

## THUẬT GIẢI DI TRUYỀN (GAs) VÀ ỨNG DỤNG TRONG QUẢN LÝ XÂY DỰNG

LẠI HẢI ĐĂNG \*  
LƯU TRƯỜNG VĂN \*\*  
NGUYỄN TẤN BÌNH \*\*\*

### I) TỔNG QUAN

- Thuật giải di truyền (Genetic Algorithms- Viết tắt là GAs), do John Holland (1975) và Goldberg (1989) đề xuất và phát triển, là thuật giải tìm kiếm dựa trên cơ chế chọn lọc và di truyền tự nhiên. Thuật giải này sử dụng các nguyên lý di truyền về sự thích nghi và sự sống các cá thể thích nghi nhất trong tự nhiên.

- Tập hợp tất cả các lời giải trong không gian tìm kiếm được gọi là kiểu hình. Các kiểu hình này khi mã hoá gọi là kiểu gen. Toán tử di truyền sẽ được thực thi trên đối tượng này. Một ánh xạ từ kiểu hình sang kiểu gen gọi là quá trình mã hoá. Mỗi cá thể trong kiểu gen có nhiều nhiễm sắc thể. Trong mỗi nhiễm sắc thể có chứa nhiều gen. Mỗi đặc trưng di truyền cụ thể được quy định bởi giá trị và vị trí của gen trong nhiễm sắc thể. Độ thích nghi là thước đo khả năng sống sót và phát triển của cá thể trong môi trường

- Toán tử xác định cá thể trong thế hệ hiện tại được giữ lại trong thế hệ kế tiếp được gọi là chọn lọc. Toán tử kết hợp ngẫu nhiên hai cá thể được chọn gọi là lai ghép. Toán tử thay

đổi ngẫu nhiên cấu trúc cá thể, tức làm thay đổi giá trị của gen gọi là đột biến.

### II) CÁC TÍNH CHẤT QUAN TRỌNG CỦA THUẬT GIẢI DI TRUYỀN

- GAs lập luận mang tính chất ngẫu nhiên để tìm giải pháp tối ưu cho những vấn đề phức tạp, thay vì xác định như toán học giải tích. Tuy nhiên đây là hình thức ngẫu nhiên có hướng dẫn bởi trị số thích nghi. Chính hàm số thích nghi là giúp GAs tìm giải pháp tối ưu trong rất nhiều giải pháp có thể có.



Hình 1. Sơ đồ cấu trúc thuật giải di truyền

\* Giảng viên, Trường Đại học Bách Khoa, Đại học quốc gia TP.HCM

\*\* Thạc sĩ, Giảng viên, Trường Đại học Bách Khoa, Đại học quốc gia TP.HCM

\*\*\* Thạc sĩ, Giám đốc TT Bồi dưỡng Kế toán, Đại học Mở TP.HCM

-GAs không để ý đến chi tiết vấn đề, trái lại chỉ chú ý đến giải pháp cho vấn đề, hay tìm điều kiện tối ưu cho việc điều hành, và phân nhóm những giải pháp có được.

- GAs được sử dụng đặc biệt cho những bài toán yêu cầu tìm kiếm tối ưu toàn cục với không gian tìm kiếm lớn và không thể kiểm soát nhờ khả năng duyệt qua không gian tìm kiếm đại diện mà không thực sự đi qua từng điểm của toàn bộ không gian

### III) CƠ CHẾ THỰC HIỆN CỦA THUẬT GIẢI

#### 1. Mã hóa

- GAs bắt đầu với quần thể, tập của nhiều cá thể (nhiệm sắc thể). Sự mã hóa các biến phụ thuộc vào từng bài toán. Thông thường có các dạng mã sau: mã nhị phân, mã Gray, mã số thực và mã dạng cây. Khi mã hóa dưới dạng chuỗi nhị phân, mỗi nhiệm sắc thể được kí hiệu bằng

Nhiệm sắc thể A 

1	1	0	0	1	0	1
---	---	---	---	---	---	---

Nhiệm sắc thể B 

0	1	0	1	1	1	0
---	---	---	---	---	---	---

các chuỗi bit 0 và 1.

Hình 2. Mã hóa dưới dạng nhị phân

- Giả sử muốn tối ưu hàm  $n$  biến  $f(x_1, x_2, \dots, x_n)$ , trong đó mỗi biến  $x_i$  thuộc miền  $D=[a_i, b_i]$  là tập con của tập số thực  $R$  và yêu cầu độ chính xác là  $k$  chữ số thập phân cho các giá trị biến  $x_i$ . Để đạt được độ chính xác như vậy miền  $[a_i, b_i]$  được phân cắt thành  $(b_i - a_i) \cdot 10^k$  miền con bằng nhau. Gọi  $m_i$  là số nguyên nhỏ nhất sao:  $(b_i - a_i) \cdot 10^k \leq 2^{m_i} - 1$

- Như vậy mỗi biến  $x_i$  thuộc  $[a_i, b_i]$  được biểu diễn bằng một chuỗi nhị phân có chiều dài  $m_i$ . Phép ánh xạ biến nhị phân thành biến thực  $x_i$  được tính theo công thức:

$$x_i = a_i + \text{decimal}(\text{string}_2) \cdot \frac{b_i - a_i}{2^{m_i} - 1}$$

Trong đó  $\text{decimal}(\text{string}_2)$  biểu diễn giá trị thập phân của chuỗi nhị phân  $\text{string}_2$ . Bây giờ mỗi nhiệm sắc thể (là một lời giải) được biểu diễn bằng chuỗi nhị phân có chiều dài  $m = \sum_{i=1}^n m_i$

$m_1$  bit đầu tiên biểu diễn các giá trị trong khoảng  $[a_1, b_1]$ ,  $m_2$  bit kế tiếp biểu diễn giá trị trong khoảng  $[a_2, b_2]$  và nhóm  $m_n$  bit cuối cùng biểu diễn giá trị trong khoảng  $[a_n, b_n]$ .

#### 2. Chọn lọc cá thể.

- Đối với tiến trình chọn lọc (chọn quần thể thỏa phân bố xác suất dựa trên độ thích nghi), ta dùng bánh xe rulét với các rãnh được định kích thước theo độ thích nghi như sau:

Tính độ thích nghi  $eval(v_i)$  của mỗi nhiệm sắc thể  $v_i$  ( $i=1..pop\text{-}size$ ), với  $pop\text{-}size$  là kích thước của quần thể:

$$eval(v_i) = \frac{f(v_i)}{\sum_{i=1}^{pop\text{-}size} f(v_i)} \quad \text{với } f(v_i) \text{ là hàm mục tiêu}$$

Tìm tổng giá trị thích nghi  $F$  cho toàn quần thể:

$$F = \sum_{i=1}^{pop\text{-}size} eval(v_i)$$

Tính xác suất chọn  $p_i$  cho mỗi nhiệm sắc thể  $v_i$

$$p_i = \frac{eval(v_i)}{\sum_{i=1}^{pop\text{-}size} eval(v_i)}$$

Tính xác suất tích lũy  $q_i$  cho mỗi nhiệm sắc thể  $v_i$

$$q_i = \sum_{i=1}^{pop\text{-}size} p_i$$

- Tiến trình chọn lọc được thực hiện bằng cách quay bánh xe rulét  $pop\text{-}size$  lần. Mỗi lần chọn ra một nhiệm sắc thể từ quần thể hiện hành vào quần thể mới theo cách sau:

Phát sinh 1 số ngẫu nhiên  $r$  trong khoảng  $[0, 1]$

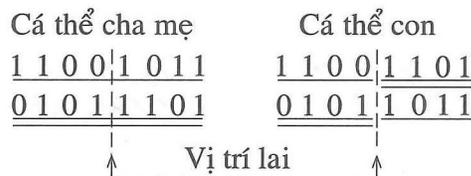
Nếu  $r < q_1$  thì chọn nhiệm sắc thể đầu tiên  $v_1$ , ngược lại chọn nhiệm sắc thể  $v_i$  sao  $q_{i-1} < r < q_i$

### 3. Lai ghép

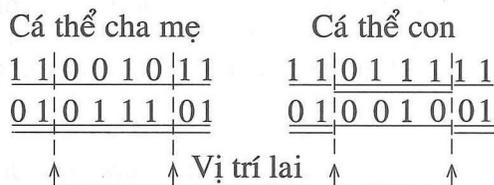
- Toán tử tác động trên các cá thể cha mẹ để tạo ra những con lai tốt được gọi là lai ghép. Các cặp cha mẹ được chọn lựa lai ghép với xác suất pc. Có 3 dạng lai ghép cơ bản: lai một vị trí, lai nhiều vị trí và lai đều. Với 3 loại trên, xác suất cá thể tạo ra do lai ghép vẫn là hằng số.

- Với mỗi nhiễm sắc thể trong quần thể:  
 Phát sinh 1 số ngẫu nhiên r trong khoảng [0,1 ]  
 Nếu r < pc thì chọn nhiễm sắc thể đó để lai ghép

- Sau đó ghép các nhiễm sắc thể đã được chọn một cách ngẫu nhiên. Đối với mỗi cặp nhiễm sắc thể được ghép đôi, lại phát sinh ngẫu nhiên một số nguyên pos trong khoảng [ 0,m ] (m là tổng số bit trong một nhiễm sắc thể). Số pos cho vị trí điểm lai. Hai nhiễm sắc thể (b1b2..bposbpos+1..bm) và (c1c2..cposcpos+1..cm) được thay bằng cặp con của chúng (b1b2..bposcpos+1..cm) và (c1c2..cposbpos+1..bm)



Hình 3. Lai đơn

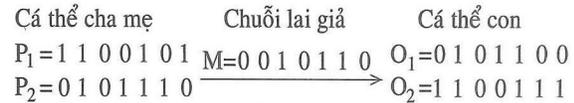


Hình 4. Lai bội

- Trường hợp lai đều thì mỗi gen của cá thể con được chọn một cách ngẫu nhiên gen tương ứng với cá thể bố hoặc mẹ. Cách tiến hành lai đều được tiến hành như sau:

+ Tạo một chuỗi lai giả M có chiều dài bằng chiều dài chuỗi bố, mẹ. Các bit được tạo ngẫu nhiên

+ Chuỗi con O được tạo ra bằng cách lấy từng gen từ cá thể cha, mẹ. Nếu bit thứ l trong chuỗi lai giả M là 1 thì lấy gen tương ứng của cá thể P1, ngược lại lấy gen tương ứng cá thể P2



Hình 5. Lai đều

### 4. Đột biến

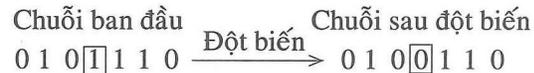
- Đột biến nhằm tạo ra những thông tin mới trong quần thể lai tạo tại các vị trí bit nào đó trong mỗi nhiễm sắc thể. Với xác suất đột biến trong quần thể là pm thì số lượng nhiễm sắc thể bị đột biến sẽ là pm\*pop-size. Mỗi bit trong nhiễm sắc thể có cơ hội đột biến như nhau và được thay đổi từ 0 thành 1 hay ngược lại:

- Với mỗi nhiễm sắc thể trong quần thể và mỗi bit trong nhiễm sắc thể:

Phát sinh 1 số ngẫu nhiên r trong khoảng [0,1 ]

Nếu r < pm , tiến hành đột biến bit đó

- Trong GAs, đột biến xảy ra với xác suất rất nhỏ, thường nằm trong khoảng 0.001 đến 0.01. Đột biến nhằm loại trừ sự nhầm lẫn do các tối ưu cục bộ



Hình 6. Đột biến

- Để quá trình đột biến có hiệu quả thì xác suất đột biến thường được chọn tỉ lệ nghịch với kích thước gen. Một xác suất đột biến thường được sử dụng là 1/N (N là kích thước gen). Hơn nữa xác suất đột biến nên độc lập với kích thước quần thể. Nghĩa là số lượng cá thể trong quần thể tăng hay giảm không ảnh hưởng đến khả năng đột biến cá thể trong quần thể .

### 5. Hàm thích nghi

- Vì hàm thích nghi phải nhận giá trị không âm, do đó phải xây dựng ánh xạ hàm mục tiêu đang xét trong bài toán sang hàm thích nghi thông qua một hay nhiều lần ánh xạ. Nếu bài toán tối ưu là cực tiểu một hàm mục tiêu g(x) thì việc chuyển hàm g(x) này sang hàm thích nghi f(x) để sử dụng trong GAs như sau:

$$f(x) = C_{max} - g(x) \text{ khi } g(x) < C_{max} ;$$

Ngược lại  $f(x)=0$

- Trong đó  $C_{max}$  là tham số đầu vào. Có thể lấy  $C_{max}$  là giá trị  $g(x)$  lớn nhất trong quần thể hiện tại, hoặc lớn nhất sau  $k$  vòng lặp

- Khi hàm mục tiêu gốc tăng hoặc bài toán đang xét cực đại của hàm  $u(x)$ , hàm thích nghi có thể được chuyển sang như sau

$$f(x) = C_{min} + u(x) \quad \text{khi } u(x) + C_{min} > 0$$

Ngược lại  $f(x)=0$

- Trong đó  $C_{min}$  là tham số đầu vào, có thể là trị tuyệt đối của  $u$  bé nhất trong quần thể hiện tại, hoặc trong  $k$  vòng lặp cuối cùng, hoặc là một hàm của biến quần thể.

- Trong một vài trường hợp hàm thích nghi có thể là nghịch đảo của hàm mục tiêu hoặc là sai số trung bình bình phương của các tập mẫu trong quần thể.

## 6. Xử lý các ràng buộc

- GAs thích hợp cho các bài toán tìm kiếm tối ưu với điều kiện không ràng buộc. Tuy nhiên thực tế bài toán có thể chứa một hoặc nhiều ràng buộc phải thỏa. Lời giải nhận được từ chiến lược tìm kiếm tối ưu nhất thiết phải nằm trong vùng khả thi, tức phải thỏa tất cả các ràng buộc. Thông thường có thể xử lý các ràng buộc bằng hàm phạt. Một số hàm phạt thường được dùng như sau:

$$\text{- Hàm phạt tĩnh: } f_p(x) = f(x) + \sum_{i=1}^m C_i \alpha_i$$

Trong đó :

$f(x)$  là hàm mục tiêu;  $\alpha_i = 1$  nếu ràng buộc thứ  $i$  vi phạm;  $\alpha_i = 0$  nếu ràng buộc thứ  $i$  thỏa

$C_i$ : hệ số phạt tương ứng khi ràng buộc thứ  $i$  vi phạm

- Hàm phạt động:

$$f_p(x,t) = f(x) + \sum_{i=1}^m S_i(t) d_i^k$$

$$\text{Với } d_i^k = \begin{cases} \alpha_i g_i(x) & i = \overline{1, q} \\ |h_i(x)| & i = \overline{q+1, m} \end{cases}$$

$g_i(x)$ ,  $h_i(x)$  là các phương trình ràng buộc

$$g_i(x) \leq 0 \quad \text{với } i=1, \dots, q$$

$$h_i(x) = 0 \quad \text{với } i=q+1, \dots, m$$

$$S_i(t) = (C_i t)^m$$

$t$ : thế hệ vào thời điểm tính toán

$m$ : hằng số, có giá trị bằng 1 hoặc 2

- Hàm phạt thích nghi:

$$f_p(x,k) = f(x) + \sum_{i=1}^m \lambda_k d_i^k$$

$$\text{Với } \lambda_{k+1} = \begin{cases} \lambda_k / \beta_1, & \overline{B}(i) \in F : \forall t-k+1 \leq i \leq t \\ \lambda_k / \beta_2, & \overline{B}(i) \notin F : \forall t-k+1 \leq i \leq t \end{cases}$$

Trong đó :

$F$ : phần khả thi của không gian tìm kiếm.

$\overline{B}(i)$ : biểu diễn cá thể tốt nhất trong thế hệ thứ  $i$ .

$\beta_1, \beta_2 > 1$  và  $\beta_1 \neq \beta_2$

## 7. Điều kiện kết thúc lặp của GAs

- Để kết thúc vòng lặp GAs, thường có thể chỉ định trước số thế hệ cần tạo ra sau đó kiểm tra lại độ thích nghi những phần tử tốt nhất bằng cách so sánh với bài toán ban đầu.

## IV) CÁC ỨNG DỤNG CỦA GAS TRONG LĨNH VỰC QUẢN LÝ XÂY DỰNG

- Mawdesley [4] đã nghiên cứu và ứng dụng GAs trong công tác thiết lập tổng mặt bằng xây dựng. Trong nghiên cứu này, Mawdesley đã giải quyết các yếu tố tồn tại trong công tác thiết kế tổng mặt bằng cho 1 nhà máy ở Anh:

+ Chọn phương án thi công xây dựng công trình tạm với điều kiện vật tư, nhân công hạn chế.

+ Chọn vị trí và diện tích tối ưu các địa điểm cung cấp nguyên vật liệu như kho chứa xi măng, sắt thép, ván khuôn, trạm đổ bê tông. Thiết kế lộ trình cung cấp vật tư từ các địa điểm này đến nơi tiêu thụ và đường di chuyển của máy đào đất sao cho đạt được hiệu quả nhất.

- Liên quan đến lĩnh vực tối ưu hóa tổng mặt bằng, Tam [5] đã sử dụng GAs để tối ưu

hóa các vị trí cung cấp nguyên vật liệu xung quanh cần trục tháp cho một tòa nhà cao tầng ở Hồng Kông

- Trong các công trình xây dựng lớn như: xây đập, đường cao tốc, khoan hầm, công tác đất (đào đất, vận chuyển và đổ đất) chiếm một tỉ trọng lớn chi phí của công trình. Vì vậy việc lựa chọn máy móc thiết bị, phương pháp thi công để tối ưu hóa công tác đất có ý nghĩa quan trọng để giảm chi phí xây dựng công trình. Marzouk [6][7] đã nghiên cứu GAs kết hợp với mô phỏng trên máy tính về đề tài tối ưu công tác đất trong luận án tiến sĩ và báo cáo khoa học trên tạp chí xây dựng ASCE (Mỹ). Trong nghiên cứu này Marzouk đã áp dụng GAs để tối ưu việc lựa chọn các tổ máy móc thi công tham gia công tác vận chuyển 2.5 triệu m<sup>3</sup> đất ra khỏi công trình xây dựng đến nơi chứa cách xa 15 km. Kết hợp với công tác lựa chọn máy móc thi công, trình tự và biện pháp thi công sẽ được mô phỏng trên máy tính sao cho chi phí xây dựng nhỏ nhất.

- Một nhóm nghiên cứu tại Đại học South Bank [8] (Anh) đã nghiên cứu và ứng dụng GAs để tối ưu hóa năng suất máy móc phục vụ công trình đào hầm theo phương pháp đúc mở; Kết quả của nghiên cứu đã được thực nghiệm qua 4 công trình trong thực tế để kiểm tra tính xác thực của đề tài này.

- Một ứng dụng quan trọng của GAs là tối ưu hóa tiến độ thi công với các ràng buộc về tài nguyên và nhân vật lực. Đây là lĩnh vực được nhiều nhà quản lý dự án quan tâm nhất. Đã có rất nhiều nhà nghiên cứu tìm hiểu về lĩnh vực tối ưu hóa tiến độ cùng với việc cải tiến GAs để đạt được mô hình hiệu quả nhất. Với sự đa dạng các yếu tố tác động đến tiến độ, trong tương lai sẽ còn nhiều nhà nghiên cứu tìm hiểu về lĩnh vực này. Hegazy [9] ứng dụng GAs để lập tiến độ với điều kiện nhân lực bị giới hạn. Senouci [10] sử dụng GAs để thiết lập mô hình tiến độ có xét đến các yếu tố như mối quan hệ giữa các công tác, cấu trúc thành phần tổ đội, cân bằng nguồn lực theo tiến độ, cực

tiểu hóa chi phí dự án sao cho thời gian hoàn thành công trình là ngắn nhất. Li và Love [11] nghiên cứu GAs để tối ưu hóa tiến độ với điều kiện thỏa hiệp về thời gian - chi phí

- Trong quá trình thi công, việc cung cấp tài nguyên (vật tư, nhân công) là thường xuyên. Để tránh rủi ro khan hiếm vật liệu, vật liệu tăng giá, do tiến độ thi công bị đẩy nhanh đột xuất làm cho thiếu hụt vật tư, nhân công, nhà thầu cần dự trữ những vật liệu gì? Số lượng bao nhiêu? Vào thời gian nào? Đây chính là bài toán tối ưu hóa quản lý tài nguyên. Lew [12] đã ứng dụng GAs để tối ưu hóa kế hoạch sử dụng tài nguyên lao động theo tiến độ. Hegazy [13] ứng dụng GAs để lựa chọn và bố trí cấu trúc thành phần tổ đội thi công các công tác sao cho đạt được năng suất chung cao nhất. Ngoài ra, nghiên cứu về lĩnh vực tối ưu hóa kế hoạch sử dụng tài nguyên lao động còn có nhiều nhà nghiên cứu khác như: Chan [14], Syswerda[15], Satyanarayana [16]

- Đối với các công ty thi công, vấn đề kiểm soát điều khiển nhiều công trình thi công cùng lúc là một điều rất khó khăn. Để hoàn thành các công trình đúng tiến độ với điều kiện bị hạn chế về vật tư, nhân công, thiếu vốn, đòi hỏi các nhà quản lý phải tính toán lựa chọn tối ưu sao cho đạt được kết quả tốt nhất. GAs có thể giúp giải quyết các bài toán tối ưu hóa sử dụng nguồn tài nguyên (dòng tiền tệ, nhân công, máy móc, vật tư) phân phối cho nhiều công trình khác nhau trong một công ty xây dựng. Đề tài này đã được East [17] nghiên cứu và kiểm nghiệm qua các dự án của trường Đại học Illinois. Lam và Gao [18] ứng dụng GAs để tối ưu hóa dòng tiền phân phối cho nhiều công trình ở các công ty xây dựng của Trung Quốc.

Tóm lại, bên cạnh lĩnh vực tối ưu hóa trong quản lý xây dựng, GAs còn có thể được ứng dụng trong lĩnh vực quản lý quy hoạch đô thị, đầu tư xây dựng và tài chính doanh nghiệp. Hy vọng trong tương lai sẽ có nhiều nhà nghiên cứu Việt Nam tìm hiểu về lĩnh vực tối ưu hóa bằng thuật giải di truyền.

**TÀI LIỆU THAM KHẢO**

1. Nguyễn Đình Thú..(2001). Lập trình tiến hóa, Nhà xuất bản giáo dục.
2. Michlaewicz, Z.(1992). Genetic Algorithms + Data Structures= Evolutionary Programs. Second Edition, Springer-Verlag, Cambridge, Mass.
3. David E. Goldberg.(1989). Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Publishing Company, Inc.
4. Michael J. Mawdesley; Saad H. Aljbi. Genetic Algorithms for Construction Site Layout in Project Planning. ASCE, Journal of Construction Engineering and Management.
5. C.M.Tam.(2001). Genetic Algorithms for Optimizing Supply Locations around Tower Crane. City University of HongKong. HongKong.
6. Mohamed Marzouk; Osama Moselhi.(2004). Multiobjective Optimization of Earthmoving Operations. ASCE, Journal of Construction Engineering and Management.
7. Mohamed Marzouk.(2002).Optimizing earthmoving operations using computer simulation. PhD thesis, Concordia Univ.Motreal.
8. A.Haidar;S.Naoum;R.Howes.(1999). Genetic Algorithms Application and Testing for Equipment Selection.South Bank University.UK
9. Tarek Hegazy.(1999). Algorithm for Scheduling with Multiskilled Constrained Resources. University of Waterloo. Canada.
10. Ahmed B. Senouci.(2004). Use of Genetic Algorithms in Resource Scheduling of Construction Projects.ASCE, Journal of Construction Engineering and Management.
11. H.Li; Peter Love.(1997). Using Improved Genetic Algorithms to Facilitate Time-Cost Optimization. ASCE, Journal of Construction Engineering and Management.
12. Sou-Sen Lew.(1999). GA-Based Multicriteria Optimal Model for Construction Scheduling. National Taiwan University of Science and Technology, Taiwan.
13. Tarek Hegazy.(1999). Optimization of Resource Allocation and Leveling Using Genetic Algorithms. University of Waterloo.Canada.
14. Chan, W.T.Chua.(1996). Construction resource scheduling with genetic algorithms. ASCE, Journal of Construction Engineering and Management.
15. Syswerda,G.,and Palmucci,J.(1991). The Application of Genetic Algorithms to Resource Scheduling.Proc, 4 th Int. Conf on Genetic Algorithms, R.K Belew and L.B.Booker,Morgan Kaufman.San Mateo, California. USA.
16. Satyanarayana.(1993).Optimum resource allocation in construction project using genetic algorithms.Proc,3rd Ind.Conf, on the Application of AI to Civil and Structure Engineering, Edinburgh, UK.
17. E.W.East(1998). Dynamic, Multi-project Scheduling Under Limited Resources with Uncertain Project Demand. Construction Engineering Research Laboratories, ATTN. USA.
18. K.C.Lam, H.Gao(2003). Optimizing Multi-project Cash Flow for Chinese Construction Firms. City University of HongKong. HongKong.