A Web Based Belief Rule Based Expert System to Predict Flood

Raihan Ul Islam  
Department of Computer Science, Electrical and Space Engineering  
Luleå University of Technology  
SE-931 87 Skellefteå, Sweden  
raihan.ul.islam@ltu.se

Karl Andersson  
Department of Computer Science, Electrical and Space Engineering  
Luleå University of Technology  
SE-931 87 Skellefteå, Sweden  
karl.andersson@ltu.se

Mohammad Shahadat Hossain  
Department of Computer Science and Engineering  
University of Chittagong  
University-4331, Bangladesh  
hossain_ms@cu.ac.bd

ABSTRACT
Natural calamity disrupts our daily life and brings many sufferings in our life. Among the natural calamities, flood is one of the most catastrophic. Predicting flood helps us to take necessary precautions and save human lives. Several types of data (meteorological condition, topography, river characteristics, and human activities) are used to predict flood water level in an area. In our previous works, we proposed a belief rule based flood prediction system in a desktop environment. In this paper, we propose a web-service based flood prediction expert system by incorporating belief rule base with the capability of reading sensor data such as rainfall, river flow on real time basis. This will facilitate the monitoring of the various flood-intensifying factors, contributing in increasing the flood water level in an area. Eventually, the decision makers would be able to take measures to control those factors and to reduce the intensity of flooding in an area.

Categories and Subject Descriptors
H.4 [Information Systems Applications]: Decision support

General Terms
Design, Measurement

Keywords

1. INTRODUCTION
Even in this twenty first century after so many technological innovations humans are helpless in the hand of natural disasters. There are different natural disasters like, floods, volcanic eruptions, earthquakes, and tsunamis. Flood is considered as the most catastrophic among the other natural disasters [1]. Flood causes the highest number of fatalities and greater economic damage in comparison to other natural disasters.

As for example, a devastating flood occurred in China during 1931 caused by overflowing of water from Huang He (Yellow River) River. This flood is considered as one of the most destructive flood occurred in the world. Eighty-eight thousands square km of land completely inundated, eighty million people were homeless and around four million people died due to this flood[1]. According to Geoscience Australia flood is a general and temporary condition of partial or complete inundation of normally dry land areas from overflow of inland or tidal water from the unusual and rapid accumulation or runoff of surface water from any source [2]. Predicting flood will help us to take necessary steps for human evacuation and other entities. Several types of data are used for predicting floods. These are: the amount of rainfall, rainfall duration, the rate of change in river flow, river water level, the characteristics of a river's drainage basin and human activities. Some of these data are quantitative in nature and others are qualitative in nature. Hence, we need an integrated framework, which is able to process both qualitative and quantitative data in a single integrated framework. In this paper, a belief rule based expert system has been used which has the capability to process both qualitative and quantitative data in a single integrated framework to predict flooding in an area.

Sensors can be used to automatically collect different types of environmental data necessary for predicting flood and transmit these data to central system. Nowadays, due to the cost efficiency and protocol standardization [3] low-powered sensors are easily deployed in large scale for different systems. We can collect data for different environmental parameters like rainfall, water level, humidity and temperature by using different types of sensors. An efficient heterogeneous wireless sensor network (WSN) is needed for collecting and transmitting data as sensors are deployed in harsh environment. In [4], we proposed an efficient heterogeneous WSN network combination of ZigBee, Ethernet, Wi-Fi connectivity and SMS suitable for flood prediction.

Knowledge base, inference engine and user interface are the most essential components of an expert system. There are different types of uncertainty like ignorance, incompleteness, ambiguity, vagueness and imprecision in data. These uncertainties can be caused by faulty sensor data or lack of human knowledge. For the design and implementation of expert systems, it is necessary to consider uncertainty. Belief Rule base knowledge representation schema captures all the above type of uncertainty [5]. Other knowledge representation schema such as Propositional Logic (PL) and First-order Logic (FOL) are not capable of capturing uncertainty. Fuzzy logic is capable of handle uncertainty due to ambiguity and vagueness but not the others. Forward Chaining and Backward Chaining are two common methods used as the

\[ http://global.britannica.com/ \]
inference mechanisms. However, they are not equipped to handle uncertainty of data. Evidential Reasoning (ER) is used as the inference methodology in the Web based Belief Rule Based Expert System as it can efficiently handle different types of uncertainty mentioned above.

Web-based applications are easily accessible for the users as the users do not need to install any additional software. Therefore, a web-based expert system will allow users to easily access and use the system. Web-based applications are deployed and maintained in one location usually. Web services provide a standardized way of communicating between two applications over the Internet Protocol using Extensible Markup Language (XML), Simple Object Access Protocol (SOAP), Web Service Definition Language (WSDL) and Universal Description, Discovery, and Integration (UDDI). There are two major groups of web services. One is Representational State Transfer (REST)-compliant Web services; another is Simple Object Access Protocol (SOAP) based web services. They provide the flexibility of communicating with different sensor aggregation platforms. This will allow other applications to communicate with expert systems using the aforementioned protocols. Therefore, developing a web-based expert system will allow user accessibility, easy deployment and maintenance of applications. We prefer to use REST compliant web services as they are more lightweight than SOAP and easier to implement by PHP and JavaScript. REST supports to provide output in Command Separated Value (CSV), JavaScript Object Notation (JSON), and XML. SOAP supports XML only. REST also does not use WSDL or UDDI, which makes it simpler.

Previously, we developed a Belief Rule Based Expert System (BRBES) by using Visual Basic and Microsoft SQL [4]. The application was running on a single PC. Limited people have access to it and in addition, pes had limited computational power. For these reasons, it was necessary to upgrade the expert system into a web service based application. We like to name the new application as Web-BRBES. This will allow to run the application on a server with more computational power and memory, easy access for users, communicate with more sensor platforms and allow other applications to get data from this application. The previous system was not equipped to use sensor data rather it used manual data collected from the field. Therefore, the real time observation of the natural phenomenon (such as effect of intense rainfall in an area) is difficult to observe in that version.

The remainder of this article is structured as follows: Section 2 covers related work on different expert Systems for flood prediction, Section 3 provides a brief overview on Belief Rule Base Expert System; Section 4 describes the Web-based Belief Rule Base Expert System. Section 5 presents the results and discussion, while Section 6 concludes the paper and indicates future work.

2. RELATED WORK

Flood prediction is becoming more accurate and faster by using WSN technologies [7][8][9][10][11][12][13]. Ahmad et al. [7] performed a detail analysis of flood prediction techniques based on GIS using Ad hoc wireless Sensor Network Architecture. They used ArcGIS simulation tool for pre and post risk analysis of flood disaster. In the paper, authors proposed a mathematical model for flood risk analysis and prediction based on the rate of change of different environmental values. In their proposed model they put emphasis on input parameters like soil moisture, air pressure, direction of wind, humidity and rainfall. This model is also useful for rescue and emergency authorities to take necessary action to save life before the critical conditions occurs. They also used underwater wireless sensors for monitoring underwater seismic activities.

Seal et al. [8] described a real time flood-forecasting model. Simple and fast calculation and cost effective implementations are the main principle of this model. Proposed model is independent of number of parameters. Therefore, parameters can be added or removed based on the on-site requirement. It also considered the anomaly in sensors data and storage capacity of the nodes and provided techniques to cope with these situations, but the proposed system needs to be deployed in real life. The proposed system uses multiple linear regression equations. Hence, it may contain uncertainty due to randomness.

Khedo [9] also proposed a system for real-time monitoring of the hydrological conditions of rivers for flood forecasting using Wireless Sensor Networks. Their system is called RTFMS. The system monitors water level, water flow and rainfall, in the region of the Rivière du Rempart of Mauritius. Water depth is monitored using a hydrostatic level sensor. A hydrostatic level sensor will be used to measure the pressure created by a liquid. Using this sensor software components provide warnings when the water level of the riverbank approaches critical. The rainfall sensor in RTFMS will perform measurement using Reed magnetic switches, which cause an interrupt after every 1 mm of rainfall. A submersible area velocity sensor will be used to measure average velocity directly, without the need for time-consuming and costly flow profiling. The sensor will need to be sealed to withstand submergence and prolonged surcharge conditions. Twenty-Two sensors were deployed along 2 km of the river. These sensors were divided into three clusters that are connected to a gateway. The gateway node receives all the data from the deployed sensors, and acts as a data-logger and transmits sensor readings back to the Central Monitoring Office (CMO) using GSM network. Each mote sensor transmits data to the gateway in a multi-hop fashion at an interval of fifteen minutes, which is the maximum sampling frequency. RTFMS uses Grid-to-Grid model to predict flood. The system is not a web-based application. However, it has a web interface and it does not use any expert system. In this paper, authors did not address the anomaly of data.

Moreover, Basha et al. [10] proposed and evaluated model-based monitoring for early warning flood detection. Their proposed sensor network tiers are able to withstand the flood event and remain functional over longer periods when there is no flood. The sensor network is distributed over larger regional areas and supports different types of sensors. They also incorporated self-monitoring for failure and adapting measurement schedules.

Shukla et al. [11] described how to design an architectural model for flood monitoring using wireless sensor networks. The proposed system architecture is composed of sensor network, processing/transmission unit and a server. Based on the first two years data, the sensor-deployed region is divided into three regions. These are green, yellow and red. Green means there are no possibilities of flood in these areas. Yellow marks the areas, which have medium risk level for flood. Red marks the areas, which have high risk of flood. The sensors are programmed to act differently on different marked areas. As for example, the rates of sending data by sensors in green areas are relatively lower than the sensors in red area. All these systems focused on how to collect data using sensors but are not concerned with addressing the issues like uncertainty with the data or nor with using the data
with decision making or expert system nor with the dissemination of data in a web environment.

Furquim et al. [12] described a motivating case from Brazil where an accurate flood-forecasting model using WSNs and chaos theory was studied in a real deployment. Finally, Ishida et al. [13] proposed a real time disaster damage information sharing system for disaster countermeasures headquarters at the time of large-scale natural disaster. The municipal employees and fire corps volunteers can report the environmental damage by sending picture to the disaster countermeasures headquarters in real time, by the Post Damage Picture Smartphone Application and the Post Damage Picture Facebook Application. The paper also shows the importance of the system by conducting a survey among general populations. This paper shows an approach of using participatory sensing for disaster information sharing.

Grove [14] provides a review of several Internet based expert systems and discuss about design issues, development tools and languages. It also provides a case study named Reptile Identification Helper (RIH). Web-based expert systems usually consist of client-server architectures and web browser based interfaces. The inference engine is usually part of server. Forward and backward chaining used as inference engine for most of the expert systems. There are several concerns for scalability of Internet based expert system. One is the size and complexity of rule base. Theoretically, the process of matching rules and knowledge grows exponentially with size. Using modular knowledge programming and efficient interference engine, this problem can be addressed. Another is the growth of the user based of the Internet based expert system. Adding more processing power to the servers and more bandwidth speed can be a solution for the growth of the user base. The Internet based expert system also provides challenges for the developers to cope with new upcoming Internet based technologies. Internet based expert system provides several advantages like easy access, common interfaces and several development tools.

Mathew et al. [15] proposed a web-based expert system for diagnosis of neurologic disorders. Case and Rule based reasoning techniques were used in this expert system.

Kong et al. [16] presented a belief rule based clinical decision support system for patients with cardiac chest pain. The prototype can automatically update its knowledge base by using belief rule based learning model. One thousand patient data records were used to simulate the prototype. This paper shows an expert system based on belief rule base. The system also provides an example of using web technologies with expert system.

In our previous work [4], we proposed a new architecture for building decision support systems using heterogeneous wireless sensor networks. Using this architecture, we described a smart risk assessment system using belief-rule based decision support systems in [16] and [18] where a case study was carried out on the location-dependent service, aimed for tourists. In our new work we incorporate the use of dynamic data (like sensor data) in the expert system. For better user access and utilization of more computational power, the new system is based on layered modular system architecture with web technologies.

Ahmad et al. [7], Seal et al. [8], Khedo [9], Basha et al. [10], and Shukla et al. [11] mostly proposed different mathematical model based flood predict systems using wireless sensor networks. Among the aforementioned papers Seal et al. [8] tried to handle the uncertainty of data. None of these systems were based on expert system as well as were not fully based on web technologies. Grove [14] discussed about the pros and cons of incorporating Internet technologies to Expert System. Mathew et al. [15] and Kong et al. [16] proposed expert systems for identification of neurologic disorders and clinical decision support system for patients with cardiac chest pain. Kong et al. [16] tried to look into the uncertainty of data using Belief Rule Based system.

3. BELIEF RULE BASED EXPERT SYSTEM

Belief Rule Based Expert System (BRBES) consists of a knowledge base and an inference engine. Belief Rule Base has been used to build the knowledge base while Evidential Reasoning is considered as the inference engine. Evidential Reasoning is a multi-criterion decision analysis procedure and can handle both qualitative and quantitative data [5]. Various types of uncertainty such as ignorance, incompleteness, randomness, vagueness, and imprecision exist in sensor data that can be processed with ER and BRB [20][21].

The Belief Rule Base (BRB) is an extension of traditional IF-THEN rule base whereas a belief rule has antecedent part and consequent part. Antecedent attribute takes referential values and possible belief degrees are associated with the consequent of a belief rule. The rule weights, antecedent attribute weight, and belief degrees are knowledge representation parameters used in BRB to capture the uncertainty. A belief rule can be defined as:

\[ R_c \text{: IF } \text{Rainfall is Medium AND Rainfall Duration is High} \]

THEN Meteorological Condition is

\[ \{(\text{Severe}, 0), (\text{Moderate}, 0.4), (\text{Low}, 0.6)\} \]

In the above rule ‘Rainfall’ and ‘Rainfall Duration’ are the antecedent attributes, while ‘Medium’ and ‘High’ are the referential values. ‘Meteorological Condition’ is the consequent attribute with referential values such as ‘severe’, ‘moderate’, and ‘low’. This rule is complete because the summation of degree of belief associated with each referential value of the consequent attribute is one. If the summation is less than one then the rule is considered as incomplete, which may be due to incomplete information or ignorance. The relationship between antecedent attributes and the consequent attribute is non-linear, which is linear in case of IF-THEN rule.

The inference procedures consist of various steps including input transformation, rule activation, belief update, and rule aggregation using Evidential Reasoning approach. The task of input transformation consists of distributing the input data over the referential values of the attribute of a rule, which is called matching degree. Once the matching degree is calculated, the rules are called antecedent and they become active and reside in the short-term memory while the rule base resides in the long-term memory. The calculated matching degree is used to calculate the activation weight of each rule. It is interesting to note that each rule does not have the same weight in calculating the referential values of the consequent attribute. The summation of the rule activation weight of a rule base should be one.

When an input data for any of the leaf nodes is ignored then the belief degree associated with each rule in the rule base should be updated. This is done by the procedures mentioned in [5]. Furthermore, the aggregation of the rules is carried out by using either analytical or recursive evidential reasoning algorithm. The fuzzy value can be converted into crisp value by using the utility
score associated with each referential value to obtain the final result.

4. SYSTEM ARCHITECTURE
In this section, the system architecture is discussed in detail. The main components of Web-based Belief Rule Based Expert System (Web-BRBES) are Knowledge Base, Input Module, BRB main module, BRB UI module, Configuration Module and Knowledge Base Driver Module.

4.1 Knowledge Base Module
There are different models like propositional logic, fuzzy logic, and rule-base for representing knowledge. Most of these rules are not capable of representing the uncertainties such as imprecision or incompleteness. Therefore, we used belief rules for representing uncertain knowledge, as mentioned previously, storing the rules in a MySQL database. MySQL is the most popular RDMS system for web applications. Facebook, Twitter or Wikipedia are using MySQL [6]. MySQL is specially designed for web application. It has high-performance query engine, tremendously fast data insert capability, and strong support for specialized web functions like fast full text searches. Other features like main memory tables, B-tree and hash indexes, and compressed archive tables that reduce storage requirements by up to eighty percent. The query engine helps us to perform faster queries to get the rules. MySQL’s faster insert capabilities help us to insert new rules to the database faster. Finally, it is open source and free.

4.2 Knowledge Base Driver Module
This module is consisting of different database connectivity drivers, e.g. MySQL, PostgreSQL, and SQLite, though our system uses MySQL it also supports different databases. Users can incorporate separated table structures and database specific SQL commands.

4.3 Input Module
The Input Module is responsible for connecting with different types of sensors or sensor platforms. Different modules, specific to sensors or sensors platform, can be incorporated easily using the input interface. They are capable of communicating using RESTful interfaces. The Input Module can also be used to retrieve data from Comma Separated Value (CSV) files.

4.4 BRB Main Module
The BRB Main Module is the core module of the expert system and consists of several sub modules. It fetches data from the input module or user interface and maps it into antecedents of the rule base. Afterwards, the inputs are transformed using utility functions based on the reference values. Rules are created using the antecedents and consequent whereas Activation Weight and Matching Degree are computed. Afterwards the belief rule update is performed to take into consideration of the incompleteness of the data. Then the aggregation of the rules is performed. It is also connected with the Knowledge Module using the Knowledge Drive Module for storing and retrieving rules from knowledge base. The Configuration Module is also connected to it. Moreover, different types of configuration parameters are fetched from The Configuration Module. Among different configuration parameters Input Module related and Knowledge Base Driver Module related configuration parameters are very important.

4.5 Configuration Module
To make the system dynamic, different parameters are read from files, so that the user can change the system behaviour without the needs of changing the source code. This module manages different parameters, which are configurable. Different input modules use different types of parameters like URL, user identity, and password. The Knowledge Base Driver Module uses database server name, IP address, database name, user identity, and password. These configuration parameters are collected from the configuration files. There are different files containing different configuration parameters. This module reads the files and passes the parameters to the BRB Main Module and also validates the configuration files. Therefore, erroneously formatted parameters cannot be passed to the BRB Main Module. The Knowledge Base Connectivity sub module reads the parameter Driver Name, IP address, database name, user identity and password from the configuration file. The Knowledge Base Driver Module retrieves the above-mentioned parameters from the Knowledge Base Connectivity sub module. The Rule Base Configurations sub module provides the number of reference values of Antecedent and Consequent and default attribute weight to the BRB Main Module retrieving that information from the configuration file. Similarly, the Configuration for Inference Using ER sub module reads the default parameters for the inference engine. Lastly, the Input Modules Configuration sub module provides a list of available input interfaces to the Input Module. It also provides necessary parameter like IP address, URI, user identity, password, and file location to connect to sensors or sensor platforms or file locations.

4.6 BRB UI Module
The BRB UI (User Interface) Module operation can be divided into two categories. One is for static data, while the other is for dynamic data. Static data related operations are used when input data are given manually. Dynamic data related operations are used for getting data from sensors. These two categories are described below.

4.6.1 Processing Static Data
The BRB UI Module provides an interactive user interface for the user. Figure 2 shows the main interface of Web-BRBES. In the middle the input variables are represented as a tree structure. The legends of the input variables are provided to the left. There are various flood affecting parameters, which can be categorized as meteorological (X8), geological (X9), river discharge (X10), topographical (X11), and human activities (X12). The middle level of the tree contains these parameters and can be called drivers as they influence the flood water level. These drivers depend on multiple factors, e.g. geological (X9) driver depends on
soil type (X19), saturation limit of the soil type (X20) and soil infiltration rate (X21). These are called leaf nodes of the tree. Some of the data of the leaf nodes can be collected using sensors, while others are static over time. Onset rainfall (X22) is collected by sensors, while unplanned infrastructure (X27) is more likely constant or static. The leaf nodes data can also be divided as quantitative and qualitative, e.g. rainfall data is quantitative while unplanned infrastructure data is qualitative.

Various types of flood prediction systems use qualitative and quantitative data separately, but our system can use both of them together [22][23].

In the BRB tree structure the user clicks on the nodes and a new window will appear. The new window is shown in Figure 3. Top of the screen provides the antecedents. The user provides the attribute weight, reference values, and input values. The user can select default reference values or can provide customized reference values. The user can also mention if the input will come from sensors by selecting the check box “Get Data from Sensor”. Similarly, for Consequent reference values the user can use the default values or provide customized values.

At the bottom of Figure 3 there are several buttons to perform the computation. The user has to press the “Calculate Rule Base” button to create the initial rule based on the utility factors. Figure 4 shows the initial rule base.

After that, the user needs to click on the “Input Transformation” button. This will distribute the input data over the referential values. As next step, the user needs to click the “Activation Weight”. During this step the matching degree is calculated and the rules are called packet antecedent or active. The calculated matching degree is used to calculate the activation weight of each rule. The summation of the rule activation weight of a rule base should be one. Figure 5 shows the calculation results of matching degrees and activation weights.

The user continues by clicking “Update”. If any of the input data for the leaf nodes is ignored, then the belief degree associated with each rule in the rule base needs to be updated. The update function performs this task. By clicking “Aggregation” the user performs the aggregation of the rules whereas analytical or recursive uses evidential reasoning algorithm to aggregate the rules.

Figure 6 shows the calculation result of the rules aggregation. By applying the ER algorithm, the rules of the “meteorological” sub rule base have been aggregated by taking into account the input data as shown in Figure 3. The combined degree of belief associated with the referential values (Severe, Average, Low) of the consequent attribute (meteorological) of this sub rule base obtained as {(Severe, 0.2977), (Average, 0.6342), (Low, 0.0732)}. The fuzzy values can be converted into a crisp value by using the utility score associated with each referential values of the meteorological factor. Therefore, the crisp value stands at 0.6098.

Similarly, the user has to click all the driver nodes and then click the root node flood water level (X7). The combined degree of beliefs associated with the referential values of the flood water level (X7) are {(High, 0.9563), (Medium, 0.0436), (Low, 0)}. The fuzzy values can be converted into a crisp value to obtain flood water depth, which is in this case is 97.81 cm as shown in Figure 7.
4.6.2 Processing Dynamic Data

If the user wants to use sensor data, then he/she has to click only the “Calculate Sensor Data”. The application will perform the above tasks one by one and get the final output as shown in Figure 9 for each driver where the Input box shows the input values from the sensors, while the Output graph box shows the crisp output values for the corresponding sensor data. Finally, the user has to click on all the driver nodes and then click the root node flood water level (X7) to get the ultimate result.

In this way, we can generate a flood scenario for different areas of a region. In summary, our system is based on a layered and modular architecture consisting of three layers, being the presentation layer, the application layer and the data access layer following the Model–View–Controller (MVC) pattern for our web application. The Input and Knowledge Base Driver modules are modular since different sub modules can be easily plugged in and out from the main module.

5. RESULT AND DISCUSSION

The new Web-BRBES provides better user interface with more visualization of data compared to the previous desktop-based application as shown in Figure 8. The new system is also capable of getting data from sensors. It provides easier access for the users than the desktop-based version, as the user does not need to install any additional software except a web browser, which is commonly available nowadays. As the expert system is web based the user does not have to be concerned with the software update process. Therefore, users and IT support personnel are relieved from problems of software deployment. Moreover, the web application now can use more computational power as it can easily be deployed in high-end servers.

Now using cloud technologies more processing power, memory, and network bandwidth can easily be added to the server. This will allow the web application to compute more rules and complex knowledge base efficiently and faster. It also provides better visual interface and nice representation of the data compared to the previous desktop version. This system also has the capability of connecting with several sensors or sensor data collection frameworks. A comparison of features between Web- and Desktop-based BRBES is presented in Table 1.

Figure 9 provides a detailed view of the input and output data. In this scenario the data fetched from sensors is considered as inputs for X23 and X22. The X-axis shows the data in chronological order while the Y-axis shows values of the data gathered by sensors during each time interval shown in the graph name “input”. Similarly, in the graph name “output”, we can view the output generated for the inputs during the same time interval. If we consider the input and output graphs we will find that X22 has greater influence on the output than X23. In this way, the web-
The next step will be to stress test the amount of sensor data the system can support. We also plan on conducting a user survey to look into the usability of the web interface. In addition, we need to look into the security aspect of the system. Figure 10 shows the location of a probable case study area named Bakkhali river, located in Cox’s bazar district of Bangladesh, which is the southern part of Bangladesh and covering an area of approximately five square miles. The area consists of plain land and also with a number of canals, facilitating the drainage of rainwater. The livelihood of the people depends upon the production of various crops in the soil of this land. Therefore, the regular prediction of flood water level or the generation of flood scenarios by taking different input variables into account is very essential.

7. ACKNOWLEDGMENTS
This work has been supported by the Swedish Research Council under grant 2014-4251.

8. REFERENCES


