

Knowledge of the Elements in Groups Question

N.Uma Rani

Department Of Computer Science and Engineering

ABSTRACT:

In this paper, as shown by the data in order to enable the introduction of the inevitability of a problem with the identification information stored in legacy systems into account. It can be generated from the potential of this decision entity, data mining, and speech processing and automated strategy / development of data analysis. Legacy systems based on the inevitability of data to improve the quality of data that is created in the representation of peremptory such as Flickr, Picasa, as well as Web applications that are compatible with pre-existing and. We have two different data processing and selection of questions trigger tasks explored in the context of this problem determination. We like the approach the threshold or higher -1 commonly used for such applications determination lead to sub-optimal performance to explain it. Instead, we are aware of the strategy developed with a question and extensive research on the assessment of real and synthetic data sets show the advantages of existing solutions.

Index Terms:

Determinization, data quality, query workload, uncertain data.

I.INTRODUCTION:

Identify potential data. As we studied in this paper addresses directly the question of the data report is not aware of any prior work, a lot of work on this project, which is linked to ours. They are determined to explore how to answer a question about the database. End-use applications that take input peremptory only continue to exist, so that, in contrast to the information (this is not an answer to this question) is the best deterministic representation. The difference between the two issues in a variety of settings to appeal. A group of objects in such a problem in order to achieve improvements in the quality of the author of the query, the answer is clear. Ours is to improve the value of the query workload, but when their mission is to improve the quality of a single query.

The different models have been proposed in the past to improve the data. Our focus, however, it is possible that this type of model is attributed enough to sign the picture and speech output possible objects, and determining. We and / or data stored in the tree to assess the potential of a more sophisticated model and should be interesting. To deal with the complexities of our business, as well as the expansion of the data is still the future direction of interest. For document retrieval and document addressing the problem in terms of the research effort index. The term describes a way of pruning-based audio to each posting an ad hoc query when they appear under the influence of the individual scores for each word retains the top post. The text, which is based on the author in terms of coverage classification, and progress on the subject of the term. This research effort is focused importance - the documents that are most appropriate in terms of getting the right group, either. Our problem is probably related in terms of a set of documents and their importance is already offered through other data processing techniques. Thus, our goal in terms of the importance of the documents to explore, but also documents, as well as the trigger / answer from a certain set of values that are optimized in terms of representation to determine the keywords issues.

II.RELATED WORK:

Introduction of cloud computing and the rapid growth of applications on the Internet, people often save their data in a variety of applications on the World Wide Web. Often, the user information stored in a variety of applications on the Internet before the signal processing, query and analysis / enrichment through a variety of strategies that are automatically generated. As well as often being many images of modern camera by the user identification, then the speech, which the microphone to speak out as a descriptive sentence in the outdoors indoors / different scene, landscape / portrait Kbesmh to generate a vision of DSLR modern cameras also support analysis for the establishment of a set of labels to identify add images [2]. Images (including labels), such as Flickr, and Web applications can be found in real-time using a wireless connection.

Automatically generated content, such as objects with often unpredictable attributes, and possibly can be, because this information is a big challenge in the Web application. For example, analysis of the feasibility of [3], [4], and, automatic as well as speech recognition (ASR) can be tagged with “determinized or a list of possible questions must be talking N of this type can create a network confusion” stored by the in the old web applications. Our problem is the same determinization deterministic representation of the potential of the data indicate a problem with mapping. determinization a lot of problems like this approach can be achieved. There are two main ways to top 1 and strategy, we have more potential value / non-probability, respectively, to attribute all the possible choices. For example, one solution that generates a speech recognition system / strategy for each word used here can be seen as one of the top to 1. Last τ technique with the possibility of greater value for each attribute can be found in τ and choose the threshold. However, these approaches are often suspected of the latest being that the results are optimal. The best approach is to maximize the value of the order is determinized choose strategies that are designed for determinization allocated. For example, triggers / alerts on automatic generation of content to support the application, please consider. Examples of such end applications (such as video on this information), as well as people in the index words (egGujarat earthquake) in the form of subscription and guess where Google Alert on a database system and has subscribed / Publications. Google Alert Subscribe to the user based on all relevant information provided. Now, for such a About video and upload them to YouTube Gujarat earthquake. By automatically processing of video and / or use of information retrieval techniques determine that there is a set of tags Speech text engine.

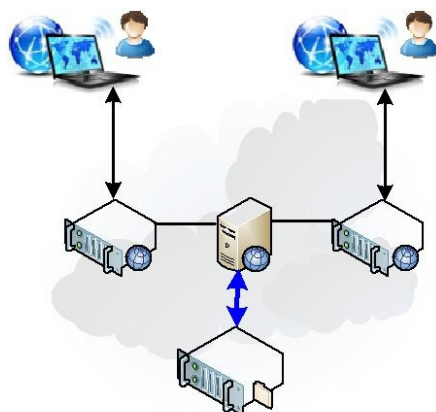


FIG 1: System Architecture

III. SYSTEM PREMELIRIES

A.Determinizing Probabilistic Data:

The problem of determining the information that I do not know of any previous work directly as we studied in this paper addresses, however, is linked to the work that is ours [1], [6]. Determinize possible answer to a question about the search in the database. End-use application that can be accessed only for the introduction of the imperative in order to keep all of the information contained above, we are concerned deterministic representation. The difference between the two issues in a variety of settings to appeal. A problem like this in the list of objects [7] The author answers the query in order to achieve development of the best in its class, to be fair. Ours is to maximize the value of the query workload, however, the purpose is to get the best value and a single query. Also, the focus is on how to choose the most excellent set, we determinize automatically because all the things from each selected object is cleaned by refining. This difference effectively to the challenges of optimization. It is estimated another relative in the area to set the graph model [7], [8], which is defined by the model along with the opportunity to discover that it is possible to be assigned to each variable. It can be seen problems determinization report to the expectations of metric map on the basis of cost as a case of a problem. In this way we can look at the issue, by testing equivalent to us NP- problem difficult to resolve is the development of a code fast and accurate high-value.

B.Probabilistic Data Models:

A range of highly developed data models have been proposed in the past. Our focus however was determinizing probabilistic objects, example image tags and speech output, for which the probabilistic attribute model suffices. We observe that determining probabilistic data stored in more highly advanced probabilistic models such as tree might also be interesting and can be possible [1]. Furthermore, our work to deal with data of such high complexity is an interesting future direction of work. There are many research efforts related that deals with the problem of selecting terms to number a document for document retrieval.

C.Key term Selection :

There are many research efforts related that deals with the problem of selecting terms to number a document for

document retrieval. A term-centric pruning method explained in keeps topmost postings for each and every term according to the individual score impact that each and every posting will have if the term is seen in an for the function search query [1]. We propose a scalable term selection for categorization of text, which is based upon coverage of the terms coverage of the terms The focus of these research efforts is based on relevance – that is, finding the correct set of terms that are most relevant to document. In our problem, a set of possibly relevant terms and their relevance to the document are already given by other data dealing out techniques. Thus, our goal is not to find the relevance of terms to documents, but to find and select keywords from the given set of terms to represent the document, such that the quality of answers to triggers/queries is optimized.

D.Query intent disambiguation:

Query information in such type of works is used to calculate many appropriate terms for queries, of queries. However, our aim is not to guess correct terms, but to find the correct keywords from the terms that are automatically generated by automated data generation tool[1].

E.Query and tag suggestions:

Another related explore area is that of query suggestion and tag suggestion [11]–[13]. On the basis of query-flow graphical representation of query information, authors in [11] develop a measure of semantic similarity between queries, which is used for the task of producing diverse and useful recommendations. Rae et al. [12] introduces an extendable structure of tag suggestion, using co-incidence examination of tags used in user detailed contents such as personal, social contact, social group and non user specific contents. The main objective of this is on how to make similarities and correlations between queries/tags and recommend queries/tags based on those information. However, our aim is not to measure similarity between object tags and queries, but to select tags from a given set of uncertain tags to optimize certain quality metric of answers to multiple.

III.CONCLUSION :

Hence, from this paper we have considered problem of determinizing uncertain objects in order to organize and

store such data in already existing systems example Flickr which only accepts deterministic value. Our aim is to produce a deterministic depiction that optimizes the quality of answers to queries/triggers that execute over the deterministic data representation .As in future work, we plan to perform project on efficient determinization algorithms that are orders of scale faster than the enumeration based best solution but achieves almost the same excellence as the optimal solution and search determinization techniques as per the application context, wherein users are also involved in retrieving objects in a ranked order.

REFERENCES :

- [1] K. Jie Xu, Dmitri V. Kalashnikov, and Sharad Mehrotra, “Query Aware Determinization of Uncertain Objects,” IEEE Trans. Knowl. Data Eng., vol. 27, no.1, Jan. 2015.
- [2] D. V. Kalashnikov, S. Mehrotra, J. Xu, and N. Venkatasubramanian, “A semantics-based approach for speech annotation of images,” IEEE Trans. Knowl. Data Eng., vol. 23, no. 9, pp. 1373–1387, Sept. 2011.
- [3] J. Li and J. Wang, “Automatic linguistic indexing of pictures by a statistical modeling approach,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 9, pp. 1075–1088, Sept. 2003.
- [4] C. Wang and F. Jing, L. Zhang, and H. Zhang, “Image annotation refinement using random walk with restarts,” in Proc. 14th Annu. ACM Int. Conf. Multimedia, New York, NY, USA, 2006.
- [5] B. Minescu, G. Damnati, F. Bechet, and R. de Mori, “Conditional use of word lattices, confusion networks and 1-best string hypotheses in a sequential interpretation strategy,” in Proc. ICASSP, 2007.
- [6] R. Nuray-Turan, D. V. Kalashnikov, S. Mehrotra, and Y. Yu, “Attribute and object selection queries on objects with probabilistic attributes,” ACM Trans. Database Syst., vol. 37, no. 1, Article 3, Feb. 2012.
- [7] V. Jojic, S. Gould, and D. Koller, “Accelerated dual decomposition for MAP inference,” in Proc. 27th ICML, Haifa, Israel, 2010.



International Journal of Research in Advanced Computer Science Engineering

A Peer Reviewed Open Access International Journal

www.ijracse.com

[8] D. Sontag, D. K. Choe, and Y. Li, "Efficiently searching for frustrated cycles in map inference," in Proc. 28th Conf. UAI, 2012.

[9] S. Bhatia, D. Majumdar, and P. Mitra, "Query suggestions in the absence of query logs," in Proc. 34th Int. ACM SIGIR, Beijing, China, 2011.

[10] C. Manning and H. Schutze, Foundations of Statistical Natural Language Processing, Cambridge, MA, USA: MIT Press, 1999.

[11] I. Bordino, C. Castillo, D. Donato, and A. Gionis, "Query similarity by projecting the query-flow graph," in Proc. 33rd Int. ACM SIGIR, Geneva, Switzerland, 2010.

[12] A. Rae, B. Sigurbjörnsson, and R. V. Zwol, "Improving tag recommendation using social networks," in Proc. RIAO, Paris, France, 2010.

[13] B. Sigurbjörnsson and R. V. Zwol, "Flickr tag recommendation based on collective knowledge," in Proc. 17th Int. Conf. WWW, New York, NY, USA, 2008.