



Big Data: A Computing Model for Knowledge Extraction on Insurgency Management

Adeniji Oluwashola David

Department of Computer Science, University of Ibadan, Ibadan, Nigeria
od.adeniji@mail.ui.edu.ng

Abstract—The study investigates applications of big-data in Africa using Boko Haram and Al shabaab insurgency and their effect in the society. A framework on Big Data was developed to extract knowledge based on the activities of the insurgency. The crucial information that concerns the activities of insurgent majorly cannot be easily access due to inaccurate data set. The research employs traditional computing models that were used to limit volume, velocity and variety as the major component of Big Data. The analysis of the frame work are based on different methods of attacks employed by these terrorists, development of sophisticated apparatus or technology, the relative increase in the spate of attack, and success rate in terms of causalities recorded. Big data requires exceptional technologies to efficiently process large quantities of data within tolerable elapsed time. The knowledge extraction in this research is to analyse the data sets by presenting a methodology from the application of big data to solve and offer analytical solutions using an integrated Boko Haram and Al shabaab data set.

Keywords: *computing models; insurgency; knowledge extraction; component of big data*

I. INTRODUCTION

The Concept of Big Data has led to the variance in the deployment of different technologies. The deployment variance in technology will provide data-set for extraction. Big Data can be defined as collection of data-set from various sources which are analyzed when certain invent took placed. The occurrence of these invent are associated with data at rest and data in motion. The interaction and activities of people, process and data will provide real time data set for knowledge extraction and discovery. The investigation of “Boko Haram” and “Al-shabaab” in this study is timely due to the challenges faced by Nigeria and Somalia. Insurgencies are resources that need to be managed. The effect of insurgencies management cannot be emphases and their implication on investments such as Schools, Churches Mosque and Health care centers are prime indexes for sustainable development of any Nation. The core interest in this research is to develop an information system basically on Boko Haram (Nigeria) and Al-shabaab (Somalia) insurgency in Africa that will contain timely data set collected from several sources which will give opportunities for new data discovery, value creation and important decision making on how to curb insurgencies in Africa.

Traditional data analytics is the traditional method of managing structured data which includes a relational database and schema to manage the storage and retrieval of the dataset. This paper is organized as follows: Section 2 explains the Component of Big Data and insurgency. In section 3 Computing models and knowledge extraction for the model was described while in section 4 discussion of result was provided. The conclusion of paper was made in Section 5.

II. COMPONENT OF BIG DATA

The Complexity of handling data-set in Big Data using on-hand management tools was describe by [1]. The study based it assessment on traditional way of processing data. However Big data can be analyzed with common software tools which can predict, mine data, analyze text and provide statistical analysis as discussed by [2]. The data set in big data are faced with challenges such as capturing, storing, searching, sharing and visualization. The three major indices for characterizing Big Data are Volume, Velocity and Variety. Although these three indices made Big Data to be different from the traditional way of analyzing data. Variety manages the complexity of multiply data types which vary from structure and unstructured data. Formatted, modeled and organized data are refer to structured data while emails, audio, video and images are unstructured data. Velocity is the major characteristic of big data. Velocity provides the speed of generation of data or how fast the data is generated and processed in order to meet specific task. Volume refers to the quantity of data that is generated. There are some other characteristic of big data such as variability, veracity, value and complexity.

A. Insurgency

Basically the act of rebellion directly or indirectly is referred to insurgency [3]. The ideal definition of insurgencies can be ambiguous depending on the usage [4]. The review by some researcher considers insurgency has association that is illegitimate or not lawful by the law of the land. The act that is illicit is called Terrorism but it is more than mere criminality [5]. Terrorism is the basic part of the first part of the three phases of revolutionary warfare. Boko Haram is an indigenous Salafist group which turned to a Salafist Jihadist group in 2009 while al-Shabaab stems from

a Salafi- Wahhabi strand of Sunni Islam, Muhyadin [6]. The information that was gathered in the militant salafi group 1990 shows that the forerunner of al-Shabab, was Al-Ittihad Al-Islami (AIAI, or “Unity of Islam”). The fall of the Siad Barre military regime (1969-1991) during the outbreak of civil war. The internet was used by Boko Haram and al-Shabaab to catalyze and promote their ideology in order to celebrate their martyrs. These two group communicate with their audience through the internet using blog and email. They maintain an email at; nigjihadist@yahoo.com to communicate with intending members.

They propagate Boko Haram and their interactions with the Western World is forbidden [7]. The information gathered shows that their interaction are against the Muslim establishment and the government of Nigeria [8]. Their agenda was to accuse the government of political corruption and weak judicial structure.

III. MATERIALS AND METHODS

The computing model is made up of Use case model, Data flow Diagram and E-R diagram. The UML diagrams was used to specify the functionality and non-functionality of the system. The process in the modeling is inherently iterative. Figure 1 shows level 0, data flow diagram.

The system requirement in the computing model developed consists of program and database which can adapt to modification. The operational approach of the model deployed consists of parallel change over, direct change over and pilot/phased change over. The system was tested to ensure validation and invalidation data in order to ascertain the correctness, effectiveness and efficiency of the system, [9]. This is to evaluate that the complete, integrated system complies with its specified requirement.

IV. RESULT AND DISCUSSION

Knowledge extraction will allow users to view, sort or search by year, country or location of the information needed. The forms that were design provide information that is sent to the database. It also prevents duplication of records in the database by the administrator. The extraction could be data of terror attacks plotted against fatalities by year of attack. Figure 2 presents the line chart of the statistical data of terror attacks, Figure 3 presents pie chart of the statistical data of terror attacks, and Figure 4 presents bar chart statistical data of terror attacks.

The extraction of knowledge from the figures can be summarized in table I.

The analysis presented in the table above provide the nature and type of attack .The rate of fatalities on the line chart increases from 2009 to 2014 and suddenly dropped at 2015. Consider 9.5k on the line chart in 2014 terror attack, it

could be deduce that a specific location in the region was massively attacked that lead high destruction of life and properties. From the table 1 also there could be more of Bombing and explosion, likewise the data that was extracted on the pie chart shows 59.3% terror attack in 2014. In 2009 the terror attacks was 4.7%, 2010 was 0.5%, 2012 was 11.8%, 2013 was 19.9% while in 2015 there was a drop of 0.1% in the attack. The justification from the data gathered shows that, number of attacks increases with years showing chronicles of cases of terror attacks.

V. CONCLUSION

The crucial information that concerns the activities of Boko Haram and Al shabaab are vulnerable. Big Data technology will assist the government to extract information and provide solution on how to eradicate the insurgent. When developing the application of big data for analysis. Different methods of attacks were employed and deployed which clearly revealed the success rate in terms of casualties recorded. However the analysis shows that the fatalities in attack periodically increase and suddenly dropped. The applications of big-data analysis have high effect in the society because of the numerous attacks the society is facing.

VI. ACKNOWLEDGMENT

The author wish to thank the Department of Computer Science, University of Ibadan for support in this research work.

REFERENCES

- [1] R. GUTH, “Deriving Intelligence from Big Data in Hadoop: A Big Data Analytics Primer [Z],” Karmasphere.com, 2013
- [2] R. Magoulas and B. Lorica, “Big data: Technologies and techniques for large scale data,” Jimmy Guterman, Nov 2009.
- [3] B. Fall, “The Theory and Practice of Insurgency and Counterinsurgency”, Naval War College Review, 17, 2010.
- [4] J. Bughin, M. Chui, and J. Manyika “Clouds, big data, and smart assets,” McKinsey Quarterly, 1, August 2010.
- [5] J. Brock, “Nigerian Islamist sect claims bomb attack,” 17(06), 2011
- [6] M. A. Roble. “Al-Shabaab Razes Somali Forests To Finance Jihad,” Terrorism Monitor, vol 8, issue 42, Nov 2010.
- [7] A. Murtada, “Boko Haram: Its Beginnings, Principles and Activities in Nigeria,” May, 2014
- [8] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, “Big data: The next frontier for innovation, competition, and productivity,” May 2011.
- [9] E. G. Anderson, “A Proof-of-Concept Model for Evaluating Insurgency Management Policies Using the System Dynamics Methodology.” Strategic Insights, vol VI, issue 5, August 2007.

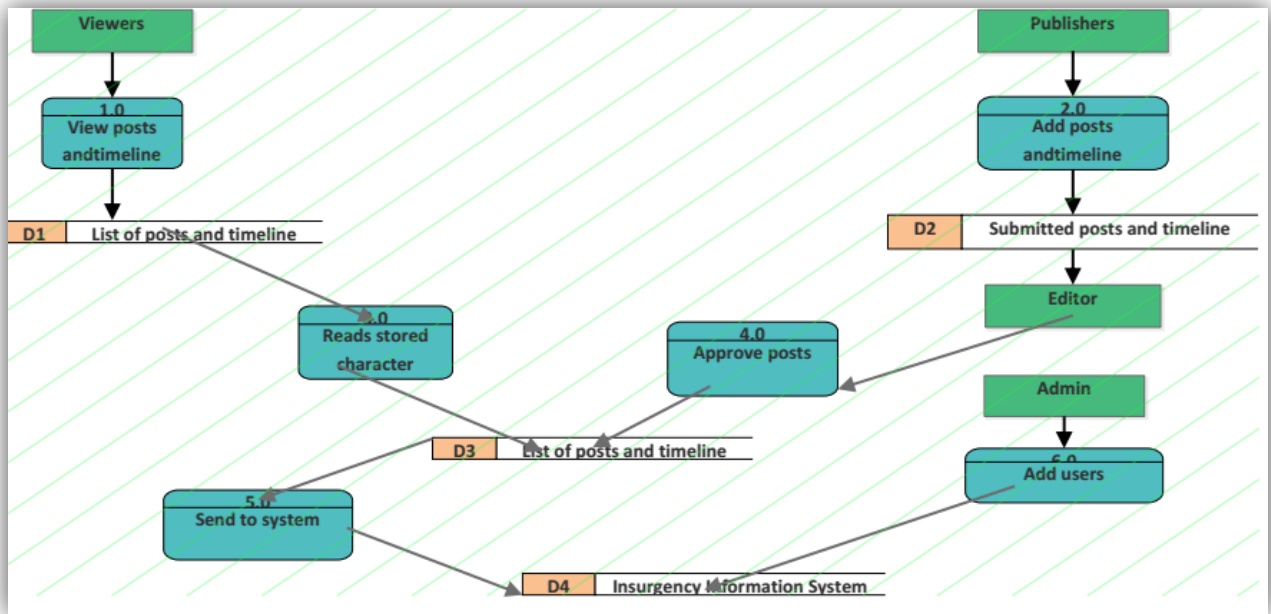


Figure 1. Shows Level 0, Data Flow Diagram

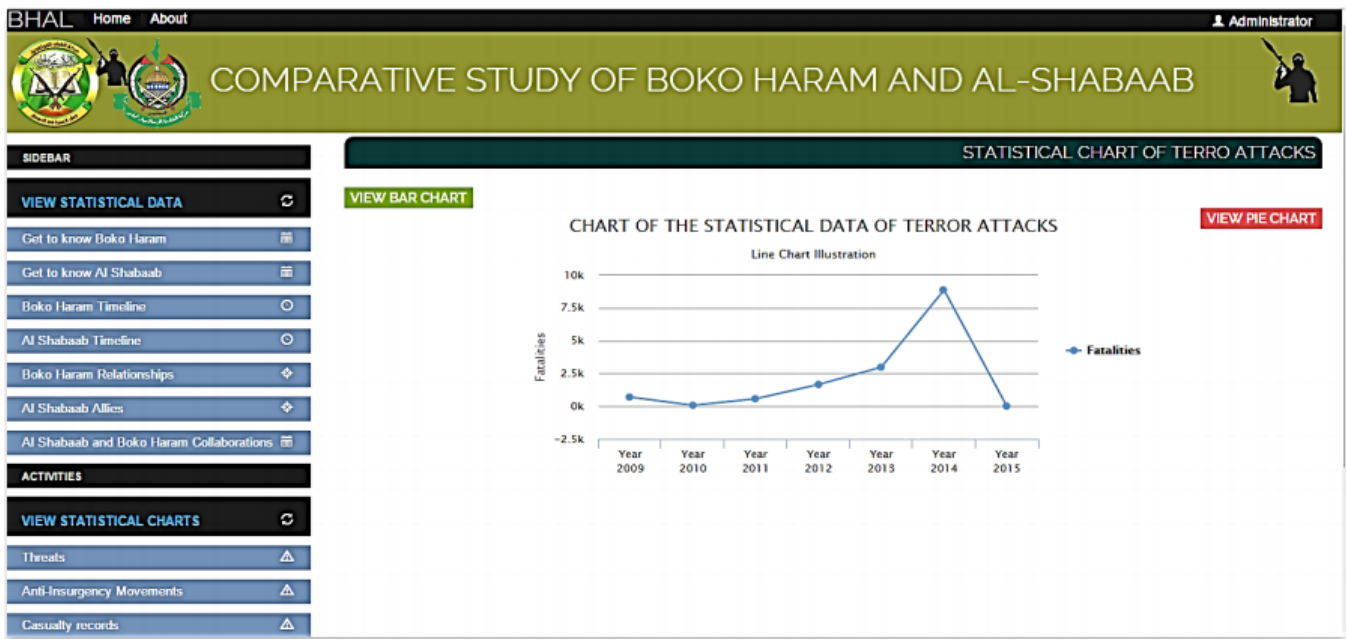


Figure 2. Line chart of the statistical data of terror attacks

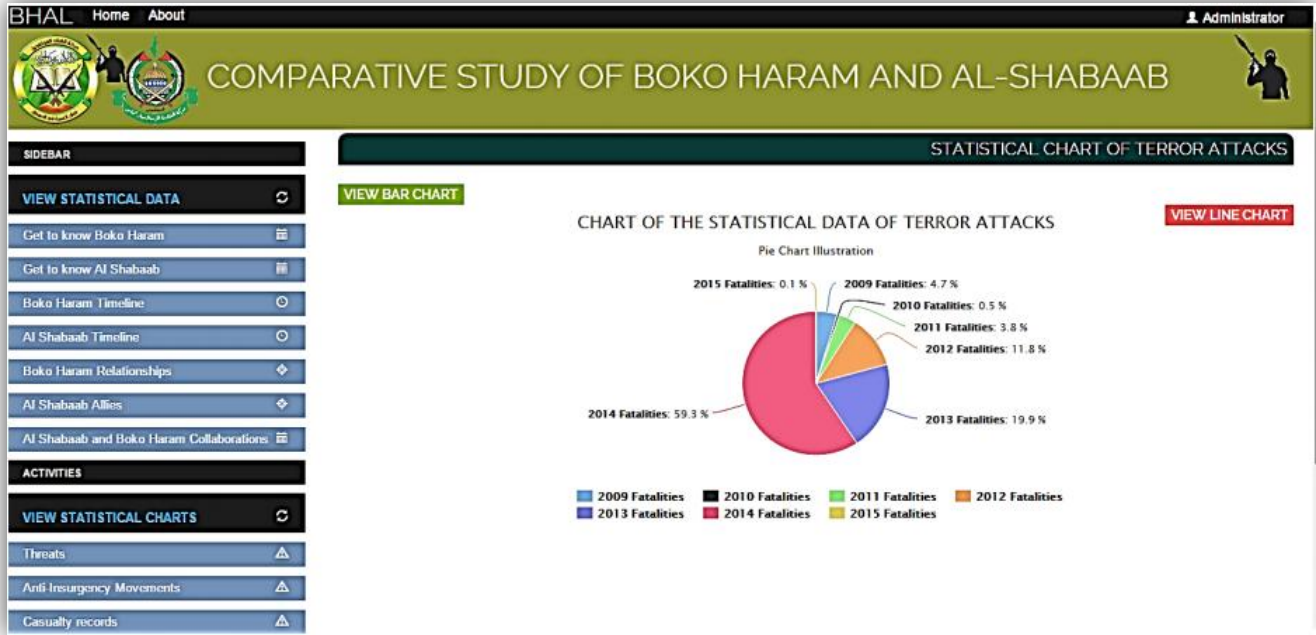


Figure 3. Pie chart of the statistical data of terror attacks

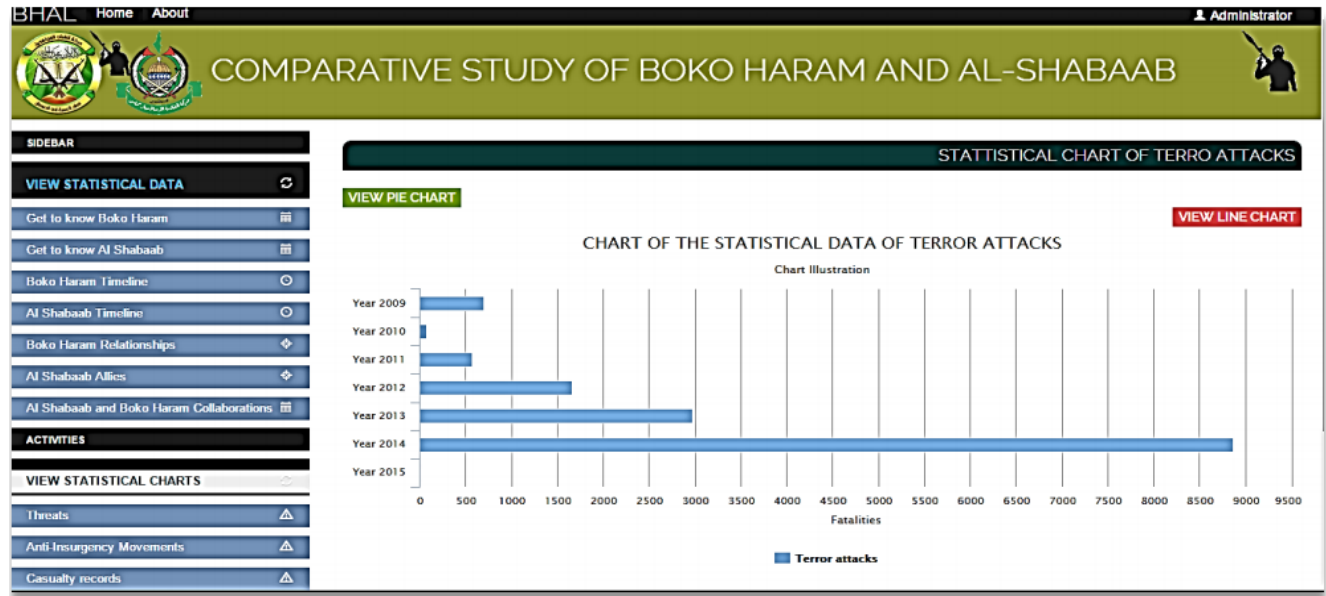


Figure 4. Bar chart statistical data of terror attacks

TABLE I. ANALYSIS SUMMARY OF THE FRAMEWORK BASED ON METHOD OF ATTACK, TECHNOLOGY USED, INCREASE IN ATTACK RATE AND SUCCESS RATE 2009 TO 2015

Years of Terror	2009	2010	2011	2012	2013	2014	2015
Line Chart	1k	0.8k	0.9k	2k	2.6k	9.5k	0.1k
Pie Chart	4.7%	0.5%	3.8%	11.8%	19.7%	59.3%	0.1%
Bar Chart	800	300	700	1,800	3,000	8,880	100