}\left[\frac{-P_S.\sigma^2_{SD}\ .\sigma^2_{R2D}}{(P_S.P_2.\sigma^2_{SD}\ .\sigma^2_{R2D})^2}\right] + \lambda = 0 \quad (19)$$

$$\frac{\partial J}{\partial \lambda} = P_S + P_1 + P_2 - P_T = 0 \quad (20)$$

After some algebraic manipulations the above equations are used to solve the value of source node power P_S, relay1 node power P_1 and relays2 node power P_2.

$$P_S = \begin{cases} \dfrac{A - 4B + \sqrt{A^2 + 8AB}}{4(A-B)}P_T \\ \dfrac{2}{3}P_T \end{cases}$$

$$P_1 = \frac{\sigma_{2D}}{\sigma_{2D} + \sigma_{R1R2}}(P_T - P_S)$$

$$P_2 = P_T - P_1 - P_S$$

Where, in the above equation $A=(\sigma_{R2D} + \sigma_{R1R2})^2.\sigma^2_{R1R2}.\sigma^2_{SR1}$ and $B=\sigma^2_{R2D}.\sigma^3_{R1R2}$

IV. Simulation Results

Fig. 2, compares the performance of OPA with that of EPA in terms of ABER for a cooperative network having 2 relays. The results show that OPA outperforms EPA when employed to a cooperative relay network having similar configuration. The performance gap remains same for the increasing values of SNR. It was also an observation that the performance of a system employing EPA can be increased only if the relays were placed symmetrically but OPA performs better for both symmetric and asymmetric configurations.

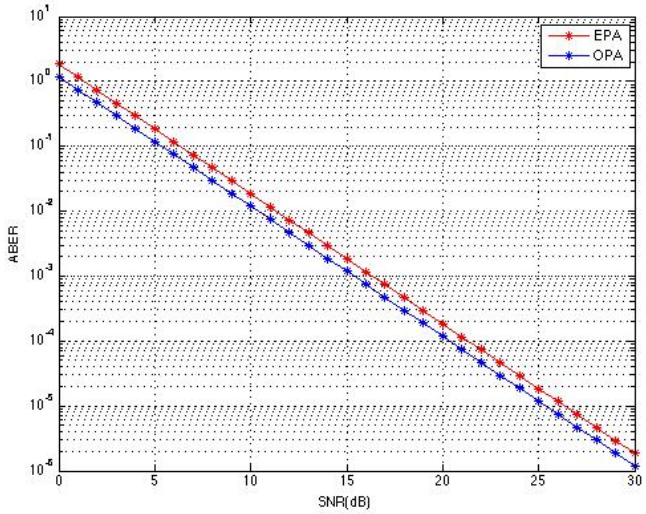


Figure 2. Equal power allocation and Optimal power allocation

Fig. 3, compares the performance of case 1: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) with case 2: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) in terms of ABER for a cooperative network having 2 relays. Here, channel link quality is denoted by "10" if the two communicating nodes are closer to each other while if the two nodes are far from each other channel link quality is "1". For this network simulation, case 2:($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) shows the best performance because in this case relays2 is closer to destination node and source node is closer to relay1 node. Case 1: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) shows slightly worse performance because only relay2 node is closer to the destination node.

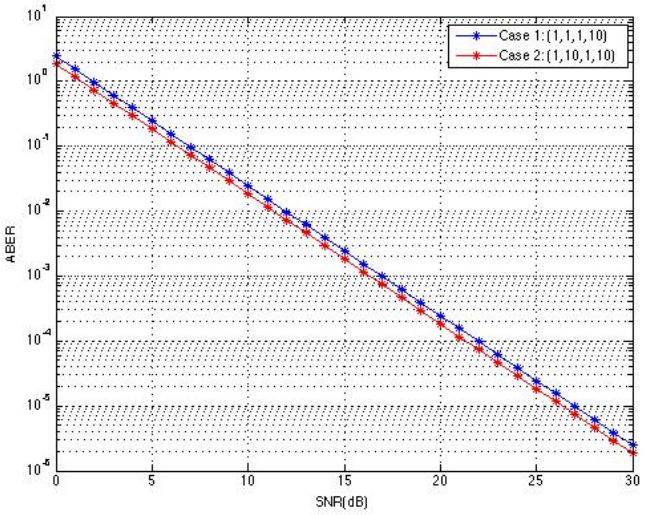


Figure 3. Cooperative Transmission Scenarios (Case1 & Case2)

Fig. 4, compares the performance of case 3: ($\sigma^2_{SD} = 10;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 1$) with case 4: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 10;\ \sigma^2_{R2D} = 10$) in terms of ABER for a cooperative network having 2 relays.

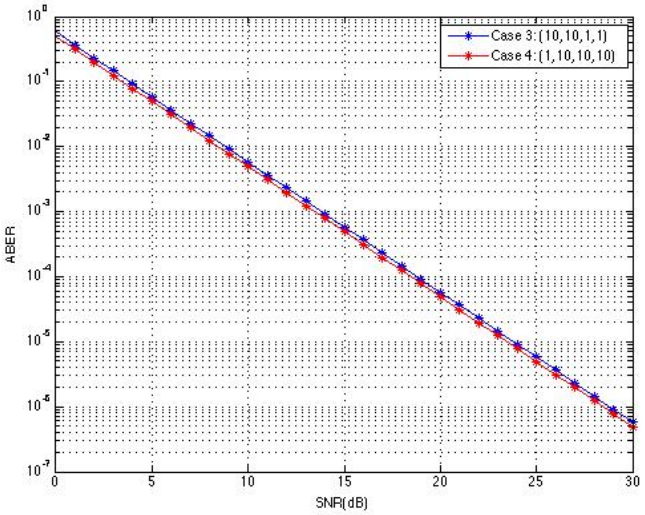


Figure 4. Cooperative Transmission Scenarios (Case3 & Case4)

For this network simulation, case 4:($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 10;\ \sigma^2_{R2D} = 10$) shows the best performance because in this case the channel link quality for every transmisson hop is strong. Case 3: ($\sigma^2_{SD} = 10;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 1$) shows slightly worse performance because only relay1 node is closer to source node.

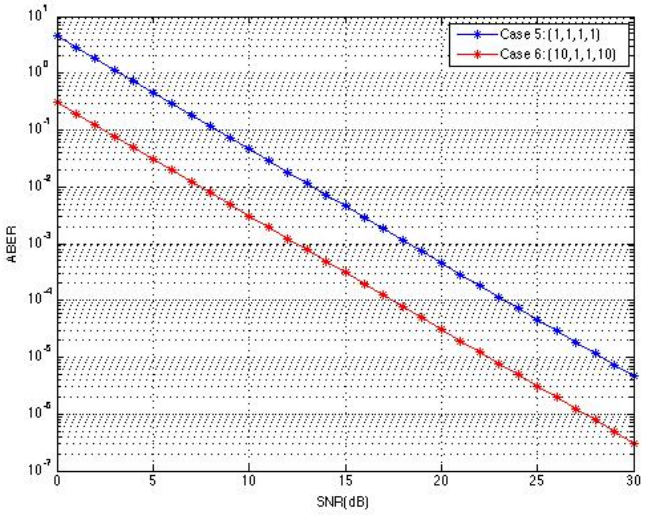


Figure 5. Cooperative Transmission Scenarios (Case5 & Case6)

Fig. 5, compares the performance of case 5: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 1$) with case 6: ($\sigma^2_{SD} = 10;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) in terms of ABER for a cooperative network having 2 relays. Out of all simulations, case 5: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 1$) is the worst one as every tranmsission link is weak . On the contrary, Case 6: ($\sigma^2_{SD} = 10;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) is the best possible scenario because not only relay2 node is closer to destination node but also the direct link between source and destination is also strong.

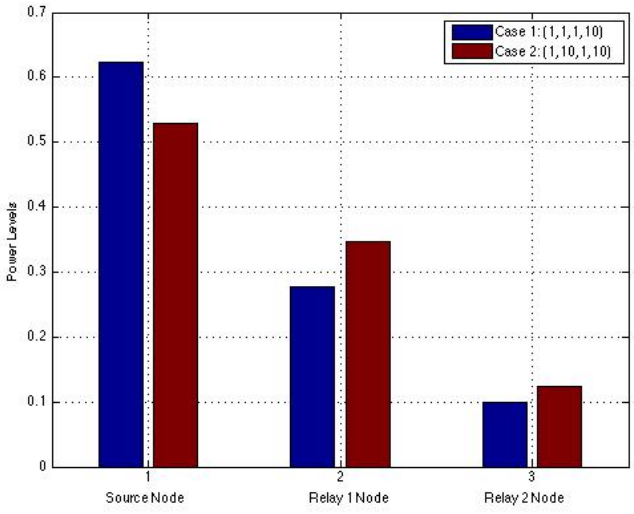


Figure 6. Allocated Power Levels (Case1 & Case2)

Fig. 6, compares the transmission power values for case 1: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) represented by blue bar and case 2: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) represented by red bar.

For case 1: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) the only strong link is between relay2 and destination node. So it is visible from the above graph that the value of transmission power (0.0986) allocated to relay2 node is the lowest when compare to the power values of source node (0.6236) and relay1 node (0.2778). For case 2: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) the power allocation algorithm assigns a transmission power of (0.529) to the source node, a power of (0.348) to the relay1 node and a tranmission power of (0.123) to the relay2 node, satisfying the total power constraint of "1", as defined in the derived mathematical expressions.

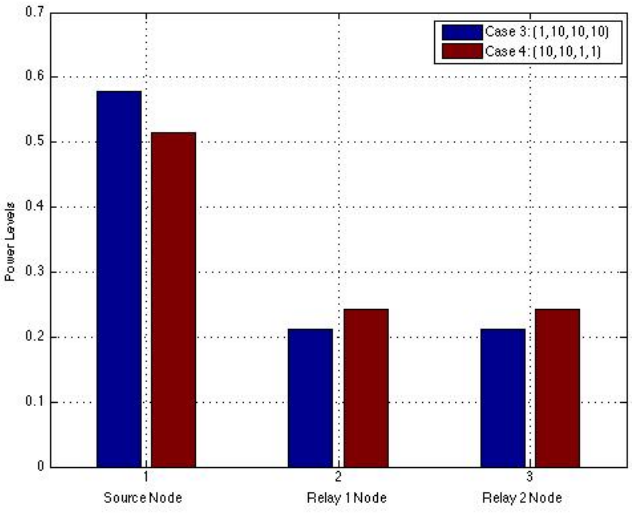


Figure 7. Allocated Power Levels (Case3 & Case4)

Fig. 7, compares the transmission power values for case 3: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 10;\ \sigma^2_{R2D} = 10$) represented by blue bar and case 4: ($\sigma^2_{SD} = 10;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 1$) represented by red bar. For case 3: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 10;\ \sigma^2_{R2D} = 10$) the channel link quality for every transmisson hop between the source and destination is strong so the value of tranmission power allocated to the the relay1 and relay2 node will be equal (0.2113) and source node will be allocated the maximum value of (0.5774). For case 4: ($\sigma^2_{SD} = 10;\ \sigma^2_{SR1} = 10; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 1$) the power allocation algorithm assigns a transmission power of (0.5144) to the source node, a power of (0.2428) to the relay1 node and a tranmission power of (0.2428) to the relay2 node, satisfying the total power constraint of "1".

Fig. 8, compares the transmission power values for case 5: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 1$) represented by blue bar and case 6: ($\sigma^2_{SD} = 10;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) represented by red bar. For case 5: ($\sigma^2_{SD} = 1;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 1$) the channel link quality for every transmisson hop between the source and destination is weak so the value of tranmission power allocated to the the relay1 and relay2 node will be equal (0.2113) and source node will be allocated the maximum value of (0.5774). For case 6: ($\sigma^2_{SD} = 10;\ \sigma^2_{SR1} = 1; \sigma^2_{R1R2} = 1;\ \sigma^2_{R2D} = 10$) the power allocation algorithm assigns a transmission power of (0.6236) to the source node, a power of (0.2778) to the relay1 node and a tranmission power of (0.0986) to the relay2 node, satisfying the total power constraint of "1".

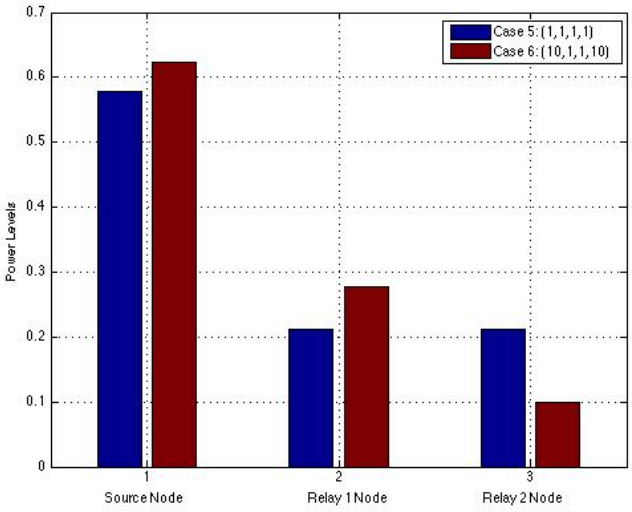


Figure 8. Allocated Power Levels (Case5 & Case6)

V. CONCLUSION

In this paper, we investigated the effect of optimal power allocation and the effect of relay location for a transmission using cooperative relays. The results show that instead of using the traditional EPA if OPA is employed in a cooperative relay network then the system performance can be enhanced substantially. Moreover, OPA is vital for these infrastructure-less networks having un-balanced communication links. With help of the above results it has been found out that the communication links between the source node (S) and the first-relay (R_1) and the link between the last-relay (R_N) and the destination node (D) have the most significant effect on the system performance. Other communication links have slightly less effect on the system performance. Finally as a future direction, optimum relays should be determined first based on their location. Then these relays should be used for transmission, by doing so the system performance can be further enhanced.

ACKNOWLEDGMENT

The authors would like to thank Vaasa University Foundation for the financial support provided by them.

REFERENCES

[1] A. Scaglione and Y.-W. Hong, "Opportunistic large arrays: cooperative transmission in wireless multi-hop ad hoc networks for the reach back channel," *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 2082–2092, Aug. 2003.

[2] Hong, Yao-Win; Wan-Jen Huang; Fu-Hsuan Chiu; Kuo, C.-C.J., "Cooperative Communications in Resource-Constrained Wireless Networks," *Signal Processing Magazine, IEEE* , vol.24, no.3, pp.47,57, May 2007

[3] J. N. Laneman, G. W. Wornell. "Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks" [J]. *IEEE Trans. on Information Theory*, 2003, 49(10): 2415-2425.

[4] J. N. Laneman, David N. C. Tse, G. W. Wornell. "Cooperative diversityin wireless network: efficient protocols and outage behavior" [J]. *IEEE Trans. on Information Theory*, 2004, 50(12): 3062-3080.

[5] Donglin Hu; Shiwen Mao, "Cooperative Relay in Cognitive Radio Networks: Decode-and-Forward or Amplify-and-Forward?," *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE* , vol., no., pp.1,5, 6-10 Dec. 2010

[6] Borade, S.; Lizhong Zheng; Gallager, R., "Amplify-and-Forward in Wireless Relay Networks: Rate, Diversity, and Network Size," *Information Theory, IEEE Transactions on* , vol.53, no.10, pp.3302,3318, Oct. 2007

[7] Fangzhen Li; Fei Yang; Xuefen Zhang, "An optimal power allocation algorithm in Distributed Sensor networks," *Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on* , vol., no., pp.414,418, 16-18 April 2010

[8] Weifeng Su, Ahmed K.Sadek and K.J.Ray Liu. "Cooperative communication protocols in wireless networks: performance analysis and optimum power allocation" [J]. *Wireless Personal Commun*., 2008, 44(2): 181-217.

[9] Yan Wang, Fei Lin. "SEP performance analysis and power allocation for amplify-and-forward relay networks" [C].*The 3rd IEEE International Conference on Wireless Communications, Networking and Mobile Computing*, September 21-23, 2007, Shanghai China (Wicom 2007):1274-1277.

[10] Mulugeta K. Fikadu, Mohammed Elmusrati, and Reino Virrankoski, "Power Allocation in Multi-node Cooperative Network in Rician Fading Channels". *IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)* 2012: 496-501

[11] Sarmad Sohaib and Daniel K. C. So, "Power Allocation for Multi-Relay Amplify-and-Forward Cooperative Networks". *IEEE Conference on Communications (ICC)*, 2011: 1-5

[12] Nan Zhang, Jian Hua Ge and Feng Kui Gong , "SEP analysis and optimal power allocation of multinode amplify-and-forward cooperation systems". 2008, 1 – 5

[13] Jia Liu, Ness B. Shroff, and Hanif D. Sherali. "Optimal Power Allocation in Multi-Relay MIMO Cooperative Networks: Theory and Algorithms." *IEEE Journal on Selected Areas in Communications*. 2012, 30(2): 331-340

[14] Khan, M.H.D.; Elmusrati, M.S.; Virrankoski, R., "Optimal power allocation in multi-hop cooperative network using non-regenerative relaying protocol," *Advanced Communication Technology (ICACT), 2014 16th International Conference on* , vol., no., pp.1188,1193, 16-19 Feb. 2014

[15] Vien, N.H.; Nguyen, H.H., "Performance Analysis of Fixed-Gain Amplify-and-Forward Relaying With MRC," *Vehicular Technology, IEEE Transactions on* , vol.59, no.3, pp.1544,1552, March 2010

Muhammad H. D. Khan (S'13) was born in Rawalpindi, Pakistan, in 1986. He received the B.E. degree in electrical engineering from the COMSATS Institute of Information Technology, Islamabad, Pakistan, in 2010, and the M.Sc. degree in technology from University of Vaasa, Vaasa, Finland, in 2013.In 2010, he joined the Department of Avionics Engineering, Air University, Pakistan, as a Research Assistant, and in 2011 became a Research Associate. His current research interests include wireless communication systems, sensor networks and resource allocation management.

Mohammed Elmusrati (S'00-M'04-SM'12) received the B.Sc. (with honors) and M.Sc. (with high honors) degrees in telecommunication engineering from the Electrical and Electronic Engineering Department, Benghazi University, Libya, in 1991 and 1995, respectively, and the Licentiate of Science in technology (with distinction) and the Doctor of Science in Technology degrees in control engineering from Aalto University - Finland, in 2002 and 2004, respectively. Currently, Elmusrati is full professor and head of communications and systems engineering group at University of Vaasa - Finland. His main research interests include Radio resource management in wireless communication, wireless networked control, game theory, and smart grids.

Road Side Unit Assisted Stochastic Multi-hop Broadcast Scheme for Instant Emergency Message Propagation

Xing Fan*, Bo Yang*, Ryo Yamamoto**, Yoshiaki Tanaka*,***

** Global Information and Telecommunication Institute, Waseda University*

1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-0051 Japan

*** Graduate School of Information Systems, The University of Electro-Communications*

1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan

**** Research Institute for Science and Engineering, Waseda University*

17 Kikuicho, Shinjuku-ku, Tokyo, 162-0044 Japan

fan.xing@asagi.waseda.jp, yangbo_youhaku@ruri.waseda.jp, ryo_yamamoto@is.uec.ac.jp, ytanaka@waseda.jp

***Abstract*— VANET (Vehicular ad hoc Network) is a special kind of ad hoc wireless network where every single node is a vehicle moving in a relatively high velocity, which leads to exclusive challenges like rapid changing topologies, safety and privacy concerns. In this specific network, broadcasts tend to be carrying important messages such as car accident notification, disaster alert or extreme traffic condition. Thus the propagation of broadcasted emergency messages could be critical to save human lives and property. Many researchers have proposed routing or broadcast protocols to solve this problem in VANET. With consideration of the other common issues. The objective of this paper is to propose a broadcast scheme in VANET that is not likely to cause broadcast storm problem with a reasonable delay and high delivery rate. Since VANET is an attack-prone network and any kind of malicious behavior in VANET might cause serious loss or even death in reality, we should also refrain from using beacons to exchange privacy-sensitive information in V2V (Vehicle to Vehicle). In this paper, a multi-hop broadcast scheme that makes use of RSU and V2I (Vehicle to Infrastructure) communication is proposed. The simulation result shows that the proposed scheme outperforms static stochastic broadcast scheme in terms of delivery rate. Comparing to flooding, we offer a better delay and less network usage.**

***Keyword*— VANET, Broadcast, Percolation, Stochastic, Wireless**

Manuscript received April 21, 2014.

X. FAN is with the Global Information and Telecommunication Institute, Waseda University, Tokyo, 169-0051 Japan. phone: 080-9442-3698; e-mail: fan.xing@asagi.waseda.jp.

B. YANG is with the Global Information and Telecommunication Institute, Waseda University, Tokyo, 169-0051 Japan. phone: 090-0495-246104; fax: 090-0495-246104; e-mail: yangbo_youhaku@ruri.waseda.jp.

R. YAMAMOTO is with the Graduate School of Information Systems, The University of Electro-Communications, Tokyo, 182-8585, Japan. e-mail: ryo_yamamoto@is.uec.ac.jp.

Y. TANAKA is with the Global Information and Telecommunication Institute, Waseda University, Tokyo, 169-0051 Japan. He is also with the Research Institute for Science and Engineering, Waseda University, Tokyo, 162-0044 Japan. e-mail: ytanaka@waseda.jp.

I. INTRODUCTION

WIRELESS ad hoc network is a network that does not rely on existing infrastructure to relay packets and uses all nodes both as clients and as routers. In this way, it does not require any type of access points or routers to be pre-installed in the network. It's decentralized, dynamic and easy-to-deploy characteristics make this kind of network a very active and challenging research area.

One of the most promising and practical topics is vehicular ad hoc network (VANET) [1]. Research in this area is significant because the increasing car accident and safety related problems [2] are threatening human lives and property, and intelligent transportation system (ITS) is designed to help solving these problems. In developing countries, motorization is happening in urban area with high population and cars are moving at high rather speed in these regions, which is considerably dangerous. Vehicle manufactures are developing various ITS systems to assist and protect driver's life and their property. VANET is one of the most basic systems to be implemented. Due to its wireless ad hoc nature, it could sustain extremely low connectivity and being established quickly, easily even in extreme condition.

VANET is simply wireless ad hoc network formed by vehicles, which allows nodes to perform vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) communication to obtain traffic condition, alert messages and benefit from navigation system. Despite of the simple concept of VANET, it introduces a lot of exclusive problems, mainly because cars are nothing like mobile devices in MANET (Mobile ad hoc network) or stationary monitors in WSN (Wireless sensor network). Its specialty also gives the following unique characteristic to VANET:

- Nodes move at rather high velocity
- Nodes move in an organized, predictable way
- Power consumption is usually negligible for vehicles

- Nodes are assumed to have modern on-board-unit (OBU) installed, including GPS, WIFI connectivity, etc.
- A special kind of stationary nodes could exist, which is called road-side-unit (RSU), as is shown in Figure 1

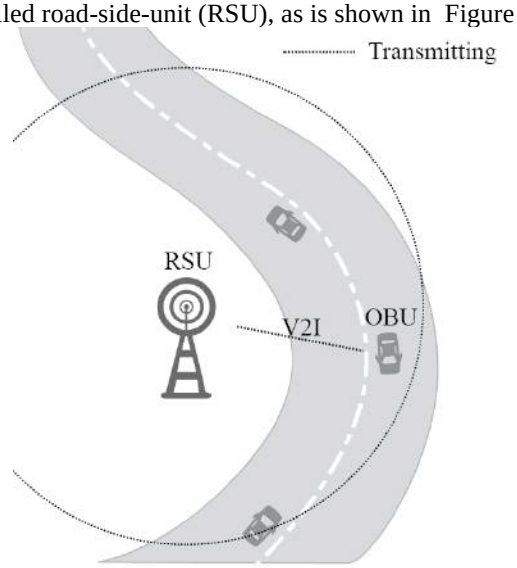


Fig. 1. RSU and V2I in VANET

The previous listed characteristics makes every applicable protocol and scheme for VANET requires much more detailed design rather than directly import from currently existing ad hoc network technique such as mobile ad hoc network(MANET) [3]. In addition to other ad hoc network, researchers need to take the following risks into consideration in VANET:

- The usage of privacy-sensitive information
- Easily fragmented network topology
- Security
- Scalability

Also, the usage of beacons and geography information services should be limited. Beacons should not carry privacy sensitive information such as geo-location, vehicle's direction or speed, because beacons are prone to interception. We also shouldn't make assumption of location services because emergency broadcast should be working independently from any infrastructure-dependently network, to make sure its availability even in extreme condition.

Scalability is another major concern due to the unpredictable size of ad hoc network. In modern VANET, the size of the network could expand across cities, or even counties connected by highways or railways. Among the connections, the network condition is also not unified. For example cars might be using different OBUs and the distribution or even existence of RSUs might vary according to culture and corresponding policies. If a scheme can only work with strict dependencies it is said not scalable.

Since most of concerns in VANET are safety of the vehicles and their drivers, it is not difficult to infer that emergency broadcast is one of the most critical areas in VANET communication. Every car accident, natural disaster and other alerts need to be propagated in the network very fast with reliability. Optimization of emergency message propagating includes many areas such as reliability research, broadcast schemes in VANET and privacy protection. In these areas, broadcast schemes are the most direct contributors. For this reason, this paper mainly focused on emergency broadcast schemes in VANET. As a common methodology in propagating emergency messages, broadcast is widely used in many areas. Pure flooding is one of the most simple and effective methods, namely every node must retransmit the packet when it receives the one. Although this scheme is good enough for simple situation, it is not robust to be applied for practical usage. This is because flooding packets without constraints could potentially cost too many resources in the high density area and increase the duplicate messages while risking congest the network, which could finally lead to collisions or even unresponsiveness (broadcast storm problem). As we stated actual VANET can have a quite large scale, it might span across cities or even countries, in this case emergency messages are no longer interested by distant audience but the cost of propagating those messages can grow exponentially. Thus, further design and constraints must apply.

In order to solve this problem, many approaches were proposed [4]. In counter-based approaches, number of rebroadcasting is limited and nodes would not retransmit a packet after the packet reaches a certain counter. In location-based approaches, only nodes in a specified location retransmit a packet. In distance-based approaches, only nodes with furthest distance to a source node can retransmit a packet. In cluster-based approaches, packet retransmitting relies on certain nodes within a specified structure. In probability-based approaches, every node has a possibility to retransmit packet when they receive one.

Each solution proposed above has its pros and cons. Counter-based approaches can somehow control the transmission range in terms of hop count but are prone to malicious modification since the counters are usually stored in the message. Also the hop count and actual distance are not directly associated, it needs to refer the node density and sometimes network topology. Location-based approaches specified location area in packets so that it could limit the transmission area. Thus, it is also prone to attacks since attackers can easily obtain information of broadcasting area. And a stable access to location service is required for each node, which is rather hard to guarantee in VANET and might have negative impact on performance. Distance-based approaches are the state of the art broadcast scheme because it tends to select the furthest nodes as relays. Thus, it maximizes the dissemination ranges. On the other hand, it does not have any control scheme once a packet is broadcasted to networks. Thus, priority or counter based mechanism must be applied to limit the transmission. Probability-based approaches are flexible schemes since all their behaviours rely on how to decide the probability, depending on different scenarios researchers use different approaches to adjust the rebroadcast probability accordingly. Most of the pragmatic schemes are hybrids of the four basic schemes or with modifications.

The effectiveness of broadcast also heavily depends on a technique of packet transmission. Dedicated short-range communication (DSRC) [5] allows high-speed connection in both V2V and V2I. It uses 5.9 GHz band and transmission range could reach 1 km. This allows us to transmit packet over 3 times further than conventional IEEE 802.11 [6] and also mitigates a fragmentation of network in low node density

area. The power consumption of this module is also negligible since cars are generating power on the run.

As is mentioned previously, there are special class of nodes called RSU (road side unit) exist exclusively in VANET. These nodes are usually installed in an intersect area on a highway, gas station and the other area according to the node density. Usually to form an infrastructure network or directly connected to one. They can also boost the connectivity in their own area. However, these RSUs rarely contribute to conventional VANET broadcast. One of the potential optimization for broadcast scheme is to use RSUs as a resources.

This paper is organized as follows: introduction of some existing broadcast schemes are in Section II, the proposed network architecture and usage of RSU are explained in Section III, detailed design of broadcast scheme is written in Section IV, simulation and result are recorded in Section V, Section VI conclude this paper.

II. Related Works

Most of the practical broadcast schemes are either variation or hybrid of the 4 basic broadcasts: Counter, Location, Distance and Probability [4].

Smart broadcast [7] is a variation of distance-based broadcast, the area around a source node is divided into different sections as shown in Figure 2, which allows every node to have their own contention window. In this way, the outer most nodes in S2 have the highest priority to access the channel. When broadcast starts, other nodes around the source node will start trying to gain the access to the broadcast channel but only the node with

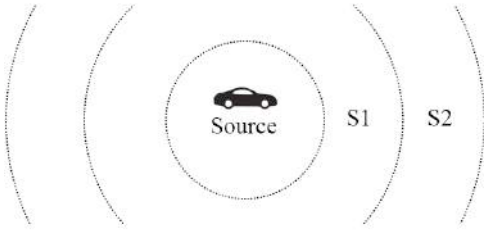


Fig. 2. Sections in smart broadcast

higher priority (further distance to the source node) will have it. This scheme provides a way to select relay for rebroadcasting using distance-based broadcasting scheme and it is practical and it guarantees maximum dissemination range.

Another interesting broadcasting is called stochastic broadcast for VANET [8], which is a combination of distance-based broadcast and stochastic broadcast. In the paper, every node must decide their own rebroadcast probability (the probability of being a relay). The paper proposed two ways to determine a retransmit probability: one is fixed; another dynamically determined according to a one-hop distance. That research has following 2 advantages. First, every node determines a probability by itself, which means transmission is not used except broadcasting and the relay selection process is considerably safer without bringing in impact to the performance. This eliminates privacy issues and provides scalability. Second, it uses a percolation study to estimate a theoretically maximum connectivity and calculate the optimal node density to form a connected graph. However, this paper does not mention about dynamical adjustment of the probability according to a network condition changes. In this paper, the proposed scheme utilizes RSUs to deal with the rapid fragmented network.

The objective of this paper is to propose a broadcast that could propagate emergency messages with high deliver rate, reasonable delay and does not overwhelm a network with broadcast storm. Privacy is also taken into consideration by limit a usage of beacons.

III. Network Architecture

VANET could be simplified into two scenarios: city with high node density, highway with sparse density and easily fragmented. Oversimplify the VANET could lead to assume the network is connected, which ignores the fact that VANET is easily fragmented due to its nature [9]. In this paper, instead of making such assumptions, we use RSUs to mitigate network fragmentation and use percolation to help us study fragmentation.

Network fragmentation is a common problem when network connection between several intermediate nodes are lost, these nodes often hold connection to a cluster, and as a result, the connection to another whole cluster of nodes are lost. In wireless network, fragmentation could appear when node density is extremely low or a sparse distribution. The consequences are equally severe: it causes delivery rate to drop significantly.

To deal with fragmentation problems in VANET, this paper proposes a usage of RSUs to assist a broadcast. This paper assumes that every node equips an OBU with DSRC and GPS enabled. Moreover, limited number of RSUs was installed along the road. Although how and where to install RSUs is out the scope this paper, we assume a generic situation. That is, RSUs are deployed in a scattered manner that cannot form a connected network and every RSU works independently.

In a multi-hop broadcast, when a source node initializes a broadcast, it simply broadcasts a message and utilizes some mechanism to let other nodes relay it. In conventional broadcast schemes, there are no roles for RSUs to participate in this process. Theoretically, putting RSUs into broadcast process by letting them rebroadcast with magnified signal and further transmitting range is beneficial. Here, we proposed a mechanism to be installed on RSUs that could still work independently to participate and contribute to the broadcast process even if they are not connected.

The proposed scheme for RSU in VANET broadcast could be divided into following two parts. For the first part, RSU should use a maximum transmission range to rebroadcast a message when it receives the one. For the second part, to mitigate the broadcast storm, it should suppress a further rebroadcast in the inner region of its broadcast range by using

certain approach as shown in Figure 3. The detailed broadcast scheme for RSU side is described as follows:

(1) When RSU receives a broadcast message, check if it is duplicate message.
(2) If it is not duplicated, RSU immediately preform a rebroadcast burst with a message that contains the original message and also two addition scalar values: broadcast suppression range R_s and minimum broadcast suppression P_m.

Note that R_s and P_m are used to suppress the rebroadcast of nodes in the transmitting range of RSU.

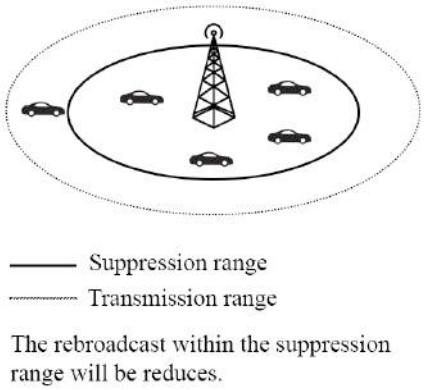


Fig. 3. RSU broadcast suppression

IV. Proposed Broadcast Scheme

The proposed broadcast scheme was developed on top of the basic probability-based broadcast, which is explained as follows:

(1) Every node maintains a value P_c, which indicates the current rebroadcast probability.
(2) While receiving a non-duplicate broadcast message, a node enters a probability check process that would cause the node has the probability P_c to rebroadcast the message after a default contention window.
(3) After a probability check, no matter the node retransmits the message or not, P_c is reset to a default value P_d.

Note that P_c was provided depending on different broadcast schemes, which will be discarded every probability check. P_d is the default broadcast probability, which enables every node to have P_d even if P_c was not set.

To align the behaviour of OBUs to the proposed scheme, this paper proposes the following scheme for OBU to deal with broadcasts from RSU:

(1) When an OBU receives a broadcast message from RSU, it checks the message whether it is a duplicate message or not.
(2) If it is not duplicated, OBU uses either signal strength or location services to get its distance to the corresponding RSU, which is denoted as R_{ro}.
(3) If R_{ro} < broadcast suppression range R_s, set the current rebroadcast probability to suppression minimum P_m.
(4) Proceed to probability check.

Considering the broadcast burst from RSU is likely to cover more area with stronger signal strength, the proposed scheme utilizes the scheme mentioned above to ensure that broadcasts from RSU suppress other broadcast within the radius of R_s to suppression minimum value P_m.

Although this paper focused on using V2I to maximum the effectiveness of the broadcast, it is imperative to keep in mind that V2V is still doing a heavy lifting in the broadcast propagating process. Therefore, this paper proposes a dynamic scheme for V2V communication, which is backed by percolation study.

The scheme we proposed for V2V communication could work independently, even without the support of RSUs, this is because in large scale network it is not appropriate to assume there is a reasonable amount of RSUs installed in each area, thus we developed this self-organized scheme for V2V for better scalability.

Wireless ad hoc network could be simplified into a model that consists of infinite nodes generated by a Poisson process using certain density. For example, in flooding every node is connected to the other node only if it's within the transmission range, this is because in broadcasting the only connection is to transmit the broadcast message, if the receivers successfully received a message and rebroadcast is then it's called a connection. So for a constant transmission range and constant area, only node density could affect the chance of forming a connected graph. With the density increased, the chance of all the nodes forming a connected graph is also increased. Continuum percolation is a study that dedicates to analyse this model [10].

On the other hand, if a node does not send any message during a broadcast (does not connect to any other node), it can be treated as if it was not there. In Figure 4 we demonstrate two scenarios when node S broadcast the message to node D through relay node I, the scenario 1) is equivalent to scenario 2) because nodes that did not rebroadcast the message will not add new connection. In percolation study we only consider nodes with connections so those nodes are negligible. Thus setting a rebroadcast probability to all nodes basically reduce the active node density, which makes the network an analogue of continuum percolation.

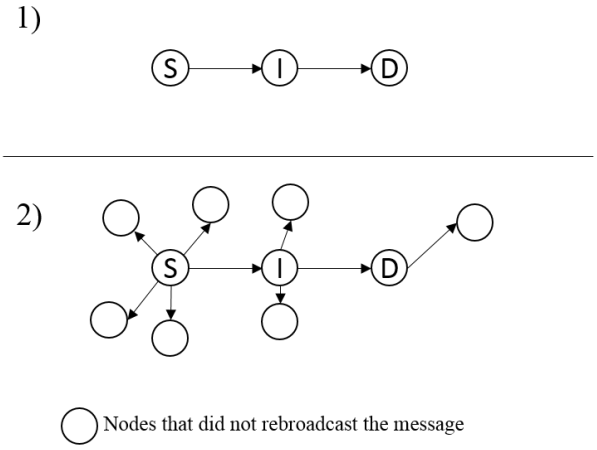


Fig. 4. Nodes did not rebroadcast messages is negligible in wireless broadcast

If nodes are distributed in a space by density ρ, we could denote an expected neighbourhood of a certain node as λ is calculated using following Eq.(1) [11]:

$$\lambda = \rho\pi r^2 \quad (1)$$

where r denotes a node transmission range. With higher λ, it is more likely that infinite nodes are in one connected cluster as is shown in Figure 5. The author of [8] used λ_c=4.508 to 4.515 and verified this value to have 99.99% probability to form one connected cluster, which means no minors such as shown in Figure 5. In this paper, we try to align the proposed scheme to match this value in order to maximize the broadcast effectiveness.

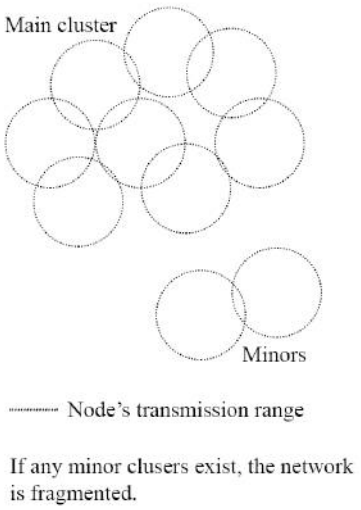


Fig. 5. Percolation

As is explained previously, if a node does not participate rebroadcasting in the broadcast process, the node is considered to be not in this network [8]. In stochastic broadcast, we have Eq.(2):

$$\lambda_s = P\rho\pi r^2 \quad (2)$$

where P denotes the rebroadcast probability. For example, if λ=10 and the rebroadcast probability for every node is 50%, λ_s=5 is good enough to ensure the network to be connected, that is, the broadcast can be propagated. As explain in continuum percolation, please note that the ideal situation is λ_s=4.5, which means rebroadcast probability should be roughly 45%.

In order to maintain λ_s, it is necessary to adjust P dynamically while the network topology changes. Thus, every node needs to maintain a neighbour node list using either beacons or location services. In this paper, we assume every node maintains a neighbour list using a simple one-hop beacon and total neighbourhood count to be N.

The V2V communication scheme is explained as follows:

- Once the node receives a non-duplicate broadcast message from another vehicle, set P_c using the following Eq.(3):

$$P_c = \begin{cases} 0 & (N = 0) \\ 1 & (0 < N < 5) \\ \dfrac{4.5}{N} & (N \geq 5) \end{cases} \quad (3)$$

(1) Reset the neighbour list.
(2) Proceed to probability check.
(3) Reset P_c to P_d

In this way, P_c could be aligned to a reasonable value using N. Since N is a dynamic value and constantly changing, P_c is theoretically always an optimal value.

V. Performance Evaluation

To evaluate the performance of the proposed scheme, a simulation was conducted in two different scenarios using different sets of parameters. In this simulation, static stochastic broadcast scheme, pure flooding scheme and the proposed is evaluated in both city and highway environment. General parameter used in this simulation is listed in Table I.

Please note that is this simulation we intended to the followings:

- Suppress the further broadcast near the RSUs within 80% of its transmission range.
- Suppressed rebroadcast probability is not set to 0 because if the node density is high, there might still be rebroadcast to cover the delivery failure caused by collisions.
- We set the RSUs with stronger presence by extending the communication range to simulate further transmission range.
- Replace actual nodes as RSUs for propose scheme. This is to avoid the performance boost brought by addition nodes (RSUs).

In the simulation, the performance of a broadcast scheme was evaluated using three metrics: message delivery rate, network delay and number of broadcasted packets.

Message delivery rate is the basic performance evaluation, which indicates if the message was successfully delivered to all nodes. Network delay basically shows how long it takes for the message to propagate to all possible nodes and it evaluates the effectiveness of each rebroadcast. As message travels along many hops, the delay increases. Number of broadcasted packets indicates how many broadcast packets are used in the broadcast process, it evaluates the network usage and potential broadcast storm problems.

Since pure flooding rebroadcast every message it receives, it can achieve the theoretically highest deliver rate and generate the most packets and create the infamous broadcast storm problem, we only simulation this scheme to compare it with the other broadcast schemes so please note that it is not for practical usage.

As shown in Table II, the parameters used in the first simulation aim to recreate a broadcast on highway scenario with the aid of RSUs installed along the road. The result was shown in Figures 5-7.

TABLE I
GENERAL STOCHASTIC BROADCAST SIMULATION PARAMETERS

Parameter	Value
R_s	0.8×RSU transmission range
P_m	0.05
P_d	0.2
Communication range V2V	1 km
RSU burst	3 times without contention window
RSU burst range	1.5 km
Beacon contention	≤1ms
Rebroadcast contention	≤2ms

TABLE II
SIMULATION 1 PARAMETERS

Parameter	Value
Area	20km × 0.1km
Number of RSU	5
RSU distribution	Uniform, installed along the road
Mobility	Random waypoint
Speed	50km/h to 80 km/h
Number of nodes	15 to 45 (10 to 40 for RSU-Assisted broadcast)

From Figures 6 and 7, the proposed broadcast scheme achieved better deliver rate and lower delay at the same time in high node density due to higher RSU's coverage. Since deliver rate has been increased, number of broadcasted packets will increase accordingly in order to cover more nodes. In Figure 8, we can see that number of broadcasted packets is increasing but please note there is an overhead cause by proposed method.

As shown in Table III, the parameters we used in the next simulation are about recreating a broadcast in city scenario where RSUs are randomly installed in the city and cars moving to random way point.

In city scenario the proposed broadcast scheme got a performance boost in the sense of delay and packet usage due to the sparse distributed RSUs comparing to highway scenario. From Figures 9 and 10 the proposed broadcast scheme achieved higher deliver rate with almost the same delay comparing with static stochastic broadcast. If comparing Figure 11 with Figure 8, it is obvious that in city scenario packet usage also decreased. Achieving high deliver rate with lower delay and lower packet usage means every broadcast covers more nodes, which is our intention to mitigate broadcast storm problem and boost the effectiveness of the broadcast scheme.

TABLE III
SIMULATION 2 PARAMETERS

Parameter	Value
Area	9km × 9km
Number of RSU	5
RSU distribution	Random
Mobility	Random waypoint
Speed	10km/h to 30 km/h
Node count	15 to 45 (10 to 40 for RSU-Assisted broadcast)

Since RSUs are distributed randomly in city scenario, message deliver rate dropped a little comparing to the highway scenario from Figure 9 and Figure 6. In general, the proposed broadcast scheme outperforms the static stochastic broadcast and went toward the upper bound, without bring heavy traffic to the network and cause broadcast storm problem (comparing to flooding). Also, there's no need to acquire the network condition in order to calculate the broadcast probability before the broadcast (comparing to static stochastic broadcast).

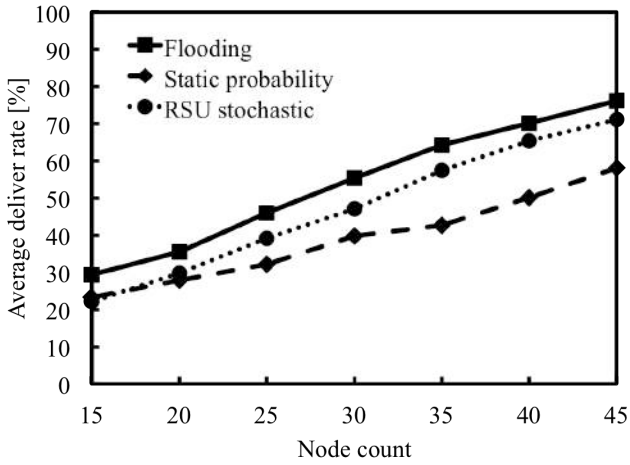


Fig. 6. Message deliver rate in highway simulation scenario

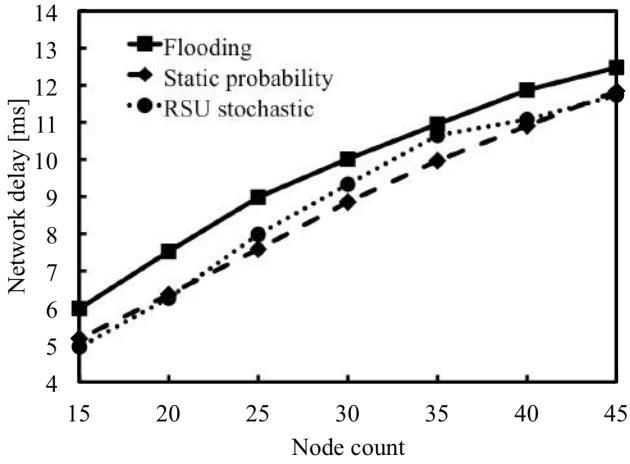


Fig. 7. Delay in highway simulation scenario

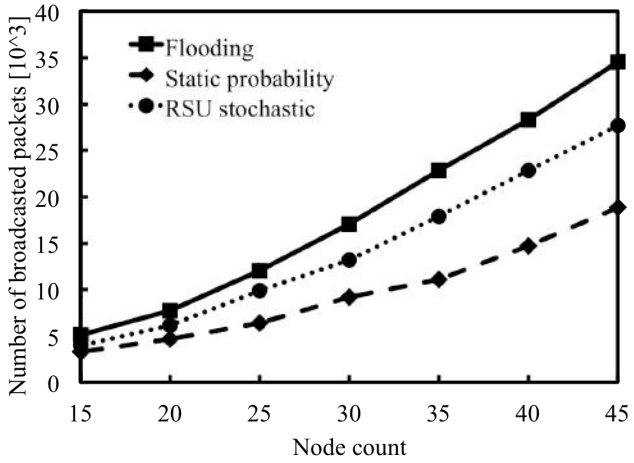


Fig. 8. Packet usage in highway simulation scenario

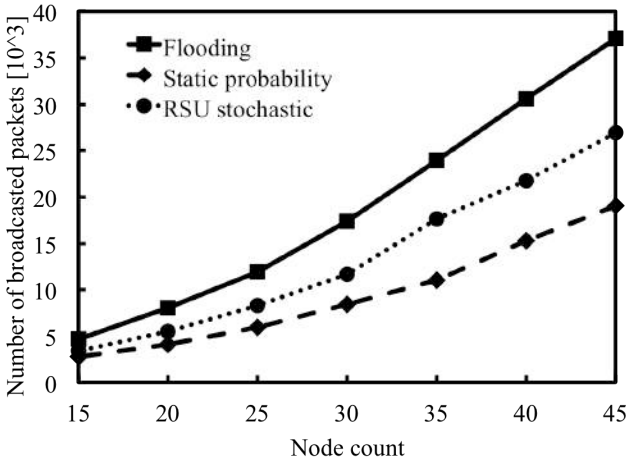


Fig. 11. Packet usage comparison in city simulation scenario

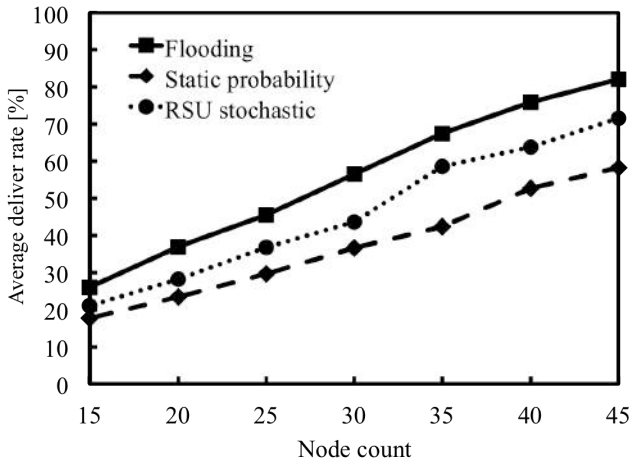


Fig. 9. Message deliver rate comparison in city simulation scenario

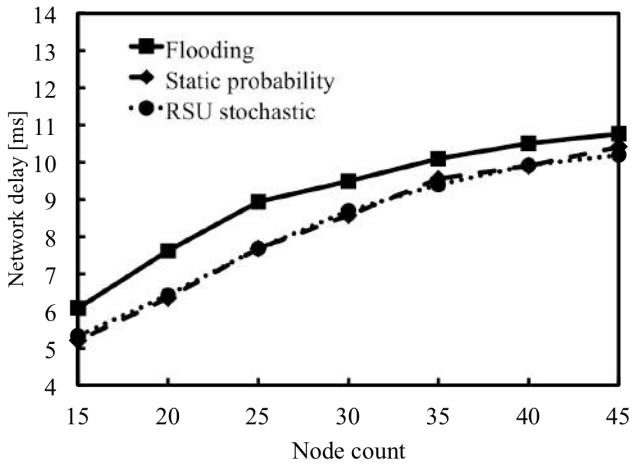


Fig. 10. Delay comparison in city simulation scenario

VI. Conclusion

This paper introduced the challenges in VANET emergency message broadcast and its contribution of protecting human lives and properties. It also summarized our objectives to contribute to emergency message broadcast in VANET research. This paper also categorized different broadcast schemes, analysed their strength and weakness. On top of that, this paper introduced two broadcast schemes that related to the study.

In the discussion of exclusive characteristics of VANET, this paper analysed some major concerns in VANET such as privacy, safety and scalability. In order to boost performance and mitigate network fragmentation, we proposed the usage of RSUs to assist to broadcasting process, explained the scheme for RSUs to work actively and individually in the broadcast.

By simplify VANET model using percolation theory, this paper developed another scheme to be used in V2V broadcasting. In this discussion, this paper explained percolation study and how to apply it to stochastic broadcast instead of flooding broadcast along with how to achieve the theoretically highest performance using stochastic model and apply the result in the proposed schemes.

In order to study the performance of the proposed scheme, this paper conducted two simulations. From the result, the proposed scheme achieves our goals by increasing deliver rate and does not bring huge impact on delay and network usage, especially in city scenario.

The proposed scheme take privacy and scalability into consideration by not using beacons to send personal data, not using location services and every part of the scheme works independently (V2V scheme will work without RSUs, RSUs do not form a connected graph and they are not connected to an infrastructure-depended network), which guarantees scalability even being deployed in large scale network.

The amount and position of installed RSUs may have impact on the performance. In the simulation we replace the nodes with RSUs when RSUs are needed, so that they will not benefit from higher actual node density. Moreover, we avoid assuming connections to location services but we are still using beacons, this is not absolutely the ideal solution. We still have some future works such as evaluating the

performance impact of RSUs and finding alternative, safer way to calculation node density without using beacons.

REFERENCES

[1] M. L. Sichitiu; M. Kihl, "Inter-vehicle communication systems: a survey," *IEEE Commun. Surveys & Tutorials*, vol. 10, no. 2, pp. 88-105, Second Quarter 2008.

[2] Bureau of Transportation Statistics, "National transportation statistics," *US Department of Transportation* [Online]. Available: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_statistics/index.html, Accessed Oct. 2013.

[3] B. Williams and T. Camp. "Comparison of broadcasting techniques for mobile ad hoc networks." *3rd ACM Int. Symp. Mobile Ad Hoc Netw. & Comput*., pp. 194-205, June 2002.

[4] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, and J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network," in *5th annual ACM/IEEE Int. Conf. Mobile Comput. and Netw*. pp. 151-162. Aug. 1999.

[5] D. Jiang, V. Taliwal, A. Meier, W. Holfelder, and R. Herrtwich, "Design of 5.9 GHz DSRC-based vehicular safety communication", *IEEE Trans. Wireless Communication*, vol. 13, no. 5, pp. 36-43, Oct. 2006.

[6] X. Ma and X. Chen, "Performance analysis of IEEE 802.11 broadcast scheme in ad hoc wireless LANs," *IEEE Trans. Vehicluar Technology*, vol. 57, no. 6, pp. 3757-3767, Nov. 2008.

[7] E. Fasolo; A. Zanella, and M. Zorzi, "An effective broadcast scheme for alert message propagation in vehicular ad hoc networks," in *2006 IEEE International Conference on Communications*, vol.9, pp. 3960-3965. June 2006.

[8] M. Slavik and I. Mahgoub, "Stochastic broadcast for VANET," in *7th IEEE Consumer Commun. and Netw. Conf.*, pp. 1-5. Jan. 2010.

[9] N. Wisitpongphan, F. Bai, P. Mudalige, V. Sadekar, and O. K. Tonguz, "Routing in sparse vehicular ad hoc wireless networks," *IEEE J. Sel.Areas Commun*., vol. 25, no. 8, pp. 1538-1556, Oct. 2007.

[10] B. Bollobas and O. Riordan, *Percolation*, Cambridge Univ. Press, 2006.

[11] P. Balister, B. Bollobas, and M. Walters, "Continuum percolation with steps in the square or the disc," *Random Struct. Alg.*, , vol. 26, no. 4, pp. 392-403, April 2005.

Xing Fan was born in Shanxi, China on June 30th. Xing Fan earned bachelor of engineering in Beijing University of Technology in July 2012. Xing Fan is a master student majored in ad hoc network at Waseda University, Japan while presenting this paper.

Bo Yang received his B. E. degree in computer science and technology from Xi Dian University, Xi'an, China, in 2009. He received his second M.E. degree in Information Communication from Waseda University, Tokyo, Japan, in 2012. Currently, he is working toward the Ph. D degree in the Global Information and Telecommunication Studies, Waseda University, Tokyo, Japan. He won the ICACT best paper award in Feb. 2012.

Ryo Yamamoto received his B.E. and M.E. degree in electronic information systems from Shibaura Institute of Technology, Tokyo, Japan, in 2007 and 2009. He received D.S. in global telecommunication studies from Waseda University, Tokyo, Japan, in 2013. He was a research associate at Graduate School of Global Information and Telecommunication Studies, Waseda University, from 2010 to 2014, and has been engaged in researching in wireless communication networks. He is presently an assistant professor at Graduate School of Information Systems, The University of Electro-Communications. He received the IEICE young researcher's award in 2010, the IEICE Network System Research Award in 2014. His current research interests are mobile ad hoc networks and cross-layered protocols.

Yoshiaki Tanaka received the B.E., M.E., and D.E. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1974, 1976, and 1979, respectively. He became a staff at Department of Electrical Engineering, the University of Tokyo, in 1979, and has been engaged in teaching and researching in the fields of telecommunication networks, switching systems, and network security. He was a guest professor at Department of Communication Systems, Lund Institute of Technology, Sweden, from 1986 to 1987. He was also a visiting researcher at Institute for Posts and Telecommunications Policy, from 1988 to 1991, and at Institute for Monetary and Economic Studies, Bank of Japan, from 1994 to 1996. He is presently a professor at Department of Communications and Computer Engineering, Waseda University, and a visiting professor at National Institute of Informatics. He received the IEEE Outstanding Student Award in 1977, the Niwa Memorial Prize in 1980, the IEICE Achievement Award in 1980, the Okawa Publication Prize in 1994, the TAF Telecom System Technology Award in 1995 and in 2006, the IEICE Information Network Research Award in 1996, in 2001, in 2004, and in 2006, the IEICE Communications Society Activity Testimonial in 1997 and in 1998, the IEICE Switching System Research Award in 2001, the IEICE Best Paper Award in 2005, the IEICE Network System Research Award in 2006, in 2008, and in 2011, the IEICE Communications Society Activity Award in 2008, the Commendation by Minister for Internal Affairs and Communications in 2009, the APNOMS Best Paper Award in 2009 and in 2012, and the IEICE Distinguished Achievement and Contributions Award in 2013. He is a Fellow of IEICE.

Greening Potential Estimation of Data Network Equipment

Yuhwa Suh*, Kiyoung Kim**, Jongseok Choi*, Yongtae Shin*

*Department of Computer Science and Engineering, Soongsil Universiety, Seoul, South Korea

** Department of Computer Software, Seoil University, Seoul, South Korea

zzarara@ssu.ac.kr, ganet89@seoil.ac.kr, jschoi@ssu.ac.kr, shin@ssu.ac.kr

***Abstract*—Green Internet is becoming a major concern recently. The main concept of it is to improve energy efficiency in the Internet for reduction of unnecessary energy consumption. Generally, energy consumption of data network equipment in the Internet is unknown although they use a substantial amount of energy. In this area, there is a lack of deeper related studies with a special focus on wired networking and it still remains to be many challenges. This paper aims at exploring an impact of data network equipment for greening the Internet. We first introduce backgrounds and motivations of green networking. Secondly, we estimate energy consumption, costs, and energy savings potential of data network equipment in detail. Thirdly, we assess impacts of it based on IP traffic type and propose new viewpoint on energy efficiency focused on the quality of service (QoS). Lastly, we propose the future works for green networking from the perspective of QoS.**

***Index Terms*—Energy efficiency, green networking, QoS, data network equipment**

I. Introduction

TODAY, a huge numbers of electronics are used and most of them have network connectivity. As setting up ubiquitous computing environments ubiquitous various devices tend to embed computing and networking functionalities. Thus the Internet has been expended quickly with growing greenhouse gas (GHG) emissions and electrical requirements from it. However, these GHG emissions and energy consumption are becoming key factors of limiting expansion of the Internet.

Manuscript received May 15, 2014. This research was supported by the MSIP (Ministry of Science, ICT & Future Planning), South Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2014-H0301-14-1010) supervised by the NIPA (National IT Industry Promotion Agency).

Yuhwa Suh is with the Department of Computer Science and Engineering, Soongsil University, 369 Sangdo-Ro, Dongjak-Gu, Seoul 156-743, South Korea (phone: +82-2-826-0690; e-mail: zzarara@ssu.ac.kr).

Kiyoung Kim is with the Department of Computer Software, Seoil University, 28 Yongmasan-ro 90-gil, Jungnarg-gu, Seoul 131-702, South Korea (e-mail: ganet89@seoil.ac.kr).

Jongseok Choi is with the Department of Computer Science and Engineering, Soongsil University, 369 Sangdo-Ro, Dongjak-Gu, Seoul 156-743, South Korea (e-mail: jschoi@ssu.ac.kr).

Yongtae Shin is with the Department of Computer Science and Engineering, Soongsil University, 369 Sangdo-Ro, Dongjak-Gu, Seoul 156-743, South Korea (e-mail: shin@ssu.ac.kr).

To overcome these problems, greening of the Internet has begun to be studied with the goal of improving energy efficiency of the current Internet. There are various approaches found on energy-saving mechanisms and power management criteria for greening the Internet, but they have mainly focused on edge devices such as end-user computers and data-centers. However, the energy consumption associated with network functionality is a big issue that we need to focus on, due to growing IP traffic from excess network-connectivity. This issue usually referred to as ***green networking*** [1], which energy managing mechanisms exploiting network-specific features like ad-hoc network have to be added to the wired data network for boosting the network energy efficiency.

The use of the Internet is a part of daily life, but a volume of energy used for providing network connectivity is largely unknown. In particular, there are few reliable and practical figures of greening potential of wired data network equipment for greening the Internet.

The Internet will eventually be constrained by energy density limitations rather than by the bandwidth of the physical components because of its high dependency on electronics. So energy efficiency improvement of the current Internet is a critical issue. But, a subtle trade-off between energy saving and performance exists because energy efficiency must be treated together with increasingly diversified demands of services and sophisticated the quality-of-service (QoS) support.

This paper aims at estimating energy consumption and savings potential in the today's wired data network equipment, understanding impacts of them and providing comprehensive view on green networking. In particular, we are devoted to enterprise network equipment. The paper is organized as follows. Section II states detailed background of the concerns and reasonability of study on greening networking. Section III includes an estimate of energy consumption, costs, and savings potential for wired data network equipment that primarily switches and routes IP packet from a source to a destination. Section IV further gives estimates of impacts of network equipment depending on traffic type in an aspect of energy efficiency considering for QoS and presents directions for future works related to our estimations. Finally, the conclusions are drawn in section V.

II. Procedure for Paper Submission

Today, the reduction of energy consumption is the world's major concern. Residential and commercial buildings account for approximately 32% of global energy use and almost 10% of total direct energy-related CO_2 emissions. Energy demands from the buildings sector will be more than double by 2050 [2]. Much of this growth is fuelled by the rising number of residential and commercial buildings in response to the expanding global population.

In the building sectors, the huge potential of improvements in energy efficiency remains untapped, and it will have the largest impact on energy savings and CO_2 emission mitigation through energy-efficiency technologies. By the International Energy Agency (IEA) 2DS (2°C scenario) [2], which assumed policy action consistent with limiting the long-term global temperature increase to 2°C, buildings sectors can contribute to 2DS objectives about 18% of CO_2 reduction shares by 2020. The Information and Communication (ICT) devices occupy over 50% [3] of the electronics used in buildings, and the amount of greenhouse gas (GHG) emissions from their energy use is significant as in Fig. 1. In 2005, CO_2 emissions produced in the ICT sector amount to 98.3MtCO_2e (carbon dioxide equivalent) or 1.9% share of the total EU-25 CO_2 emissions and these are projected to double by 2020 as each 187.7MtCO_2e or 4.5% in BAU (Business As Usual) scenario assuming no significant efforts to reduce emissions. However, in ECO-scenario adopting energy-saving solutions, CO_2 emissions are expected to decline to 132.1MtCO_2e or about 3%, this is the amount decreased by about 30% of BAU-scenario.

Fig. 2 shows GHG emissions footprint of networks from 2002 to 2020. In 2011, this amounted to 22 % of the total ICT footprint and represents 5% compound annual growth rate (CAGR) by 2020. Overall, wired networks' emissions are estimated to grow at a 4% CAGR from 2011 to 2020 to reach 0.14 GtCO_2e. The GHG emission footprint from wired network equipment is to amount to approximately 11 % of the total ICT footprint and approximately 0.25% of total GHG emissions [5]. This is a non-neglected level, so the current wired network infrastructure needs to challenge for energy efficiency.

IPCC 2007 climate change report [6] presents top-down studies that indicate substantial economic potential for the mitigation of global GHG emissions over the coming decades. As shown in TABLE I, in 2030, global economic mitigation

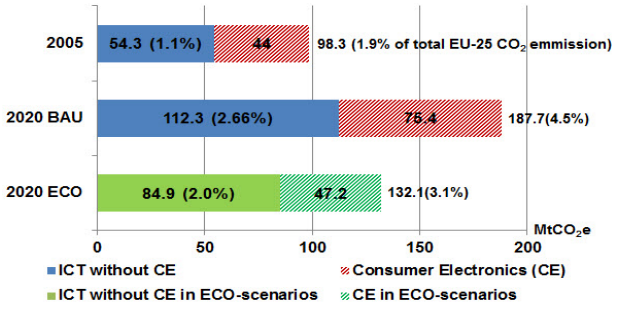


Fig. 1. GHG emission and savings potential (EU-25) of ICT : source [4]

Fig. 2. Global mobile and wired networks emissions [5]

potential estimated and assessed the economy-wide potential of mitigation options. TABLE I assumed two scenarios, SRES A1B and SRES B2 (SRES: IPCC Special Report on Emissions Scenarios). They assumed that GHG emissions in SRES A1B are about 68 GtCO_2e and them in SRES B2 is about 49 GtCO_2e in 2030. They are measured economic potential to US$ by different GHG mitigation rates. In TABLE I, on assumption that GHG emission is 68GtCO_2e/year and GHG mitigation rate is 13~27% (9~18 GtCO_2e/year) in 2030, it is possible to save 20 US$/t$CO_2$e.

With reference to [5][6], we estimated economic potential of GHG emission mitigation of network-connected equipment to US$ from 2020 to 2030. Assuming 2% CAGR in BAU, GHG emissions of global network-connected equipment are expected to emit about 1.55 GtCO_2e in 2030. According to [5], if network-connected equipment is introduced to other sectors, global GHG abatement potential enabled by these technologies is between five and seven times larger than network-connected equipment's own footprint. Therefore, global GHG emissions is likely to reduce annually about 7.8~10.9 GtCO_2e, using TABLE I, we can estimate economic savings potential at approximately 156~218 billion US$. The energy consumption of network-connected equipment in the USA reflects the world's trend because the USA accounts for the largest amount of the world's (40% of world, in 2008 [7]).

Appliances take a great share of building electricity use, and are also of growing importance in the service sector in the form of office equipment. Globally, network connectivity is being added to appliances which would not previously have had such functionality, and demand for the availability of traditionally network-connected equipment is increasing.

In the USA which occupies the largest proportion of global electricity consumption (27%, 2008 [3]), as shown in Fig. 3, buildings accounts for over 70% of the total electricity consumption in 2006, and electronics account for about 11% of

TABLE I Global Economic Mitigation Potential in 2030 Estimated from Top-Down Studies [6]

Carbon price (US$/t$CO_2$e)	Economic potential (GtCO_2e/yr)	Reduction relative to SRES A1B (68GtCO_2/yr)%	Reduction relative to SRES B2 (49GtCO_2e/yr)%
20	9-18	13-27	18-37
50	14-23	21-34	29-47
100	17-26	25-38	35-53

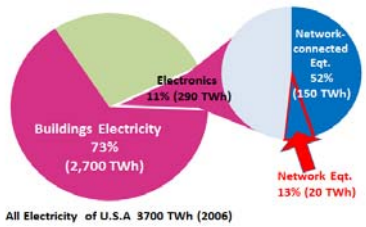
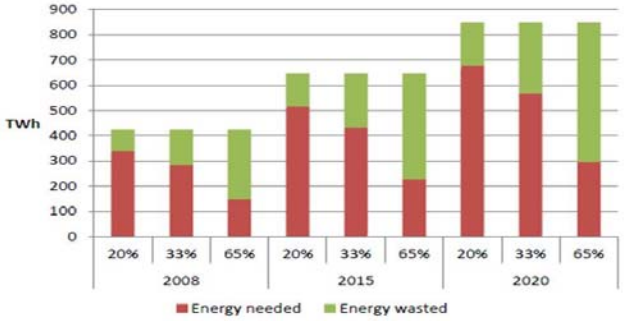


Fig. 3. Electricity consumption of the USA (2006), source : [3]

building electricity use. Electricity consumption of network-connected equipment forms about 52% of electronics and about 0.6 ~ 0.7% of the total buildings'.

Reference [7] presents global energy consumption of network-connected equipment in buildings (e.g., PC, Printer, Phones, TV, game console, audio receiver, media player, STB, router, gateway etc.) and savings potential of them. This is estimated by the degree of market penetration of energy efficiency technologies (e.g. 20.08 ~ 65%) as shown in Fig. 4.

It is estimated that in 2008, total energy consumption by network-connected equipment was 423.85TWh and it accounts for 2.3% of the global electricity consumption (18,603TWh). Based on market trends, this is projected to increase to 646TWh in 2015 and 849TWh in 2020-double the 2008 level. The green parts in Fig. 4 implies that how much energy saving is possible by energy efficient technologies. The lower- end estimate of wasted energy is around 20%, as a result of excess connectivity and/or the use of sub-optimal technologies instead of cost-effective improved technology. This 20% of energy could be saved by means of implementation of power management and power-level reduction policies. The maximum estimate (technical potential) is around 65% of energy, assuming all network-connected products and components are 1W power state lower, and implementation of effective power management policies are on them. The amount of energy wasted by excessive connectivity is estimated between 85TWh and 275TWh in 2008, rising to between 130TWh and 420TWh in 2015, and between 170TWh and 551TWh in 2020, an



	Energy Consumption (TWh)	Energy Wasted (TWh)		
		20.08% Savings potential	33% Savings potential	65% Savings potential
2008	423.85	85.11	140.29	275.36
2015	646	129.78	213.92	419.90
2020	849	170.40	280.87	551.31

Fig. 4. Projected energy consumption and wasted by network-connected equipment worldwide, source : [7]

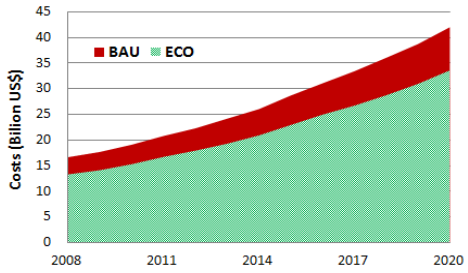


Fig. 5. The USA energy costs for network-connected equipment

amount that is superior to the entire electricity consumption of all network-connected equipment in the USA in 2008 (about 169TWh).

We made an effort to look at a scale of expenditure on energy use of network-connectivity equipment in BAU and ECO scenarios. With reference to [7][22], we presented roughly an estimate of the costs of network-connected equipment in the USA between 2008 and 2020 (Fig. 5). We supposed that the energy consumption of network-connected equipment grows annually at 7.3% from 2008 to 2015 and at 6.3% from 2016 to 2020 the same as [7] and ECO scenario has 20% energy savings potential (as the lower limit) over BAU. As a result, it is possible to save 3~9 billion US$ every year from and accumulated saving is about 71 billion US$.

This energy consumption is caused by the absence of effective power management of network-connected equipment. Greater quantities of new network-connected products are entering the market and products are spending more time in higher power modes due to network-related requirements and a lack of effective power management strategies. Power consumption in lower power modes is also increasing because network interfaces require more power in standby modes in order to maintain a network link and networks are tending towards faster speeds and higher bandwidth, which increases power in the absence of effective power management. So, it is important to consider energy efficiency related to networking functionality in electronics. Moreover, it is useful to distinguish between edge devices and network equipment (where the main function is to maintain network links) because their function and energy-saving potential are quite different. So far, edge devices have become an object of attention in terms of energy efficiency and many related researches have carried out, but network equipment has not. In this paper, we therefore discuss relatively unknown energy consumption and savings potential by wired data network equipment.

III. Energy Consumption and Savings Potential for Data Network Equipment

In this paper, we specifically deal with the energy consumption and the implication of wired data network equipment whose primary purpose is to transport, route, switch, or process network traffic, excluding edge devices such as PCs, servers, other sources and sinks of IP traffic.

The reasons that we should note to energy conservation of the wired network equipment are the followings. Firstly, there is inefficiency of the current networking system. Networking infrastructure involves high-performance and high-availability machines. They therefore relay on powerful devices, which are organized in an over-provisioned and redundant architecture [1]. Traditionally, networking system are designed to endure peak load and degraded conditions, they are dimensioned with extra capacity to allow for unexpected events. As a result, during low traffic periods, over-provisioned networks are also over-energy-consuming. For resiliency and fault-tolerance, networks are also designed in a redundant manner. Devices are added to the infrastructure with the sole purpose of taking over the duty when another device fails, which further add to the overall energy consumption however these objectives are under-utilized in normal operation [1]. Moreover, network equipment expends a great deal of energy even when it's at idle (they are powered on 24/7). Unlike monitors or other computing equipment that satisfy Energy Star recommendations by going into various energy saving states when it's at idle, network equipment typically does not (there are no Energy Star recommendations for them). This is because maximizing network throughput and minimizing latency are the primary driving factors in network design [8], so network areas leave a large room for energy savings.

Secondly, to support new generation network infrastructures and related services for a rapidly growing customer population, telcos and ISPs need an ever larger number of devices, with sophisticated architectures able to perform increasingly complex operations in a scalable way. According to [8], high-end IP routers, which provide more and more network functionalities, continue to increase their capacities, with an increase factor of 2.5 every 18 months. At the same time, silicon technologies The sole introduction of novel low consumption silicon technologies cannot clearly cope with such trends, and be enough to draw current network equipment towards a greener future Internet. Thus, there is much likely as in other areas where energy efficiency is a concern, it is required more approaches to networking area for greening the Internet. Today's network relies very strongly on electronics, despite the great progresses of optics are in transmission and switching. Therefore, how energy consumption of the network equipment is a key factor of growing importance and it will become more and more significant to consider energy efficiency in terms of networking.

Thirdly, in many parts of the world, electricity is a scarce resource and poses one of the barriers to widespread Internet deployment. In addition, frequent power outages reduce the uptime of the deployed Internet. If energy consumption of the Internet devices is reduced, we can deploy more devices for the same energy cost and, given the same UPS capacity, have more of them up and running during periods of power outage thus improving overall network reliability [9].

Currently, the major solution of the paradigms for green networking is proportional computing. Many green network approaches follow this solution. It depicts different energy consumption profiles by its utilization level as shown in Fig 6. Energy-agnostic devices, whose energy consumption is constant, independently of their utilization, represent the worst case: such devices are either on and consume the maximum amount of energy, or off and inoperative. In contrast, fully energy-aware devices exhibit energy consumption figures proportional to their utilization level. Between these two extreme situations, there exist an infinite number of possible intermediate profiles. Today's many green networking techniques have been committed to make energy consumption figures proportional against their utilization level. This strategy can be introduced into individual devices or components, a networked system and protocols.

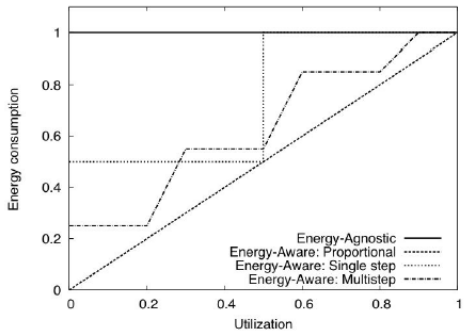


Fig. 6. Energy consumption as a function of the utilization [10]

According to [8], the typical access, metro and core device density and energy requirements in today's typical networks. The power consumption of transport and core network represents about 30% of the overall network requirement, and access devices weigh for 70%. Even if power consumption of an access device is approximately one-sixth of a core device, the number of access devices is much more than core devices. Thus power consumption of the today's overall network is driven by access networks.

To look at a scale of local and global energy consumption of access network equipment, we estimated it by (1).). $\boldsymbol{E_{g\_ne}}$ is the global energy consumption of network equipment in BAU (assuming that continuity is maintained considering the current situation and trends) and ECO scenarios (assuming that there is a push for ICT-based energy efficient solutions) and it is sum of $\boldsymbol{E_{ne_i}}$ which is each country's total energy consumption of network equipment $\boldsymbol{E_{ne_i}}$ can be calculated with (2). $\boldsymbol{E_i}$ is a country's total consumption of electricity [11], $\boldsymbol{B_i}$ is the share of buildings in the total electricity of a country [12], $\boldsymbol{N}$ is the share of network equipment in the buildings electricity, which is assumed 0.7% based on the USA [3]. $\boldsymbol{n}$ in (1) is the number of countries included and its value is 136.

$$E_{g\_ne} = \sum_{i=1}^{n} E_{ne_i} \quad (1)$$

$$E_{ne_i} = E_i \times B_i \times N \times w_i \quad (2)$$

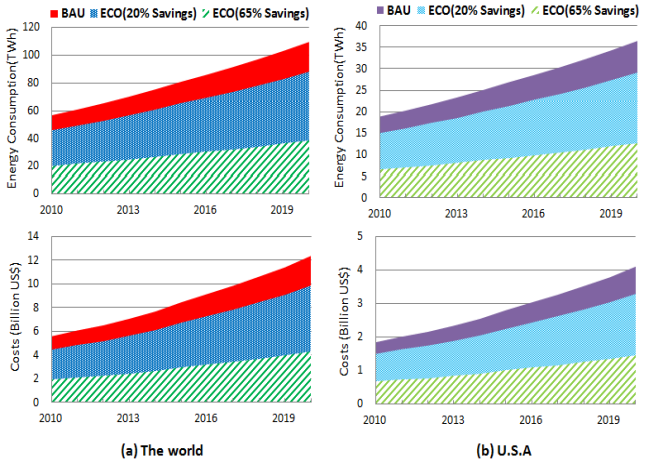


Fig. 7. The energy consumption and costs by access network equipment

$$w_i = I_{sb_i} \times \frac{1}{I_{sb_{us}}} \tag{3}$$

$\boldsymbol{w_i}$ is weight, which the value is used for taking into account levels of IT development varied by country. This weight is determined by the IDI access sub-index in 2010. IDI (The ICT Development Index) is one benchmark measure that serves to monitor and compare developments in information and communication technology (ICT) across countries. It was developed by the International Telecommunication Union (ITU) and aims to capture the evolution of the information society as it goes through its different stages of development, taking into consideration technology convergence and the emergence of new technologies [13]. $\boldsymbol{I_{sb}}$ access sub-index of the IDI captures ICT readiness, and includes infrastructure and access indicators. We use $\boldsymbol{I_{sb_i}}$ the IDI access sub-index of a country adjusted by $\boldsymbol{I_{sb_{us}}}$, the IDI access sub-index of the USA, 5.75 in 2010, which the USA accounts for the largest amount of the world's electricity use.

Therefore the energy consumption of network equipment is :

$$E_{g\_ne} = \sum_{i=1}^{n} \left(E_i \times B_i \times N \times I_{sb_i} \times \frac{1}{I_{sb_{us}}} \right) \tag{4}$$

The results are shown in Fig. 7. According to that, the global energy consumption by access network equipment is around 57TWh in 2010. Approximately 9TWh costs in the order of 0.9 billion dollars per year (at 9.8 cents per kWh from all sector average in 2010 [23]) and requires one nuclear reactor unit. A typical average nuclear power plant have two reactors, each of which generates an average of 9TWh electricity per year [9] (an average life of a nuclear reactor is 40~60 years [24]). As shown in Fig. 7, in 2010, the USA access network equipment's energy consumption per year is an enormous amount that required more than two nuclear reactors. Moreover, to look at a scale of savings by energy efficient technologies just at a glance, we estimated local and global energy consumption of it in BAU and ECO scenarios between 2010 and 2020. In this work, as the growth of network equipment is correlated with that of the network-connected equipment, similar growth rates can reasonably be applied to network equipment in general. So, the yearly growth rate of the energy consumption of it is assumed to be the same those in Fig. 4. In here, we supposed that the energy consumption of access network equipment grows annually at 7.3% from 2008 to 2015 and at 6.3% from 2016 to 2020 and ECO scenarios are assumed to has 20% and 65% energy savings potential (as the lower and upper limit) over BAU.

As shown in Fig. 7, the 2010 energy consumption of the world's access network equipment accounted for 0.3% (56.6TWh) of the total world's electricity consumption (17,839TWh [11]) and it is forecasted to show growing trend continually. In 2020, it is forecasted to increase by about 109.7TWh and this doubles in 2010. If the energy consumption of the world's access network equipment was possible to save by 20% over BAU, it could save 11~22TWh per year from 2010 to 2020. In the USA, it is possible to save by 4~7TWh per year, then this is an amount needed almost one nuclear reactor.

Access network equipment mainly consists of office network switches and residential equipment. They account of approximately 0.7% (2006) [3] and 1% (2008) [14] of power use in buildings in the USA. In particular, the enterprise network equipment is largely responsible for about 60% (42% at switches, 6% at routers, 10% at security appliances and 2% at WLAN [14]) of energy use in the global access network. As a typical case, According to the energy use distribution in Berkeley campus LAN reported by LBNL [14], switches and routers are account for 87% of total energy use of campus LAN. So, switches would account for most of energy use in general enterprise network. This clearly implies why switches are targeted leadingly for energy efficiency.

IV. An Impact of Data Network Equipment Based on Traffic Types

We devoted to analyze energy consuming factors in the switching equipment which is the major cause (over 80%) of energy consumption in a general enterprise network. According to [15], typically, 60 % of power consumption in networking equipment is associated with packet-processing part and packet-processing support part such as memories

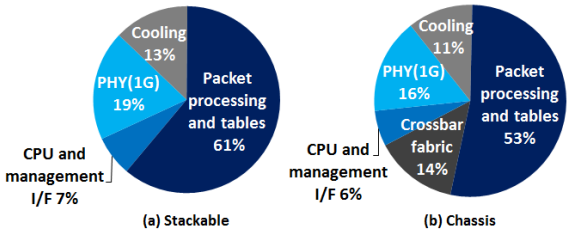


Fig. 8. Typical switch power allocation [15]

(dynamic random access memory (DRAM) and ternary content access memory (TCAM)). Fig. 8 presents the maximum switch power consumption happens at the packet processing subsystem. This subsystem processes vast traffic with functionality such as buffering, lookup, forwarding, switching fabric and I/O. To increase power savings, the best place to start is to reduce the power consumption associated with packet processing.

The traditional network equipment does not consume energy proportional to traffic loads, because they consume constant energy even during idle period. Total energy consumption of the current typical switches (non-green network switches) depends on the number of active ports, capacity of each port, but does not on port utilization. Current Ethernet switches require both transmitters and receivers to operate continuously on a link, thus consume energy all the time, regardless of the amount of data exchanged. That is to say, they consume large energy even at low loads, which means very low energy efficiency. As solution to this problem, green technologies are devoted to make its energy consumption nearly proportional to its traffic load by approaches such as resource consolidation and selective connectedness. Representatively, energy efficiency Ethernet (EEE) [25] of adaptive link rate (ALR) solutions is already well advanced and addressed by IEEE Std. 802.3az and has shown important energy savings at low loads. Network equipment in today's Internet is being rapidly replaced to high-end ones and EEE shows better energy efficiency over high speed link. With such the trend, energy consumption of switches would become proportional to its traffic load. However, green approaches of resource consolidation and selective connectedness are required more sophisticated mechanisms in terms of the quality-of-service (QoS). The current green networking techniques have hardly taken QoS into account.

In this context, that switches comprehend a volume and characteristic of traffic on them and manipulate appropriately could improve energy efficiency. Moreover, at the line consuming second-largest energy, it is important that switches consider a volume and characteristic of traffic to decide sensitive factors like states (sleeping or idle), timers in green techniques such as ALR. Further, the green networking techniques need to focus on the quality of the user experience. The current green networking techniques have hardly taken into account QoS but should consider energy efficiency with view of QoS in the future. Processing traffic by green levels classified on the basis of characteristic of traffic could further improve energy efficiency. For this, switches need to understand and manage information about traffic characteristic in an aspect of energy efficiency. This enables to consider further QoS, and thereby could provide more sophisticated energy efficiency. Also this information can be utilized for smart decisions and exact predictions in the resource consolidation and dynamic adaptation for green networking, either.

Fig. 9 shows global IP traffic growth from 2011 to 2016. In 2016, about 93% of the global IP traffic will be comprised of

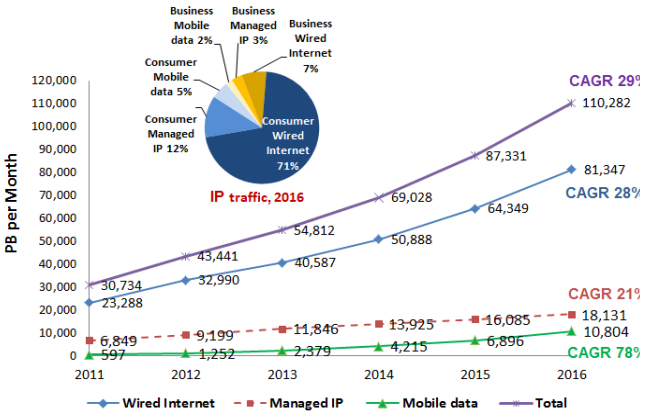


Fig. 9 Global IP traffic, 2011-2016, source : [16]

the consumer traffic which includes IP traffic generated by households, university population, and Internet cafes and the consumer wired Internet traffic of them will account for the majority (over 70%) of the IP traffic. Thus switching equipment will require tremendous energy for processing these huge amounts of IP traffic. Internet video, file sharing, and web/email data account for 99% of total. Especially, Internet video streaming and downloads of them are beginning to take a larger share of bandwidth and will grow to over 54% of all consumer Internet traffic in 2016 [16].

Depending on the traffic types, volumes of traffic have very big difference. Their characteristics and needs are different in terms of energy efficiency as well. Thus, we put effort to figure out impacts on network equipment by traffic types in an aspect of greening network. We focused on consumer the wired Internet traffic which is the majority of the Internet traffic. We classified based on traffic types and estimated their impacts from the viewpoint of greening.

In this work, we limited network equipment to enterprise switching equipment which account for the majority of energy consumption of access networks and used the forecast in 2016 in [16]. Also, with reference to a power benchmarking for network devices in [15][17][18], we supposed that a switch has a line card with 48 full-duplex 1Gbps ports and it consumes average 3.3mWatts/1Mbps when it transmits at fully load with the aggregate bandwidth of 48000Mbps.

In (5), E_{en} is the total enterprise switching equipment's energy consumption and Se is the share of it in the total access

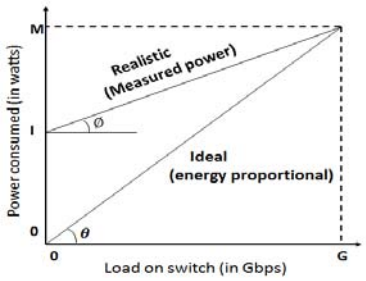


Fig. 10. Ideal and realistic power characteristic of network equipment

network equipment, therefore ***1 - Se - r*** is a share of the residential customer premises equipment (r is a share of routers, security appliances, WLAN and etc. of the enterprise network devices and it is about 15% in 2008 [14]). We assumed that S_e is assumed about 40% by section III. ***T*** is the amount of the traffic generated per hour and E_{MB} is the energy consumption per megabyte. ***p*** is the energy proportionality formed by Fig. 10.

$$E_{en} = T \times E_{MB} \times S_e \times \frac{1}{p} \tag{5}$$

$$= \left(\frac{MB\ per\ Month}{30\ days \times 24\ hours} \right) \times 0.003Wh \times 0.4 \times \frac{\tan\theta}{\tan\phi}$$

$$G_{en} = E_{en} \times g_{kwh} \tag{6}$$

Fig. 10 shows the power consumed by network equipment against the load on it with a maximum load of ***G*** bps. Ideally, the power consumed should be proportional to the load, with the maximum power, M watts, being as low as possible. In practice however, the behaviour of network equipment follow the line marked realistic, with the equipment consuming ***I*** watts even under idle (no load) conditions. Thus, ***p*** is the difference between the ideal and measured lines in Fig. 10 and we defined as $p = \frac{\tan\phi}{\tan\theta}$. In here, θ and ϕ are angles at the origin for the ideal and measured power. ***p*** = **1** implies that the equipment has perfect energy proportionality, and ***p*** = **0** implies that the energy consumed by the equipment is completely agnostic to offered load.

In (6), G_{en} is the amount of CO_2 emission for processing the traffic and g_{kWh} is the CO_2 emission per kWh and 0.524lb [19].

We estimated the global energy consumption on switches against the energy proportionality (***p***) by (5), as shown in Fig. 11. The more ***p*** approaches to 1, the more the energy consumption decreases and it drops sharply by ***p*** = **0.2**. Although energy efficiency improves lightly, it can achieve much energy saving. Based on traffic types, we also estimated impacts on the total switches in enterprise network by (5) and (6) (Fig. 12 and Fig. 13). As shown Fig.12 and Fig 13, the Internet video traffic is critical for greening network. When the energy proportionality drops below 0.2, energy consumption and GHG emission of it have strong impacts on network.

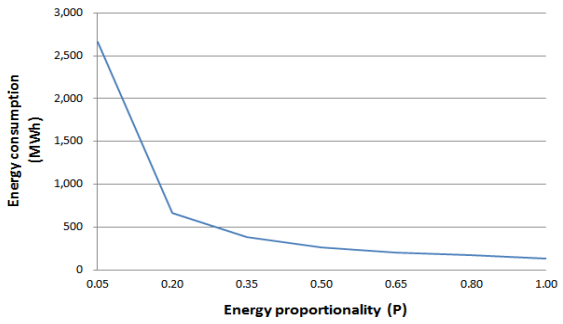


Fig. 11. Energy Consumption against Energy Proportionality on Switching Equipment

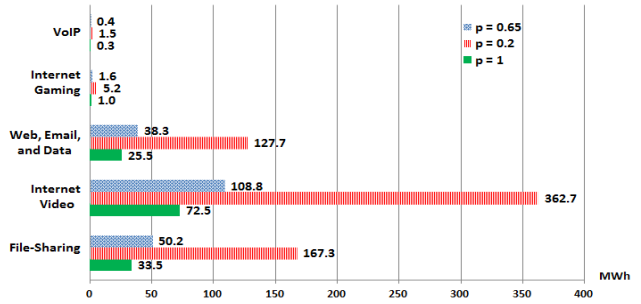


Fig. 12. Energy Impacts of Traffic Types on Switching Equipment

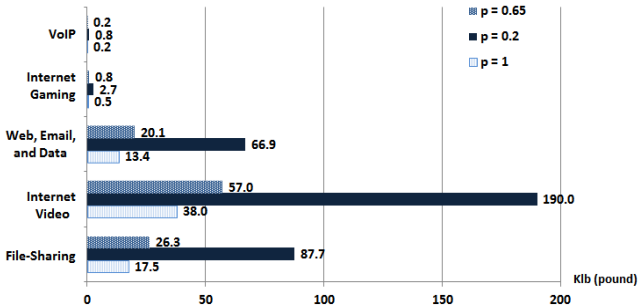


Fig. 13. GHG Impacts of Traffic Types on Switching Equipment

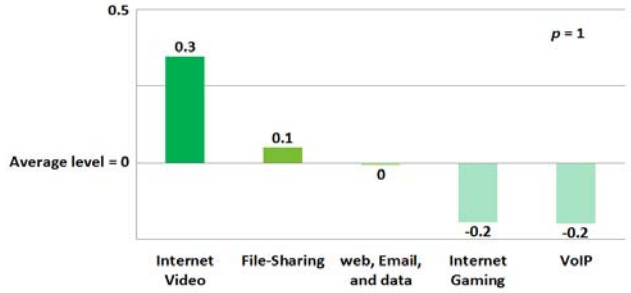


Fig. 14. The greening impacts based on traffic type

We normalized the greening impacts in Fig 14 on a basis of Fig. 12 and Fig 13. In Fig. 14, we set the average level is 0 at average traffic volume and each figures are the relative values of it. Internet traffic is possible to be more specifically distinguished to application-level classification by IP address, port number, connection duration, packet size, packet arrival time, and others, and they can be classified by aspects of energy efficiency as Fig. 12, Fig. 13, and Fig 14. This view is possible to classify with more sophisticated green QoS levels based on traffic type.

As we have said above, the traditional Internet has large energy savings potential and this field remains to be challenged a lot. A major shift is truly needed in research and development to introduce energy-awareness in the network design, without compromising either the QoS or the network reliability. The ultimate goal of networking is to provide services to end-users,

so the quality of the user experience is topics that span all branches of the green networking. So, the energy gain from energy efficiency techniques must not come at the price of a network performance loss.

However this delicate tradeoff arises from opposite principles: networked systems have traditionally been designed and dimensioned according to principles such as over-provisioning and redundancy, green networking approaches praise opposite practices such as resource consolidation and selective connectedness. The challenge lays in this case in applying the latter principles in a way that is as transparent as possible to the user. Most of the previous green studies focused more on the achievable energy gain, and QoS has at least partly been taken into account. Furthermore they have been studied in isolation and that could constitute serious threats to the QoS. Also, the studies and the developments related to green networking have been concentrated on controlling interface on Layer 2.

The techniques based on L2 such as EEE need more delicate synchronization mechanism for link termination and they have to take the trade-off between QoS and energy savings into consideration. Additionally, diverse transitional techniques will have been required until the current network equipment is completely replaced with the green network equipment.

In terms of these, the estimated data in Fig. 12, Fig. 13, and Fig 14 can be used as knowledge for building energy efficient virtual network. As an example, they can be available to build VLANs, which provide logical segmentation based on broadcast domains so reduce unnecessary traffic, improve network performance and offer independency of network. VLAN supports 8 levels as 802.1Q [26] tag user priority level for QoS. These priority levels are possible to be changed by management policy, so they could be classified under the green priority levels including information such as energy consumption, CO2 emissions and costs of network equipment based on traffic types. Traffic could be processed by the levels and the network resource could be allocated by them. If VLANs are mapped by sorts of applications, using mapping based on protocol types, it is possible to make policies by green levels such as Fig 12, Fig 13 and Fig 14, and dynamically build and manage networks by green factors.

Moreover, the above traffic classification based on energy efficiency can help the current green techniques with the dynamic adaptation approaches involving only local decisions make precisional decisions and predictions, as aware of information about patterns and loads of traffic and it can support sophisticated energy management by using knowledge in an aspect of QoS. This approach is possible to introduce to building virtual networks at global level because of Ethernet which has extended to backbone network. We remain to develop specifically advanced these in the long term.

V. CONCLUSION

In this paper, we analyzed impacts of data network equipment for the green Internet and presented future works for the green networking. For this work, we introduced motivation and background of the concern about greening the Internet in view points of energy resource and also presented importance and reasonability of study on energy efficiency of data network equipment. Moreover, we explored detailed energy consumption and savings potential of data network equipment by top-down approach. Then, we stated the current research trends for the green networking and presented limitations in aspects of the quality-of-service (QoS) and a lack of the global view point in them. In this context, we further estimates impacts of data network equipment depending on traffic types and proposed directions for future works related to our estimates. So far, the studies of energy efficiency in the green networking field have focused on dimensioning discreetly components in hardware or controlling network interface by local decision. Moreover few works exist from the global viewpoint which have based on network topology and traffic characteristic. Our traffic classification based on energy efficiency is possible to treat more sophisticatedly trade-off between energy efficiency and QoS. Also these could help precise decisions and predictions of existing green techniques as using the knowledge in an aspect of QoS. Further, this classification can be used in building L2/L3 virtual networks considering for both energy efficiency and QoS. For this work, we need to undertake more precise measurement and modeling, and they remain as our future works.

Greening the Internet is inevitable due to with the energy depletion and environmental threats. Therefore, the green networking solution should be transparent to users and be undertaken to do more elaborate works for QoS support. It should also be developed with a comprehensive view at both local and global - which opens a number of interesting questions that are so far all unexplored.

REFERENCES

[1] A.P. Bianzino, C. Chaudet, D. Rossi, J.-L. Rougier, “A Survey of Green Networking Research”, *IEEE Communications Surveys & Tutorials*, VOL. 14, NO. 1, pp.3-20, Jan., 2012.

[2] IEA, “Energy Technology Perspectives 2012”, 2012.

[3] B. Nordman “What the Real World Tells Us about Saving Energy in Electronics”, presentation, at the 1st Berkeley Symposium on Energy Efficient Electronic Systems(E3S), Jun. 11, 2009.

[4] European Commission DG INFSO, “Impacts of Information and Communication Technologies on Energy Efficiency,” final report, Sep. 2008.

[5] GeSI(Global e-Sustainability Initiative), “SMARTer2020: The Role of ICT in Driving a Sustainable Future”, Dec. 2012.

[6] IPCC(the Intergovernmental Panel on Climate Change),.“2007: Climate Change 2007: Mitigation”.

[7] BIO Intelligence, “Estimate of energy wasted by network-connected equipment Final Report”, Jun., 2011.

[8] R. Bolla, R. Bruschi, F. Davoli, F. Cucchietti, “Energy Efficiency in the Future Internet: A Survey of Existing Approaches and Trends in Energy-Aware Fixed Network Infrastructures”, *IEEE Communications Surveys & Tutorials*, VOL. 13 , NO. 2, pp. 223 – 244, May., 2011.

[9] M. Gupta, S. Sigh, “Greening of the Internet”, in Proc. *The ACM conference on Applications, technologies, architectures, and protocols for computer communications (SIGCOMM '03)*, Karsrhue, Germany, pp. 19-26, Aug., 2003.

[10] L. A. Barroso and U. H¨olzle, “The Case for Energy-Proportional Computing,” *IEEE Computer*, vol. 40, pp. 33 – 37, Dec. 2007.

[11] IEA, " electricity Information 2012", 2012.
[12] The United States in consultation with MEF Partners, "Technology Action Plan Buildings Sector Energy Efficiency", Dec., 2009.
[13] ITU "Measuring the Information Society 2012", 2012.
[14] Lawrence Berkeley National Laboratory, "Data Network Equipment Energy Use and Savings Potential in Buildings", 2010 ACEEE Summer Study on Energy Efficiency in Buildings, Jun., 2010.
[15] Hewlett-Packard Development Company, "Energy Efficient Networking Business white paper", May, 2011.
[16] Cisco, "Cisco Visual Networking Index: Forecast and Methodology, 2011-2016", May,. 2012.
[17] P. Mahadevan, P. Sharma, S. Banerjee, P. Ranganathan, "A Power Benchmarking Framework for Network Devices", *LNCS*, Vol.5550, *(in proc. IFIP International Federation for Information Processing 200*9) pp.795–808, 2009.
[18] GeSI, "Fixed Network Operators Energy Efficiency Benchmark", Mar., 2012.
[19] Cisco White Paper, "Cisco Ethernet Power Study of Cisco and Competitive Products", 2008.
[20] K. Hooghe, M. Guenach, "Towards energy-efficient packet processing in access nodes", in Proc. *Global Telecommunications Conference (GLOBECOM 2011)*, Houston, Texas, USA, pp. 1 – 6, Dec., 2011.
[21] P. Reviriego, J.A. Maestro, J.A. P. Hernández, D. Larrabeiti, "Burst Transmission for Energy-Efficient Ethernet", *IEEE Internet Computing*, Vol 14, No. 4, pp.50-57, Jul.-Aug., 2010.
[22] K. He, Y. Wang, X. Wang, W. Meng, B. Liu, "GreenVLAN : An Energy-Efficient Approach for VLAN Design", in proc. *Computing, Networking and Communications (ICNC 2012)*, pp. 522-526, Jan. 30 - Feb. 2, 2012.
[23] US Energy Information Administration, Electricity Supply, Disposition, Prices and Emisssions, AEO 2012 Reference case. Available : http://www.eia.gov/oiaf/aeo/tablebrowser/#release=AEO2014ER&subject=0-AEO2014ER&table=3-AEO2014ER&region=1-0&cases=full2013-d102312a,ref1914er-d102413a
[24] World Nuclear News. Is there life after sixty? Available: http://www.world-nuclear-news.org/RS_Is_540there_life_after_sixty_0302121.html
[25] IEEE P802.3az Energy e_cient Ethernet task force. Available : http://www.ieee802.org/3/az/index.html.
[26] WG802.1 - Higher Layer LAN Protocols Working Group. Available : http://standards.ieee.org/findstds/standard/802.1Q-2011.html.

Yuhwa Suh is a doctor student in the Internet Convergences and Networking Laboratory at Soongsil University in South Korea. She received B.S. and M.S. degrees in computer science and engineering from Soongsil University in 2003 and 2005, respectively. She worked as a researcher at National IT Industry Promotion Agency in South Korea. Her research interests include Greening the Internet, VLAN, Ad-hoc/Sensor Networks and Multicast.

Ki-Young Kim has been working as a professor at the Department of Computer Software of Seoil University in South Korea, since 2004. He received his M.S. and Ph. D. degrees in computer science and engineering from Soongsil University in 1999 and 2003, respectively. He worked as a researcher at Trigem Information & Communication Co. in South Korea from 1995 to 1997. His research interests include Mobile Computing, Multicast, ITS, Network Security.

Jongseok Choi is a doctor student in the Internet Convergences and Networking Laboratory at Soongsil University in South Korea. He received M.S. degree in computer science and engineering from Soongsil University in 2012 and B.S. degree in Computer Science from Bucheon University in South Korea in 2010. His research areas include Greening the Internet, WLAN, Mobile Network.

Yongtae Shin has been working as a professor at the School of Computer Science and Engineering of Soongsil University in South Korea, since 1995. He is also the founder of DigiCAP Inc., which is a leader of the DRM and CAS solution and service providers in South Korea. He received M.S. and Ph.D. degrees in computer science from the University of Iowa. His research areas include DRM, BCN, Wireless Networks, QoS, Network Security.

Volume 3 Issue 3, May 2014, ISSN: 2288-0003

ICACT-TACT JOURNAL

Global IT Research Institute

1713 Obelisk, 216 Seohyunno, Bundang-gu, Sungnam Kyunggi-do, Republic of Korea 463-824

Business Licence Number : 220-82-07506, Contact: secretariat@icact.org Tel: +82-70-4146-4991