IOP Conf. Series: Earth and Environmental Science **32** (2016) 012021 doi:10.1088/1755-1315/32/1/012021

The Semantic Network of Flood Hydrological Data for Kelantan, Malaysia

[1]Aziyati Yusoff, [1]Norashidah Md Din, [2]Salman Yussof and [3]Samee Ullah Khan

[1]College of Engineering, Universiti Tenaga Nasional (UNITEN), Malaysia
[2]College of Information Technology, Universiti Tenaga Nasional (UNITEN), Malaysia
[3]Department of Electrical and Computer Engineering, North Dakota State University (NDSU), United States of America

Email: aziyati@mohr.gov.my, norashidah@uniten.edu.my , salman@uniten.edu.my , samee.khan@ndsu.edu

Abstract. Every year, authorities in Malaysia are putting efforts on disaster management mechanisms including the flood incidence that might hit the east coast of Peninsular Malaysia. This includes the state of Kelantan of which it was reported that flood is just a normal event occurred annually. However, the aftermath was always unmanageable and had left the state to struggle for its own recoveries. Though it was expected that flood occurred every year, among the worst were in 1967, 1974, 1982 and recently in December 2014. This study is proposing a semantic network as an approach to the method of utilising big data analytics in analysing the huge data from the state's flood reading stations. It is expected that by using current computing edge can also facilitate mitigating this particular disaster.

1. Introduction

The study on the patterns of flood incidence had triggered further studies to be done for an effective warning system and an efficient evacuation system. Researchers had proposed numerous approaches on the flood disaster management. The need for precise computation had challenged the experts in Geographical Information System (GIS) to be updated with the latest technology in the world of computing. For this research, we are interested to study on the flood patterns that were reported to occur annually in the state of Kelantan, Malaysia.

In most cases, the governmental agencies had been using renowned computer software for flood data and mapping. Among others, ArcGIS software by Esri Inc. is always preferable for mapping and geospatial measurement Other prominent geospatial mapping tools include Global Precipitation Measurement (GPM) [3], SERVIR Labs [4], MODIS NRT Flood Mapping [5], Global Flood Monitoring System (GFMS) [6], Global Disaster Alert and Coordination System (GDACS) [7], and Dartmouth Flood Observatory [8].

The flood management is under the administration of Department of Irrigation and Drainage (DID), Ministry of Natural Resources and Environment (NRE). According to DID, every year it is estimated that 29,800 kilometers of Malaysian land is prone to flood incidence. In year 2006 and 2007, Malaysia was facing a very bad flood disaster that total loss was estimated as RM1.1 billion nationally and RM776 million for regional state assets excluding the personal loss of the affected communities nationwide [9]. However, the authority body that manage disaster occurrence in this country is the National Disaster Management Agency (APBN), Prime Minister's Department. This agency has the responsibility to coordinate the management of national disaster, and to ensure that all of the national disaster policies and mechanisms are adhered and implemented at all level of disaster management units [10].

As an example, one of the worst hit flood had occurred in December 2014. Kelantan had the highest figures with losses and damages. Report from the officials of Malaysia [8] had been recorded and the profile of the incidence as illustrated in Table 1.

Table 1. Profile of the latest flood incidence in the state of Kelantan (2014)

Facts	**Figures**
Country	Malaysia (States worst-hit: Kelantan, Terengganu, Pahang)
Date	21 to 31 December 2014
Death	At least 14 persons
Injured	8 persons
Affected institutions	1,415 villages, 237 schools, 32 religious sites, 17 government offices
Evacuees	132,000 (21,686 for Kelantan state)
Affected infrastructures	Neighbourhoods turned into islands surrounded by a sea of murky brown water. Streets doubled as rivers. Cars were swept away and toppled over. Strong currents and interrupted power supply. Clean water cannot be delivered.
Affected national production	184,500 hectares (455,000 acres) of palm oil production. Output dropped about 20 per cents (1.4 million metric tons).

2. Semantic network of hydrology data

This research is to study on a feasible method for handling bulk hydrology data that being archived by the state authorities, of which were never being fully utilised for the sake of effective flood management. As a result, an approach of semantic network is proposed for this research. Figure 1 has been developed to represent the classes and relationships that can be abstracted from this study.

The study is proposing that its main class is 'big data analytics'. The sub-classes for this analytics will involve the 'hydrology data', 'GIS data' and data from 'NASA Earth Observatory (NEO)', which is downloadable from its web portal.

As illustrated in Figure 1, the hydrology data has the property of sub-classes of: (i) rainfall data, (ii) water level data, and (iii) reading stations. They have equal relationship of determining the type of data being read and stored in the class of 'hydrology data'. From the classes of 'rainfall data' and 'water level data', statistical inferences were made. The null hypothesis represented here is that the occurrence of flood incidence is determined by these two (2) variables. As a consequence, the study on the two variables are reflected from alternative hypotheses that can be deduced from the data readings and were computed as the analysis in the big data research.

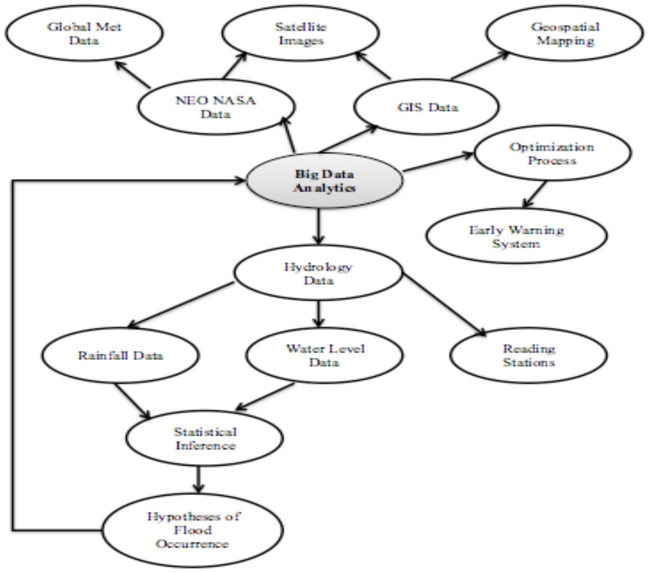

Figure 1. The semantic network of big data analytics for hydrology data.

In addition, the main class of ‘big data analytics’ should also be able to facilitate relative variables from sub-classes of ‘NEO NASA data’ and ‘GIS data’. Though the type of data readings from the three sub-classes of ‘big data analytics’ might be in different format, it is expected that the proposed early warning flood system would be able to accommodate each class of data type. It is also expected that, the hydrology data was meant to represent the comma separated values (csv) data, the GIS data was to represent the Geography Markup Language (GML) or GeoRSS feed, and the NEO NASA data was represented by Geography Tag Image File Format (GeoTIFF).

3. The Knowledge Abstraction of Semantic Network

From a semantic network, we are able to define the knowledge representations and predicate calculus abstracted from the relationships and properties of each class and subclasses identified. The biggest contribution for this step is the fact that we are able to do decision making and knowledge management [11].

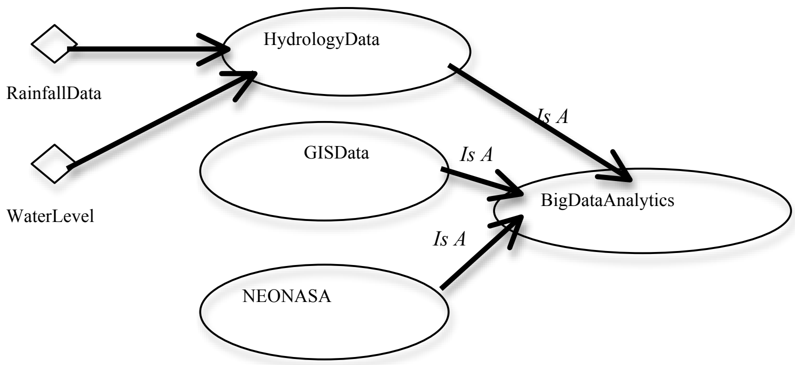

Figure 2. The class and instances relationship

The property of 'hydrology data' which is the subclass of 'big data analytics' is also illustrating on the fact that our limitations of study are contained within these two main classes. Hence other subclasses of the main class of 'big data analytics' hold TRUE to its identified classes until they are defined otherwise. The predicate calculus for this semantic includes: (i) *isA* (*HydrologyData, BigDataAnalytics*), (ii) *isA* (*GISData, BigDataAnalytics*), and (iii) *isA* (*NEONASA, BigDataAnalytics*), so on and so forth. This is as illustrated in Figure 2. From the figure, two instances were given for the class 'HydrologyData' consisting of 'RainfallData' and 'WaterLevel'. Instances can be further extended from other classes too. It is nevertheless, from this semantic that it is expected to design an effective data management and build efficient expert systems for flood management in Kelantan.

4. Conclusion

This research is proposing on using semantic network to facilitate the classes and relationships to the study of big data analytics of flood hydrological data. The study is specifically for the state of Kelantan, Malaysia. An introduction to the state flood history and previous reports were primarily highlighted for identifying the classes and relative parameters. As a result, a semantic network was designed and the class and instances relationship of the predicate calculus were also developed.

5. References

[1] A.S. Zakaria, "The Geomorphology of Kelantan Delta (Malaysia)." Catena 2 (1975): 337-349.
[2] TCPD. 2002. Malaysia Structure Plan, Malaysia Department of Town and Country Planning. Unpublished report. TCPD: Kuala Lumpur.
[3] Global Precipitation Measurement (GPM): Precipitation Measurement Missions, URL: "*http://pmm.nasa.gov/GPM*", Jul 2015
[4] SERVIR: Crest Viewer, "*http://ags.servirlabs.net/crestviewer/*", Jul 2015
[5] NRT Global Flood Mapping, "*http://oas.gsfc.nasa.gov/floodmap/*", Jul 2015
[6] Global Flood Monitoring Services (GFMS), "*http://flood.umd.edu/*", Jul 2015
[7] Global Disaster Alert and Coordination System, "*http://www.gdacs.org*", Jul 2015
[8] Dartmouth Flood Observatory, "*http://floodobservatory.colorado.edu/*", Jul 2015
[9] Pengurusan Bencana Tanggungjawab Bersama. URL: "*http://portalbencana.mkn.gov.my/Portal/Board/Detail?board=141&entity=7524*", 2014
[10] Majlis Keselamatan Negara. "Arahan No. 20 (Semakan Semula): Dasar dan Mekanisme Pengurusan Bencana Negara". Jabatan Perdana Menteri, Malaysia. 2012
[11] E. Turban, R. Sharda, D. Delen, and T. Efraim, "Decision Support And Business Intelligence Systems". Pearson Education India, 2007.

Acknowledgments

The authors wish to express the gratitude to the research team of Early Flood Warning System and Big Data Analytics of College of Engineering, UNITEN under the JICA-SATREPS Project on Research and Development for the reduction geo-hazard damage in Malaysia caused by landslide and flood and also to Big Data research team of North Dakota State University (NDSU). Heartiest appreciation goes to Public Service Department of Malaysia (JPA) as the main sponsor for this research, and Department of Irrigation and Drainage (JPS) Malaysia for providing the flood big data.