Time domain features in combination with a support vector machine classifier for constructing the termite detection system

M A Nanda¹, K B Seminar¹, D Nandika² and A Maddu³
¹Department of Mechanical and Bio-system Engineering, Bogor Agricultural University, Bogor 16680, West Java, Indonesia
²Department of Forest Products, Bogor Agricultural University, Bogor 16680, West Java, Indonesia
³Department of Physics, Bogor Agricultural University, Bogor 16680, West Java, Indonesia
E-mail: m.achirulnanda@gmail.com

Abstract. Over the last decade, it has been broadly reported that wooden buildings have been massively degraded due to termite attacks. The termite detection system based on the acoustic signal has been proposed to overcome such termite attacks. In this study, we investigate the implementation of the support vector machine (SVM) at the termite detection system. In this work, the pine wood as the medium for termite infestation was divided into two groups, i.e., the wood infested by termites Coptotermes curvignathus (‘infested’) and the normal wood (‘uninfested’). The acoustic signal from each group was analyzed to produce the acoustic features, i.e., energy (E) and entropy (H). Subsequently, the acoustic features were included to construct the SVM Classifier. According to the numerical results, the SVM classifier achieved an accuracy of 93.21 ± 2.58%.

1. Introduction
Temperature and humidity (i.e., 28-30 °C and 70-90 %) in Indonesia are ideal for termite life, no surprise that there are 200 termite species that spread in various regions [1]. This certainly has a significant adverse effect on the damage of wooden buildings. Over the last decade, it is broadly reported that wooden buildings have experienced the degradation of their shelf life due to termite attacks. Arinana et al. [2] indicated that termites were able to destroy the wooden buildings up to the 33rd floor in Jakarta, Indonesia. As a result of this attack, Nandika et al. [1] estimated that the economic loss to house buildings in Indonesia reached IDR. 8.7 trillion in 2015, which is expected to increase continuously along with the enhancement in community settlements.

Among the 200 termite species, there is one that is considered to inflict the most significant impact loss, called Coptotermes curvignathus [3]. This termite is categorized as a subterranean termite species, the Rhinotermitidae family. This termite is widely reported to have damaged the apartments, books and buildings in Indonesia University [4]. One way to overcome the termite attack is with initial detection. Due to their cryptic behavior, termite attack on wood or wood products, it is difficult to detect visually.

Various approaches based on termite behavior when attacking the wood have been developed to detect termites including acoustic, temperature, moisture and gas related approaches. Termites produce the acoustic signal during foraging and moving. Therefore, this approach has been used to design termite
detection systems by many researchers [5, 6]. However, Manzoor [7] reported that the termite detection systems still faces challenges to separate the acoustic signals generated by termites or noise from the environment. One way to solve this problem is to build a classification model. Therefore, we propose a support vector machine (SVM) classifier to construct a termite detection system.

2. Materials and Method

2.1. Timber Test

Pine wood, used as a medium for the feeding activity test of termites Coptotermes curvignathus. This wood has a size of 20 (l) x 9.5 (w) x 2.5 (h) cm and the inside has a hole with size 12 (l) x 6 (w) x 0.5 (h). Firstly, we group the wood into two classes, i.e., positive class (+1), the wood infested by 220 termites C. curvignathus ('infested'), and the negative class (-1), the normal wood that is not infested by termites ('uninfested'). Each class consists of four wood samples. In addition, the termite infestation is carried out for seven days.

2.2. Data Acquisition and Time Domain Features

The equipment used for the data acquisition is two electret microphone sensors by ITEad Studio, with a sensitivity of -56 dB (frequency of 0.1-10 kHz), which is connected to the microcontroller. Furthermore, the acoustic signal will be displayed directly on the personal computer. Figure 1, a diagram of the acoustic signal to measure the acoustic signals in both classes is shown. The data acquisition is carried out for 3 seconds (repeated 10 times) on each wood.

![Figure 1. Diagram of acoustic signal sensing in the wood](image)

The time domain feature is generated without requiring any a signal transformation first. This feature is easier and faster than features in other domains, in terms of computation, which is the frequency domain [8], because we extract its feature from the acoustic signal sample directly [9]. In this study, we propose the two time domain features, energy and entropy.

Furthermore, the energy \( E_i \) and entropy \( H \) can be calculated following equations 1 and 2 respectively [10]. Where, \( E_i \) is energy, \( H \) is entropy, \( W_L \) is length of frame size, \( e_j \) is the energy ratio of sub-frames with the total energy total and \( X_i \) is value of signal, \( n=1, \ldots, W_L \).

\[
E_i = \frac{1}{W_L} \sum_{n=1}^{W_L} |X_i(n)|^2
\]

\[
H = -e_j \log_2(e_j)
\]

2.3. Support Vector Machine (SVM)

SVM is a binary classification method developed by Vladimir Vapnik in 1995 [11, 12]. The basic concept of SVM is to find the best decision boundary (hyperplane) that can separate perfectly the d-dimensional data into two classes (class +1 and class -1). As can be seen in figure 2, the hyperplane should be able to maximize the geometric distance from the hyperplane to its support vector (SV); this
distance is well-known as margin \((r)\). In addition, \(\beta \) and \(\beta_0\) are a vector going through the center and the offset parameter sequentially.

\[
\alpha_i y_i \exp(-\gamma \|x_i - x_d\|^2) + b
\]

Where, we classify the data \(x_d\) into the class (i.e., +1 or -1) based on the calculation of several parameters, namely: \(\alpha_i\) which is a lagrange multiplier, \(y_i\) which is the membership of the class (+1, -1), \(x_i\) which represents the support vector data, \(\gamma\) as the gamma and \(b\) is intercept, with \(i = 1, 2, 3, \ldots, n\).

Grid search is the step to take the best parameter value which is needed to construct the SVM classifier [14]. The required parameters in the RBF kernel are the cost parameter \((C)\) and the gamma parameter \((\gamma)\). The cost parameter can be viewed as a way to control the over-fitting of each support vector [15, 16]. In table 1, we set the lower and upper limit to find the best parameter value (i.e., \(C\) and \(\gamma\)).

| Kernel function | Cost \((C)\) | Gamma \((\gamma)\) |
|------------------|-------------|------------------|
| RBF              | \(2^{-20}\) | \(2^2\)   | \(2^{-15}\) | \(2^2\) |

\(a\) LL is the lower limit.  
\(b\) UL is the upper limit.

2.4. Classifier Performance
Measurement of the classification performance was performed by taking 40 observations randomly from the training data set, and was repeated seven times. Furthermore, this performance produces 2 x 2 confusion matrix, including \(TP\), \(TN\), \(FP\) and \(FN\), where, true positives \((TP)\) were the actual infested class that was predicted correctly. True negatives \((TN)\) were the uninfested class that was predicted correctly. False positives \((FP)\) were the infested class that was predicted as an uninfested class. False negatives
(FN) were the uninfested class that was incorrectly classified. With a simple formula, we can measure the accuracy (AC) of the proposed classifier as follows [17]:

\[
AC = \left( \frac{TP + TN}{TP + FP + FN + TN} \right) \times 100(\%)
\]

3. Results and Discussion

3.1. Time Domain Features Distribution

Two proposed features in the time domain are energy \((E_i)\) and entropy \((H)\). As can be seen in figure 3, based on numerical results, we can plot the two dimensional data distribution between \(E_i\) and \(H\) in each class. Clearly, the features in both class (i.e., infested and uninfested) overlap; thus, they are difficult to separate with linear hyperplane. Therefore, kernel function is required to overcome this problem.

Figure 3. Time domain features distribution in the both groups

3.2. RBF Kernel Implementation

A common kernel is the radial basis function (RBF); this kernel is implemented into our termite detection system. Two important parameters need to be optimized to maximize the margin distance between two classes, i.e. cost \((C)\) and gamma \((\gamma)\). As a result, table 2 shows the best parameters for \(C\) and \(\gamma\) are 4 and 4 sequentially.

Table 2. The best parameters value using the grid search method

| RBF kernel   |          |
|--------------|----------|
| Cost \((C)\) | 4        |
| Gamma \((\gamma)\) | 4        |

\[
f(x_d) = \sum_{i=1}^{49} \alpha_i y_i \exp(-4\|x_i - x_d\|^2) + 0.162
\]

Based on the numerical calculation, this SVM classifier needs 49 support vectors. Furthermore, we can classify the new observation \((x_d)\) into the class (i.e., infested or uninfested) following equation 5. The
rule of this classifier is, if \( f(x_d) > 0 \), then \( x_d \) will be classified into the infested class. However, if \( f(x_d) < 0 \), then \( x_d \) will be classified into the uninfested class.

3.3. Performance of SVM Classifier
Based on the procedure described in Chapter 2.4, it produces the \( TP, TN, FP \) and \( FN \). The performance of the SVM classifier can be seen in Table 3, where our proposed classifier that is implemented into the termite detection system has an accuracy of \( 93.21 \pm 2.58\% \). This classifier is built using two acoustic features, namely energy \( (E_i) \) and entropy \( (H) \) in the time domain.

| Repetition | TP  | TN  | FP  | FN  | AC (%) |
|------------|-----|-----|-----|-----|--------|
| 1          | 17  | 21  | 0   | 2   | 95     |
| 2          | 16  | 22  | 1   | 1   | 95     |
| 3          | 18  | 17  | 2   | 3   | 87.5   |
| 4          | 17  | 20  | 2   | 1   | 92.5   |
| 5          | 17  | 21  | 1   | 1   | 95     |
| 6          | 16  | 22  | 2   | 0   | 95     |
| 7          | 18  | 19  | 0   | 3   | 92.5   |
| mean ± stdev | 93.21 ± 2.58 |

In previous research, Lewis et al. [18] tested the performance of the termite detection tool; this has been tested on different treatments, i.e., various wood and the location of the sensor. Based on numerical results, that tool has an accuracy of 80%. Therefore, the opportunity of the further research is to test this classifier on various types of wood.

4. Conclusion
Detecting termites in wood is a real challenge in the context of integrated pest management, which aims to avoid higher damage of wood. In this paper, we have proposed an effective framework of a termite detection system based on acoustic signal by application of the SVM classifier. In this classifier, we implemented the radial basis function (RBF) kernel and optimized the parameters (i.e., cost and gamma) to achieve the maximum accuracy. According to the numerical results, our termite detection system in combination with the SVM classifier achieved an accuracy of \( 93.21 \pm 2.58\% \).

Acknowledgment
This work is supported by the accelerated master's program leading to doctoral research grants (PMDSU), Indonesia. In addition, the first author thanks Prof. Kudang Boro Seminar and Prof. Dodi Nandika, who provided the facility during the research at the Bioinformatics Laboratory and Termite Laboratory, Bogor Agricultural University.

References
[1] Nandika D, Rismayadi Y and Diba F 2015 Rayap: Biologi dan Pengendaliannya, edisi ke-2 (Termite: Biology and its Control, 2nd Edition) (Surakarta: Universitas Muhammadiyah Surakarta)

[2] Arinana, Philippines I, Koesmaryono Y, Nandika D, Aunu R, Harahap I S, Sumertajaya I M and Bahtiar E T 2016 Coptotermes curvignathus Holmgren (Isoptera: Rhinotermitidae) capability to maintain the temperature inside its nests J. of Entomology 13 (5-6) 199-202
[3] Sukartana P, Sumarni G and Broadbent S 2009 Evaluation of chlorfluazuron in controlling the subterranean termite Coptotermes curvignathus (Isoptera: Rhinotermitidae) Indonesia J. of Tropic. For. Sci. 21 13-8
[4] Primanda A, Ischak T M and Basukriadi A 2003 Termite species richness on the campus of Universitas Indonesia, Depok Makara J. of Sci. 7 9-14
[5] Mankin R, Hagstrum D, Smith M, Roda A and Kairo M 2011 Perspective and promise: A century of insect acoustic detection and monitoring American Entomologist 57 30-44
[6] Lewis V, Leighton S, Tabuchi R and Haverty M 2011 Seasonal and daily patterns in activity of the western drywood termite, Incisitermes minor (Hagen) Insects 2 555-63
[7] Manzoor F 2013 Biosensors for termite control. In: IOP Conference Series: Materials Science and Engineering, (Indus University, Pakistan: IOP Publishing) pp 1-3
[8] Griffin J 1989 An alternating frequency/time domain method for calculating the steady-state response of nonlinear dynamic systems J. of App. Mech. 56 149
[9] Rivai M and Tasrripan 2015 Fuel qualification using quartz sensors ARPN J. of Eng. and App. Sci. 10 6737 – 43
[10] Giannakopoulos T and Pikrakis A 2014 Introduction to Audio Analysis: A MATLAB® Approach (USA: Academic Press)
[11] Bakheet S 2017 An SVM Framework for Malignant Melanoma Detection Based on Optimized HOG Features Computation 5 4
[12] Karatzoglou A, Meyer D and Hornik K 2005 Support vector machines in R J. of Stat. Software 15 1-28
[13] Sakthivel N, Saravanamurugan S, Nair B B, Elangovan M and Sugumaran V 2016 Effect of kernel function in support vector machine for the fault diagnosis of pump J. of Eng. Sci. and Tech. (JESTEC) 11 826-38
[14] Santos P, Villa L F, Reñones A, Bustillo A and Maudes J 2015 An SVM-based solution for fault detection in wind turbines Sensors 15 5627-48
[15] Arun C and Durairaj D C 2017 Identifying medicinal plant leaves using textures and optimal colour spaces channel J. Ilmu Komp. dan Informasi 10 19-28
[16] Nanda M A, Seminar K B, Nandika D and Maddu A 2018 A comparison study of kernel functions in the support vector machine and its application for termite detection Information 9 5
[17] Ciptohadijoyo S, Litananda W, Rivai M and Purnomo M 2016 Electronic nose based on partition column integrated with gas sensor for fruit identification and classification Comp. and Electronics in Agric. 121 429-35
[18] Lewis V R, Power A B and Haverty M I 2004 Surface and subsurface sensor performance in acoustically detecting western drywood termites in naturally infested boards Forest Products J. 54 57-62