Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7778477

ACCELERATING THE PERFORMANCE OF MULTIPLE SEQUENCE ALIGNMENT USING HYBRID SWARM ALGORITHM

Karamjeet Kaur[1],Anil Kumar Sagar[2], Sudeshna Chakraborty[3]
[1]School of Engineering and Technology, Sharda University, Greater Noida, U.P., India
E-mail: karam_7378@yahoo.com
[2]School of Engineering and Technology, Sharda University, Greater Noida, U.P., India
E-mail: aksagar22@gmail.com
[3]Galgotias University, Greater Noida, U.P., India
E-mail: sudeshna2529@gmail.com
*Corresponding author: karam_7378@yahoo.com

Abstract
Multiple sequence alignments are the core of bioinformatics and exploring life science. The emerging methods of information technology explode the area of RNA sequencing. The approach of sequencing moves to next-generation sequencing (NGS). Recently various researchers proposed genome sequencing methods which are based on CUDA, parallel programming, and swarm intelligence-based algorithms. The incremental approach of algorithms for genome sequencing increases score points and reduces the execution time of analysis. This paper proposed hybrid methods for accelerating multiple sequence alignment. The proposed algorithm combines particle swarm optimization and a genetic algorithm. This algorithm is very efficient in terms of score points of alignment and reduces the execution time of patterns. The proposed algorithm was tested on MATLAB tools and validated with three datasets: chimpanzee, mouse, and dog. The experimental results suggest that the proposed algorithm is very efficient than existing algorithms based on CUDA and GA.
Keywords: - MSA, CUDA, GA, PSO, CPU, bioinformatics, swarm intelligence

I. Introduction

Genome sequencing is an important phase of computational biology for understanding data in life science. The genome sequencing process is a very challenging task due to its high computational value regarding time. The applicability of genome sequencing is increased in every field, such as developing new medicine, vaccine, and the nurturing of seeds [1,2,3]. The utility of genome sequencing brings the attention of various researchers to the development of algorithms for the fast processing of sequence alignment. Recently CUDA programming-based algorithm has been the most important algorithm for multiple sequence alignment. The CUDA algorithm is a set of parallel programming that increase the utilization of the central processing unit and graphical processing unit of a computer machine [4].By gradually aligning the nearest pair, the progressive procedure is the most extensively used heuristic method for creating MSA. Its fundamental flaw stems from the algorithm's greedy character, which causes discovered alignments to become stuck in local optima. This implies that errors in intermediate matches can't be fixed later if more sequences are thrown into the mix. Moreover, there is no function for evaluating sequence alignment performance. Another approach is to employ a multiple

objective extension like the Carrillo-Lipman approach, MSA, DCA, and so on [5,6] to align many sequences at the same time. These algorithms, on average, yield better outcomes than progressive algorithms. However, these algorithm's complexity in terms of running time and memory requirements are drawbacks.As a consequence, they should only be employed to solve problems with a restricted quantity of iterations. Several studies have analyzed the alignment accuracy [7, 8], processing time and memory use, and other aspects of the MSA method's performance using these benchmark datasets. According to the findings of these investigations, each MSA approach has its own set of strengths and shortcomings, and no MSA method is flawless on all benchmark datasets. These, on the other hand, indicated that all of the MSA techniques had certain similar flaws. One disadvantage of MSA approaches that cannot be overlooked is uncertainty. According to a study, ambiguity in MSA alignment can cause various issues, including different alignment approaches yielding different outcomes. When utilizing MSA approaches, the presence of a large number of highly varied sequences has been found to influence the alignment of sequences with a modest genetic distance. Compared to PSA approaches, negative impacts on clustering results were another downside. According to 16 s rRNA datasets study, MSA-based clustering approaches performed worse than PSA-based clustering methods[10,11,12]. This paper focuses on multiple sequence alignment using the hybrid algorithm of swarm intelligence. Particle swarm optimization and a genetic algorithm are used in the hybrid algorithm. The genetic algorithm is a very convincing algorithm for a limited set of data, but the length of data increases the processing of the genetic algorithm and behaves just like standard searching algorithms. The PSO algorithm is a meta-heuristic algorithm that handles all data variants in manners small and large. The process of sequence selection function generated by genetic algorithm and the processing of optimal results selection using PSO. The main objective of this study is to decrease the execution time of multiple sequence alignment and increase the score point of the MSA algorithm. The algorithm's processing was tested on three genome sequencing datasets such as chimpanzee, mouse, and dog. The rest of the paper is organized as in section II related work, and in section III proposed methodology, section IV describes the experimental analysis of the proposed algorithm, and section V concludes the paper.

II. Related Work

Despite several studies of MSA algorithms, most authors considered the execution time and score point as a significant issue of multiple sequence alignment. The incremental approach of algorithms development moves on to next-generation computational biology data analysis. Some recent contribution describes here.In this [1] researchers investigate the original MOPSO's confluence. With probability theory, the developed convergence metric is applied to the original MOPSO's convergence. The original MOPSO is not a global convergence method, according to their findings, because it cannot ensure global convergence with probability one.In this [2] researchers apply a co-evolutionary multi-colony ant colony optimization to robot path planning. To begin, a DGCA is created, which is then combined with ACS and MMAS to produce a multi-population heterogeneous structure. To increase the overall performance of optimization,each subpopulation co-evolves and complements one another. In this [3] researchers propose an assembly model which is used to include different assembly constraints and rules. In contrast to the ACO algorithm, the hybrid SOS-ACO technique provides the best and closestplanning in shorter cycles. SOS-ACO is also successful in determining the

appropriate construction sequence in nearly every experiment. Finally, studies are conducted to determine how effectively SOS-ACO operates when ACO settings are changed.In this [4] researchers propose a new hybrid refining phase based on progressive alignment has been included. Researchers can conclude that their method can refine the MSA produced by GA-based algorithms and provide better biologically meaningful alignments. Furthermore, the statistic hypothesis test reveals that the differences between the acquired findings, as well as the improvements, are statistically significant.In this [5] researchers address the VNF placement and routing problem, the fuzzy heuristic-based ant colony optimization is introduced as a fuzzy knowledge-based version of ACO. Simulation findings on US and Pan-European network topologies demonstrate that the suggested FH-ACO method outperforms current techniques.In this [6] researchersanalyse the incidence of the inherent disorder in their proteins, researchers were able to investigate SARS-CoV-2, bat SARS-like, and human SARS-CoVthat have dark proteomes. In this study, researchers used data from the SARS-CoV-2 genome and the translated proteome from GenBank to conduct a comprehensive computational analysis of the prevalence of the intrinsic disorder in SARS-CoV-2 proteins.In this [7] researchers improve sequence homology searches, protein profile calculation, and hence homology-based function annotation, template-based, and even de novo protein structure prediction by including sequences from under-sampled eukaryotic lineages. As a consequence, more research on eukaryotic activity in varied environments will be possible.In this [8] author introduce the bio info portal gateway in detail, including its structure, features, and integration with the CS grid middleware, which is used to operate the Brazilian national high-performance computing environment. The performance analysis of application demonstrates how machine learning was utilized to increase the functionality of bio info portal by recommending predictive models for resource allocation that achieved a performance efficiency of over 75%.In this [9] researchers propose a multimodal approach combining radiosurgery and radiation therapy is becoming increasingly popular for achieving a satisfactory functional outcome and tumor control. Advances in technology, genetics, and radionics could lead to improved tumor biology profiling and, as an outcome, treatment modification based on precision medicine concepts.In this [10] researchthe secondary structure of mitochondria RNA was looked as a reaction of the secondary mitochondrial genome after changes. The mitochondrial DNA sequences of LHON patients were extracted using Sanger sequencing from six different bloodlines. To assess the impact of polymorphisms in mitochondrial RNA genes, researchers conducted sequence conservation analyses in silico models of secondary structural and regional 3D structures.In this [11] researchers examine the most up-to-date methods for detecting pathogenic variations in autoimmune disorders, as well as accessible sequencing alternatives and methodologies and tactics in bioinformatics to realize the capability of diagnostic and therapeutic programs in which person variant intelligence is used to improve clinical practice, the development of accurate and robust sequencing and computational techniques to find deleterious variations is crucial.In this [12] researchers employed RNA oligonucleotides as viral baits in a hybridization capture technique combined with high throughput scanning to detect appropriate and relevant new mammalian viruses from eDNA/DNA. This strategy will be beneficial in supplementing or replacing invasive approaches in contexts where invasive approaches are not possible.In this [13] researchers proposed a comprehensive and versatile Python-based toolkit for DNA, RNA, and protein sequences that integrates the functionality of feature extraction, clustering,

normalization, selection, dimensionality reduction, predictor construction, best descriptor/model selection, ensemble learning, and results from visualization. In this [14] author presentsiLearnas a Python-based toolset for extracting features from DNA, RNA, and protein sequences, as well as grouping, dimensionality reduction, predictor creation, optimal descriptor/model selection, ensemble learning, and outcomes presentation. iLearn includes several DNA, RNA, and protein descriptors, as well as four feature output formats that make it easier to use direct output or interface with other computational tools.In this [15] researchers propose the DE-CQPSO method which is based on the particle diversity of genetic algorithm's crossover operators and the quick convergence of differential evolution algorithms. The crossover probability is updated using a parameter adaptive control method to improve the optimization results. In addition, to deal with the difficulty of multi-objective optimization, a penalty factor is used.In this [16] author revealed the coronavirus RNA polymerase which is modelled, validated, and then targeted using anti-polymerase medications that have been approved for use against a variety of viruses. Because they bind strongly to the virus's RDRP, studies demonstrate that Ribavirin, Reformulations, Sofosbuvir, and Tenofovir are effective treatments against SARS-COV-2.In this [17] researchers access the whole mt DNA sequencing data, which is constantly increasing, thanks to the constantly improved next-generation sequencing methods. Researchers employed an integrated method that began with RNA rather than DNA to maximize the novelty of their study. With the same quantity of sequencing mt DNA data, RNA-Seq would result in substantially poorer coverage since it is more difficult to cover the full-length mt DNA, including all regulatory areas deletions and rearrangements that could be missed from the transcriptomics data. In this [18] researchers give a brief explanation of how AI methods and microalgae bioinformatics could be used to improve present tactics for manufacturing desirable microalgae products by extracting important information and anticipating molecular interactions from genetic data in gene sequencing and editing. Recent advances in computer vision and machine learning algorithms have made exact strain species screening and categorization possible, resulting in high-quality microalgae photos that may be used for further research.In this [19] researchers establish for the first time that DNA barcoding is a reliable tool for distinguishing between Tribulus species that are closely related. The taxonomic issues in the genera were also resolved as a result of this research.

In this [20] researchers propose METTL3-m6A-YTHDC1 axis which governs the production of DNA-RNA hybrids at DSB locations, which in turn attract RAD51 and BRCA1 for HR-mediated repair. Cancer cells and mouse xenografts are substantially more vulnerable to DNA damage-based therapy as a result of METTL3 knockdown. These findings shed information on METTL3 and YTHDC1's functions in HR-mediated DSB repair, which could have consequences for cancer treatment.In this [21] researchers show the SARS-CoV-2 neutralizing antibodies might be chosen only based on their predicted CDR3H structures, which are identical to those of SARS-COV neutralizing antibodies. Overall, the researchers showed that high-throughput single B-cell sequencing might be utilized to swiftly discover human-neutralizing antibodies in the event of a pandemic.In this [22] researchers used a multi-objective optimization framework to create 2D-FGPs gender-fluid functionally graded plates with varying thicknesses. Superior solutions were developed in contrast to alternatives by improving the 2D-FGP material and thickness distributions at control locations using the approved GIGA– IMOPSO approach.In this [23] researchers propose pairing OD with IPSO

to eliminate these problems. OD is used to arrange the position of particles and to pick just the particles with the best solution for subsequent iterations, resulting in a significant reduction in computational cost. The bridge's model was updated to reduce the disparities between numerical and experimental outcomes.

III. Proposed Methodology

The proposed algorithm of multiple sequence alignment is based on particle swarm optimization and genetic algorithm. The process of the proposed algorithm describes as initialization, crossover, mutation, fitness selection, and optimal multiple sequence alignment using particle swarm optimization. The process of GA and PSO (particle swarm optimization) algorithm is mentioned in figure 1.

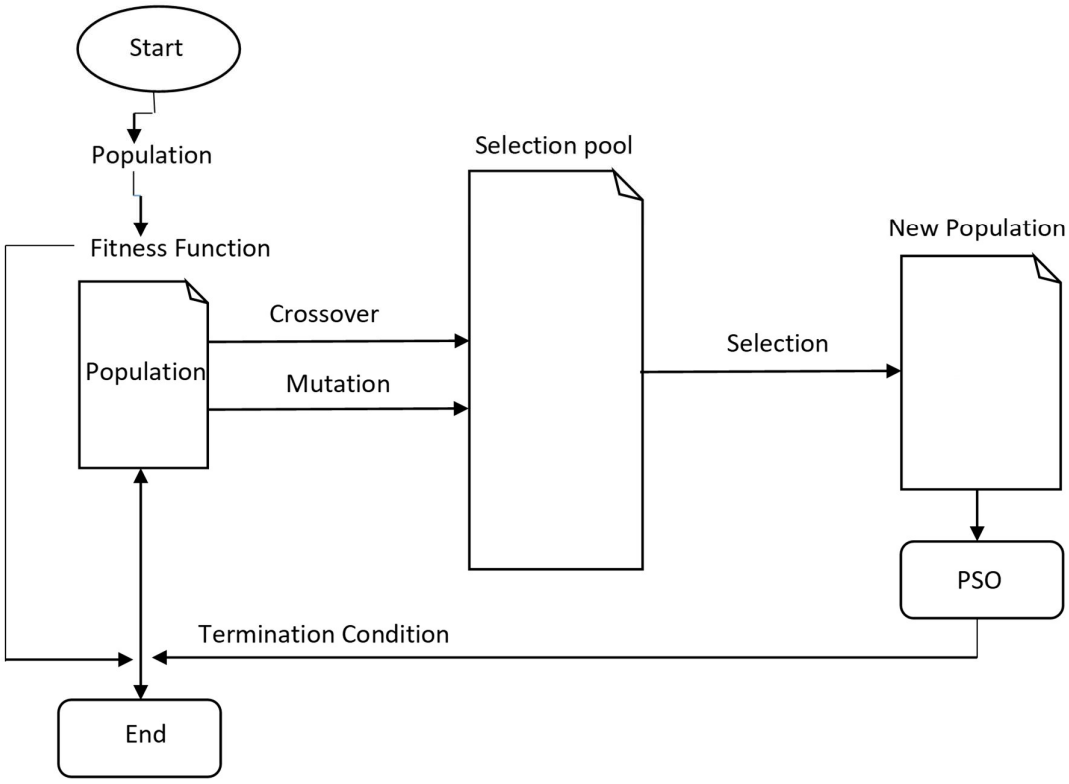

Figure 1:Proposed algorithm for multiple sequence alignment using GA-PSO

The processing of the proposed algorithm represents the K sequence, the sequence of cells isa variable length,the alignment of parents as a matrix, each sequence encoded as a row with an alphabet set.

The population of the algorithm is generated randomly with their encoded rows with length l. After the processing of the population, the selection task is performed by fitness selection function, the fitness selection function formulates as:

$$Fitness(xi, xj) = \sum_{1 \le p \le N} C(xip, xjp) \dots\dots\dots\dots\dots\dots (1)$$

Where xi and xjare two sequences of alignment, now alignment selection as

$$fitness(A) = \sum_{i=1}^{k-1} \sum_{j=j=1}^{k} fitness(xi, xj) \dots\dots\dots\dots (2)$$

The process of mutation done by the parent1 and parent2 using a probabilistic mutation operator value as 0.07. After the mutation of the genetic algorithm, the new population set passes through the particle swarm optimization.

The optimal sequence alignments choose the best placements of the population in multiple sequence alignment. With particle swarm optimization, the procedure of alignment optimal sequence estimates. A well-known swarm intelligence algorithm is particle swarm optimization. The acceleration function, constant factor, initial velocity, ultimate velocity, and particle location are used to process particle swarm optimization [18]. The algorithm's processing is described here.

The description of optimization of coefficient describes as S- feature space vector of particle. Now the space of particle is $\mathrm{xi} = (\mathrm{xi1}, \mathrm{xi2}, \mathrm{xis}, \ldots\ldots\ldots\ldots\ldots\ldots\ldots, \mathrm{XiS})$. The vector of velocity space is represented as $\mathrm{Vi} = (\mathrm{Vi1}, \mathrm{Vi2}, \ldots\ldots\ldots \mathrm{Vis})$.The processing of features changes the position of particle vector $\mathrm{Pi} = (\mathrm{pi1}, \mathrm{pi2}, \ldots\ldots\ldots, \mathrm{pis})$. The estimation of updated position as

$$\mathrm{vis} = \mathrm{Vis} + \mathrm{C1a1}(\mathrm{Pis} - \mathrm{xis}) + \mathrm{c2a2}(\mathrm{Pgs} - \mathrm{xis}) \ldots\ldots (3)$$

$$\mathrm{xis} = \mathrm{xis} + \mathrm{Vis} \qquad \mathrm{s} = 1, \ldots\ldots \mathrm{s}, \ldots\ldots\ldots\ldots \mathrm{S} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

the range of factors is [0,1]

the selection coefficient as K =C1+C2

$$k = \begin{cases} \dfrac{2x}{L(p) - 2 + \sqrt{x^2 - 4x}} & for\ X > 4 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (5) \\ x, otherwise \end{cases}$$

The value of K updated the region of the feature coefficient

$$\mathrm{Vis} = \mathrm{K}\big(\mathrm{vis} + \mathrm{c1a1}(\mathrm{Pis} - \mathrm{xis}) + \mathrm{c2a2}(\mathrm{Pgs} - \mathrm{xis})\big) \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

IV. Experimental Analysis

To validate the proposed algorithm for multiple sequence alignment, simulate in MATLAB tools. The simulation process was carried out on i7 processor and 16GB RAM with Linux operating system. The three datasets have been tested, chimpanzee, mouse, and dog protein sequence. r = 0.1, c = 0.1, q0 = 0.8, b =2, number of particles = 100, population size = 500, and iteration = 5000 are the default parameters. The crossover probability (Pc) is set to 0.5 and the mutation probability (Pm) is set to 0.05 for genetic operators. Following that, this is likely to be equipped with spermatogenesis and recombination technicians and choose the corresponding operators. The performance of algorithms is estimated as score points and execution time[20,21,22,23].

Figure 2,3 and 4 shows that hybrid swarm algorithm reduces the execution time by varying sequence length as compared to Genetic and CUDA based algorithm when datasets of chimpanzee, dog and mouse are used.

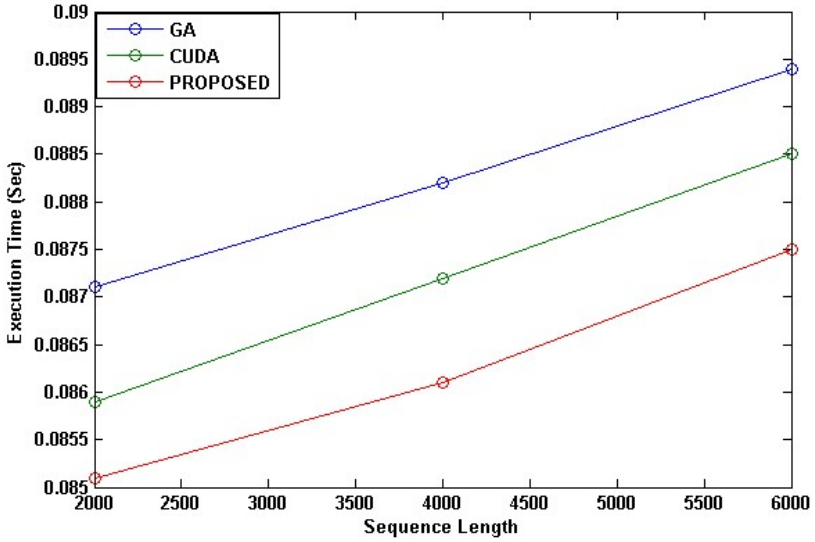

Figure2: Comparative analysis of chimpanzee dataset using GA, CUDA, and proposed model with the help of execution time and sequence length parameter. Here we observe that the execution time of the proposed method takes less time than the other two techniques GA, CUDA.

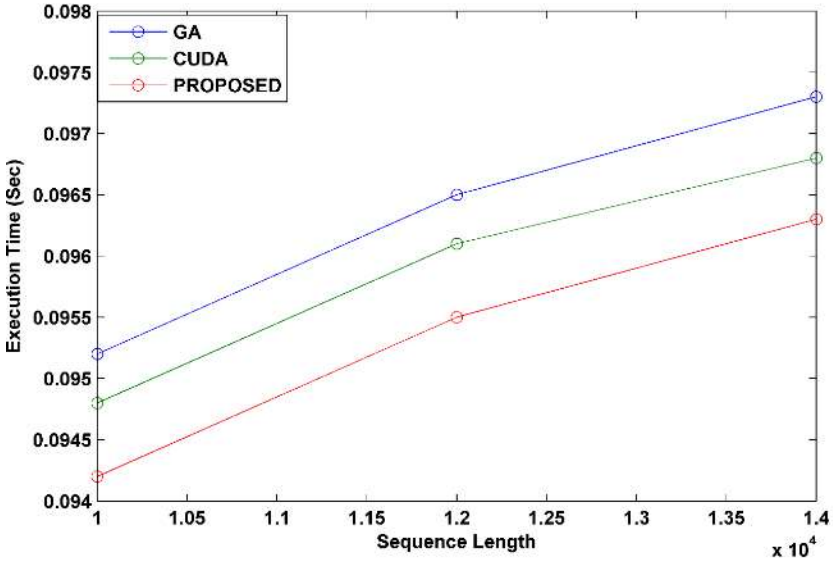

Figure3: Comparative analysis of dog dataset using GA, CUDA, and proposed model with the help of execution time and sequence length parameter. Here we observe that the execution time of the proposed method takes less time than the other two techniques GA, CUDA.

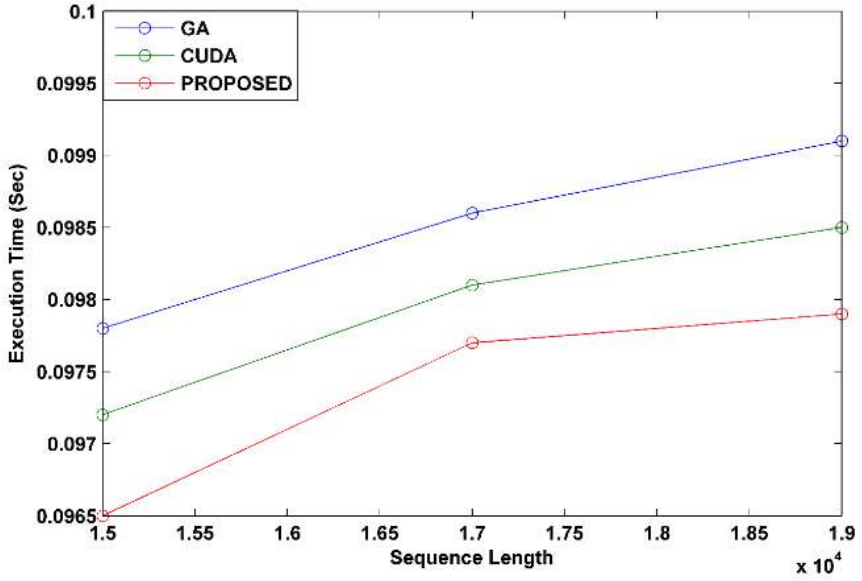

Figure4: Comparative analysis of mouse dataset using GA, CUDA, and proposed model with the help of execution time and sequence length parameter. Here we observe that the execution time of the proposed method takes less time than the other two techniques GA, CUDA.

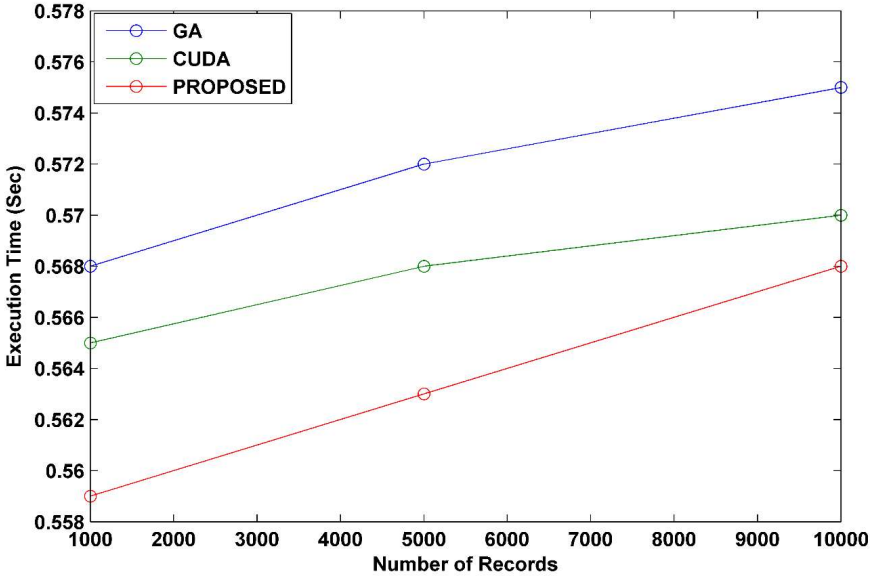

Figure5: Comparative analysis of chimpanzee dataset using GA, CUDA, and the proposed model with the help of execution time and the number of records parameters. Here we observe that the execution time of the proposed method takes less time than the other two techniques GA, CUDA.

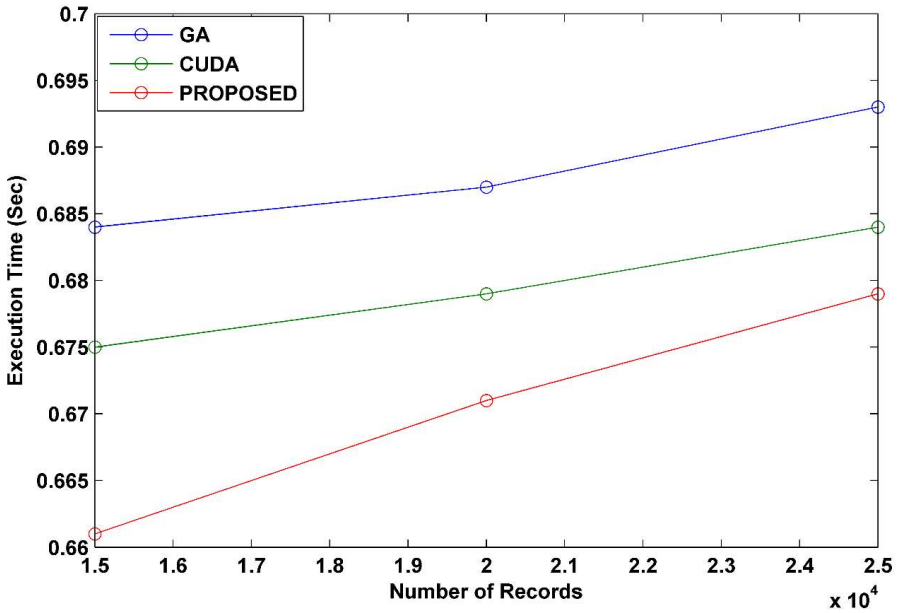

Figure6: Comparative analysis of dog dataset using GA, CUDA, and proposed model with the help of execution time and the number of records parameters. Here we observe that the execution time of the proposed method takes less time than the other two techniques GA, CUDA.

Figure 5, 6 and 7 shows that hybrid swarm algorithm reduces the execution time by varying number of records as compared to Genetic and CUDA based algorithm when datasets of chimpanzee, dog and mouse are used.

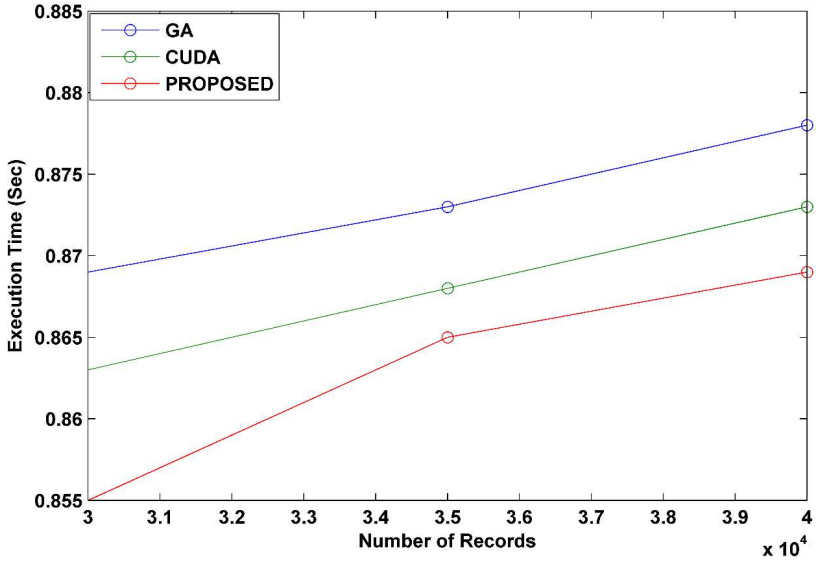

Figure7: Comparative analysis of mouse dataset using GA, CUDA, and proposed model with the help of execution time and the number of records parameters. Here we observe that the execution time of the proposed method takes less time than the other two techniques GA, CUDA.

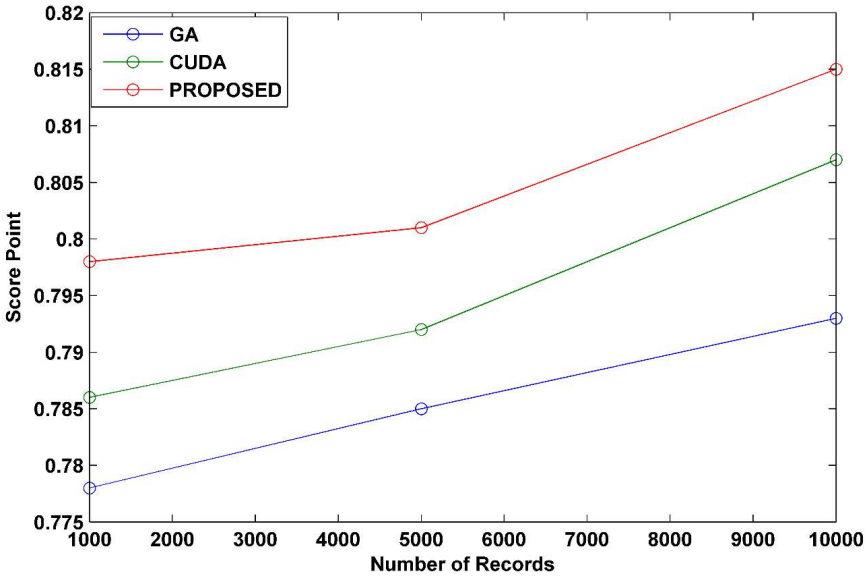

Figure8: Comparative analysis of chimpanzee dataset using GA, CUDA, and proposed model with the help of score point and the number of records parameters. Here we observe that the score point of the proposed method is more than the other two techniques GA, CUDA.

Figure 8, 9 and 10 shows that hybrid swarm algorithm increase the score point by varying number of records as compared to Genetic and CUDA based algorithm when datasets of chimpanzee, dog and mouse are used.

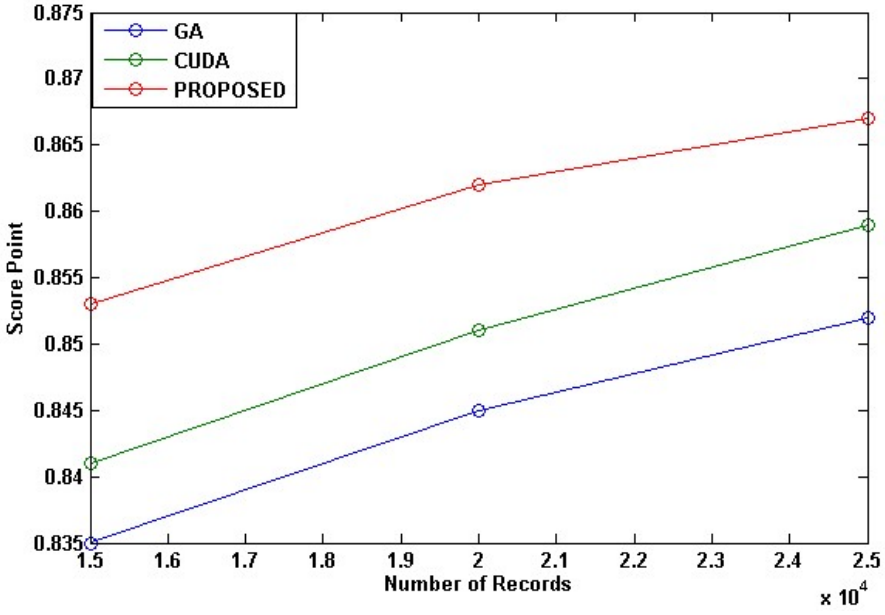

Figure9: Comparative analysis of dog dataset using GA, CUDA, and proposed model with the help of score point and the number of records parameters. Here we observe that the score point of the proposed is more than the other two techniques GA, CUDA.

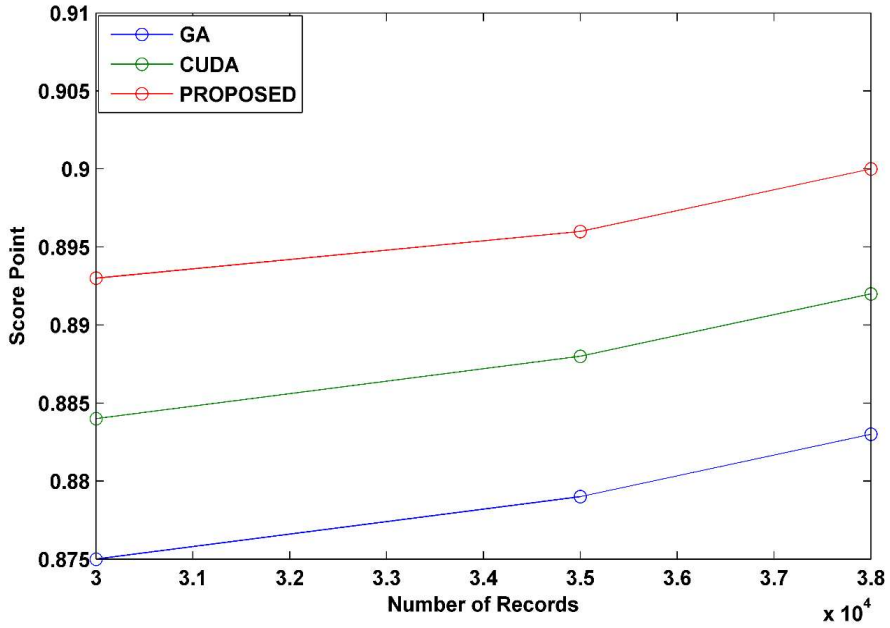

Figure10: Comparative analysis of mouse dataset using GA, CUDA, and proposed model with the help of score point and the number of records parameters. Here we observe that the score point of the proposed is more than the other two techniques GA, CUDA.

V. Conclusion & Future Work

This paper proposed a technique for aligning multiple sequences. The new technique is a hybrid of genetic and PSO technique. The suggested approach solves the scaling problem in protein databases. We evaluated hybrid GA–PSO technique to two other approaches, including guided and unguided. On genuine protein sequences, we tested our method. According to the findings, the GA strategy was unable to outperform the CUDA sequential approach for short sequences. However, when employing the hybrid strategy, the performance improvement was 4 % times faster than when using the sequential approach. We also tested the approach's scalability by randomly creating sequences of various lengths with a particular resemblance.

References

[1]. Xu, Gang, Kun Luo, Guoxiu Jing, Xiang Yu, XiaojunRuan, and Jun Song. "On convergence analysis of multi-objective particle swarm optimization algorithm." *European Journal of Operational Research* 286, no. 1 (2020): 32-38.

[2]. Li, Shundong, Xiaoming You, and Sheng Liu. "Co-evolutionary Multi-Colony Ant Colony Optimization Based on Adaptive Guidance Mechanism and Its Application." *Arabian Journal for Science and Engineering* (2021): 1-19.

[3]. Han, Zunpu, Yong Wang, and De Tian. "Ant colony optimization for assembly sequence planning based on parameters optimization." *Frontiers of Mechanical Engineering* (2021): 0.

[4]. Zafalon, Geraldo Francisco Donegá, Vitoria Zanon Gomes, Anderson Rici Amorim, and Carlos Roberto Valêncio. "A Hybrid Approach using Progressive and Genetic Algorithms for Improvements in Multiple Sequence Alignments." (2021).

[5]. Shokouhifar, Mohammad. "FH-ACO: Fuzzy heuristic-based ant colony optimization for joint virtual network function placement and routing." *Applied Soft Computing* 107 (2021): 107401.

[6]. Giri, Rajanish, Taniya Bhardwaj, Meenakshi Shegane, Bhuvaneshwari R. Gehi, Prateek Kumar, KundlikGadhave, Christopher J. Oldfield, and Vladimir N. Uversky. "Understanding COVID-19 via comparative analysis of dark proteomes of SARS-CoV-2, human SARS and bat SARS-like coronaviruses." *Cellular and Molecular Life Sciences* 78, no. 4 (2021): 1655-1688.

[7]. Karin, Eli Levy, MilotMirdita, and Johannes Söding. "MetaEuk—sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics." *Microbiome* 8, no. 1 (2020): 1-15.

[8]. Ocaña, Kary ACS, Marcelo Galheigo, Carla Osthoff, Luiz MR Gadelha Jr, Fabio Porto, Antônio Tadeu A. Gomes, Daniel de Oliveira, and Ana Tereza Vasconcelos. "BioinfoPortal: A scientific gateway for integrating bioinformatics applications on the Brazilian national high-performance computing network." *Future Generation Computer Systems* 107 (2020): 192-214.

[9]. Ius, Tamara, Alessandro Tel, Giuseppe Minniti, Teresa Somma, Domenico Solari, Michele Longhi, Pasquale De Bonis et al. "Advances in Multidisciplinary Management of Skull Base Meningiomas." *Cancers* 13, no. 11 (2021): 2664.

[10]. Rovcanin, Branislav, JasnaJancic, Janko Samardzic, MarijaRovcanin, Blazo Nikolic, Nikola Ivancevic, Ivana Novakovic, and Vladimir Kostic. "In silico model of mtDNA mutations effect on secondary and 3D structure of mitochondrial rRNA and tRNA in Leber's hereditary optic neuropathy." *Experimental Eye Research* 201 (2020): 108277.

[11]. Field, Matt A. "Detecting pathogenic variants in autoimmune diseases using high-throughput sequencing." *Immunology and Cell Biology* 99, no. 2 (2021): 146-156.

[12]. Alfano, Niccolò, Anisha Dayaram, Jan Axtner, Kyriakos Tsangaras, Marie-Louise Kampmann, Azlan Mohamed, Seth T. Wong, M. Thomas P. Gilbert, Andreas Wilting, and Alex D. Greenwood. "Non-invasive surveys of mammalian viruses using environmental DNA." *Methods in Ecology and Evolution* 12, no. 10 (2021): 1941-1952.

[13]. Shoji, Kan, and Ryuji Kawano. "Recent Advances in Liposome-Based Molecular Robots." *Micromachines* 11, no. 9 (2020): 788.

[14]. Chen, Zhen, Pei Zhao, Fuyi Li, Tatiana T. Marquez-Lago, André Leier, Jerico Revote, Yan Zhu et al. "iLearn: an integrated platform and meta-learner for feature engineering, machine-learning analysis and modeling of DNA, RNA and protein sequence data." *Briefings in bioinformatics* 21, no. 3 (2020): 1047-1057.

[15]. Xin-gang, Zhao, Liang Ji, Meng Jin, and Zhou Ying. "An improved quantum particle swarm optimization algorithm for environmental economic dispatch." *Expert Systems with Applications* 152

(2020): 113370.
[16]. Elfiky, Abdo A. "Ribavirin, Remdesivir, Sofosbuvir, Galidesivir, and Tenofovir against SARS-CoV-2 RNA dependent RNA polymerase (RdRp): A molecular docking study." *Life sciences* 253 (2020):
117592.
[17]. Donato, Luigi, Concetta Scimone, Simona Alibrandi, Alessandro Pitruzzella, Federica Scalia, Rosalia
D'Angelo, and Antonina Sidoti. "Possible A2E mutagenic effects on RPE mitochondrial DNA from
innovative RNA-seq bioinformatics pipeline." *Antioxidants* 9, no. 11 (2020): 1158.
[18]. Teng, Sin Yong, Guo Yong Yew, KateřinaSukačová, Pau Loke Show, VítězslavMáša, and Jo-Shu
Chang. "Microalgae with artificial intelligence: A digitalized perspective on genetics, systems and
products." *Biotechnology advances* (2020): 107631.
[19]. Selvaraj, Dhivya, and Sathishkumar Ramalingam. "Identification of morphologically similar species
of Tribulus (Zygophyllaceae) by employing DNA barcodes and rRNA secondary
structures." *Ecological Genetics and Genomics* 18 (2021): 100072.
[20]. Zhang, Canfeng, Liping Chen, Di Peng, Ao Jiang, Yunru He, Yanru Zeng, Chen Xie et al. "METTL3
and N6-methyladenosine promote homologous recombination-mediated repair of DSBs by
modulating DNA-RNA hybrid accumulation." *Molecular Cell* 79, no. 3 (2020): 425-442.
[21]. Cao, Yunlong, Bin Su, Xianghua Guo, Wenjie Sun, Yongqiang Deng, Linlin Bao, Qinyu Zhu et al.
"Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell
sequencing of convalescent patients' B cells." *Cell* 182, no. 1 (2020): 73-84.
[22]. Wang, Chao, Jin Ming Koh, Tiantang Yu, Neng Gang Xie, and Kang Hao Cheong. "Material and
shape optimization of bi-directional functionally graded plates by GIGA and an improved multi-
objective particle swarm optimization algorithm." *Computer Methods in Applied Mechanics and
Engineering* 366 (2020): 113017.
[23].Tran-Ngoc, H., Leqia He, Edwin Reynders, Samir Khatir, T. Le-Xuan, Guido De Roeck, T. Bui-Tien, and M. Abdel Wahab. "An efficient approach to model updating for a multispan railway bridge using orthogonal diagonalization combined with improved particle swarm optimization." *Journal of Sound and Vibration* 476 (2020): 115315.