

바이오의약품 위탁생산 일정계획 수립을 위한 유전자 알고리즘

A Genetic Algorithm for Production Scheduling of Biopharmaceutical Contract Manufacturing Products

김지훈 · 김정현 · 김재곤[†]

인천대학교 산업경영공학과

요약

바이오의약품 위탁생산 사업(CMO)에서 다양한 고객의 주문에 대해 납기를 만족시키는 생산 일정계획을 수립하는 것은 사업이 경쟁력 측면에서 매우 중요하다. CMO 공정에서 각 주문은 다수의 배치로 구성되어 있으며 복수 개의 생산라인에 소량의 배치 단위로 할당되어 병렬 생산할 수 있다. 본 연구는 동종 병렬 설비를 보유하고 있는 CMO 공정에서 주문의 총 납기 지연을 최소화하는 일정계획을 수립하기 위한 메타 휴리스틱 알고리즘을 제안한다. 제안된 알고리즘은 생물학적 진화에서 영감을 받아 염색체와 같은 데이터 구조를 무작위로 생성하여 특정 문제를 해결하며, 교차 및 돌연변이와 같은 작업을 사용하여 다양한 솔루션을 효과적으로 탐색한다. 국내 CMO 기업체에서 제공한 현업 데이터를 기반으로 컴퓨터 실험을 진행하여 제안한 알고리즘이 기업체에서 사용하고 있는 전문가 알고리즘과 상용 최적화 패키지보다 우수한 일정계획을 적절한 계산시간 내에 도출하는 것을 확인하였다.

■ 중심어 : 유전자 알고리즘, 바이오의약품 위탁생산, 순서에 따른 작업 준비시간, 총 납기 지연 최소화, 작업 분할

Abstract

In the biopharmaceutical contract manufacturing organization (CMO) business, establishing a production schedule that satisfies the due date for various customer orders is crucial for competitiveness. In a CMO process, each order consists of multiple batches that can be allocated to multiple production lines in small batch units for parallel production. This study proposes a meta-heuristic algorithm to establish a scheduling plan that minimizes the total delivery delay of orders in a CMO process with identical parallel machine. Inspired by biological evolution, the proposed algorithm generates random data structures similar to chromosomes to solve specific problems and effectively explores various solutions through operations such as crossover and mutation. Based on real-world data provided by a domestic CMO company, computer experiments were conducted to verify that the proposed algorithm produces superior scheduling plans compared to expert algorithms used by the company and commercial optimization packages, within a reasonable computation time.

■ Keyword : Genetic Algorithm, Biopharmaceutical Contract Manufacturing, Sequence Dependent Setup Times, Total Tardiness, Job Splitting

I. 서론

바이오의약품 위탁생산 사업(Contract Manufacturing Organization; CMO) 전문기업에서 바이오의약품을 생산할 경우, 의뢰기업이 요구한 생산량과 납기를 만족하는 효율적인 일정 계획이 중요하다[1]. 대다수의 바이오의약품 생산기업은 효율적인 생산을 위해 CMO 전문기업에 생산을 위탁한다[2]. CMO 전문기업이 새로운 의뢰기업의 위탁생산 주문을 받은 경우, 기존 의뢰기업의 주문을 고려한 일정계획을 토대로 수락 여부를 결정한다. 효율적인 일정계획을 통해 CMO 전문기업들은 기존보다 많은 의뢰를 수락하여 기업 이익을 증대할 수 있다.

CMO에서 바이오의약품 생산 일정계획의 경우 작업자는 각 생산라인(Site)에서의 의약품 생산 순서와 생산량을 결정해야 한다. 의뢰기업의 주문을 하나의 작업으로 볼 때, 각 작업은 여러 개의 배치(batch)로 구성되어 있으며 분할되어 복수 개의 Site에서 병렬 생산될 수 있다. 따라서, CMO에서의 바이오의약품 생산은 작업 분할과 순서를 같이 고려하는 조합 최적화 문제이다. 일반적으로 병렬 설비 내의 일정계획 문제는 조합 최적화 문제 중에서도 NP-hard로 알려져 있다[3,4,5]. 또한, 작업 분할이 같이 고려되어야 하는 병렬 설비 내의 일정계획 문제도 NP-hard로 알려져 있다[6,7]. 작업 분할을 고려하는 병렬 설비 내의 일정계획 문제는 바이오의약품 위탁생산 공정 이외에 다양한 공정에 적용될 수 있으며, 이러한 이유로 최근까지 다수의 연구자에 의해 연구가 진행되었다[8,9].

기존 연구에서는 NP-hard의 해결 방안 중 하나로써 유전자 알고리즘(Genetic Algorithm; GA)을 사용하고 있다[10]. GA는 자연 진화를 모방한 확률적 탐색기법으로서 Holland(1975)에 의해 처음 제안된 메타 휴리스틱 알고리즘이다. 무작위로 생성된 초기 해를 유전 연산자를 통해

점진적으로 개선하고 적합도 함수를 활용하여 해의 품질을 평가하는 작업을 거치며, 적절한 해를 찾을 때까지 이를 반복한다. 다양한 조합 최적화 문제에서 다수의 연구자에 의해 그 유용성이 입증되었으며, NP-hard 문제에서 근사 최적해를 도출하기에 적합한 알고리즘이라는 것이 확인되었다[12,13].

본 연구에서는 CMO에서의 바이오의약품 생산 일정계획에서 총 납기 지연을 최소화하는 유전자 알고리즘을 제안한다. 의뢰기업이 주문한 의약품의 요구 생산량과 납기를 고려하여 총 납기 지연의 발생을 최소화하는 일정계획을 도출하는 것이 CMO에서의 바이오의약품 생산 일정계획에서 가장 중요한 요소임을 실제 국내 CMO 전문기업인 M사로부터 확인하였다. 또한 M사의 공정 데이터를 기반으로 문제 및 제약을 구성하였으며, 수리모형의 결과와 비교실험을 진행하여 그 성능을 입증한다.

본 논문은 5장으로 구성되어 있으며 각 장의 구체적인 내용은 다음과 같다. 제2장에서는 기존 연구자들에 의해 진행된 관련 연구를 설명하고, 제3장에서는 구체적인 문제 정의 및 수리 모델을 소개한다. 제4장에서는 문제 해결을 위한 알고리즘을 소개한다. 제5장에서는 제안한 알고리즘과 현업 알고리즘의 성능 비교 실험을 진행하였으며, 제6장에서는 실험 결과를 바탕으로 결론 및 추후 연구에 대하여 논의하였다.

II. 기존 연구

바이오의약품 위탁생산 일정계획을 최적화하기 위한 기존 연구들은 다수의 연구자에 의해 이루어졌다. Lakhdar 등(2007)은 수율이 존재하는 CMO에서의 바이오의약품 생산 문제에서 수리모형을 기반으로 하는 다목적 용량 계획 결정 모델을 제안하였으며, Oyebolu(2017) 등은 수율이 존재하는 CMO에서의 바이오의약품 생산 문

제에서 다목적 휴리스틱 알고리즘을 제안하였다. Sahling과 Hahn(2019)은 유통기한이 존재하는 바이오의약품 생산 계획 문제를 다루고 있으며, 휴리스틱 알고리즘을 제안하고, 실제 바이오의약품 업계의 생산 문제를 이용한 실험을 통해 제안한 알고리즘의 성능을 입증하였다. 바이오의약품 생산 계획을 다룬 기존 연구들이 있으나, 문제 구조가 본 연구에서 다루는 바이오의약품 생산공정과 맞지 않아 기존 연구를 적용하기에는 어려움이 있다. CMO에서의 의약품 생산 일정계획을 최적화하기 위해서는 의약품의 배치를 각 Site에 분할 할당하는 일정계획이 필요하다. 이는 작업 분할이 가능한 작업에 대한 병렬 설비에서의 일정계획 최적화 연구와 밀접한 연관성을 갖는다.

Kim 등(2004)은 작업 분할이 가능한 동종 병렬 설비 내의 일정계획 문제에 대해 총 납기 지연을 최소화하는 방법을 제시하였다. 두 단계에 걸친 휴리스틱 알고리즘을 사용하여 문제를 해결하였으며 당시 기존 연구에 비해 제안된 알고리즘이 높은 성능이 보임을 확인하였다. Logendran과 Subur(2004)은 총 납기 지연 최소화를 목표로 하는 이종 병렬 설비 내에서 최대 두 개의 하위작업으로 분할될 수 있는 일정계획을 다룬다. 디스패치 규칙 기반의 네 가지 방법을 통해 초기 해를 형성한 후 타부 서치 기반의 휴리스틱 알고리즘을 사용하는 해결책을 제시하였다. 실험 결과로 특정 복합 디스패치 휴리스틱 알고리즘이 타부 서치 기반 휴리스틱이 최적해에 도달하는 속도를 증가시키는 것을 확인하였다. Eroglu 등(2014)은 1,100개의 작업 수와 적격성 제약이 존재하는 작업의 작업 분할을 고려하는 이종 병렬 설비 내의 일정계획 문제에서 총작업 완료 시간을 최소화하는 문제를 다루었다. 작업순서에 따른 작업 준비시간이 존재하며 설비에 작업 적격성 제약이 존재하는 문제를 해결하기 위한 유전 알고리즘을 제시하였다. Kim 등(2020)은

작업 분할이 가능한 작업의 총지연을 최소화하는 동종 병렬 설비 내의 일정계획 문제에 대해 다루고 있으며, 이를 해결하기 위해 두 가지 인코딩 방식과 세 가지 디코딩 방식을 가진 유전 알고리즘을 제안하였으며, 디코딩 방식에서 휴리스틱 알고리즘을 통한 작업 분할 방식을 사용하였다. Lee 등(2021)은 설비 간 작업 준비 작업이 동시에 수행될 수 없으며 동일 설비 내 작업 분할이 불가능한 제약사항 내에서 총작업 완료 시간 최소화를 목적으로 하는 병렬 설비에서의 작업 분할이 가능한 일정계획 문제에 대해 다루고 있다. 문제 해결을 위해 수리모형을 통해 해를 도출한 후 반복 작업 분할 휴리스틱 알고리즘을 통해 해를 개선하였다. Zheng 등(2022)은 작업 적격성 제약이 존재하는 병렬 설비 내에서 작업 분할이 가능한 작업 총 주문 완료 시간과 기계의 작업 처리비용의 가중 합계를 최소화하는 수리모형을 제시하였으며 대규모 문제를 해결하기 위해 두 가지 휴리스틱 알고리즘을 제시하였다.

기존 연구는 일정계획에서 작업의 할당 순서와 작업 분할을 동시에 고려하지 않거나 작업 분할을 휴리스틱 알고리즘에 의존한 경우가 대부분이다. 작업 분할과 작업순서를 동시에 고려하지 않은 일정계획의 경우 전역 해 공간을 탐색하지 못해 지역 최적해에 빠질 위험이 있다. 또한 휴리스틱 알고리즘이 매개 변수를 사용하는 경우 선택된 매개 변수에 따라 성능이 달라질 수 있다. 본 연구에서는 유전자 표현 방식에서 작업의 순서와 분할을 동시에 고려하는 GA 기반의 CMO에서의 바이오의약품 생산 일정의 총 납기 지연 최소화 알고리즘을 제시한다.

III. 문제 정의

본 연구의 문제는 작업을 분할 할 수 있는 순서에 따른 작업 준비시간이 존재하는 동종 병렬

설비 내의 일정계획 문제로 나타낼 수 있다. CMO 전문기업 M사는 의뢰기업의 의뢰주문에 따라 주문량만큼 의약품을 Site에서 생산한다. Site에서 한 번에 생산할 수 있는 용량이 정해져 있기에, 생산 용량에 따라 나누어지는 배치 생산 방식을 통해 바이오의약품을 생산한다. 본 연구의 일정계획 문제는 각 Site가 같은 생산 용량을 갖는다고 가정한다. Site에 할당된 배치는 종 배양(Seed culture), 본 배양(Main culture), 정제(Purification) 세 단계의 하부공정을 거친다. 각 하부공정은 Site 내 별도의 설비에서 처리되며 이전 하부공정이 완료되면 즉시 후속 하부공정이 진행된다. 기존 배치가 완료되지 않아도 하부공정이 겹치지 않는 한, 후속 배치를 같은 Site에 할당할 수 있다.

그림 1은 의뢰주문 A, B, C의 일정계획에 대한 간트차트 예시이며, 의뢰주문 표기 뒤의 숫자는 몇 번째 배치인지를 나타낸다. 이전 배치의 하부공정이 종료되지 않은 시점에 다음 배치가 할당되는 것을 확인할 수 있다. M사의 각 Site는 하부공정을 처리하기 위한 동종 설비를 보유하고 있으며 의약품 오염을 방지하기 위해 하부공정이 완료 후에는 세척 및 소독을 진행한다. M사의 경우 같은 의뢰주문의 배치가 본 배양과 정제에서 연속적으로 할당되는 경우 반대의 경우에 비해 상대적으로 긴 시간 동안 세척 및 소독이 이루어진다. 종 배양의 세척 및 소독 시간은 의뢰주문의 종류에 영향을 받지 않는다. 본 배양과 정제의 시작 시각은 종 배양 시작 시

각에 따라 결정되기에 각 배치의 작업 준비시간은 종 배양의 세척 및 소독 시간뿐 아니라 현재 배치와 후속 배치의 처리시간과 세척 및 소독 시간을 고려하여 결정한다. 같은 의뢰주문의 배치를 연속 할당하는 경우, 세척 및 소독으로 인해 다른 의뢰주문의 배치가 할당되는 경우보다 작업 준비시간이 적다. 같은 의뢰주문의 배치를 연속으로 할당하여 생산 시간을 단축할 수 있으며, Site 내에서 같은 의뢰주문의 배치를 분할 할당하는 것은 납기 준수에 이점이 없다.

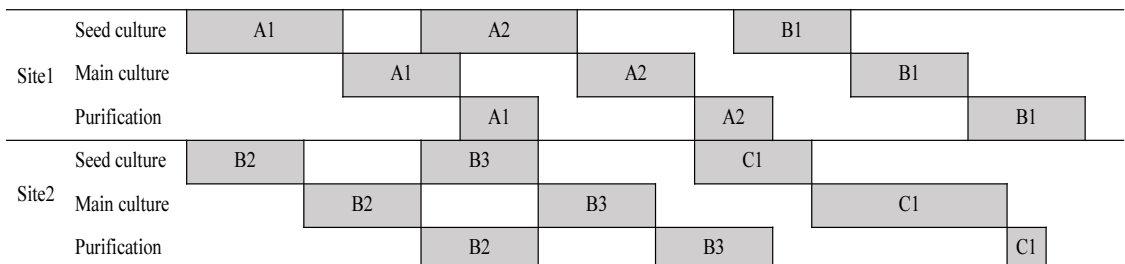
본 연구는 의뢰받은 주문량 내에서 의약품을 배치 단위로 각 Site에 분할 및 할당하는 일정계획 문제로 나타낼 수 있다. CMO에서의 바이오 의약품 생산 일정계획 총 납기 지연 최소화 문제는 정수 선형 계획법을 통해 수리모형으로 나타낼 수 있다. 수리모형과 논문 전반에 사용되는 의사결정 변수 및 파라미터는 표 1과 같다. 본 연구에서는 Sim(2024)의 수리모형을 참고하여, 목적 식 (1)과 제약 식 (2)~(15)로 문제를 정의하였다. 수리모형에서 의뢰주문 0은 가상의 더미 주문으로 수리모형 설계의 편의상 도입하였다.

[P]

$$\text{Minimize } \sum_{i=1}^I \tau_i \tag{1}$$

Subject to

$$\sum_{i=0}^I z_{i,l,t} = 1 \quad \forall l = \{1, \dots, L\} \\ \forall t = \{1, \dots, T\} \tag{2}$$



〈그림 1〉 CMO에서의 바이오의약품 생산 일정계획 간트 차트

〈표 1〉 의사결정 변수와 파라미터

Parameter	Explanation
I	Number of orders
O	Number of operations
L	Number of Sites
T	Number of total batches
i, j	Index of orders
l	Index of Sites
o	Index of operations
t	Index of positions
V	Very large number
$p_{i,o}$	Processing time of operation o of order i
D_i	Due date of order i
Q_i	Size(number of batches) of order i
$b_{i,j}$	Setup time required to process batches of order j just after those of order i
Decision variable	Explanation
$s_{l,t}$	Start time of a batch assigned to t -th position of Site l
$c_{l,t}$	Completion time of a batch assigned to t -th position of Site l
C_i	Completion time of order i
τ_i	Tardiness of order i
$z_{i,l,t}$	=1 if order i is assigned to t -th position on Site l and 0, otherwise
$\gamma_{i,j,l,t}$	=1 if order i assigned to t -th position precedes order j (order j is sequenced just after order i) on Site l and 0, otherwise

$$\sum_{l=1}^L \sum_{t=1}^T z_{i,l,t} = Q_i \quad \forall i = \{1, \dots, I\} \quad (3) \quad c_{i,t} - V(1 - z_{i,l,t}) - D_i \leq \tau_i \quad \begin{matrix} \forall i = \{1, \dots, I\}, \\ \forall l = \{1, \dots, L\}, \\ \forall t = \{1, \dots, T\} \end{matrix} \quad (9)$$

$$\gamma_{i,j,l,t} \leq 0.5(z_{i,l,t} + z_{j,l,t+1}) \quad \begin{matrix} \forall l = \{1, \dots, L\}, \\ \forall t = \{1, \dots, T-1\}, \\ \forall i, j = \{1, \dots, I\} \end{matrix} \quad (4) \quad V(1 - z_{0,l,t}) \geq \sum_{i=1}^I z_{i,l,t+1} \quad \begin{matrix} \forall l = \{1, \dots, L\}, \\ \forall t = \{1, \dots, T-1\} \end{matrix} \quad (10)$$

$$\gamma_{i,j,l,t} \geq z_{i,l,t} + z_{j,l,t+1} - 1 \quad \begin{matrix} \forall l = \{1, \dots, L\}, \\ \forall t = \{1, \dots, T-1\}, \\ \forall i, j = \{1, \dots, I\} \end{matrix} \quad (5) \quad s_{l,t} \geq 0 \quad \begin{matrix} \forall l = \{1, \dots, L\} \\ \forall t = \{1, \dots, T\} \end{matrix} \quad (11)$$

$$s_{l,t} - s_{l,t-1} \geq \sum_{i=1}^I z_{i,l,t-1} p_{i,1} + \sum_{i=1}^I \sum_{j=1}^I \gamma_{i,j,l,t-1} b_{i,j} \quad (6) \quad c_{l,t} \geq 0 \quad \begin{matrix} \forall l = \{1, \dots, L\} \\ \forall t = \{1, \dots, T\} \end{matrix} \quad (12)$$

$$s_{l,1} = 0 \quad \forall l = \{1, \dots, L\} \quad (7) \quad \tau_i \geq 0 \quad \forall i = \{1, \dots, I\} \quad (13)$$

$$s_{l,t} + \sum_{i=1}^I \sum_{o=1}^O z_{i,l,t} p_{i,o} = c_{l,t} \quad \begin{matrix} \forall l = \{1, \dots, L\}, \\ \forall t = \{1, \dots, T\} \end{matrix} \quad (8) \quad z_{i,l,t} \in \{0, 1\} \quad \begin{matrix} \forall i = \{0, \dots, I\}, \\ \forall l = \{1, \dots, L\}, \\ \forall t = \{1, \dots, T\} \end{matrix} \quad (14)$$

$$\gamma_{i,j,l,t} \in \{0,1\} \quad \begin{matrix} \forall i,j = \{1, \dots, I\}, \\ \forall l = \{1, \dots, L\}, \\ \forall t = \{1, \dots, T\} \end{matrix} \quad (15)$$

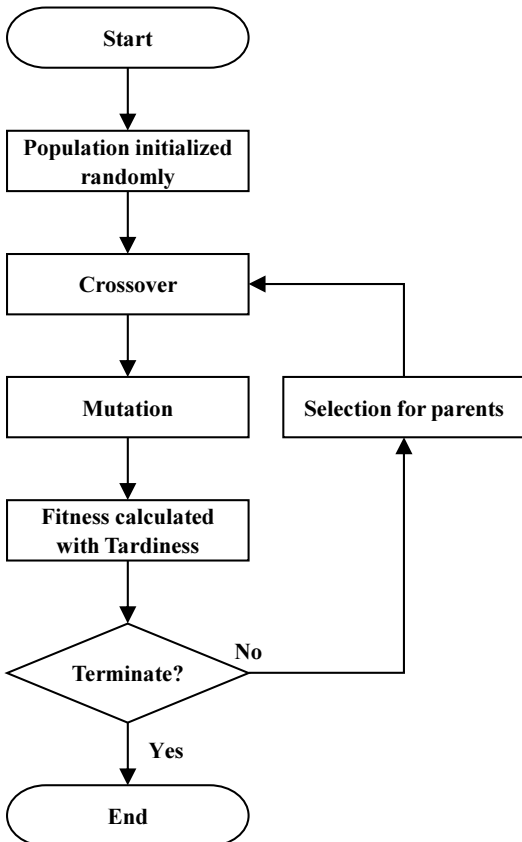
식(1)은 [P]의 목적함수로서 프로젝트의 총 납기 지연 최소화를 나타낸다. 식(2)는 같은 Site에 두 개 이상의 배치를 동시에 할당할 수 없음을 보장한다. 식(3)은 의뢰기업의 주문 생산량을 만족해야 함을 보장한다. 식(4), (5)는 배치 할당 순서에 따라 달라지는 할당 간 최소 시간 간격을 확인한다. 식(6)은 할당 간 최소 시간 간격을 고려한 배치의 할당 시간을 나타낸다. 식(7)은 각 사이트의 첫 번째 할당된 배치는 시작시간이 0임을 보장한다. 식(8)은 할당한 배치의 처리 완료 시간을 계산한다. 식(9)는 의뢰기업의 주문에

대한 납기 지연을 계산한다. 식(10)은 더미 의뢰 주문이 Site 내에서 실제 주문의 할당이 끝난 후 할당됨을 보장한다. 식(11)-(15)은 의사결정 변수가 특정 값 범위 내에 제한되도록 한다.

수리모형 [P]를 이용해 소규모 문제에 대해서는 최적해를 도출할 수 있으나, 본 연구에서 다루는 실제 대규모 문제의 경우 다항 시간 내에 최적해를 도출할 수 없다[24]. 따라서 본 연구에서는 현실 세계의 문제를 빠르고 효과적으로 해결하기 위한 GA를 제시한다.

IV. 문제 해결 알고리즘

GA는 진화 이론의 원리에 기반한 확률적 검색 및 최적화 기법으로 어려운 조합 최적화 문제를 해결하여 상당한 관심을 받았다. 본 연구에서는 염색체(Chromosome)를 통해 배치의 할당 순서와 할당 설비를 결정한 후, 설비를 기준으로 동일 제품의 배치를 결합하여 작업 분할을 고려하는 GA를 제안한다. 제안된 GA의 흐름도는 그림 2에서 나타낸다. 본 알고리즘은 일정 시간 이상 지난 경우, 또는 특정 세대 이상에서 가장 우수한 해가 동일한 경우 알고리즘의 실행을 종료한다.



〈그림 2〉 GA 흐름도

4.1 해의 표현

본 연구에서는 0-1 사이의 범위 내에서 균일 분포로 염색체 랜덤 키(Random key) 생성한 후 생성된 염색체의 표현을 통해 의약품 배치의 할당 순서와 할당 Site를 결정한다. 랜덤 키 생성 GA는 Bean(1994)에 의해 처음 제안되었으며 염색체를 이진법 대신 실수로 표현하는 방식이다. 랜덤 키 생성 GA는 다수의 조합 최적화 문제 해결에 유용함이 입증되었다[26,27]. 염색체의 인자는 배치 또는 Site 변경의 기준점을 의미하며

염색체는 $\sum_{i=1}^L Q_i + L - 1$ 길이를 가진 양의 실수로 이루어진 1차원 벡터이다. 인자의 오름차순으로 요소의 순서가 결정된다. 만약 선택된 인자가 Site 변경을 의미하는 경우 이후 배치의 할당은 다음 Site에서 이루어진다. Site 내 배치의 할당 순서가 결정된 후, 각 주문 중 가장 먼저 할당된 배치 기준으로 최종 주문 순서가 결정된다. Site 내에서는 주문의 분할은 일정계획에 이점이 없으므로 Site 내 할당된 동일 주문의 배치 수가 해당 Site에 할당된 주문량이 된다. 예를 들어 그림 3에서는 Site 1에 주문 2를 1 배치 할당하였고, Site 2에 주문 3, 1, 4를 1 배치, 주문 2를 2 배치 할당하였다. 그리고 Site 3에 주문 1, 3을 2 배치 할당하였고, 주문 4를 1 배치 할당하였다.

4.2 유전자 연산자

유전 연산자로는 교차 연산자(crossover operator)와 돌연변이 연산자(mutation operator)가 있다. 교차 연산자는 부모 개체로부터 정보를 혼

합하여 새로운 해를 탐색하는 데 도움을 준다. 다양성을 유지하고, 해 공간 탐색 능력을 향상하는 역할을 한다. 대표적인 교차 연산자로는 1개 또는 2개의 교차점을 두고 그 지점을 기준으로 유전자를 교환하는 1점 교차 방법과 2점 교차 방법이 있다. 본 연구의 교차 연산은 각 유전자 위치에서 두 부모 중 어느 부모의 염색체를 물려받을지 무작위로 결정하는 균일 교차 방법을 사용하였다. 돌연변이 연산자는 개체의 유전자를 무작위로 변이시켜 다음 세대의 다양성을 유지하고 새로운 해를 탐색하는 데 도움을 준다. 일반적으로 돌연변이 연산은 매우 작은 확률로 적용된다. 이는 전체 개체 집단 중에 매우 적은 수의 개체에만 돌연변이가 적용되도록 하여, 전체적인 개체의 품질을 유지하면서도 새로운 해를 발견할 수 있도록 한다. 대표적인 돌연변이 연산자로는 임의의 두 원소를 교환하는 교환방법(exchange method)과 임의의 원소를 다른 위치에 삽입하는 삽입 방법(insert method)이 있다. 본 연구에서는 삽입 방법을 통해 선택된 임의의

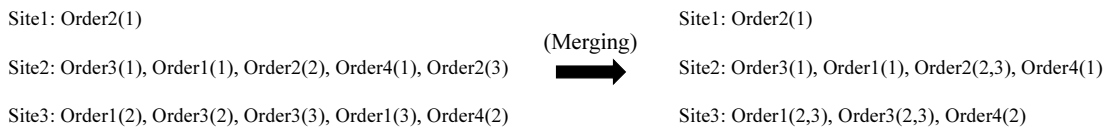
Step 1 (Encoding)

Batch of order1	Batch of order1	Batch of order1	Batch of order2	Batch of order2	Batch of order2	Batch of order3	Batch of order3	Batch of order3	Batch of order4	Batch of order4	Site change	Site change
0.29	0.62	0.98	0.11	0.44	0.30	0.78	0.84	0.23	0.38	0.99	0.59	0.17

↓ (Sorting)

0.11	0.17	0.23	0.29	0.30	0.38	0.44	0.59	0.62	0.78	0.84	0.98	0.99
Batch of order2	Site change	Batch of order3	Batch of order1	Batch of order2	Batch of order4	Batch of order2	Site change	Batch of order1	Batch of order3	Batch of order3	Batch of order1	Batch of order4

Step 2 (Decoding)



〈그림 3〉 염색체 부호화 및 해석

원소를 0에서 1 사이의 임의의 실수로 바꾸는 삽입 방법을 사용한다.

4.3 적합도 평가 및 유전자 선별

유전자 선별은 다음 세대를 위해 좋은 염색체를 선택하는 연산자로서 선별 과정에서 편향을 조절하는 것이 중요하다. g 세대에서 선택 전 k 번째 염색체의 적합도를 F_g^k 로 표기한다. 본 연구에서는 식(16)과 같이 알고리즘에 의해 얻어진 해결책의 적합도를 계산하였으며 Liu(2013)의 연구를 참고하였다.

$$F_g^k = w + \max_{s \in \{1, \dots, e\}} T_g^s - T_g^k \quad (16)$$

식(16)에서 T_g^s 는 g 세대의 s 번째 염색체의 총 납기 지연을 의미하며 e 는 g 세대의 염색체 수이다. 총 납기 지연이 적을수록 염색체의 적합도는 크다. w 는 임의의 상수로서 모든 염색체의 적합도가 0이 되는 것을 방지한다. 일반적으로 염색체의 적합도가 증가할수록 부모로 선택될 기회가 더 많은 것이 좋다. 본 연구에서는 다음 세대의 부모 해를 선별하는 방법으로 룰렛 휠 샘플링 방식을 사용한다. 룰렛 휠은 각 개체가 선택될 확률을 개체의 적합도에 비례하여 할당하는 방식이다. 이것은 각 개체를 룰렛 휠 상의 한 부분으로 간주할 수 있다. 적합도가 높을수록 해당 부분이 커지고, 따라서 선택될 확률이 높아진다. 휠은 부모 해의 염색체 수만큼 회전하며 회전에서 휠의 마커 아래에 있는 염색체가 다음 세대의 부모 염색체로 선택된다.

V. 실험 및 결과분석

본 논문의 GA의 성능 확인을 위해 수리모형 [P] 및 전문가 휴리스틱 (Expert Heuristic)과 비

교한다. 전문가 휴리스틱은 M사 작업자의 일정 계획 방식에 따라 일정계획을 도출하는 휴리스틱 알고리즘이다. 전문가 휴리스틱과 비교하여 제안된 GA가 M사의 일정계획을 얼마나 개선했는지 확인할 수 있다. 비교를 위해 문제 규모에 따라 의뢰주문 수(J), Site 수(L) 그리고 의뢰주문의 배치 생성 분포의 차이를 두어 3개의 테스트 세트를 생성하였으며 이를 표 2에서 나타낸다. 이 외 의뢰주문의 총 배양, 본 배양, 정제의 처리시간 및 작업 준비시간은 M사의 공정 데이터에 기반하여 생성하였다. 모의실험을 통해 GA의 적절한 파라미터를 결정하였으며 이를 표 3에서 나타낸다. 각 테스트 세트당 5개의 문제를 실험하였으며 문제별 총 납기 지연 및 결과 도출 시간을 확인하였으며 총 납기 지연에 대하여 상대 지연 지수(Relative Delay Index; RDI)를 통해 차이를 나타내었다. RDI는 0에서 1 사이의 상대적 성능 평가지표로, 0에 가까울수록 좋은 솔루션임을 의미한다.

알고리즘 성능 실험을 위해 AMD Ryzen-7 5800X CPU와 32GB RAM이 장착된 PC를 사용

〈표 2〉 실험 데이터 셋

Test set	Number of Orders	Number of Sites	Number of Batches
Small	30	3	$X \sim U(1,5)$
Medium	50	5	$X \sim U(1,10)$
Large	100	10	$X \sim U(1,20)$

〈표 3〉 GA의 하이퍼파라미터

Parameter	Value
Population size	100(Small), 200(Medium), 300(Large)
Crossover probability	0.9
Mutation probability	0.5
Arbitrary constant for stopping criterion	100
Maximum execution time	3600sec

〈표 4〉 추론 시간(단위 : 초) 및 총 납기 지연(단위 : 일) 비교

Test set	Problem	Total tardiness(RDI)			Execution time		
		CPLEX(1Hr)	Expert Heuristic	GA	CPLEX(1Hr)	Expert Heuristic	GA
Small	1	1333(0.865)	1508(1.000)	215(0.000)	3600.000	0.001	91.938
	2	953(0.735)	1287(1.000)	25(0.000)	3600.000	0.001	74.003
	3	2516(1.000)	1211(0.438)	195(0.000)	3600.000	0.001	68.246
	4	3906(1.000)	1864(0.394)	539(0.000)	3600.000	0.001	59.323
	5	6540(1.000)	1342(0.156)	380(0.000)	3600.000	0.001	128.821
	Average	3049.6(0.920)	1442.4(0.597)	270.8(0.000)	3600.000	0.001	84.4662
Medium	1	16333(1.000)	7740(0.371)	2671(0.000)	3600.000	0.002	1190.445
	2	11481(1.000)	6415(0.487)	1602(0.000)	3600.000	0.003	1072.210
	3	17770(1.000)	8767(0.381)	3213(0.000)	3600.000	0.002	915.666
	4	17291(1.000)	8690(0.426)	2287(0.000)	3600.000	0.003	1211.674
	5	13200(1.000)	7833(0.486)	2745(0.000)	3600.000	0.002	812.349
	Average	15215(1.000)	7889(0.431)	2503.6(0.000)	3600.000	0.0024	1040.469
Large	1	n.a.	38556(1.000)	24615(0.000)	n.a.	0.010	3600.000
	2	n.a.	36476(1.000)	24107(0.000)	n.a.	0.010	3600.000
	3	n.a.	35129(1.000)	22080(0.000)	n.a.	0.009	3600.000
	4	n.a.	40847(1.000)	26230(0.000)	n.a.	0.010	3600.000
	5	n.a.	37132(1.000)	21145(0.000)	n.a.	0.011	3600.000
	Average	n.a.	37628(1.000)	23635.4(0.000)	n.a.	0.010	3600.000

하였다. 수리모형의 추론은 IBM ILOG CPLEX 12.6.3.0을 통해 진행하였으며, NP-hard임으로 1시간 이내의 추론 결과와 비교하였다. GA와 전문가 휴리스틱은 Python3.9 환경에서 진행되었다. 실험 결과는 표 4에서 나타낸다.

실험 결과 GA가 모든 테스트 세트에서 CPLEX 및 전문가 휴리스틱 결과보다 적은 총 납기 지연을 발생시키는 일정계획을 도출하였다. GA 실험 결과와 비교하여 테스트 세트 Small에서 CPLEX는 평균 0.920, 전문가 휴리스틱은 0.597의 RDI값을 보였다. 테스트 세트 Medium에서는 CPLEX는 평균 1, 전문가 휴리스틱은 0.431의 RDI값을 보였다. 테스트 세트 Large에서는 메모리의 한계로 CPLEX는 1시간 이내에 가능해를 도출해내지 못하였으며, 테스트 세트 Large에서 전문가 휴리스틱 대한 GA의 총 납기 지연의 감

소율을 확인한 결과 평균 37.19%의 감소를 확인하였다. 결과 도출 시간 확인 결과, GA는 테스트 세트 Small에서 평균 84.4662초, 테스트 세트 Medium에서는 평균 1040.469초의 추론 시간을 기록했다. 테스트 세트 Large에서는 모든 문제에서 1시간 이내에 정지조건을 만족하지 못하였기 때문에 전부 3,600초를 기록했다. 이는 GA의 하이퍼파라미터의 변경을 통해 정지조건을 수정하거나, 연산자의 교차 연산과 돌연변이 연산의 확률을 조정하여 해결할 수 있으며, 이를 통해 더 나은 GA의 성능 향상을 기대할 수 있다.

VI. 결론 및 향후 연구 방향 제시

본 연구는 CMO에서 바이오의약품 생산 일정 계획 문제를 해결하기 위한 GA를 제안하였으며, 제안된 GA의 성능을 평가하기 위해 수리모

형의 CPLEX 추론 결과와 기존 M사의 작업 방식을 모방한 전문가 휴리스틱의 결과와 비교하였다. GA는 모든 데이터 세트에서 총 납기 지연 최소화 측면에서 CPLEX보다 더 나은 성능을 보였으며 빠른 수렴 특성으로 인해 실제 작업자나 CPLEX보다 빠르게 양질의 결과를 도출할 수 있다. 본 연구는 하위작업으로 분할이 가능한 일정계획 문제에 널리 적용될 수 있다. 이러한 일정계획 문제로는 조립 부품 생산 문제, 컴퓨터 자원 할당 문제가 있으며, 본 연구에서는 CMO에서의 바이오약품 생산 문제만 다루었지만, 다양한 일정계획 문제에도 적용되어 기업의 이익을 증대할 것으로 예상된다. 그러나 규모가 큰 문제에서는 GA가 정해진 시간 내에 충분히 수렴하지 못하는 결과를 보인다. 이러한 이유로는 작업의 순서와 설비에 할당할 양을 결정하기 위해 탐색체 부호화 단계에서 각 작업을 배치로 분할 하기 때문으로 보인다. 배치로 분할된 작업을 부호화하기 때문에 탐색체의 길이가 길어지고 두 개 이상의 탐색체가 같은 적합도를 가질 경우가 발생하여 불필요한 연산이 늘어나는 경우가 발생한다. 이는 문제 규모가 커질수록 GA의 수렴을 방해하여 성능을 저해하는 요소로, 향후 연구에서는 효율적인 탐색체 표현 또는 휴리스틱 알고리즘을 통한 양질의 초기 해등을 통해 해결할 예정이다.

참 고 문 헌

- [1] Harjunkoski, I., C. T. Maravelias, P. Bongers, P. M. Castro, S. Engell, I. E. Grossmann, ... and J. Wassick, "Scope for industrial applications of production scheduling models and solution methods," *Computers & Chemical Engineering*, Vol. 62, pp. 161-193, 2014.
- [2] Pandya, E. J. and K. V. Shah, "CONTRACT MANUFACTURING IN PHARMA INDUSTRY," *Pharma Science Monitor*, Vol. 4, No. 3, pp. 123-144, 2013.
- [3] Rinnooy Kan, A. H. G., "Machine scheduling problems: classification, complexity and computations," Springer Science & Business Media, 2012.
- [4] Lenstra, J. K., and A. H. G. Rinnooy Kan, "Complexity of scheduling under precedence constraints," *Operations Research*, Vol.26, No.1, pp.22-35, 1978.
- [5] Garey, M. R. and D. S. Johnson, "A Guide to the Theory of NP-Completeness," *Computers and intractability* pp.37-79, 1990.
- [6] Tahar, D. N., F. Yalaoui, C. Chu, and L. Amodéo, "A linear programming approach for identical parallel machine scheduling with job splitting and sequence-dependent setup times," *International journal of production economics* Vol.99, No.(1-2), pp.63-73, 2006.
- [7] Eroglu, D. Y. and H. C. Ozmutlu, "Mip models and hybrid algorithms for simultaneous job splitting and scheduling on unrelated parallel machines," *The Scientific World Journal* 2014, 2014.
- [8] Tutumlu, B., and T. Saraç, "A MIP model and a hybrid genetic algorithm for flexible job-shop scheduling problem with job-splitting," *Computers & Operations Research*, Vol. 155, pp.106222, 2023.
- [9] Avgerinos, I., I. Mourtos, S. Vatikiotis, and G. Zois, "Scheduling unrelated machines with job splitting, setup resources and sequence dependency," *International Journal of Production Research* Vol.61, No.16, pp. 5502-5524, 2023.
- [10] Liu, C., "A Hybrid Genetic Algorithm to Minimize Total Tardiness for Unrelated Parallel Machine Scheduling with Precedence Constraints,"

- Mathematical Problems in Engineering, pp.537127, 2013.
- [11] Holand, J. H., “*Adaptation in Natural and Artificial Systems*,” Ann Harbor: University of Michigan Press, 1975.
- [12] Kumar, M., D. M. Husain, N. Upreti, and D. Gupta, “Genetic algorithm: Review and application,” Available at SSRN, Vol.2, No.2, pp.451-454, 2010.
- [13] Muñoz, A. and F. Rubio, “Evaluating genetic algorithms through the approximability hierarchy,” *Journal of Computational Science*, Vol. 53, 101388, 2021.
- [14] Lakhdar, K., J. Savary, L. G. Papageorgiou and S. S. Farid, “Multiobjective long-term planning of biopharmaceutical manufacturing facilities,” *Biotechnology progress*, Vol.23, No.6, pp.1383-1393, 2007.
- [15] Oyebolu, F. B., J. van Lidth de Jeude, C. Siganporia, ... and J. Branke, “A new lot sizing and scheduling heuristic for multi-site biopharmaceutical production,” *Journal of Heuristics*, Vol.23, pp.231-256, 2017.
- [16] Sahling, F. and G. J. Hahn, “Dynamic lot sizing in biopharmaceutical manufacturing,” *International Journal of Production Economics*, Vol.207(C), pp.96-106, 2019.
- [17] Kim, Y. D., S. O. Shim, S. B. Kim, Y. C. Choi and H. M. Yoon, “Parallel machine scheduling considering a job-splitting property,” *International Journal of Production Research*, Vol.42, No.21, pp.4531-4546, 2004.
- [18] Logendran, R. and F. Subur, “Unrelated parallel machine scheduling with job splitting,” *IIE Transactions* Vol.36, No.4, 359-372, 2004.
- [19] Eroglu, Y. D., H. C. Ozmutlu, and S. Ozmutlu, “Genetic algorithm with local search for the un-
- related parallel machine scheduling problem with sequence-dependent set-up times,” *International Journal of Production Research* Vol.52, No.19, pp.5841-5856, 2014.
- [20] Kim, J. G., S. W. Song, and B. J. Jeong, “Minimizing total tardiness for the identical parallel machine scheduling problem with splitting jobs and sequence-dependent setup times,” *International Journal of Production Research*, Vol.58, No.6, pp.1628-1643, 2020.
- [21] Lee, J. H., H. Jang, and H. J. Kim, “Iterative job splitting algorithms for parallel machine scheduling with job splitting and setup resource constraints,” *Journal of the Operational Research Society* Vol.72, No.4, pp.780-799, 2021.
- [22] Zheng, F., K. Jin, Y. Xu and M. Liu, “Unrelated parallel machine scheduling with processing cost, machine eligibility and order splitting,” *Computers & Industrial Engineering*, Vol.171, pp.108483, 2022.
- [23] 심재민, “바이오의약품 제조공정에서의 생산 계획 최적화 모델” 국내 석사 학위 논문 인천대학교 공학대학원, 2024.
- [24] Xing, W. and J. Zhang, “Parallel machine scheduling with splitting jobs,” *Discrete Applied Mathematics*, Vol.103, No.(1-3), pp.259-269, 2000.
- [25] Bean, J. C., “Genetic algorithms and random keys for sequencing and optimization,” *ORSA Journal of Computing*, Vol. 6, No. 2, pp. 154-160, 1994.
- [26] De Andrade, C. E., R. F. Toso, M. G. Resende and F. K. Miyazawa, “Biased random-key genetic algorithms for the winner determination problem in combinatorial auctions,” *Evolutionary computation*, Vol.23, No.2, pp.279-307, 2015.
- [27] Gonçalves, J. F. and M. G. Resende, “Biased random-key genetic algorithms for combinatorial

optimization,” Journal of Heuristics, Vol. 17, No. 5, pp.487-525, 2011.

저자 소개



김지훈(Ji-Hoon Kim)

- 2023년 : 인천대학교 산업경영공학과(공학사)
- 2023년~현재 : 인천대학교 산업경영공학과(석사과정)
<관심분야> 기계 학습, 최적화



김정현(Jeong-Hyun Kim)

- 2022년~현재 : 인천대학교 산업경영공학과(학사과정)
<관심분야> 기계 학습, 최적화



김재곤(Jae-Gon Kim)

- 1994년 : 한국과학기술원 산업 및 시스템공학과(공학사)
- 1996년 : 한국과학기술원 산업 및 시스템공학과(공학석사)
- 2001년 : 한국과학기술원 산업 및 시스템공학과(공학박사)
- 2003년~현재 : 인천대학교 산업경영공학과(교수)
<관심분야> 산업인공지능, 데이터기반 운영최적화