JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 25, 1391-1403 (2009)

Empowering Plagiarism Detection with a Web Services Enabled Collaborative Network

BILAL ZAKA
Institute for Information Systems and Computer Media
Graz University of Technology
8010, Graz, Austria

This paper explains how collaborative efforts in terms of technology and content, can help improve plagiarism detection and prevention. It presents a web service oriented architecture, which utilizes the collective strength of various search engines, context matching algorithms and indexing contributed by users. The proposed framework is an open source tool, yet it is extremely efficient and effective in identifying plagiarism instances. By creatively using distributed processing capabilities of web services, this tool offers a comprehensive mechanism to identify pirated contents. With an aim to extend current plagiarism detection facilities, the proposed framework not only tries to reduce known deficiencies but also aims to provide plagiarism protection mechanism. The distributed indexing approach adapted in the system provides scalability to examine deep web resources. Network nodes with more focused indexing can help build domain specific archives, providing means of context aware search for semantic analysis.

Keywords: web services, plagiarism, similarity detection, IPR, collaborative web

1. INTRODUCTION

Due to mass digitization and increasing use of digital libraries, scholarly contents are more vulnerable to plagiarism and copyright infringements. Surveys conducted by various researchers [1] indicate that the use of contents without proper attribution to the original source is becoming widespread. Different prevention, detection and punishment methods are being practiced globally. Some institutes rely more on grooming and ethical motivation to fight the problem and some place more emphasis on a policing approach. However, the most effective way is a balanced combination of all methods.

Many institutes provide well documented information and rules dealing with academic misconduct and plagiarism during the enrolment phase. Honor codes are used to add moral bindings. Departments and even teachers on an individual level are undertaking efforts to educate their students, research staff and faculty. Tutorials and guides are available to explain plagiarism, citation rules, and writing standards. The other aspect of plagiarism prevention is detection and penalization. The penalties for plagiarism start from warnings, grade reductions, failing grades and can end up in suspension, expulsion or even revoking of title or degree.

Detection of plagiarism is done using a number of techniques. These techniques include stylometric analysis of writing, manual search of characteristic terms in writing and use of automation tools to compare documents for similarities, within a local corpus or across the internet. It is becoming a common practice to use software and services that

Received November 19, 2007; accepted July 19, 2008.
Communicated by Jonathan Lee, Wei-Tek Tsai and Yau-Hwang Kuo.

automate the processes of plagiarism detection. The majority of these applications are based on the document source comparison method. The detection process in such programs generally starts with submission of the suspected document to the system via a desktop application or web based form. The document is converted into plain text format removing any formatting information, images *etc*. The text information is broken down into moderately sized segments (referred to as fingerprints). The fingerprints are compared for similarity with index of available documents. Finally a report is presented highlighting the matched sources and copy percentage. The comparison index can be a corpus of local documents processed in similar way, or it may include the public/private internet.

As mentioned before, the ease with which digitized contents can be accessed accounts for the rise in plagiarism. Without a doubt, the ease of content availability is an attribution of internet usage. Naturally the most popular tools used to detect plagiarism are also built on the idea of efficiently checking for document source availability over the internet. Over the years these programs and services have indeed proven their effectiveness in educational and industrial environments. However, there is still room for considerable improvements. A survey on plagiarism [2] is a good starting point for a better understanding of various plagiarism detection strategies and strengths/weaknesses of available tools. Experimental results in the referenced survey suggest that in order to have a more precise plagiarism detection tool, the inspection system requires broader and an up-to-date content index, added semantic elements for similarity check, cross language content similarity detection and finally a mechanism to verify the findings. Existing tools following either desktop applications or software as a service approach lack these capabilities. Albert Einstein once said "The secret to creativity is knowing how to hide your sources", and yes, plagiarists today are more creative. Copied contents are often not publicly available or modified in a way which is hard to detect using existing applications and approach. Further experiments to benchmark capabilities of popular plagiarism detection services revealed that intelligent use of good search engines can greatly add value to plagiarism detection applications [3].

As an attempt to fulfill the needed requirements in plagiarism detection systems, collaborative service oriented architecture for plagiarism detection is presented. The proposed service oriented collaborative network openly available to educational community aims at extending the existing similarity check methods. It offers a seamless, combined use of multiple search services. This technique provides a broader and more context aware internet search, which proves to be more revealing than any single searching and querying approach. Collaborative authoring and indexing of document sources offered by this approach enhances the search capabilities with addition of documents not available publicly. This also provides users an option to add intellectual contents for protection against copyright infringements. Participating institutes allow access to deep web, hidden from normal search engines. The system provides multiple services for search result analysis. More services can be added to the system due to its modular nature. The user has an option to use mere text matching to deduce similarity or can apply writing structure analysis, semantic or cross language analysis. The system also offers the possibility of synonym normalization and translation in collaborative search service and peer node indexes. This adds semantic matching capabilities not possible in conventional internet searches. This system makes use of off-the-shelf tools (web services) and user

contributed contents to extend plagiarism detection. Its pluggable services constitute composite web applications offering flexibility and variety in use. Having described the basic idea behind the service oriented collaborative plagiarism detection network, the following section describes the conceptual design of the system. Section 3 describes a practical realization of the architecture and compares results of the prototype with other services. Section 4 presents future directions of work. Section 5 examines the potential of the presented approach and concludes the paper.

2. CONCEPTS BEHIND SERVICE ORIENTED COLLABORATIVE ARCHITECTURE

Service Oriented Architecture (SOA) can be described as a heterogeneous environment of applications with self describing and open components which allow inter application communication. SOA offers distributed and collaborative computing infrastructure over the network or internet. A research study for the future of flexible software [4] provides a vision of personalized, self adapting and distributed software environment. The software is structured in small simple units which co-operate through rich communication structures. The collaborative units work in a transparent way to provide a single abstract computing environment. The study shows interdisciplinary approach would be critical to developing a future vision of software. A significant proportion of software and associated data does not exist in isolation but in a political, social, economic and legal context. In order to have applications with high level of productivity and quality, it is essential that they don't have rigid boundaries but offer rich interaction and interlinking mechanisms with users as well as other applications. The service oriented approach has been in use for almost a decade and adds the aforementioned functionalities in software systems. These integration technologies exist in the form of Component Object Model (COM), Distributed Component Object Model (DCOM), Enterprise JavaBeans (EJB) and Common Object Request Broker Architecture (CORBA). However what really boosted the concept recently is the emergence of the next generation of SOA based on "web services". A framework built on top of web services will offer the extension and flexibility to plagiarism detection as described in the introductory part.

2.1 Web Service Model

"A Web service is a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-process able format. Other systems interact with the Web service in a manner prescribed by its description (WSDL) using XML based messages (SOAP or REST), typically conveyed using HTTP with an XML serialization in conjunction with other Web-related standards" [5]. Web service interaction can be either synchronous or asynchronous. Commonly available and popular internet search web service APIs use synchronous request/response communications. This approach works well in limited use environments where the web service can process a request in quickly. However, in plagiarism detection, search normally requires exploring the internet for a large number of queries (finger prints of a document) or browsing through document signatures from a number of distributed nodes.

In this scenario using asynchronous service interaction for the user is the better solution.

The proposed framework consists of a service proxy that enables asynchronous use of synchronous internet search APIs. The time independent interaction model (asynchronous) is implemented using multiple synchronous request/response web services. The first service initiates processing from the end user by sending the information parameters. The service sets an identifier of the submitted job and responds to the end user with same. The end user can then use the second service and the identifier as a parameter to check if the submitted job is complete, pending or failed [6]. The first request in asynchronous communication mode validates and acts as a buffer between the end user and the synchronous internet search service. The submitted job is processed using search and analysis services at the respective network node. The similarity detection results are stored and the job identifier status is updated for later reference of the end user. Fig. 1 shows an overview of CPDNet's service linked architecture.

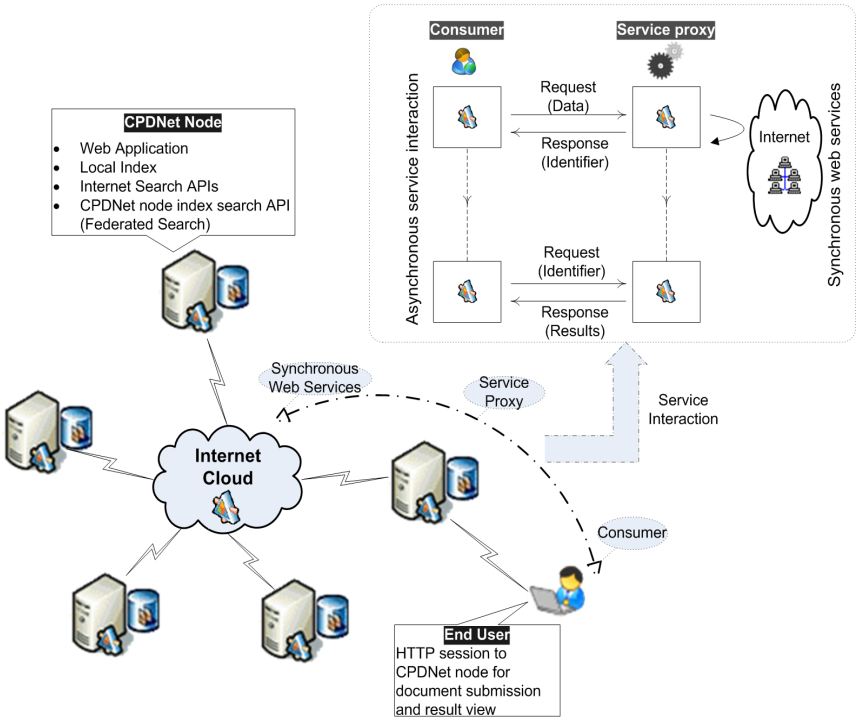

Fig. 1. Collaborative plagiarism detection network overview.

2.2 Mashup of Search and Analysis Web Services

One of the major strengths of the system is the next generation search capabilities termed "Search 2.0" by Ezzy [7]. It is defined as a search "designed to combine the scalability of existing internet search engines with new and improved relevancy models; they bring into the equation user preferences, collaboration, collective intelligence, a rich user experience, and many other specialized capabilities that make information more produc-

tive" [7]. In the concept described here, users are given the option to select a number of system compatible internet & collaborative search services. Combining the strengths and scalability of existing internet search engines broadens the web search scope compared to searching via a single source. Further mashup with collaborative search API built using full text query mechanism on user contributed finger print data and local node resources greatly add to value. The collective search is not conventional meta-search where the user might have to weed through irrelevant matches. The initial result set lists the possible matches by each search service. Analysis services applied to search results produce precise and productive output for the final report. The system has been tested using a few popular search services. The results of our experiments presented in a later section indicate that using the search services systematically can detect cut paste plagiarism more effectively then any other commercial plagiarism detection service. This is mainly because of recent open access and indexing initiatives by publishers. More and more options are becoming available to do full text search on digital libraries via a particular search engine or a library's own search mechanism. One significant example of such an initiative is Crossref search pilot. A group of 45 leading journal publishers including ACM, IEEE, Blackwell, Springer, Oxford University press and John Wiley & Sons, are providing full text search options using Google via Crossref gateway [8]. A plagiarism detection system with up-to-date search capabilities can outperform similar tools of its class. The proposed service oriented approach gives its user an option to integrate any existing search service and any upcoming more powerful search service.

The initial prototype includes an analysis services based on the vector space model [9] approach to measure cosine similarity. The queried text and search engine's returned matching snippet are converted to word vectors, based upon the vocabulary of both. The angular measure (dot product) of vectors is used as a score to determine similarity between the queried text and any searched result. The combination of the highest similarity scores of the queried text segments represents the percentage of plagiarism in a document. There is a number of other possibilities for similarity analysis within a document or with the search service's detected contents. One such analysis approach tested for the proposed framework involves a structural comparison of suspected documents. This statistical analysis service uses a measure of standard deviation in the document structures (lines, words) to determine a possible match. Another analysis planned to be part of the framework is stylometric analysis based on Jill Farringdon's CUSUM (cumulative sum) technique [10]. The CUSUM technique is based on the assumption that every person has some quantifiable writing or speaking habits. The measure of consistency of these habits can be used to determine single or multiple authorships. The numerous factors which determine authorship include checking of sentence length consistencies, checking the use of function words, nouns and other common language practice throughout the document. Future research which could be conducted on the system also includes the development of semantic analysis service that uses language ontology. The idea is further described in section 4.

2.3 Collaborative Authoring, Indexing & Searching – Access into the Deep Web

The ability of collaborative authoring and indexing at participating institute nodes of network is an important feature in extending plagiarism checks. The motive behind

collaborative indexing and search approach is the fact that conventional search engines only index the shallow internet contents and do not cover deep web contents. Shallow contents are generally static web pages linked with each other and openly available to search engine spiders. However the deep web consists of unlinked or dynamically generated pages, databases, protected sites, intranets and contents behind firewalls. These contents are invisible to the index of general internet search engines. A study by BrightPlanet in 2001 estimated that the deep web information is 500 times larger than the commonly defined World Wide Web [11]. It is very important to have access to this massive information base for thorough plagiarism checks. Educational institutes usually have a very large collection of un-linked and non-indexed local contents. Institutes and research groups within an institute also have privileged access to, and better knowledge of specific deep web resources. This access and knowledge enables them to gather resources not commonly available. Collaborative networking provides the means of creating a local searchable index of these resources. Any network node run by an institute can setup a search gateway service providing access to its invisible contents and can access to protected digital libraries. A collaborative search API consumes the local search services according to the access policy of each peer node. The collaborative search access produces limited results usable for similarity analysis services. The search results may only contain specific matching portion and associated meta information. A local index can be restricted to a specific domain *e.g.* an institute specializing in computer science articles. This means that in addition to general internet search services; the proposed system also use collaborative search service which harnesses the deep web contents of participating nodes.

The collaborative search channel is also important in terms of reducing the dependency of certain search engines. Researchers have shown concern in recent studies that the search engine monopoly gives them the role of gatekeeper in the information flow. A search engine can determine what is findable and what is kept outside the view of the common user [12]. The view restriction or any other implication a search engine may apply or is applying can be seen in the form of web search API switching from Google. Shifting from an XML standard and generic SOAP based access to a more restraining AJAX API is not seen as a welcome move by many research and development forums. It is thus imperative to have an alternate and more open channel of searching the intellectual content base. System users can contribute documents to the associated network node for indexing. User contributed content are stored either by conventional indexing and making complete documents openly available, or by storage in more abstract form (normalized). In the case of a search match, only a content snippet and meta information are sent from the index, not the complete document. Any matches found in such snippets point to the original source for further verification. Authoring resources can be tempting for a user or node administrator, because of following reasons:

(1) Contributing resources can expose the contents to all network search APIs in a protective manner. This approach helps where users cannot index complete contents in a finished formatting for the public internet.
(2) User contributed authoring acts as a "personal copyright tool" which protects against any possible piracy of articles, personal blogs, assignments, presentations *etc*. Submitted documents can be indexed with the author's meta information. Users or node

administrators may choose to receive alerts produced by any similarity matches from other sources in the future. This can help authors keep track of legal or illegal use of their contents.

(3) The envisioned framework in its mature form is based on P2P incentive based resource access scheme. Users and nodes with a higher index of shared resources will receive better access to local resources of peer nodes.

2.4 Service Publication, Discovery and Access Mechanism

Web services for end users are available as selectable index of compatible searching APIs. No technical details or WSDL is required at the end user level. User can choose any service by simply selecting or providing personal usage credentials *e.g.* API code or key. Master nodes keep a well descriptive index of available services to share and search. The system administrator of each node can incorporate the available services on a specific node and make them available to the end user. The local document source (collaboratively authored) sharing services at each node uses either an open access policy or implements restrictions on access. Peer nodes may contribute more search services and sample service consuming codes to master service index. Initial implementation uses a plain index approach and open access policy at selected test nodes. Later stages of the project include a more controlled central registry to maintain service descriptions and access mechanisms.

3. IMPLEMENTATION

Based on the abstract architecture described in the previous section, a partial implementation is developed as a proof of concept. The prototype serves as a valuable tool to benchmark and test the search engine capabilities, match detection algorithms and document source generation. Initial experiments show very promising results closely comparable (better in some cases) to already existing commercial services which detect plagiarism. Prototype named CPDNet[1] (Collaborative Plagiarism Detection Network) is available for test purposes, although it is an early stage of development. Users may register for an account with their personal Search API code to test drive the system. CPDNet currently supports Google SOAP search API, and Microsoft Live Search API. Existing CPDNet nodes use Lucene[2], an open source Java based indexing and search technology. Search result sets from various discovery services are generated in OpenSearch[3] Standard. The collaborative search client uses a SOAP interface to discover matches from all available service nodes. The process of detecting plagiarism includes the following steps:

(1) The submitted document is broken down into moderately sized text chunks also called fingerprints. This document source can also be marked for indexing in the local database, accessible via a collaborative search API. System allows the flexibility of configuring automatic fingerprint generation. The text chunks can be based on sentences or sequential word group limited by a character length. The existing prototype uses text segmentation based on character length. System split text at word boundaries (spaces) at or before specified number of characters limit. In experiments

[1] Collaborative Plagiarism Detection Network: http://www.cpdnet.org.
[2] Lucene Java: http://lucene.apache.org/.
[3] OpenSearch: http://opensearch.a9.com/.

with prototype a fingerprint length of 200 characters is used. This particular size is selected because common web search APIs produce a matching result snippet/description segment of maximum 255 characters including some joining characters. The fingerprint size can be changed in accordance with search-indexing parameters and document size. Furthermore the text segment can be refined using stopword removal and/or text normalization process. This is further explained in section 4.1.

(2) The plagiarism check is initiated by querying the internet using the fingerprint data. The selected APIs search the web. Locally indexed document sources (signature in more abstract form) and that of peer nodes can also be queried if collaborative search service is enabled.

(3) Most relevant matches obtained via the search services are passed to similarity analysis service. The existing active service uses word vector based similarity detection as described earlier.

(4) Fingerprint similarity scores of a document are calculated to find the plagiarism percentage. The document text linked with the similarity scores, matching contents and source links, is presented to the user within a final report.

The described process steps are visualized in Fig. 2.

The architecture is flexible to accommodate numerous services at each level. The running services in the current prototype can be further explored at the project portal.

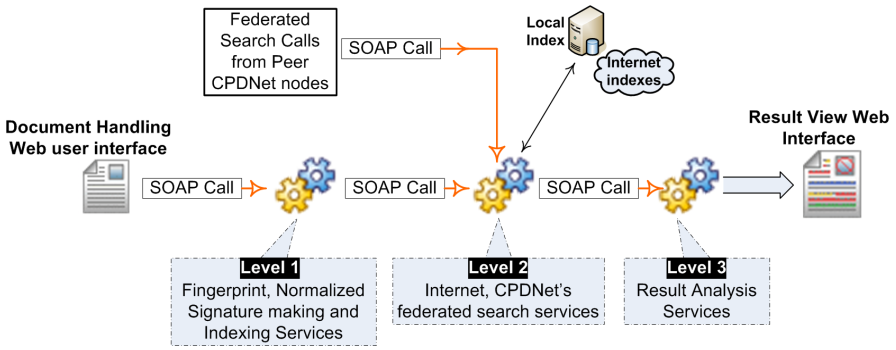

Fig. 2. Web service flow in CPDNet.

3.1 Results of Prototype

In order to benchmark the system, a document corpus is generated with various proportions of plagiarized contents from both deep and shallow internet. The test document set consist of undergraduate student assignments, and manually tailored documents that contains plagiarized contents from access restricted intellectual archives such as IEEE Xplore, ACM and SpringerLink Digital Library. Search APIs of Google, Microsoft Live, Yahoo were used for coverage of internet public index (standard check). For testing standard and conceptual plagiarism checks a local index of 1174 documents is created. The documents come from 164 issues of an online digital journal[4]. Table 1 shows statistics of standard and normalized local index maintained using Lucene based CPDNet node.

[4] JUCs: http://www.jucs.org.

Table 1. Index statistics.

Type of index	# of documents	# of terms	Size (KB)
Standard	1174	388467	23949
Normalized	1174	366013	23391

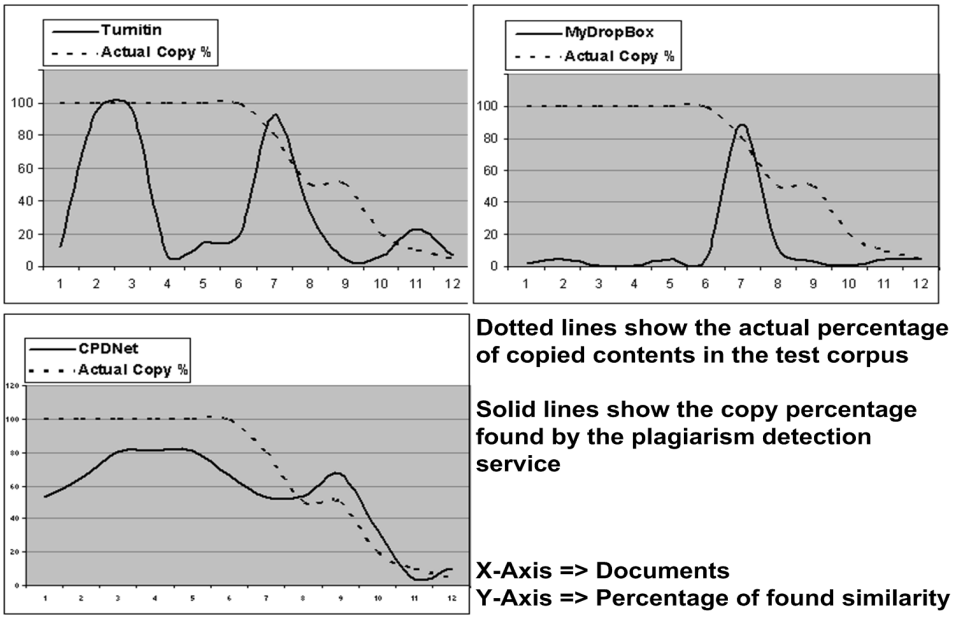

Fig. 3. Comparison of search and similarity detection capabilities.

Test results from the selected corpus show significantly better similarity detection capabilities of the system compared to other services. The graphs in Fig. 3 give an overview of the plagiarism detection capabilities of CPDNet. Better plagiarism detection by the system developed to date, is due to the enhanced SOA based searching capabilities. Compared to other systems that claim to use their own index and search mechanism this system makes use of broader more up to date discovery approach with multiple search services. Further results of an early benchmarking experiment with details of test document corpus are presented in referenced work [3].

4. TOWARDS A SEMANTIC PLAGIARISM DETECTION SERVICE

To trace paraphrased and cross language plagiarism, algorithms are required to discover similarities on the semantic level. This kind of analysis requires detecting similar word replacement (synonymizing), word deletion, word addition and translation *etc*. The application of these checks on a large scale with conventional internet indexes and current search APIs seems far fetched and computationally very expensive. However, the proposed framework provides a mean of indexing submitted contents in a normalized

form. The normalized contents which are shared at each node can be queried using a collaborative search API of peer nodes. The queried finger print is also normalized in order to determine its conceptual equivalence. The semantic level similarity check can certainly help its users in more then just plagiarism detection. The findings can also be used by knowledge workers to discover relevant contents already available on internet. In the near future, the focus of this project's research will include following:

4.1 Fingerprint Normalization into Generic Signatures

A system component for generation and indexing of signatures (semantic fingerprints) is being developed and tested. This component will normalize the submitted contents to a root level with the help of a POS (Part of Speech) tagger and language thesaurus. Initial development includes modification in crawling and indexing process of open source index server that constitutes collaborative search service nodes of network. Fig. 4 illustrates the process of text normalization in CPDNet.

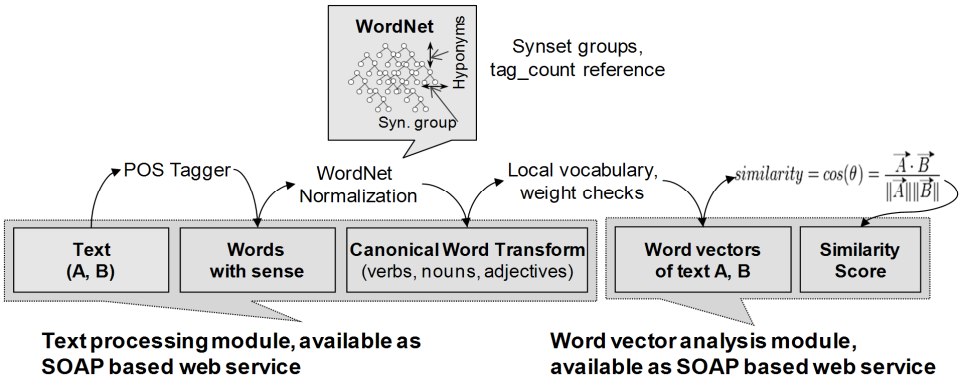

Fig. 4. Process of normalization of text.

The crawled contents are passed through a POS tagger, this process provides the exact sense of a word which then is normalized to a basic form with the help of WordNet SynSet thesaurus. The content in this basic language form is then processed for indexing. The query for such an index is again treated for normalization to develop a semantic match. The process of normalization of text adds conceptual plagiarism check capability in system, such conceptual check is available for documents that are indexed in local CPDNet nodes. Further details and example of process is available in indexing section of project portal[5].

4.2 Introduction of Translation and Normalized Signature Search Service

In order to check plagiarism across language barrier, another service at a different abstraction layer is required. This must translate the normalized indexes and queries into standard English. Each node can provide translation into and from a specific language depending on the local resources. This service will compliment the normalization on a

[5] http://www.cpdnet.org/indexer.

more global and conceptual level. Such abstraction may produce undesired and false results at some levels. However it is worth experimenting with the cross language similarity checks, because of the large availability of intellectual contents in non-English languages.

4.3 Noise Reduction with Domain Specific Searches and Citation Checks

Similarity detection on an abstract level may introduce unnecessary noise in generated matches. It would be helpful to restrict the semantic level search and analysis to a domain specific index. Subject specific searching capability will be introduced by means of setting up specialized indexes of certain domains. The participating institute's local index can be categorized based on various departments or research groups. Introduction of a service that determines the domain of the document being analyzed can help determine the right indexes to search in. One such example is the use of Yahoo Term Extraction service [13].

Another level of service is required to decrease false positives while detecting plagiarism. Some plagiarism detection services give their users the option of ignoring texts found within quotation. This approach however is not sufficient in determining proper citations. There is a need to automatically compare the referenced articles and remove any plagiarism score coming from these sources. Such automation can be achieved by scanning through the referenced articles and creating an index of referenced resources in a document. The user can then choose to ignore the similarity matches which are available in the reference index. The automated detection of proper attribution or citation in a document will save the examiner both time and effort. Developing such a reference index may require a persistent and universally accessible resource identifier associated with the citation. The increasing web publication trend and the emergence of common linking and resource identification standards like DOI [14] are encouraging factors which will lead to further investigations in this area.

5. CONCLUSIONS

Collaborative web service oriented architecture substantially extends current plagiarism detection systems. With flexible and extendable services, rich web user interface, standardized XML based inter application communication and collaborative authoring, it brings us a step closer towards Web 2.0 applications. A survey [2] pointed out that existing plagiarism detection systems fail in the following areas:

1. When systematic attempts are made to combat plagiarism tools by *e.g.* applying extensive paraphrasing through synonymizing tools, syntactic variations or different expressions for same contents.
2. When plagiarism is based on documents that are not available electronically or archive is not available to detection tool.
3. When plagiarism crosses language boundaries.

Based on experimental results, it can be safely stated that the platform presented addresses the second issue effectively. This is due to the additional support of internet

searching API mashup and the collaborative indexing approach. Moreover, the availability of various analysis services, such as vector space similarity, structural evaluation of suspicious documents and fingerprint normalization in the system is an attempt to handle issues 1 and 3. The technology industry has a rapidly growing interest in web services. Many companies and service providers already have web service components available with their applications. Almost every software and internet organization focuses on web services as a core element in future strategies. This tendency suggests that the proposed web services enabled platform is best suited to carry out multiphase plagiarism detection. It will offer the flexibility to incorporate any new processing, discovery or indexing components that may become available to its users. The user-centered collaborative nature of this system makes it an ideal choice to build specialized indexes which are capable of handling semantic considerations in the similarity detection process.

REFERENCES

1. Plagiarism.org: Statistics, http://www.plagiarism.org/plagiarism_stats.html.
2. H. Maurer, F. Kappe, and B. Zaka, "Plagiarism – a survey," *Journal of Universal Computer Science*, Vol. 12, 2006, pp. 1050-1084.
3. H. Maurer and B. Zaka, "Plagiarism – a problem and how to fight it," in *Proceedings of World Conference on Educational Multimedia*, *Hypermedia and Telecommunication*, 2007, pp. 4451-4458.
4. K. Bennett, P. Layzell, D. Budgen, P. Brereton, L. Macaulay, and M. Munro, "Service-based software: The future for flexible software," in *Proceedings of IEEE 7th Asia-Pacific Software Engineering Conference*, 2000, pp. 214-221.
5. W3C: Web Services Architecture, W3C Working Group Note 11 February 2004, http://www.w3.org/TR/ws-arch/.
6. K. Hogg, P. Chilcott, M. Nolan, and B. Srinivasan, "An evaluation of Web services in the design of a B2B application," in *Proceedings of ACM 27th Australasian Conference on Computer Science*, Vol. 56, 2004, pp. 331-340.
7. Search 2.0 vs. Traditional Search, http://www.readwriteweb.com/archives/search_20_vs_tr.php.
8. Crossref Search, http://www.crossref.org/crossrefsearch.html.
9. Vector Space Model, in Wikipedia, http://wn.wikipedia.org/w/index.php?title=Vector_space_model&oldid=113611338.
10. J. M. Farringdon with contributions by A. Q. Morton, M. G. Farringdon, and M. D. Baker, *Analyzing for Authorship: A Guide to the Cusum Technique*, University of Wales Press, Cardiff, 1996.
11. M .K. Bergman, "The deep web: Surfacing hidden value," *The Journal of Electronic Publishing*, http://www.press.umich.edu/jep/07-01/bergman.html.
12. N. Kulathuramaiyer and W. T. Balke, "Restricting the view and connecting the dots – Dangers of a web search engine monopoly," *Journal of Universal Computer Science*, Vol. 12, 2006, pp. 1731-1740.
13. Yahoo Content Analysis Web Services: Term Extraction, http://developer.yahoo.com/search/content/V1/termExtraction.html.
14. A. W. Scott, "DOIs and deeplinked e-reserves: Innovative links for the future," *Technical Services Quarterly*, Vol. 22, 2005, DOI: 10.1300/J124v22 n04_01.

Bilal Zaka holds an M.Sc. degree in Electronics from Quaid-i-Azam University, Islamabad, Pakistan. He is Manager Systems at COMSATS Institute of Information Technology, Pakistan. He received an HEC scholarship to be an overseas researcher in Austria. At present, he is research fellow and doctoral candidate in Institute for Information Systems and Computer Media, Graz University of Technology, Austria. His research interests include collaborative web applications, knowledge management, cloud computing, multimodal applications for the internet and computer networks.