DNA sequence approximate matching is one of the main challenges in Bioinformatics. Despite the evolution of new technology, there is still a need for new algorithms that accommodate the huge amount of Bioinformatics data. In this paper, a parallel n-gram approach is proposed with a method that is taking in mind the variety of DNA sequence lengths for approximate matching. The proposed approach showed a satisfiability result in terms of time complexity compared to parallel dynamic programming method.

References


Index Terms
Keywords

DNA Sequence, Longest Common Sequence, N-gram, Parallel