# Refined Ontology Model for Content Anatomy and Topic Summarization

**A. Mekala [1], Dr.C.Chandra Sekar [2]**

**[1] Manonmaniam Sundaranar University, Tirunelveli**

**[2] Department of computer science, Periyar University, Salem**

**Abstract**

When the performance of any information processing system can be enhanced by the concept of ontologies, domain specific terms enclosing wealthy and defined semantics. Research has been accomplished with the help of variety of resources on automatic ontology construction. Each of these resources has different qualities that have need of special approaches to term and relationship extraction. On the consideration of terminological resources, semantic structure of ontology construction facilitates the NLP (Natural Language Processing) that extracts terms and relationships. Generally in this phase there can be a problem in that many relationships are incorrectly defined or applied excessively. For that reason, extracting ontological relationships from documents necessitates data cleaning and refinement of semantic relationships. In our research we provide the automatic term relationship and refinement ontology construction for the content anatomy and topic summarization. Where the automatic topic extraction mechanism will be done based on the significant score computation and the highest score will be the topics. Our proposed system supports effective joint inference approach, which simultaneously constructs the ISA (is a) and HASA (Has a)-tree, while mapping Topic models to WordNet, achieves the best performance. To end with, we estimate our ontology-based topic summarization results that formulate exploit of similarity-based metrics first enlarged for automatic term relationship findings and refinement of semantic relationships. The experimental result shows that the proposed system produce the better summarization result when compared with the existing methods.

***Keywords:*** *Semantic Relationship Refinement, Noun Phrase Analysis, Semantic Web, Ontology, ISA (is a) &HASA (Has a)-tree, WordNet, Natural Language Processing, Automatic term relationship detection, Information Extraction.*

## 1. Introduction

In today's world the quantity of information generation is increasing enormously by way of each day. In electronic form the amount of information being generated there has exactly been an ignition in and exposed through the World Wide Web. In the field of research and development this is extraordinarily spot-on where the frequency of the number of papers and articles being distributed is supplementing every day. This presents the requirements for users to be capable to browse through several different documents and rapidly discover the information. A précis submits to an abbreviated or a concentrated translation of a document. It is a brief and to the point depiction of the unique document exactness the most imperative positions enclosed surrounded by it thus removing the feel like to have to understand the full text. For the research purpose have to frequently undergo many research papers with the most spontaneous mode to refer that by reading the abstract and the conclusion together summing up the whole concept. But in various documents they do not automatically contain abstracts as part of it. The abstract is representing the overview of what the document is explaining about and does not essentially list out the most important ideas in a line. So there is a need of sorting out the documents can be achieved if only the summaries can be produced involuntarily to the user who is deciding the documents is useful before reading the whole thing that avoid the time complexity. While reading the news articles situation the summaries can be useful that help the readers can browse through the most important phase of the article as an alternative of reading the whole length article from this we can say that the summarization systems needed mostly in various situations.

That the summarization is compresses the large amounts of source information at the same time they maintain the aim of producing the consolidated main contents of the documents. Next to condensing information and eliminating redundancies this can be used to summative and collecting the information from different source documents with this highlighting the similarities and differences and produce the mostly concentrated information. These advantages are balanced through the fact that summaries can act as filter for irrelevant source information that is based on the user's need. These prospective benefits have motivated in the task of

growth of involuntarily reachable information formulates the automatic summarization for the effective information retrieval mechanism. When there is a lot superior quantity of controlled data obtainable to influence sophisticated applications the

because it is both inclusive and high eminence. With the aim of efficiently utilize extracted data using the spotless and reliable ontology.

Ontology enlargements along with inhabitants are assignments of dominant consequence in semantic web applications. The physical presentation of these responsibilities is labour-exhaustive and consequently cost-intensive, and yields from a maximum level of computerization. For this purpose, the recognition and taking out of terms which is a very important first step to facilitate obtain part in an important role in the field under concern. Essential model of lot of knowledge based applications are mainly depends on the Automatic term recognition which is also known as term extraction that applications such as automatic indexing, terminology mining, knowledge discovery and monitoring, knowledge management and biomedical domains. In this research which is mainly based on the automatic term extraction that is used to construct the refined ontology. Among the term recognition and information extraction there is a comparatively observable meaning moreover appear for other types of information than terms also it may not be always focused on specific domain. Conventionally, the term recognition mainly focused on statistical method; at the same time information extraction is based on machine learning methods.

In our research paper the input topics of the documents will be split as a paragraph where the association rule between each of the topics and paragraphs is achieved. As of the Web the input documents acquired which are normally noisy, necessitating deduplication and stemming and stop word removal is done before they could be organized into ontology. Let denote the input documents as $D = \{d_1, d_2 \ldots . d_k\}$. We define the consecutive sentences of $D$ as $w$ blocks where let denote the topic as $T = \{t_1, t_2, t_3 \ldots \ldots t_m\}$ are a set of stemmed terms removal of stop words. The topic is described as $m \times n$ which is called as term block association matrix B in that the columns represented blocks which can be $\{b_1, b_2, b_3 \ldots . b_n\}$ decomposed from the topic documents. The association between each of $T$ and $D$ will be analysed and the related terms will be

automatic summarization through the quick

visualization of a Semantic Web only can be recognized. In this we can formulae the feasible automatic search scheme using the machine learning trained extractors which is logical source for extraction

considered by referring the WordNet and its structures makes it is a tool to Natural Language Processing for computational linguistics. In that the Noun phrase relationship will be analysed and forming the tree structure with ISA and HASA relationship by the way the ontology model is reconfigured. Through this the process of TSCAN model is achieved with automatic term extraction process to produce high accuracy in summarized result. The main contribution of the work is as follows:

1. First, we look for the automatic ontology construction. The association relationship is checked with each of the topics and the paragraph of the documents. In that the automatic topic extraction is based on the significant score calculation where the highest score of the terms will be extracted as topics.
2. Second, to attain the refined ontology it should contain a precise ISA tree and the HASA tree, where individual modules are semantically separate and accepted modules are well signified.
3. Third, each topic model should be defined with a well-off scheme, citation an inclusive list of instructive attributes. Topic models should be populated with numerous instances. We note that, while Topic models have rich schemata, many duplicate topic models and attributes exist. This is automatically removed.
4. At last, with the effective refined ontology model the topic summarization is examined and the experimental results are evaluated.

The rest of the paper is organized as follows: In section 1 the introduction about the paper is explained. In section 2 the related work is discussed. In section 3 the proposed work of refined ontology construction for the topic summarization is explained. The experimental results are placed in section 4. In section 5 the conclusion part is done.

## 2. Related Work

Ontology requires to be constantly modernized through new a perception which is to be a useful tool, relations and lexical resources. Consecutively we refer the several methods and techniques used to refine the domain ontology. It mainly spotlights on approaches appropriate to the researches on lexical attainment and linguistically aggravated mining. For the past, this comprises mounting algorithms and arithmetical methods for satisfying the gaps in obtainable machine legible dictionaries through seeming at the occasion models or declarations in huge content corpora. For the afterward this contains, review linguistic advances to text mining with more linguistic and semantic information useful to colonize ontology. Nicola Guarino [1] dealt with for reasons of information retrieval and extraction, the notional concerns associated to the devise and utilize of such ontologies. Later than a conversation on the character of semantic identical inside a model-theoretical structure, introduced the theme of reserved Ontology, viewing how the thinking of ingredient cover, reliability, uniqueness, and reliance can be of assist in perceptive, organizing and formalizing original ontological features where no background domain knowledge is offered. Marc Ehrig and Steffen Staab [2] believed QOM, Quick Ontology Mapping, as a technique to changeover between effectiveness and efficiency of the mapping creation algorithms also the QOM has subordinate run-time density than existing important advances.

Nenad Stojanovic [3] presented an application of the logic-based query refinement in the incisive for information in an information gateway. The refinement approach is supported on the detection of fundamental associations between queries concerning the addition relation between the responds of these queries. A formal model defined for the query-answering twosomes and employ techniques as of the inductive logic programming for the competent computation of a (lattice) sort between them. In a case revise demonstrated the reimbursement of with this approach in the conventional information retrieval missions. Chris WELTY [4] presented here OWL ontology on behalf of the essential OntoClean divisions, and a tool and method for concerning it to OWL ontologies. Here temporarily touched on the semantic problems oblique by using OWL Full syntax to distinguish the OntoClean meta-properties as properties of OWL Classes, and how that was explained to utilize an off-the-shelf OWL DL reasoner to ensure the OntoClean limitations on the classification. S'everin Lemaignan [5] presented an offer for a developed upper ontology, expected to summary an ordinary semantic mesh in mechanized domain. Convenience of ontologies for data formalization and distribution, particularly in an industrialized environment, are initially conversed. Particulars are given regarding the Web Ontology Language (OWL) and its adequation for ontologies in the modern systems.

Lina Zhou [6] provided an inclusive appraisal and conversation of foremost concerns, confronts, and opportunities in ontology learning. An innovative learning-oriented model proposed for ontology expansion and a structure for ontology learning. Furthermore, for classifying ontology learning advances identified and argued important proportions and techniques. In explanation of the conflict of pasture on choosing ontology erudition progress, recapped domain characteristics that can assist potential ontology knowledge endeavor. Yihong Ding [7] presents a generic architecture for automated ontology reuse. With our implementation of this architecture, we show the practicality of automating ontology generation through ontology reuse. Jens Dietrich [8] propose a novel approach to the formal definition of design patterns that is based on the idea that design patterns are knowledge that is shared across a community and that is by nature distributed and inconsistent. By using the web ontology language (OWL) we are able to properly define intend patterns and some connected notions such as pattern contestant, pattern refinement, and model occasion. Diana MAYNARD [9] described a method for word recognition using linguistic and statistical methods, construction exercise of appropriate information to bootstrap knowledge. Investigated afterward how term recognition techniques can be functional for the wider task of information extraction, production exploit of similarity metrics and background information. Two tools are described that have urbanized to formulate exploit of contextual information to assist the development of regulations for named entity recognition.

Achim Rettinger [10] the reasonable restrictions assumed from ontologies can be exploited to improve and manage the learning task by enforcing depiction logic satisfiability in a latent multi-relational graphical model. To show the possibility of the approach tests using real world social network data in form of SHOIN (D) ontology provided. Lei Liu [12] presented an iterative method extorting ISA relations as of large text for ontology learning. Routine acquisition of ISA relations is an essential trouble in knowledge acquisition from text. Initially, it determines a set of stretches using numerous special lexico-syntactic patterns from free text corpus. Secondly combine exterior coating elimination and in the interior coating gathering for acquiring concepts of constituting ISA relation. Asanee Kawtrakul [15] presented a hybrid advance for (semi-)automatically sensing the challenging relationships and for signifying extra specifically distinct ones. The system consists of three main components: Rule Acquisition, Detection and Suggestion, and Verification. The Refinement Rule Acquisition module is employed to acquire rules exacting by experts and throughout machine learning. The discovery and proposal constituent employs noun phrase analysis and WordNet position to intellect inaccurate relations and to propose more suitable ones supported on the application of the acquired rules. The confirmation module is a tool for proving the proposed associations. From the above mentioned methods there is always lacking in ontology reconfiguration model. These methods only considering the ISA relationship were also other relationships not considered and the automatic topic extraction is not done in all these methods. So there is need of ontology refinement with automatic topic extraction with that this will be summarized.

# 3. Proposed Methodology

## 3.1 Automatic Topic Extraction Model

In this phase of implementation model, initially the correlated keywords will be found out via WordNet tool. Calculating similarity values between the topics by means of WordNet. Commencing those topics recognize related topics, when finding document using swarm intelligence techniques those topics are exercised. In this technique applied TF-IDF as weighted values of terms and we measured relative terms during the penetrating time merely. In this approach the weight values of term is calculated with significance with other terms other than if apply the subsequent developments. This significance supports for score computation which is efficiently civilizing the tf-idf based computations. When first forthcoming an unidentified area or necessities document, to obtain a rapid clasp of what the important ideas and entities in the area are, it is frequently productive. This procedure is described as concept recognition, where the theme concept submits to a thing or notion that has an exacting consequence in the area. The most important reason of relating statistical techniques for concept recognition is to grade applicant abstractions based on an exacting measure that provides advanced attains to probable abstraction applicants.

The majority ordinary statistical technique is to suppose the implication of an applicant term from the number of periods it happens in the document. There is a meticulous dispute connected with multitopic terms because the majority methods, together with corpus-based frequency profiling, rely on recognizing individual topics, and count up these separately. There are collocation analysis techniques that can suppose lexical affinities; conversely, while mainly relationship procedures are distinct to compute the pair-wise devotion of topics $(t_i, t_j)$ merely, they cannot be trained for measuring the relationship among more than two topics. In supplies construction, it is reasonably recognizable to assemble domain words, such as software requirements measurement, that encompass more than two topics. Suitably managing such successions is consequently an imperative dispute, as a number of researchers maintain that in particular value. Although multitopic terms can be recognized is key difficulty, in abstraction recognition desire to grade words in order of the significance of their signified abstractions.

In terms of pure frequency, it is common for significant multitopic terms to occur relatively occasionally in a document. Not as good as, no normative corpus of which we are conscious encloses large numbers of multitopic terms. This is for the reason that the majorities such terms are precise to meticulous areas and therefore are doubtful to discover their method into a corpus whose position is to provide as a direct to universal practice of a speech. Therefore, while the corpus-based frequency profiling technique described above works well for terms that are single topics, in practice it doesn't help with multitopic terms. To explain this problem, synthesize a significance value for all terms using a heuristic based on the number of topics of which the term is composed, and the LL (Log-Likelihood) value for each topic. In its simplest form, the significance value for a term $T = \{t_1, t_2, t_3 \ldots\ldots t_m\}$ is specified by the formula which is as follows: $S_T = \frac{\sum_i LL_{w_i}}{l}$

Basically computes the mean of the LL values for all the part topics comprising a multitopic term. Nevertheless, conjecture that not all the topics supply evenly to the implication assessment of the multitopic term of which they are a constituent. The suggestion is supported on supposition that such a term is normally created of a head topic and one or more modifiers. Presuppose that the head topic is the majority important constituent of the term; thus the term is extra important, and the LL value of term should bring more weight than the LL of term. To put up the assumption, the implication equation is modified to incorporate a weight, $k_i$ which dispenses a weight to each topic that is a constituent of the term $S_T = \frac{\sum_i k_i LL_{w_i}}{l}$

It merges a number of obtainable natural language processing (NLP) methods in a work of fiction method to facilitate it to handle both single and multitopic terms, ranked in order of confidence. The evaluation method that employ for Relevance-driven Abstraction Identification (RAI) is one of the main offerings of the occupation, which circumvents the difficulties associated with using expert creature judgment for evaluating how fit the terms revisited against the trouble domain's fundamental abstractions. The significance based topic identification is realized by the subsequent practice:

1. Each topic in the domain document is interpreted with the use of Tree dragger (or) Stanford parser.

2. The set of topics is sorted to take away frequent topics doubtful to suggest concepts.

3. The left over topics are lemmatized to decrease them to their dictionary structure, to subside inflected shapes of topics to a support form.

4. Each theme is dispensed LL rate by concerning the approach described above, using the number of topics which is collected from Wikipedia.

5. Syntactic patterns are functional to the text to categorize multitopic terms.

6. A significance score is derivative for each term by applying the formula of:

$$S_T = \frac{\sum_i k_i LL_{w_i}}{l}$$

7. Recognized topics are sorted which depends on their significance score and the consequential list is revisited.

## 4. Refined Ontology Construction For Topic Summarization

Ontology refinement intends to fine adjust the ontology and is one of the reconfiguration method. In this step, all compositional sources are present to populate the ontology and refinement has to rely on amorphous sources like documents. Where the assessment procedure relies on information taken out from text for which IE (Information Extraction) is utilized. The subsequent segment specifies these IE tasks and presents the state of the art approaches. The specific necessary tasks are, if a specified word is a synonym of an existing perception occupies to find out the identification of the perception signified by the term and the aspirant synonym notions. If no specific perception is establish, a new impression is fashioned. To find the close relative of a given model we move onto the significant score calculation. If there is a relation between a concept and any other concept in the ontology we will move onto the ISA and HASA relationship. These tasks rely on recognizing portions of information applicable for a specified circumstance defined by the word we are meting out. An assortment of quotations is completed by IR (Information Retrieval). With this, the sentences belonging to the recovered documents will be graded by consequence each time an appropriate form can be created. This model will depend on the mining patterns that we propose to exercise. If a suitable approach for sentence ranking is not establish, a Boolean expression maintained on the words in the representation is employed to recover only on probable suitable sentences and speed up the extraction system.

In this model from the topic extraction phase the each topic will be associated with the set of paragraphs of the input document. There can be the associative block matrix will be formed. After that the related term of will be found out from the WordNet tool. From that the ontology (and the procedure utilized to generate it) will be formed and it must convince numerous criteria. First, we look for automatic ontology construction. Second, the ontology should contain a well-defined ISA hierarchy and HASA hierarchy, where individual classes are semantically separate and ordinary classes are well signified. Third, each class should be defined with a wealthy schema, listing a complete list of instructive attributes. Classes should be populated with numerous instances. Where the relationship and the frequency of each topic is found out within itself and the higher values of topic will be the parent and other will be formed based on that as a child. Then the ISA and HASA relationship tree structure is formed. In this way we are refining the ontology model. By computing a mapping from WordNet concept nodes, for illustration, if both c and d have completely equivalent nodes in WordNet and one WordNet node includes the other (say isaFT(c,d,isaWN)& say has aft(WN)), then this is probable to be extremely predictive for a beginner. Because computing the mapping to WordNet is involved. This procedure is given in Fig 1.

IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 1, January 2013
ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814
www.IJCSI.org


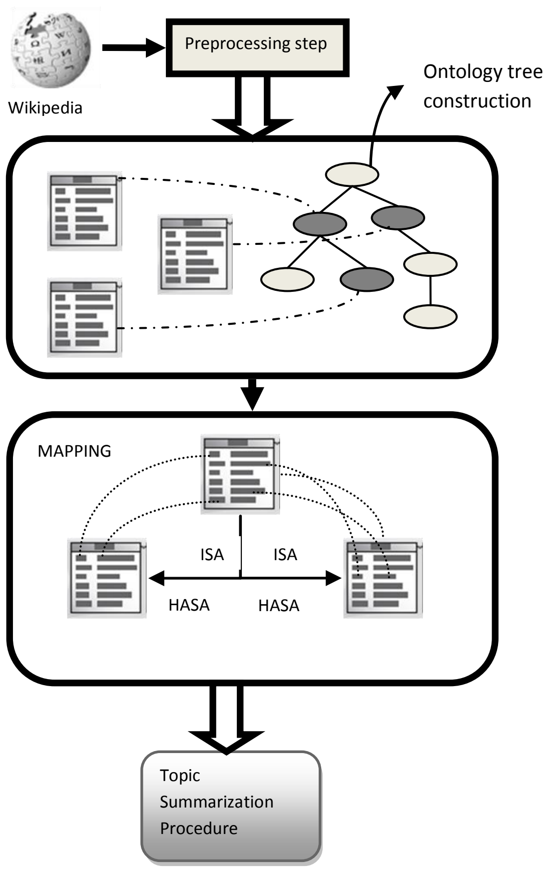


Fig 1: The Overall Architecture Diagram

The procedure is given as algorithm as follows:

$Refinement\ (t_1 \ldots, t_m, Rel)$

$Input: \{t_1, t_2, t_3 \ldots\ldots t_m\}, rel$

$Output: NewRel$

$For\ (i = 0; i < m; i++)$

$\{$

$If (Rel = isa\ |\ Rel = hasa)$

$Then\ if\ \ Agree\ except\ (t_i, Rel)$

$Return\ new\ refined\ relationship\ (NewRel)$

$else\ if\ parent\ is\ compatable(t_i)$

$then\ return\ parent\&child\ relationship$

$else\ if\ is\ wordnet\ hyername\ path(t_i)$

$then\ return\ \ parent\&child\ relationship$

$else\ if\ agree\ revision\ rules(t_i, Rel)$

$Then\ return\ new\ relationship$

$\}$

$End\ the\ result$

## 5. Experimental Results

An experiment testing with the training rules technique using some examples for few semantic relationships. Using topics which are automatic extraction from the documents, done the noun phrase analysis, using WordNet for the experimental tests. Where the precision, recall rate and Fmeasure is measured and analyzed with existing methods. Each will be analyzed and described as follows.

### 5.1 Precision Rate

We analyze and compare the performance offered by Ontology & NonOntology with Automatic Topic Extraction and Refined ontology with Automatic Topic Extraction. Here if the number of document size increased the precision accuracy also increased linearly. The precision accuracy of the proposed method is high. Based on the comparison and the results from the experiment shows the proposed approach works better than the other existing systems with higher rate. The values are represented in the Table 1.

Table 1: Precision vs. Number of Documents

| S. No | No. of Documents | Ontology & NonOntology with Automatic Topic Extraction | Refined ontology with Automatic Topic Extraction |
|---|---|---|---|
| 1 | 10 | 0.75 | 0.85 |
| 2 | 20 | 0.69 | 0.75 |
| 3 | 30 | 0.58 | 0.64 |
| 4 | 40 | 0.45 | 0.56 |
| 5 | 50 | 0.39 | 0.49 |
| 6 | 60 | 0.33 | 0.47 |

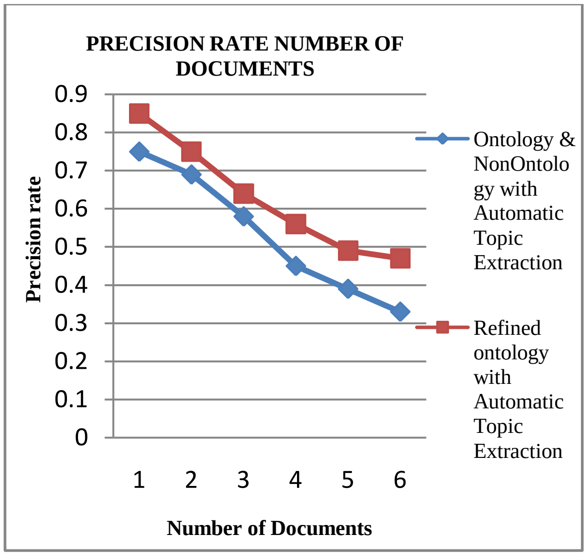


Fig 2: Precision Rate Number of Documents

In this graph we have chosen two parameters called number of Document and precision which is help to analyze the existing system and proposed systems. The precision parameter will be the Y axis and the number of document parameter will be the X axis. The blue line represents the proposed system and the red line represents the existingsystem. From this graph we see the precision of the proposed system is higher than the existing system. Through this we can conclude that the proposed system has the effective precision rate.

## 5.2 Recall Rate

This graph shows the recall rate of existing and proposed system based on two parameters of recall and number of Document. From the graph we can see that, when the number of number of Document is improved the recall rate also improved in proposed system but when the number of number of Document is improved the recall rate is reduced in existing system than the proposed system. From this graph we can say that the recall rate of proposed system is increased which will be the best one. The values of this recall rate are given below:

Table 2: Recall vs. Number of Documents

| SNO | Number of Documents | Ontology & NonOntology with Automatic Topic Extraction | Refined ontology with Automatic Topic Extraction |
|---|---|---|---|
| 1 | 10 | 0.75 | 0.87 |
| 2 | 20 | 0.67 | 0.77 |
| 3 | 30 | 0.58 | 0.69 |
| 4 | 40 | 0.48 | 0.58 |
| 5 | 50 | 0.41 | 0.51 |
| 6 | 60 | 0.38 | 0.48 |

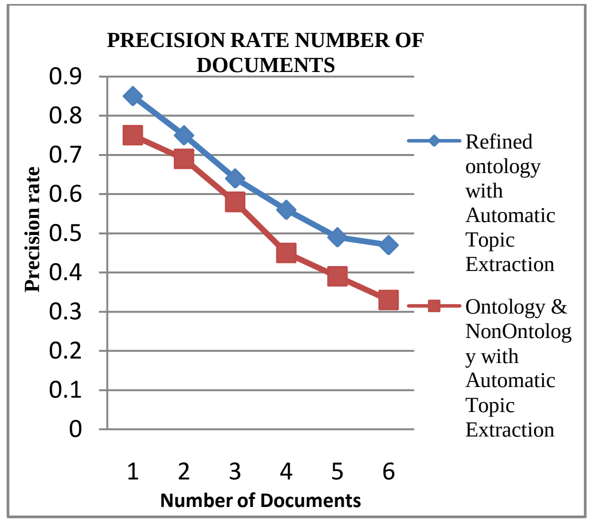


Fig 3: Recall Vs. Number Of Documents

In this graph we have chosen two parameters called number of Document and recall which is help to analyze the existing system and proposed systems on the basis of recall. In X axis the Number of document parameter has been taken and in Y axis recall parameter has been taken. From this graph we see the recall rate of the proposed system is in peak than the existing system. Through this we can conclude that the proposed system has the effective recall.

## 5.3 Fmeasure Rate

This graph shows the Fmeasure rate of existing and proposed system based on two parameters of Fmeasure and number of Document. From the graph we can see that, when the number of number of Document is improved the Fmeasure rate also improved in proposed system but when the number of number of Document is improved the Fmeasure rate is reduced in existing system than the proposed system. From this graph we can say that the Fmeasure rate of proposed system is increased which will be the best one. The values of this Fmeasure rate are given below:

Table 3: Fmeasure vs. Number of Documents

| SNO | Number of Documents | Ontology & NonOntology with Automatic Topic Extraction | Refined ontology with Automatic Topic Extraction |
|---|---|---|---|
| 1 | 10 | 0.67 | 0.75 |
| 2 | 20 | 0.56 | 0.69 |
| 3 | 30 | 0.48 | 0.62 |
| 4 | 40 | 0.4 | 0.53 |
| 5 | 50 | 0.34 | 0.45 |
| 6 | 60 | 0.23 | 0.39 |

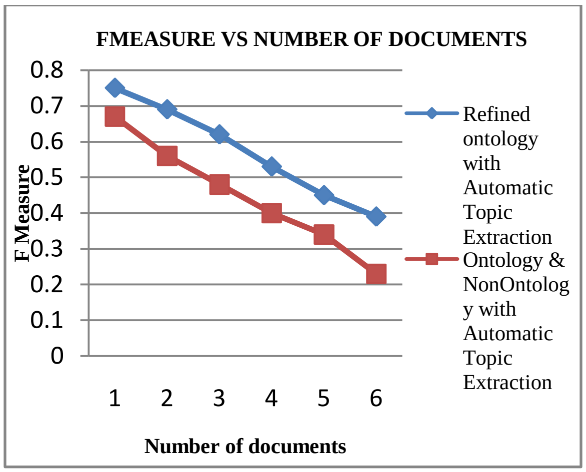


Fig 4: Fmeasure vs. Number of Documents

In this graph we have chosen two parameters called number of Document and recall which is help to analyze the existing system and proposed systems on the basis of Fmeasure. In X axis the Number of document parameter has been taken and in Y axis Fmeasure parameter has been taken. From this graph we see the Fmeasure of the proposed system is in peak than the existing system. Through this we can conclude that the proposed system has the effective Fmeasure.

## 6. Conclusion

This manuscript presents the three methodologies for data preprocessing and semantic relationship refinement and ontology refinement to resolve the difficulty of producing well-defined semantics from inadequately distinct or underspecified semantics in documents. The organization refines the semantic associations although noun phrase analysis, WordNet alignment, and semantic relationship topics, a number of produced by authorities and others produced from explained illustrations by an inductive statistical machine learning arrangement. Ontologies with accurate semantic are central for improving retrieval systems, for automating procedures from side to side machine reasoning, and for the Semantic Web. The refinement of ontology is very difficult to form and this methodology is efficient to make the configuration of ontology and automatic extraction of topics automatically through significant score calculation. The experimental results show that the proposed method is more efficient than the existing method.

**REFERENCES**

[1] Nicola Guarino," Semantic Matching: Formal Ontological Distinctions for Information Organization, Extraction, and Integration", LADSEB-CNR, National Research Council, Corso Stati Uniti 4, I-35127 Padova-1997.

[2] Marc Ehrig and Steffen Staab," QOM - Quick Ontology Mapping" Institute AIFB, University of Karlsruhe, 2004.

[3] Nenad Stojanovic,"On the Query Refinement in Searching a Bibliographic Database", Internationale Tagung irtschaftsinformatik (WI2005) 23.-25. February 2005 in Bamberg.

[4] Chris WELTY,"OntOWLClean: Cleaning OWL ontologies with OWL", IBM Watson Research Center, NY, USA 2006.

[5] S´everin Lemaignan, Ali Siadat, Jean-Yves Dantan, Anatoli,"MASON: A Proposal For An Ontology Of Manufacturing Domain", Proceedings of the IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06) 0-7695-2589-X/06 $20.00 © 2006 IEEE

[6] Lina Zhou, "Ontology learning: state of the art and open issues", Inf Technol Manage (2007) 8:241–252 DOI 10.1007/s10799-007-0019-5, Published online: 24 March 2007.

[7] Yihong Ding, Deryle Lonsdale, David W. Embley, Martin Hepp, and Li Xu,"Generating Ontologies via Language Components and Ontology Reuse", This work was partially funded under National Science Foundataion Information and Intelligent Systems grant IIS-0414644 2007.

[8] Jens Dietrich, Chris Elgar,"Towards a web of patterns", Massey University, Institute of Information Sciences and Technology, Palmerston North, New Zealand 2007.

[9] Diana MAYNARD, Yaoyong LI and Wim PETERS,"NLP Techniques for Term Extraction and Ontology Population", Corresponding Author: Diana

Maynard: Dept. of Computer Science, University of Sheffield, 211 Portobello St, Sheffield, UK; 2008.

[10] Achim Rettinger, Matthias Nickles, and Volker Tresp,"Statistical Relational Learning with Formal Ontologies', Technische Universitat Munchen, Germany, University of Bath, United Kingdom, Siemens AG, CT, IC, Learning Systems, Germany,2009.

[11] Elena Beisswanger,"Exploiting Relation Extraction for Ontology Alignment", Jena University Language and Information Engineering (JULIE) Lab and Friedrich-Schiller-Universitat Jena, 2010.

[12] Lei Liu, Sen Zhang, Luhong Diao, Cungen Cao."An Iterative Method of Extracting Chinese ISA Relations for Ontology Learning", Journal of Computers, College of Applied Sciences, Beijing University of Technology, Beijing, China, Vol. 5, NO. 6, June 2010.

[13] Javier Lacasta, Javier Nogueras-Iso, Jacques Teller, and Gilles Falquet,"Transformation of a keyword indexed collection into a semantic repository: applicability to the urban domain" International Conference on Theory and Practice of Digital Libraries 2011 - TPDL 2011.

[14] Farheen Siddiqui, M. Afshar Alam,"Web Ontology Language Design and Related Tools: A Survey", Journal of Emerging Technologies in Web Intelligence, Vol. 3, No. 1, February 2011 Academy Publisher doi:10.4304/jetwi.3.1.47-59

[15] Asanee Kawtrakul , Aurawan Imsombut , Aree Thunyakijjanukit c, Dagobert Soergel, Anita Liang, Margherita Sini, Gudrun Johannsen g, and Johannes Keizer, "Automatic Term Relationship Cleaning and Refinement for AGROVOC",March 2011.

[16] Ilaria Corda, Vania Dimitrova, Brandon Bennett,"An Ontological Approach to Unveiling Connections between Historical Events", School of Computing, University of Leeds, United Kingdom ilaria,vania,brandon@comp.leeds.ac.uk, September 30, 2012