# The Dempster-Shafer Theory Algorithm and its Application to Insect Diseases Detection

Andino Maseleno and Md. Mahmud Hasan

*Computer Science Program, Universiti Brunei Darussalam*
*andinomaseleno@pasca.gadjahmada.edu, mahmud.hasan@ubd.edu.bn*

## Abstract

*This paper presents Dempster-Shafer Theory for insect diseases detection. Sustainable elimination of insect diseases as a public-health problem is feasible and requires continuous efforts and innovative approaches. In this research, we used Dempster-Shafer theory for detecting insect diseases and displaying the result of detection process. Insect diseases which include babesiosis, dengue fever, lyme, malaria, and west nile. We describe six symptoms as major symptoms which include fever, red urine, skin rash, paralysis, headache, and arthritis. Dempster-Shafer theory to quantify the degree of belief, our approach uses Dempster-Shafer theory to combine beliefs under conditions of uncertainty and ignorance, and allows quantitative measurement of the belief and plausibility in our identification result.*

*Keywords: Insect diseases, detection, Dempster-Shafer theory*

## 1. Introduction

There are many insects that are the primary or intermediate hosts or carriers of human diseases. Mosquitoes are perhaps the best known invertebrate vector and transmit a wide range of tropical diseases including malaria, dengue fever and yellow fever. Malaria is the world's most devastating disease and kills more people than any other communicable disease except Tuberculosis. Between 2000 and 2009, out of 36 countries listed as endemic, 24 received the exclusive support of WHO (World Health Organization) either to assess the epidemiological status of HAT (human African trypanosomiasis) or to establish control and surveillance activities [1]. Another large group of vectors are flies. Sandfly species transmit the disease leishmaniasis, by acting as vectors for protozoan Leishmania species, and tsetse flies transmit protozoan trypansomes (Trypanosoma brucei gambiense and Trypansoma brucei rhodesiense) which cause African Trypanosomiasis (sleeping sickness). Ticks and lice form another large group of invertebrate vectors. The bacterium Borrelia burgdorferi, which causes Lyme Disease, is transmitted by ticks and members of the bacterial genus Rickettsia are transmitted by lice. For example, the human body louse transmits the bacterium Rickettsia prowazekii which causes epidemic typhus. Some systems for diagnosis in insect diseases have been developed which were expert system for identifies forest insects and proposes relevant treatment [2], and expert system of diseases and insects of jujube based on neural networks [3]. Actually, according to researchers knowledge, Dempster-Shafer theory of evidence has never been used for built an system for detecting insect diseases.

## 2. Dempster-Shafer Theory

The Dempster-Shafer theory [7] or the theory of belief functions is a mathematical theory of evidence which can be interpreted as a generalization of probability theory in which the elements of the sample space to which nonzero probability mass is attributed are not single points but sets [8]. The sets that get nonzero mass are called focal elements. The sum of these probability masses is one, however, the basic difference between Dempster-Shafer theory and traditional probability theory is that the focal elements of a Dempster-Shafer structure may overlap one another. The Dempster-Shafer theory also provides methods to represent and combine weights of evidence.

Considering a finite set (frame of discernment) $\Theta$, a basic probability assignment is a function $m: 2\Theta \rightarrow [0, 1]$ so that $m(\emptyset) = 0$, $A \subseteq \Theta$ $m(A) = 1$ and $m(A) \geq 0$ for all $A \subseteq \Theta$. The subsets of $\Theta$ which are associated with nonzero values of $m$ are known as focal elements and the union of the focal elements is called core. The value of $m(A)$ expresses the proportion of all relevant and available evidence that supports the claim that a particular element of $\Theta$ belongs to the set $A$ but to no particular subset of $A$. This value pertains only to the set $A$ and makes no additional claims about any subsets of $A$. From this kind of mass assignment, the upper and lower bounds of a probability interval can be defined. Shafer defined the concepts of belief and plausibility as two measures over the subsets of $\Theta$ as follows.

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B) \tag{1}$$

$$\text{Pl}(A) = \sum_{B \cap A \neq \emptyset} m(B) \tag{2}$$

A basic probability assignment can also be viewed as determining a set of probability distributions $P$ over $\Theta$ so that $\text{Bel}(A) \leq P(A) \leq \text{Pl}(A)$. It can be easily seen that these two measures are related to each other as $\text{Pl}(A) = 1 - \text{Bel}()$. Therefore, one needs to know only one of the three values of $m$, $\text{Bel}$, or $\text{Pl}$ to derive the other two. Dempster's rule of combination can be used for pooling of evidence from two belief functions $\text{Bel}1$ and $\text{Bel}2$ over the same frame of discernment, but induced by different independent sources of information. The Dempster's rule of combination for combining two sets of masses, $m1$ and $m2$ is defined as follows.

$$m_{12}(\emptyset) = 0 \tag{3}$$

$$m_{12}(A) = \frac{1}{1-k} \sum_{B \cap C = A \neq \emptyset} m_1(B) m_2(C) \tag{4}$$

$$k = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{5}$$

Here $k$ is a measure of the amount of conflict between two evidences. If $k = 1$ the two evidences cannot be combined because their cores are disjoint. This rule is commutative, associative, but not idempotent or continuous.

In this research, we used Dempster-Shafer theory for detecting insect diseases and displaying the result of detection process. Insect diseases detection used Dempster-Shafer theory for decision support process. Flowchart of insect diseases detection shown in Figure 1.
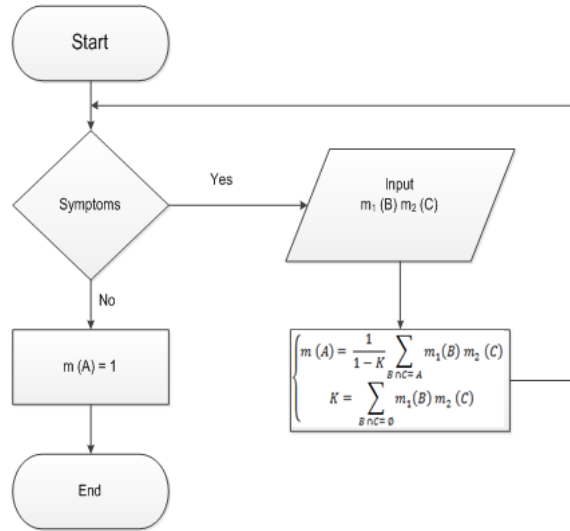
**Figure 1.  Flowchart of Insect Diseases Detection**

Insect diseases which include babesiosis, dengue fever, lyme, malaria, and west nile. We describe six symptoms as major symptoms which include fever, red urine, skin rash, paralysis, headache, and arthritis. Basic probability assignment for each symptoms can be seen in Table 1.

**Table 1. Basic Probability Assignment**

| No. | Symptom | Diseases | Basic Probability Assignment | | | | |
|-----|---------|----------|------------------------------|--|--|--|--|
| | | | Condition 1 | Condition 2 | Condition 3 | Condition 4 | Condition 5 |
| 1 | Fever | Babesiosis | 0.45 | 0.65 | 0.65 | 0.65 | 0.65 |
| | | Dengue Fever | | | | | |
| | | Malaria | | | | | |
| | | West Nile | | | | | |
| 2 | Red urine | Babesiosis | 0.55 | 0.45 | 0.65 | 0.65 | 0.65 |
| 3 | Skin rash | Lyme | 0.45 | 0.55 | 0.45 | 0.65 | 0.65 |
| 4 | Paralysis | Lyme | 0.45 | 0.45 | 0.55 | 0.45 | 0.65 |
| 5 | Headache | Malaria | 0.55 | 0.45 | 0.45 | 0.55 | 0.45 |
| 6 | Arthritis | Dengue Fever | 0.65 | 0.65 | 0.65 | 0.65 | 0.65 |

## 3. Detecting Insect Diseases

The following will show detecting insect diseases using Dempster-Shafer Theory. Insect diseases which include babesiosis {B}, dengue fever {DF}, malaria {M}, west nile {WN}, and lyme {L}.

### 3.1  Symptom 1: Fever

Fever is a symptom of babesiosis, dengue fever, lyme, malaria, and west nile with a bpa of 0.45, so that:

$m_1$ { B, DF, M, WN} =  0.45

$m_1$ { $\Theta$} = 1 – 0.45 = 0.55

### 3.2 Symptom 2: Red Urine

Red urine is a symptom of babesiosis with a bpa of 0.55, so that:

$m_2 \{B\} = 0.55$

$m_2 \{\Theta\} = 1 - 0.55 = 0.45$

With red urine symptom then required to calculate the new bpa values for some combinations ($m_3$). Combination rules for the $m_3$ can be seen in the Table 2.

**Table 2. Combination of Symptom 1 and Symptom 2**

| {B} | | 0.55 | $\Theta$ | | 0.45 |
|---|---|---|---|---|---|
| {B, DF, M, WN} | 0.45 | {B} | 0.25 | {B, DF, M, WN} | 0.20 |
| $\Theta$ | 0.55 | {B} | 0.30 | $\Theta$ | 0.25 |

$m_3 \{B\} = \dfrac{0.25 + 0.30}{1 - 0} = 0.55$

$m_3 \{B, DF, M, WN\} = \dfrac{0.20}{1 - 0} = 0.20$

$m_3 \{\Theta\} = \dfrac{0.25}{1 - 0} = 0.25$

### 3. 3 Symptom 3: Skin Rash

Skin rash is a symptom of lyme with a bpa of 0.45, so that:

$m_4 \{L\} = 0.45$

$m_4 \{\Theta\} = 1 - 0.45 = 0.55$

With skin rash symptom then required to calculate the new bpa values for some combinations ($m_5$). Combination rules for the $m_5$ can be seen in the Table 3.

**Table 3. Combination of Symptom 1, Symptom 2 and Symptom 3**

| {L} | | 0.45 | | $\Theta$ | 0.55 |
|---|---|---|---|---|---|
| {B} | 0.55 | Ø | 0.25 | {B} | 0.30 |
| {B, DF, M, WN } | 0.20 | Ø | 0.09 | {B, DF, M, WN } | 0.11 |
| $\Theta$ | 0.25 | {L} | 0.11 | $\Theta$ | 0.14 |

$m_5 \{B\} = \dfrac{0.30}{1 - (0.25 + 0.09)} = 0.45$

$m_5 \{B, DF, M, WN\} = \dfrac{0.11}{1 - (0.25 + 0.09)} = 0.17$

$m_5 \{L\} = \dfrac{0.11}{1 - (0.25 + 0.09)} = 0.17$

$m_5 \{\Theta\} = \dfrac{0.14}{1 - (0.25 + 0.09)} = 0.21$

### 3.4 Symptom 4: Paralysis

Paralysis is a symptom of lyme with a bpa of 0.45, so that:

$m_6 \{L\} = 0.45$

$m_6 \{\Theta\} = \ 1 - 0.45 = 0.55$

With paralysis symptom then required to calculate the new bpa values for some combinations ($m_7$). Combination rules for the $m_7$ can be seen in the Table 4.

**Table 4. Combination of Symptom 1, Symptom 2, Symptom 3 and Symptom 4**

| {L} | | | 0.45 | $\Theta$ | | 0.55 |
|---|---|---|---|---|---|---|
| {B} | 0.45 | Ø | 0.20 | {B} | | 0.25 |
| {B, DF, M, WN} | 0.17 | Ø | 0.08 | {B, DF, M, WN} | | 0.09 |
| {L} | 0.17 | {L} | 0.08 | {L} | | 0.09 |
| $\Theta$ | 0.21 | {L} | 0.09 | $\Theta$ | | 0.11 |

$$m_7 \{B\} = \frac{0.25}{1 - (0.20 + 0.08)} = 0.35$$

$$m_7 \{B, DF, M, WN\} = \frac{0.09}{1 - (0.20 + 0.08)} = 0.12$$

$$m_7 \{L\} = \frac{0.08 + 0.09 + 0.09}{1 - (0.20 + 0.08)} = 0.36$$

$$m_7 \{\Theta\} = \frac{0.11}{1 - (0.20 + 0.08)} = 0.15$$

### 3.5 Symptom 5: Headache

Headache is a symptom of malaria with a bpa of 0.55, so that:

$m_8 (M) = 0.55$

$m_8 \{\Theta\} = \ 1 - 0.55 = 0.45$

With headache symptom then required to calculate the new bpa values for some combinations ($m_9$). Combination rules for the $m_9$ can be seen in the Table 5.

**Table 5. Combination of Symptom 1, Symptom 2, Symptom 3, Symptom 4 and Symptom 5**

| {M} | | | 0.55 | $\Theta$ | | 0.45 |
|---|---|---|---|---|---|---|
| {B} | 0.35 | Ø | 0.19 | {B} | | 0.16 |
| {B, DF, M, WN} | 0.12 | {M} | 0.07 | {B, DF, M, WN} | | 0.05 |
| {L} | 0.36 | Ø | 0.19 | {L} | | 0.16 |
| $\Theta$ | 0.15 | {M} | 0.08 | $\Theta$ | | 0.07 |

$$m_9 \{B\} = \frac{0.16}{1 - (0.19 + 0.19)} = 0.26$$

$$m_9 \{B, DF, M, WN\} = \frac{0.05}{1 - (0.19 + 0.19)} = 0.07$$

$$m_9 \{L\} = \frac{0.16}{1-(0.19+0.19)} = 0.26$$

$$m_9 \{M\} = \frac{0.07+0.08}{1-(0.19+0.19)} = 0.21$$

$$m_9 \{\Theta\} = \frac{0.07}{1-(0.19+0.19)} = 0.09$$

### 3.6 Symptom 6: Arthritis

Arthritis is a symptom of dengue fever with a bpa of 0.65, so that:

$$m_{10} \{DF\} = 0.65$$

$$m_{10} \{\Theta\} = 1 - 0.65 = 0.35$$

With bleeding around the bite symptom then required to calculate the new bpa values for some combinations ($m_{11}$). Combination rules for the $m_{11}$ can be seen in the Table 6.

**Table 6. Combination of Symptom 1, Symptom 2, Symptom 3, Symptom 4, Symptom 5 and Symptom 6**

|  | {DF} |  | 0.65 | Θ | 0.35 |
|---|---|---|---|---|---|
| {B} | 0.26 | Ø | 0.17 | {B} | 0.09 |
| {B, DF, M, WN} | 0.07 | {DF} | 0.04 | { B, DF, M, WN} | 0.02 |
| {L} | 0.26 | Ø | 0.17 | {L} | 0.09 |
| {M} | 0.21 | Ø | 0.14 | {M} | 0.07 |
| Θ | 0.09 | {DF} | 0.06 | Θ | 0.03 |

$$m_{11} \{B\} = \frac{0.09}{1-(0.17+0.17+0.14)} = 0.17$$

$$m_{11} \{ B, DF, M, WN\} = \frac{0.02}{1-(0.17+0.17+0.14)} = 0.04$$

$$m_{11} \{L\} = \frac{0.09}{1-(0.17+0.17+0.14)} = 0.17$$

$$m_{11} \{M\} = \frac{0.07}{1-(0.17+0.17+0.14)} = 0.13$$

$$m_{11} \{DF\} = \frac{0.04+0.06}{1-(0.17+0.17+0.14)} = 0.19$$

$$m_{11} \{\Theta\} = \frac{0.03}{1-(0.17+0.17+0.14)} = 0.06$$

## 4. Result

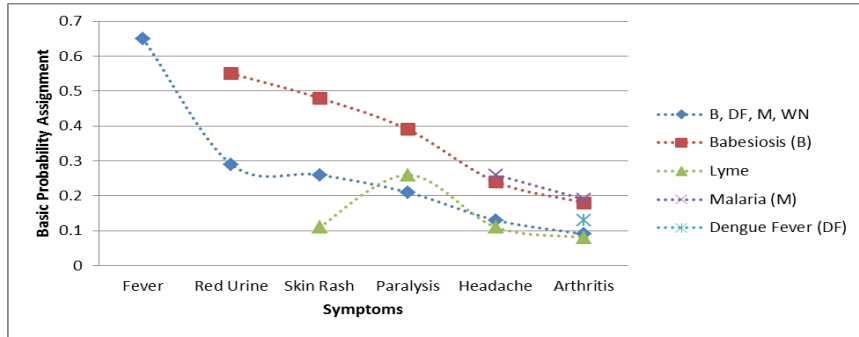Figure 2 through figure 6 are shown graphic of detection from each condition.



**Figure 2. Condition 1**

Figure 2 shows the graphic of Condition 1, we get the highest basic probability assignment is dengue fever that is equal to 0.15 which shows from the last calculation of Dempster-Shafer on symptom 6 which means the possibility of a temporary disease is dengue fever.
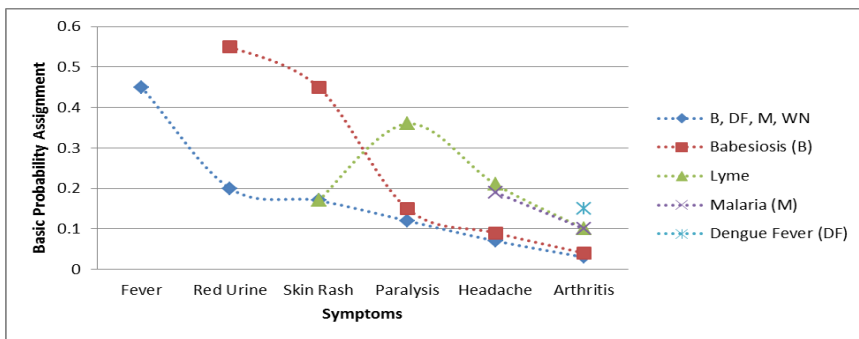


**Figure 3. Condition 2**

Figure 3 shows the graphic of Condition 2, we get the highest basic probability assignment is malaria that is equal to 0.19 which shows from the last calculation of Dempster-Shafer on symptom 6 which means the possibility of a temporary disease with symptoms of fever, red urine, skin rash, paralysis, headache, and arthritis is malaria.
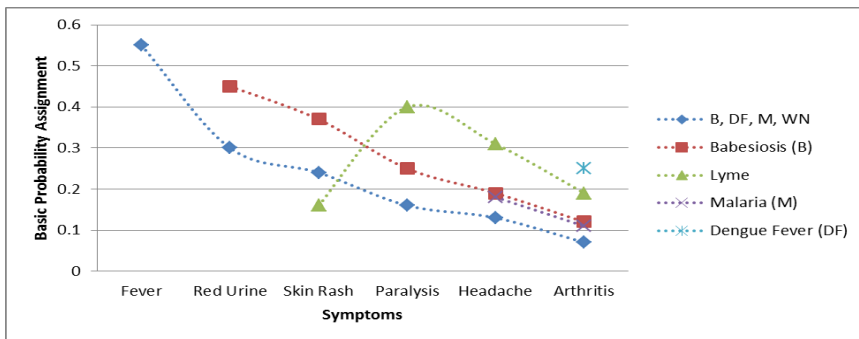


**Figure 4.  Condition 3**

Figure 4 shows the graphic of Condition 3, we get the highest basic probability assignment is dengue fever that is equal to 0.25 which shows from the last calculation of Dempster-Shafer on symptom 6 which means the possibility of a temporary disease with symptoms of fever, red urine, skin rash, paralysis, headache, and arthritis is dengue fever.
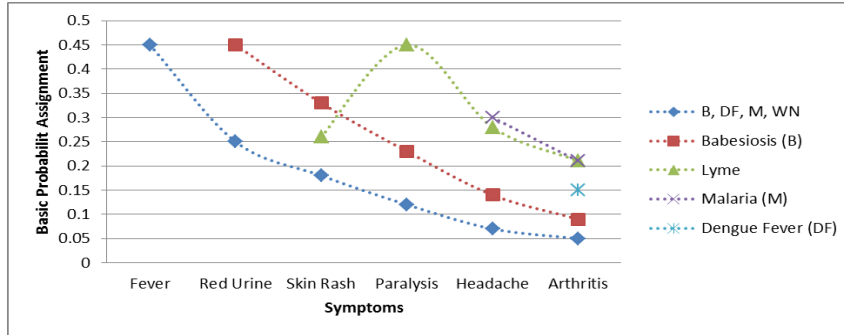


**Figure 5. Condition 4**

Figure 5 shows the graphic of Condition 4, we get the highest basic probability assignments are lyme and malaria that is equal to 0.21 which shows from the last calculation of Dempster-Shafer on symptom 6 which means the possibility of a temporary disease with symptoms of fever, red urine, skin rash, paralysis, headache, and arthritis are lyme and malaria.
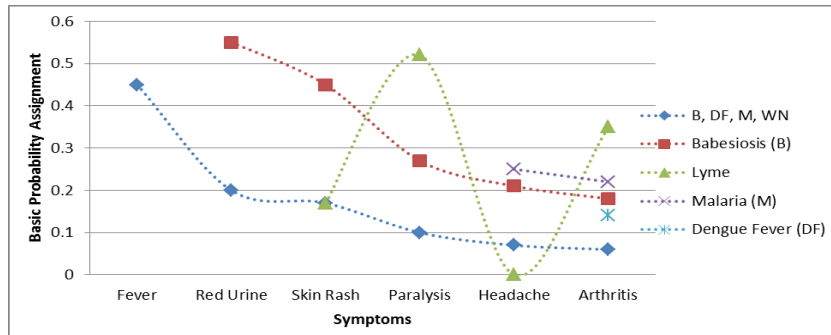


**Figure 6. Condition 5**

Figure 6 shows the graphic of Condition 5, we get the highest basic probability assignment is lyme that is equal to 0.35 which shows from the last calculation of Dempster-Shafer on symptom 5 which means the possibility of a temporary disease with symptoms of fever, red urine, skin rash, paralysis, headache, and arthritis is lyme.

## 5. Conclusion

In this paper we describe six symptoms as major symptoms which include fever, red urine, skin rash, paralysis, headache, and arthritis. The simplest possible method for using probabilities to quantify the uncertainty in a database is that of attaching a probability to every member of a relation, and to use these values to provide the probability that a particular value is the correct answer to a particular query. The knowledge is uncertain in the collection of basic events can be directly used to draw conclusions in simple cases, however, in many

cases the various events associated with each other. Knowledge based is to draw conclusions, it is derived from uncertain knowledge. Reasoning under uncertainty that used some of mathematical expressions, gave them a different interpretation: each piece of evidence may support a subset containing several hypotheses. This is a generalization of the pure probabilistic framework in which every finding corresponds to a value of a variable. Identification of insect diseases can be performed using Dempster-Shafer Theory.

## References

[1] P. P. Simarro, A. Diarra, J. A. R. Postigo, J. R. Franco and J. G. Jannin, "The Human African Trypanosomiasis Control and Surveillance Programme of the World Health Organization 2000–2009: The Way Forward," PLoS Negl Trop Dis., vol. 5, no. 2, **(2011)**, pp. e1007**.**

[2] S. Kaloudisa, D. Anastopoulosa, C. P. Yialourisb and N. A. Lorentzosb, "Sideridis, Insect identification expert system for forest protection", Expert Systems with Applications, vol. 28, no. 3, **(2005)**, pp. 445–452.

[3] J. Song, S. Wu, J. Ren and X. Yu, "Research on expert system of diseases and insects of jujube based on neural networks", Computer Science & Education (ICCSE), 2011 6th International Conference, **(2011).**

[4] A. P. Dempster, "A Generalization of Bayesian inference", Journal of the Royal Statistical Society, vol. B(30), **(1968)**, pp. 205-247**.**

[5] G. Shafer, "A Mathematical Theory of Evidence", Princeton University Press, New Jersey, **(1976).**