A NOVEL CONTENT BASED ZERO WATERMARKING ALGORITHM FOR TAMPER-PROOFING PLAINTEXT DOCUMENT

SAKSHIKHULLAR & BIJENDRA SINGH

NSIT, Delhi, India

ABSTRACT

With the advent of internet, mobiles & other communication media, it has become extremely important to protect digital data against illegal copying, tampering, forgery, illicit redistribution etc. Digital watermarking provides authentication and tamper protection for multimedia contents over the internet. Text is an important medium travelling over the internet, since most digital content are as plain text. In this paper, we propose a novel content based zero-watermarking algorithm for tamper proofing text documents. A watermark gets generated based on the occurrence of vowels in the text. It is zero based since the algorithm generates watermark using the content of the text itself which gets registered with a certifying authority.

The plaintext file is divided into blocks. Within this block, the vowel count is calculated for each block. The algorithm automatically chooses the first two vowels it sees in the block. Keeping one constant, it measures the distance to each occurrence of second vowel's distance & averages it. The vowel sum is divided by this average distance to obtain W0 which is the first position in the watermark. And likewise the process is repeated for the remaining blocks to obtain the final Watermark WM. This watermark is then registered by the original owner to the CA. A trusted certifying authority is an essential requirement in this algorithm with whom, the original copyright owner registers his/her watermark. Whenever the content/text ownership is in question, this trusted third party acts as a decision authority.

Experimental results demonstrate the effectiveness of the algorithm against tampering attacks by identifying watermark distortion rates on 9 different samples. It has provided very satisfactory results against insertion, deletion, re-ordering and substitution attacks.

KEYWORDS: Authentication, Tampering, Text Document Security, Watermarking, Watermark Embedding, Watermark Extraction, Zero Watermarking