GRAPH SPECTRAL DOMAIN BLIND WATERMARKING

Hiba Al-khafaji and Charith Abhayaratne

Department of Electronic and Electrical Engineering, The University of Sheffield, United Kingdom.

Email: h.alkhafaji@sheffield.ac.uk, c.abhayaratne@sheffield.ac.uk, Web: http://svc.group.shef.ac.uk

For More Information
http://tiny.cc/demogwm
@charithmetic

1. Introduction

- Recently, data recorded on non-Cartesian grids have been increased in applications like, sensor networks. Therefore, the authentication and protection of these data have become an important issue.
- The existing approaches are based on the node domain data hiding or embedding of watermarks. This can distort the graph data values and might not be robust to many attacks as in the case of pixel-domain watermarking for images.
- This work proposes a graph spectral domain blind watermarking algorithm with a low embedding distortion and high robustness against attacks.
- The main contributions of this work are:
  - Proposal of a model for choosing the graph spectral coefficients for minimising the embedding distortion.
  - Proposal of a model for enhancing the robustness of hidden data against attacks.

2. The proposed method

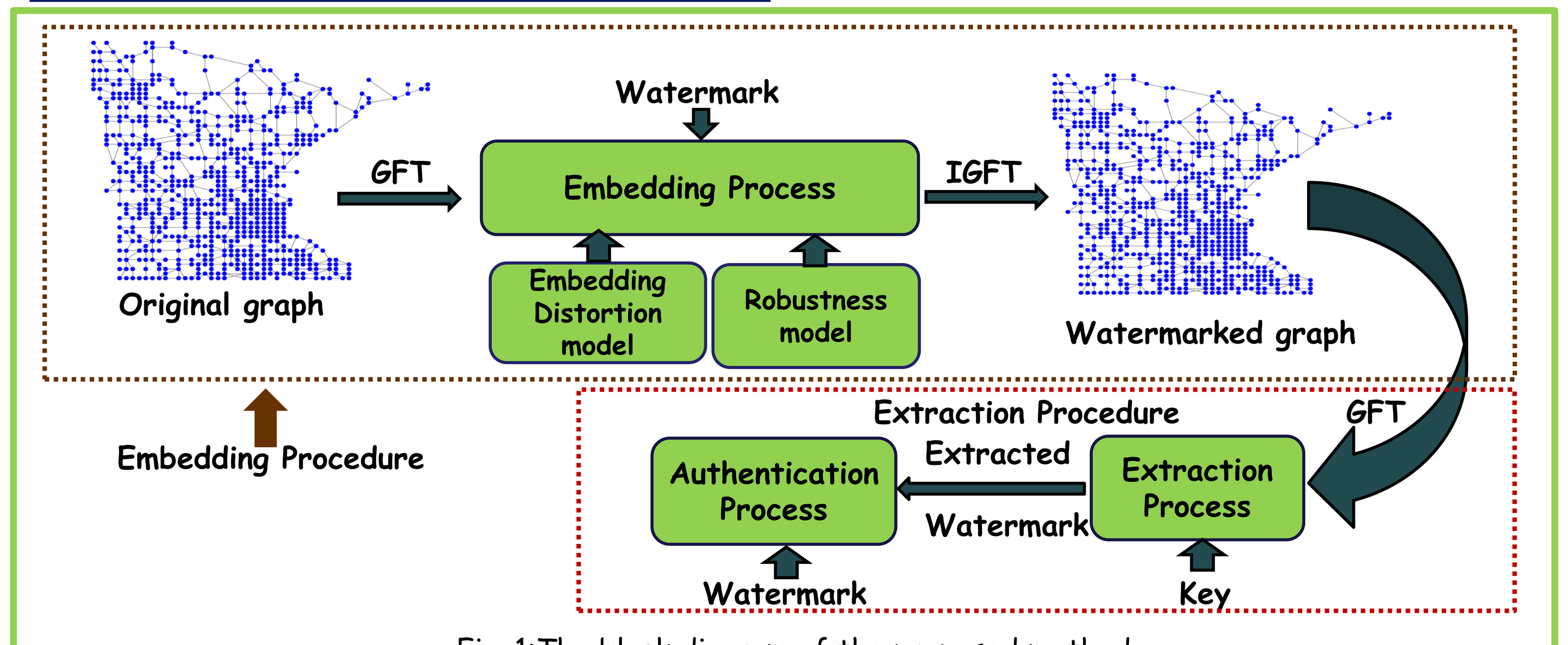

Fig. 1:The block diagram of the proposed method

2.1-Graph Fourier Transform

- Let G = {V,E,A}, is an undirected graph, where V is the set of N vertices, E is the set of edges and A is the adjacency matrix. The combinatorial graph Laplacian matrix, L, is defined as L = D−A, where D is the diagonal matrix of vertex degrees.

-An eigenvalue decomposition of L matrix as follows:

Eigenvalues of L — Eigenvectors of L

$$L = U\lambda U^t = \sum \lambda_\ell u_\ell u_\ell^t.$$

- The Graph Fourier Transform (GFT) and its inverse (IGFT) are defined as follows:

GFT coefficient
$$X(\ell) = \sum_{i=0}^{N-1} x(i)\, u_\ell(i)\,.$$

Graph signal
$$x(i) = \sum_{\ell=0}^{N-1} X(\ell)\, u_\ell^t(i).$$

2.2-GFT domain blind watermarking

- Compute the GFT coefficients and sorted in descending order.
- A non-overlapping 3 x 1 running window is passed through the sorted GFT coefficients, $X_s(m)$, to embed the watermark in the median coefficient as follows:

$$X_{sw}(m) = \left|\frac{X_s(m-1)+X_s(m+1)}{2}\right| + w. \quad \text{watermark}$$

- The watermark is extracted by passing 3 x 1 running window through the sorted watermarked GFT coefficients as follows:

Extracted watermark
$$w' = X_w(m) - \left|\frac{X_s(m-1)+X_s(m+1)}{2}\right|.$$

- The extracted watermark bit b' is determined based on a threshold T, where T=(w_0+ w_1)/2.

2.3-Embedding distortion minimisation

- To minimise the embedding distortion, the relationship between the error distortion using mean square error and the selected Graph Fourier coefficients is established to embed the watermark.

MSE
$$\mu \propto \left(\sum \left|\frac{X_s(m-1)+X_s(m+1)}{2}\right| - X_s(m)\right)^2.$$

- This leads to the model: For each embedding coefficient triple,
  $|0.5(X_s(m-1)+X_s(m+1))|$-$X_s(m)$ must be close to 0
- In other words the gradient difference, $[(X_s(m-1)-X_s(m)) - (X_s(m)-X_s(m+1))]$ must be close to 0
- This means the error distortion is increased when the gradient difference increases and the minimum MSE (μ) is obtained when the gradient difference is close to 0.

$$\mu \propto [(X_s(m-1)-X_s(m)) - (X_s(m)-X_s(m+1))].$$

2.4-On enhancing robustness

- To improve the watermarking robustness against attacks, the relationship between the watermark extraction and the effect of the attacks, namely, additive noise and nodes data deletion is established.
- Three scenarios of the watermark bits are considered: embedding only '0' bits, embedding only '1' bits and embedding '0' and '1' bits as follows:

Embed '0': To extract the correct watermark after embedding '0' bits, the GFT coefficients chosen for watermark embedding should be in the range:

$$\left|\frac{X'_w(m-1)+X'_w(m+1)}{2}\right| + w_0 \le X'_w(m) < \left|\frac{X'_w(m-1)+X'_w(m+1)}{2}\right| + T.$$

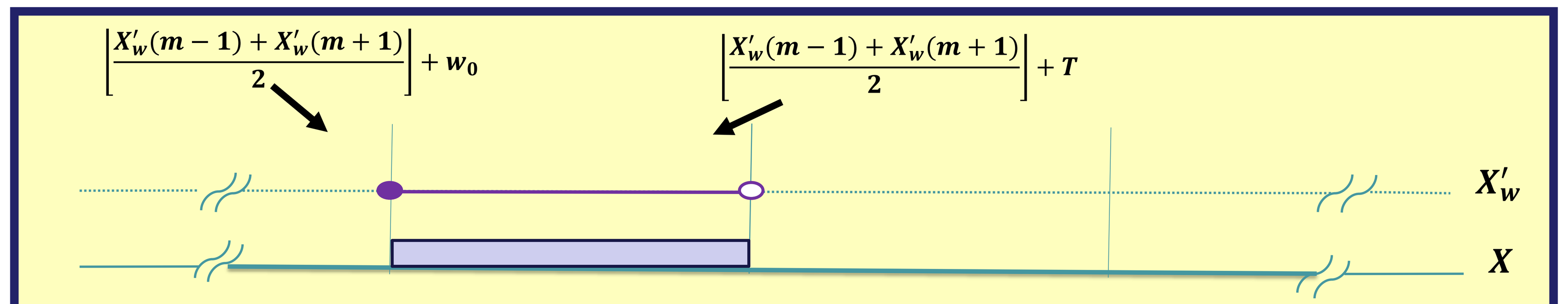

Fig. 2: The range of the GFT coefficients capable of extracting the watermark bits w ='0' correctly.

Embed '1': To extract the correct watermark after embedding '1' bits, the GFT coefficients chosen for watermark embedding should be in the range:

$$\left|\frac{X'_w(m-1)+X'_w(m+1)}{2}\right| + T \le X'_w(m) < \left|\frac{X'_w(m-1)+X'_w(m+1)}{2}\right| + w_1.$$

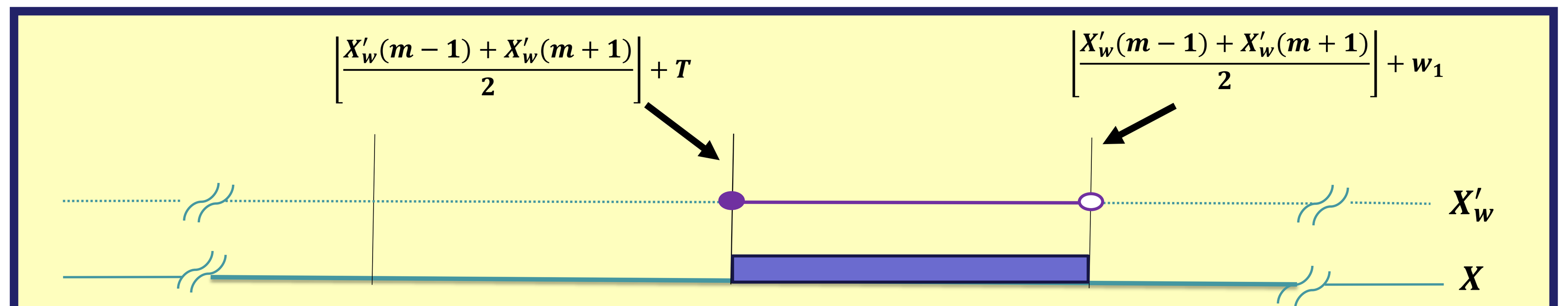

Fig. 3: The range of the GFT coefficients capable of extracting the watermark bits w ='1' correctly.

Embed '0' and '1': By combining the two cases above, we can find the condition of correct detection of the watermark bits when embedding '0' and '1'. The range of the GFT coefficients which retain the watermark bits correctly is:

$$\left|\frac{X'_w(m-1)+X'_w(m+1)}{2}\right| + w_0 \le X'_w(m) < \left|\frac{X'_w(m-1)+X'_w(m+1)}{2}\right| + w_1.$$

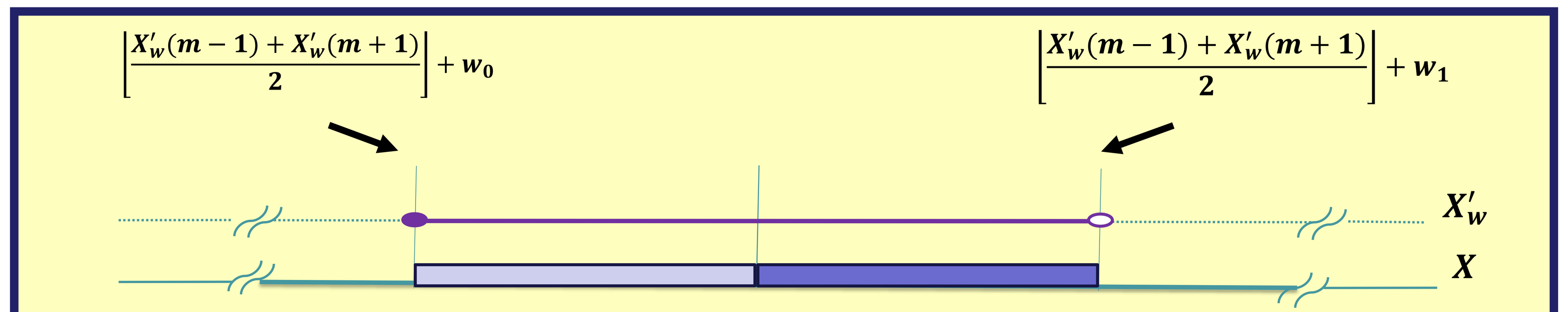

Fig. 4: The range of the GFT coefficients capable of extracting the watermark bits w ='0 &1' correctly.

3. Verification of the proposed models

3.1-Verification of the embedding distortion model

- The embedding distortion model was verified by comparing the MSE of the watermarked graph using the original blind algorithm and with the embedding distortion model for 5 embedding scenarios w='0', w='0.1', w='0.2', w='0.3' and w='0.4'.

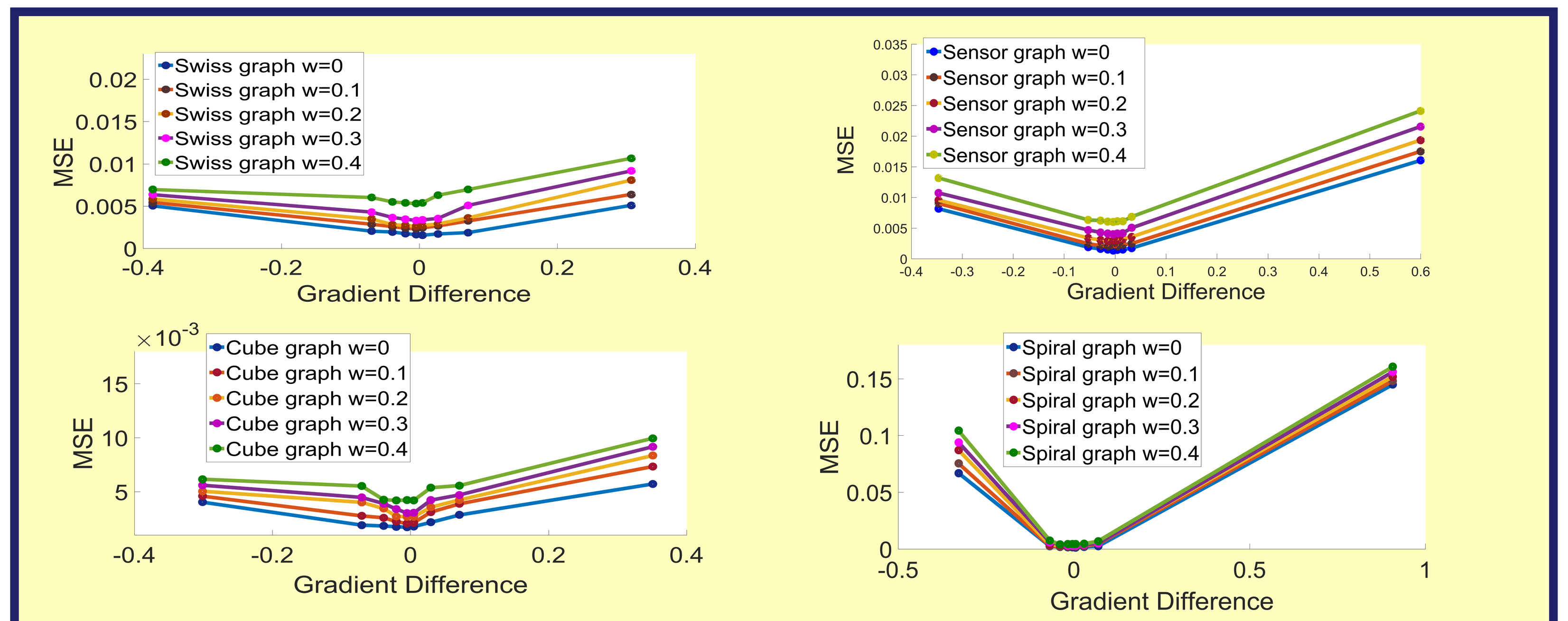

Fig. 5: Verification of the embedding distortion minimisation model for various watermark values.

3.2-Verification of the robustness model

- The robustness model was verified by comparing the Hamming distance (HD) of the extracted watermark using the original blind algorithm and with the robustness model after additive noise and deleting random nodes data for 3 embedding scenarios: Embed '0' only; Embed '1' only; and embed '0' and '1'.

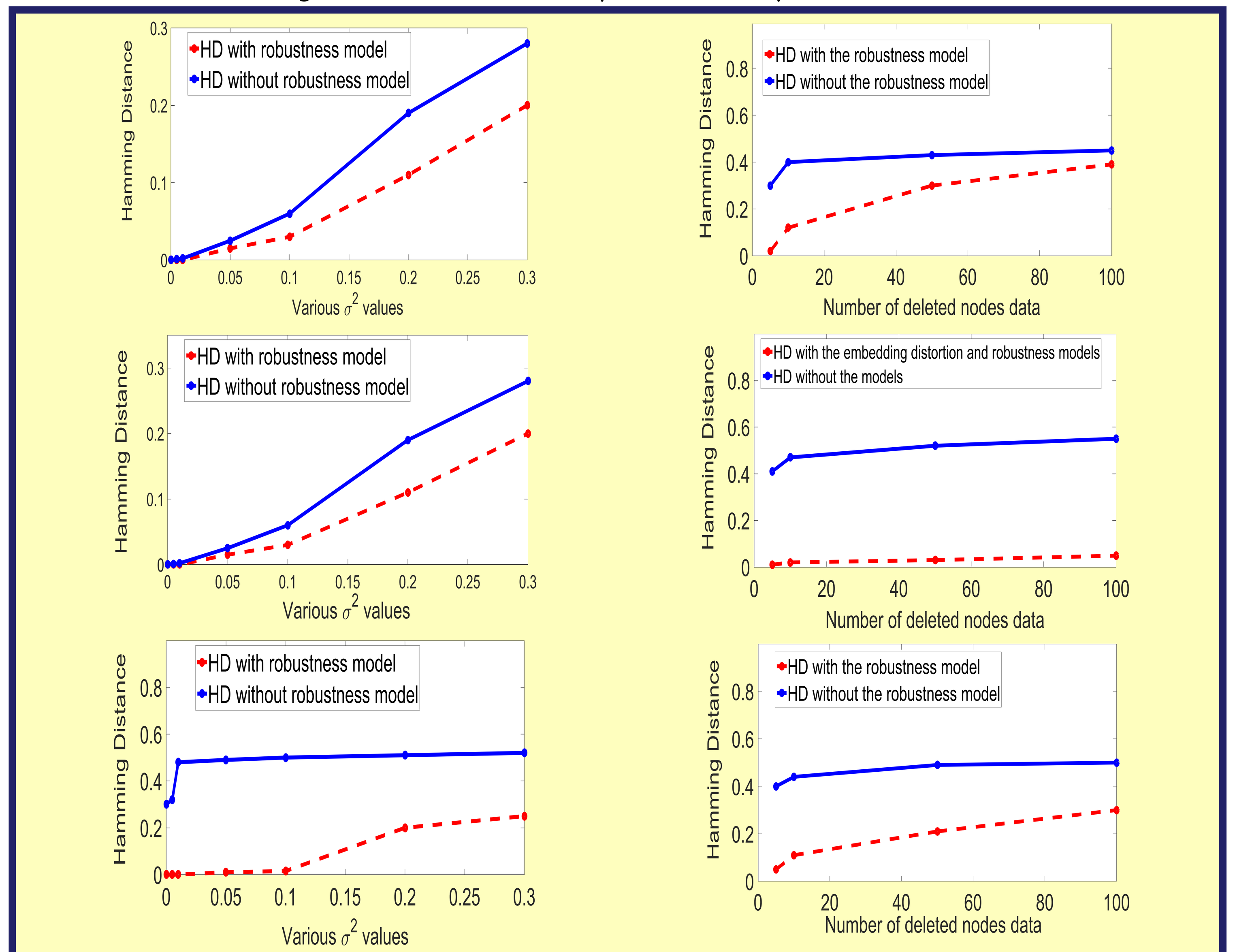

Fig. 6: Verification of the robustness model after additive noise (Column 1) and deleting random nodes data (Column 2) for 3 embedding scenarios Embed '0' (Row 1), embed w '1' (Row 2) and embed ='0 and 1' (Row 3) with choosing w_0 = '0.1' and w_1 ='0.3' for embedding '0' and '1' respectively.

4. Performance Evaluation

4.1-Embedding distortion performance

- The embedding distortion was evaluated by calculating the MSE of the watermarked graph using the original blind algorithm and with the embedding model for various embedding capacities (bits).

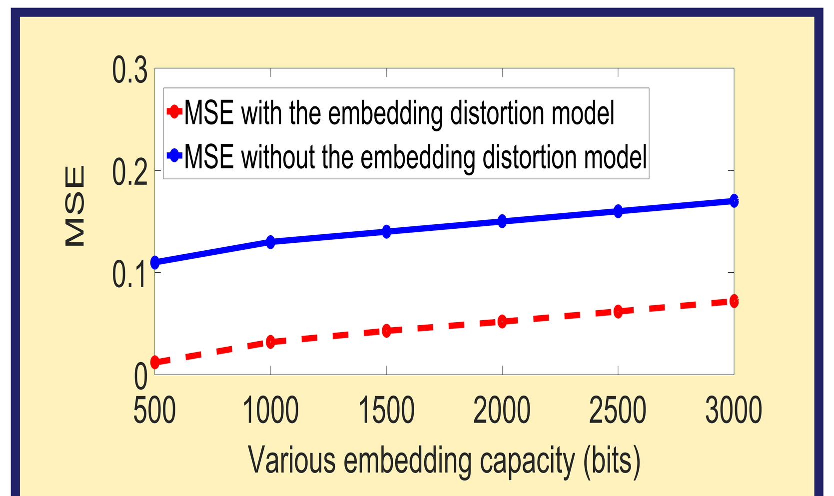

Fig. 7: Embedding distortion performance.

4.2- Robustness performance

- The robustness performance was evaluated by calculating the Hamming Distance of the extracted watermark using the original blind algorithm and with the embedding model after additive noise withσ^2 =0.1 and deletion 10 random nodes data for various embedding capacities (bits).

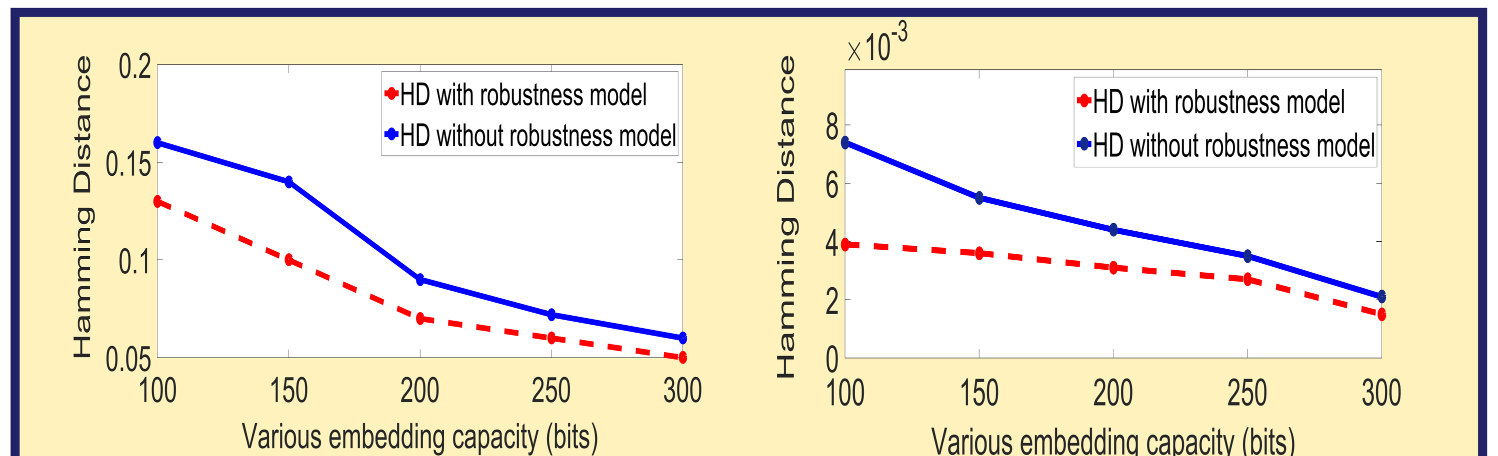

Fig. 8: Robustness performance. Column 1: After additive noise (0.1). Column 2: After deletion nodes data for (10) nodes

5. Conclusions

- We have proposed a novel graph spectral domain blind watermarking for unstructured data.
- It includes two new models for choosing GFT coefficients to embed the watermark minimising the distortion and enhancing the robustness to attacks, respectively.
- The proposed models were experimentally verified individually and together within the proposed graph spectral domain blind watermarking