World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:6, 2007

DNA Computing for an Absolute 1-Center Problem: An Evolutionary Approach

Zuwairie Ibrahim, Yusei Tsuboi, Osamu Ono, Marzuki Khalid

Abstract**—**Deoxyribonucleic Acid or DNA computing has emerged as an interdisciplinary field that draws together chemistry, molecular biology, computer science and mathematics. Thus, in this paper, the possibility of DNA-based computing to solve an absolute 1-center problem by molecular manipulations is presented. This is truly the first attempt to solve such a problem by DNA-based computing approach. Since, part of the procedures involve with shortest path computation, research works on DNA computing for shortest path Traveling Salesman Problem, in short, TSP are reviewed. These approaches are studied and only the appropriate one is adapted in designing the computation procedures. This DNA-based computation is designed in such a way that every path is encoded by ligonucleotides and the path's length is directly proportional to the ength of oligonucleotides. Using these properties, gel electrophoresis s performed in order to separate the respective DNA molecules ccording to their length. One expectation arise from this paper is that t is possible to verify the instance absolute 1-center problem using NA computing by laboratory experiments.

Keywords**—**DNA computing, operation research, 1-center roblem.

I. Introduction

IN 1965, Gordon E. Moore [1] observed an exponential growth in the number of transistor per integrated circuit gainst time. Currently, this is the definition of Moore's Law, neaning that more and more transistors can be crammed into a hip until the silicon itself reaches its finite atomic scale imitation. Since the traditional silicon-based computer is estricted by its fundamental physical limitation, researchers ave been searching for alternative medium for computation nd DNA would turn out to be the answer.

A new computing paradigm based on DNA molecules is appeared in 1994 when Leonard M. Adleman [2] launched a novel approach to solve the so-called Hamiltonian Path Problem (HPP) with seven vertices by DNA molecules. While in conventional silicon-based computer, information is stored as binary numbers in silicon-based memory, he encoded the information of the vertices by a randomly DNA sequences. The computation is performed in biochemical reaction fashion involving hybridization, denaturation, ligation, Polymerase Chain Reaction (PCR), etc. The output of the computation, also in the form of DNA molecules can be read and printed by electrophoretical fluorescent method.

DNA molecules are composed of single or double DNA fragments or often called oligonucleotides or strands. Nucleotides form the basis of DNA. A single-stranded fragment has a phospho-sugar backbone and four kinds of bases denoted by the symbols A, T, G, and C for the bases adenine, thymine, guanine, and cytosine respectively. These four nucleic acids, which can occur in any order combined in Watson-Crick complementary pairs to form a double strand helix of DNA. Due to the hybridization reaction, A is complementary with T and C is complementary with G. As an example, any sequence oligonucleotides, such as 5'-ACCTG-3' has a complementary sequence, 3'-TGGAC-5'. Digits 5' and 3' denote orientation of DNA oligonucleotides.

The facility location optimization has many applications reported to date mostly in operations research and network design problems. It has been used in placement of routes [3] and internet web proxies [4], agglomeration of traffic [5-6], and web server replications in a content distribution network [7]. It can be largely divided into two main problems, namely MINSUM and MINMAX. MINSUM problem is to minimize the sum of the weighted distances to the nearest source facility. An example is p-medium problem [8]. On the other hand, MINMAX problem is to minimize the maximum weighted distance from the demand destinations to the nearest source facility. It is also called p-center problems [9]. Even if the number p=1 but the network is large, the problems are almost NP-complete. It is then become relevant to develop a new method based on DNA computing where polynomial-time computation may be developed. For these reasons, the authors concentrate on solving an absolute 1-center problem using DNA-based computing.

Gigantic memory capacity and massively parallelism inherent in DNA computing can be exploited to make a brute force search on a big problem space in constant or polynomial-time. That is why DNA computing is well known very suitable for NP-complete problems. After Adleman's

Manuscript received December 21, 2003.

Zuwairie Ibrahim was with the Department of Microelectronics and Computer Engineering, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Johor Darul Takzim, Malaysia. He is now pursuing his PhD at the Institute of Applied DNA Computing, Meiji University, Kanagawa, 214-8571 Japan (phone: +81-44-934-7377; fax: +81-44-934-7909; e-mail: zuwairie@isc.meiji.ac.jp).

Yusei Tsuboi is with the Institute of Applied DNA Computing, Meiji University, Kanagawa, 214-8571 Japan (phone: +81-44-934-7377; fax: +81-44-934-7909; e-mail: tsuboi@isc.meiji.ac.jp).

Osamu Ono is the Director of Institute of Applied DNA Computing, Meiji University, Kanagawa, 214-8571 Japan (phone: +81-44-934-7377; fax: +81-44-934-7909; e-mail: ono@isc.meiji.ac.jp).

Marzuki Khalid is the Director of Center for Artificial Intelligence and Robotics (CAIRO), University Technology of Malaysia, Jalan Semarak, 54100 Kuala Lumpur, Malaysia (phone: +6-03-2691-3710; fax: +6-03-2697-0815; e-mail: marzuki@utmkl.utm.my).

novel discovery is published, various combinatorial problems have been chosen to express the possibility of DNA computing in many applications. It is reported that DNA computing is appropriate to compute the maximal clique problem [10], to solve satisfiability (SAT) problem [11] and even to break the Discrete Encryption Standard (DES) [12]. However, one of the combinatorial problems not being solved by DNA computing yet is the facility location optimization, in particular, an absolute 1-center problem. Nevertheless, this problem has been solved extensively by non-molecular or mathematical approaches and performed on sequential silicon based computer as computing platform. Some examples are greedy method [13], approximation algorithms [14] and by solving a series of set covering problems [15].

So far, nobody has proposed a molecular algorithm for performing 1-center problem. Hence, the authors confidently claim that the proposed approach is the first effort to solve the problem by DNA computing. In this paper, a new computational design is developed to solve an absolute 1-center problem by DNA-based computing approach. This idea is motivated from the previous research works carried out to compute the shortest path of TSP. The procedures are adapted and the overall design are modified and enhanced in order to solve the selected problem. The proposed computation necessarily composed of combination of DNA bio-chemical operation such as DNA synthesis, PCR, ligation, gel electrophoresis, annealing, melting and hybridization. More importantly, the proposed DNA-based computational approach may be easier to accomplish because the tedious affinity-purification operation will be never been used during the computation.

II. Molecular Computation for Shortest Path Problem

At first, Adleman pioneered the molecular computation by solving seven nodes of HPP. The goal of HPP is to determine whether a path exists that will commerce at the start city, finish at the end city and pass through each city of the remaining cities exactly once. While this solution was very promising, the problem did not involve shortest path computation at all.

A variation of HPP, shortest path TSP tackled by Narayanan and Zorbalas [16] to solve by DNA algorithms. A constant increase of DNA strands according of the actual length of distance is occupied. A drawback of this method is that for some concatenations, concatenated DNA strand of two distances, while less numerically than the longest distance, may be longer than the DNA strands for that longest distance. This may leads to errors in computing shortest path. This scheme however was not been realized by laboratory experiments.

Yamamoto et al. presented concentration-controlled DNA computing for accomplishing a local search for shortest path problem [17]. Although DNA computing with concentration control method enables local search among all the candidate solutions, it is not guaranteed that the most intensive band is the DNA representing the shortest path in the given graph. In addition, it is technically difficult to extract a single optimal solution from the most intensive band.

J. Y. Lee and his colleagues proposed temperature gradient-based DNA computing for shortest path TSP problem [18]. They introduce melting temperature (T_m) for this purpose. At certain temperature, DNA strands of correct solutions will be denatured at that temperature, and be amplified by PCR process. As the denaturation temperature increases, other DNA strands also will be amplified. But the amount of correct solutions will be exponentially increased cycle by cycle, and occupy the major part of the solution. However, after hybridization, the temperature for the effective PCR process is not provided.

It is better to note that the absolute 1-center problem, even it involved shortest path computation as part of the problem, it requires computing the longest path among a set of shortest paths. It also requires computing the shortest path among a set of longest paths. Thus, the authors possibly found a molecular computation approach to do such computation. The computing strategy is motivated from [16] but we did a lot of modifications in order to make it suitable for computational needs. In this approach, the edge weights are handled by the DNA strands and the actual weights are approximated by the directly proportional length DNA strands. Hence, the important information during the computation is DNA length. The computation only considers about the starting node and end node encoded by DNA strands. The actual nodes encoded within the DNA strands will not be computed. As a result, the affinity-purification process is not required during the computation.

III. Problem Formulation

The 1-center problem may be phrased as follows. As given by Cheriyan and Ravi [19], the problem is to locate a facility on the points of a given network or graph so as to minimize the maximum cost of a path between a demand node and its nearest facility. A weighted undirected network or graph $G = (V, E)$ with 5 vertices, 7 edges and 7 weights is used to model the problem where V and E are the vertices and edges of the graph respectively. If a path exists between the two nodes, a positive weight is assigned to that path. The graph is depicted in Fig. 1 [20].

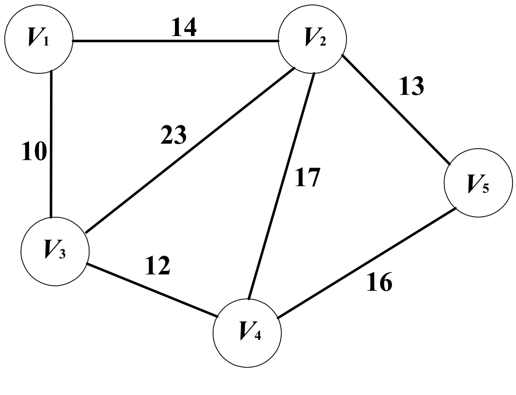


Figure 1: A weighted undirected graph $G = (V, E)$.

Normally, in order to solve this problem, one can note the shortest path between a particular node to its nearest node by a matrix as in Table 1 and then choose a node such that the maximum entry in its row in the matrix is smallest among the maximum entries of all rows. The underlined bold numbers highlights the maximum entries for each row. According to Table 1, it is clear that the right answer of this problem is to locate at node V_4 with a maximum distance of 22.

TABLE I
NODE-TO-NODE DISTANCE BY A MATRIX

	V_1	V_2	V_3	V_4	V_5
V_1	0	14	10	22	**27**
V_2	14	0	**23**	17	13
V_3	10	23	0	12	**28**
V_4	**22**	17	12	0	16
V_5	27	13	**28**	16	0

IV. DNA BIOCHEMICAL OPERATIONS

There are a number of feasible DNA bio-chemical operations in DNA computing. As such, one can manipulate or generate a solution containing DNA strands by DNA synthesis, polymerase chain reaction (PCR), ligation and gel electrophoresis. In this section, these essential tools are described in more details as follows:

A. *DNA Synthesis*

At present, it is possible to get a test tube containing approximately 10^{18} DNA molecules with a desired sequence. Some commercial DNA synthesis companies are available, which provide a reasonable price for this reason. As in 1998, Adleman [21] said that the cost for sequences of length 20 is just about $25.

B. *Denaturation*

Double stranded DNA molecules can be separated without breaking the single strands by applying heat to the solution at about 95°C. The double stranded molecules will come apart because the hydrogen bonds between complementary nucleotides are much weaker than the covalent bond between the adjacent nucleotides in the same strands.

C. *Annealing, Concatenation, or Hybridization*

As oppose to denaturation, by cooling down a DNA solution from approximately 95°C to 55°C, single Watson-Crick complementary DNA sequences will combined to form double stranded molecules. Also, based on this behavior, concatenation of two identical strands would be happened with the presence of a connective strand as depicted in Fig. 2. Concatenation is crucial in Adleman's demonstration and in this proposed approach as well.

D. *Polymerase Chain Reaction*

PCR is an incredible sensitive copying machine for DNA. It also can be used for DNA detection. Given a site-specific single molecule DNA, a million or even billion of similar molecules can be created by PCR process. In n steps, it can produces 2^n copies of the same molecules. PCR needs a number of sub-sequence strands called 'primers', which is usually about 20 base long to signal a specific start and end site at a template for replication. PCR normally runs for 20-30 cycles of 3 steps as shown in Fig. 3. Each cycle consist of three phases: separating base pair strands of DNA at about 95°C, annealing at 55°C and extension at 75°C [22]. It takes about 3 hours normally to complete the cycles.

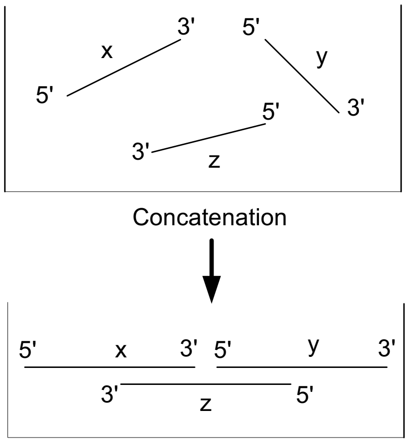


Figure 2: Concatenation of strand 5' – x – 3' and 5' – y – 3' with a presence of connective strand 3' – z – 5'. If 5' – x – 3' = 5' – $x_1 x_2$ – 3', 5' – y – 3' = 5' – $y_1 y_2$ – 3', then 3' – z – 5' = 3' – $\overline{x}_2\overline{y}_1$ – 5', where $\overline{x}_2$ and $\overline{y}_1$ indicate the complementary strand of x_2 and y_1 respectively.

E. *Ligation*

Ligation often invoked after the single DNA strands are annealed and concatenated to each other. Many single-strand fragments will be connected in series and ligase is used as 'glue' to seal the covalent bonds between the adjacent fragments as shown in Figure 4 [23].

F. *Gel Electrophoresis*

DNA strands in a solution can be separated in term of its length by means of gel electrophoresis. In fact, the molecules are separated according to their weight, which is almost proportional to their length [24]. This technique is based on the fact that DNA molecules are negatively charged [25]. Hence, by putting them in an electric field, they will move towards the positive electrode at different speed. The longer molecules will remain behind the shorter ones. The speed also depends on the gel porosity. Agarose gel is frequently used and by varying the porosity of the gel used, the sensitivity of this length separation operation can be altered. The precision is high even molecules which differ by one nucleotide can still be distinguished between each other. An example of gel electrophoresis process [26] and its output [27] are well depicted in Figure 5 and Figure 6 respectively. This technique can be used to "print" the results of DNA computation as well. Normally, at the end of this process, the gel is photographed for convenience.

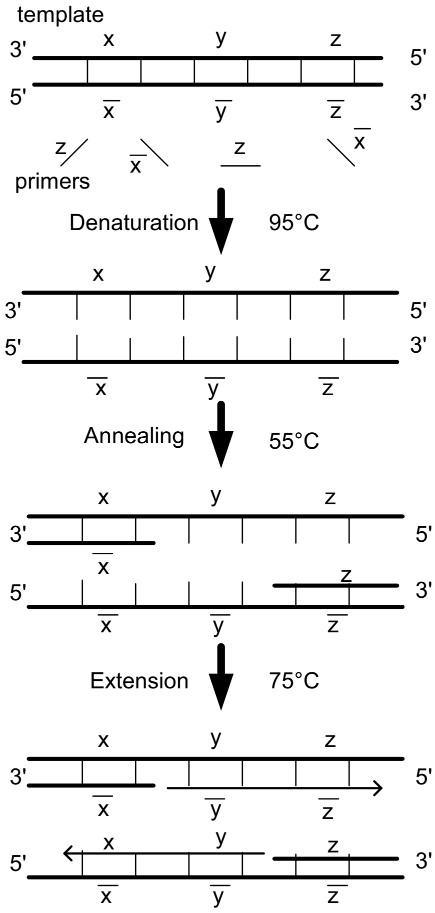


Figure 3: One PCR cycle. For 30 cycles of PCR, these 3 steps are repeated for 30 times.

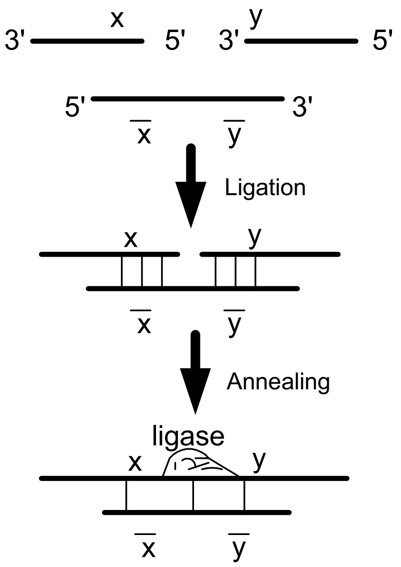


Figure 4: Ligation reaction.

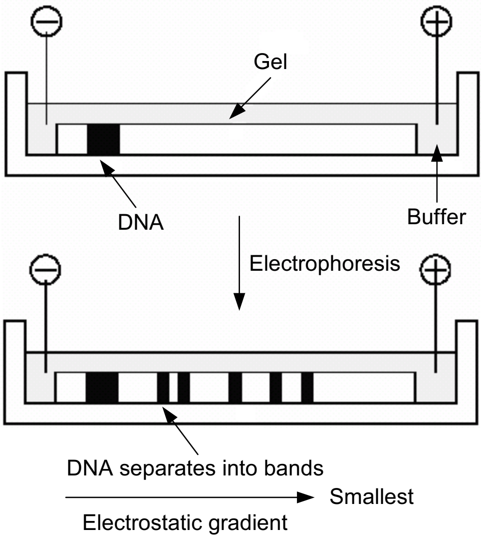


Figure 5: Gel electrophoresis process.

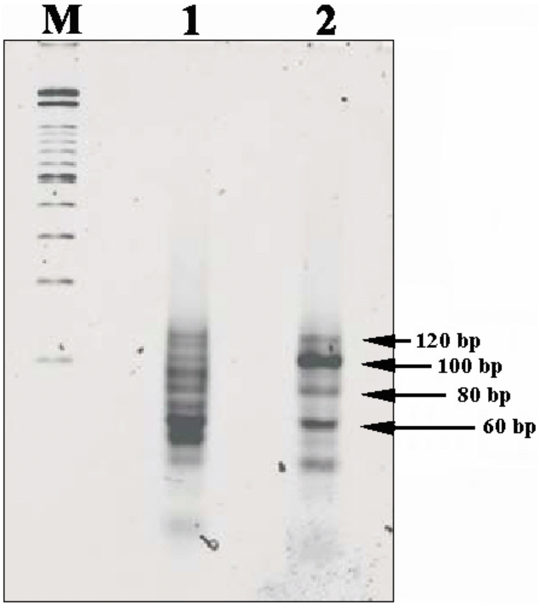


Figure 6: Gel electrophoresis output where lane M is DNA size marker. Lane 1 and 2 are used for the tested DNA molecules.

V. Computing With DNA

The overall computation described in this section consists of five steps altogether. Step 1 through step 3 is crucial for the generating initial pool, random route formation, whereas step 4 through step 6 is designed for the computation process and applied to the initial pool of solution.

A. Step 1

Let *V* be the total number of nodes in the graph and *E* the total number of paths in the graph. The random sequence strands correspond to all nodes (and its complements) and distances (and its complements) are created. Let *O_i (i = 1,.....,* *V)* and $\overline{O}\_i$ *(i = 1,....., V)* be the fixed length random sequences correspond to all nodes in the graph and its complements respectively. In the same way, let *O_d (d = 1,....., E)* and $\overline{O}\_d$ *(d = 1,....., E)* be the variable length random sequences correspond to all the distances and its complements respectively. The length *l* of *O_d* and $\overline{O}\_d$ will be directly proportional to the weight distances. For instance, let say, for a constant proportional increase of 2, all the random sequences are placed in Table 2 and Table 3. For the sake of convenience, in Table 2 and Table 3, in order to avoid clumsy presentation, we will use symbolic representation of the distance sequences. Recall that these sequences would be occurred in any order of nucleotides and its length will be directly proportional to the distances. At the end of this step, oligonucleotides $\overline{O}\_d$ and $\overline{O}\_i$ are synthesized.

TABLE II
FIXED LENGTH NODE RANDOM SEQUENCES

Node	DNA Code (5'-3')	Complements (3'-5')
V_1	TATT	ATAA
V_2	GCGG	CGCC
V_3	GTCG	CAGC
V_4	AGAA	TCTT
V_5	TCCG	AGGC

TABLE III
VARIABLE LENGTH DISTANCE RANDOM SEQUENCES

Edge Cost	DNA Code (5'-3')	Complements (3'-5')
D {10}	$\boxed{20}$	$\boxed{\overline{20}}$
D {12}	$\boxed{24}$	$\boxed{\overline{24}}$
D {13}	$\boxed{26}$	$\boxed{\overline{26}}$
D {14}	$\boxed{28}$	$\boxed{\overline{28}}$
D {16}	$\boxed{32}$	$\boxed{\overline{32}}$
D {17}	$\boxed{34}$	$\boxed{\overline{34}}$
D {23}	$\boxed{46}$	$\boxed{\overline{46}}$

B. Step 2

Let *i* be the start node of a path, *d* the path's distance and *j* the end node of that path. Oligonucleotides representing every path between two nodes in the graph are synthesized as follows:

synthesize oligonucleotides
O_i-d-j_end as **HR** *O_i* + **ALL** *O_d* + **ALL** *O_j*
O_j-d-i_end as **HR** *O_j* + **ALL** *O_d* + **ALL** *O_i*
O_i-d-j as **HR** *O_i* + **ALL** *O_d* + **HL** *O_j*
O_j-d-i as **HR** *O_j* + **ALL** *O_d* + **HL** *O_i*

where '**ALL**' represents the whole DNA strand, '**HL**' the left half part, '**HR**' the right half part and '+' is a join. For each path, connecting two identical nodes, each of the nodes could be a starting node, an end node or a connective node in a path. Thus, four identical oligonucleotides will be created for each edge.

As such, if *O_i* = **TATT** then $\overline{O}\_i$ = **ATAA**; if *O_j* = **GCGG** then $\overline{O}\_j$ = **CGCC** and if *O_d* = **ACGACGTT** then $\overline{O}\_d$ = **TCGTGCAA**. Then the path *O_i-d-j_end* is **TTACGACGTTGCGG** (HR *O_i* + ALL *O_d* + ALL *O_j*), the path *O_j-d-i_end* is **GGACGACGTTTATT** (HR *O_j* + ALL *O_d* + ALL *O_i*), the path *O_i-d-j* is **TTACGACGTTGC** (HR *O_i* + ALL *O_d* + HL *O_j*), and lastly the path *O_j-d-i* is **GGACGACGTTTA** (HR *O_j* + ALL *O_d* + HL *O_i*).

At the end of this step, 28 identical edge strands will be created by a combination of the nodes sequences and distance sequences. Because of this combination, the edges strands *O_i-d-j*, *O_i-d-j_end*, *O_j-d-i*, and *O_j-d- i_end* listed in Table 4 are influenced by the additional nodes bases and hence, unable to handle the weight of each edges. The additional bases would cause errors during the computation because the additional bases composed in the edge sequences reflect the weight value of edges. To compensate the additional bases, the strands corresponds to distance sequences are shorten according to the additional bases and the node sequences in the edge strands are left unchanged. All the modified oligonucleotides *O_i-d-j*(*modified*), *O_i-d-j_end*(*modified*), *O_j-d-i*(*modified*), and *O_j-d-i_end*(*modified*) are listed in Table 5. This modification however, decreases the generality of the approach because the weight of edges less than 3 is not allowed.

C. Step 3

All modified oligonucleotides *O_i-d-j*(*modified*), *O_i-d-j_end*(*modified*), *O_j-d-i*(*modified*), and *O_j-d-i_end*(*modified*) followed by all complement oligonucleotides ($\overline{O}\_i$ and $\overline{O}\_d$) are inserted into a test tube and DNA ligase reaction is performed in which random routes through the graph are formed. In this way, the initial solution, N_0, which contains a large set of random routes of the graph are generated. After that, in order to make sure that no sticky end strands exist in the solution, polymerase enzyme is used for the extension where the sticky end strands are altered to become blunt end strands.

At this moment, an initial solution N_0 is already obtained and it is time to filter out the right routes among the vast random routes of the graph. Unlike the conventional filtering, it is not merely throwing away the unwanted strands but rather copying the right strands exponentially by incredibly sensitive PCR process.

D. Step 4

Produce another five solutions from the initial solution, N_0. Each solution represents exactly one row in Table 1.

TABLE IV
EDGE SEQUENCES

Edges	DNA Code (5'-3')
$V_3 - V_1$	CG $\boxed{20}$ TATT
	TT $\boxed{20}$ GTCG
	CG $\boxed{20}$ TA
	TT $\boxed{20}$ GT
$V_4 - V_3$	AA $\boxed{24}$ GTCG
	CG $\boxed{24}$ AGAA
	AA $\boxed{24}$ GT
	CG $\boxed{24}$ AG
$V_2 - V_5$	GG $\boxed{26}$ TCCG
	CG $\boxed{26}$ GCGG
	GG $\boxed{26}$ TC
	CG $\boxed{26}$ GC
$V_1 - V_2$	TT $\boxed{28}$ GCGG
	GG $\boxed{28}$ TATT
	TT $\boxed{28}$ GC
	GG $\boxed{28}$ TA
$V_5 - V_4$	CG $\boxed{32}$ AGAA
	AA $\boxed{32}$ TCCG
	CG $\boxed{32}$ AG
	AA $\boxed{32}$ TC
$V_4 - V_2$	AA $\boxed{34}$ GCGG
	GG $\boxed{34}$ AGAA
	AA $\boxed{34}$ GC
	GG $\boxed{34}$ AG
$V_2 - V_3$	GG $\boxed{46}$ GTCG
	CG $\boxed{46}$ GCGG
	GG $\boxed{46}$ GT
	CG $\boxed{46}$ GC

TABLE V
MODIFIED EDGE SEQUENCES

Edge	DNA Code (5'-3')
$V_3 - V_1$	CG $\boxed{14}$ TATT
	TT $\boxed{14}$ GTCG
	CG $\boxed{16}$ TA
	TT $\boxed{16}$ GT
$V_4 - V_3$	AA $\boxed{18}$ GTCG
	CG $\boxed{18}$ AGAA
	AA $\boxed{20}$ GT
	CG $\boxed{20}$ AG
$V_2 - V_5$	GG $\boxed{20}$ TCCG
	CG $\boxed{20}$ GCGG
	GG $\boxed{22}$ TC
	CG $\boxed{22}$ GC
$V_1 - V_2$	TT $\boxed{22}$ GCGG
	GG $\boxed{22}$ TATT
	TT $\boxed{24}$ GC
	GG $\boxed{24}$ TA
$V_5 - V_4$	CG $\boxed{26}$ AGAA
	AA $\boxed{26}$ TCCG
	CG $\boxed{28}$ AG
	AA $\boxed{28}$ TC
$V_4 - V_2$	AA $\boxed{28}$ GCGG
	GG $\boxed{28}$ AGAA
	AA $\boxed{30}$ GC
	GG $\boxed{30}$ AG
$V_2 - V_3$	GG $\boxed{40}$ GTCG
	CG $\boxed{40}$ GCGG
	GG $\boxed{42}$ GT
	CG $\boxed{42}$ GC

E. Step 5

For each solution (or each row);

i. Select a particular node as a starting point and find the shortest path from the starting node to other nodes.
ii. Store the strands representing each shortest path into a solution.
iii. Gather all together the solution representing all shortest paths.
iv. Search the longest strand among the shortest strands, which indicates the maximum path of a particular row in Table 1.

F. Step 6

Gather all together the solutions that represent all the longest paths of each row. Search the shortest strand. Find the starting node of that path to get the answer of the absolute 1-center problem. The overall procedures are summarized by a flow chart given in Fig. 7.

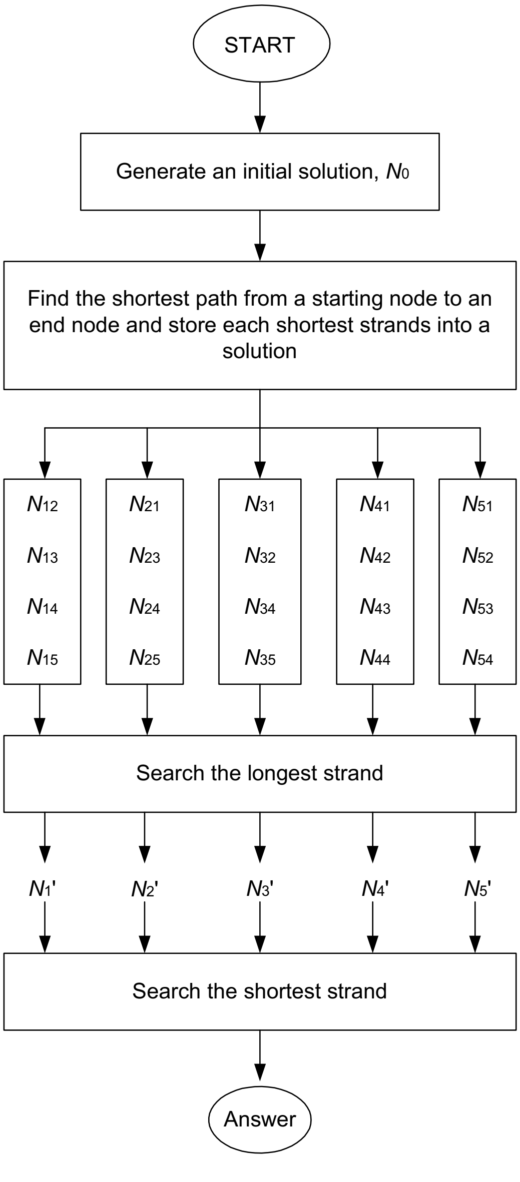

START
Generate an initial solution, N_0
Find the shortest path from a starting node to an end node and store each shortest strands into a solution
N_{12} N_{13} N_{14} N_{15}
N_{21} N_{23} N_{24} N_{25}
N_{31} N_{32} N_{34} N_{35}
N_{41} N_{42} N_{43} N_{44}
N_{51} N_{52} N_{53} N_{54}
Search the longest strand
N_1' N_2' N_3' N_4' N_5'
Search the shortest strand
Answer

Figure 7: Overall procedures.

VI. Discussions

The fourth step could be done by merely divides the solution N_0 equally into five solutions and marks the solutions as N_1, N_2, N_3, N_4, and N_5. In this work, it is assumed that all the strands are spread to all areas in the solution.

It is possible to carry out the fifth step by the subsequent procedure. Since the dividing operation is a volume decreasing operation, it is important, therefore, to increase the volume of solutions to a sufficient amount. The only operation, which is able to increase the volume of the solutions, is the polymerase chain reaction (PCR) operation. Besides, this operation selects the routes that begin with a particular node for each test tube obtained in the previous step. Five types of primers will be employed where one primer, encoding the 3'-half of starting node is assigned to one particular test tube as shown in Fig. 8.

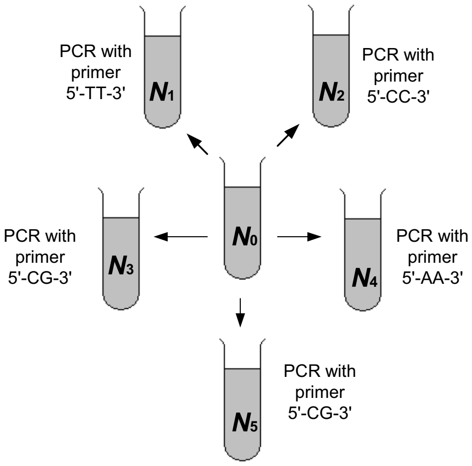


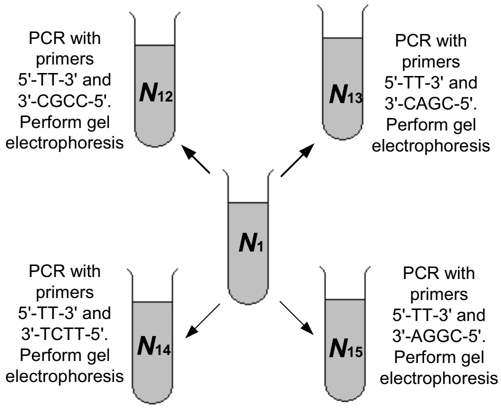
Figure 8: The initial solution N_0 is divided equally into 5 solutions. PCR process is executed to each solution with 5 different primers representing 5 starting nodes.

Next, closely similar to the previous operation, by assuming that all the strands are spread uniformly in the solution, each solution obtained from previous operation, N_1, N_2, N_3, N_4, and N_5 are divided equally into another 4 test tube. As a result, there should be 20 test tubes all together. Next, DNA strands with a particular end node are amplified and consequently, the volume of solutions is increase as well. PCR process is employed for each test tube by primers encoding the 3'-half of starting node and complement of the end nodes. This operation is well depicted in Fig. 9.

For better understanding, take a look closely at the molecules of solution N_{12} only. After two PCR operations are accomplished, there should be a numerous number of strands representing the beginning node, V_1 and the end node, V_2. Some of them are shown in Fig. 10. Next, gel electrophoresis technique will separates these strands in term of length and the shortest length is chosen in order to obtain the shortest path. As a result, the strands $V_1 \rightarrow V_2$ will be chosen. The output solution is called N_{12}'.

The same procedure will be executed to the solution N_{13}, N_{14}, and N_{15} results in 3 solutions, N_{13}', N_{14}', and N_{15}' consisting the shortest strands $V_1 \rightarrow V_3$, $V_1 \rightarrow V_3 \rightarrow V_4$, and $V_1 \rightarrow V_2 \rightarrow V_5$ respectively. These strands are shown in Fig. 11.

All these four solutions, N_{12}', N_{13}', N_{14}', and N_{15}' will be gathered together and due to the previous gel electrophoresis outputs, the longest strand will be selected among the shortest strands. Meaning that the strand N_{15}' will be selected. The resultant solution will be called as N_1'. The same procedures for N_1 is repeated to the solution N_2, N_3, N_4, and N_5 and the output solution N_2', N_3', N_4', and N_5' respectively will be produced. At this stage, the sixth step is already completed.


PCR with primers 5'-TT-3' and 3'-CGCC-5'. Perform gel electrophoresis
N12
N13
PCR with primers 5'-TT-3' and 3'-CAGC-5'. Perform gel electrophoresis
N1
PCR with primers 5'-TT-3' and 3'-TCTT-5'. Perform gel electrophoresis
N14
N15
PCR with primers 5'-TT-3' and 3'-AGGC-5'. Perform gel electrophoresis

igure 9: For instance, N_1 is divided into another four test tubes amely N_{12}, N_{13}, N_{14}, and N_{15}. PCR process is executed again to ach solution with primers representing the 3' half of start node nd complement of end nodes. Next, for each test tube, the gel lectrophoresis is applied for separating the strands by length.

TT	22	GCGG
AA	$\overline{22}$	CGCC

(a)

TT	14	GT	CG	42	GCGG
AA	$\overline{14}$	CA	GC	$\overline{42}$	CGCC

(b)

TT	14	GT	CG	20	AG	AA	30	GCGG
AA	$\overline{14}$	CA	GC	$\overline{20}$	TC	TT	30	CGCC

(c)

Figure 10: (a) Strand $V_1 \rightarrow V_2$ (b) Strand $V_1 \rightarrow V_3 \rightarrow V_2$, and (c) Strand $V_1 \rightarrow V_3 \rightarrow V_4 \rightarrow V_2$.

Lastly, the sixth step begins with collecting the solution N_1', N_2', N_3', N_4', and N_5'. According to the previous gel electrophoresis results, the solution representing the shortest strand will be chosen manually. This solution indicates the actual answer of the absolute 1-center problem. The selection involved based on gel electrophoresis is shown in Fig. 12.

TT	14	GTCG
AA	$\overline{14}$	CAGC

(a)

TT	14	GT	CG	20	AGAA
AA	$\overline{14}$	CA	GC	$\overline{20}$	TCTT

(b)

TT	22	GC	GG	22	TCCG
AA	$\overline{22}$	CG	CC	$\overline{22}$	AGAA

(c)

Figure 11: (a) Strand $V_1 \rightarrow V_3$ (b) Strand $V_1 \rightarrow V_3 \rightarrow V_4$, and (c) Strand $V_1 \rightarrow V_2 \rightarrow V_5$.

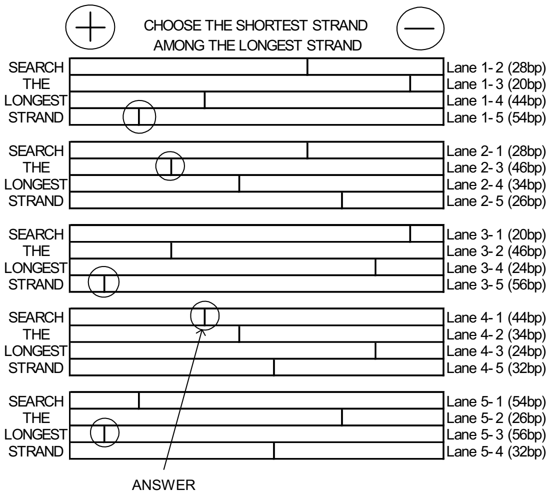


Figure 12: The selection based on expectation result of gel electrophoresis.

Heuristically, at this stage, there is only one type of strand exist. The DNA strands indicating the path from node V_4 to V_1 (or explicitly the path from node V_4 to V_3 to V_1). This expected result shows that one has to locate the facility on point V_4 in order to minimize the maximum cost of a path between a demand node to its nearest facility.

VII. Conclusions

This objective of this paper is to certify the feasibility of DNA computing in an absolute 1-center problem. The authors carried out the first attempt to compute 1-center problem by DNA-based computing approach. The contents of this paper also imply a possibility to verify the proposed procedures by experimental work. Since the computing platform is now different, it is not relevant to make a comparison between the proposed approach and other existing approaches. Moreover, this computing platform is still infant compared to that matured conventional silicon-based computation. Despite all the limitations, the usefulness and strong potentials of DNA-based computational approach provides a compelling reason to explore a lot of possible applications of DNA computing and solving an absolute 1-center problem is one of them.

REFERENCES

[1] G. E. Moore, "Cramming more components onto integrated circuits," *Electronics*, vol. 38, no. 8, 1965.

[2] L. Adleman, "Molecular computation of solutions to combinatorial problems," *Science*, vol. 266, 1994, pp. 1021-1024.

[3] S. Guha, A. Meyerson, and K. Munagala, "Hierarchical placement and network design problems," *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, 2000.

[4] B. Li, M. Golin, G. Italiano, X. Deng, and K. Sohraby, "On the optimal placement of web proxies in the internet," *Proceedings of 18th Annual Joint Conference of the IEEE Computer and Communications Societies*, 1999, pp. 1282-1290.

[5] M. Andrews and L. Zhang, "The access network design problem," *Proceedings of the 39th Annual IEEE Symposium on Foundations of Computer Science*, 1998.

[6] S. Guha, S. Meyerson, and K. Munagala, "Improve combinatorial algorithm for single sink edge installation problems," *Technical Report STAN-CS-TN00-96*: Stanford University, 2000.

[7] S. Jamin, C. Jin, Y. Jin, D. Raz, Y. Shavitt, and L. Zhang, "On the placement of internet instrumentations," *Proceedings of 19th Annual Joint Conference of the IEEE Computer and Communications Societies*, 2000, pp. 26-30.

8] E. S. Correa, M. T. A. Steiner, A. A. Freitas, and C. Carnieri, "A genetic algorithm for the p-medium problem," *Proceedings of Genetic and Evolutionary Computation Conference*, 2001, pp. 1268-1275.

9] K. Jain, M. Mahdian, and A. Saberi, "A new greedy approach for facility location problems," *Proceedings of 3rd ACM Symposium on Theory of Computing*, 2002, pp. 731-740.

10] O. Ouyang, P. D. Kaplan, L. Shumao, and A. Libchaber, "DNA solution of the maximal clique problem," *Science*, vol. 278, 1997, pp. 446-449.

11] R. S. Braich, N. Chelyapov, C. Johnson, P. W. K. Rothemund, and L. Adleman, "Solution of a 20-variable 3-SAT problem on a DNA computer," *Science*, vol. 296, 2002, pp. 499-502.

12] L. M. Adleman, P. W. Rothemund, S. Rowies, and E. Winfree, "On applying molecular computation to the data encryption standard," *Journal of Computational Biology*, vol. 6, no. 1, 1999, pp. 53-63.

13] T. Gonzalez, "Clustering to minimize the maximum inercluster distance," *Theoretical Computer Science*, vol. 38, 1985, pp. 293-306.

14] J. Mihelic and B. Robic, "Approximation algorithms for k-center problem: an experimental evaluation," *Proceedings of Operations Research*, 2002.

15] M. S. Daskin, *Network and Discrete Location: Models, Algorithms, and Applications*, New York: Wiley, 1995.

16] A. Narayanan and S. Zorbalas, "DNA algorithms for computing shortest paths," *Proceedings of Genetic Programming*, 1998, pp. 718-723.

17] M. Yamamoto, A. Kameda, N. Matsuura, T. Shiba, Y. Kawazoe, and A. Ohochi, "Local search by concentration-controlled DNA computing," *International Journal of Computational Intelligence and Applications*, vol. 2, no. 4, 2002, pp. 447-455.

18] J. Y. Lee, S. Y. Shin, S. J. Augh, T. H. Park, and B. T. Zhang, "Temperature gradient-based DNA computing for graph problems with weighted edges," *Lecture Notes in Computer Science*, vol. 2568, 2003, pp. 73-84.

[19] J. Cheriyan and R. Ravi, *Lecture Notes on Approximation Algorithms for Network Problems*, Canada: University of Waterloo, 1998.

[20] E. A. Cabral, *Lecture Notes on Distribution Managament (MGTSC 461)*, Canada: University of Alberta, 2000.

[21] L. M. Adleman, "Computing with DNA," *Scientific American*, 1998, pp. 34-41.

[22] J. P. Fitch, *An Engineering Introduction to Biotechnology*, Washington: SPIE, 2002.

[23] M. Zucca, *DNA Based Computational Models,* PhD Thesis, Politecnico Di Torino, Italy, 2000.

[24] C. S. Calude and G. Paun, *Computing with Cells and Atoms: An Introduction to Quantum, DNA, and Membrane Computing,* New York: Taylor & Francis Inc, 2001.

[25] G. Paun, G. Rozenberg, and A. Salomaa, *DNA Computing: New Computing Paradigms*, Heidelberg: Springer-Verlag, 1998.

[26] M. Amos, *DNA Computation*, PhD Thesis, The University of Warwick, UK, 1997.

[27] M. Yamamoto, A. Kameda, N. Matsuura, T. Shiba, Y. Kawazoe, and A. Ohochi, "A separation method for DNA computing based on concentration control," *New Generation Computing*, vol. 20, no. 3, 2002, pp. 251-262.

Zuwairie Ibrahim received the B.Eng (Mechatronics) and M.Eng. (Image Processing) from Universiti Teknologi Malaysia, Malaysia, in 2000 and 2002 respectively. Since 2002, he engaged with Department of Mechatronics and Robotics, Universiti Teknologi Malaysia as a lecturer. He is currently pursuing his PhD at the Institute of Applied DNA Computing, Meiji University, Kanagawa, Japan. He is a student member of Institute of Electrical and Electronics Engineers (IEEE), International Computational Intelligence Society (ICIS), and International Signal Processing Society (ISPC). His research interests include signal and image processing, automated visual inspection, evolutionary and unconventional computing such as molecular or DNA computing.

Yusei Tsuboi received the B. Eng and Master degree from Meiji University, Japan, in 2000 and 2002 respectively. Currently, he is pursuing the PhD degree also at Meiji University, Japan. Presently, his research interest includes robotics and DNA computing for artificial intelligence.

Osamu Ono received the Bachelor, Master and Doctor Degree in Engineering all from Waseda University, Tokyo, in 1974, 1976, and 1979 respectively. He is a Professor of the Department of Electrical and Electronic Engineering, Meiji University, Japan. He is the Director of Tokyo Branch of Japan Institute Electrical Engineering and a committee member of Japan Society for Simulation Technology (JSST). His research interest includes large scale industrial process, mechatronics, advanced mobile robotics and image processing. Currently, he is interested on the application of DNA computing in engineering field.

Marzuki Khalid is currently a Professor in Intelligent Control and Director at the Centre of Artificial Inteligence and Robotics (CAIRO) of Universiti Teknologi Malaysia. His current research interest is in the field of artificial intelligence with applications to control. He has been appointed to the Editorial Advisory Board of the International Journal of Engineering Applications of Artificial Intelligence published by Elsevier Science and an Associate Editor of the Journal of Systems and Control Engineering published by the Institute of Mechanical Engineers, United Kingdom. He is currently the IEEE Student Activities Chair for Region 10 (Asia Pacific) and also the Founding Member of the Asian Control Professors Association.