International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-3 Number-2 Issue-10 June-2013

Data Authentication Using Cryptography

Sagar Chouksey[1], Rashi Agrawal[2], Dushyant Verma[3],Tarun Metta[4]

Abstract

We present a novel approach using cryptography for data authentication. The key idea is to provide a encoded quantized data projection as authentication data. This can be correctly decoded with the help of an authentic data using as side information. Cryptography source coding provides the desired robustness against legitimate variations while detecting illegitimate modification. Additional adjustments might not change the meaning of the content, but could be misclassified as tampering. Users might also be interested in localizing tampered regions. Distinguishing legitimate encodings with possible adjustments from tampering and localizing tampering are the challenges addressed in this paper. We apply cryptography source coding and statistical methods to solve the data authentication problem. Experimental results have been presented for data authentication.

Keywords

Image Authentication, Encryption, Decryption, Cryptography.

1. Introduction

Today's leading enterprises utilize state of the art of integrated solutions and technologies an attempt to obtain the authenticated data. Multilayer network architectures, scalable web services, custom applications, distributed services and heterogeneous server platform environments, form a small sample of the infrastructure's complexity in modern organizations. These complex architectures in the core network infrastructure, result in large and more difficult than ever security demands in order to keep data and information assets secure [1]. Additionally to this recently added system and authentication complexity,criminal organizations have formulated their hacking procedures in a try to break into networks and harm the organization with every possible way [2].

In these systems P2P file sharing, BitTorrent [3] etc. each user not only receives the requested content but also acts as a relay forwarding the received portions to the other users. The un-trusted intermediaries might tamper with the media for a variety of reasons, such as interfering with the distribution of particular files, piggy backing unauthentic content, or generally discrediting a particular distribution system. A 2005 survey indicates that more than 50% of popular songs in KaZaA are corrupted, e.g., replaced with noisy or different songs. Distinguishing legitimate encoding versions from maliciously tampered ones is important in applications that deliver media content through un-trusted intermediaries. The problem is more challenging if some legitimate adjustments, such as cropping and resizing an data, are allowed in addition to lossy compression. Additional adjustments might not change the meaning of the content, but could be misclassified as tampering. Users might also be interested in localizing tampered regions. Distinguishing legitimate encodings with possible adjustments from tampering and localizing tampering are the challenges addressed in this paper. We apply distributed source coding and statistical methods to solve the data authentication problem such that whether the transmitted data from sender is legitimate or tampered received by the receiver. Data Authentication System presented below:

The data authentication based on Slepian-Wolf coding is shown in figure 1. In which we have shown that we apply Slepian-Wolf coding on legitimate data X with certain code rate say R and also we provide this information to Slepian-Wolf decoder with Y as side information.

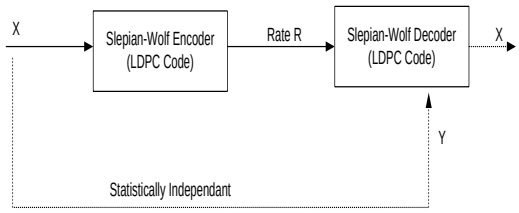

Figure 1: The source X and side information Y are statistically dependent, but Y is available only at the decoder.

Most companies and institutes work diligently to maintain an effective data security policy, implementing the latest products and services to prevent fraud, sabotage, and information leakage. However this proactive up-to-date approach does not

result in a successful security policy. The problem is that they still do not know whether and where they are vulnerable. Unfortunately, the up-to-date security approach is not adequate because it does not detect mis-configured settings An organization that truly wants to adopt a proactive approach, aggressively seeks out all types of vulnerabilities by using relevant methods.

2. Literature Review

This paper first reviews about the introduction of distributed source coding techniques and low density parity check codes. A new method of distributed source coding for binary sources using low density parity check codes is to be developed. This new scheme for distributed source coding of binary sources using low density parity check codes is developed and implemented in the probability domain as opposed to the log domain presented by Liveris et. al. [4]. The performance of these schemes is analyzed by comparing the bit error rate and symbol error rate for different correlations between two randomly generated binary and non-binary sources respectively. Gallager's [5] low density parity check codes are defined by sparse parity check matrices, usually with a random construction. Such codes have near Shannon limit performance when decoded using an iterative probabilistic decoding algorithm. Low density parity check codes are also shown to be useful for communicating over channels which make insertions and deletions as well as additive (substitution) errors. LDPC codes can also be extended over non binary sources for channel coding by Davey et. al. [6]. These codes were shown to have a 0.6 db improvement in signal to noise ratio for a given bit error rate. The use of cyptography codes for distributed source coding was suggested by Liveris et. al. [4]. They developed distributed source coding scheme for binary sources in the log domain. This paper deals with the development of distributed source coding for non-binary sources using cyptography codes.

The first scheme is the implementation of distributed source coding scheme proposed in the paper [4]. The implementation in this dissertation is carried out using Slepian-Wolf coding in the probability domain described in [4]. This was necessitated as the development of distributed source coding for binary sources over a log domain could not be directly extended for non-binary sources. Two binary sources with different correlations are considered. One of the sources is assumed to be available at the decoder and is acting as the side information. The other source is compressed and sent to the decoder. The decoder decodes the compressed source by using the side information available to it from the other source. The performance of this scheme is evaluated for different correlations between the two binary sources. The bit error rates are computed and plotted.
One of the aims of this LDPC codes is transmission of video frames using the above mentioned scheme. Video frames are highly correlated which lend themselves well for transmission using distributed source coding. The above mentioned scheme is for binary domain. This necessitated the conversion of the video frames into its binary equivalent [7]. On conversion of the video frames into binary sequences, it was observed that the correlation between the video frames decreased drastically thereby rendering the above scheme unsuitable for video frame transmission. This led to the investigation of modifying non binary LDPC for distributed source coding which would allow the transmission of video frames without conversion to binary sequences thereby preserving the correlation. Thus the second scheme developed in this paper provides an ideal way of implementing distributed source coding for non-binary sources like video sources. The scheme devised for distributed source coding in the non-binary domain is implemented for different Galois fields.The performance of this scheme is evaluated for different correlations between the two non-binary sources. The above scheme is implemented for Galois field 2 (GF(2)) and Galois field 4 (GF(4)) binary and non-binary fields. Higher order Galois fields and video frame transmissions were not simulated due to memory constraints. The performance of the above scheme is quantified by using sources with different correlations as well as different compression rates at the encoder [6].

3. Proposed Method for Data Authentication

Algorithm for data authentication problem contains following steps:
1: let us consider input data x.
2: Apply Transformation on data x that yield an gray scale data $x_T$of size 336 × 336 .
3: Divide the data $x_T$into 16×16 Block and then A Projection is applied on each block

which gives a projected data X which is quantized to yield X_q.
4: on the projected data $X_q$performed encoding that produces bit-stream $S(X_q)$.
5:on the X_q(Result of step 3) conversion is performed that convert X_q into an data digest $X_{D.}$
6:on this data digest X_D performs Private key cryptography Algorithm that produces an Encrypted data X_E. The original data x is then sent to the receiver along with encoded data $S(X_q)$ and encrypted data X_E as for authentication of data.

1:At the receiver, the user seeks to authenticate the data y with authentication data $S(X_q)$ and X_E.
Step2:Apply transformation on data y that yield an gray scale data y_T.
3: Divide the data $y_T$into 16×16 Block and then projection is applied on each block which give a projected data Y which is quantized to yield Y_q.
4:Perform decoding on the LDPC bit-stream $S(X_q)$ using Y as a side information that constructs X_q'
5: On the X_q' conversion is performed that convert X_q' into a data X_D'.
6:Performs same Private Key cryptography Algorithm on the Encrypted data $X_E$thatproducesan decrypted data X_D.
7: Compare the results from step 5 and step6. If these two data digests do not match, the receiver recognizes that data is tampered.

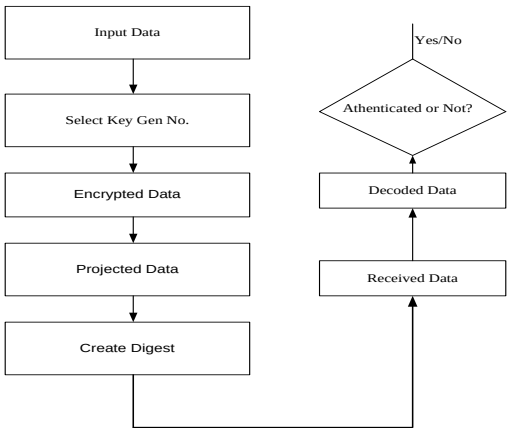

Figure 2: Flow Chart for data authentication algorithm

4. Experimental Results

It is not uncommon that a data authentication based on proposed scheme has undergone some additional adjustments, some of these we might want to accept as legitimate image adjustments. Some results based on proposed method discussed above are presented here:

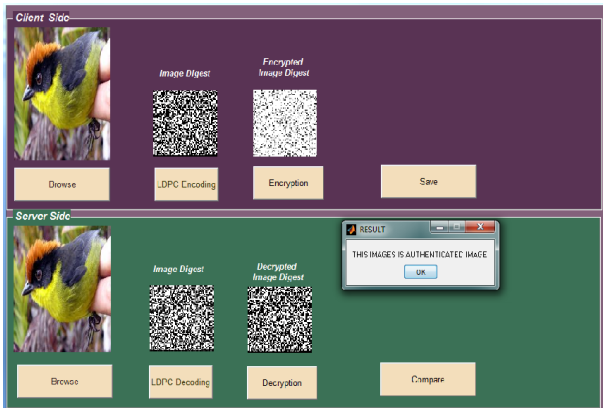

Figure 3: shows that after received the data via

The transmitter side receiver starts authentication process of the received data. For this process receiver apply decoding, and decryption process of encoded data if since image is not tampered so that result displayed that received image is authenticated.

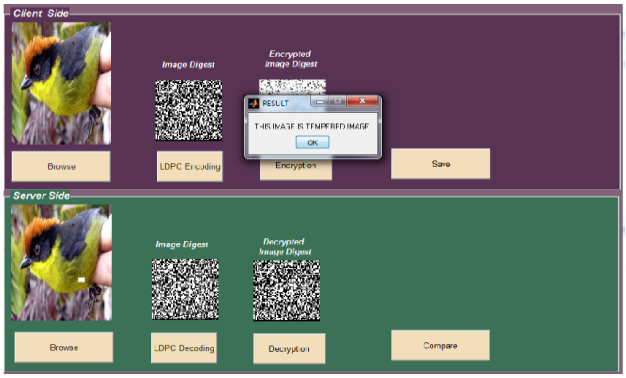

Figure 4: shows that after received the data via the transmitter side receiver start authentication process of the received data. For this process receiver apply decoding and decryption process of encoded data if since image is tampered so that result displayed that received image is not authenticated

5. Conclusion and Future Work

In this work we have shown that binary encoding can be used to data authentication using distributed source coding. Two sources using a virtual binary symmetric channel and q-ary symmetric channel. Also in this work we presents and investigates a

novel data authentication scheme that distinguishes legitimate encoding variations of an data from tampered versions based on distributed source coding and statistical methods. The decoder is extended using authentication methods to address target data. Spatial models are applied to exploit the spatial correlation of the tampering. Distributed source coding is an ideal tool for the image authentication problem in which the data sent for authentication are highly correlated to the information available at the receiver.

In the future, the proposed scheme can be extended for the compression of video and audio data. It can be used for practical implementation of distributed source coding in video communication and sensor networks. It can also be extended to more than two sources and for implementation in the time domain. Estimation of the adjustment parameters requires the target image and the original image projections, but the latter is not available before decoding.

References

[1] Khidzir, N.Z., Mohamed, A. and Arshad, N.H.H., "Information Security Risk Management: An Empirical Study on the Difficulties and Practices in ICT Outsourcing", NETAPPS 2010.

[2] Kshetri, N., "The simple economics of cybercrimes", IEEE Security and Privacy, vol. 4, no. 1, 2006.

[3] J. Liang, R. Kumar, Y. Xi, and K. W. Ross, "Pollution in P2P file sharing systems," in Proc. IEEE Information communication, Mar. 2005, vol. 2, pp. 1174-1185.

[4] A.D. Liveris, Z.Xiong and C.N.Georghiades. "Compression of binary sources with side information at the decoder using LDPC codes". IEEE Commum. Lett., vol. 6, pp.440-442 .

[5] R.G.Gallager, "Low density parity check codes". PhD thesis, MIT, Cambridge, Mass., September 1960.

[6] Matthew. C. Davey and D.J.C. Mackay. "Low density parity check codes over GF(q)". IEEE Commum. Lett., vol. 2, pp.165-167, June 1998.

[7] Y.C. Lin, "Image Authentication Using Distributed Source Coding," Ph.D. dissertation, Stanford University, Stanford, CA, 2010.

[8] D. Varodayan, A. Aaron, and B. Girod, "Rate-adaptive codes for distributed source coding," EURASIP Signal Process. J., Special Section on Distributed Source Coding, vol. 86, no. 11, pp. 3123-3130, Nov. 2006.

Sagar Chouksey received BE degree inElectronics and Telecommunication from the Rajiv Gandhi University, Bhopal in 2011. He joined Infosys Technologies in 2011and held the position of System Engineer and worked in different Financial Insurance and Banking Projects and worked in Bank of America and BRITT Insurance Projects during 2011-2013.Now he is working as a Senior Consultant and developer at Daffodil Technologies Pvt . Ltd.