

Estimasi Parameter Super Pairwise Alignment pada Kombinasi Virus Dengue Menggunakan Particle Swarm Optimization

Estimation of Super Pairwise Alignment Parameters on Combination of Dengue Virus Using Particle Swarm Optimization

Dinita Rahmalia¹, Arif Rohmatullah², M. Syaiful Pradana³

^{1,2,3}Program Studi Matematika, Universitas Islam Darul Ulum Lamongan

E-mail: dinitarahmalia@gmail.com, arifrohmatullah@unisda.ac.id, syaifulp@unisda.ac.id

Abstrak

Di Indonesia terdapat empat jenis virus dengue atau demam berdarah. Untuk melihat tingkat kesamaan (similarity) antara dua sekuens virus, dibutuhkan proses pensejajaran pada sekuens virus. Metode yang digunakan untuk pensejajaran pada dua sekuens virus adalah Super Pairwise Alignment (SPA). Nilai fungsi objective pada SPA adalah nilai penalty antara dua sekuens virus. Karena nilai fungsi objective tergantung pada parameter SPA, maka pada penelitian ini nilai parameter SPA akan diestimasi menggunakan metode heuristik seperti Particle Swarm Optimization (PSO). Simulasi diterapkan pada enam kombinasi virus dengue untuk proses estimasi parameter SPA. Berdasarkan hasil simulasi pada enam kombinasi virus dengue, PSO dapat menemukan parameter SPA yang optimal secara pendekatan. Parameter SPA yang optimal juga dapat mengetahui posisi dan panjang dari unit sekuens yang mengalami penambahan atau penghapusan.

Kata kunci: Estimasi Parameter, Super Pairwise Alignment, Pensejajaran Sekuens, Particle Swarm Optimization

Abstract

In Indonesia, there are four dengue virus types. To see the similarity between two virus sequences, it is required alignment process to virus sequences. The method used for aligning on two virus sequences is Super Pairwise Alignment (SPA). Objective function on SPA is the penalty score between two virus sequences. Because the value of objective function depends on SPA parameters, then in this research, SPA parameters will be estimated by heuristic methods such as Particle Swarm Optimization (PSO). Simulations are applied in six combinations of dengue virus for estimation SPA parameters. Based on simulation results on six combinations of dengue virus, PSO can find optimal SPA parameters in approaching. Optimal SPA parameters also can give positions and lengths of sequences unit occurring insertions or deletions.

Keywords: Parameter Estimation, Super Pairwise Alignment, Sequence Alignment, Particle Swarm Optimization

1. PENDAHULUAN

Penyakit dengue atau demam berdarah adalah penyakit yang disebabkan oleh virus dengue dan ditularkan melalui nyamuk *Aedes aegypti*. Gejala penyakit muncul selama 3-14 hari setelah terinfeksi seperti demam tinggi, sakit kepala, muntah, nyeri otot dan sendi, dan karakteristik ruam kulit. Penyakit ini tersebar terutama pada daerah iklim tropis termasuk Indonesia. Pada penelitian sebelumnya, di Indonesia terdapat empat jenis virus dengue atau demam berdarah [1]. Untuk melihat tingkat kesamaan (similarity) antara dua sekuens virus,

dibutuhkan proses penjejajaran pada sekuens virus [2]. Metode yang digunakan untuk penjejajaran pada dua sekuens virus adalah Super Pairwise Alignment (SPA).

Pada masalah penjejajaran sekuens menggunakan SPA, terdapat parameter SPA yang ditentukan secara trial and error. Nilai fungsi objective pada SPA adalah nilai penalty antara dua sekuens virus. Pada fungsi objective, semakin kecil nilainya maka tingkat kesamaan (similarity) antara dua sekuens virus semakin besar. Karena nilai fungsi objective tergantung pada parameter SPA, maka pada penelitian ini nilai parameter SPA akan diestimasi menggunakan metode heuristik seperti Particle Swarm Optimization (PSO).

Particle Swarm Optimization (PSO) ditemukan oleh Kennedy dan Eberhart pada tahun 1995. PSO adalah metode optimisasi yang terinspirasi dari perilaku populasi ikan atau unggas dalam mencari sumber makanan dimana suatu individu disebut partikel dan populasi disebut swarm [3]. Algoritma PSO telah banyak diaplikasikan antara lain : optimasi dengan kendala [4], optimisasi pada masalah distribusi [5], estimasi parameter pada masalah kontrol [6], optimisasi pada peramalan [7].

Simulasi diterapkan pada enam kombinasi virus dengue untuk proses estimasi parameter SPA yaitu : dengue virus 1 terhadap dengue virus 2, dengue virus 1 terhadap dengue virus 3, dengue virus 1 terhadap dengue virus 4, dengue virus 2 terhadap dengue virus 3, dengue virus 2 terhadap dengue virus 4, dan dengue virus 3 terhadap dengue virus 4. Berdasarkan hasil simulasi pada enam kombinasi virus dengue, PSO dapat menemukan parameter SPA yang optimal secara pendekatan. Parameter SPA yang optimal juga dapat mengetahui posisi dan panjang dari unit sekuens yang mengalami penambahan atau penghapusan.

2. SEKUENS BIOLOGIS

2.1 DNA, RNA, dan Sekuens Protein

Pada sekuens biologis, terdapat istilah : DNA, RNA, dan sekuens protein. DNA (*deoxyribonucleic acid*) sekuens memiliki empat huruf alfabet $\{A, C, G, T\}$, dimana A, C, G, T adalah : *nucleic acid adenine, cytosine, guanine, dan thymine*. Sekuens DNA terdiri dari dua untaian nukleotida ke dalam bentuk double helix. Protein-coding gen digunakan untuk menghasilkan protein dimana polimer dari 20 asam amino dihubungkan oleh ikatan peptida. Sekuens protein memiliki dua puluh huruf alfabet $\{A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y\}$. RNA adalah untaian tunggal dan mirip dengan dengan DNA. Sekuens RNA memiliki empat huruf alfabet $\{A, C, G, U\}$ dengan the U adalah *nucleic acid uracil* [8].

Sekuens DNA (atau RNA) memiliki struktur nukleotida sebagai berikut :

$$A = (a_1, a_2, \dots, a_{n_a}), B = (b_1, b_2, \dots, b_{n_b}), C = (c_1, c_2, \dots, c_{n_c})$$

dimana A, B, C adalah sekuens, dan a_i, b_i, c_i adalah tiap unit dari sekuens di posisi i dimana elemennya diperoleh dari himpunan $V_q = \{0, 1, \dots, q-1\}$.

$q = 4$ dan $V_4 = \{A, C, G, T\}$ jika A, B, C adalah DNA

$q = 4$ dan $V_4 = \{A, C, G, U\}$ jika A, B, C adalah RNA

$q = 20$ dan $V_{20} = \{A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y\}$ jika A, B, C adalah sekuens protein

2.2 Pensejajaran Sekuens

Misalkan terdapat dua sekuens yang saling berhubungan karena berasal dari moyang yang sama selama proses evolusi, maka dua sekuens tersebut bersifat homolog. Dua sekuens disebut homolog jika terdapat hubungan tingkat kesamaan. Langkah awal untuk mendeteksi homolog adalah menemukan tingkat kesamaan antar sekuens dengan cara mensejajarkannya.

Contoh : Terdapat dua sekuens, masing-masing memiliki tujuh unit

$X : TACCAGT$

$Y : CCCGTAA$

Jika terdapat gap, terdapat banyak pensejajaran yang mungkin seperti pada (1) dan (2) sehingga akan dipilih pensejajaran yang terbaik [8].

Kasus 1 :

$$\begin{aligned} X &: TACCAGT \text{ --} \\ Y &: C \text{ --} CC \text{ --} GTAA \end{aligned} \quad (1)$$

Kasus 2 :

$$\begin{aligned} X &: TACCAGT \text{ --} \\ Y &: \text{ --} \text{ --} CCCGTAA \end{aligned} \quad (2)$$

2.3 Proses Mutasi

Secara biologis, mutasi dari sekuens A dapat mengakibatkan sekuens A berubah menjadi sekuens B. Sekuens B disebut sebagai mutasi dari sekuens A. Mutasi sekuens dapat dibagi ke dalam empat jenis yaitu [9] :

Tipe I : Mutasi yang disebabkan oleh nukleotida yang berubah dari bentuk satu ke bentuk yang lain, contoh : “a” berubah menjadi “g”.

Tipe II : Mutasi yang disebabkan oleh segmen nukleotida mengalami permutasi (berubah posisi), contoh : segmen “accgu” berubah menjadi “guacc”.

Tipe III : Mutasi yang disebabkan adanya penambahan segmen ke dalam sekuens, contoh : penyisipan “aa” ke dalam “ggugg” sehingga menjadi segmen baru “gguaaugg”.

Tipe IV : Mutasi yang disebabkan adanya penghapusan segmen ke dalam sekuens, contoh : penghapusan “ag” dari “acaguua”, sehingga menjadi segmen baru “acuua”.

Mutasi tipe I dan tipe II tidak mengubah panjang sekuens sedangkan mutasi tipe III dan tipe IV mengubah panjang sekuens sehingga mutasi tipe III dan tipe IV adalah mutasi yang dapat bergeser.

2.4 Struktur Pergeseran pada Sekuens

Jika sekuens B adalah mutasi dari sekuens A melalui mutasi tipe III dan tipe IV, seperti:

$$A = (a_1, \dots, a_{n_a}) \xrightarrow{III} C = (c_1, \dots, c_{n_c}) \xrightarrow{IV} B = (b_1, \dots, b_{n_b})$$

Dengan menggunakan fungsi mutasi T , dan misalkan $A \xrightarrow{T} B$, maka fungsi mutasi T adalah :

$$T = \{(i_k, l_k), k = 1, 2, \dots, k_a\} \quad (3)$$

Nilai k_a di T adalah jumlah pergeseran mutasi pada sekuens A. Notasi (i_k, l_k) adalah posisi dan panjang dari pergeseran mutasi ke-k, dimana i_k adalah posisi mutasi pergeseran, dan l_k adalah informasi jenis dan panjang dari pergeseran mutasi ke-k. Mutasi tipe III terjadi jika $l_k > 0$, dan mutasi tipe IV terjadi jika $l_k < 0$. $|l_k|$ adalah panjang dari segmen yang disisipkan atau dihapus.

3. METODE PENELITIAN

3.1 Perbandingan pada Dua Sekuens

Masalah pensejajaran sekuens adalah menemukan bentuk pensejajaran sehingga semua sekuens memiliki tingkat perbedaan (difference) yang minimum atau tingkat kesamaan (similarity) yang maksimum. Secara umum tingkat perbedaan dinotasikan oleh matriks penalti yang menjelaskan tingkat perbedaan dari masing-masing sekuens. Bentuk dari matriks penalti adalah sebagai berikut :

$$W = [w(a,b)]_{a,b \in V_5} \quad (4)$$

Matriks penalti juga digunakan untuk perhitungan hasil pensejajaran sekuens. Elemen pada matriks penalti adalah :

$$w(a,b) = \begin{cases} 0, & \text{if } a = b \in V_5 \\ 1, & \text{yang lainnya} \end{cases} \quad (5)$$

Misalkan dua sekuens A dan B pada V_5

$$A = (a_1, a_2, \dots, a_n) \text{ dan } B = (b_1, b_2, \dots, b_n)$$

Fungsi optimisasi dihitung menggunakan persamaan (6).

$$w(A,B) = \sum_{j=1}^n w(a_j, b_j) \quad (6)$$

Fungsi korelasi lokal dari sekuens A dan sekuens B berdasarkan matriks penalti dihitung menggunakan persamaan (7).

$$w(A,B,i,j,n) = \sum_{k=1}^n w(a_{i+k}, b_{j+k}), \quad i+n \leq N_A, \quad i+n \leq N_B \quad (7)$$

Fungsi Sliding Window dari sekuens A dan sekuens B adalah :

$$w(A,B,i,j,n) = \frac{1}{n} \sum_{k=1}^n w(a_{i+k}, b_{j+k}) \quad (8)$$

3.2 Algoritma Super Pairwise Alignment (SPA)

Algoritma SPA digunakan untuk melihat tingkat kesamaan (similarity) antara dua sekuens. Misalkan (A,B) adalah dua sekuens dengan sekuens B adalah mutasi dari sekuens A. Parameter SPA yang digunakan adalah n, θ, θ' dengan $n > 0, 0.3 \leq \theta \leq 0.4$, dan $\theta < \theta' \leq 1$. Langkah-langkah SPA adalah sebagai berikut [9] :

1. Estimasi posisi mutasi pertama \hat{i}_1 di T

Inisialisasi $i = j = 0$ dan hitung $w(A,B,i,j,n)$.

- a. If ($w(A,B,i,j,n) \geq \theta'$)

$$\hat{i}_1 = 0$$

End

- b. If ($\theta < w(A,B,i,j,n) < \theta'$)

Misalkan $i \leftarrow i+h, j \leftarrow j+h$

Hitung $w(A,B,i,j,n)$.

Jika ($\theta < w(A,B,i,j,n) < \theta'$), ulangi langkah 1(b) dengan mengubah

$i \leftarrow i+h, j \leftarrow j+h$ sampai ($w(A,B,i,j,n) > \theta'$), kemudian simpan nilai i dan

j sebagai \hat{i}_1

End

c. If $(w(A, B, i, j, n) \leq \theta)$

Misalkan $i = j = k_1(n - \tau)$ dengan k_1 adalah bilangan bulat

Hitung $w(A, B, i, j, n)$.

Jika $(w(A, B, i, j, n) \leq \theta)$, ulangi langkah 1(c) dengan mengubah $k \leftarrow k + 1$ sampai $(w(A, B, i, j, n) > \theta)$ dan $(w(A, B, i, j, n) > \theta')$, kemudian simpan nilai i dan j sebagai \hat{i}_1

End

2. Estimasi \hat{l}_1 berdasarkan nilai estimasi \hat{i}_1 pada posisi mutasi pertama di T .

$$w(A, B, \hat{i}_1 + l, \hat{i}_1, n), \quad w(A, B, \hat{i}_1, \hat{i}_1 + l, n), \quad l = 1, 2, 3, \dots, l_{\max}$$

a. Jika $(w(A, B, \hat{i}_1 + l, \hat{i}_1, n) \leq \theta)$, maka $\hat{l}_1 = -l$ dan sisipkan l virtual symbol (gap) ke dalam sekuens B dari posisi \hat{i}_1 , sementara sekuens A tidak berubah.

b. Jika $(w(A, B, \hat{i}_1, \hat{i}_1 + l, n) \leq \theta)$, maka $\hat{l}_1 = l$ dan sisipkan l virtual symbols (gap) ke dalam sekuens A dari posisi \hat{i}_1 , sementara sekuens B tidak berubah.

Melalui kedua langkah ini, diperoleh estimasi mutasi lokal $T_1 = \{(i_1, l_1)\}$, dan pensejajaran lokal (C_1, D_1) sebagai berikut :

$$C_1 = (C_{1,1}, A_{2,1})$$

$$D_1 = (D_{1,1}, B_{2,1})$$

Misalkan panjang $C_{1,1}$ dan $D_{1,1}$ adalah $|\hat{i}_1 + |\hat{l}_1||$. Ketika tidak ada mutasi pergeseran yang terjadi pada posisi n yang pertama dari $A_{2,1}$, $B_{2,1}$, maka $L_1 = \hat{i}_1 + |\hat{l}_1| + n$ adalah titik awal untuk pensejajaran berikutnya.

3. Setelah memperoleh estimasi (\hat{i}_1, \hat{l}_1) , lanjutkan mengestimasi \hat{i}_2 berdasarkan (C_1, D_1) . Inisialisasi $i = j = L_1$ dan hitung $w(A, B, i, j, n)$ dengan mengulangi langkah 1.

4. Estimasi \hat{l}_2 berdasarkan estimasi $\hat{i}_1, \hat{l}_1, \hat{i}_2$.

$$w(A, B, \hat{i}_1 + l, \hat{i}_1, n), \quad w(A, B, \hat{i}_1, \hat{i}_1 + l, n), \quad l = 1, 2, 3, \dots, l_{\max}$$

Ulangi langkah 2 untuk mengestimasi \hat{l}_2 dan pensejajaran lokal (C_2, D_2)

5. Ulangi langkah 1-4 di atas sampai sekuens (\hat{i}_k, \hat{l}_k) dan pensejajaran lokal (C_k, D_k) untuk semua $k = 1, 2, 3, \dots$. Proses berakhir sampai k_0 sehingga $C_{k_0} = (C_{1,k_0}, A_{2,k_0})$ and $D_{k_0} = (D_{1,k_0}, B_{2,k_0})$ memiliki mutasi pergeseran yang terjadi di (A_{2,k_0}, B_{2,k_0}) .

6. Hitung fungsi objective pada persamaan (9).

$$w(C_{k_0}, D_{k_0}) = \sum_{j=1}^n w(c_j, d_j) \quad (9)$$

Dengan nilai elemen adalah :

$$w(c_j, d_j) = \begin{cases} 0, & \text{if } c_j = d_j \\ 1, & \text{yang lainnya} \end{cases} \quad (10)$$

Berdasarkan persamaan (9), semakin kecil nilai $W(C_{k_0}, D_{k_0})$ maka semakin besar tingkat kesamaan (similarity) pada dua sekuens.

3.3 Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) ditemukan oleh Kennedy dan Eberhart pada tahun 1995. PSO adalah metode optimisasi yang terinspirasi dari perilaku populasi ikan atau unggas

dalam mencari sumber makanan dimana suatu individu disebut partikel dan populasi disebut swarm [3].

PSO diinisialisasi dengan sekumpulan partikel sebagai kandidat solusi pada posisi acak. Setiap partikel diberikan posisi awal dan kecepatan awal. Ketika suatu partikel menemukan arah sumber makanan, partikel lainnya akan mengikuti menuju sumber makanan [3].

Pada estimasi parameter SPA menggunakan PSO, representasi partikel yang digunakan adalah parameter SPA yaitu $X = (n, \theta, \theta')$ dengan $n > 0$, $0.3 \leq \theta \leq 0.4$, dan $\theta < \theta' \leq 1$ dan fungsi objective adalah persamaan (9). Pada fungsi objective, semakin kecil nilainya maka tingkat kesamaan (similarity) antara dua sekuens semakin besar. Langkah-langkah algoritma PSO adalah sebagai berikut [10] :

1. Inisialisasi posisi partikel $X^s(0)$, $s = 1, 2, \dots, \text{max_swarm}$
2. Inisialisasi kecepatan partikel $V^s(0)$, $s = 1, 2, \dots, \text{max_swarm}$
3. Tentukan local best particle

$$Pbest^s = X^s(0), s = 1, 2, \dots, \text{max_swarm} \quad (11)$$

4. Hitung fitness function $f(Pbest^s)$, $s = 1, 2, \dots, \text{max_swarm}$ beserta nilai penalti $p(X^s(0))$
5. Tentukan global best particle

$$Gbest = \arg \min_{Pbest} (f(Pbest^s), s = 1, 2, \dots, \text{max_swarm}) \quad (12)$$

6. Update posisi dan kecepatan partikel

for $t = 0 : \text{max_t} - 1$

for $s = 1 : \text{max_swarm}$

- a. Hitung kecepatan partikel

$$V^s(t+1) = \omega V^s(t) + c_1 r_1 (Pbest^s - X^s(t)) + c_2 r_2 (Gbest - X^s(t)) \quad (13)$$

Dengan nilai ω adalah inersia yang bernilai antara 0.9-1.2, nilai $c_1 = c_2 = 2$, sedangkan nilai r_1 dan r_2 adalah bilangan acak antara (0,1).

- b. Update posisi partikel

$$X^s(t+1) = X^s(t) + V^s(t+1) \quad (14)$$

- c. Hitung fitness function posisi partikel $f(X^s(t+1))$ beserta nilai penalti $p(X^s(t+1))$

- d. Tentukan local best particle

$$Pbest^s = \arg \min_X (f(X^s(0)), f(X^s(1)), \dots, f(X^s(t+1))) \quad (15)$$

end

- e. Tentukan global best particle

$$Gbest = \arg \min_{Pbest} (f(Pbest^s), s = 1, 2, \dots, \text{max_swarm}) \quad (16)$$

End

4. HASIL DAN PEMBAHASAN

Berdasarkan penelitian sebelumnya [6], di Indonesia terdapat empat jenis virus dengue. Dari keempat virus tersebut, akan dibentuk enam macam kombinasi untuk proses estimasi parameter SPA, yaitu : dengue virus 1 terhadap dengue virus 2, dengue virus 1 terhadap dengue virus 3, dengue virus 1 terhadap dengue virus 4, dengue virus 2 terhadap dengue virus 3, dengue virus 2 terhadap dengue virus 4, dan dengue virus 3 terhadap dengue virus 4.

Data yang digunakan dalam simulasi diperoleh dari National Center for Biotechnology Information (NCBI) yang diakses pada tanggal 2 Mei 2019. Data dari empat jenis sekuens protein virus dengue dapat dilihat pada Tabel 1.

Tabel 1 Data Sekuens Protein Virus Dengue

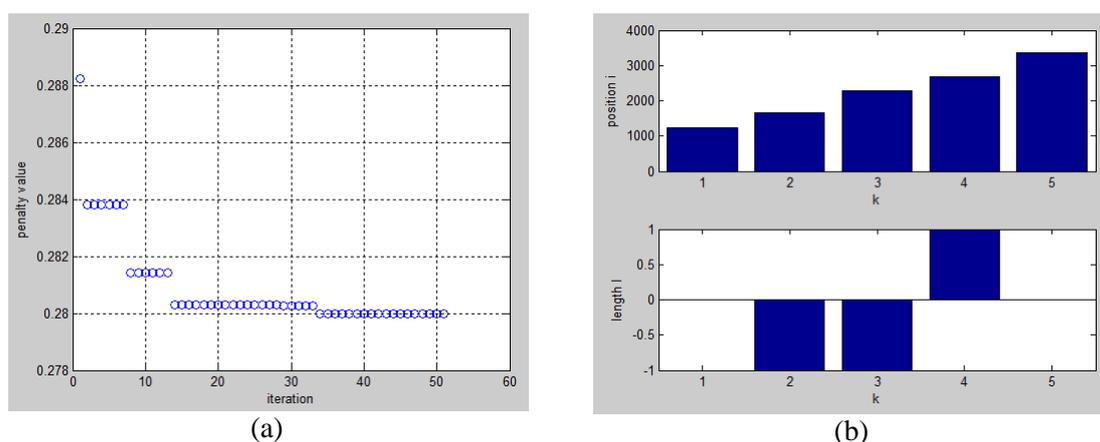
| Jenis | Kode Akses | Panjang (bp) |
|----------------|------------|--------------|
| Dengue Virus 1 | AHG06327 | 3392 |
| Dengue Virus 2 | AHG06364 | 3391 |
| Dengue Virus 3 | AHG06376 | 3390 |
| Dengue Virus 4 | AHG06382 | 3387 |

Setiap kombinasi pada dengue virus, parameter SPA akan diestimasi menggunakan algoritma PSO. Parameter PSO yang digunakan adalah : jumlah populasi adalah 10 dan iterasi maksimum adalah 50.

4.1 Hasil Simulasi pada Sekuens Protein Virus Dengue 1 terhadap Sekuens Protein Virus Dengue 2

Hasil simulasi PSO dapat dilihat pada Gambar 1(a). Pada awal iterasi, partikel memilih parameter SPA secara acak kemudian partikel diupdate sehingga menghasilkan pendekatan parameter SPA optimal yaitu nilai penalti sebagai nilai fitness adalah 0,2800 dengan parameter SPA optimal adalah $(n, \theta, \theta') = (39; 0,3603; 0,6724)$ dan waktu komputasi adalah 37,10 detik.

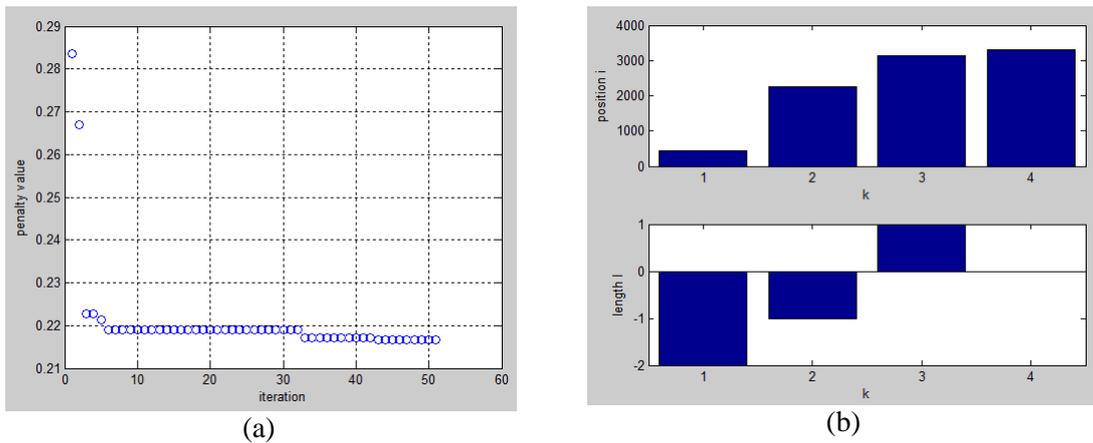
Gambar 1(b) adalah hasil dari parameter SPA yang optimal yaitu posisi i_k dan panjang l_k , dengan $k = 5$. Jika $l_k > 0$ maka terjadi penambahan pada panjang unit protein sekuens. Jika $l_k < 0$ maka terjadi penghapusan pada panjang unit protein sekuens. Jika $l_k = 0$ maka tidak ada perubahan pada unit protein sekuens.



Gambar 1 Perbandingan Sekuens Protein Virus Dengue 1 terhadap Sekuens Protein Virus Dengue 2. (a) Hasil Optimisasi PSO. (b) Hasil pada Parameter SPA Optimal

4.2 Hasil Simulasi pada Sekuens Protein Virus Dengue 1 terhadap Sekuens Protein Virus Dengue 3

Hasil simulasi PSO dapat dilihat pada Gambar 2(a). Pada awal iterasi, partikel memilih parameter SPA secara acak kemudian partikel diupdate sehingga menghasilkan pendekatan parameter SPA optimal yaitu nilai penalti sebagai nilai fitness adalah 0,2166 dengan parameter SPA optimal adalah $(n, \theta, \theta') = (87; 0,3747; 0,5765)$ dan waktu komputasi adalah 35,01 detik. Gambar 2(b) adalah hasil dari parameter SPA yang optimal yaitu posisi i_k dan panjang l_k , dengan $k = 4$. Jika $l_k > 0$ maka terjadi penambahan pada panjang unit protein sekuens. Jika $l_k < 0$ maka terjadi penghapusan pada panjang unit protein sekuens. Jika $l_k = 0$ maka tidak ada perubahan pada unit protein sekuens.

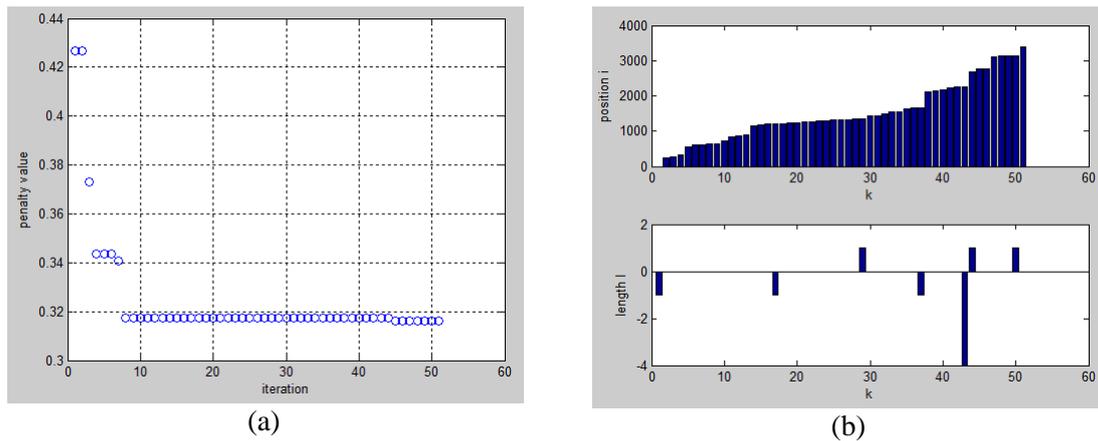


Gambar 2 Perbandingan Sekuens Protein Virus Dengue 1 terhadap Sekuens Protein Virus Dengue 3. (a) Hasil Optimisasi PSO. (b) Hasil pada Parameter SPA Optimal

4.3 Hasil Simulasi pada Sekuens Protein Virus Dengue 1 terhadap Sekuens Protein Virus Dengue 4

Hasil simulasi PSO dapat dilihat pada Gambar 3(a). Pada awal iterasi, partikel memilih parameter SPA secara acak kemudian partikel diupdate sehingga menghasilkan pendekatan parameter SPA optimal yaitu nilai penalti sebagai nilai fitness adalah 0,3164 dengan parameter SPA optimal adalah $(n, \theta, \theta') = (9; 0,3607; 0,6268)$ dan waktu komputasi adalah 42,77 detik.

Gambar 3(b) adalah hasil dari parameter SPA yang optimal yaitu posisi i_k dan panjang l_k , dengan $k = 51$. Jika $l_k > 0$ maka terjadi penambahan pada panjang unit protein sekuens. Jika $l_k < 0$ maka terjadi penghapusan pada panjang unit protein sekuens. Jika $l_k = 0$ maka tidak ada perubahan pada unit protein sekuens.

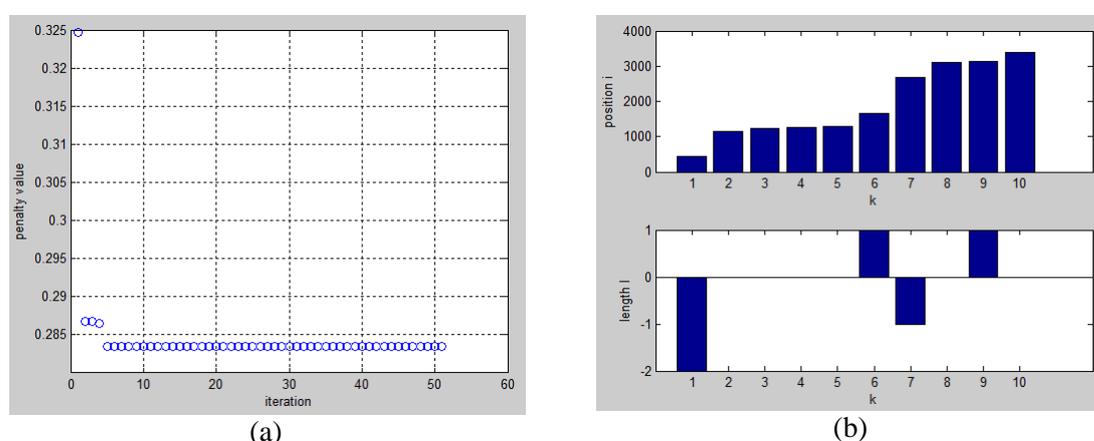


Gambar 3 Perbandingan Sekuens Protein Virus Dengue 1 terhadap Sekuens Protein Virus Dengue 4. (a) Hasil Optimisasi PSO. (b) Hasil pada Parameter SPA Optimal

4.4 Hasil Simulasi pada Sekuens Protein Virus Dengue 2 terhadap Sekuens Protein Virus Dengue 3

Hasil simulasi PSO dapat dilihat pada Gambar 4(a). Pada awal iterasi, partikel memilih parameter SPA secara acak kemudian partikel diupdate sehingga menghasilkan pendekatan parameter SPA optimal yaitu nilai penalti sebagai nilai fitness adalah 0,2835 dengan parameter SPA optimal adalah $(n, \theta, \theta') = (34; 0,3718; 0,6572)$ dan waktu komputasi adalah 33,53 detik.

Gambar 4(b) adalah hasil dari parameter SPA yang optimal yaitu posisi i_k dan panjang l_k , dengan $k = 10$. Jika $l_k > 0$ maka terjadi penambahan pada panjang unit protein sekuens. Jika $l_k < 0$ maka terjadi penghapusan pada panjang unit protein sekuens. Jika $l_k = 0$ maka tidak ada perubahan pada unit protein sekuens.

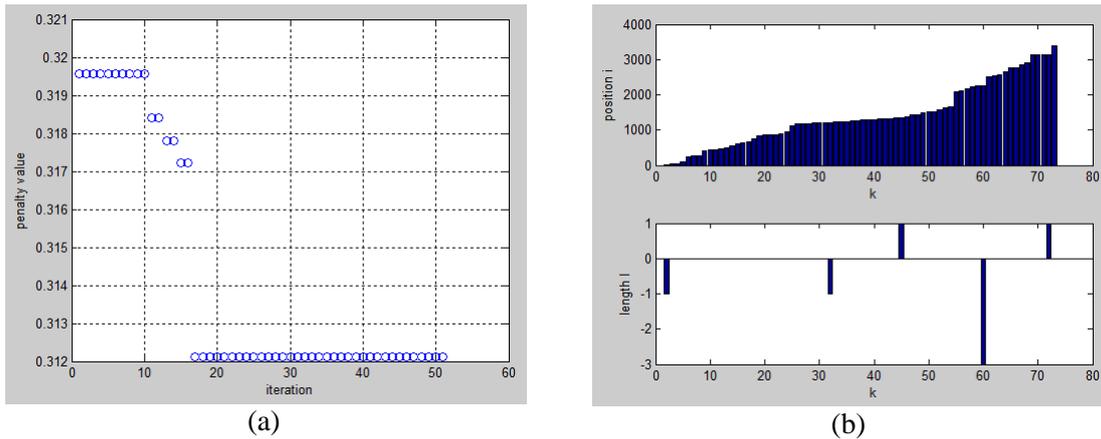


Gambar 4 Perbandingan Sekuens Protein Virus Dengue 2 terhadap Sekuens Protein Virus Dengue 3. (a) Hasil Optimisasi PSO. (b) Hasil pada Parameter SPA Optimal

4.5 Hasil Simulasi pada Sekuens Protein Virus Dengue 2 terhadap Sekuens Protein Virus Dengue 4

Hasil simulasi PSO dapat dilihat pada Gambar 5(a). Pada awal iterasi, partikel memilih parameter SPA secara acak kemudian partikel diupdate sehingga menghasilkan pendekatan parameter SPA optimal yaitu nilai penalti sebagai nilai fitness adalah 0,3121 dengan parameter SPA optimal adalah $(n, \theta, \theta') = (8; 0,3641; 0,5615)$ dan waktu komputasi adalah 35,15 detik.

Gambar 5(b) adalah hasil dari parameter SPA yang optimal yaitu posisi i_k dan panjang l_k , dengan $k = 73$. Jika $l_k > 0$ maka terjadi penambahan pada panjang unit protein sekuens. Jika $l_k < 0$ maka terjadi penghapusan pada panjang unit protein sekuens. Jika $l_k = 0$ maka tidak ada perubahan pada unit protein sekuens.

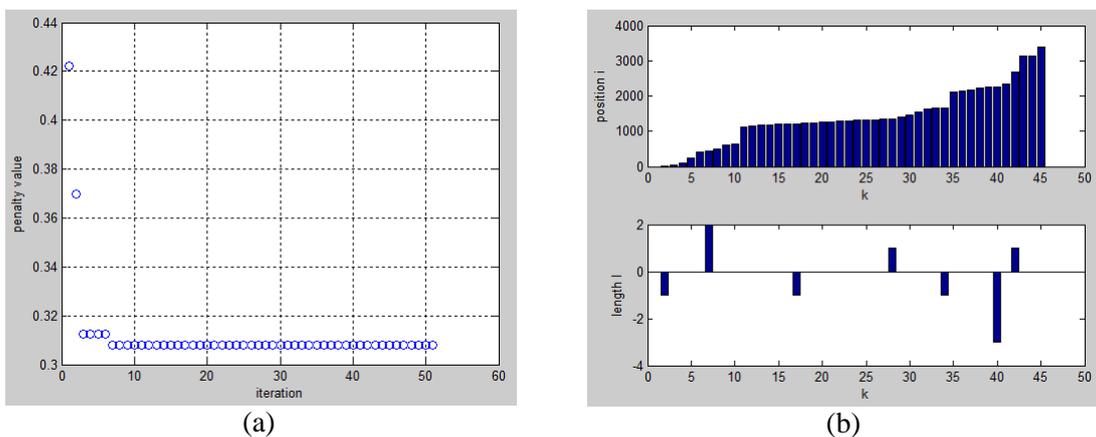


Gambar 5 Perbandingan Sekuens Protein Virus Dengue 2 terhadap Sekuens Protein Virus Dengue 4. (a) Hasil Optimisasi PSO. (b) Hasil pada Parameter SPA Optimal

4.6 Hasil Simulasi pada Sekuens Protein Virus Dengue 3 terhadap Sekuens Protein Virus Dengue 4

Hasil simulasi PSO dapat dilihat pada Gambar 6(a). Pada awal iterasi, partikel memilih parameter SPA secara acak kemudian partikel diupdate sehingga menghasilkan pendekatan parameter SPA optimal yaitu nilai penalti sebagai nilai fitness adalah 0,3079 dengan parameter SPA optimal adalah $(n, \theta, \theta') = (9; 0,3732; 0,6117)$ dan waktu komputasi adalah 48,36 detik.

Gambar 6(b) adalah hasil dari parameter SPA yang optimal yaitu posisi i_k dan panjang l_k , dengan $k = 45$. Jika $l_k > 0$ maka terjadi penambahan pada panjang unit protein sekuens. Jika $l_k < 0$ maka terjadi penghapusan pada panjang unit protein sekuens. Jika $l_k = 0$ maka tidak ada perubahan pada unit protein sekuens.



Gambar 6 Perbandingan Sekuens Protein Virus Dengue 3 terhadap Sekuens Protein Virus Dengue 4. (a) Hasil Optimisasi PSO. (b) Hasil pada Parameter SPA Optimal

5. KESIMPULAN DAN SARAN

Metode SPA dapat digunakan untuk menentukan tingkat kesamaan antara dua sekuens dengan cara mensejajarkan antara dua sekuens. Pada SPA, nilai dari penalty value tergantung dari parameter SPA. Pada penelitian ini, parameter SPA dioptimisasi menggunakan PSO.

Berdasarkan hasil simulasi pada enam kombinasi virus dengue, PSO dapat menemukan parameter SPA yang optimal secara pendekatan. Parameter SPA yang optimal juga dapat mengetahui posisi dan panjang dari unit sekuens yang mengalami penambahan atau penghapusan. Penelitian lebih lanjut dapat digunakan pada penentuan tingkat kesamaan pada multiple sekuens.

UCAPAN TERIMA KASIH

Terima kasih disampaikan pada Kementerian Riset Teknologi dan Pendidikan Tinggi atas dana Penelitian Dosen Pemula tahun 2019.

DAFTAR PUSTAKA

- [1] Pradana, M.S., Amiroch, S., 2019, Protein Sekuens Analysis of the Zika Virus and the Dengue Virus Using Smith Waterman Algorithm, *AIP Conference Proceeding*, vol. 2084, Depok, Agustus.
- [2] Shahab, M.L, Irawan, M.I., 2017, Sekuens Alignment Using Nature-Inspired Metaheuristic Algorithm, *International Journal of Computing Science and Applied Mathematics*, vol 3, hal 27-31.
- [3] Kennedy, J., Eberhart, R.C., 1995, Particle Swarm Optimization, *Proceedings IEEE Int. Conf. Neural Network*
- [4] Rahmalia, D., 2018, Teknik Penalti pada Optimisasi Berkendala Menggunakan Particle Swarm Optimization, *JMPM : Jurnal Matematika dan Pendidikan Matematika*, vol 3, hal 44-52.
- [5] Rahmalia, D., 2017, Particle Swarm Optimization-Genetic Algorithm (PSOGA) on Linear Transportation Problem, vol 1867, *AIP Conference Proceeding*, Surabaya, November.
- [6] Herlambang, T., Rahmalia, D., Yulianto, T., 2019, Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) for Optimizing PID Parameters on Autonomous Underwater Vehicle (AUV) Control System, *Journal of Physics : Conference Series*, vol. 1211, Jember, November.
- [7] Rahmalia, D., Herlambang, T., 2017, Prediksi Cuaca Menggunakan Algoritma Particle Swarm Optimization-Neural Network (PSO-NN), *Prosiding Seminar Nasional Matematika dan Aplikasinya*, Oktober.
- [8] Isaev, A., 2004, *Introduction to Mathematical Methods in Bioinformatics*, Springer, Berlin.
- [9] Shen, S.N., Tuszynski, J.A., 2008, *Theory and Mathematical Model for Bioinformatics*, Springer, Berlin.
- [10] Rao, S.S., 2009, *Engineering Optimization Theory and Practice*, John Wiley and Sons, New Jersey.