Scheduling Diet for Diabetes Mellitus Patients using Genetic Algorithm

M F Syahputra¹, V Felicia¹, R F Rahmat¹, and R Budiarto²

¹Department of Information Technology, Faculty of Computer Science and Information Technology, University of Sumatera Utara, Medan, Indonesia
²Department of Information System, College of Computer Science and Information Technology, Albaha University, Saudi Arabia

Email: nca.fadly@usu.ac.id, vanesa_felicia92@students.usu.ac.id, romi.fadillah@usu.ac.id, rahmat@bu.edu.sa

Abstract. Diabetes Mellitus (DM) is one of metabolic diseases which affects on productivity and lowers the human resources quality. This disease can be controlled by maintaining and regulating balanced and healthy lifestyle especially for daily diet. However, nowadays, there is no system able to help DM patient to get any information of proper diet. Therefore, an approach is required to provide scheduling diet every day in a week with appropriate nutrition for DM patients to help them regulate their daily diet for healing this disease. In this research, we calculate the number of caloric needs using Harris-Benedict equation and propose genetic algorithm for scheduling diet for DM patient. The results show that the greater the number of individuals, the greater the more the possibility of changes in fitness score approaches the best fitness score. Moreover, the greater the created generation, the more the opportunities to obtain best individual with fitness score approaching 0 or equal to 0.

1. Introduction

Scheduling is resources allocation needed to do tasks in certain time to generate operations [1]. It is used to minimize required time so it can improve productivity. For every Diabetes Mellitus (DM) patient, scheduling diet, like regulating balanced and healthy diet, is required to heal the disease. Diabetes Mellitus is the symptoms caused by elevated levels of blood glucose due to a decrease in progressive insulin secretion motivated by the insulin resistance [2]. Based on Basic Health Research [3], around 1.5-2.3% of population aged > 15 years old in Indonesia who suffers Diabetes Mellitus II. According to World Health Organization (WHO), the number of DM patients will increase to 366 million people in 2025. It is caused by lifestyle changes, like unhealthy or irregular dietary habit, and lack of information in consuming food. Scheduling diet and consuming healthy food with appropriate nutrition certainly can prevent this disease from human body. Therefore, a scheduling diet system is required to help DM patients get much information about a proper diet in accordance with the caloric needs of patients in a day based on several parameters, such as age, sex, weight, height, level of activity and allergy.

Many researchers have done research about diet for diabetics, but they did not include complete required parameters for calculation of calorie needs in a day. They classified allowed food for diabetics using various method. In 2010, Phanich, et al. use Self-Organizing Map (SOM) and K-meand Clustering to classify allowed food for diabetics [4].
The next research in 2011, scheduling diet for cancer patient had been conducted by Husain, et al. using Genetic Algorithm (GA) [5]. They succeeded to schedule diet with high accuracy, but they limited the data of food and generated it only per day.

Previously, GA was often used for recommendation research in various kind of problem. Another related research using GA was done by Tias, et al. for determination of healthy food menu. From their testing result, they obtained healthy food menu solutions in generation 100 with the fitness score of 0.449 [6].

In 2010, GA was used for solving the optimization of function problem with encoding integer [7]. They concluded that the high number of population is not necessarily going to generate better fitness score. It is caused by the weakness of GA, probabilistic, always done randomly. Scheduling courses for students and lecturer in Department of Information System ITS was also done by Puspaningrum in 2014 using GA [8]. They reached 10 of best fitness score and 275 of generation number.

Another related research about DM and GA was done by Kushardiana in 2013, discussed about the composition of diet for DM patients. They succeeded to determine the diet, however they only used age parameter for determining the calorie needs in a day.

In this research, we propose genetic algorithm for scheduling diet for Diabetes Mellitus patients. The rest of the paper will be organized as follows: Section II describes the proposed method. Results and discussions are presented in Section III. Section IV provides summary and suggestions for future research.

2. Research methodology

In our proposed method, there are several steps for scheduling diet for diabetics, namely collecting data of food nutrition and patient data; calculation of calorie needs using Harris-Benedict equation; determination of diet type; implementation of genetic algorithm, including initializing population, calculating fitness score, selecting the best individuals, performing crossover, performing mutation to change the missing genes and generate new chromosome. The obtained output of this system is a schedule menu in a week for DM patients. Each of steps will be explained in the next subsection. Fig. 1 shows the general architecture of our proposed method.
2.1. Data
The used data in this research consists of two data types, namely DM patient data and food nutrition data. The data of patients is used to calculate calorie needs then used to determine the type of diet. The used parameters in this research are age, sex, weight, height, level of activity and allergy. The data of food nutrition is ingredients data in 100 grams obtained from Nutrisurvey Indonesia in 2007. This data is used to generate a menu in accordance with the calorie needs. The used parameters are food name, type of food (staple food, side dishes, vegetables, fruits, and complement), energy (kcal), protein (gr), fat (gr) and carbohydrate (gr).

2.2. Calorie Needs Calculation
Calorie needs can be calculated using Harris-Benedict equation which estimates resting energy expenditure (REE). The calculation of calorie needs is obtained from weight, height, sex, age and level of activity. These following steps are used to calculate the calorie needs.

**Step 1**: Calculate Ideal Body Weight (kg) based on rule in table 1.

### Table 1. Rule of Ideal Body Weight calculation.

| Sex     | Height    | Formula                |
|---------|-----------|------------------------|
| Male    | < 160 cm  | \((\text{Height} - 100) \times 1\,\text{kg}\) |
| Female  | < 150 cm  | \((\text{Height} - 100) \times 1\,\text{kg}\) |
| Male    | \(\geq 160\,\text{cm}\) | \((\text{Height} - 100) \times 1\,\text{kg}\) \times 90\% |
| Female  | \(\geq 150\,\text{cm}\) | \((\text{Height} - 100) \times 1\,\text{kg}\) \times 90\% |
Step 2: Calculate Basal Calorie (kcal) based on rule in table 2.

**Table 2. Rule of Basal Calorie calculation.**

| Sex       | Formula               |
|-----------|-----------------------|
| Male      | Ideal Body Weight x 30 kcal |
| Female    | Ideal Body Weight x 25 kcal |

Step 3: Calculate the subtraction for activity factor (kcal) based on rule in table 3.

**Table 3. Rule of subtraction for the activity calculation.**

| Category    | Activity                              | Formula       |
|-------------|---------------------------------------|---------------|
| Light       | teaching, reading, walking, working in the office, fishing, etc. | Basal Calorie x 10% |
| Medium      | cycling, bowling, gardening, house cleaning, etc. | Basal Calorie x 30% |
| Hard        | aerobic, jogging, dancing, etc.        | Basal Calorie x 40% |

Step 4: Calculate the subtraction for age factor (kcal) based on rule in table 4.

**Table 4. Rule of subtraction for the age calculation.**

| Age         | Formula       |
|-------------|---------------|
| 40 - 59     | Basal Calorie x 5% |
| 60 - 69     | Basal Calorie x 10% |
| > 70        | Basal Calorie x 20% |

Step 5: Calculate calorie needs (kcal) using equation (1)

\[
Calorie = Basal Calorie - (Activity + Age)
\]  

(1)

Step 6: Calculate Body Mass Index using equation (2)

\[
BMI = \frac{Weight}{Height^2} \times 100\%
\]  

(2)

Step 7: Calculate total calorie needs based on Body Mass Index based on rule in table 5.

**Table 5. Rule of total calorie needs calculation.**

| Body Weight Status | Body Mass Index | Formula       |
|--------------------|-----------------|---------------|
| Obesity            | BMI > 28        | Calorie – 700 kcal |
| Fat                | BMI > 25        | Calorie – 500 kcal |
| Normal             | 18 < BMI < 25   | Calorie       |
| Thin               | BMI < 18        | Calorie + 500 kcal |

2.3. Diet Type Determination for DM Patient

Diet for DM patients is classified into eight types. Table 6 shows the nutrition ingredients, such as energy, protein, fat, carbohydrate, cholesterol, fiber, sodium and sucrose, for each type [10].
Table 6. Diet type classification.

| Diet Type | Energy (kcal) | Protein (g) | Fat (g) | Carbohydrate (g) | Cholesterol (mg) | Fiber (g) | Sodium (mg) | Sucrose (mg) |
|-----------|---------------|-------------|---------|------------------|------------------|-----------|-------------|--------------|
| I         | 1100          | 43          | 30      | 172              | 303              | 37        | < 2300      | < 5          |
| II        | 1300          | 45          | 35      | 192              | 303              | 37        | < 2300      | < 5          |
| III       | 1500          | 51.5        | 36.5    | 235              | 303              | 37        | < 2300      | < 5          |
| IV        | 1700          | 55.5        | 36.5    | 275              | 303              | 37        | < 2300      | < 5          |
| V         | 1900          | 60          | 48      | 299              | 303              | 37        | < 2300      | < 5          |
| VI        | 2100          | 62          | 53      | 319              | 303              | 37        | < 2300      | < 5          |
| VII       | 2300          | 73          | 59      | 369              | 303              | 37        | < 2300      | < 5          |
| VIII      | 2500          | 80          | 62      | 396              | 303              | 37        | < 2300      | < 5          |

2.4. Genetic Algorithm Model

2.4.1. Genetic Algorithm Structure. Genetic algorithm is based on an analogy with the genetic structure, which consists of three main component, namely individuals, chromosome and genes. In this research, there are three chromosomes in each of individual, namely breakfast, lunch and dinner. Each of chromosome has two values, namely fitness score and the total of calorie. One individual consists of 15 genes which contains the total of calorie for each kind of food, namely staple food, side dishes, vegetables, fruits, and complement. The value in each of gene is called by allele. Figure 2, figure 3 and figure 4 sequentially shows the structure of individual, chromosome and genes.

![Individual Structure](image)

**Figure 2. Individual Structure**

![Chromosome Structure](image)

**Figure 3. Chromosome Structure**

![Gen Structure](image)

**Figure 4. Gen Structure**

2.4.2. Initialization. Genetic algorithm always starts from randomly generating the initial population or individuals. The population size depends on the nature problem and types of implemented genetic operators. In this research, the individual is food menu from food database. The greater the size of individual, the more the generated diet variation. Several steps below are done to initialize the individual.

**Step 1.** Normalize food menu which aims to adjust the generated diet.

a. Divide the total of calorie needs

Total of calorie needs is divided into 3 chromosomes, namely breakfast, lunch and dinner. It aims to make the diet schedule consider the balance of total calorie for breakfast, lunch and dinner. It is done as formulated in equation (3).
\[
\text{Total calorie per chrom} = \frac{\text{Total of calorie needs}}{3}
\]

\[k = \text{fitness score for each individual in a week (1-7)}\]

b. Calculate total calorie for every kind of food

Each of chromosome consists of 5 genes of food type, namely staple food, side dishes, vegetables, fruits, and complement, which respectively has a weight of 45%, 20%, 15%, 15%, and 20%. It can be calculated using equation (4) until equation (8).

\[
\begin{align*}
\text{staple food} &= 45\% \times \text{total calorie per chrom} \\
\text{side dishes} &= 20\% \times \text{total calorie per chrom} \\
\text{vegetables} &= 15\% \times \text{total calorie per chrom} \\
\text{fruits} &= 15\% \times \text{total calorie per chrom} \\
\text{complement} &= 20\% \times \text{total calorie per chrom}
\end{align*}
\]

C. Calculate each food weight in 100 gram

Each of food in database is expressed in units of 100 grams. In the database, there are a few of menu that has calorie more than needed calorie for the food. To generate the number of calories and weight that corresponds to the total calorie needs of DM patient, we do the calculation as formulated in equation (9).

\[
\text{Food weight (gr)} = \frac{\text{Calorie needs}}{\text{Food calorie} \times 100}
\]

Step 2: Random and generate 15 normalized food menu as an individual.

Step 3: Input the calorie of food into chromosome. The generated food menu from previous step will be divided into 3 chromosomes, namely breakfast, lunch and dinner, so each of chromosome has 5 genes which contain the calorie of staple food, side dishes, vegetables, fruits, and complement.

2.4.3. Fitness Function Value Evaluation. Fitness score represents the ability of an individual assigned to generated solution. It also can be used as reference in achieving optimal value in genetic algorithm. In this research, GA is used to find an individual with the most minimum fitness score. Fitness score can be calculated using equation (10), while the average of fitness score for each generation can be calculated using equation (11).

\[
\text{Fitness score} = \left[ \text{total calorie needs} - \text{total food calorie in a day} \right]
\]

\[
\text{fitness average} = \frac{\sum_{k=1}^{7}(a_k)}{7}
\]

where:

- \(a_k\) = fitness score for each individual
- \(k\) = individual in a week (1-7)

2.4.4. Selection. Selection aims to give preference to better individual, allowing them to pass on their genes to the next generation. The goodness of each individual depends on its fitness score. The higher the fitness score of individual, the higher the possibility of selecting the individual. In this research, we use Rank Based Fitness Assignment method to do selection. It sorts the individual based on the fitness score ranged from lowest to highest. It serves to make each individual have the possibility to be selected. The illustration of Rank Based Fitness Assignment method is shown as figure 5.
2.4.5. **Crossover**. After selecting the individuals, then it will be used to do crossover. We set 0.1 as the percentage of crossover to determine how many parents are selected to do crossover. After that, determination of gen cut point is done randomly using two point crossover method. It almost same as one point crossover method, the difference is that there are two cut points randomly generated in this method [11]. These following steps will describe the two point crossover method.

**Step 1**: Input the parent or selected individuals to do crossover

**Step 2**: Determine two random number of cut point along the bit number of chromosome in each individual

**Step 3**: For each gen in parents chromosome, do:

   a. For the initial index is smaller than the final index, do:
      • For genes located before initial index and after the final index, copy gen value from first parent to be given to the next generation (offspring).
      • For genes located between the initial and the final index, genes in offspring will be genes from the second parent.
   b. For the initial index is higher than the final index, do:
      • Crossover starts from the initial index to the final index
      • For genes located before initial index and after the final index, copy gen value from first parent to be given to the next generation (offspring).
      • For genes located between the initial and the final index, genes in offspring will be genes from the second parent

2.4.6. **Mutation**

Mutation, done after doing crossover, aims to change the new offspring by exchange 0 to 1 and 1 to 0. It purposes to maintain diversity within the population and prevent the local optima in search space. In this research, we use random method to implement the mutation. Before doing mutation, the number of genes, which will do mutation, can be calculated based on Probability of Mutation (Pm) and number of gen as formulated in equation (12).

\[
Pm \times total\ population \times total\ gen
\]

Equation (12) will generate the number of gen which will do mutation, and then its position will be randomly determined. The selected gen will perform mutation with the gen which has the same type in the next chromosome. These following steps will describe the process of mutation.

**Step 1**: Determine the number of gen to do mutation

**Step 2**: Random a number as gen position from 0 until the number of chromosome.

**Step 3**: Perform mutation by replacing the gen value into the value of gen in selected index

   a. If the selected index is on chromosome 1 (one) or 2 (two), then the value is replaced with the value of gen that has the same type in the next chromosome.
   b. If the selected index is on chromosome 3 (three), then the value is replaced with the value of gen that has the same type in the previous chromosome.
GA will be repeated in accordance with the number of individual and generation. The result from this method is schedule of diet from 7 (seven) better individuals with lower fitness score. The schedule will be displayed in a week (7 days) for breakfast, lunch and dinner menu. Each of schedule consists of five types of food, namely staple food, side dishes, vegetables, fruits, and complement.

3. Result and discussion
The result of this research is schedule of diet for DM patient. The testing is conducted twelve times with different number of individual and the number of generation. The number of individual, generation and fitness score is generated from the patient data, then calculated the calorie needs using Harris-Benedict equations. Table 2 shows the testing result of the proposed method and the correlation among the individual, generation and fitness score is shown as figure 11.

![Fig 11. Correlation among individual, generation and fitness score](image)

| Table 2. F-Score value after evaluation |
|----------------------------------------|
| Total Individu | Total Generation | Best avg fitness value |
|----------------|------------------|------------------------|
| 10             | 50               | 57                     |
| 10             | 70               | 36                     |
| 10             | 100              | 31                     |
| 20             | 50               | 31                     |
| 20             | 70               | 31                     |
| 20             | 100              | 23                     |
| 30             | 50               | 16                     |
| 30             | 70               | 15                     |
| 30             | 100              | 8                      |
| 40             | 50               | 4                      |
| 40             | 70               | 2                      |
| 40             | 100              | 0                      |

Based on Fig. 11, it can be informed that the number of individual and generation affect on fitness score. The greater the number of individual, the greater the generated individuals, which affects on fitness score. The greater the generated individual, the more the possibility of fitness score changes into approaching best fitness score. It also can be informed that the greater the created generation, the more the possibility to get best individual with fitness score approaching 0 or equal to 0. Best individual has the smallest fitness score.

4. Conclusion
Based on all testing results, we conclude that genetic algorithm can be implemented to find optimal solution for scheduling diet for DM patients and works well. The weakness of genetic algorithm is probabilistic, caused by randomly initializing the individual. The greater the number of individual, the more the possibility of fitness score changes into approaching best fitness score. The greater the created generation, the more the possibility to get best individual with fitness score approaching 0 or equal to 0.

Future work can be done by adding generated food data, so it can be more varied and appropriate. Another research possibly done is scheduling diet for DM patient who has another complication of disease. Adding another parameter to calculate the calorie needs, such as metabolic stress level, is also possibly done.

5. References
[1] Baker K R 1974 Introduction to Sequencing and Scheduling *John Wiley & Sons*
[2] Soegondo S, Soewondo P and Imam S 2009 *Penatalaksanaan Diabetes Melitus Terpadu, Edisi Kedua* (Jakarta: Balai Penerbit FKUI)

[3] Badan Penelitian dan Pengembangan Kesehatan 2007 *Riset Kesehatan Dasar (RISKESDAS 2007)* Kementrian Kesehatan Republik Indonesia Jakarta

[4] Phanich M, Pholkul P and Phimoltares S 2010 Food recommendation system using clustering analysis for diabetic patients *Int. Conf. on Information Science and Applications (ICISA)* pp 1–8

[5] Husain W, Wei, L J, Cheng S L and Zakaria N 2011 Application of data mining techniques in a personalized diet recommendation system for cancer patients *IEEE Colloquium on Humanities, Science and Engineering Research (CHUSER 2011)* pp 239–244

[6] Tias R A, Rahman M A and Dewi C Implementasi algoritma genetika untuk optimasi 0/1 multi-dimensional knapsack problem dalam penentuan menu makanan sehat *Jurnal Mahasiswa PTIIK UB*

[7] Indrianingsih Y 2010 *Algoritma genetika untuk menyelesaikan masalah optimasi fungsi berkendala dengan pengkodean bilangan bulat* Thesis Sekolah Tinggi Teknologi Adisutjipto

[8] Puspaningrum W A, Djunaidy A and Vinarti R A 2014 Penjadwalan mata kuliah menggunakan algoritma genetika di jurusan sistem informasi ITS *Jurnal Teknik POMITS* 2 127-133

[9] Kushardiana R 2013 *Penetuan Komposisi Menu Makanan untuk Penderita Diabetes Mellitus Menggunakan Algoritma Genetika* Thesis Universitas Brawijaya

[10] Almatser S 2004 *Penuntun Diet* (PT. Gramedia Pustaka Utama: Jakarta)

[11] Kusumadewi S and Purnomo H 2005 Penyelesaian Masalah Optimasi dengan Teknik-Teknik Heuristik (Yogyakarta: Graha Ilmu)