DNA sequences compression by GP2R and selective encryption using modified RSA technique

ABSTRACT

Humans, by nature, have always been fascinated by the possibility of being able to acquire more information in minimum possible time and space. The effective lossless compression method, effective data structure, and DNA (Deoxyribonucleic Acid) data searching are quite essential as they provide a stimulus to easy accessibility and communication. The proposed algorithm is a new Lossless Compression algorithm, which compresses data, based on two tiers. Firstly, it searches for the exact Genetic Palindrome(GP), Palindrome(P) and Reverse(R)[GP 2 R] and the substring is reported, which is replaced by the corresponding ASCII character creating a Library file. By using the ASCII code, the Library file acts as a signature as well as provides the security of data. Secondly, modified RSA technique is proposed for the selection encryption purpose. This selection encryption of the modified RSA technique is an approach to lessen computational resources for greatly sized DNA facts. The experimental work shows 44% to 45% original sequence is encrypted where above 95% of the original file is damaged by using this method. This technique can find out the 3.851273 bits per base of the compression rate. The O(n) is the complexity of this algorithm. The running time is a few seconds of this algorithm. This is a hybrid approach to the compression & encryption process. For reducing the compression rate, the first pass output is again compressed by the second pass but it is lossy, This experiment is performed on benchmark DNA order.

Keyword: Reverse; Genetic palindrome; Palindrome; Compression; Ratio; Rate; Encryption; Speed; Genetic palindrome; Palindrome; Reverse (GP2R)