International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 03 | Mar-2016 www.irjet.net p-ISSN: 2395-0072

Review Paper On An Ontology-Based Text Mining To Generate D-Matrix For Fault Detection and Diagnosis

Priyanka D. Raut [1], Prof. M. M. Bartere [2]

[1] *SGBAU, Amravati University, G.H. Raisoni College of Engg & Management, Anjangaon Bari Road, Amravati, 444701, India*
[2]*Professor M. M. Bartere, Dept. of Computer Science & Engineering, G.H. Raisoni College of Engg & Management, Anjangaon Bari Road, Amravati, 444701, India*

Abstract - *Fault dependency (D)-matrix is a systematic diagnostic model to capture the hierarchical system-level fault diagnostic information consisting of dependencies between recognizable symptoms and failure modes associated with a system. There is need to gather information regarding servable symptoms and failure modes to modify the fault dependency matrix which can be helpful to build Accurate and efficient fault diagnosis. Dependency matrix is an organized diagnostic model to pinch the graded system-level fault diagnostic information. Organizing this information is typically based on the previously known knowledge and research. Constructing a D-matrix from first principles and updating it with the domain knowledge is a labor intensive and time consuming task. Here, an ontology based text mining method for automatically constructing and updating a D-matrix by mining hundreds of thousands of repair verbatim is described. In proposed approach, firstly construct the fault diagnosis ontology consisting of concepts and relationships commonly observed in the fault diagnosis domain. Next, hire the textual content mining algorithms that employ ontology concept to become known of the wanted artifacts, inclusive of symptoms, failure modes, and their dependencies from the unstructured restore verbatim textual content.*

Key Words: **Data Mining, fault analysis, fault diagnosis, text processing, information retrieval.**

1.INTRODUCTION

A complex system interacts with its surrounding to perform a set of tasks by maintaining its performance within an acceptable range of tolerances. Any deviation of a system from its acceptable execution is treated as a fault. The fault detection and diagnosis (FDD) is carried out to find out the root causes of a fault because there may be a possibility that other interconnected subsystems may also give fault indications that may possibly hide the root cause. Identification of faults and its correction is a subfield of control designing which relate itself with managing a system, recognizing when a fault has happened, what are the reasons behind it and discover the sort of fault and its location [1].On-board diagnostics (OBD) systems like diagnostic sensors, internet, etc. the process of FDD has been a challenging task in case of component or system malfunction.

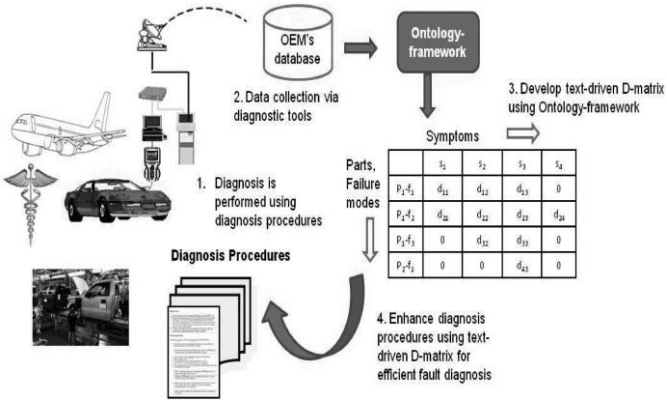

Fig. 1. Scope of the ontology-based text mining framework.

Text mining [2] is the process of gaining a serious attention due to its ability to automatically discover the useful or interesting knowledge from unstructured text. Various techniques have been proposed for text mining such as association rule mining, decision trees, rule induction methods, conceptual structure and episode rule mining.

Previously, Ontology learning systems for concept retrieval were based on words. Before keywords were identified from the unstructured text, the identified words are typically single-word terms and they are considered to be the concepts. Later, combining the extracted keywords possible multi-word terms are formed. As a result, the multi-word terms generated were not natural and thus only the single-word terms were formed from most of the extracted concepts. So while processing documents using the NLP component, most noun terms found in the text was multi-word terms. As it was also shown in the unstructured text, in that 85% of the terms were multi-word terms, so traditional systems focusing on single-word term extraction will therefore miss many concepts.

2. LITERATURE SURVEY

An Ontology-based fault detection and diagnosis for power transformers by Prof. D. Wang presented a methodology to transformer fault diagnosis that focused around the thought of formal and machine open descriptions of importance by using the Semantic Web[3]. An ontology model is created for precise and effective fault diagnosis for power transformers. Through the use of this model, different transformer fault symptomatic strategies can be incorporated to depict and derivation among fault phenomena, fault sources and reasons for faults. Previously, the new concepts can be added into an ontology based on domain [4], but it requires some ways for evaluating and updating the result. So, it is very lengthy process.

According to Schuh et al., 2013, this approach is to remove the faults and provide an ontology-guided data

mining and data transformation[6]. But Discovery is lost because result is not in the form of matrix.

According to Singh and Dhir, 2013, this paper presents transaction reduction for finding item sets based on tags and shows result in matrix[7]. This method does not give an accurate result. Its search is based on tags only. This paper does not make use of ontology.

M. Gaeta, F. Orciuoli, S. Paolozzi, and S. Salerno, provide an easy to use interface that generates relevant sequences of data in meaningful context and retrieve and display similar information but it only shows similar information not an accurate result in the form of D-MATRIX.

Till now, there is no accurate mechanism available for the data retrieval using text information. Existing systems depend on the title given to the Files or Data. Title of each File is used as a main parameter for sorting the number of Data against the search query.

3. PROPOSED WORK

This paper presents An Ontology-Based Comprehensive D-Matrix generation .The system comprises of the developments of D-matrix from the repair verbatim data. First, the D-Matrices are created from different datasets[8], after that the graph is generated for each D-Matrix. Then, the graphs are combined in such a way that only the common patterns are merged from the generated heterogeneous D-Matrices to construct a single, generic D-Matrix.

To construct the D-matrix , following steps need to be followed

1. The fault diagnosis Ontology by using dataset.

Ontology is a mechanism that describes the concepts and also the relationships between those concepts observed in the domain of vehicle fault diagnosis. To develop an ontology system for the fault diagnosis of automobile systems, it is essential to analyze numerous concepts and relationships exhibited.

2. Ontology-based Text Mining.
Ontology-based Text Mining consists of three steps, Document Annotation, Term Extraction and Phrase Merging.

i. Document annotation

The document annotation helps to filter out the information that is not relevant to our analysis and it provides a specific context [5] for the consistent comprehension of the data.

ii. Term extraction

Having annotated the terms, the vital terms desirable for the generation of a D-matrix, such as, symptoms and failure modes are extracted by using the term extractor algorithm.

iii. Phrase merging

During phrase merging, the same failure mode phrases that are generally written with the help of an conflicting vocabulary are merged into a single, consistent failure mode phrase using phrase merging algorithm.

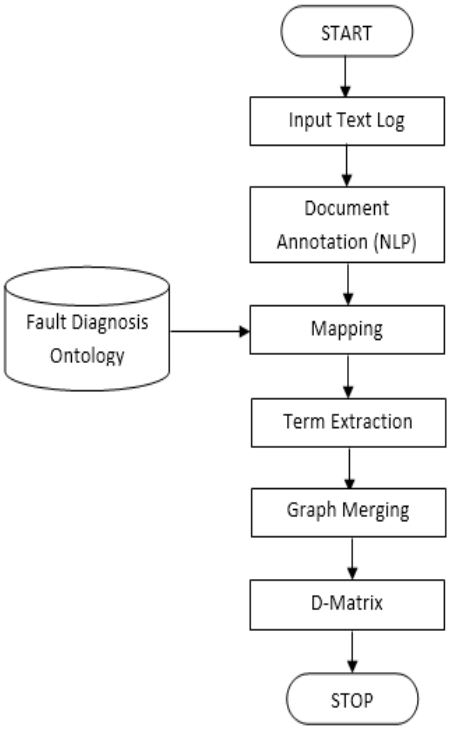

Fig. 2: Dataflow Diagram

4. CONCLUSION

This paper has proposed an Ontology-based text mining approach. It constructs the D-matrices through routinely mining the unstructured restore verbatim data collected in the event of fault diagnosis. This approach has provided a framework that converts data-rich unstructured documents into structured documents. In this approach, we develop d-matrices from different datasets. Then, we represent each d-matrix in the form of graph and by using graph merging algorithm, we combine only the common patterns of generated graphs into new graph. This newly generated graph is then used to construct D-matrix which is comprehensive and combination of two d-matrixes.

REFERENCES

[1].O. Benedittini, T. S. Baines, H. W. Lightfoot, and R. M. Greenough,"State-of-the-art in integrated vehicle health management," *J. Aer. Eng*.,vol. 223, no. 2, pp. 157–170, 2009.

[2] T. Hearst, "Untangling text data mining," in *Proc. 37th Annu. Meeting Assoc. Comput. Linguist.*, 1999, pp. 3–10.

[3].Wang, W. H. Tang, and Q. H. Wu, "Ontology-based fault diagnosis for power transformers," in Proc. IEEE Power Energy Soc.Gen. Meeting, 2010, pp.

[4].Rung-Ching Chen, Ya-Ching Lee, Ren-Hao Pan," Adding New Concepts On The Domain Ontology Based On Semantic Similarity," Department of Information Management, Chaoyang University of Technology 168, Jifong East Road, Wufong Township, Taichung County 41349, Taiwan ROC.

[5] M. Agosti and N. Ferro, "Annotations as context for searching documents," in *Proc. Int. Conf. Concept. Library Inf. Sci.—Context: Nature, Impact Role, LNCS*, 2005, pp. 155–170.

[6]. Schuh M., Sheppard J. W., Strasser S., Angryk R., and Izurieta C., 2013. A Visualization tool for knowledge discovery in maintenance event sequences.In July 2013.

[7]. Singh H. and Dhir R., 2013.ANew efficient Matrix based frequent Item sets Mining Algorithm with Tags. In August 2013.

[8]. Dnyanesh G. Rajpathak, "An Ontology-Based Text Mining Method to Develop D-Matrix from Unstructured Text"*Member, IEEE* and Satnam Singh, *Senior Member, IEEE.*

BIOGRAPHY

Priyanka D. Raut received the B.E. degree in Computer Science and Engineering from Prof. Ram Meghe Institute of Technology & Research, Badnera in 2013. She now with G.H. Raisoni College of Engineering and Management, Amravati.