Abstract

In the study of content authentication and tamper detection of digital text documents, a novel English text zero-watermarking approach based on probabilistic patterns proposed in this paper. Hidden Markov model was presented, and the core algorithm of the watermark generation and detection was designed in this paper. In the proposed approach, Hidden Markov model letter-level based of order one was constructed for text analysis and watermark generation based on interrelationship between contents of host text document by using the two-dimensional matrix coordinate of probabilistic weights of states and transitions. However, we can extracted this watermark later using extraction and detection algorithm to identify the status of text document such as authentic, tampered, etc. Furthermore, the effectiveness and feasibility of our proposed algorithm was proved with algorithm experiments and comparative with other recently approach under random insertion and deletion attacks in localized and
dispersed form on 6 text datasets of varying length. Results showed that our proposed algorithm is more secure, had better robustness, and always detects tampering attacks occurred randomly on text even when the tampering volume is low or high.

References


**Index Terms**

Computer Science  
Emerging Trends in Technology

**Keywords**

Digital Watermarking  
Hidden Markov Model