Zero-Watermarking Algorithm for Content Authentication of Chinese Text Documents

Xin-Min Zhou[1,2,*], Wei-Dong Zhao[1], Zhi-Cheng Wang[1], and Gang Wei[1]

[1] Research Center of CAD

Tongji University

Siping Road 1239, Shanghai, China

{529_xavman, wd, zhichengwang, weigang}@tongji.edu.cn

[2] Information School

Hunan University of Commerce

Changsha, Hunan, China

zhouxinmin@yahoo.cn

Received 15 February 2009; Revised 1 April 2009; Accepted 3 April 2009

Abstract. Most watermarking schemes for text documents are usually implemented by embedding a mark in a host document itself for intended purposes. The existing techniques for text watermarking are confronted with the problems of perceptible quality degradation and the inherent conflict between imperceptibility and robustness, which introduced by the watermark embedding. In this paper, we propose a novel text zero-watermarking scheme for content authentication. This method can solve the problems mentioned above through embedding the authentication information in the constructed binary pattern of Chinese Word document, not in host document itself. In this scheme, the hashing value of the text document's content calculated by MD5 algorithm and the structure knowledge of Chinese characters are exploited for constructing the binary pattern, and the chaotic encrypting algorithm is used to enhance the security of the watermarking scheme. When any modification of the text document's content is made, the extracted result will be completely different to the original authentication information. The experimental results indicate that the proposed zero-watermarking scheme is efficient and secure.

Keywords: zero-watermarking, text watermarking, content authentication, Chinese text documents

References

[1] P.V.K. Borges and J. Mayer, "Text Luminance Modulation for Hardcopy Watermarking," *Signal Processing*, Vol. 87, pp. 1754-1771, 2007.

[2] H. Frank and K. Martin. "Multimedia Watermarking Techniques," *Proceedings of the IEEE*, Vol.87, pp.1079-1107, 1999.

[3] J. Sang, X. Liao, M.S. Alam, "Neural-network-based Zero-watermark Scheme for Digital Images," *Optical Engineering*, Vol.45, 2006.

[4] N. Chen and J. Zhu, "A Robust Zero-Watermarking Algorithm for Audio," *EURASIP Journal on Advances in Signal Processing*, 2008.

[5] C. Hanqiang, X. Hua, L. Xutao, L. Miao, Y. Sheng, W. Fang, "A Zero-watermarking Algorithm Based on DWT and Chaotic Modulation," in *Proceedings of SPIE*, pp. 1–9, 2006.

[6] L. Jing and F. Liu, "Double Zero-watermarks Scheme Utilizing Scale Invariant Feature Transform and Log-polar Mapping," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 2118–2121, 2007.

* Correspondence Author

[7] Y. Hu, S. Zhu, D. Zhang, "A Novel Zero-watermark Algorithm in Image Subspace Domain," in *Proceedings of 2007 IEEE International Conference on Control and Automation*, pp. 2744-2748, 2007.

[8] Q. Wen, T. Sun, S. Wang, "Concept and Application of zero-Watermark," *Acta Electronica Sinica*, vol. 31, pp. 214-216, 2003.

[9] X.M. Sun, H.W. Chen, L.H. Yang, Y.Y. Tang, "Mathematical Representation of a Chinese Character and Its Applications," *International Journal of Pattern Recognition and Artificial Intelligence*, Vol.16, pp. 735-747, 2002.

[10] Chinese Text Computing. http://lingua.mtsu.edu /Chinese-computing/statistics/, 2009-1-20.

[11] X. Zhou, W. Zhao, Z. Wang, R. Peng, G. Wei, "A Robust Digital Watermarking of Chinese Texts Based on Watermarking Structure and Embedding Strategy," in *Proceedings of CISP'08*, pp. 635–639, 2008.

[12] D. Shank, R. Mark, T. Myers, "Microsoft Office 2000 Visual Basic Programmer's Guide," Microsoft Press, 1999.