

A Hybrid Knowledge-based Prediction Method for Avian Influenza Early Warning

Jie Zhang, Jie Lu and Guangquan Zhang

Faculty of Engineering and Information Technology, University of Technology, Sydney
PO Box 123, Broadway, NSW 2007, Australia
{jiezhang; jielu; Zhangg}@it.uts.edu.au

Abstract—High pathogenic avian influenza remains rampant and the epidemic size has been growing in the world. The early warning system (EWS) for avian influenza becomes increasingly essential to militating against the risk of outbreak crisis. An EWS can generate timely early warnings to support decision makers in identifying underlying vulnerabilities and implementing relevant strategies. This paper addresses this crucial issue and focuses on how to make full use of previous events to perform comprehensive forecasting and generate reliable warning signals. It proposes a hybrid knowledge-based prediction (HKBP) method which combines case-based reasoning (CBR) with the fuzzy logic technique. The method can improve the prediction accuracy for avian influenza in a specific region at a specific time. An example is presented to illustrate the capabilities and procedures of the HKBP method.

Keywords—Case-based reasoning, knowledge-based systems, fuzzy logic, avian influenza, early warning systems

I. INTRODUCTION

Early warning systems (EWSs) have developed into obligatory requirements worldwide. According to the results presented by Ebi and Schmier [1], an EWS includes the four principal components: identification and forecasting of the event; prediction of the possible health outcomes; an effective and timely response plan; and an ongoing evaluation of the system and its elements. The prerequisites for EWSs are to provide effective predictions of infectious diseases among which avian influenza is hard to forecast and has devastating damages in human and other animals for its virulence.

The highly pathogenic strains of avian influenza A H5N1 has dispersed throughout Asia first and then to Europe and Africa [3]. By the end of 2008, totally 393 people had been confirmed to be infected by the H5N1 virus, and among them 248 died. (http://www.who.int/csr/disease/avian_influenza/country/cases_table_2009_01_19/en/index.html). Birds had been culled worldwide and this has led to billions of poultry trade loss. Researchers in the area have been trying to explore the nature of this virus, observing and predicting the trends and discovering the effective vaccine against this epidemic disease. However, there are some difficulties conducting and finding effective solutions because of the following reasons: (1) The virus is special and hard to extinct [4]. The virus is varied by the hosts and its origin and has been changing continuously. Also the virus can cross different species, for example, from birds to mammals, but it is lucky to believe that it cannot spread among mammals [4]. (2) Poultry farm, bird trade and wild birds' migration make it spread through all the possible channels [3] [4] [5]. (3) Some other

undiscovered reasons also contribute to these difficulties [3], such as the unreported and illegal trade of both wild birds and domestic poultry will shield the truth. These reasons have sparked a timely need for EWSs which can monitor and track new, unknown diseases, produce signals to detect possible crises at an early stage [6]. The problem is how to take advantage of existing cases, experiences and knowledge to improve the prediction and convey reliable early warnings to decision makers and the society.

This paper proposes a hybrid knowledge-based prediction (HKBP) method to solve this issue, which can overcome the predicting difficulties by making full use of previous cases and existing knowledge. This proposed HKBP method combines case-based reasoning (CBR) techniques with fuzzy logic to make the prediction more effective and the results more accuracy. The rest of the paper is organized as follows. Section II reviews related research in this area. Section III describes the HKBP method in details. Section IV uses an example to illustrate the capability of this HKBP method and finally conclusions and discussions are in Section V.

II. RELATED WORK OF EPIDEMIC DISEASE PREDICTION

Epidemic disease prediction methods vary in different virus transmission ways. Some viruses can transmit among people directly, but others are not, such as the avian influenza virus. Various prediction methods have been reported in literature in the epidemic diseases warning research area. This section will discuss these research methods in details.

A. Transmission of avian flu viruses and its prediction

Avian flu A viruses are mostly transmitted among birds, pigs, and occasionally from birds to humans. The highly pathogenic avian influenza viruses H5 or H7 are widespread among birds, mammals and humans [4]. It was believed that wild birds migration is the main reason of the virus dispersal, but new evidences were found that human activities of farming and marketing of poultry are highly related to the spread of the influenza A viruses [4]. Waterfowl whether wild or human raised are key factors of the avian influenza ecology [4]. Avian influenza viruses are transmitted by birds including both wild birds and farm poultry. Because birds share with human the same capacities of travelling a long distance within short time, so it makes this prediction problem more complicated. Wild bird migration routes and birds number data are aggregated and analyzed to predict possibility of virus spreading in Western Mediterranean Wetlands where wild birds can move from different places such as central Asia, Northern and Eastern Europe, western Africa [7]. Other findings show that short

distance migration is more important than the long distance migration for the ecologic immunology [5]. But Kevin Winker [8] discovers evidences that the risk of virus transmitted by wild birds is low. Some research findings show that poultry farming and trading activities are the major factors responsible for the virus spreading [4]. Others show that waterfowl is the key element of the virus spread and in Southeast Asia free-grazing ducking is the major risk of persisting and transmitting the virus [9]. Also in specific regions, the factors making major contributions to the outbreak are special [9]. It is natural that researchers try to apply all these factors to predict the spread of the H5N1 avian influenza [3].

Both overt factors and covert factors lead to the wide spread of avian influenza viruses [3]. In order to fully address the prediction issue, both two factors must be investigated closely. Some findings are evidenced by researchers, which can be developed into expert rules; others are history events which may have potential knowledge can be transformed into case base for the CBR method.

B. Case-based reasoning method in EWSs

There have been a variety of investigations of EWS on epidemic diseases. ProMED system [10] applied knowledge discovery, data mining and CBR techniques to help forecasting epidemics spread. TeCoMed [6] is another epidemic EWS which can provide communicable disease early warning. Among the AI techniques used in these systems, the CBR seems more powerful and promising as a case-based knowledge acquisition solution. Though many factors will affect the outbreak of the avian flu and some of the reasons are not clear yet, we can still record the events already happened as evidences for the future prediction.

Case-based reasoning employs past experiences which are cases stored in the case base to support the current decision making with the similar situations. This technique is most suitable for the medical decision-making [11] because medical decision makers often apply their previous experience to reason and judge new cases and a CBR system can keep knowledge in its case base. The CBR cycle consists of four steps [12]: (1) retrieving former similar cases, (2) reusing similar cases solutions as references, (3) revising the solutions make it adapt to the new case and (4) retaining new learned cases.

The advantage of CBR is its knowledge acquisition capabilities which can resolve the undiscovered knowledge problem. However, many cases contain uncertain values and uncertain risk levels; therefore, we will introduce fuzzy sets and fuzzy rules created from former case findings to solve these problems.

III. A HYBRID KNOWLEDGE-BASED PREDICTION METHOD

This section will present the HKBP method. This HKBP method contains two main components: an CBR component for avian influenza and rule-based knowledge generation component. The HKBP method contains four main steps.

A. Knowledge body for avian influenza

We first provide the basic knowledge of the factors that can affect the avian influenza virus spread and the case-base structure.

1) *Factors causing the outbreak* : There are many factors which can lead to the disease outbreaks. Some of them we have already identified, but still many other covert factors need to be discovered in the future. Fig. 1 highlights main factors that can affect the virus transmission and outbreak in birds. These factors fall into five groups: wild birds, farm poultry, bird trade, climate and region factors.

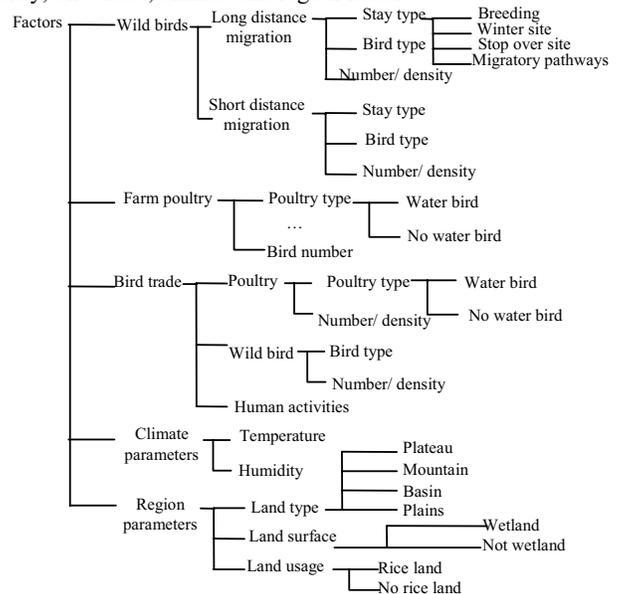


Figure 1. factors associated with the outbreak

2) *knowledge base structure* : One basic part of knowledge is the case base which is the database file of all the bird outbreak events happened before. In the file the attributes are outbreak-time, outbreak-region, region-weather, region latitude, region geographical properties (plains, plateau, mountain, basin), region surface properties (wetland or not), region poultry density, region transport condition, distance to the nearest cities outbreak level and etc. In addition to this basic data, there are database files about poultry trade information, wild bird migratory information, rice land usage information. These basic attributes can be induced from the basic invent data file, and other information can be calculated from the other three data files. The other important part is knowledge base which consists of fuzzy rule sets which reflect the expert knowledge about the bird flu transmission. We need to apply the CBR method and rule based method together to obtain better prediction and a more accurate early warning.

B. HKBP method processes

HKBP method has two main components and four steps. *Step 1 Problem information preprocessing*: In order to predict avian influenza spread trend in a particular area and at a particular time, we need first collect some data about this particular area at that time. We then transform non numerical data into numerical values and change them into format for the CBR case matching. For fuzzy rule applying, we transform

some case parameters (e.g. distance of the nearest infected farm) from numerical values to fuzzy values such high, low, near, far and others for the prerequisites of the rules.

Step 2 Using CBR technique to find the best matches to the new case: This is a typical CBR process. First, we transform and normalize all the data and then follow the 4R cycle, retrieve, reuse, revise and retain. Details of the process are as follows:

a) *Data preprocessing and normalization:* The new case data must be preprocessed. It include two steps: the first is to change some linguistic data, such as normal data (e.g. plateau) and ordinal data, into numerical data; the second step is to normalize the data into [0,1] interval.

b) *Case retrieval:* After the preprocessing and normalization, the casebase is retrieved to find the most matching cases. We calculate the distance of case i and case j by

$$DC_{i,j} = \frac{\sum_{k=1}^n w_k^c \frac{|A_{i,k} - A_{j,k}|}{\max(A_k) - \min(A_k)}}{\sum_{k=1}^n w_k^c}, \quad (1)$$

where $A_{i,k}$ is i_{th} case's k_{th} attribute, w_k^c is weight of the k_{th} attribute of the case when calculate the distances, $\max(A_k) - \min(A_k)$ is the largest scope of the k_{th} attribute.

c) *Case reuse and revision:* A matched case is the case which has the distance less than a small value ε . The value of ε can be calibrated based on specific requirements of a real world situation. All the matched cases can be reused, the final result $Risk^c$ is aggregated from each fuzzy number risk level \tilde{R}_i^c which is the risk level of i_{th} matched case among all n matched ones. We have five risk levels, each risk level belongs to a set of linguistic terms which can be described by triangular fuzzy numbers, such as normal (N), low risk (LR), middle (M), high risk (HR), very high risk (VH) as shown in Fig. 2.

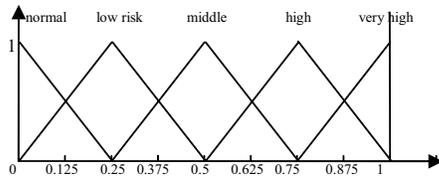


Figure 2. Fuzzy membership functions of risk levels

If we have n matched cases, then the final aggregated risk level can be calculated by

$$\tilde{Risk}^c = \frac{\sum_{i=1}^n (1 - DC_{i,new}) \tilde{R}_i^c}{\sum_{i=1}^n (1 - DC_{i,new})} \quad (2)$$

where $DC_{i,new}$ is the distance between the i_{th} matched case and the new case. We can notice that \tilde{Risk}^c is also a fuzzy number.

d) *Case retain:* Normally if the prediction is finally confirmed, then the new case can be retained in the case base. Besides the cases can be obtained from real happened events. The CBR method can be calibrated by changing weights in (1)

and the matching value ε . These parameters can be adjusted dynamically according to the newly obtained real world data. In addition, we can seek to other valuable expert knowledge, which presents as patterns and rules.

Step 3: Applying the patterns/rules: First each rule set's prerequisites are used to find out if the new case parameters fit or not. Sometimes one case can fit for more than one rule set. If so, all rule sets which can match up the cases are applied and final risk level is aggregated. After this aggregation, another risk value from fuzzy rules is created.

a) *Rule-based knowledge generation:* There are lots of findings about the HPAI avian influenza outbreaks. The virus can spread based on the poultry [13], by the duck during the rice cultivation [14] or maybe through the migration of birds [15, 16]. It is of great significance to find out the way of integrating all these patterns or rules to the prediction of an outbreak. Now we present the brief version of some typical patterns and rules as follows.

b) *Poultry transmission pattern and rules:* The pattern are represented by rule set which can be divided into two parts. The first part consists of a set of the preconditions of fixed values or fixed interval values, which are the prerequisites to the applying conditions of the rules. If the contexts of a new case are the same or similar as the first set of preconditions, then this rule is appropriate for the new case. The second part consists of a set of fuzzy rules in which there are two most important conditions of the poultry density and distance of the nearest infected farm and the result of the risk level. The fuzzy rules are shown in Fig. 3. Before reasoning, numeric values will be changed to fuzzy values by fuzzy membership functions. For example, the distance between two farms can be defined as in Fig. 4. In second part of this rule, if the poultry density and the farm distance are known then the final risk level can be predicted by fuzzy logic.

```

If
Region. Latitude= [N50, N55];
Region. Month= [2-3];
Region. Temperature=[0-8];
Region. Land_type= plain;
And if
Region. Poultry_density= X1 (very high/high/medium /low /very low);
Region. Distance-to-kernel= X2 (near/medium/far);
Then
The outbreak risk = R (very high/high/medium/low/very low)

```

Figure 3. Poultry density rule set

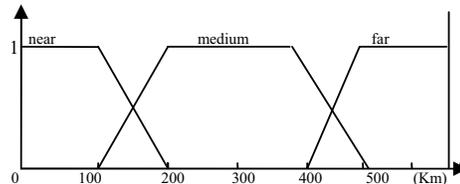


Figure 4. Fuzzy membership function of distance

c) *Duck and rice cultivation pattern and rules:* This pattern indicates that in some regions the free ranged water poultry such as ducks are the main reason responsible for the outbreak of the virus H5N1 during the rice harvest season. Just the same structure as the first rule set, this one also has two parts. In order to apply the rules that the first part of the prerequisites must be fulfilled and then according to the land cultivate status and the density of the duck in this area the final risk can be deduced.

```

If
Region. latitude= [N10, N20]
Region. month=[12,01];
Region. temperature=[20-30];
Region. surface= wetland;
Region. usage = rice_land;
And if
Region. Land_usage.status = X1 (harvest/no harvest);
Region. Poultry_density.free_water_foultry_density = X2 (very
high/high/medium /low /very low);
Then
the outbreak risk =R (very high/high/medium /low /very low).

```

Figure 5. Free range duck rule set

d) *Wild birds migratory Pattern and rules:* This pattern is more complicated. Firstly, the preconditions must be fulfilled just the same as previous rules. And then the wild bird status must be investigated in detail such as winter sites, stopover sites, migratory pathway and etc. From this information, the interval of the birds stay can be estimated. The longer birds stay and the higher poultry density is, the more possible the infectious interface can be established. The density of the local poultry will also affect the final result.

```

if
Region. latitude= [N32-N38]
Region. month=5;
Region. temperature=[0-5];
Region. surface= wetland;
Region. Nearest_migratory_place. Status = true
And if
Region. Nearest_migratory_place. distance =X1(near/medium/far);
Region. Nearest_migratory_place. Stay_interval = X2 (long/ short );
Region. Nearest_migratory_place. Wild_bird_density=X3 (very
high/high/medium/low/very low);
Region. Poultry_density = Y (very high/high/medium/low/ very low);
Then
The outbreak risk = R (very high/high/medium/low /very low)

```

Figure 6. Wild bird migratory rule set

e) *Rules aggregation:* In avian influenza prediction research, factors affecting the final results are complicated. Some of the rules or patterns play essential roles in one particular environment, while others take great effect in some other environments. The reason lies in the transmission mechanism of avian influenza virus is not fully discovered. When a new case appears, even though the preconditions are much similar as the rule applying conditions, we should also consider that maybe some other rules also function under the same context. While this concern is reasonable, there are some limits need to be considered. That is to say, if the precondition

difference between the rule set and the case is great enough, then it is not necessary to apply the rule set. We must set some measures to observe how much the rule preconditions and the case parameters can match each other.

Suppose we have p_i preconditions in the i_{th} rule set, among them there are q_i numeric scope values and $p_i - q_i$ single values. Each rule set's precondition has its corresponding new case attribute value. The distance between the case parameters and the preconditions of the i_{th} rule are calculate by (3).

$$DR_{i,new} = \frac{\sum_{k=1}^{q_i} \frac{w_k^i \text{dist}_k^i}{\max(A_k) - \min(A_k)} + \sum_{k=q_i+1}^p \frac{w_k^i |A_{i,k} - A_{new,k}|}{\max(A_k) - \min(A_k)}}{\sum_{k=1}^p w_k^i} \quad (3)$$

$$\text{dist}_k^i = \begin{cases} 0, & A_{new,k} \in [r \min_{i,k}, r \max_{i,k}] \\ \text{Min}(|r \min_{i,k} - A_{new,k}|, & A_{new,k} \notin [r \min_{i,k}, r \max_{i,k}] \\ |r \max_{i,k} - A_{new,k}|) & \end{cases} \quad (4)$$

where w_k^i represents the i_{th} rule set's k_{th} precondition' weights. $A_{i,k}$ is the i_{th} rule set's k_{th} precondition's value and $A_{new,k}$ is the k_{th} numeric attribute value of the new case relevant of $A_{i,k}$. $\max(A_k) - \min(A_k)$ is a constant, which means the scope of the k_{th} attribute value. $[r \min_{i,k}, r \max_{i,k}]$ is the i_{th} precondition's value scope of the k_{th} rule set. If a binary rule set precondition value is nominal, we transform it into binary number and calculate its normalized Hamming distance.

For each rule set, $DR_{i,new}$ represents how much the new case parameters fit for preconditions of this rule set. Here we induce a threshold θ . If $DR_{i,new} \leq \theta$, then we apply the rule set. Otherwise, the new case is not appropriate for applying the rule set.

When we identify which rule set are appropriate for the new case, some additional information for applying these rules must be obtained, such as migratory information when applying rule three. When all the information is obtain, we can use second part of the rule set to deduce the result. After applying a fuzzy rule set we can obtain a group of weighted fuzzy numbers of risk level (e.g. $\{0.37/N, 0.63/LR, 0/M, 0/H, 0/VH\}$). These fuzzy numbers are integrated by (5) to generate a fuzzy number which represents the risk level deduced from this rule set.

$$\tilde{R}_i^r = \frac{\sum_{i=1}^s \lambda_i \tilde{R}_i^r}{\sum_{i=1}^s \lambda_i}, \quad (5)$$

where λ_i is the i_{th} λ cut value of fuzzy number \tilde{R}_i^r .

Finally we aggregate all these rule set's results. Actually, we obtain risk level value represented by a fuzzy number from each rule set. If finally m results are obtained, then they can be aggregated by (6).

$$\tilde{Risk}^r = \frac{\sum_{i=1}^m (1 - DR_{i,new}) \tilde{R}_i^r}{\sum_{i=1}^m (1 - DR_{i,new})}, \quad (6)$$

where \tilde{R}_i^r is the risk value of i_{th} rule's risk value. The final risk value \tilde{Risk}^r , is also a fuzzy number.

Step 4 Synthesis the results: The final result can be obtained by synthesize the outcomes of Step 2 and Step 3 by

$$\tilde{Risk} = Max(\tilde{Risk}^r, \tilde{Risk}^c), \quad (7)$$

where \tilde{Risk}^r is risk level of rules outcome and \tilde{Risk}^c is risk level from matching cases. For safe reason, we only obtain the maximum value of the two as the final risk level. In order to find out which risk level is the nearest, we can apply (8) to calculate

the distance between final risk fuzzy number and each risk level [17],

$$d(\tilde{a}, \tilde{b}) = \sqrt{1/3[(a_0^- - b_0^-)^2 + (a - b)^2 + (a_0^+ - b_0^+)^2]} \quad (8)$$

where a_0^- , a , a_0^+ are the left, middle and right points of a triangular fuzzy number.

IV. AN AVIAN INFLUENZA EARLY WARNING EXAMPLE

Now we use an example to show the application of this proposed hybrid knowledge-based prediction method. Related data is shown in the Table I. We suppose there are four cases in the case base in the example. We will illustrate that when a new case arrives, how the proposed method conducts reasoning to identify if this case is with a risk, and if yes, what is its risk level.

TABLE I THE CASES EXAMPLES

region_id	region_type	region_surface	region_usage	latitude	temperature	humidity	poultry_density	nearest_farm	nearest_breakout_city	transportation_easily	poultry_trade	wild_bird_trade	wild_bird_migration	Risk level
1	plain	wetland	rice land	N40	0	70	3000	100	500	yes	500	0	no	H
2	basin	not wetland	not rice land	N30	1	60	2000	300	300	yes	400	0	no	VH
3	mount	not wetland	not rice land	N20	25	80	3000	100	300	no	600	0	no	M
4	plateau	wetland	not rice land	N35	5	35	100	900	900	yes	100	0	no	N
x	plateau	wetland	not rice land	N37	10	30	20	1000	1000	yes	20	0	yes	?
1	11	1	1	40	0	70	3000	100	500	1	500	0	0	H
2	10	0	0	30	1	60	2000	300	300	1	400	0	0	VH
3	01	0	0	20	25	80	3000	100	300	0	600	0	0	M
4	00	1	0	35	5	35	100	900	900	1	100	0	0	N
x	00	1	0	37	7	30	20	1000	1000	1	20	0	1	?
Scope	2	1	1	90	80	100	7000	1000	1000	1	2000	1	1	
weight	1	1	1	1	1	1	1	1	1	1	1	1	1	

Step1 Problem information preprocessing: When the new case information is obtained, we need to do the transformation. Firstly, all the ordinal, nominal data must be transformed into the numerical ones. For example, the land type can be transformed into binary variables shown in Table I. The logical variable all transform yes to number "1" and no to "0". After the transformation, Hamming distance are used to calculate the binary variables.

Step2 Using CBR technique to find the best matches: After transformation, we can use (1) to calculate distances between the new case and all the old cases. Here we present all the distance values between the new case and the old cases. The results are shown in Table II. The distance of case 4 and the new case is 0.11 and if we use the $\epsilon = 0.15$, so the $DC_{4,new} < \epsilon$. we use case four's risk level to conclude the new case's level is $\tilde{Risk}^c = (0,0,0.25)$, that is to say, the risk level is "normal".

TABLE II. DISTANCES AMONG DIFFERENT CASES

Case ID	1	2	3	4	x
x	0.47	0.40	0.57	0.11	0.00

Step 3 Applying the patterns/rules: First, we calculate the distance between preconditions of the rules and the case parameters. The brief information is listed in Table III. Only Rule 1 and Rule 3 are satisfied the rule applying conditions. During the process, additional data of migratory birds is gathered and we can find there is wild bird breeding place in only 100km, moreover the bird will stay there for a month, and wild bird density is {0.67/high, 0.33/very high}. After obtaining enough data, we can apply Rule 1 (Table. IV) and Rule 3 (Table V) directly to obtain the risk value. After calculating the final results are shown in Table VI. $\tilde{Risk}^r = (0, 0.12, 0.37)$, it means that aggregated risk level value by rules deducing is not "normal".

TABLE III. RULES PRECONDITIONS AND CASE PARAMETERS

no	region_type	region_surface	region_usage	latitude	Month	temperature	wild_bird_migration
r1	plain	-	-	50-55	2-3	0-8	-
r2	plain	wetland	rice land	10-20	12-01	20-30	-
r3	plateau	wetland	-	32-38	5	0-5	y
x	plateau	wetland	not rice land	N37	4	7	y

TABLE IV. RULE SET ONE

		poultry density				
		v low 2/3	Low1/3	medium	high	very high
distance to farm	near	LR	M	HR	VH	VH
	middle	LR	M	M	HR	HR
	far(1)	N(2/3)	LR(1/3)	LR	M	M

TABLE V. PART OF RULE SET THREE

rule id	distance to w-bird	interval	w-birds density	poultry density	risk level
1	near(1)	long(1)	(2/3) high	very low(2/3)	N(2/3)
2	near(1)	long(1)	(1/3)very high	very low(2/3)	LR(1/3)
3	near(1)	long(1)	High (2/3)	low(1/3)	LR(1/3)
4	near(1)	long(1)	(1/3)very high	low(1/3)	LR(1/3)

TABLE VI. RESULTS OF APPLYING FUZZY RULES

rule no	$DR_{i,new}$	$\Theta=0.3$	\tilde{R}_i^r (λ -cut)	\tilde{R}_i^r (vertices)
rule1	0.25	y	[0.08 λ , 0.33-0.25 λ]	(0,0.08,0.33)
rule2	0.42	n	-	
rule3	0.02	y	[0.15 λ , 0.40-0.25 λ]	(0,0.15,0.4)
w1r1+w3r3			[0.12 λ , 0.37-0.25 λ]	(0,0.12,0.37)

Step 4 Synthesis the results: Since $\tilde{Risk}^r \geq Risk^c$, we use $\tilde{Risk}^r = (0, 0.12, 0.37)$ as final result. The distances between \tilde{Risk}^r and the five risk levels can be calculated and the results are listed in Table VII. The distances between the new case risk value and N LR are similar $|0.099-0.105| < 0.01$. Therefore we can choose the higher risk level to describe the situation as LR.

TABLE VII. DISTANCES TO THE RISK LEVEL

R-Level	N	LR	M	HR	VH
Fuzzy number	0,0,0.25	0,0.25,0.5	0.25,0.5,0.75	0.5,0.75,1	0.75,1,1
d	0.099	0.105	0.341	0.589	0.760

V. CONCLUSIONS AND FURTHER STUDY

To predict future trends and evaluate the risk level of avian influenza is really a complex task. There are many transmission drivers to be identified and many virus dispersal laws created from different assumptions. In this paper, we advocate adopting combining CBR method and fuzzy expert rules together to overcome these difficulties.

This proposed HKBP method has three advantages: (1) making full use of the previous cases data and expert rules altogether to improve the effectiveness of the prediction; (2) combine the CBR and the fuzzy logic method to integrate the whole knowledge body in a warning system; and (3) great ease of expanding both kind of knowledge by simply adding cases and rules.

Our further study includes collecting more events and generating more rules to finalize the proposed method. The system is under developing. The data we used in this paper are fractions of real data of previous events from Europe and Asian, and rules are brief versions from findings of real epidemic researchers. So we hope to implement the system and perfect our method and idea in the future.

ACKNOWLEDGMENT

The work presented in this paper was supported by Australian Research Council (ARC) under the discovery project DP088739.

REFERENCES

- [1] K. L. Ebi and J. K. Schmier, "A Stitch in Time: Improving Public Health Early Warning Systems for Extreme Weather Events," *Epidemiol Rev*, vol. 27, pp. 115-121, 2005.
- [2] A. T. Peterson and R. A. J. Williams, "Risk mapping of highly pathogenic Avian Influenza distribution and spread," *Ecology and Society* vol. 13, 2008. <http://www.ecologyandsociety.org/vol13/iss2/art15/>
- [3] A. M. Kilpatrick, A. A. Chmura, D. W. Gibbons, R. C. Fleischer, P. P. Marra, and P. Daszak, "Predicting the global spread of H5N1 avian influenza," *Proceedings of the National Academy of Sciences*, vol. 103, pp. 19368-19373, 2006.
- [4] M. GAUTHIER-CLERC, C. LEBARBENCHON and F. THOMAS, "Recent expansion of highly pathogenic avian influenza H5N1: a critical review," *IBIS*, Vol. 149, issue 2, pp. 202-214, 2007.
- [5] TP Weber, NI Stilianakis, "Ecologic immunology of avian influenza (H5N1) in migratory birds" *Emerg Infect Dis*. vol. 13, no. 8, pp.1139-1143, 2007.
- [6] L. Gierl and R. Schmidt, "Geomedical warning system against epidemics," *Intl. J. of Hygiene and Environmental Health*, vol. 208, pp. 287-297, 2005.
- [7] E Jourdain, M. Gauthier-Clerc, DJ Bicout, P. Sabatier. "Bird migration routes and risk for pathogen dispersion into western Mediterranean wetlands," *Emerg Infect Dis*. 2007 [viewed on 30/12/2008]. Available from <http://www.cdc.gov/EID/content/13/3/365.htm>
- [8] K. Winker, KG McCracken, DD Gibson, CL Pruet, R Meier, F Huettmann, et al. Movements of birds and avian influenza from Asia into Alaska". *Emerg Infect Dis* [serial on the Internet]. 2007 [viewed on 30/12/2008]. Available from <http://www.cdc.gov/EID/content/13/4/06-1072.htm>
- [9] M. Gilbert, X. Xiao, D. U. Pfeiffer, M. Epprecht, S. Boles, C. Czamecki, P. Chaitaweesub, W. Kalpravidh, P. Q. Minh, M. J. Otte, V. Martin, and J. Slingenbergh, "Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia," *Proceedings of the National Academy of Sciences*, vol. 105, pp. 4769-4774, 2008.
- [10] M. Bull, G. Kundt, and L. Gierl, *Discovering of health risks and case-based forecasting of epidemics in a health surveillance system*, 1997
- [11] R. Schmidt, S. Montani, R. Bellazzi, L. Portinale, and L. Gierl, "Case-Based Reasoning for medical knowledge-based systems," *International Journal of Medical Informatics*, vol. 64, pp. 355-367, 2001.
- [12] A. Aamodt and E. Plaza, "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches," *AI Communications*, vol. 7, pp. 39-59, 1994.
- [13] G. J. Boender, T. J. Hagenaars, A. Bouma, G. Nodelijk, A. R. W. Elbers, M. C. M. de Jong, and M. van Boven, "Risk Maps for the Spread of Highly Pathogenic Avian Influenza in Poultry," *PLoS Comput Biol*, vol. 3, p e71, 2007.
- [14] M. Gilbert, X. Xiao, P. Chaitaweesub, W. Kalpravidh, S. Premasithira, S. Boles, and J. Slingenbergh, "Avian influenza, domestic ducks and rice agriculture in Thailand," *Agriculture, Ecosystems & Environment*, vol. 119, pp. 409-415, 2007.
- [15] Gilbert M, Xiao X, Domenech J, Lubroth J, Martin V, Slingenbergh J. Anatidae migration in the western Palearctic and spread of highly pathogenic avian influenza H5N1 virus. *Emerg Infect Dis*. 2006. Available from <http://www.cdc.gov/ncidod/EID/vol12no11/06-0223.htm>
- [16] Winker K, McCracken KG, Gibson DD, Pruet CL, Meier R, Huettmann F, et al. Movements of birds and avian influenza from Asia into Alaska. *Emerg Infect Dis*. 2007 Apr Available from <http://www.cdc.gov/EID/content/13/4/06-1072.htm>
- [17] C.-T. Chen, "Extensions of the TOPSIS for group decision-making under fuzzy environment," *Fuzzy Sets and Systems*, vol. 114, pp. 1-9, 2000.