Ebola Virus Disease Detection using Dempster-Shafer Evidence Theory

Andino Maseleno, Fauzi, Muhammad Muslihudin
STMIK Pringsewu
Pringsewu- Lampung, Indonesia
Email: andinomaseleno@mail.ru, fauzistmik@yahoo.co.id, muslih.udin@ymail.com
http://maseleno.blogspot.com

***Abstract*—This research presents Ebola virus disease detection using Dempster-Shafer evidence theory. The Dempster-shafer evidential theory is a method about uncertainty reasoning, and this theory reduces the requirements of the knowledge of prior probability and conditional probability. Between 1976 and 2013, the World Health Organization had reported a total of 24 outbreaks involving 1,716 cases. The existing methods used for detecting Ebola virus disease are complex, time consuming, can only be performed under laboratory conditions, and often require highly trained lab workers and time-intensive procedures, as well as a highly sterile experimental environment. The main contribution of this research is to consider Dempster-Shafer evidence theory for Ebola virus disease detection by combined each symptom. The result reveals that Ebola virus disease detection using Dempster-Shafer evidence theory obtained degree of belief of 0.85.**

Keywords: ebola virus disease; uncertainty reasoning; Dempster-Shafer evidence theory

I. Introduction

Ebola virus disease is a disease of humans and other primates caused by ebola viruses. Signs and symptoms typically start between two days and three weeks after contracting the virus with a fever, sore throat, muscular pain, and headaches. Early symptoms of ebola virus disease may be similar to those of other diseases common in Africa, including malaria and dengue fever [1]. The symptoms are also similar to those of Marburg virus disease and other viral hemorrhagic fevers. The complete differential diagnosis is extensive and requires consideration of many other infectious diseases such as typhoid fever, shigellosis, rickettsial diseases, cholera, sepsis, borreliosis, leptospirosis, scrub typhus, plague, Q fever, candidiasis, histoplasmosis, trypanosomiasis, visceral leishmaniasis, measles, and viral hepatitis. Non-infectious diseases that may result in symptoms similar to those of ebola virus disease include acute promyelocytic leukemia, hemolytic uremic syndrome, snake envenomation, clotting factor deficiencies/platelet disorders, thrombotic thrombocytopenic purpura, hereditary hemorrhagic telangiectasia, Kawasaki disease, and warfarin poisoning. Some researchers have studied ebola virus disease detection. Ebola virus disease diagnostics for acute filoviral infection are based mainly on reverse-transcription polymerase chain reaction (RT-PCR) technology and antigen-capture ELISA [3], [4], [5]. Baca et al. [2] studied surface acoustic wave sensors to detect Ebola antigens at the point-of-care without the needfor added reagents, sample processing, or specialized personnel. The existing methods used to detect ebola virus disease are complex, time consuming, and can only be performed under laboratory conditions, often require highly trained lab workers and time-intensive procedures, as well as a highly sterile experimental environment.Dempster-Shafer evidence theory is a powerful tool for combining accumulative evidence and changing prior knowledge in the presence of new evidence [6]. Dempster-Shafer evidence theory allows one to combine evidence from different sources and arrive at a degree of belief which is represented by a belief function that takes into account all the available evidence [6]. The paper is organized as follows. Section 2 presents uncertainty reasoning. Section 3 presents fundamentals of Dempster-Shafer evidence theory. Section 4 describes implementation of Dempster-Shafer evidence theory to ebola virus disease detection. Result and discussion are presented in section 5. Finally, section 6 presents some concluding remarks.

II. Uncertainty Reasoning

Uncertainty is a general concept that reflects human lack of sureness about something or someone, ranging from just short of complete sureness with an almost complete lack of conviction about an outcome. The fundamental structure of uncertainty model contains following three components which include the description of information uncertainty or rules; the description of evidence uncertainty or facts; and the spread of uncertainty [7]. Reasoning theories are divided into certainty reasoning theories and uncertainty reasoning theories. Certainty thinking was once and will be still prevailing in different disciplines. In the Cartesian philosophy, mathematics was the only accurate knowledge learning to provide. With the combination of mathematics and physics, all sorts of natural and social phenomena could be explained in science. Symbolic language of science could construct the universal logic and logical calculus, and all phenomena could be clearer. Newton's absolute time-space sure that all the observable physical quantity in principle could be infinite accurate measurement and its foundation was the uncertainty of physical laws. An unknown world was deterministic for perfectly rational policy makers in the traditional decision science view. A man could get the effect of maximization as long as according to the principle which marginal benefit equals marginal cost decision. However, the world is uncertain [8].

Decisions are often taken on the basis of imperfect information and knowledge (imprecise, uncertain, incomplete) provided by several more or less reliable sources and depending on the states of the world: decisions can be taken in certain, risky or uncertain environment [9]. The lack of certainty is ubiquitous and happens in every single event people encounter

978-1-4673-9088-0/15/$31.00 ©2015 IEEE

in the real world. Uncertainty distinguishes from a certainty in the degree of belief or confidence. If certainty is referred to as a perception or belief that a certain system or phenomenon can experience or not, uncertainty indicates a lack of confidence or trust in an article of knowledge or decision. Uncertainty is a term used in subtly different ways in a number of fields, including philosophy, physics, statistics, economics, finance, insurance, psychology, sociology, engineering, and information science. It applies to predictions of future events, to physical measurements that are already made, or to the unknown. Uncertainty arises in partially observable and/or stochastic environments, as well as due to ignorance and/or indolence [10]. According to the Cambridge Dictionary, "uncertainty is a situation in which something is not known, or something that is not known or certain."[11]

Uncertainty arises from different sources in various forms, and is classified in different ways by different communities. According to the origin of uncertainty, it is categorized into aleatory uncertainty or epistemic uncertainty. Aleatory uncertainty derives from the natural variability of the physical world. It reflects the inherent randomness in nature. It exists naturally regardless of human knowledge. For example, in an event of flipping a coin, the coin comes up heads or tails with some randomness. Even if researchers do many experiments and know the probability of coming up heads, researchers still cannot predict the exact result in the next turn. Aleatory uncertainty cannot be eliminated or reduced by collecting more knowledge or information. No matter whether people know it or not, this uncertainty stays there all the time. Aleatory uncertainty is sometimes also referred to as natural variability, objective uncertainty, external uncertainty, random uncertainty, stochastic uncertainty, inherent uncertainty, irreducible uncertainty, fundamental uncertainty, real world uncertainty, or primary uncertainty. Epistemic uncertainty origins from human's lack of knowledge of the physical world and lack of the ability of measuring and modelling the physical world. Unlike aleatory uncertainty, given more knowledge of the problem and proper methods, epistemic uncertainty can be reduced and sometimes can even be eliminated. Epistemic uncertainty is sometimes also called knowledge uncertainty, subjective uncertainty, internal uncertainty, incompleteness, functional uncertainty, informative uncertainty, or secondary uncertainty. Dempster-Shafer mathematical theory of evidence can deal with both aleatory and epistemic uncertainty.

III. Fundamentals of Dempster-Shafer Evidence Theory

Dempster-Shafer theory of evidence can be implemented as generalization of probability theory [14], [15], [13]. The Dempster-Shafer theory [13] assumes that there is a fixed set of mutually exclusive and exhaustive elements called hypotheses or propositions and symbolized by the Greek letter θ. θ = Disease 1, Disease 2,... , Disease n, where symptom is called a hypothesis or propositions. A hypothesis can be any subset of the frame, in example, to singletons in the frame or to combinations of elements in the frame. θ is also called frame of discernment [13]. A basic probability assignment (bpa) is represented by a mass function $m : 2\theta \rightarrow [0, 1]$ [13]. Where 2θ is the power set of θ. The sum of all basic probability assignment of all subsets of the power set is 1 as shown in equation 2, which embodies the concept that total belief has to be one [16]. The value of the bpa for a given set A (represented as $m(A), A \in 2\theta$, expresses the proportion of all relevant and available evidence that supports the claim that a particular element of θ (the universal set) belongs to the set A but to no particular subset of A. The value of m(A) pertains only to the set A and makes no additional claims about any subsets of A. Any further evidence on the subsets of A would be represented by another bpa, in example B, m(B) would the bpa for the subset B. Formally, this description of m can be represented with the following two equations 1 and 2 [13]:

$$m(\emptyset) = 0 \tag{1}$$

$$\sum_{A \in 2^{\Theta}} m(A) = 1 \tag{2}$$

From the primitive of evidence theory or mass function, the upper and lower bounds of an interval can be defined. This interval contains the precise probability of a set of interest and is bounded by two non additive continuous measures called Belief function and Plausibility function. Evidence theory uses two measures of uncertainty, belief function and plausibility function, expressed as $Bel()$ and $Pls()$ respectively. Given a basic probability assignment m, the corresponding belief function measure and plausibility function measure are determined for all sets $A \in 2\theta$ and $B \in 2\theta$. by equations 3 and 4 [13]:

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{3}$$

$$Pls(A) = \sum_{B \cap A \neq \emptyset} m(B) \tag{4}$$

The support function or belief, Bel, is the total belief of a set and all its subsets. The lower bound Belief for a set A is defined as the sum of all the basic probability assignments of the proper subsets (B) of the set of interest (A) ($A \subset B$). The plausibility function of a proposition, Pls, is the sum of the masses of all propositions in which it is wholly or partially contained. The plausibility function is defined as the degree to which the evidence fails to refute A. These two functions, which have been sometimes referred to as lower and upper probability functions, have the following properties are given by equations 5 and 6:

$$Bel(A) \leq Pls(A) \tag{5}$$

$$Pls(A) = 1 - Bel(\bar{A}) \tag{6}$$

Where $\bar{A}$ is the complementary hypothesis of A, $A \cup \bar{A} = \Theta$ and $A \cap \bar{A} = \emptyset$. The plausibility Pls (A) is defined as the degree to which the evidence fails to refute A. This term is given by the equation 7:

$$Pls(A) = 1 - Bel(\bar{A}) = 1 - \sum_{B \subseteq \bar{A}} m(B) \tag{7}$$

Dempster-Shafer theory provides a method to combine the previous measures of evidence of different sources to deal with evidence conflicts and multi-attribute decision making problems which are subjectively uncertain. This rule assumes that these sources are independent. Dempster's rule of combination [13] given in equation 8 below.

$$(m_1 \oplus m_2)(A) = \begin{cases} 0; & \text{A} = 0 \\ \dfrac{\displaystyle\sum_{B_i \cap B_j = A} m_1(B_i)m_2(B_j)}{1 - \displaystyle\sum_{B_i \cap B_j \neq 0} m_1(B_i)m_2(B_j)}; & \text{A} \neq 0 \end{cases} \tag{8}$$

IV. IMPLEMENTATION

In this implementation, assume that the basic probability assignments of ebola virus disease detection in which already known is available as shown in Table I.

TABLE I. BASIC PROBABILITY ASSIGNMENT OF SYMPTOM

Symptom	Basic Probability Assignment
Fever	0.75
Myalgia	0.70
Vomitting	0.65
Headache	0.60
Sore throat	0.85

In this implementation, ebola virus disease detection describes five symptoms which include fever, myalgia, vomitting, headache and sore throat. The following will shown the process of ebola virus detection using Dempster-Shafer theory of evidence.

1) Symptom 1 is Fever.
Fever is a symptom of dengue fever {DF}, ebola virus disease {E}, hemorrhagic fever {HF} and malaria {M}.
$m_1\{DF, E, HF, M\} = 0.75,$
$m_1\{\theta\} = 1 - 0.75 = 0.25$

2) Symptom 2 is Myalgia.
Myalgia is a symptom of dengue fever {DF}, ebola virus disease {E}, hemorrhagic fever {HF}, and malaria {M}.
$m_2\{DF, E, HF, M\} = 0.70,$
$m_2\{\theta\} = 1 - 0.70 = 0.30$

Table II shows The first combination of ebola virus disease detection.

TABLE II. THE FIRST COMBINATION OF EBOLA VIRUS DISEASE DETECTION

	$m_2(\{DF, E, HF, M\}) = 0.70$	$m_2(\{\theta\}) = 0.30$
$m_1(\{DF, E, HF, M\})$ = 0.75	{DF, E, HF, M} = 0.525	{DF, E, HF, M} = 0.225
$m_1(\{\theta\}) = 0.3$	{M, I, G, LD} = 0.15	$\{\theta\} = 0.15$

The first two bpas m_1 and m_2 are calculated to yield a new bpa m_3 by a combination rule as follows
$m_3\{D, E, HF, M\} = 0.525 + 0.225 + 0.175/(1 - 0) = 0.925$
$m_3\{\theta\} = 0.075/(1 - 0) = 0.075$

3) Symptom 3 is Vomiting.
Vomiting is a symptom of ebola virus disease {E} and malaria {M}.
$m_4\{E, M\} = 0.65$
$m_4\{\theta\} = 1 - 0.65 = 0.35$

Table III shows the second combination of ebola virus disease detection.

TABLE III. THE SECOND COMBINATION OF EBOLA VIRUS DISEASE DETECTION

	$m_4(\{E, M\}) = 0.65$	$m_4(\{\theta\}) = 0.35$
$m_3(\{DF, E, HF, M\})$ = 0.925	{E, M} = 0.601	{DF, E, HF, M} = 0.324
$m_3(\{\theta\}) = 0.075$	{E, M} = 0.049	$\{\theta\} = 0.026$

The first two bpas m_3 and m_4 are calculated to yield a new bpa m_5 by a combination rule as follows
$m_5\{E, M\} = 0.601 + 0.049/(1 - 0) = 0.650$
$m_5\{DF, E, HF, M\} = 0.324/(1 - 0) = 0.324$
$m_5\{\theta\} = 0.026/(1 - 0) = 0.026$

4) Symptom 4 is Headache.
Headache is a symptom of ebola virus disease {E} and Malaria {M}.
$m_6\{E, M\} = 0.60$
$m_6\{\theta\} = 1 - 0.60 = 0.40$

Table IV shows the third combination of ebola virus disease detection.

TABLE IV. THE THIRD COMBINATION OF EBOLA VIRUS DISEASE DETECTION

	$m_6(\{E, M\}) = 0.60$	$m_6(\{\theta\}) = 0.40$
$m_5(\{E, M\}) = 0.650$	{E, M} = 0.39	{E, M} = 0.26
$m_5(\{DF, E, HF, M\})$ = 0.324	{E, M} = 0.194	{DF, E, HF, M} = 0.13
$m_5(\{\theta\}) = 0.026$	{E, M} = 0.016	$\{\theta\} = 0.01$

The first two bpas m_5 and m_6 are calculated to yield a new bpa m_7 by a combination rule as follows
$m_7\{E, M\} = 0.39 + 0.194 + 0.016 + 0.26/(10) = 0.88$
$m_7\{DF, E, HF, M\} = 0.13/(10) = 0.13$
$m_7\{\theta\} = 0.01/(1 - 0) = 0.01$

5) Symptom 5 is sore throat.

Sore throat is a symptom of ebola virus disease {E}.
$m_8\{E\} = 0.85$
$m_8\{\theta\} = 1 - 0.85 = 0.15$

Table V shows the fourth combination of ebola virus disease detection.

TABLE V. THE FOURTH COMBINATION OF EBOLA VIRUS DISEASE DETECTION

	$m_8(\{E\}) = 0.85$	$m_8(\{\theta\}) = 0.15$
$m_7(\{E, M\}) = 0.86$	{E} = 0.73	{E, M} = 0.129
$m_7(\{DF, E, HF, M\}) = 0.13$	{E} = 0.11	{DF, E, HF, M} = 0.019
$m_7(\{\theta\}) = 0.01$	{E} = 0.009	$\{\theta\} = 0.002$

The first two bpas m_7 and m_8 are calculated to yield a new bpa m_9 by a combination rule as follows
$m_9\{E\} = 0.73 + 0.11 + 0.01/(1-0) = 0.85$
$m_9\{E, M\} = 0.129/(1-0) = 0.129$
$m_9\{DF, E, HF, M\} = 0.019/(1-0) = 0.019$
$m_9\{\theta\} = 0.002/(1-0) = 0.002$

Finally, the final ranking of the degree of belief is 0.85 > 0.129 > 0.019. The final ranking is Ebola virus disease > Ebola virus disease, Malaria > Dengue fever, Ebola virus disease, Hemorrhagic fever, Malaria .

V. Result and Discussion

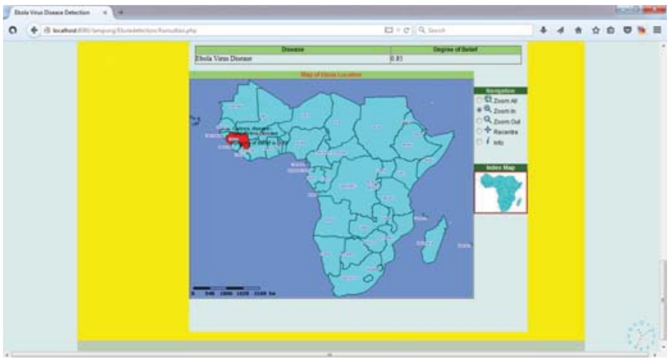

Fig. 1. Result of ebola virus disease detection

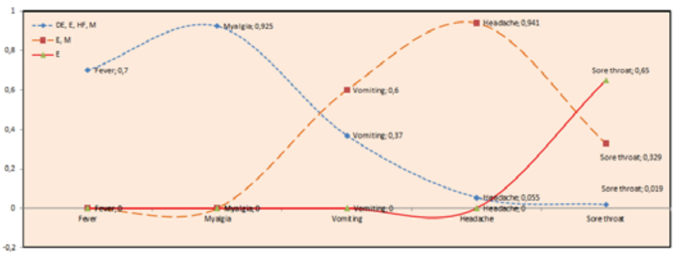

Fig. 2. Degree of belief of ebola virus disease

Figure 1 shows the result of detection of ebola virus disease. Figure 2 shows degree of belief of ebola virus disease. The final ranking of the degree of belief is 0.85 > 0.129 > 0.019. It can be seen from Figure 2 that the final ranking is Ebola virus disease > Ebola virus disease, Malaria > Dengue Fever, Ebola virus disease, Hemorrhagic Fever, Malaria.

VI. Conclusion

Early warning system of Ebola virus disease detection is important in interrupting the transmission cycle of the parasite and progress of the disease to the late stage. Early warning is the provision of timely and effective information, through identifying institutions, that allows individuals exposed to a hazard to take action to avoid or reduce their risks and prepare for effective responses. Therefore, cost-effective, simple, rapid, robust and reliable methods, are urgently needed. There is also an urgent need for accurate tools for the diagnosis of the disease spreading, a new initiative for the development of new diagnostic tests to support the control of disease. In this research, Ebola virus disease detection using uncertainty reasoning is proposed. In the proposed method, a basic probability assignment is assigned to each symptom. Symptoms are combined using Dempster-Shafer evidence theory. The result reveals that Ebola virus disease obtained degree of belief of 0.85. The Dempster-Shafer evidence theory approach described in this research is conceptually straightforward, simple to implement, efficient to compute, and performs well in the detection of Ebola virus disease.

References

[1] World Health Organization. *Neglected Tropical Diseases*. Switzerland: World Health Organization; 2009.

[2] Dietmar S. The history of African Trypanosomiasis. *Parasites and Vectors* 2008; **1**: 1 - 8.

[3] Molyneux DH, Pentreath V, Doua F. African trypanosomiasis in man. In: Cook GC, editors. *Manson's Tropical Diseases 20th edition*. London: W.B. Saunders Company Ltd; 1996. p. 1171-1196.

[4] Simaro PP, Diarra A, Postigo JAR, Franco JR, Jannin JG. The Human African Trypanosomiasis Control and Surveillance Programme of the World Health Organization 2000-2009: The Way Forward. *PLoS Neglected Tropical Diseases* 2011; **5**.

[5] Tarwater PM, Martin CF. The Effects of population density on the spread of disease. *Complexity* 2001; **6**: 29 - 36.

[6] Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. Global trends in emerging infectious diseases. *Nature*, 2008; **451**:990-993.

[7] Li, Y., Fan, X., Wang, J., He, Q., Yuan, Y.,: Uncertainty Reasoning on Fuzziness and Randomness in Challenged Networks. In: 2010 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, pp. 466 - 471. Hongkong (2010).

[8] Yanfei, S.,: Emotional Intelligence and Affective Computing in Uncertainty Situation. In: 2013 Third International Conference on Intelligent System Design and Engineering Applications, pp. 705-708. Hongkong (2013).

[9] Tacnet, J.M., Deert, J.,: Cautious OWA and evidential reasoning for decision making under uncertainty. In: 2011 Proceedings of the 14th International Conference on Information Fusion (FUSION). pp. 1-8. Chicago, USA (2011).

[10] Norvig, P., Thrun, S.,: Introduction to Artificial Intelligence (2013).

[11] Cambridge,: Cambridge Dictionaries Online. Cambridge University Press (2013).

[12] Zadeh LA. Fuzzy sets. *Information and Control* 1965; **8**:338-353.

[13] Shafer G. *A Mathematical Theory of Evidence*. New Jersey: Princeton University Press; 1976.

[14] Dempster AP. Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Stat.*, 1967; **38**: 325-339.

[15] Dempster AP. A Generalization of Bayesian inference. *Journal of the Royal Statistical Society*, 1968; **B 30**: 205-247.

[16] Yager R. Arithmetic and other operations on Dempster-Shafer structures. *International Journal of Man-Machine Studies*, 1986; **25**: 357-366.