

COMPARATIVE ANALYSIS OF WATERMARKING TECHNIQUES (Review Paper)

Ahmed Khan¹

¹Department of Information Security – Military College of Signals, NUST, RWP.

Corresponding Author: ahmedkhaniub@gmail.com

ABSTRACT- Copyright assurance now a the very first moment's of primary security issues, which obscure fate of Text watermarking advances, is point by point assessment of existing checking plans. Numerous watermarking strategies have been proposed in writing to address these reasons. This paper displays an overview of some of late research in watermarking strategies for content archives. review methodologies are grouped into three classes: picture based methodology, syntactic methodology and semantic methodology. This paper talks about primary commitments, points of interest and disadvantages of diverse strategies utilized for content watermarking as a part of past.

INTRODUCTION

Indeed content is broadly utilized channel of correspondence current over Internet. Along these lines significant segments of sites, ebooks, daily papers, articles, authoritative archives are straightforward plain content. So plain content requires most noteworthy assurance and security from copyright insurance.

As of late, various advanced watermarking calculations have been proposed for pictures, sounds, and recordings; however computerized watermarking calculations for plain content are lacking and ineffective. process in which embeddings and uprooting an advanced watermark to and from a computerized content record which solely distinguishes unique copyright proprietor of that content is called Digital Text Watermarking.

Content watermarking takes by same guidelines as picture, sound, or video watermarking. Truth be told numerous strategies or systems use to make information secure. However as of late different sight and sound substance are used as host to install payload [1], content reports keep on being a basic decision of host because of its all inclusive reality in computerized area. Rightly or wrongly, numerous information concealing routines [2-3] are utilized for record following or copyright security. Others are utilized for confirmation and as of late, for steganography [1].

Freely talking, content based information concealing system in advanced space can be separated into Seagram's and open codes (Kessler, 2004). Seagram implants payload by method for changing appearance of host-content, for example, including additional spaces or tabs while utilizing some pre-characterized information representation(e.g., one space character \rightarrow '0' and two space characters \rightarrow '1') [1].

The open space system proposed by [12,14] is one of soonest examples of content Seagram's. Here, mystery message is encoded by controlling white spaces in a record in which case between sentence, between word and end-of-line spacings are considered [1].

On other hand, Zhou and Shi [3] partition a column of writings in two arrangements of word (say ΦA and ΦB) so that entirety of between word spaces (i.e., number of pixels between two continuous words) [1] in both sets is roughly same. Between word spaces in every set are either contracted or extended such that sure conditions are fulfilled to encode data [1]. That is, $|\Phi A| > |\Phi B| + \epsilon$, is forced to insert '0', and the other way around, where ΦB is absolute number of between word spaces in set ΦB and $\epsilon_r > 0$. Further [7] enhance this work by utilizing multi step adjustment procedure to

accomplish higher bearer limit. These distinctions offer meet people's high expectations in Text watermarking also [1].

With the goal that content watermarking procedures ought to embed selective and imperceptible watermarks in content archives which remain complete after changed altering assaults of insertion, erasure, and re requesting.

Computerized watermarking answers for content make it simple to send and get content over Internet, extranet, and copy. Reports can be assessed for content security and copyright assurance. Identification of any altering made should likewise be possible making so as to utilize advanced content watermarking systems it carefully designed.

MATERIALS AND METHODS

Truth be told that Text watermarking can be utilized for substantial number of utilizations in certifiable. With expanding and across the board utilization of Internet all over world for data sharing, content watermarking has increased more significance.

Developing ideas of computerized libraries, e-business, e-learning, and e-government, ebooks, has made content watermarking a need [2]. Authoritative reports, endorsements, sites, strategies for success, books, articles, verse, organization archives, private substance, SMS, and messages, can be ensured by content checking calculations.

A. AUTHENTICATION

Truth be told that Text watermarking can be utilized for substantial number of utilizations in certifiable. With expanding and across the board utilization of Internet all over world for data sharing, content watermarking has increased more significance.

Developing ideas of computerized libraries, e-business, e-learning, and e-government, ebooks, has made content watermarking a need [2]. Authoritative reports, endorsements, sites, strategies for success, books, articles, verse, organization archives, private substance, SMS, and messages, can be ensured by content checking calculations.

B. COPYRIGHT PROTECTION AND PREVENTION

For copyright assurance, it is extremely important to secure copyrights of web substance, ebooks, research papers, diary articles and different archives containing plain content. content proprietor can implant a watermark speaking to copyright data of his information. This watermark can later be removed to demonstrate proprietorship if any contention of copyright case emerges in future [2].

It is possibly most imperative utilization of computerized content water stamping. As per abuse of law or illicit

duplicating and conveyance of content can likewise kept away from by content watermarking calculations.

Watermarked data can straightforwardly control advanced duplicating gadget which can be a printer or basic a duplicate glue order [2]. Implanted key can speak to a duplicate consent bit stream that is recognized by recording programming which then choose if replicating methodology ought to go (on the off chance that it is permitted) or not (in the event that it is disallowed by substance proprietor) [2].

C. COVERT COMMUNICATION

Primary thought of content watermarking is to be securing correspondence or sending or getting information. Incognito correspondence along these lines implies embedding a mystery message into a guiltless looking content in a way that would keep any informal individual to recognize it and proposed beneficiary would have the capacity to get it [4].

D. FINGERPRINTING

Ordinarily creator or proprietor have mystery key duplicate and gave numerous clients for information securing. For proprietor, implanting a special serial number-like watermark is a decent approach to identify clients who break their permit assention by duplicating secured information and supplying it to an outsider [2]. Numerous organizations use unique mark watermarks to recognize copyright violators.

E. DIFFERENT TYPES OF ATTACKS ON TEXT

In PC world is very little amped up for content watermarking advances. fundamental reason is to un-revealed watermarking strategies and absence of strength towards assaults. It is workable for an aggressor to perform fractional assaults regardless of the possibility that they are not ready to do it totally. So it is important to examine every kind of assault. Watermark assaults are depicted as takes after: truth be told, normally use in content experiences propagation, equivalent word substitution, and reformatting, rewording, and syntactic changes assaults. Every one of these assaults can be set in taking after classes:

1) Unauthorized Insertion

Under this type of assault, words and sentences are added to content to make it look changed and at times to keep another message\watermark of any assailant. An aggressor infrequently embeds some content to unique content to add some extra data. This sort of assault happens when an assailant is intrigued to include some false data for instance if there should be an occurrence of authoritative reports and cases [2]. Such assaults can be maintained a strategic distance from by consolidating an affirming power in watermarking architectures which timestamps substance in name of creator with ebb and flow date and time. At whatever point, a disagreement about copyright claim emerges, this timestamp is utilized to recognize creator who enrolled content first.

2) Unauthorized Detection and Deletion

In a few applications, capacity to identify ought to be limited. Its possible that capacity of a foe only distinguishes regardless of whether imprint is available in a given Work will debilitate security of a watermarking framework [3].

As indicated by this assault implies irregular cancellation of words and sentences from unique content. assailant erases some data to diminish peruser and conceal character of

unique proprietor of content. Its important to keep an aggressor from recouping unique, however its more imperative to keep expulsion of watermark from content.

4) Re-ordering

The aggressor rearranges and reorders words and sentences of content to make it look changed and to decimate watermark [2]. If there should be an occurrence of content, aggressor rethinks and replaces certain words with their equivalent words. goal by and large is to obliterate written work style, essence, and in some cases importance of content [3].

F. NATURE, LOW AND HIGH VOLUME

Truth be told when aggressor is intrigued to add or erase some data to and from content, then volume of assault will be low. In any case, if aggressor is keen on utilizing some piece of content as a part of own content, then volume of assault will be high. The nature thought is joined insertion, erasure, and re-requesting assault is termed as altering assault [3].

Altering can be made at any arbitrary area in content report. Altering can be made in two courses including: In which Localized altering implies, insertion or erasure; of words or sentences at a solitary area in content. Area can be in starting, at end or anyplace in content, contingent upon aggressor's aim of utilization.

In which Dispersed insertion and erasure of sentences and words can be made at numerous areas in unique content. Assailants attempting to make content look changed makes scattered altering in content. This sort of assault for the most part happens in exploration counterfeiting and scholarly written work [3].

G. TEXT WATERMARKING ISSUE`S

Actually Plain content, being most straightforward method of data, conveys numerous difficulties with regards to copyright security. Content has restricted limit for watermark installing following there is no absence of movement in content as can be found in pictures, sound, and recordings. Twofold nature with clear partition in the middle of front and foundation, piece/line/word designing, semantics, structure, style, and dialect tenets are some of famous properties of content which are should have been tended to in any content watermarking calculation [4].

Plus, characteristic properties of a non specific watermarking plan like intangibility, Robustness, and security additionally should be fulfilled. Any change on content ought to save meaning, office, grammaticality, and estimation of content. importance of content is its worth, and it ought to be rationed through watermarking all together not to irritate correspondence [3]. Certainty is required to speak to significance of content in a reasonable and smooth way, all the more vitally in legendary compositions. Procedure of installing ought to see with syntax tenets of dialect, so as to protect decipherability of content. Protecting style of writer is essential in a few spaces, for example, writing composing or news channels [4]. Touchy nature of a few reports, for example, authoritative archives, verse, and cites don't permit us to make semantic changes haphazardly in light of the fact that in these types of content basic change in some cases wrecks both semantic essence and estimation of content [3].

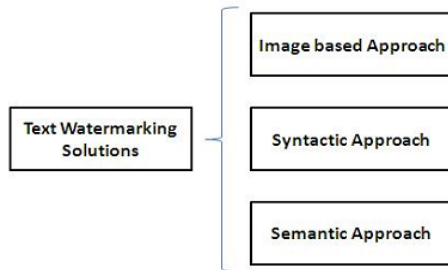


Fig 1 Text Watermarking Solutions [8]

H. RELATED WORK ON TEXT WATERMARKING

The depiction of work done in every classification of content watermarking is as per the following: actually Text watermarking appeared in Fig 1 is region of exploration that has risen after improvement of Internet and correspondence advancements. initially reported push to secure copyright of content was made in 1994 by Brassil et al. [5-6], when IEEE Journal on Selected Areas in Communications issue was planned to be distributed, for Secure Electronic Publishing Trial. There were more than 1,200 enlisted clients in first month, and every duplicate of every paper has been enrolled and watermarked with beneficiary [7] and it is right now an exceptionally dynamic examination territory with various specialists chipping away at content watermarking for English dialect and also Persian, Turkish, Korean, Urdu, and Arabic.

He past work on advanced content watermarking can be grouped in taking after classes [7]; a picture based methodology, a syntactic methodology, and a semantic methodology. Portrayal of every class and work done likewise are following.

A. IMAGE-BASED APPROACH

In this methodology towards computerized content watermarking, content report picture is utilized to implant watermark [8]. Content is hard to watermark in view of its straightforwardness, affectability, and low limit for watermark installing.

At first endeavors in content watermarking attempted to regard content as picture. Watermark was inserted in format and appearance of content picture. Brassil, et al. proposed a couple of routines to watermark content archive by utilizing content picture [9-11]. In the first place technique proposed by Brassil was line-shift moving so as to code calculation which modifies report picture lines upward or descending (left or right) contingent upon paired sign (watermark) to be embedded as appeared in fig 2. Recognition calculation is non-blind in which unique archive ought to be accessible. Second technique was word-shift coding calculation which moves words inside of content on a level plane consequently growing spaces to insert watermark. calculation can work both in non-visually impaired and dazzle modes [8].

Third strategy is highlight coding calculation which somewhat changes components, for example, pixel of characters, length of end lines in characters to encode watermark bits in content. All these proposed strategies debilitate un-approved conveyance by inserting every report with a novel code word [3]. Among three introduced routines, line-shift coding is most powerful arrangement

under assorted assaults yet this can likewise be effortlessly vanquished. Maxemchuk, et al. [13-15] dissected execution of aforementioned strategies. connection and centroid-based techniques [16] are additionally proposed which regards profiles as a discrete time flag and search for bearing of

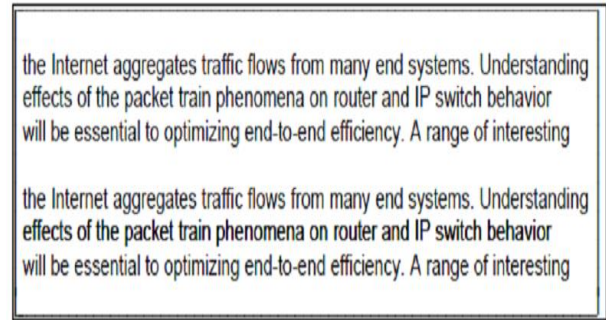


Fig 2. Line shift coding [12]

furthermore, Yan [17] proposed a calculation in light of a normal between word separation in every line. separations are balanced by wave of a particular stage and recurrence. highlight and pixel level calculations were additionally created which check reports by adjusting stroke elements, for example, width or serif [18]. Calculation which uses dim scale picture of content was additionally created [19]. Another calculation which watermarks content record picture utilizing edge course histogram was additionally proposed [20]. Youthful Won Kim et al. proposed a content watermarking calculation in light of word arrangement and between word space insights [17]. In this methodology, all words in a content archive are characterized relying upon some content components and afterward contiguous words include a portion and that section is ordered relying upon class names of words inside of fragment. data is encoded by adjusting a few measurements of between word spaces of sections fitting in with same class.

A few favorable circumstances over traditional word-shift calculations are talked about. Adnan M. AL attar et al. proposed a calculation [15] to watermark electronic content archives containing legitimized passages and sporadic line dispersing. calculations which misuse printed content record to recognize source printer were likewise created [8]. These systems use print quality deformities as an inborn mark of a printer demonstrates banding components of a content report. These elements can recognize particular make and model of gadget which made it.

Cox et al., [14] portrayed various utilizations of advanced watermarking and their normal properties like heartiness alter resistance, constancy, computational expense, and false positive rate. Yang and Kot [10] proposed a technique for watermarking on content report pictures to confirm proprietor or approved client is proposed. proposed technique makes utilization of coordinated bury character and word spaces for watermark installing. Qadir and Ahmad [23] proposed an original thought based upon a clever encoding plan in universe of content watermarking which has no impact on modification of sentence structure of record and additionally format. In this way giving a design/position autonomous system in which data inside of content is controlled to shroud certain data. Abdullah and Wahab [22]

displayed a content watermarking plan focusing on an item based environment. heart of proposed arrangement depicts idea of watermarking an article based content record where every single content string is entertained as a different item having its own arrangement of properties. Exploiting z-requesting of articles, watermark is connected with z-pivot letting zero constancy aggravations to message. Villan et al. [8] examined hypothetical useful parts of content information covering up in printed archives.

Mikkilineni *et al.* [3,7-8] attempted to upgrade information concealing and watermark implanting limit of printed paper archives. Micic et al. [20] Proposed calculation for validation of content report utilizing advanced watermarking. Content report pictures were contrasted with assess changes. Xing Ming Sun with his group proposed a part based advanced watermarking calculation for Chinese writings [16]. Li and Dong proposed a calculation for Chinese content watermarking in view of Chinese characters structure. Another content watermarking calculation utilizing Eigen qualities is likewise proposed [15]. Zhou *et al.* [24] Presented a zero-watermarking calculation for substance verification of Chinese content records.

B. SYNTACTIC APPROACH

Content is comprised of characters, words, and sentences. Sentences have distinctive syntactic structures. work likewise depicts in Mikhail. J. Atallah, et al. Initially proposed characteristic dialect watermarking plan utilizing syntactic structure of content [8,17-18] where the syntactic tree is assembled and changes are connected to it to insert watermark saving every single innate property of content. They created procedures for installing a strong watermark in content by various data confirmation and security systems with cutting edge and assets of characteristic dialect handling. For watermark installing, they utilized handlings of TMR (Text meaning representation, for example, appending, trimming, and swap. These routines are safe towards numerous assaults yet change content to an expensive degree. Subsequently can't be connected to content of a delicate nature like verse, authoritative records, transcripts, and contracts. Common Language Processing (NLP) systems are utilized to break down syntactic and semantic structure of content while performing any changes to implant watermark bits.

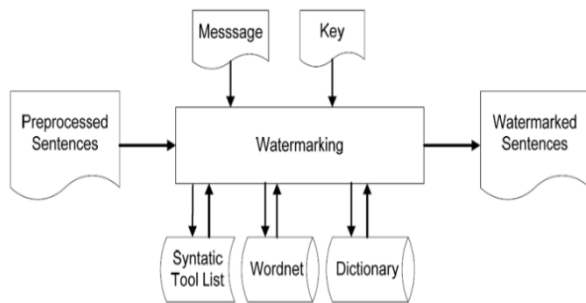


Fig 3. Syntactic sentence level watermarking [8]

Kankanhalli and Hau [16] proposed a technique to watermark electronic content reports in utilizing ASCII characters and accentuation as a part of content. Hassan et al. proposed normal dialect performing so as to watermarking calculation

Morpho-syntactic changes to message [8]. content is initially changed into a syntactic tree outline where progressions and utilitarian conditions are made express and watermark is installed. watermarking procedure is appeared in fig 3. creator expressed that agglutinative dialects like Turkish is simpler to watermark than English dialect. watermarking answers for agglomerative dialects like Turkish, Korean, Arabic, and Urdu are productive since these dialects give space to watermark inserting. Notwithstanding, syntactic answers for English dialect are very little sufficient. Hassan et al., additionally proposed 21 syntactic instruments for content watermarking [3] and MI Young Kim [7] as of late proposed a calculation for content watermarking utilizing syntactic investigation of plain content. Kim [10] likewise proposed a characteristic dialect watermarking calculation for Korean dialect utilizing word intensifying removal. Helga Hoehn proposed a characteristic dialect watermarking calculation which utilizes rather semantic and linguistic changes of unique content substance as opposed to altering content [7]. Murphy and Vogel [3] present three normal dialect stamping calculations utilizing shallow parsing strategies, lexical substitutions, and swapping. They additionally investigated noteworthiness of computerized and reversible syntactic changes to conceal information in plain text[8].

C. SEMANTIC APPROACH

In this watermarking plans concentrate on utilizing semantic structure of content to install watermark. Content substance, verbs, things, words and their spellings, acronyms, sentence structure, punctuation rules, and so on have been misused to embed watermark in content, however none of these turned out to be strong and debase nature of content to a substantial degree. Atallah *et al.* were first to propose semantic watermarking plans in 2000 [17,18]. Later, the equivalent word substitution technique was proposed in which watermark is inserted by supplanting certain words with their equivalent words [19]. Xing Ming, et al. Proposed thing verb based method for content watermarking [7] which misuses things and verbs in a sentence parsed with a linguistic use parser utilizing semantic systems. Mercan et al., proposed a sentence based content watermarking calculation [8] which depends on different components of every sentences and misuses thought of orthogonally between elements. Later Mercan, et al., proposed a calculation of content using so as to watermarking mannerisms to install watermark [20]. calculations make astute utilization of writing blunders, acronyms, and contractions that are regular in superficial content like messages, online journals, talk, SMS and so on. Calculations were produced to watermark content observing so as to utilize phonetic semantic wonders of presuppositions [21,22] talk implications and representations. Presupposition is verifiable data considered also known. Presuppositions are recognized and after that changes like passivation, topicalization, additional circumstance, and offering are connected to implant watermark in content. content pruning and joining calculations were additionally created in past. calculation taking into account content importance representation (TMR) strings has as of late been proposed [9]. Shirali-Shahreza et al. [8] proposed another strategy for

mystery trade of data through SMS by utilizing shortened form content steganography with utilization of designed dialect of SMS-messaging. They likewise proposed a strategy for steganography in English writings. In this system US and UK spellings of words substituted with a specific end goal to shroud information in an English content. In English a few words have diverse spelling in US and UK [8]. Later, Rafat [2] proposed an upgraded technique for SMS steganography utilizing SMS-messaging dialect, by evacuating static nature of word-contraction list and presenting computationally light weighted XoR encryption.

There are two fundamental sorts of information securing from programmers and aggressors. Initial one is Public and other one is Private so that numerous work done on these systems. depiction of work done in every class of content watermarking is as per the following: Furthermost of existing watermarking methods in writing are private, implying that they depend on some private parameters(e.g. a mystery key). Just approved individuals (e.g. Content proprietors) who know these private parameters can check watermark and demonstrate their responsibility for in the event of any illicit redistribution, false possession claim, and burglary and so forth. Then again, private watermarking methods experience the ill effects of uncover of private parameters to degenerate individuals once watermark is checked in vicinity of open.

With access s to private standard ammeters, assailants can without much of a stretch discredit watermark discovery either by adding so as to expel watermarks from ensured information or a false watermark to non-watermarked information. Interestingly, if there should arise an occurrence of open watermarking methods any end client can confirm installed watermark the same number of times as vital without having any earlier learning about any of private parameters to guarantee that they are utilizing right (not altered) information originating from unique source. For example, when a client utilizes touchy data, for example, coin trade rates or stock costs, it is imperative for him to guarantee that information are right and originating from unique source.

DISCUSSIONS

Each recorded has numerous restrictions so that a few constraints are depict content watermarking calculations utilizing paired content pictures are not vigorous against retyping and content propagation assaults. With expanding effective utilization of OCR (Optical Character Recognition) now days, these routines are absolutely a disappointment [3]. OCR use can devastate changes made by moving words upward and descending, to record edges, to textual styles, serif, and elements of content. Additionally, watermarking can without much of a stretch annihilated by a straightforward duplicate glue to scratch pad assault [8]. In semantic content watermarking strategies definitively enhance data modifying so as to conceal limit of English content granularity of significance of individual sentence yet semantic content watermarking plans are exceptionally theoretical and unreasonable [4]. equivalent word based methods are not strong to irregular equivalent word substitution assaults. There may be situations where wrong words get chose for equivalent word substitution. Besides,

equivalent word based techniques require an expansive synonymy lexicon and an immense collocation database [3]. At last, message watermarking in view of semantics, is dialect subordinate where dialect is not something static. With entry of time, dialect shifts and subsequently security and copyright arrangement gave by advanced watermarking in light of semantic will have constrained quality. semantic strategies for advanced watermarking use characteristic dialect preparing calculations to investigate content significance and to perform change [8]. NLP is a youthful region of exploration; we can say that content watermarking utilizing semantics does not give a functional and complete content watermarking arrangement questions still remains an open issue for future examination.

CONCLUSION

In this paper we review momentum condition of specialty of distinctive watermarking strategies for Text information. In this field, work done on content watermarking is exceptionally constrained and particular. Content watermarking calculations utilizing twofold content picture are not strong against multiplication assaults and have restricted appropriateness. We portray a wide range of use and Attack and primary arrange all methods in light of (i) Image based (ii) syntactic (iii) Semantic. At long last, we watch that past strategies are computationally costly and non-hearty. Content confronting colossal insertion, erasure and reordering assaults should be secured, and proficient picture and content watermarking calculations are required.

REFERENCES

- [1] KokSheik Wong, Kok Onn Chee Lip Yee Por, "UniSpaCh: A text-based data hiding method using Unicode space characters," Journal of Systems and Software, pp. 1-8, December 2011.
- [2] Zunera Jalil, "Copyright Protection of Plain Text," FAST National University of Computer and Emerging Sciences, Islamabad, A thesis submitted in partial fulfillment of requirements for degree of Doctor of 2010.
- [3] Zhou, W. Zhao, Z. Wang, and L. Pan, "Security theory and attack analysis for text watermarking", in Proceedings of International Conference on E-Business and Information System Security,(EBISS 2009), May 2009, pp. 1-6.
- [4] M. Topkara, G. Riccardi, D. Hakkani-Tur, and M. J. Atallah, "Natural language watermarking: Challenges in building a practical system", Proceedings of SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents, San Jose, CA, Jan. 2006.
- [5] J. T. Brassil, S. Low, N. F. Maxemchuk, and L. O. Gorman, Electronic Marking and Identification Techniques to Discourage Document Copying, IEEE Journal on Selected Areas in Communications, vol. 13, no. 8, pp. 1495-1504, 1995.
- [6] J. T. Brassil, S. Low, N. F. Maxemchuk, L. O. Gorman, "Hiding information in document images", in Proceedings of 29th Annual Conference on Information

- Sciences and Systems, Johns Hopkins University, 1995, pp 482-489.
- [7] N. F. Maxemchuk and S. Low, "Marking text documents," in Proceedings of IEEE International Conference on Image Processing, Washington, DC, Oct. 1997, pp. 13-16.
- [8] MAnwar Mirza ,Zunera Jalil, "A Review of Digital Watermarking Techniques for Text Documents," in International Conference on Information and Multimedia Technology, Islamabad,2009, pp. 1-5.
- [9] Khan and Anwar M. Mirza, Genetic Perceptual Shaping: Utilizing Cover Image and Conceivable Attack Information Using Genetic Programming, Information Fusion, vol. 8, no. 4, pp. 354-365, 2007.
- [10] Khan, Intelligent Perceptual Shaping of a digital Watermark, PhD Thesis, Faculty of Computer Science and Engineering, GIK Institute, Pakistan, 2006.
- [11] J. T. Brassil, S. Low, N. F. Maxemchuk, and L. O. Gorman, Electronic Marking and Identification Techniques to Discourage Document Copying, IEEE Journal on Selected Areas in Communications, vol. 13, no. 8, pp. 1495-1504, 1995.
- [12] J. T. Brassil, S. Low, and N. F. Maxemchuk, "Copyright protection for electronic distribution of text documents", in Proceedings of IEEE, vol. 87, no. 7, 1999, pp.1181-1196.
- [13] N. F. Maxemchuk and S. Low, "Marking text documents," in Proceedings of IEEE International Conference on Image Processing, Washington, DC, Oct. 1997, pp. 13-16.
- [14] N. F. Maxemchuk, S. H. Low, Performance Comparison of Two Text Marking Methods, IEEE Journal of Selected Areas in Communications (JSAC), vol. 16, no. 4, pp. 561-572, 1998.
- [15] N. F. Maxemchuk, Electronic Document Distribution, AT&T Technical Journal, vol. 6, Sept. 1994, pp. 73-80.
- [16] S. H. Low, N. F. Maxemchuk, and A. M. Lapone, "Document identification for copyright protection using centroid detection," IEEE Transactions on Communications, vol. 46, no.3, March 1998, pp 372-381.
- [17] S. H. Low and N. F. Maxemchuk, "Capacity of text marking channel," IEEE Signal Processing Letters, vol. 7, no. 12, pp. 345 -347, Dec. 2000.
- [18] T. Amano and D. Misaki, "A feature calibration method for watermarking of document images," Proceedings of Fifth International Conference on Document Analysis and Recognition (ICDAR), 1999, pp.91-94.
- [19] Bhattacharjya and H. Ancin, "Data embedding in text for a copier system",Proceedings of 1999 International Conference on Image Processing (ICIP), vol. 2, 1999, pp. 245-249.
- [20] Y. Kim, K. Moon, and S. Oh, "Watermarking text document images using edge direction histograms", Pattern Recognition Letters, Elsevier. vol. 25, pp. 1243-1251, 2004.
- [21] Y. Kim, K. Moon, and I. Oh, "A text watermarking algorithm based on word classification and inter-word space statistics", in Proceedings of Seventh International Conference on Document Analysis and Recognition (ICDAR), vol. 2, Aug. 2003, pp. 775-779.
- [22] M. Abdullah and F. Wahab, "Key based text watermarking of e-text documents in an object based environment using z-axis for watermark embedding", World Academy of Science, Engineering and Technology, vol. 46, 2008, pp. 199-202.
- [23] M. A. Qadir and I. Ahmad, "Digital text watermarking: secure content delivery and data hiding in digital documents", IEEE Aerospace and Electronic Systems Magazine, vol. 21, no. 11, 2006, pp. 18-21.
- [24] X. Zhou, W. Dong, Z. Wang, and G. Wei, Zero-watermarking algorithm for content authentication of Chinese text documents, Journal of Computers, vol. 20, no. 1, 2009.