CHAPTER 1: INTRODUCTION

1.1 Rationale

The rapid advancement of digital technology and the increasing need for secure communication have fueled a growing demand for efficient and reliable data-hiding techniques[1]. The era of digital advancement has ushered in an imperative for secure data transmission, a pivotal aspect in modern communications. In this digital age, ensuring the confidentiality and integrity of sensitive information emerges as a critical challenge, particularly with the widespread adoption of digital communication methods. Amidst various data security techniques, Reversible Data Hiding (RDH) has garnered significant attention. This technique uniquely allows data to be embedded in a way that enables the complete restoration of the original image after extracting the hidden data, proving invaluable in contexts where data integrity is crucial, such as military operations, medical imaging, and governmental documentation[3].

Historically, RDH has relied heavily on techniques like Pattern Substitution (PS). However, the Adaptive Pattern Substitution (PS) method, while effective, confronts notable limitations concerning embedding capacity and image distortion. These constraints become particularly problematic in high-stakes environments that demand high image accuracy and quality. To overcome these challenges, this study introduces an innovative approach the Two-Phase Adaptive Pattern Frequency Replacement (PFR) method. This method aims to enhance embedding capacity without significantly compromising image quality, positioning itself as a substantial improvement over existing techniques.

Previous works of reversible data hiding, such as the basic pattern substitution approach, often struggle with inherent limitations in capacity and distortion, which can compromise the quality and usability of the hidden data[2]. These limitations are particularly critical in sensitive applications like military mapping, medical imaging, and governmental documentation[3], where the fidelity and security of the data are paramount. The proposed research introduces the Two-Phase Adaptive PFR (Pattern Frequency Replacement) method as an innovative approach to address these challenges. By leveraging an adaptive pattern substitution mechanism and a two-phase embedding process, this method aims to significantly enhance data hiding capacity.

The Two-Phase Adaptive PFR method builds upon existing reversible data hiding techniques but introduces critical improvements through adaptive selection of patterns, guided by context window scanning. This adaptiveness allows the method to dynamically adjust to the most frequent, least frequent, and second least frequent patterns within the binary image, optimizing the embedding process to minimize visual changes. The first phase of embedding utilizes the "embedding including LM" method[4], which strategically embeds secret data by targeting infrequent patterns without disrupting the overall structure of the image. Following this, the binary image undergoes a reverse transformation before the second embedding phase is applied, further increasing embedding capacity by re-evaluating pattern selection in the reversed context.

1.2 Theoretical Framework

The theoretical framework of this research is anchored in the principles of pattern substitution methods, specifically focusing on the Most Frequent Pattern (PM), Least Frequent Pattern (PF), and Second Least Frequent Pattern (PFR). These methods form the basis for enhancing data-hiding techniques in binary images. A critical operation in this framework is the XOR transformation, which converts the binary image into a difference matrix, setting the stage for pattern substitution. The adaptive nature of the method is further facilitated by context window scanning, where a fixed length of 10 (L = 10) allows for a dynamic selection of patterns. This process ensures that PF and PFR are closely aligned within the same group class as PM, creating an optimized substitution pattern list that maintains the integrity of the image while enhancing data hiding capacity[5].

1.3 Conceptual Framework

The conceptual framework of this research outlines a systematic approach for implementing the Two-Phase Adaptive PFR method in binary images. The first step involves applying the XOR operation to transform the binary image into a difference matrix, which prepares the data for subsequent pattern analysis and substitution. The next step, context window scanning, involves examining the image with a context window of length 10 to identify the most frequent (PM), least frequent (PF), and second least frequent (PFR) patterns. PM is the pattern that appears most frequently within the context window and is typically left unchanged during the embedding process to preserve the image's original appearance. In contrast, PF, being the least frequent pattern, is ideal for substitution with secret data due to its minimal impact on the image's visual quality. Similarly, PFR, the second least frequent pattern, provides additional opportunities for embedding secret data without significantly altering the image’s structure[6].

The embedding process is executed in two distinct phases. In the first phase, embedding is performed using the “embedding including LM” method, which targets the least frequent patterns for substitution. This is followed by a reverse transformation of the binary image, where the order of pixels is flipped from last to first, setting the stage for the second phase of embedding. The second phase again employs the “embedding including LM” method, but this time on the reversed image, further maximizing the embedding capacity.

1.4 Statement of the Problem

The Adaptive Pattern Substitution (PS) method, integral to RDH, is adept at embedding data within binary images while maintaining the reversibility of the process. Yet, it faces substantial challenges: it is restricted in the volume of data it can embed without inducing significant degradation of image quality, and it frequently causes high image distortion during the embedding process especially when altering or substituting patterns. Such distortions are particularly detrimental in fields requiring high data integrity and pristine image quality, such as military mapping and medical imaging. In these applications, any distortion can lead to critical inaccuracies, underscoring the need for more refined RDH techniques.

1.5 Hypothesis

The primary aim of this research is to refine RDH techniques by introducing a novel approach that not only increases the capacity for data embedding but also minimizes the distortions introduced during the data embedding process. The specific objectives of this study are to design and implement the Two-Phase Adaptive Pattern Frequency Replacement (PFR) method for binary images, evaluate its effectiveness in enhancing embedding capacity over the Adaptive PS method, analyze its impact on image distortion, particularly assessing metrics like PSNR and SSIM, and examine its resilience against various noise conditions including salt & pepper noise and physical damage.

This study hypothesizes that the Two-Phase Adaptive PFR method will substantially increase the embedding capacity without degrading the image quality as significantly as the Adaptive PS method. However, it anticipates certain compromises in terms of distortion, especially under conditions where images undergo flip mirroring, potentially leading to greater pattern disruptions and diminished PSNR and SSIM values.

1.6 Assumption

This research assumes that the binary images used for testing represent typical use cases in data hiding applications. It is also assumed that the context window length of 10 (L = 10) is optimal for adaptive pattern selection, encoder, and decoder using binary bit and knows what is PM, PF, and PFR patterns and where the location of secret data & location map is embedded.

1.7 Scope and Delimitation

This research focuses on binary images and implementing the Adaptive PS method. The scope of the study includes algorithm design, security analysis, and practical applications in the context of image processing. The research findings include XOR operation, context window scanning, pattern selection, and a three-round embedding process. The research does not cover other types of images (e.g., color images) or alternative data-hiding techniques. The study is limited to the specific parameters and methods described.

1.8 Importance of the Study

The findings of this research are poised to contribute significantly to the field of steganography and reversible data hiding by offering a more efficient method for embedding secret data into binary images. The Two-Phase Adaptive PFR method not only aims to increase the embedding capacity but also to enhance the overall security of the embedded data, making it especially suitable for use in high-risk environments such as military mapping, governmental documentation, and sensitive medical imaging.