*Computer-Aided Civil and Infrastructure Engineering* **16** (2001) 79–93

# A Hybrid AI-Based System for Site Layout Planning in Construction

Emad Elbeltagi & Tarek Hegazy*

*Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1*

**Abstract:** *The layout of temporary facilities on a construction site is a complex and experience-based process that aims at minimizing the travel distance among facilities and improving site productivity and safety. To address the difficulties associated with the process, this article presents a hybrid system for construction-site layout planning. The originality of the proposed system stems from its integration of three components, each is developed based on a different artificial intelligence (AI)–based technique: (1) a knowledge base that stores and uses information related to identifying necessary facilities and their sizes, (2) a fuzzy quantifier to address the vagueness inherent in the project manager's assessment of facilities' closeness relationships, and (3) an improved genetic algorithm to search for an optimal layout solution. Details of the system are described, and a case study is used to demonstrate its capabilities. Possible extensions are then outlined.*

## 1 INTRODUCTION

The basic consideration in an effective site layout plan is the smooth and low-cost flow of materials, labor, and equipment within the site, in addition to satisfying various work constraints and safety requirements. In general, layout planning can be viewed as a complex optimization problem that has many engineering applications ranging from the layout of manufacturing plants to the design of computer chips. Since the early 1960s, the problem has been analyzed extensively in the industrial engineering (IE) and operational research (OR) communities. However, due to the unique nature of construction sites, models have been developed specifically for the construction domain. Early models were based solely on mathematical optimization techniques and, due to the complexity of problem formulation, were successful in laying out only a single or a limited number of facilities. An example is the optimization model of Warszawski and Peer[25] to place storage facilities and construction equipment on the site, with the objective of minimizing transportation, maintenance, and installation costs. Another example is the model of Rodriguez-Ramos and Francis[22] to position a crane within a construction site.

Heuristic approaches and knowledge-based systems also have been used to solve larger size problems of site layout. Hamiani[10] and Tommelein,[24] for example, used a rule-based system to place temporary facilities on site, one at a time, through a constraint-satisfaction search. Cheng[3] also used a similar knowledge-based approach linked to a geographic information system (GIS) to provide a dynamic layout solution. In general, heuristic solutions attempt to satisfy spatial relationships among facilities and have been reported to produce good but not optimal solutions.[28]

Due to the complexity of the site layout problem, nontraditional optimization techniques based on artificial intelligence (AI) have been applied to solve the problem. Yeh,[26] for example, used a simulated annealing neural network to find an optimal site layout by having a set of predefined facilities continuously exchange their positions within a predetermined set of locations on the site until a layout of minimum cost is obtained. Other researchers[12,15,18] have used the powerful random-search capabilities of genetic algorithms (GAs) to search for optimal solutions. While each GA model involved a different formulation, all studies reported the benefits of the GA technique in large-scale problems, for which traditional mathematical optimization is likely to fail.

In most construction-related models, such as those described earlier, the site and the facilities are often limited

*To whom correspondence should be addressed. E-mail: *tarek@uwaterloo.ca.*

to rectangular shapes, with each facility represented by a single block. Also, facilities are assumed to be predetermined, and the models provide few guidelines regarding identifying necessary facilities and their area requirements. Furthermore, decisions related to the desired interrelationships among facilities are left to the user, and the models do not incorporate mechanisms to address the fuzziness of such decisions. These issues represent modeling difficulties that need to be addressed in a realistic site layout planning model.

In an attempt to address the preceding challenges, this article presents a hybrid system for construction-site layout planning. The developments made in this article build on a previous basic model developed by Hegazy and Elbeltagi in 1999.[12] The basic model required user inputs of all facilities and their relationships and included a rudimentary GA for layout optimization. To add practicality and supplement the user's subjective judgment, three main components have been incorporated into the basic model to form the hybrid system presented in this article: (1) a knowledge base for facility identification and area determination, (2) a fuzzy quantifier to address the vagueness in the assessment of facilities' interrelationships, and (3) a modified GA to speed the search for an optimal layout solution. Details of these components are described along with their implementation. A prototype of a hybrid system is then demonstrated and its performance validated.

## 2 PROPOSED HYBRID SYSTEM

### 2.1 Model components

The proposed hybrid model aims at supporting the three most difficult decisions related to site layout planning, with each lending itself well to a different solution mechanism. First, the process of identifying the necessary temporary facilities for a construction site and their appropriate areas is mainly knowledge-dependent, and as such, a knowledge-based system can aid in this process. Second, determining a crisp assessment of the desirability of having facilities close or apart from each other is a problem that involves vagueness and lends itself well to fuzzy-set application. Third, the optimization of facility placement on the site is a difficult problem that lends itself well to genetic optimization. Integrating these three components together forms a hybrid AI-based system for construction-site layout planning (Figure 1), with each component dealing with one of the subproblems using a different AI-based technique. To facilitate the integration among these three components, the unified representation of a site and its facilities used in the earlier basic model[12] (discussed in the next subsection) is maintained. Also, to build on previous developments, Microsoft Excel software[8] is used for implementing the hybrid system to take advantage of its ease of use, simple interface, and powerful programmability features.

### 2.2 Site and facility representation

Any irregular site shape is modeled in this study using a two-dimensional grid. A facility is also modeled using a number of adjacent grid units (as opposed to a single block representation). The area of each grid unit is calculated as the greater common divisor (GCD) of all facilities' areas, and the position of a facility on the site is defined by its location reference on the grid (Figure 2). The location reference is formulated using the column and row boundaries of the site and is used to define the starting position at which a facility is to be placed on the site, as follows:

$$\text{Location reference} = (\text{row position} - 1) \times \text{total columns} + \text{column position} \tag{1}$$

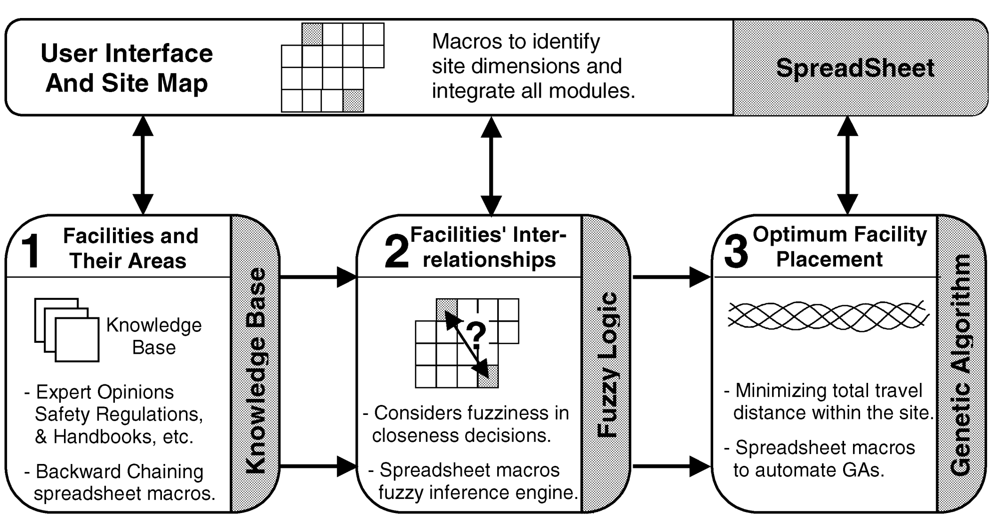


**Fig. 1.** Components of the proposed model.

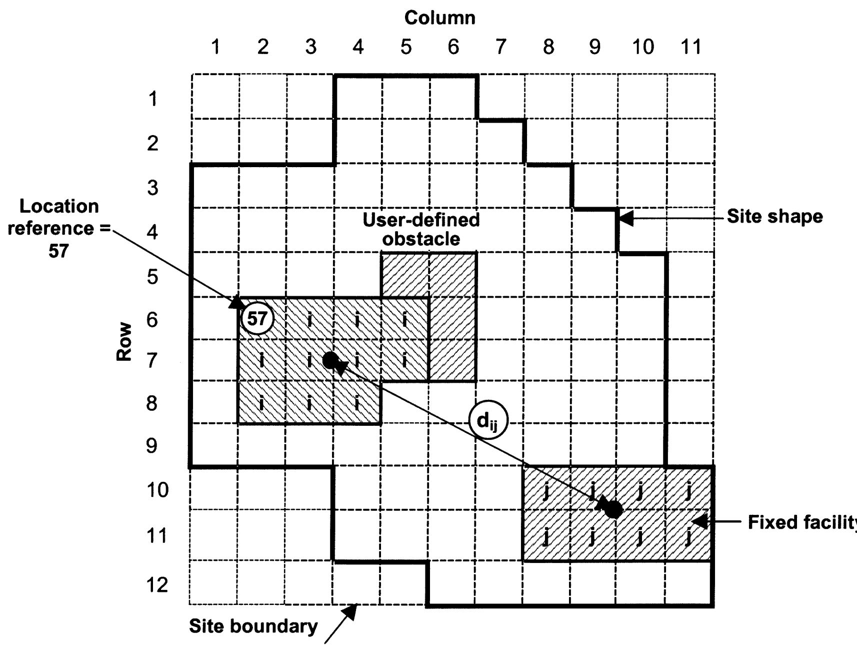


**Fig. 2.** Site and facility representation.

For flexibility, the present model deals with three types of facilities: normal facilities, fixed facilities, and obstacles (see Figure 2). The difference between a fixed facility and an obstacle is that the former can have defined relationships with other facilities, such as an existing building, a subproject, or an area on the site where some activity is being conducted. A facility can be placed on the site grid, starting from its location reference, in three ways: horizontal, vertical, or rectangular.[12]

Having defined the site and facility representations, the three AI-based components of the proposed hybrid system and their implementation are discussed in the following sections.

## 3 FACILITY IDENTIFICATION

Identifying the required temporary facilities for a specific project and determining their areas are difficult decisions that require thorough consideration of the project conditions and local regulations. In current practice, layout planning is often done in a speedy manner by adjusting previous plans based mainly on the project manager's experience and common sense.[4] In many situations, some facilities that are required by local bylaws, such as a site first-aid office or a fire route, may be omitted. Accommodating these facilities later can be costly and can cause loss of site productivity. While some information related to facility identification has been documented in the literature based on site characteristics and safety requirements,[3,10,11,19,20,24] this information, however, is scattered and is not readily usable.

To facilitate and automate facility identification, a knowledge-based module has been developed as an integrated component of the proposed model (component 1 of Figure 1). Similar to most knowledge-based expert systems (KBES), it consists of (1) a knowledge base, expressed in a set of IF-THEN rules, (2) a context, which describes the values of the problem attributes (factors affecting facility determination), and (3) an inference engine, which is a mechanism for searching the knowledge base to continuously modify the context until a solution is reached.

The IF-THEN representation of the knowledge related to facility identification and sizing is suitable and valid. As shown in Figure 3, the decision on the area of a batch plant depends on several factors, including the quantity of concrete, site space, availability of cheaper alternatives, and potential of plant reuse. This knowledge, as expressed in Figure 3, takes a hierarchical form suitable for rule-based representation. Also, the separation of the knowledge base from the processing mechanism, which is a main benefit of KBES, also simplifies the updating of the knowledge base.

In order to develop the knowledge base for the facility identification, 139 rules were compiled from different sources, including construction safety and health manuals,[13] company handbooks,[17] published dissertations,[3,10,24] and technical articles.[11,19,20] The knowledge-base rules correspond to 22 temporary facilities (Table 1) that possibly can be used on a construction site. The rules determine the size of these facilities based mainly on the personnel requirements, estimated quantity of work, production rate of resources, availability of site space, and cost.

**Table 1**
A list of temporary facilities

| *No.* | *Facility name* | *No.* | *Facility name* |
|---|---|---|---|
| 1 | Offices | 12 | Batch plant |
| 2 | First aid | 13 | Sampling/testing lab |
| 3 | Information and guard | 14 | Piping yard |
| 4 | Toilet on site | 15 | Parking lot |
| 5 | Engineer's/staff dormitory | 16 | Tank |
| 6 | Labor's dormitory | 17 | Long-term lay-down yard |
| 7 | Labor's rest area | 18 | Machine room |
| 8 | Maintenance shop | 19 | Shops |
| 9 | Rebar fabrication/storage yard | 20 | Scaffold storage yard |
| 10 | Carpentry shop | 21 | Material warehouse |
| 11 | Cement warehouse | 22 | Welding shop |

## 3.1 Inference mechanism

A backward-chaining mechanism is used to search the knowledge base for solutions. The mechanism starts from a goal that requires a solution and works backward to determine its subgoals. Consider the chain of knowledge in Figure 3, for example; the goal to be determined (right side) is the area of the batch plant. Before a final conclusion is made, the subgoals have to be determined first. The process moves backward to evaluate the subgoals and their sub-subgoals following the knowledge chain. Following this process, the rules (numbered 1 to 6 in Figure 3) fire sequentially, and accordingly, the area for the batch plant is determined.

## 3.2 Implementation

To facilitate the integration of the knowledge-base component into the hybrid system, this component was modeled in a spreadsheet format. Backward chaining was coded using a combination of spreadsheet functions and VBA macros that underlay simple interface elements for updating the knowledge base (Figure 4) and for consultation. The IF-THEN rules also were put in a separate worksheet to facilitate manual updating. During consultation, the 22 facilities in Table 1 one by one are considered as goals. The backward-chaining subroutine then searches the knowledge base and asks the user questions relevant to the rule being considered. Accordingly, one or more rules will fire, determining the area requirement of a facility (or zero

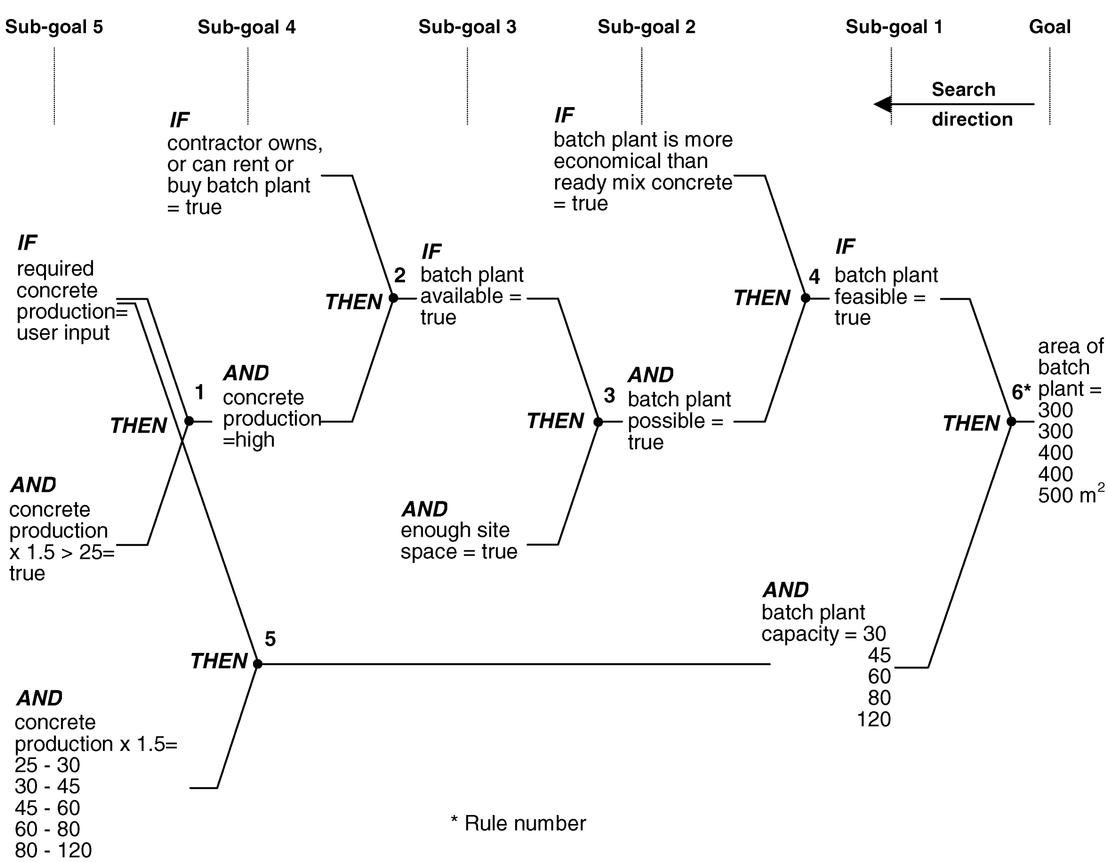


**Fig. 3.** Knowledge representation and the backward-chaining process.

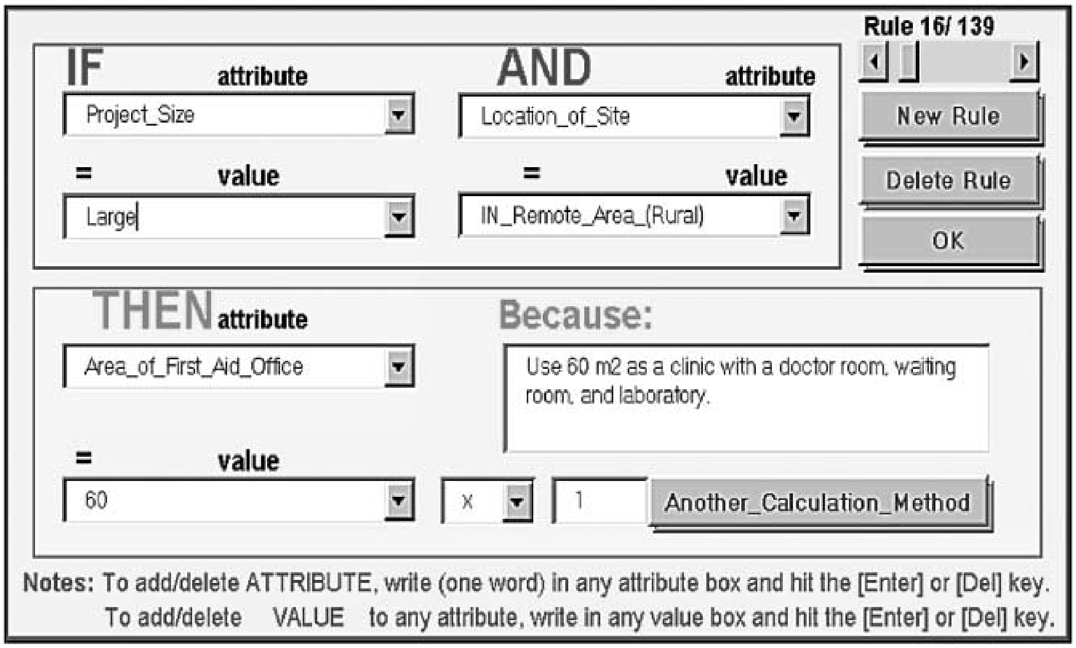


**Fig. 4.** Knowledge-base edit screen.

if the facility is not required) and then proceeding to the next goal. A typical consultation session on the 22 goals (facilities) takes about 3 minutes (assuming reasonable user response time to the questions during consultation).

## 4 FACILITIES' CLOSENESS RELATIONSHIPS

The objective function in site layout planning is to minimize the total travel distance within the site, which is a direct function of the desired closeness relationships among facilities.[12] These relationships represent the user's preference in having the facilities close or apart from each other. Such relationships can be expressed as user-specified weight values, representing his or her subjective assessment of the desired relationship between each two facilities, or can be the actual amount of material, equipment, or personnel flow between the two facilities. Often, however, the user's decisions on facilities' relationships involve some degree of fuzziness and ambiguity. In large, complicated sites, the user may be able to specify facilities' relationships only in the form of linguistic expression such as "as far as possible" or "very close." It is difficult, therefore, to accurately quantify a closeness value (weight) between each two facilities. To address this problem, a fuzzy quantifier (component 2 of Figure 1) has been developed.

The development of this component used the concept of fuzzy-set theory originated by Zadeh[27] and the concepts of fuzzy control developed by Mamdani.[16] Fuzzy decision making has been applied successfully in many areas such as project scheduling,[14] evaluating alternative construction technologies,[2] and industrial plant layout.[7] Background material on fuzzy-set theory can be found in many references (for example, Ross[23]).

### 4.1 Fuzzy output

In site layout planning, a fuzzy linguistic variable "closeness rating" ($R$) is a fuzzy variable that represents the output variable. This linguistic variable can be represented by a family of linguistic terms (fuzzy sets $A$, $E$, $I$, $O$, $U$, and $X$, as shown in Figure 5). These six fuzzy sets cover the space of closeness-rating solutions ranging from "absolutely important" for $A$ to "undesirable" for $X$. Each of these six sets (e.g., set $A$) has a triangular membership function, with some overlap among them, as shown in Figure 5. It is noted that the ranges shown in Figure 5 for the different membership functions (e.g., set $A$ ranges from 2500–7500) were designed to exhibit an exponential increase in the closeness values. This gives high weight values when facilities are required to be close to each other (e.g., set $A$), thus enforcing this relationship later during layout optimization. These membership functions are used to quantify a crisp value for the closeness relationship between each two facilities, as discussed in the following subsections.

### 4.2 Fuzzy inputs

In plant layout planning, various factors have been considered by researchers to affect the closeness rating among departments, including equipment flow, material flow, personnel flow, and information flow.[7] In construction sites,

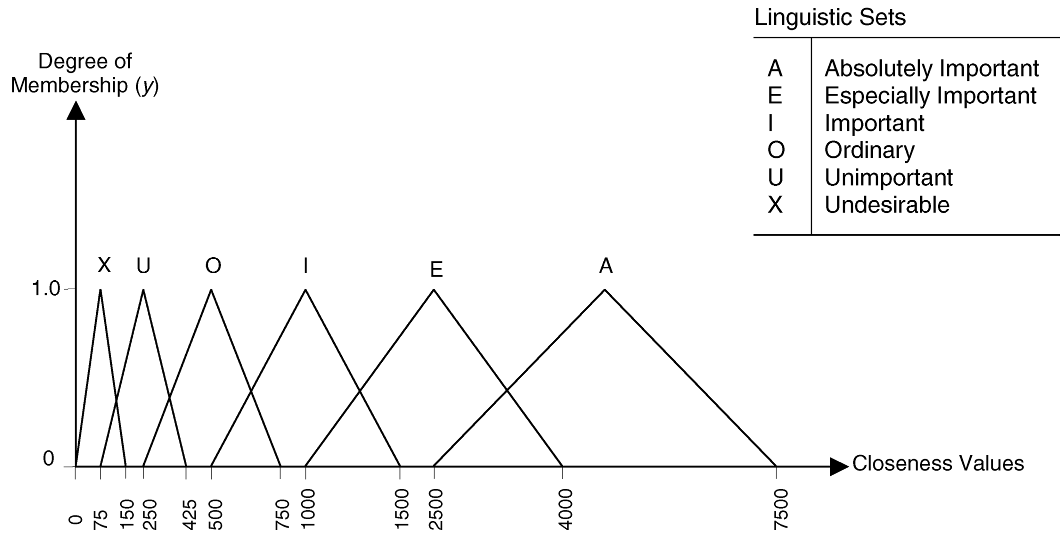


**Fig. 5.** Fuzzy sets for the output variable "closeness rating."

however, other important factors used by experts in site layout planning come into play, such as safety/environmental concerns, supervisory requirements, and user preference. In the present development, therefore, three factors have been considered in determining the closeness weights between each two facilities: (1) the level of work flow between the two facilities, (2) the level of safety/environmental hazard, and (3) the user's preference. The work flow between two facilities greatly affects site productivity and encompasses the total flow of material, equipment, personnel, and information between the two facilities. The level of safety/environmental hazard also represents any concerns that may arise when the two facilities are close to each other that may affect site workers by increasing the likelihood of accidents, noise, uncomfortable temperature, and/or pollution. The third factor represents the project manager's desirability of having the facilities close or apart from each other. This factor becomes important when the project manager desires to have the two facilities close to each other despite the little or no work flow between them.

Based on the preceding discussion, the problem at hand involves three fuzzy input variables: "work flow" (*WF*), "safety/environmental concerns" (*SE*), and "user's preference" (*UP*). These three variables affect the "closeness rating" fuzzy output variable identified earlier. A family of fuzzy sets has been formulated for the three fuzzy variables, and for simplicity, each variable was limited to three membership functions "low" ($L$), "medium" ($M$), and "high" ($H$) (Figure 6). The shape and range of values of the three membership functions ($L$, $M$, and $H$) were determined through experimentation. Accordingly, triangular and trapezoidal shapes were adopted (Figures 5 and 6). These two shapes are also the most frequently used in the literature.[5,6]

The values shown on the membership functions are proposed by the writers based on their experience. Work flow is assumed to vary from 0 to 200 daily trips of material, equipment, personnel, and information. The shape of the *WF* membership functions is symmetrical and centers around 100 trips per day. Modifying the membership function values requires an intensive survey among practitioners, which is the subject of future research by the writers. The *SE* membership functions, on the other hand, are biased toward the "high" safety/environmental concerns. Therefore, while the *SE* indicator ($x$ axis) has a range from 0 to 10, the "high" membership function has values starting from 4 to 10. The third membership function *UP* is similar to that of the *WF*, but with a different scale, ranging from 0 to 10.

### 4.3 Fuzzy decision rules

So far the "closeness rating" desired to be determined is governed by three fuzzy variables, *WF*, *SE*, and *UP*. Since each of these fuzzy variables has three membership functions $L$, $M$, and $H$, there could be a total of $3^3$ (27) different combinations of preconditions that affect the closeness rating. These preconditions have to be stored in the form of rules (called *fuzzy rules*) along with the decision maker's preference in their associated closeness rating. An example rule is

Rule 6:

IF Work flow (*WF*) is low ($L$)
AND Degree of safety (*SE*) is medium ($M$) (2)
AND User's preference (*UP*) is high ($H$)
THEN Closeness rating ($R$) is important ($I$)

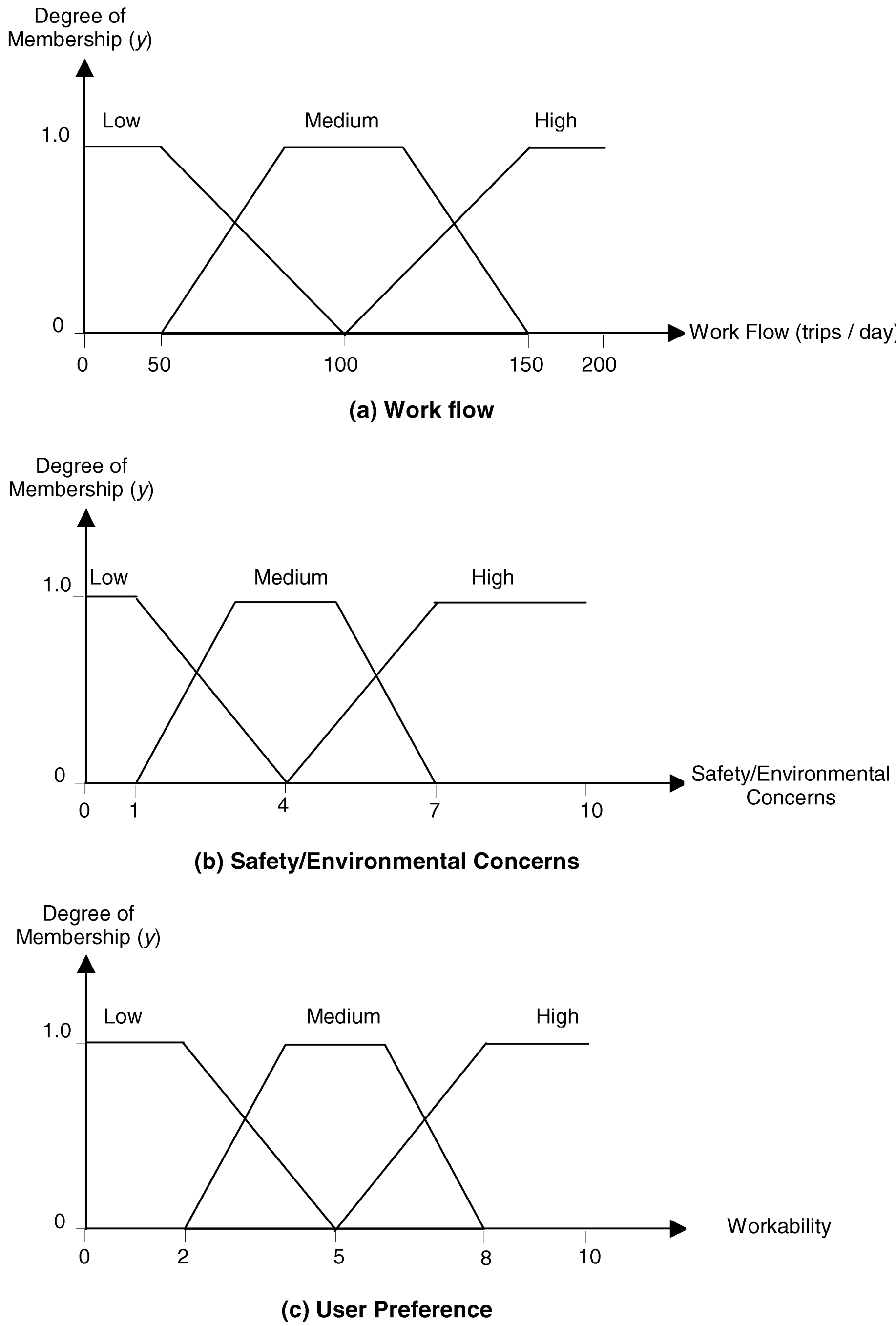


**Fig. 6.** Fuzzy sets of the input variables.

As shown in this rule, the THEN part refers to one of the six membership functions associated with the fuzzy output variable "closeness rating."

In developing the fuzzy rules for the problem at hand, a systematic approach was used to determine the appropriate membership function ($A$, $E$, $I$, $O$, $U$, or $X$) to associate with the three preconditions of each rule. For each input variable, a score of 3, 2, or 1 was given to the "high," "medium," or "low" linguistic term, respectively, of the *WF* and *UP* variables. On the other hand, a score of 1, 2, or 3 was given to the "high," "medium," or "low" linguistic term, respectively, of the *SE* variable. Considering the fuzzy rule in Equation (2), for example, the three preconditions of the rule have a total score of 6 (1 for $WF+2$ for $SE+3$ for *UP*). Once the total score was calculated, it was compared with a preset value of 3, 4, 5, 6, 7, and 8 that

**Table 2**
Fuzzy decision rules

| *Rule no.* | *Work flow* | *Safety/environmental concerns* | *User's preference* | *Closeness rating* |
|---|---|---|---|---|
| 1 | Low ($L$) | Low ($L$) | Low ($L$) | Ordinary ($O$) |
| 2 | Low ($L$) | Low ($L$) | Medium ($M$) | Important ($I$) |
| 3 | Low ($L$) | Low ($L$) | High ($H$) | Especially important ($E$) |
| 4 | Low ($L$) | Medium ($M$) | Low ($L$) | Unimportant ($U$) |
| 5 | Low ($L$) | Medium ($M$) | Medium ($M$) | Ordinary ($O$) |
| 6 | Low ($L$) | Medium ($M$) | High ($H$) | Important ($I$) |
| 7 | Low ($L$) | High ($H$) | Low ($L$) | Undesirable ($X$) |
| 8 | Low ($L$) | High ($H$) | Medium ($M$) | Unimportant ($U$) |
| 9 | Low ($L$) | High ($H$) | High ($H$) | Ordinary ($O$) |
| 10 | Medium ($M$) | Low ($L$) | Low ($L$) | Important ($I$) |
| 11 | Medium ($M$) | Low ($L$) | Medium ($M$) | Especially important ($E$) |
| 12 | Medium ($M$) | Low ($L$) | High ($H$) | Absolutely important ($A$) |
| 13 | Medium ($M$) | Medium ($M$) | Low ($L$) | Ordinary ($O$) |
| 14 | Medium ($M$) | Medium ($M$) | Medium ($M$) | Important ($I$) |
| 15 | Medium ($M$) | Medium ($M$) | High ($H$) | Especially important ($E$) |
| 16 | Medium ($M$) | High ($H$) | Low ($L$) | Unimportant ($U$) |
| 17 | Medium ($M$) | High ($H$) | Medium ($M$) | Ordinary ($O$) |
| 18 | Medium ($M$) | High ($H$) | High ($H$) | Important ($I$) |
| 19 | High ($H$) | Low ($L$) | Low ($L$) | Especially important ($E$) |
| 20 | High ($H$) | Low ($L$) | Medium ($M$) | Absolutely important ($A$) |
| 21 | High ($H$) | Low ($L$) | High ($H$) | Absolutely important ($A$) |
| 22 | High ($H$) | Medium ($M$) | Low ($L$) | Important ($I$) |
| 23 | High ($H$) | Medium ($M$) | Medium ($M$) | Especially important ($E$) |
| 24 | High ($H$) | Medium ($M$) | High ($H$) | Absolutely important ($A$) |
| 25 | High ($H$) | High ($H$) | Low ($L$) | Ordinary ($O$) |
| 26 | High ($H$) | High ($H$) | Medium($M$) | Important ($I$) |
| 27 | High ($H$) | High ($H$) | High ($H$) | Especially important ($E$) |

relates to the use of membership functions $X$, $U$, $O$, $I$, $E$, and $A$, respectively. Following this process, the fuzzy rules were formulated as shown in Table 2.

### 4.4 Determining facility closeness using a fuzzy rule–based system

With the membership functions and fuzzy rules formulated, it is possible to use them with specific values of the input variables (numeric, not linguistic) to compute a numeric value of the output variable. This process is known as *fuzzy rule–based inferencing*. Figure 7 shows the typical steps used in a fuzzy rule–based system[2] to define the facilities' closeness relationships.

As shown in Figure 7, the process first requires the user to input numeric values for the *WF*, *SE*, and *UP* variables between each two facilities. The process then fuzzifies these values through the membership functions of the input variables. For example, assume the user inputs *WF*, *SE*, and *UP* values 90, 3, 6, and respectively, between two specified facilities. These values are applied on the 27 rules, one by one, to determine the firing strength of each rule and how much it contributes to the output value. Figure 8 shows the calculations in an intermediate rule (rule 6), which was described in Equation (2). According to the rule, the *WF* value of 90 was applied to its $L$ membership function, the *SE* value of 3 was applied to its $M$ membership function, and the *UP* value of 6 was applied to its $H$ membership function. The intersection of these values with the membership functions provided membership values $\omega 1, \omega 2$, and $\omega 3$ of 0.2, 0.67, and 0.33, respectively (see Figure 8). The firing strength of that rule is then calculated using the minimum operator, which is the smallest of $\omega 1, \omega 2$, and $\omega 3$ membership values of the rule (0.2).

The determined firing strength (0.2) was used to truncate the membership function for the output, thus forming the shaded area (Area6) that defines the contribution of this rule to the overall output. Once these calculations are completed for all rules, the union operator is used to aggregate the consequences (Area1 to Area27) of the 27 rules to form an overall membership function (see Figure 8). This

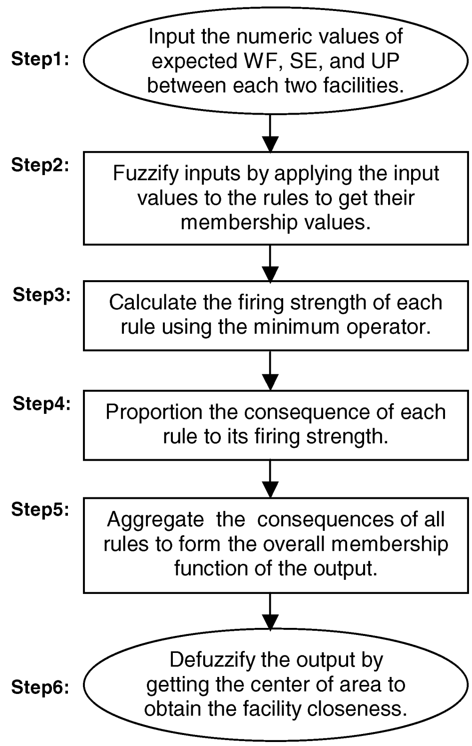


**Fig. 7.** Steps of fuzzy rule–based calculations.

overall membership function is then converted into a crisp (nonfuzzy) value through a defuzzification process. Various methods can be used to defuzzify the overall membership function, among which the center-of-area method is most common. Using this method, a crisp value of 2100 was obtained. In a similar fashion, a closeness rating between any two facilities can be calculated based on user input of the *WF*, *SE*, and *UP* values.

### 4.5 Implementation

The fuzzy quantifier was modeled on a spreadsheet that is set up with cell formulas encoding all computations as a function of a group of cells containing the values of the user input variables. Fixed data related to the fuzzy membership functions and the fuzzy rules were put in a separate worksheet to be used for the calculations. Also, VBA macros were written to guide the consultation and acquire user inputs. During its use, the user is prompted to enter the daily number of trips (work flow), any safety/environmental concerns (a value ranging from 0 to 10, where 0 means no concerns and 10 means completely unsafe), and the user's preference (a value ranging from 0 to 10, where 0 means far apart, 10 means close, and 5 means indifferent) between each two facilities. Accordingly, a closeness rating value between the two facilities will be calculated automatically and used in site layout optimization.

## 5 IMPROVED GENETIC OPTIMIZATION

Having defined the required facilities with their areas and closeness relationships, a genetic algorithm (GA) procedure (component 3 of Figure 1) has been used to optimally place facilities on site. The procedure is a random search for the optimal location reference of each nonfixed facility so that closeness relationships are optimally maintained. A representative score for the total travel distance associated with a layout is calculated based on the $d_{ij}$ distances among the facilities (see Figure 2), as follows:

$$\text{Objective function} = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} d_{ij} R_{ij} \tag{3}$$

where $R_{ij}$ is the desired closeness weight value between facilities $i$ and $j$, and $n$ is the total number of facilities (fixed + nonfixed).

The GA procedure used in this study is a modified version of that used in Hegazy and Elbeltagi,[12] with improved gene representation to reduce computations and speed the processing. This improved gene formulation has been a result of a different manner of dealing with fixed facilities. As opposed to the previous version, the positions of fixed facilities have been removed from the gene formulation, making the gene shorter and faster to process (Figure 9). Consequently, the GA code was modified to directly read the positions of the fixed facilities from the user-drawn site map. Fixed facilities are then kept in their place without taking part during GA optimization. As shown in Figure 9b, the modified gene structure has been set as a string of elements, each corresponds to the location reference of a nonfixed facility, and the gene length equals the number of nonfixed facilities to be placed on the site. Such improvement has resulted in a more robust GA and noticeable convergence speed. Full details of GA processing have not been included in this article but can be found in other references.[1,9,12]

### 5.1 Implementation

The revised GA was coded using the VBA macro language of Excel and then integrated with the spreadsheets of the proposed system. It accepts user input of the population size (number of genes) and the number of offspring generations. During processing, it draws possible solutions directly on the site map. All potential solutions are retained and sorted according to their score, and the user can view any desired solution.

## 6 PROTOTYPE

A prototype of the complete site layout planning system was compiled by integrating the KBES, the Fuzzy quantifier, and the modified GA. The development of the prototype involved substantial effort in coding and testing the

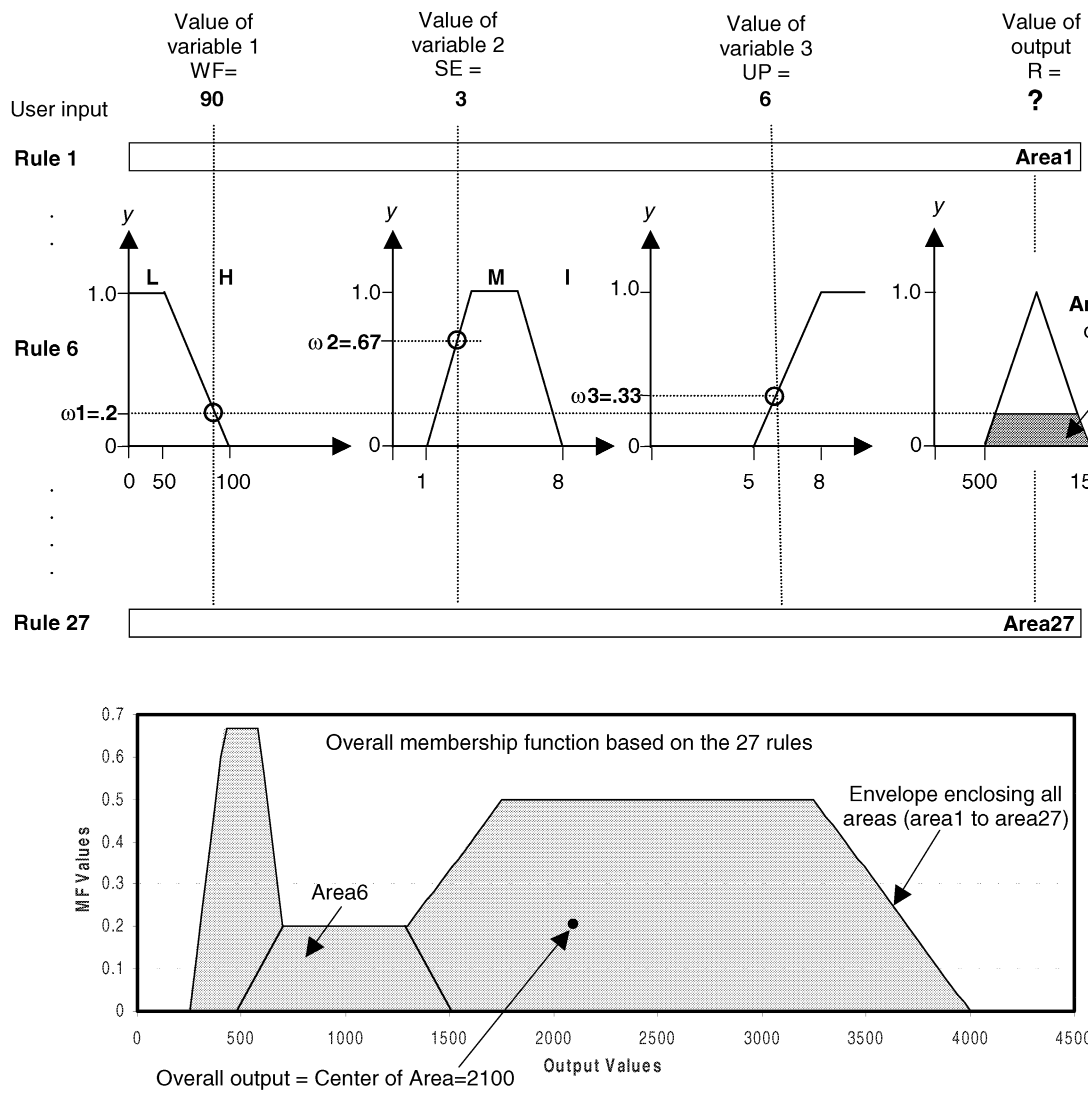


**Fig. 8.** Calculating the overall membership function for the output variable.

different components, developing a unified user interface, and experimenting with several case studies. The prototype is an Excel workbook that contains nine worksheets, including a main menu sheet (Figure 10), four sheets for facilities' data, facilities' closeness relationships, and site-map drawing, storage for possible solutions, and other sheets for internal calculations. To demonstrate the prototype capabilities, a case study is used, as described in the following subsection.

### 6.1 Prototype capabilities: A case study

Since no work in the literature can be used for comparison purposes of the knowledge base and the fuzzy quantifier, a hypothetical case study is used to demonstrate the capabilities of the developed system. The case study is a project with a site area of 3880 $m^2$. The user follows the steps shown in Figure 10. To identify the necessary facilities, a consultation session was conducted, and a total of 23 questions were asked; the responses entered to the system are given in Table 3. The data in Table 3 represent the project and site characteristics based on which the KBES identified the relevant temporary facilities needed for the site and their areas. At the end of consultation, the user was presented with the generated list of facilities and their areas (Figure 11), which he or she could modify.

The greater common divisor (GCD) of all facility areas (20 $m^2$) was then calculated automatically, and accordingly, each facility size is converted to site units and presented to the user for approval and/or modification. The user at this stage can set the desired maximum width and the preferred method of arrangement of each facility (horizontal, vertical, or rectangular). In addition, the user may set some of the facilities as fixed. In the present case study, four facilities were set as fixed. The final list of facilities (fixed first, sorted descendingly by facility area) and their data are shown in Table 4.

Once facilities are defined, the user proceeds with defining the closeness weights either directly using desired closeness weights among facilities or using the fuzzy logic option to calculate the closeness weights. When the fuzzy logic option is used, the user is prompted to enter the expected daily amounts of work flow, safety/environmental concerns, and user's preference among each two facilities. Accordingly, the closeness weights between the two facilities are computed automatically, and the process continues with other facilities. Following this process in the present case study, the closeness weights among the facilities were computed as shown in Figure 12. Notice that facilities "parking lot" and "offices," for example, are desired to be close to each other, whereas facilities "first aid" and "cement warehouse," for example, are desired to be far from each other. Afterwards, the site map was drawn on the spreadsheet using copies of the colored cell in Figure 13, where each cell is 20 m$^2$ ($\sim$ 4.5 $\times$ 4.5 m). Fixed facilities number 11, 8, 9, and 3 also were placed on the site map and their numbers written on the cells of their positions. Similarly, obstacles were drawn on the site map using copies of another colored cell. Afterwards, the site became ready for layout optimization.

At the start of the GA optimization, the user inputs the population size and number of offspring generations. For the present case study, a population of 100 genes and offspring genes of 200 were used. The evolutionary process proceeds through cycles of offspring generation and population improvement to search for optimal placement of facilities according to their required method of placement (horizontal, vertical, or rectangular). The score of each solution is then calculated automatically using Equation (3), which represents the total travel distance associated with the site layout. At the end of the process, the best solution is presented on the site map, and the user may view it or any of the other layout solutions saved during the evolutionary process. The final solution for the case study is shown in Figure 13 with its representative score of the total travel distance being 138,784.

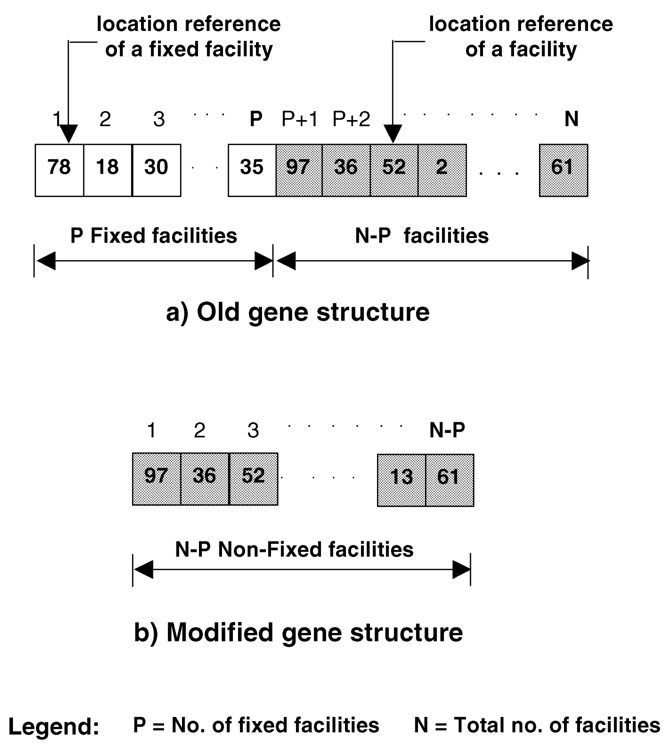


**Fig. 9.** Gene formation.

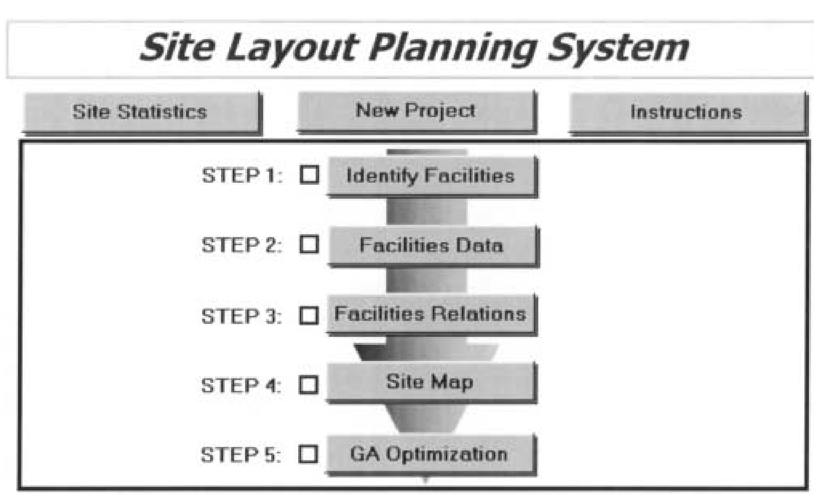


**Fig. 10.** Site layout planning steps.

**Table 3**
User input for facility identification and area determination

| *No.* | *Given question* | *Options given to user* | *User response* |
|---|---|---|---|
| 1 | Number of staff/engineers? | Input box | 3 |
| 2 | Number of owner's representatives? | Input box | 2 |
| 3 | Number of subcontractor's staff/engineers? | Input box | 0 |
| 4 | Number of laborers? | Input box | 50 |
| 5 | Number of mechanics? | Input box | 0 |
| 6 | Amount of daily reinforcing bars (tons)? | Input box | 3 |
| 7 | Peak amount of daily cement usage (tons)? | Input box | 25 |
| 8 | Project total cost ($ million)? | Input box | 6 |
| 9 | Maximum estimated hourly concrete requirements? | Input box | 15 $m^3$/h |
| 10 | Number of carpenters? | Input box | 0 |
| 11 | Site space availability? | Abundant, restricted | Abundant |
| 12 | Location of site? | Urban, suburb, rural | Suburb |
| 13 | Location of laborers' permanent residence from site? | Local, nearby, faraway | Local |
| 14 | Do you own cutting/bending machines? | Yes, no | No |
| 15 | Are preformed bars available and costly feasible? | Yes, no | Yes |
| 16 | Average duration steel bars stocked on site (weeks)? | No, 1, 2, 3 weeks | 3 weeks |
| 17 | Are lumber units fabricated on site? | Yes, no | No |
| 18 | Average duration forms stocked on site (weeks)? | No, 2, 4, 6 weeks | No |
| 19 | Average duration cement stocked on site (days)? | No, 7, 14, 21, 28 days | 21 days |
| 20 | Does the contractor own testing equipment? | Yes, no | Yes |
| 21 | Are steel workers required for this project? | Yes, no | No |
| 22 | Can you provide transportation to laborers? | Yes, no | Yes |
| 23 | Type of scaffolds. | Wooden, metallic | Metallic |

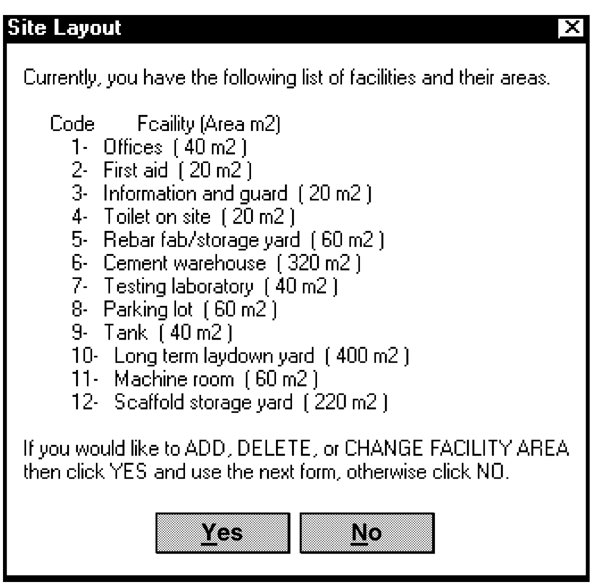


**Fig. 11.** The selected list of facilities and their areas.

## 7 COMMENTS AND FUTURE EXTENSIONS

The developed site layout planning system has been demonstrated to work effectively on the example applications presented in this article. Various other problems with different site shapes and different conditions were experimented with, and the system performed well. The current system can be used to construct a new layout, improve an existing layout, or optimally place a new facility in a given site layout.

There are a number of possible extensions to the system currently being pursued by us:

- Extending the study to dynamic site layout planning. This involves considering the changes in the facilities areas and locations throughout the project life.
- Integrating the current system with scheduling software to provide a direct link to the start and finish dates of project activities, thus automatically determining their time-dependent facility requirements.
- Incorporating multiobjectives in the optimization, including not only the travel distance between facilities but also an indication of the goodness of the resource circulation and movement around the empty spaces on the site.
- Creating an Internet application from the current system to enable users to benefit from it.
- Conducting a survey among construction practitioners to refine the membership functions in order to determine accurate closeness weights among facilities.

**Table 4**
The final list of facilities and their data

| *Facility code* | *Facility name* | *Area ($m^2$)* | *No. of units** | *Max. width (units)* | *Facility type* | *Method of arrangement†* |
|---|---|---|---|---|---|---|
| 11 | Machine room | 60 | 3 | N/A | Fixed | N/A |
| 8 | Parking lot | 60 | 3 | N/A | Fixed | N/A |
| 9 | Tank | 40 | 2 | N/A | Fixed | N/A |
| 3 | Information and guard | 20 | 1 | N/A | Fixed | N/A |
| 10 | Long-term lay-down yard | 400 | 20 | 4 | Normal | 1 |
| 6 | Cement warehouse | 320 | 16 | N/A | Normal | 2 |
| 12 | Scaffold storage yard | 220 | 11 | 2 | Normal | 0 |
| 5 | Rebar Fab./storage yard | 60 | 3 | 3 | Normal | 1 |
| 1 | Offices | 40 | 2 | 1 | Normal | 0 |
| 7 | Testing laboratory | 40 | 2 | 2 | Normal | 1 |
| 2 | First aid | 20 | 1 | 1 | Normal | 0 |
| 4 | Toilet on site | 20 | 1 | 1 | Normal | 0 |

*Number of units = area/GCD = area/20.
†0 = horizontal, 1 = vertical, and 2 = rectangular.

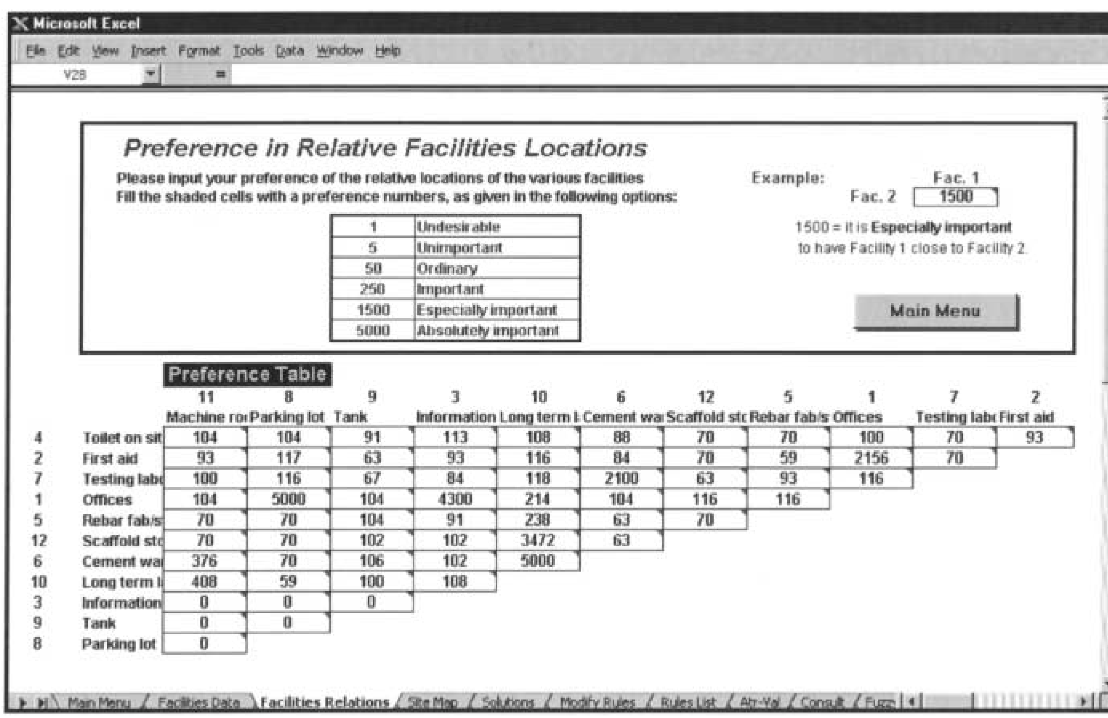


**Fig. 12.** Facilities' closeness relationship weights.

- Extending the knowledge base for facility identification and area determination.

## 8 SUMMARY AND CONCLUSION

In this article, a construction site layout planning system was developed incorporating a knowledge base to identify and size the required facilities on the site, a fuzzy quantifier to identify the facilities' closeness weights, and a modified genetic algorithm to optimally place facilities on the construction site. All system components proved to work efficiently in support of site layout planning decisions. The knowledge base stores expert knowledge related to the type and size of temporary facilities as a function of

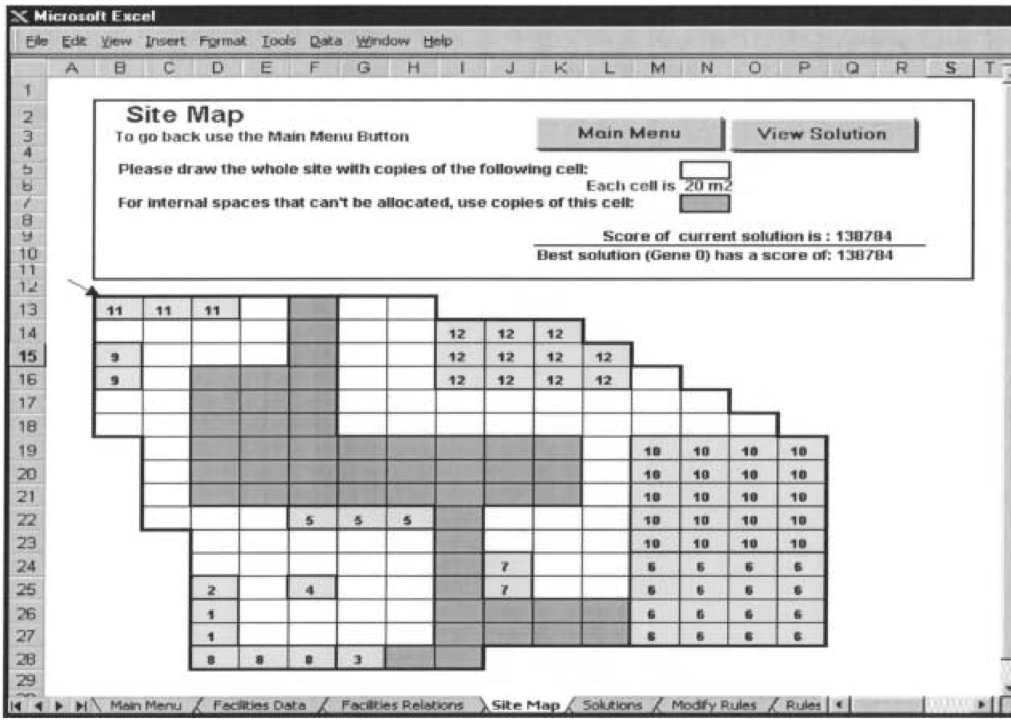


**Fig. 13.** Site map and final site layout plan of case study.

work requirements, site characteristics, and the resources engaged in a project. This helps in preventing omissions of important facilities and provides guidance for area computations. The fuzzy logic quantifier, on the other hand, effectively translates the fuzziness in the facilities' closeness relationships into crisp numbers that account for the work flow, safety concerns, and user preference of having facilities close or apart from each other. Lastly, the improved genetic algorithm proved to work successfully for practical size problems to find near-optimal layout of facilities in less time. To facilitate use of the model by practitioners, the model was implemented as a prototype spreadsheet system that is easy to use. The prototype allows the user to draw the site using spreadsheet cells and automates the layout optimization. A case study was used to demonstrate the capabilities of the system. Future extensions to dynamic site layout planning were then outlined.

## REFERENCES


1. Bishop, J. M., Bushnell, M. J., Usher, A. & Westland, S., Neural networks in the color industry, in *Applications of AI in Engineering VI*, G. Rzevski and R. A. Adey (eds.), Elsevier Applied Science, New York, 1991, pp. 423–34.
2. Chao, L. & Skibniewski, M. J., Fuzzy logic for evaluating alternative construction technology, *Journal of Construction Engineering and Management*, ASCE, **124** (4) (1998), 297–304.
3. Cheng, M. Y., Automated site layout of temporary construction facilities using geographic information systems (GIS), Ph.D. thesis, University of Texas, Austin, TX, 1992.
4. Cheng, M. Y. & O'Connor, J. T., ArcSite: Enhanced GIS for construction site layout, *Journal of Construction Engineering and Management*, ASCE, **122** (4) (1996), 329–36.
5. Cheng, S. J. & Hwang, C. L., *Fuzzy Multiple Attribute Decision Making: Methods and Applications* (Lecture Notes in Economics and Mathematical Systems), Springer-Verlag, Berlin, 1992.
6. Dubois, D. & Prade, H., *Possibility Theory: An Approach to Computerized Processing of Uncertainty*, Plenum Press, New York, 1998.
7. Dweiri, F. & Meier, F. A., Application of fuzzy decision making in facilities layout planning, *International Journal of Production Research* **34** (11) (1996), 3207–25.
8. *Excel Reference Manual*, Microsoft Co., One Microsoft Way, Redmond, WA, 1996.
9. Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Publishing Co., Reading, MA, 1989.
10. Hamiani, A., CONSITE: A knowledge-based expert system framework for construction site layout, Ph.D. thesis, University of Texas, Austin, TX, 1987.
11. Hanada, V. & Lang, B., Construction site planning, *Construction Canada* **30** (3) (1988), 43–9.
12. Hegazy, T. & Elbeltagi, E., EvoSite: An evolution-based model for site layout planning, *Journal of Computing in Civil Engineering*, ASCE, **13** (3) (1999), 198–206.

13. ILO, *Safety and Health in Construction*, International Labor Office, Geneva, 1992.
14. Larterapong, P. & Moselhi, O., Project-network analysis using fuzzy sets theory, *Journal of Construction Engineering and Management*, ASCE, **122** (4) (1996), 308–18.
15. Li, H. & Love, P., Site-level facilities layout using genetic algorithms, *Journal of Computing in Civil Engineering*, ASCE, **12** (4) (1998), 227–31.
16. Mamdani, E. H., Application of fuzzy algorithms for simple dynamic plant, *IEEE Proceedings* **121** (1974), 1585–8.
17. Obayashi Corporation, *Company Handbook for Temporary Facility Site Layout*, Japan, 1975.
18. Philip, M., Mahodevan, N. & Varghese, K., Optimization of construction site layout: A genetic algorithm approach, in *Proceedings of 4th Congress on Computing in Civil Engineering*, ASCE, Philadelphia, 1997, pp. 710–7.
19. Popescu, C., Construction site population: Estimating the size and structure, in *Proceeding of the 8th CIB Congress*, Oslo, Norway, 1980, pp. 633–5.
20. Rad, P. F. & James, B. M., The layout of temporary construction facilities, *Cost Engineering* **25** (2) (1983), 19–26.
21. Rodriguez-Ramos, W. E., Quantitative techniques for construction site layout planning, Ph.D. thesis, University of Florida, Gainesville, FL, 1982.
22. Rodriguez-Ramos, W. E. & Francis, R. L., Single crane location optimization, *Journal of Construction Division*, ASCE, **109** (4) (1983), 387–97.
23. Ross, T. J., *Fuzzy Logic with Engineering Applications*, McGraw-Hill, New York, 1995.
24. Tommelein, I. D., SightPlan: An expert system that models and augments human decision making for designing construction site layout, Ph.D. thesis, Stanford University, Stanford, CA, 1989.
25. Warszawski, A. & Peer, S., Optimizing the location of facilities on a building site, *Operational Research Quarterly* **24** (1) (1973), 35–44.
26. Yeh, I. C., Construction site layout using annealed neutral network, *Journal of Computing in Civil Engineering*, ASCE, **9** (3) (1995), 201–8.
27. Zadeh, L. A., Fuzzy sets, *Information and Control* **118** (1965), 338–53.
28. Zouein, P. P., MoveSchedule: A planning tool for scheduling space use on construction sites, Ph.D. thesis, University of Michigan, Ann Arbor, MI, 1995.