Adaptation of Evolutionary Algorithms for Decision Making on Building Construction Engineering *(TSP Problem)*

Raniyah A. Wazirali, Arwa D. Alzughaibi, and Zenon Chaczko

Abstract—The report revolve on building construction engineering and management, in which there are a lot of requirements such as well supervision and accuracy and being in position to forecast uncertainties that may arise and mechanisms to solve them. It also focuses on the way the building and construction can minimise the cost of building and wastages of materials. The project will be based of heuristic methods of Artificial Intelligence (AI). There are various evolution methods, but report focus on two experiments Pattern Recognition and Travelling Salesman Problem (TSP).

The Pattern Recognition focuses Evolutionary Support Vector Machine Inference System for Construction Management. The construction is very dynamic are has a lot of uncertainties, no exact data this implies that the inference should change according to the environment so that it can fit the reality, therefore there a need of Support Vector Machine Inference System to solve these problems. TSP focus on reducing cost of building construction engineering and also reduces material wastages, through its principals of finding the minimum cost path of the salesman.

Keywords—TSP, genetic algorithms, GA, Support Vector Machines, SVM

I. Introduction

IN the construction engineering is involve of a lot of disparate activities, which usually relate to one another and have impacts to one another. The construction building is affected by various uncertainties, such as geological aspects, weather, aspect of human judgement and market fluctuation. Due to these facts professional construction engineering profession is very important. The construction is a process of achieving construction objective through application of available materials or resources. Due to uncertainties in the construction engineering, practical engineering aspect is very complicated and ill structured [1]. The process of determining mathematical model to solve this problem is very complicated and costly at the same time. The viable alternative is use of proximate inference which is fast and not expensive. Due to the fact that information changes the inference process should change accordingly. The construction is very dynamic are has a lot of uncertainties, no exact data this implies that the inference should change according to the environment so that it can fit the reality. In this case it must emulate the human mind aspect, where human has the capability to solve complex problems, even in time of uncertainty, imprecision and incomplete [1].

The aspect of inference process to emulate the human brain is the best solution to construction engineering. The artificial intelligent (AI) which is branch of computer science design can be adopted by inference process. AI emulates the human mind in all aspects and solves things intelligently using logics and pattern recognition [2].

Support Vector Machines (SVM) and fast messy genetic algorithms (fm GA) are applied to solve this construction problem. SVM can be trained by the user by selecting a kernel function such as linear, polynomial.

II. Support Vector Machines Approach

On the SVM the learning theory has brought alternative training technique of polynomial, radial based function classifiers. SVM is oriented on structural risk minimisations (SRM) reduction principle [3] which tries to mitigate generalised errors to limited defined boundary. For the classification of the case, RVM determines s separate hyper plane that mitigate the margin between the two classes. Maximisation of the margin is a quadratic programming (QP) problem that can be tackled from its two problems by incorporating Lagrangian Multipliers. The use of the linear programming is very important and vary innovative due to its flexibility when used in large dataset.

In most cases to identify a suitable hyper plane in input space is very challenging fact. The solution to this problem is to map the input space into higher level feature space and then try to identify the optimal hyper plane on the featured space. When there is no knowledge of mapping the SVM uses the dot product functions in the feature termed as kernel. The kernel based on mercer theorem is used in SVM to map data from the input to higher dimensional spaces [3]. The simple functions defined on pairs of input data patterns are used in computing dot product and a linear decision area.

The function definition:

Input space X, if there is a mapping $\phi : X \to H$ that maps any $x, z \in X$ into Hilbert space H then a kernel, $K : X \times X \to R$, is constructed as $K(x, z) = \langle \phi(x), \phi(z) \rangle H$, where $\langle \cdot, \cdot \rangle H$ is the scalar product operator in H. Kernel function k is the one that satisfy the mercer condition, the kernel matrix is created

R. A. Wazirali is with the University of Technology (UTS) Sydney, NSW, Australia (e-mail: Raniyah.A.Wazirali@uts.edu.au).

A. D. Alzughaibi is with the Taiba University, Saudi Arabia, Almadinah (e-mail: Arwa.d.alzughaibi@student.uts.edu.au).

Z. Chaczko is with the University of Technology (UTS) Sydney, NSW, Australia (e-mail: zenon.chaczko@eng.uts.edu.au).

by limiting k to definite subset of $X \; is \; +VE$ semi defined and this make to be called the mercer kernel and mercer condition is very important to kernel design, when we apply regression function and we give it the following training data set $\{(x_1, y_1), \ldots, (x \; n, y \; n)\} C \; N \times K$ [4].

$$|\xi|_\varepsilon = \begin{cases} 0 & if \; |\xi| \le \varepsilon \\ |\xi| - \varepsilon & otherwise \end{cases} \quad (1)$$

N denote the space patterns R^{k}, ε SVM regression with the goal to get a function $f(x)$ the ε is the deviation from the real obtained target y_i in all trained data.

We may include error less than ε and we done not allow the error greater than ε (1) is the ε insensitive function.

The regression can be made non linear by through mapping training patters(x_i) via nonlinear transformation process $\varepsilon \; \Phi : N \to F$ to dimension feature (F) as shown in the EQ. 2 where best fitting is estimated in feature space F [4].

$$f(X) = w.\Phi(x) \pm b \quad (2)$$

In order to eliminate over fitting there is need to add capacity control. In the formal SVM regression model should be written as shown the (3).

$$\min_{w,b,\xi_i,\xi_i^*} \frac{1}{2}|w|^2 + C \sum_{i=1}^{N}(\xi_i + \xi_i^*)$$

$$\text{Subject to} \quad \begin{array}{l} y_i - (w^t < I > (x_i) + b) \le \varepsilon - \xi_i \\ (w^t < I > (x_i) + b) - y_i \le \varepsilon - \xi_i^* \\ \xi_i \xi_i^* \ge \forall i. \end{array} \quad (3)$$

The constant $C > 0$ is the one that determines the relationship between the complexity of $f(x)$ and the extent to which deviation greater than ε are accepted. That interesting thing about RVM is that it can be expressed into form of functions both linear and non linear [5]. There are other functions such as polynomial, radial basis (RBF) and sigmoid kernel. In the case of kernel parameters there is need to set them properly to improve forecast accuracy in building construction.

$$k(x_i, x_j) = (1 + x_i \cdot x_j)^d \quad (4)$$

Radial basis function kernel:

$$k(x_i, x_j) = \exp(-y|x_i - x_j|^2) \quad (5)$$

Sigmoid kernel:

$$k(x_i, x_j) = \tanh(k x_i \cdot x_j - \delta) \quad (6)$$

III. FAST MESSY GENETIC ALGORITHMS APPROACH

The fm GA is based on the approach of its flexibility, it emulate genes of chromosome which are represented in pairs allele value and allele locus. Allele locus represents gene position while allele represents the value of the gene in the position [6]. Such example is ((3 1) (1 0) (2 1) (4 1) (5 0)) and ((2 1) (5 0) (3 1) (1 0) (4 1)).

In this step, the SVM uses default parameters and a training dataset to train a prediction model.
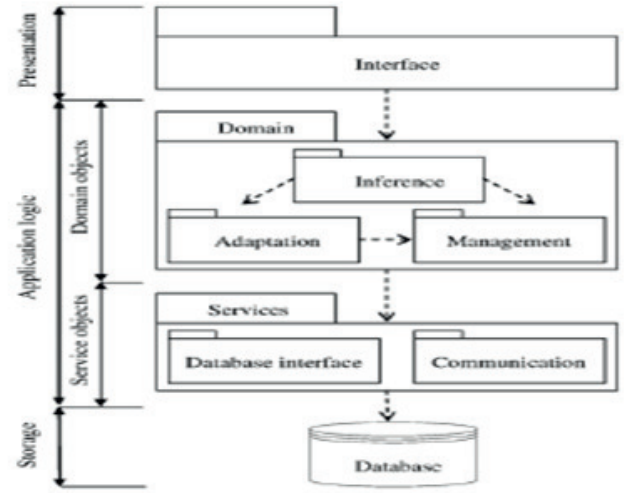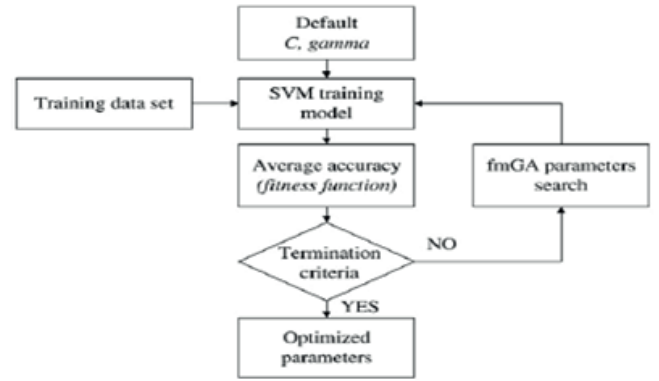

Fig. 1. Training SVM [4].

Fig. 2. RVM System Architecture [4].

IV. ADAPTATION OF TRAVELLING SALESMAN PROBLEM (TSP) TO REDUCE COST AND MATERIAL WASTAGES IN BUILDING CONSTRUCTION ENGINEERING

TSP focus on reducing cost of building construction engineering and also reduces material wastages, through its principals of finding the minimum cost path of the salesman. Due to complexity in the building construction and uncertainty due to various factors, a well designed mechanism is needed to tackle this problem [7]. The best option is to adopt TSP that is based on heuristic that emulate the human activities. TSP major on the aspect of a salesman and a set of cities. The salesman has to go or visit all the cities starting from one and return to back to the original city. The biggest challenge is how the salesman will minimise the aggregate travelling cost to visit all the cities [7]. The form definition of TSP is described as follows

$$TSP = \{(G, f, t) : \quad \begin{array}{l} G = (V, E) \text{ a complete graph,} \\ f \text{ is a function } V \times V \to Z, \\ t \in Z, \end{array} \quad (7)$$

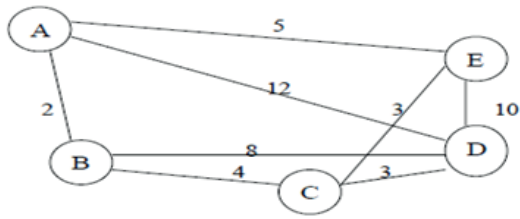The problem is how to find a minimal route passing from all the nodes. for example if you take path one from

Fig. 3. A Graph with Weights on Its Edges [7].

Fig. 4. Genetic Algorithm and Genetic Local Search [8].

$\{A, B, C, D, EA\}$ and the path two that is $\{A, B, C, E, D, A\}$ you have passed all the cities, but path one has a sum of twenty four and path two has a sum of thirty one. It implies that the path two has the longest distance and path one has the shortest distance. On Hamilton cycle, is a cycle in a graph that goes around in all nodes [7].

Definition:
$P = \{A, B, C, D, E\}$ is the Hamilton cycle. The issue is identifying cycle of Hamilton in a graph is NP-complete. Theorem of travelling salesman problem is NP complete.
Proof: Prove that TSP belongs to *NP*.

When we check a tour quality, then we check that the tour visit each vertex once. Then we have sum cost of the edges and finally we check whether the path is minimum path or has less cost. This can be termed as completed polynomial time, which implies TSP belongs to NP (The Travelling Salesman Problem, n.d).

This step is to prove or show that TSP is NP hard. To prove this is to show that Hamilton cycle $\leq_p$ TSP in the sense that Hamilton cycle problem is NP complete. Let assume that $G = (V, E)$ to be an instance cycle of Hamilton then construct an instance. The creation of complete graph to prove is needed and is $G' = (V, E')$, where $E' = \{(i, j) : i, j \in V$ and $i \neq j\}$, hence the cost function is defined as:

$$t(i, j) = \begin{cases} 0 & \text{if } (i, j) \in E, \\ 1 & \text{if } (i, j) \notin E. \end{cases} \quad (8)$$

Suppose that a Hamiltonian cycle h exists in G. It is then clear that the cost of each edge is h is o in G' because each edge belong to E. In this regard G has a Hamilton cycle it implies g' has o tour cost., let assume then that G' has a tour h' of cost of at most zero. The cost of edges in E' is zero and one by definition. Therefore each edge must have a cost of zero as the cost of h' is zero. Then we can conclude that h' contains only edges in E [7].

In that case we have proven that G has a Hamilton cycle if G' has a tour of cost zero, hence TSP is NP complete.

V. TSP APPROACH USING EVOLUTIONALLY ALGORITHMS

TSP is classical NP hard combinatorial optimization problem and there is a lot studies that has been done on it. Taking n cities and costs or distances between two cities, then we want to get the minimum cost tour that will visit each city one tim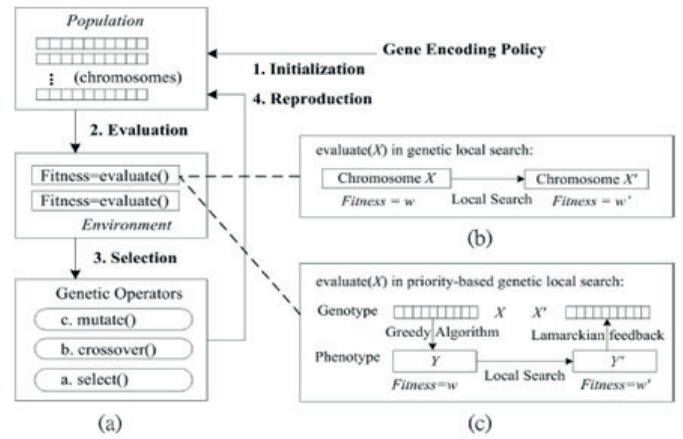e only. The assuming that $d_{i,j}$, as the cost of visiting city i and city j, then the TSP can be formulated so that we can get a permutation π of $\{1, 2, \ldots, n\}$ that is minimal [8].

$$C(\pi) = \sum_{i=1}^{n-1} d_{\pi(i), \pi(i+1)} + d_{\pi(n), \pi(1)} \quad (9)$$

(a) GA flowchart
(b) GLS is a combination of GA together with local search heuristics
(c) Priority-Based GLS and uses a greedy algorithm and also a Lamarckian feedback process to exchange between genotype and phenotype.

Regarding to symmetric TSP (STSP), $d_{i,j} = d_{j,i}$ for any two adjacent cities i and j, but in asymmetric the condition may not hold. The Euclidean TSP is a unique case of STSP, in this case cities are located in R^m distance for some m. and the cost obey the triangle inequality; $d_{i,k} + d_{k,j} \geq d_{i,j}$ for all unique i, j and k [8].

VI. CONCLUSION

The travelling salesman problem can be adopted to handle complicated work of construction engineering .the travelling salesman problem has capability to do what is needed in the construction industries with a lot of ease and maintain accuracy and associated risks. It has the capability of designing and implementing building construction projects at high professional level. The reason behind this is that it emulates human behaviours and way of reasoning. TSP focus on reducing cost of building construction engineering and also reduces material wastages, through its principals of finding the minimum cost path of the salesman.

Genetic Algorithms are important in building construction due to these facts professional construction engineering profession is very important. The construction is a process of achieving construction objective through application of available materials or resources. Due to uncertainties in the construction engineering, practical engineering aspect is very complicated and ill structured and can be solved well by Genetic Algorithms.

REFERENCES

[1] H. Li, "Case-based reasoning for intelligent support of construction negotiation," *Information & Management*, vol. 30, no. 5, pp. 231–238, 1996.

[2] I. Mareels and J. W. Polderman, *Adaptive Systems: An Introduction*. Boston, Massachusetts: Birkhauser, 1996, pp. 1–3.

[3] L. Chun-fu, "Fuzzy support vector machines," Ph.D. dissertation, Departament of Electrical Engineering, National Taiwan University, Taipei, Taiwan, 2004.

[4] B. Yang and N. Yau, "Integrating case-based reasoning and expert system techniques for solving experience-oriented problems," *Journal of the Chinese Institute of Engineers*, vol. 23, no. 1, pp. 83–95, 2000.

[5] C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121–167, 1998.

[6] J. Lin and C. Hsu, "A simple decomposition method for support vector machine," *Machine Learning*, vol. 46, no. 1-3, pp. 219–314, 2002.

[7] T. Travelling Salesman Problem, n.d.

[8] J. Wei, "Approaches to the Travelling Salesman Problem Using Evolutionary Computing Algorithms," 2008.