Received January 9, 2021, accepted January 13, 2021, date of publication January 18, 2021, date of current version January 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3052562

A Hybrid GA-SA for the Urgent Patients Disturbed Physical Examination Rescheduling Problem Considering Setup Time

DAN-DAN ZHU, JUN-QING SUN, AND YU ZHAO
Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Tianjin University of Technology, Tianjin 300384, China
Corresponding author: Dan-Dan Zhu (ddzhu0121@163.com)

This work was supported in part by the National Nature Science Foundation under Grant 71501141, and in part by the Science and Technology Ministry's Special Foundation of Developing Important and Large Equipment under Grant 2014YQ120351.

ABSTRACT In the practice of medical services, the occurrence of disturbance events will inevitably interrupt the pre-arranged patient visit sequence and medical resource arrangement, so rescheduling is essential. In this paper, in view of the disturbance event of urgent patients and the setup time of medical equipment that cannot be ignored, we studied the urgent patient disturbance physical examination rescheduling problem that considering setup time. The optimization goal is to minimize the sum of medical equipment's setup time and diagnostic completion time of all patients. In this problem, multiple patients need to be examined in multiple medical equipment, and the setup time of all patients on a medical equipment are sequence-dependent which was rarely considered in the previous medical service scheduling research. One of our contributions is that when constructing the mathematical model for the problem, we first introduced the change on the original patient's visit sequence between rescheduling and initial scheduling should be less than a given upper bound as the constraint to reduce the impact on the original patient. Another contribution is that since the problem addressed is strongly NP-hard, combined the global search performance of the Genetic algorithm (GA) and the local search performance of the Simulated Annealing algorithm (SA), we proposed a hybrid algorithm (HGA-SA) of improved GA and improved SA to solve the problem. Finally, the model and algorithm are verified through extensive simulation experiments, results show that the proposed algorithm has good performance compared with several other existing algorithms.

INDEX TERMS Urgent patient disturbance, setup time, physical examination, medical service rescheduling, GA-SA hybrid algorithm.

I. INTRODUCTION

Health care is one of the most important social and economic challenges faced by every country [1]: Today, people's expectations for medical needs are increasing, health care managers, clinicians, researchers and practitioners in other fields are facing increasing pressure. People also pay more and more attention to preventive medical care, and the requirements for medical services are constantly improving; due to population growth and aging, the number of patients has greatly increased, which has aggravated the contradiction between the limited medical resources and the rapidly growing demand for medical services. As a result, the hospital operation and management face the following challenges: Small and medium-sized hospitals are inherently inadequate in medical resources, while large hospitals have abundant medical resources, but at the same time attract more patients to visit, if the hospital is not well managed, it will lead to unreasonable allocation of medical resources and failure of scheduling. Unreasonable allocation of medical resources such as medical staff, medical equipment, etc. will not only cause problems such as the waste of medical resources, but also cause dissatisfaction of patients, which will lead to medical disputes, however, doctor-patient disputes are essentially a contradiction between medical demand and medical resource shortage [2]. At the current stage, constrained by the cost of medical equipment manufacturing and the education cost of medical students, it is impossible to rely on the investment

The associate editor coordinating the review of this manuscript and approving it for publication was Donghyun Kim.

of a large amount of manpower and material resources to solve the current dilemma in a short time. But, the reasonable scheduling of medical services can not only relieve the pressure of shortage of medical resources and avoid unnecessary waste of resources, but also reduce the contradiction between doctors and patients, and promote the harmony between doctors and patients. Therefore, the use of scientific management and scheduling methods to solve the contradiction between medical resources and medical service needs has important practical significance [3].

The German Industry 4.0 [4] and the China Manufacturing 2025 [5] have accelerated the development of modern information technologies such as computer science, big data, artificial intelligence, and cloud computing, changing the service and production fields, smart home, smart transportation, smart communities, smart medical care, etc. continue to emerge [4], [5]. Especially in the health field, the Internet of Things, big data technologies, cloud computing, and fog computing are revolutionizing electronic health and its entire ecosystem, pushing it to Healthcare 4.0, which provides powerful technical support for advanced medical service management and scheduling [3]. The continuous optimization of the medical service management scheduling method will further improve the service efficiency and quality of medical institutions, enhance the management level of medical institutions, and reduce the public medical cost [6].

Medical service scheduling is an important branch involving computer science and technology, operations research and management, optimization and control, etc. Its systematic and complete theory is the foundation of the development of intelligent medical care and has broad application prospects, therefore, it has received extensive attention from researchers in related disciplines [7]. In the practice of medical services, some unplanned events will inevitably occur, which will disturb the medical scheduling system and make the initial plan not smoothly implemented. Disturbance events can be divided into two types: patient-related and medical resource-related [8]. Among them, patient-related disturbance events include: unplanned patient arriving, urgent patient, patient missed appointments, patients arriving late or early, etc.. Resource-related disturbance events include: medical equipment failure, late arrival of medical staff, insufficient wards, and insufficient medical resources caused by public health emergencies, etc.. The new crown pneumonia epidemic we are experiencing is a public health emergency, which has caused great disturbance to the medical service system. In response to emergencies, we must minimize its interference to the initial scheduling system, so the initial scheduling needs to be repaired or rescheduled.

People are paying more and more attention to preventive medical care, so physical examination also plays an important role in the medical service system [38]. Physical examination is a medical scheduling environment involving multiple medical equipment and patients. In this paper, we will in the physical examination medical service scheduling environment, study the disturbance rescheduling problem of urgent patients considering sequence-dependent setup time (UDPERP-SDS), with the following contributions:

- The setup time was first introduced into the medical service scheduling of emergency patient disturbance. We assumes that as long as the medical equipment examines the patient, there exist setup time, and the setup time is sequence-dependent.
- A mixed integer linear programming model was established to minimize the total diagnostic completion time of all patients including the setup time, the change on the original patient's visit sequence between rescheduling and initial scheduling should be less than a given upper bound first introduced as the constraint to reduce the impact on the original patient.
- Under the premise of satisfying the disturbance constraint, two decisions need to be made, that is, to determine the diagnosis path of each patient and the order of the patients on each medical equipment.
- For the problem addressed is strongly NP-hard, combined the global search performance of the Genetic algorithm and the local search performance of the Simulated Annealing algorithm, we proposed a hybrid algorithm of improved GA and improved SA to solve the problem. The improved genetic algorithm was used for global search incorporated two new constructive heuristics as well as two different decoding schemes. The improved SA was used for local search, taking the total sequence disturbance difference of the original patient between the offspring and the parent as the annealing criterion. Extensive computational experiments were conducted to verify the proposed model and algorithm.

The research work of this paper has a certain significance: (1) The change on the original patient's visiting sequence was constrained to measure the original patient's satisfaction in the rescheduling, which provided a new idea for the study of the disturbance scheduling problem in the medical service scheduling. (2) The research on modeling and algorithm for the addressed problems has made a beneficial discussion for the theoretical methods of medical service rescheduling. (3) In practical sense, the solution of this problem provides an applicable reference method for anti-interference in the process of treating patients, improves the efficiency of hospital operation and management, and makes the doctor-patient relationship more harmonious.

The paper is organized as follows. In section II, we provide a survey of the literatures relevant to our work. In section III, we give a detailed description for the problem addressed in this paper, put forward some hypotheses and construct the mathematical model of the problem. In section IV, an improved Hybrid GA-SA algorithm is proposed to solve the problem. In section V, the improvement of genetic algorithm applied to global search is described in detail. In section VI, the improvement of SA algorithm applied to local search is described in detail, in order to reduce the search space and avoid redundancy, we take the total sequence disturbance difference of the original patient between the offspring and the parent as the annealing criterion. In section VII, we will introduce the experimental design and experimental results. In VIII, conclusion and future work.

II. RELATED WORKS

In recent years, uncertain medical service scheduling is an attractive research field, and there are also many studies, in this section, the related research works will be introduced.

A. UNCERTAIN MEDICAL SERVICE SCHEDULING

As a branch of the scheduling theory, medical service scheduling aims to effectively allocate medical resources (including medical staff, hospital beds, operating rooms, and medical equipment, etc.) to patients to optimize performance standards [9]. The service provider in the medical service system is the hospital, and the medical service demander is the patient. Whether it is a provider or a demander, there are inevitably some uncertainties in practice, and there are some corresponding studies.

1) PATIENT UNCERTAINTY

Mahmoudzadeh *et al.* considered the uncertainty in demand, proposed a mixed-integer programming model for the priority patient medical service scheduling problem. This model considers the priority of the patient, takes waiting time as the optimization goal, introduces the concept of uncertainty budget and studies the robustness of the model. Finally, a numerical comparison is made between the proposed robust model and the deterministic method [10]. Oliveira *et al.* put forward a comprehensive method to combine the priority of patients and the schedule of patients to improve the selectivity of services, and gradually increase the utility of patients according to the time the patients are on the waiting list, balancing the benefits of arranging patients according to their utility and the risk of excessive delays for low-priority patients [11]. Saure *et al.* described a model that arranges patients with different needs and emergency levels into a single resource. This model applied for the dynamic scheduling problem with multi-priority patients and random service time [7]. Marchesi *et al.* considered the uncertainty of patient arrival, a two-stage stochastic planning model was introduced to comprehensively solve the staffing and scheduling problem, so that the doctor's schedule is consistent with the patient's arrival. At the same time, minimize the number of patients waiting for treatment. The final analysis shows that the optimal plan generated by the model is robust to changes in demand and service rate [12]. Latorrenunez *et al.* considered the constraints of operating room resources and the possibility of urgent patients' arrival, studied operating room scheduling problem. They put forward an integer programming model and a meta-heuristic method based on GA, they also transformed the model into a constraint programming model [13]. Heydari *et al.* also took into account the arrival of urgent patients and adopted a two-stage stochastic plan [14]. Castaing *et al.* proposed a two-stage random integer programming to design patient appointment scheduling when the treatment time is uncertain. Aimed to minimize the patient's expected waiting time, put forward a heuristic algorithm to find an approximate solution [15]. Jiang and Tang studied the capacity allocation problem with adding capacity policy in high demand, considering the impact of the uncertainty of patients' appointments on the overload of doctors, a two-layer enumeration algorithm for solving the global joint optimal solution was designed [16]. Feldman *et al.* presented an appointment scheduling model that considered the patient's visit time preference, the patient who made an appointment can cancel the appointment or not to attend the service, a dynamic model considering the reservation status was established, and a heuristic solution method was proposed [17].

2) MEDICAL RESOURCES UNCERTAINTY

Wang *et al.* studied the operating room scheduling problem with uncertain operation time and urgent needs. A stochastic model was established to minimize the total cost of the surgery, which was transformed into a deterministic model, and proposed a heuristic algorithm based on column generation (CGBH) [18]. Zhang *et al.* studied the operating room scheduling problem consisting of multiple operating rooms and downstream surgical intensive care unit (SICU). The uncertainty of operation time and postoperative hospital stay was taken into account, and the goal was to reduce costs related to patients and costs related to hospitals. A two-stage stochastic programming model was proposed, and was transformed into deterministic integer linear programming (DILP) model, adopted heuristic algorithm based on column generation (CGBH) to solve the problem [19]. Jebali and Diabat took into account the uncertainty of the operation time and the stay time in the intensive care unit and ward. They proposed a two-stage stochastic planning, which was solved by sample average approximation technology [20]. Neyshabouri and Berg studied the problem of operating room planning in response to uncertainties related to operating time and length of stay in the intensive care unit, based on two-stage robust optimization method, they adopted column generation method to solve this problem [21]. Legrain *et al.* studied the scheduling problem of nurses under uncertainty. According to the nurse preferences, proposed online random algorithm based on the original dual algorithm of online optimization and the approximate average value of the sample [22]. Wickert *et al.* studied the problem that one or more nurses who have been scheduled cannot be present due to unforeseen events, the reschedule strategy based on different problems parameter relaxation was discussed [23]. Mohammad *et al.* considered the uncertainties of operation time, recovery and postoperative residence time to perform multi-stage and multi-resource surgical scheduling for patients under an open scheduling strategy, and established a MILP, considered three optimization criteria: free time and overtime, surgery delay. A constructive heuristic algorithm and hybrid genetic algorithm for solving medium and large-scale problems were proposed [24]. Lee and Yih modeled the operating room scheduling problem as a flexible job shop scheduling problem, taking into account the access restriction to the recovery bed, they used fuzzy numbers to consider the uncertainty of the operation time, and used two stage decision process and genetic algorithm to solve this problem [25].

By reviewing relevant research on uncertain medical service scheduling, we have noticed that there are a large number of studies on emergency patients, but the existing research has the following shortcomings: (1) Most of the existing studies give absolute priority to urgent patients, and aim to minimize the waiting time of urgent patients, but ignoring the interference of urgent patients on the initial scheduling, which ignores the satisfaction of the original patients who share the same medical resources, and causes the cost waste of scheduled medical resources. That is, in the rescheduling after urgent patients arrival, the disturbance constraints of the original patients in the rescheduling relative to the initial scheduling caused by urgent patients are not considered, which may cause the original patient's visit sequence or completion time to be unrestrictedly delayed. (2) Ignoring the setup time of medical equipment, before medical equipment examines the patient, it needs to perform some setup tasks. The patient's examination part is different and the setup time is also different. Some medical equipment starts to incur costs once they are turned on. The longer the idle time, the more waste of resources. Clearly consider setup time/cost when making scheduling decisions, which can improve scheduling efficiency, eliminate waste, and improve resource utilization.

B. SETUP TIME

In modern manufacturing and service environment, setup time scheduling plays an important role to ensure reliable service. Setup time is the time required to prepare resources (people, equipment) to perform tasks (operations, jobs). In the mid-1960s, research on scheduling problems began to consider setup time, a literature survey on scheduling problems showed that less than 10% of the scheduling problem literature considered setup time [26]. Kopanos *et al.* pointed out that the setup time appeared in a large number of industrial and service applications. For some applications, it may be effective to ignore setup time, but, it will negatively affect the quality of some other scheduling applications' solutions [27]. Regarding the setup time, Allahverdi *et al.* [28] proposed about 50 different application industries, such as printed circuit board assembly [29], wafer probing scheduling [30], [31], ship assembly operation management [32]. Chen *et al.* solved the scheduling problem in the solar cell industry and modeled the problem as a hybrid flow shop scheduling problem with setup time [33]. Ying and Bin took the actual scheduling problem of a sheet metal processing company as an object, and transformed it as a single machine scheduling problem with sequence-related setup time. Minimizing setup time is one of the goals [34]. Park and Seo studied the transporter scheduling problem of ship assembly block operation management, and transformed it into a parallel machine scheduling problem with sequence dependent setup time [35]. Cankaya *et al.* studied the parallel machine scheduling problem with family-related setup time and total weighted completion time minimization, introduced five kinds of new mixed integer linear programming [36]. Setup time also appeared in the medical service scheduling environment, for example, Kramer *et al.* pointed out that when assign patients to the operating room, the operating room must be based on the patient's condition to configure the medical resources required for surgery, generate the corresponding setup time and setup cost [37]. In these industries, it is important to explicitly consider the scheduling with different setup times.

Setup time can be divided into sequence-independent and sequence-dependent. In the medical service scheduling environment, when the setup time of a given patient depends on the previous patient on the medical equipment, it is called setup time sequence-dependent, on the contrary, called setup time sequence-independent.

III. PROBLEM DESCRIPTION AND MODEL

A. UDPERP-SDS PROBLEM DESCRIPTION

Urgent patients disturbance physical examination rescheduling problem that considering sequences-dependent setup time (UDPERP-SDS), described as follow. (1) Initial scheduling stage: m medical equipment with different functions in the hospital can be used to examine patients, n_o original patients make appointments for the use of the m different medical equipment, each original patient j needs to be examined once on each medical equipment k, each patient's physical examination items have a diagnosis time without routing constraint, that is, these patients can access the medical equipment in any order. The diagnosis time p_{jk} of the physical examination items O_{jk} of patient j on medical equipment k is a constant, the time S_{ijk} that medical device k setup for the physical examination item O_{jk} of a certain patient j depends on the physical examination item O_{ik} of the patient i that before patient j. Based on medical service efficiency and patient satisfaction, with the goal of minimizing the total setup time and completion time, sort the visit order of n_o original patients on m medical equipment to form an initial schedule. (2) Rescheduling stage: Before the actual execution of the initial schedule, n_u unplanned urgent patients send requests for the use of these m medical equipment, and n_u urgent patients will also be examined once on each medical equipment, the setup time is also sequence-dependent. In addition, suppose there is a virtual patient 0, m virtual physical examinations items of the virtual patient 0 are ranked before the first patient examined on each medical equipment, and are used to determine the setup time of the first patient. The setup time of all original patients, urgent patients and virtual patients on m equipment is $m(n_o + n_u + 1) \times (n_o + n_u + 1)$ matrix. Since the insertion of urgent patients will destroy the initial schedule, causing dissatisfaction of the original patients and waste of medical resources, although the order of the original patient in the initial optimal scheduling is allowed to be disrupted in the rescheduling, but the amount of disturbance should be minimized. The disturbance cost can be measured by the change of the original patient's visit position in the rescheduling compared to initial scheduling, expressed by $D_{jk}(\alpha, \pi)$, and we give it an upper limit ε. That is, under the constraint $D_{jk}(\alpha, \pi) \leq \varepsilon$, with the optimization goal of minimizing the total completion time $\sum_1^n C_i$ of all patients to reschedule all the patients (the examination sequence of one patient on all medical equipment, and the examination sequence of all patients on one medical equipment).

Inspired by Alizadeh *et al.*, they modeled the medical outpatient appointments scheduling problem with m doctors and n patients in a similar way to the open shop scheduling problem in production scheduling [9]. There is also a similar relationship between patients and medical equipment in the question studied in this paper. Therefore, the UDPERP-SDS problem can also be regarded as an open shop scheduling problem, and we use the three-domain expression $\alpha|\beta|\gamma$ to present the problem as $O|ST_{sd}, D_{jk}(\alpha, \pi) \leq \varepsilon| \sum C_j$, where, α domain is the scheduling environment, O represents the open shop, domain β is the constraint condition, ST_{sd} represents the setup time sequence-dependent, $D_{jk}(\alpha, \pi) \leq \varepsilon$ represents the disturbance constraint, and domain γ is the optimization goal.

B. MODEL ASSUMPTIONS AND NOTATION

1) MODEL ASSUMPTIONS

According to the characteristics of the problem studied in this paper, the basic assumptions of the model are given.

- All patients are independent.
- Both urgent and planned original patients are examined on the same medical equipment.
- All urgent patients arrive at the same time, which is the beginning of rescheduling, defined as time 0 (Otherwise, if urgent patients arrive one after another, multiple rescheduling is required, making the problem more complicated). At this time, the initial scheduling has been made, but the implementation has not yet begun, and all patients are available.
- At any time, each medical equipment can only examine at most one patient, and there are m different medical devices.
- At any time, each patient can only be examined by one medical equipment at most, that is, different physical examination items of the patient are not allowed to overlap in time.
- Each patient has to be checked once in each medical equipment, so each patient has m examination items.
- Each patient's examination item on each medical equipment has a diagnosis time without routing constraints.
- The diagnostic path for each patient is arbitrary.
- Preemption is not allowed. The patient examined on the medical equipment cannot be interrupted.
- Importance and penalty levels of patients are different. Moreover, importance and penalty levels of urgent patients are higher than that of original patients.
- Sequence-dependent setup time are considered for each physical examination item. That is, the setup time of each patient on each medical equipment is related to the medical equipment and the previous patients.
- Assume that there is a virtual patient 0, which is located before the first patient examined on each medical equipment.
- The examination time and setup time of all patients are known positive integers.

2) NOTATION

The indices, parameters and decision variables used to build the mathematical model are described as follow.

Indices

$M = \{1, 2, \cdots, m\}$	The medical equipment set, m is the number of medical equipment.
$J^o = \{1, 2, \cdots, n_o\}$	The original patient set that needs to be diagnosed in the plan.
J^u	A group of urgent patients arrived at the same time at time 0. $J^u = \{n_o + 1, \cdots, n_o + n_u\}$
$J = J_o \cup J_u$	The set of all patients.
n_o	Number of original patients.
n_u	Number of urgent patients.
$n = n_o + n_u$	Number of all patients.
O_{jk}	The examination of patient j on the medical equipment k, $\forall j \in J$, $\forall k \in M$.
0	The virtual patient, for $\forall k \in M, p_{0k} = 0$.
α	Optimal initial scheduling of the original patients.

Parameters

N	A sufficiently large positive integer.
d_j	Expected completion time of patient j, where $j \in J$.
p_{jk}	The examination time of patient j on medical equipment k,where $j \in J$, $k \in M$.
υ_i	Importance factor of patient i, equivalent to the unit completion time cost of patient i.
$\upsilon_{i\in J^o} < \upsilon_{i\in J^u}$	The unit completion time cost of urgent patients is greater than original patients
$Pos_{jk}(\alpha)$	The position of the original patient j examined on machine k in the initial schedule, $j \in J^o$, $k \in M$.
S_{ijk}	The setup time for the patient j immediately behind the patient i on the medical equipment k.
H_k	Patient setup time matrix on medical equipment k, $k = \{1, \cdots, m\}$. We have:

$$H_k = \begin{bmatrix} & 0 & 1 & 2 & \cdots & n \\ 0 & - & S_{01k} & S_{02k} & \cdots & S_{0nk} \\ 1 & - & - & S_{12k} & \cdots & S_{1nk} \\ 2 & - & S_{21k} & - & \cdots & S_{2nk} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ n & - & S_{n1k} & S_{n2k} & \cdots & - \end{bmatrix}$$

Decision variables

C_j	Physical examination completion time of patient j
T_j	Tardiness of patient j. $T_j = \max\{C_j - d_j, 0\}$.
Ts_{jk}	The starting time of patient j on medical equipment k.
ξ_{ijh}	If physical examination item O_{ij} processed before O_{ih} for patient i, then $\xi_{ijh} = 1$; otherwise, $\xi_{ijh} = 0$. Where $\forall i \in J, \forall j, h \in M, j \neq h$.
ζ_{kij}	If physical examination item O_{kj} processed before O_{ij} on medical equipment j, then $\zeta_{kij} = 1$; otherwise, $\zeta_{kij} = 0$. Where $\forall k \in J \cup \{0\}, \forall i \in J, k \neq i, \forall j \in M$.
ϑ_{kij}	If physical examination item O_{kj} immediately processed before O_{ij} on medical equipment j, then $\vartheta_{kij} = 1$; otherwise, $\vartheta_{kij} = 0$. Where $\forall k \in J \cup \{0\}, \forall i \in J, k \neq i, \forall j \in M$.
π	Rescheduling.
$P_{jk}(\alpha, \pi)$	For each original patient $j \in J^o$, define $P_{jk}(\alpha, \pi)$ as the absolute difference between its position in initial schedule α compared to reschedule π on equipment k, $P_{jk}(\alpha, \pi) = \lvert Pos_{jk}(\alpha) - Pos_{jk}(\pi)\rvert$.
$D_{jk}(\alpha, \pi)$	$D_{jk}(\alpha, \pi) = \max P_{jk}(\alpha, \pi)$

C. MATHEMATICAL MODEL

According to the above problem description and assumptions, we have established a mixed integer linear programming (MILP) model, minimizing the total weighted diagnosis completion time as optimization goal. Noted that S_{0ij} is the setup time for the physical examination item O_{ij} of the first patient i on the medical equipment j. If the patient i is the first patient on the medical equipment j, then $\vartheta_{0ij} = 1$. The mathematical proposed model as follow:

$$\min Z = \sum_{i=1}^{n} \upsilon_i C_i \tag{1}$$

$$\text{s.t. } Ts_{ij} + \sum_{l=0, l\neq i}^{n} \left(S_{lij} \times \vartheta_{lij}\right) + p_{ij} < C_i, \quad \forall i \in J, \ \forall j \in M \tag{2}$$

$$Ts_{ij} + \sum_{l=0, l\neq i}^{n} \left(S_{lij} \times \vartheta_{lij}\right) + p_{ij} - N(1-\xi_{ijq}) \leq Ts_{iq}, \quad \forall i \in J, \ \forall j, q \in M, j \neq q \tag{3}$$

$$Ts_{iq} + \sum_{l=0, l\neq i}^{n} \left(S_{liq} \times \vartheta_{liq}\right) + p_{iq} - N \times \xi_{ijq} \leq Ts_{ij}, \quad \forall i \in J, \ \forall j, q \in M, j \neq q \tag{4}$$

$$Ts_{ij} + \sum_{\substack{l=0,\\ l\neq i\neq u}}^{n} \left(S_{lij} \times \vartheta_{lij}\right) + p_{ij} - N(1-\zeta_{iuj}) \leq Ts_{uj}, \quad \forall i, u \in J, i \neq u, \ \forall j \in M \tag{5}$$

$$Ts_{uj} + \sum_{\substack{l=0,\\ l\neq i\neq u}}^{n} \left(S_{lij} \times \vartheta_{kij}\right) + p_{uj} - N \times \zeta_{iuj} \leq Ts_{ij}, \quad \forall i, u \in J, i \neq u, \quad \forall j \in M \tag{6}$$

$$\xi_{ijq} + \xi_{iqj} = 1, \quad \forall i \in J, \ \forall j, q \in M, j \neq q \tag{7}$$

$$\zeta_{iuj} + \zeta_{uij} = 1, \quad \forall i, u \in J, i \neq u, \ \forall j \in M \tag{8}$$

$$\zeta_{lij} + \vartheta_{lij} \geq 0, \quad \forall l \in J \cup \{0\}, \ \forall i \in J, l \neq i, \ \forall j \in M \tag{9}$$

$$\zeta_{lij} + \vartheta_{ilj} \leq 1, \quad \forall i, l \in J, \ \forall i \in J, l \neq i, \ \forall j \in M \tag{10}$$

$$\sum_{i=1, i\neq l}^{n} \vartheta_{lij} \leq 1, \quad \forall l \in J \cup \{0\}, \ \forall j \in M \tag{11}$$

$$\sum_{l=0, l\neq i}^{n} \vartheta_{lij} = 1, \quad \forall i \in J, \ \forall j \in M \tag{12}$$

$$C_i - d_i \leq T_i, \quad \forall i \in J \tag{13}$$

$$Ts_{ij} \geq 0, \quad \forall i \in J, \ \forall j \in M \tag{14}$$

$$\upsilon_{i\in J^o} < \upsilon_{i\in J^u} \tag{15}$$

$$\max \left|Pos_{ij}(\alpha) - Pos_{ij}(\pi)\right| \leq \varepsilon, \quad \forall j \in M, \ \forall i \in J^o \tag{16}$$

$$T_i \geq 0, C_i \geq 0, \quad \forall i \in J \tag{17}$$

$$\zeta_{lij}, \vartheta_{lij} \in \{0, 1\}, \quad \forall l \in J \cup \{0\}, \ \forall i \in J, l \neq i, \ \forall j \in M \tag{18}$$

$$\xi_{ijq} \in \{0, 1\}, \quad \forall i \in J, \ \forall j, q \in M, j \neq q \tag{19}$$

In the above mathematical model. Equation (1) is the optimization goal with smallest total weighted completion time of all patients. Constraint (2) expresses that the patient's completion time is not less than the sum of the patient's starting time, setup time and diagnosis time. Constraints (3) and (4) express the relationship does not need to be consecutive between two physical examination items of patients on two different equipment. The starting time for setup operation of physical examination items O_{iq} is no earlier than the completion time of physical examination items O_{ij}. Constraint (5) and (6) represent the sequence of physical examination items of different patients processed on the same medical equipment no need to be consecutive. After the medical equipment j completes the diagnostic task of the items O_{ij}, it can start the setup operation of the physical examination item O_{uj}. Constraint (7) represents the order relationship between any two physical examination items for the same patient, if $\xi_{ijq} = 1$, then $\xi_{iqj} = 0$; Otherwise, $\xi_{ijq} = 0$ and $\xi_{iqj} = 1$. Constraint (8) represents the sequence of any physical examination item pair (O_{ij}, O_{uj}) on the same medical equipment j, if $\zeta_{igj} = 1$, then $\zeta_{uij} = 0$; Otherwise, $\zeta_{iuj} = 0$ and $\zeta_{uij} = 1$. Constraints (9) and (10) define the constraint relationship between ζ_{lij} and ϑ_{lij}, when considering virtual patient 0, if $\zeta_{lij} = 1$, then $\vartheta_{lij} = 1$ or 0; otherwise, $\zeta_{lij} = 0$, and $\vartheta_{lij} = 1$ or 0. When virtual patient 0 is not considered, if $\zeta_{lij} = 1$, then $\vartheta_{ilj} = 0$; otherwise, $\zeta_{lij} = 0$, and $\vartheta_{ilj} = 1$ or 0. Because the virtual patient cannot be placed after any patient, it is only used to represent the relative setup time of the first patient on each medical equipment. Constraints (11) expresses that after patient l is examined by the medical equipment j, at most one patient can be examined immediately; if l is the last patient examined on medical equipment j, we have $\vartheta_{lij}= 0$. Constraints (12) means that when virtual patient 0 is considered, only one patient can be examined immediately before patient i on medical equipment j. Constraints (13) describes the tardiness for patients. Constraint (14) indicates that all patients should be available for scheduling at time 0 in the rescheduling. Constraint (15) indicates that urgent patients are more important than the original patients, which is equivalent to that the unit completion time cost of urgent patients is higher than that of original patients. Constraint (16) defines the disturbance constraint, the absolute value of the difference between the original patient's position in rescheduling and initial scheduling on any medical equipment is less than the given upper limit. Constraints (17)-(19) respectively describes binary and continuous decision variables.

In the UDPERP-SDS problem, the setup time of a patient depends on the patient previously examined. Therefore, the setup times of each patient are represented by an element of matrix S_{ijk}, the setup times for each patient and medical equipment are sequence-dependent and variable. The lower bound (LB) is proposed for the UDPERP-SDS problem, refer to equation (20).

$$\mathrm{LB} = \sum_{j=1}^{n} \sum_{k=1}^{m} (p_{jk} + \min_{i=1,\ldots,n} \{S_{ijk}\}) \tag{20}$$

This lower bound would be an integral part of the heuristic initialization population for Genetic Algorithms proposed in Section V, in Section VIII, it is also used to measure the solutions' quality. Since the problem we study is NP-hard, developing an effective approximate method to obtain optimal solutions with an acceptable amount of calculation is necessary. Next, a meta-heuristic algorithm based on the hybrid algorithm of Genetic and Simulated Annealing is proposed to address this issue.

IV. PROPOSED HYBRID GA-SA ALGORITHM

The multi-equipment and multi-patient physical examination medical service scheduling problem discussed in this paper is essentially to sort medical equipment and patient physical examination items, which is similar to the open shop scheduling problem (OSSP) in production scheduling. Since open shop scheduling problems are classified as NP-Hard problems, it is usually impractical to obtain the optimal solution of large and medium-sized problems by using precise methods, so effective meta-heuristic methods need to be applied [39]. The Genetic Algorithm (GA) is relatively mature compared to other algorithms. Since GA was used to solve the problem of shop scheduling, GA has attracted much attention because of its better optimization ability and robustness and easy combination with other algorithms. Therefore, this paper chooses genetic algorithm to solve the problem discussed. As is known that GA is an effective and universal search method to solve optimization problems, however, in some complex problems, the performance of single GA is not good. As a result, a variety of hybrid strategies are proposed, traditional heuristics (for instance Tabu Search, TS) are usually as a part of the improvement process and merged into classical GA. A common form of hybrid GA is to use the local improvement process as an additional part of the classic GA recombination and selection cycle. In other words, before the newly generated offspring being inserted into the population, a local improvement strategy is used to achieve a local optimum. Therefore, the genetic algorithm is applied for the global search of the population, and the local improvement programs is applied for the local search of the chromosome. Because of the complementary relationship between local improvement programs and GA, the performance of hybrid genetic algorithms is usually better than any method that operates alone.

We introduced the Simulated Annealing algorithm in the local search, and proposed a hybrid GA-SA search strategy (HGA-SA), under the premise of satisfying the disturbance constraint, the optimization goal is to minimize the setup time and the total diagnosis completion time, and provide high-quality solutions within the allowable calculation time. For the initial population generation, a specific construction heuristic method is proposed; the operation of the genetic operator is improved; in the generated optimal offspring, SA algorithm as the local search strategy is adopted. One difficulty is that because there is no fixed route of treatment in the physical examination service environment, the number of feasible solutions will increase greatly as the instance size increases. Another difficulty is that many redundant solutions may appear in permutation coding, therefore, for a given algorithm, it is very important to filter redundant. In addition, for the clear attention on setup time, refer to Naderi [40], for the redundant solutions in genetic operators and local search, some filters are provided.

The operations provided by the HGA-SA algorithm include: encoding and decoding scheme, population initialization, selection operator, crossover operator and mutation operator, the restart that used to avoid population premature convergence, and the SA method that is adopted in descendants for local search. The corresponding flowchart shown in Figure 1, the pseudo code is described in Algorithm 1.

V. GLOBAL SEARCH—IMPROVED GENETIC ALGORITHM

A. CODING AND DECODING STRATEGY

1) ENCODING STRATEGY

Encoding Strategy: Let the solution sequence be identifiable for algorithm, has a significant impact on the success of any algorithm [41]. Including: rank matrix and permutation list. In the rank matrix, every row represents the sequence of the physical examination items of one patient on different medical equipment, every column represents the sequence of the patients on the same medical equipment. The permutation list is a single-row sequence of all the physical examination items

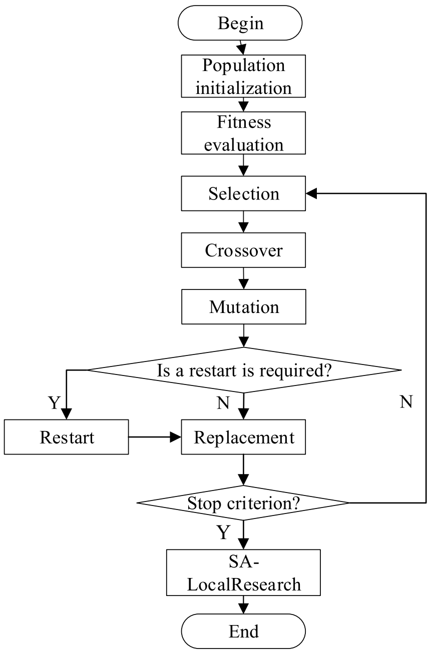


FIGURE 1. Flowchart of the proposed algorithm HGA-SA.

of all patients. In order to effectively deal with the constraints of the UDPERP-SDS problem, we applied the permutation list. The two main elements that need to be clarified are: the order of different patients' physical examination items on each medical equipment and the physical examination path of each patient. Encoding the above two elements into a single-row sequence constitutes the permutation list. A permutation list encodes these two main elements into a single-row array. The benefits of this approach are: (1) Good adaptability; (2) Rapid understandability; (3) Easy to implement in coding.

Now we specifically describe the permutation list. Since each patient should be processed once on each medical equipment, and the physical examination path of each patient is not repeated, for the UDPERP-SDS problem with n patients and m medical equipment, a feasible sequence consists of $m * n$ elements, that is, the number of elements in a feasible sequence is equal to the total number of all physical examination items, which is equivalent to a chromosome in the genetic algorithm. Starting from equipment 1, the diagnostic order of patient 1 to patient n on medical equipment 1 was indexed successively as 1 to n; the corresponding index on medical equipment 2 is $n + 1$ to $2n$, until the corresponding index on medical equipment m is $(m - 1) * n + 1$ to $m * n$. Taking 3 different medical equipment (m = 3) (Computed Tomography, Magnetic Resonance Imaging, Electrocardiogram) and 3 patients (n = 3) as an example (TABLE 1), the corresponding single-row array of 9 elements is shown in Figure2.

There are two shortcomings of permutation coding. In permutation coding, there may be a large number of possible redundant solutions. Although the position of the physical examination items in the sequence changes, the same solution will still be obtained in the end. On the other hand, the acquisition of repeated solutions obviously consumes the execution time of the algorithm. Therefore, this serious shortcoming makes the algorithm unable to make full use of the search space. It is very important to filter redundant solutions in the design of the algorithm.

Algorithm 1: HGA-SA Algorithm Pseudo Code

Data: *p, setup, popsize, MaxGen,* selection(.), crossover(.), mutation(.), *pumt,* restart(.), SA_LocalSearch(.)
Result: *A sequence* $\Pi := (O_{11}, O_{21}, \cdots .O_{nm})$

pop ← CreatePopulation (*popsize*);
fitness ← CalculateFitness(*pop*, *p*, *setup*);
while $k \leq MaxGen$ **do**
  parent1, *parent2* ← selection(*pop*, *fitness*);
  offspring ← crossover(*parent1*, *parent2*);
  if *random()*≤ *pmut* **then**
    offspring ← mutation (*offspring*);
  end
  if *the criterion for performing the restart is satisfied* **then**
    pop, *fitness* ← restart(*pop*, *fitness*);
  end
  pop, *fitness* ← replacement(*offspring*);
  $k \leftarrow k + 1$;
end
Π ← the best solution ∈ *pop*;
Π ← SA_LocalSearch (Π);

TABLE 1. Patients' diagnosis times in the instances.

Patient j	Medical Equipment i CT	MRI	EKG
Patient 1	3	5	3
Patient 2	4	5	2
Patient 3	4	6	2

Index	1	2	3	4	5	6	7	8	9
Examination Items	O_{11}	O_{21}	O_{31}	O_{12}	O_{22}	O_{32}	O_{13}	O_{23}	O_{33}
Medical Equipment	CT	CT	CT	MRI	MRI	MRI	EKG	EKG	EKG
Patient	1	2	3	1	2	3	1	2	3

FIGURE 2. Permutation list based on physical examination items.

2) DECODING STRATEGY

Since there is no priority relationship between the physical examination items of each patient in the medical examination service scheduling studied in this paper, the solution space is much larger. Therefore, the decoding operator can reduce the search area and improve the efficiency of the algorithm without excluding the optimal solution. In this paper, we mainly consider two decoding schemes: direct decoding and indirect decoding. We will first introduce the main ideas and decoding steps of the two schemes, and then compare and verify the

Index	8	7	1	2	9	3	4	6	5
Examination Items	O_{23}	O_{13}	O_{11}	O_{21}	O_{33}	O_{31}	O_{12}	O_{32}	O_{22}

FIGURE 3. **A feasible sequence for the instance (n = 3, m = 3).**

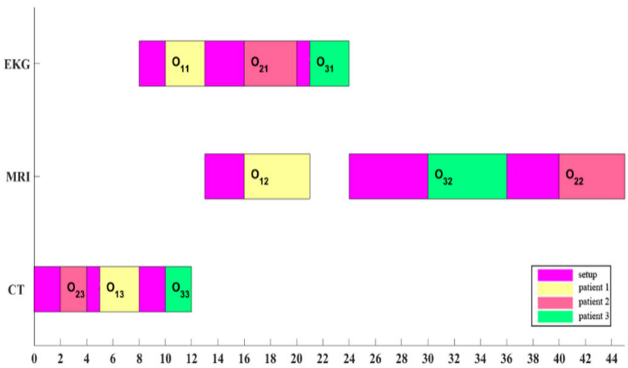


FIGURE 4. **Feasible scheduling Gantt chart based on direct decoding.**

performance of the two schemes through a large number of simulation data experiments.

Direct Decoding(also known as semi-active scheduling): In this kind of decoding, the permutation list is scanned from left to right, and the scanned operations are directly put into the scheduling. Returning to the example above again, the data in the Table 1 represents the diagnostic time of three patients on three medical equipment [42]. The setup time of each medical equipment for the patients examined on it depends on the order of the patients. Considering the virtual patient 0, the setup time of the three patients on each medical equipment is a $4 * 4$ matrix, as follows:

$$H_{CT} = \begin{bmatrix} - & 2 & 4 & 5 \\ - & - & 3 & 4 \\ - & 4 & - & 1 \\ - & 6 & 2 & - \end{bmatrix}$$

$$H_{MRI} = \begin{bmatrix} - & 3 & 2 & 4 \\ - & - & 1 & 6 \\ - & 1 & - & 1 \\ - & 2 & 4 & - \end{bmatrix}$$

$$H_{EKG} = \begin{bmatrix} - & 1 & 2 & 3 \\ - & - & 1 & 2 \\ - & 1 & - & 4 \\ - & 1 & 5 & - \end{bmatrix}$$

For this example, we randomly obtain a feasible sequence (8, 7, 1, 2, 9, 3, 4, 6, 5), as shown in Figure 3, and using the direct decoding method to decode the sequence to obtain the corresponding feasible scheduling Gantt chart, as shown in Figure 4, where the completion time of the last physical examination item O_{22} is 45.

Indirect Decoding: In terms of maximum completion time, indirect decoding reduces the search space without excluding the optimal solution [43]. The optimal solution of medical examination service scheduling should be an active scheduling, in which the order of medical treatment is: if another medical examination item is not delayed, no medical examination item can be started in advance. However, the active scheduling set with the optimization goal of minimizing the maximum completion time is usually very large and the solution quality is also poor. Delay-free scheduling is a subset of active scheduler and is much smaller than active scheduling, the scheduling quality is usually better. Before giving the steps of indirect decoding, the following definitions are made:

TPA_i : the earliest time that patient i can undergo the next physical examination item;

TEA_j : the earliest time that medical equipment j can be used again.

$$Ts_{ij} = \max\left\{0, \max_{\forall i}\left\{C_{ij}\right\} \max_{\forall j}\left\{C_{ij}\right\}\right\} \tag{21}$$

$$C_{ij} = Ts_{ij} + \sum_{k=0,k\neq i}^{n} (S_{kij} \times Z_{kij}) + p_{ij} \tag{22}$$

Indirect Decoding Procedure:

Step1: Consider a feasible sequence as shown in Figure 2. At the beginning, for $\forall i \in J$, $\forall j \in M$, let $TPA_i = 0$, $TEA_j = 0$.

Step2: Select the physical examination item that is in the first position. According to the corresponding order of the physical examination items in the feasible sequence in Figure 1, select the physical examination items of different patients and different medical equipment, such as O_{ij} and O_{kh}. Scheduling these physical examination items using equation (21) and (22).

Step3: For $\forall i \in J$, $\forall j \in M$, update TPA_i, TEA_j:

$$TPA_i = \max\left\{C_{ij}\right\}, \forall i$$
$$TEA_j = \max\left\{C_{ij}\right\}, \forall j$$

Check the items in Figure 2, if all the physical examination items are scheduled, stop.

Step4: According to the corresponding order of the physical examination items in the feasible sequence, considering an unscheduled physical examination item, such as O_{ij}. Schedule this physical examination item using equation (21) and (22), and then solve the corresponding MILP model to calculate each objective function.

Using the above indirect decoding method to decode the sequence in Figure 3 to obtain the corresponding feasible scheduling Gantt chart, as shown in Figure 5, and the completion time of the last physical examination item O_{22} is 23. From the scheduling Gantt chart of the two decoding schemes in Figure 4 and Figure 5, it can be seen that the scheduling obtained by indirect decoding is an active scheduling with higher quality.

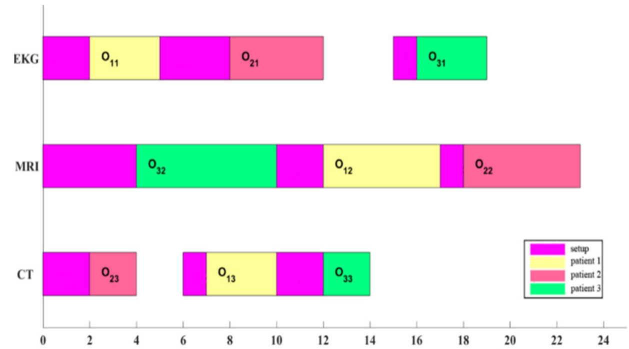


FIGURE 5. **Feasible scheduling Gantt chart based on indirect decoding.**

B. POPULATION INITIALIZATION BASED ON HEURISTIC

Randomly generating the initial population is a common method. But, some studies have pointed out that the application of heuristic methods for population initialization can obtain higher quality solutions. In this paper, we combine the BICH method proposed by Perezgonzalez [44] and the MIH method proposed by Levi [45] to generate 30% of the initial population individuals, and the remaining population individuals are randomly generated. Next, we specifically introduce how these two heuristic methods and their hybrid methods are applied to the problems addressed in this paper.

1) BICH POPULATION INITIALIZATION

The bounded insertion construction heuristic (BICH) algorithm was first used in the flow shop scheduling problem. This method constructs the solution iteratively by means of adding an unscheduled job at the end of the partial sequence of unscheduled jobs. This kind idea is very consistent with the solution of the dynamic scheduling problem after the arrival of urgent patients. After the arrival of urgent patients in a certain iteration, taking a certain physical examination item of the urgent patient as the first choice to append to the end of the partial sequence of original patients.

For the physical examination service scheduling problem with m medical equipment and n patients, constructing a BICH sequence $\Pi := (O_{11}, O_{12} \cdots, , O_{mn})$, at the beginning, Π is empty, then appending a physical examination item in each iteration $o(o = 1, 2, \cdots, n{*}m)$. The selection process of this physical examination item as follows: (1)Select the medical equipment k that first completes the diagnostic task in sequence Π; (2) During the physical examination items that have not been scheduled corresponding to medical equipment k, select the physical examination item O_{jk} that satisfies the following conditions: After adding the physical examination item O_{jk} to sequence Π, the total completion time of Π plus the lower bound (*LB*, calculated by equation(20)) of the remaining unscheduled physical examination items is the smallest. So the basic idea of BICH is to choose an indicator to measure the pros and cons of selecting a specific physical examination item and non-selected. Algorithm 2 gives the pseudo code of the BICH. Where, p is the sample instance, *LB* is the operator for calculating the lower bound, *TCT* is the operator for calculating the total diagnosis completion time.

Algorithm 2: BICH Pseudo Code

Data: TCT (.), LB, p
Result: *A sequence* $\Pi := (O_{11}, O_{21}, \cdots .O_{nm})$
$\Pi \leftarrow \{\}$;
$P \leftarrow copy(p)$;
$M \leftarrow$ *list with time cumulative in each medical equipment;*
$J \leftarrow$ *list with time cumulative in each patient;*
$\Omega_k \leftarrow$ *list with the patients allocated to medical equipment* $k, k \in \{1, \cdots, m\}$;
while $||\Pi|| < n \times m$ **do**
  Medical equipment $k \leftarrow \underset{i \in \{1,\ldots,m\}}{\operatorname{argmin}} M_i$;
  patient $j \leftarrow \underset{j \in \{1,\ldots,n\}, j \notin \Omega_k}{\operatorname{argmin}} TCT(\Pi \cup \{O_{ik}\}, p) + LB(O_{ik}, p)$;
  $\Pi \leftarrow \Pi \cup \{O_{jk}\}$;
  $\Omega_k \leftarrow \Omega_k \cup \{j\}$;
  $P_{jk} \leftarrow 0$;
  update J and M with time of physical examination item O_{jk};
end

2) MIH POPULATION INITIALIZATION

The open shop scheduling problem is usually based on the SPT (shortest processing time) priority rule that sorts operations in descending order. However, this rule will result in higher idleness of the medical equipment. In the open medical service scheduling environment, it will increase the completion time of a given solution.

In the UDPERP-SDS problem, there exist factors such as discontinuity of resource requests between the same medical device for different patients and the same patient for different medical equipment inevitably cause medical equipment to be idle. Specifically, medical equipment has different setting times for different patients and patients cannot be available in time and so on. In order to reduce the idle rate of the equipment, inspired by Levi *et al.* [45], we adopt the Minimal Idleness Heuristic (MIH) method. If a solution has a lower cumulative processing time then, under normal circumstances, this solution also has a lower completion time. The cumulative processing time of patients and medical equipment can be used to measure the idleness of medical equipment. If the cumulative processing time J_j of a given patient j is higher than the cumulative processing time M_i of a given medical equipment i, then the medical equipment i will always wait for patient j to complete the diagnosis on the previous medical equipment k, and then the medical equipment i make the diagnosis for the next physical examination item of patient j, thus the idleness is generated. On the contrary, if the patient j has already completed the diagnosis on another equipment k, its diagnosis on the current device i

Algorithm 3: MIH Pseudo Code

Data: p
Result: *A sequence* $\Pi := (O_{11}, O_{21}, \cdots .O_{nm})$
$\Pi \leftarrow \emptyset$;
$M \leftarrow$ *list with time cumulative in each medical equipment;*
$J \leftarrow$ *list with time cumulative in each patient;*
$\Omega_k \leftarrow$ *list with the patients allocated to medical equipment* $k, \forall k \in \{1, \cdots, m\}$;
while $||\Pi|| < n \times m$ **do**
  Medical equipment $k \leftarrow \underset{i\in\{1,\ldots,m\}}{\text{argmin}}\ M_i$;
    //index from medical equipment finishing first
  patient $j \leftarrow \underset{j\in\{1,\ldots,n\}, j\notin\Omega_k}{\text{argmin}}\ \Phi_{jk}$; *//the best patient not allocated in medical equipment* k*, though the MIH rule*
  $\Pi \leftarrow \Pi \cup \{O_{jk}\}$;
  $\Omega_k \leftarrow \Omega_k \cup \{j\}$;
  update J *and* M *with time of* O_{jk};
end

will not generate idleness. Φ_{ji} is the idleness resulting from assigning patient j to medical equipment i, which is assigned to medical equipment i after patient l, thus resulting in the corresponding setup time S_{lji}. Equation (23) shows how to calculate the idleness, and Algorithm3 shows the pseudo code of MIH.0

$$\Phi_{ji} = \begin{cases} J_i - M_i + S_{lji}, & \text{if } J_j > M_i \\ S_{lji}, & \text{otherwise} \end{cases} \tag{23}$$

3) BICH-MIH POPULATION INITIALIZATION

Framinan [46] adopted a heuristic method with a predictive mechanism to solve the customer order scheduling problem with a total completion time as the optimization goal. The contribution of candidate orders and unscheduled orders to the objective function can be evaluated effectively. Based on this idea, we propose a construction algorithm for the physical examination service scheduling problem, using a weighted aggregation function to combine the two goals of minimizing total completion time and minimizing idle time. In order to weigh these two criteria, we define a weight variable δ, where δ close to 1 means that the idle time is more important, and δ close to 0 means that the completion time is more important. Define Ψ_{ji} as the performance indicator of inserting the physical examination item O_{ji} in the sequence, Φ_{ji} be the expected contribution to minimize idle time. The Ψ_{ji} can be calculated by equation (24).

$$\Psi_{ji} = (1-\delta) \times \begin{pmatrix} TCT\left(\Pi \cup \left\{O_{ji}\right\}, p, setup\right) \\ +LB\left(O_{ji}, p, setup\right) \end{pmatrix} + \delta \times \Phi_{ji} \tag{24}$$

Algorithm 4: BICH-MIH Pseudo Code

Data: TCT (.), LB, p, *setup*, δ
Result: *A sequence* $\Pi := (O_{11}, O_{21}, \cdots .O_{nm})$
$\Pi \leftarrow \emptyset$;
$P \leftarrow copy\ (p)$;
$Setup \leftarrow$ copy *(setup)*;
$M \leftarrow$ *list with time cumulative in each medical equipment;*
$J \leftarrow$ *list with time cumulative in each patient;*
$\Omega_k \leftarrow$ *list with the patients allocated to medical equipment* $k, k \in \{1, \cdots, m\}$;
while $||\Pi|| < n \times m$ **do**
  Medical equipment $k \leftarrow \underset{i\in\{1,\ldots,m\}}{\text{argmin}}\ M_i$;
  patient $j \leftarrow \underset{j\in\{1,\ldots,n\}, j\notin\Omega_k}{\text{argmin}}\ \Psi_{jk}$; *//the best patient not allocated in medical equipment* k*, though the BICH-MIH rule with* δ.
  $\Pi \leftarrow \Pi \cup \{O_{jk}\}$;
  $\Omega_k \leftarrow \Omega_k \cup \{j\}$;
  $P_{jk} \leftarrow 0$;
  update J *and* M *with time of physical examination item* O_{jk};
end

C. GENETIC OPERATOR

1) SELECTION OPERATOR

The selection operator aims to select two parents for the following cross operations. We consider two selection operators:Tournament and Roulette. Tournament selection: according to the binary tournament, 4 individuals are randomly selected from the current population and paired for comparison. *parent*1 is the best of the two individuals selected first, and *parent*2 is the best of the two individuals selected last. Roulette selection: for each individual i, take the inverse of its objective function to get $F_i = 1/TCT_i$, then, calculate the probability $prob_i$ of selecting individual i in the current population, $prob_i = F_i / \sum F_i$, equivalent to the fitness ratio relative to the overall.

2) CROSSOVER OPERATOR

Crossover operators are selected according to the permutation list encoding scheme, an element in the permutation list, that is, a gene on a chromosome corresponds to a patient's physical examination items on a medical equipment. Three kinds of crossover operators are considered: OX, PMX, and CX, introduced as follows.

a: ORDER CROSSOVER (OX)

Select two cutoff point in *parent*1 randomly, put the information between the two cutoff points (including the two cutoff points) into the corresponding position of offspring G, then for *parent*2, filtering out the same information between two

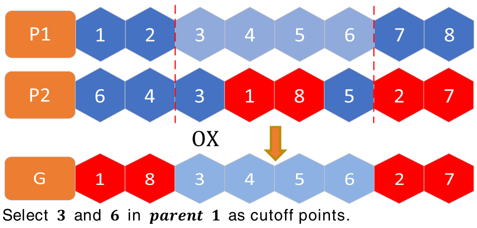


FIGURE 6. The process of OX crossover.

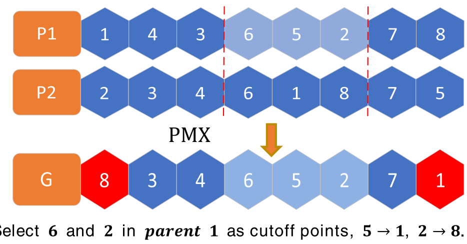


FIGURE 7. The process of PMX crossover.

cutoff points as *parent*1 and add the remaining information of *parent*2 to offspring G in order, generating a feasible offspring. Figure 6 shows the process.

b: PARTIALLY MATCHED CROSSOVER (PMX)

Include three steps, at first, select two cutoff points from *parent*1 randomly, and the offspring will inherit the partial sequence between these two cutoff points at the corresponding position. Second, the information in *parent*2 that do not coincide with *parent*1 are inherited to the offspring G at the corresponding positions. Third, the remaining positions in the offspring G directly inherit the information of the corresponding position in the *parent*2. In short, the last two steps are to correct the infeasible solution. When an infeasible sequence generated, some exchanges will be performed to obtain a feasible solution. Figure 7 illustrates an example for PMX crossover.

c: CYCLE CROSSOVER (CX)

Generating an offspring according to the cycle of the genes in the two parents, constructing solutions that retains their absolute positions. Using an example to describe it in detail, shown in Figure 8, starting from the first gene 1 of $P1$ corresponding to the first gene 8 of $P2$, then gene 8 of $P1$ corresponding to gene 5 of $P2$, and gene 5 of $P1$ corresponding to gene 2 of $P2$, by analogy, a cycle of $1 \rightarrow 8 \rightarrow 5 \rightarrow 2 \rightarrow 4 \rightarrow 7 \rightarrow 1$ is formed, marked in red, genes 3 and 6 are not in this circle, marked in blue, and offspring G inherits all genes the parent $P1$ that appear in the cycle and inherit genes from the parent $P2$ that do not appear in the cycle.

3) MUTATION OPERATOR AND RESTART

The mutation operator randomly modifies the characteristics of the parent by simulating the mutation link in the biological genetic and evolution process to produce offspring. It can

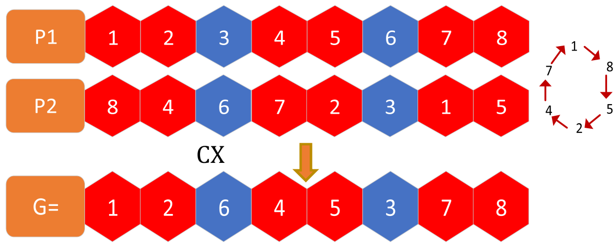


FIGURE 8. The process of CX crossover.

enhance the diversity of the population and has a mechanism to get rid of the local optimum. Mutation operation is based on a swap movement, the basic process is: For each gene of the offspring individuals generated in the crossover operation, obtain a random number P_{rand} between [0, 1], given the mutation probability parameter *pumt*, and if $P_{rand} < pumt$, perform mutation operation.

The purpose of the restart operation is to prevent the search from converging prematurely [53], so as to obtain search results with better parameters. We use the parameter *MaxGen* (maximum generations) to control whether a restart is required, restart is performed if no improvement is made after $0.4 \times MaxGen$ generations. By restarting, it can not only ensure that the new solution obtains good genetic material from the optimal solution, but also have a hopeful search space.

VI. LOCAL SEARCH–IMPROVED SA STRATEGY

There are three criteria for updating the population: the simple replacement criterion of the classical genetic algorithm (Simple), the simulated annealing criterion (SA), and the hill climbing criterion (HC). The above standards are based on elitism: the optimal value of the objective function is reserved for the next generation. Among them, simple replacement usually replaces the worst solution in the current population with the generated offspring, which can guide the obtaining of the local optimum. The HC is to replace the parent with the generated offspring only if the solution of the generated offspring is better than the solution of the parent to maintain population diversity.

For the purpose of improving the algorithm performance, we adopt a hybrid strategy that GA global search and SA local search. The simulated annealing algorithm has the advantages of simple, fast and effective, local search is performed according to the SA improvement strategy raised by Naderi *et al.* [40], it used filters to reduce the search space to avoid redundancy, and is operated in the optimal offspring produced by genetic algorithm. The applied SA can be briefly described as follows: we insert offspring according to a specific probability (follow Boltzmann distribution) related to temperature. The basic idea of SA is that beginning with an initial solution, a series of *Shift* operation are performed until meet the stopping condition. During the neighborhood of the current best parental sequence $parent_{best}$, a new sequence $P_{offspring}$ is generated through the *Shift* operator. The randomly selected operation is repositioned to the

Algorithm 5: SA_LocalSearch Pseudo Code

```
Take a solution parent_best;
Initialize t = T_0;
Counter = 0;
While counter ≤ CT do
    for i = 1 to FN do
        Generate a new neighbor P_offspring from
          current solution parent_best;
        ΔC = TSD(P_offspring) − TSD(parent_best)
        if ΔC ≤ 0 do
            parent_best = P_offspring;
            Update the best solution so far found;
        elseif exp(ΔC/t) < rand do;
            parent_best = P_offspring;
        endif
    endfor
    if the best solution improved in this temperature do
        counter = 0;
    else
        counter = counter + 1;
    endif
    t = 0.97t;
endwhile
```

given sequence by *Shift* operator. Whether to adopt the new sequence is determined by another random rule ϖ, which is controlled by the temperature parameter t, refer to equation (25), if $\Delta C \leq 0$, permutation $P_{offspring}$ is adopted. Otherwise, the probability of permutation $P_{offspring}$ being adopted is $exp(\Delta C/t)$.

$$\varpi = e^{-\Delta C/t} \tag{25}$$

where, ΔC represents the total sequence disturbance (TSD) gap between the optimal parent $parent_{best}$ and the generated offspring $P_{offspring}$ for the original patient, namely:

$$\Delta C = TSD\left(P_{offspring}\right) - TSD\left(parent_{best}\right) \tag{26}$$

Therefore, in the search process, we guarantee an alternative mechanism with selection pressure and stochastic to avoid falling into a local optimal state. At each temperature t, SA does a fixed number (FN) of neighborhood moves. The starting temperature T_0 is set to a higher value ($T_0 = 100$), as the iteration proceeds, the temperature t_i attenuates geometrically, $t_i = \rho t_{i-1}$, the attenuation rate ρ is set to 0.97 in the experiment, the SA continues until there is no improvement in the predetermined consecutive temperatures (CT). Algorithm 5 gives the corresponding pseudo code. This local search strategy reduces the search space and avoids generating redundant solutions. Based on preliminary calculation experiments, we find that the introduction of local search brings better results.

TABLE 2. Experimental parameters.

Parameters	Levels		
Maxgen	200	400	600
Popsize	30	60	100
Select	Rank	Tournament	Roulette
Pmutation	0.03	0.04	0.07
Crossover	CX	OX	PMX
Replace	SA	Simple	HC
Restart	1	0	—

VII. COMPUTATIONAL EXPERIMENTS AND RESULTS

A. EXPERIMENTAL DESIGN

In this part, the MILP model constructed and the HGA-SA algorithm proposed in this paper will be evaluated through a large number of computational experiments. We describe the experimental design of the optimal parameters for the algorithm proposed in this paper. We will introduce evaluation indicators, test instances and algorithm parameters.

1) EVALUATION INDICATORS

In addition to evaluating the proposed hybrid GA-SA algorithm, we also compared and evaluated several other existing algorithms for solving the problem of minimizing the maximum completion time. For the instance i, we use the relative percentage deviation (RPD) between the solution sol_{ik} gained by method k and the lower bound (LB_i, refer to equation (20)) of instance i as the statistic for experimental analysis. RPD calculated by equation (27).

$$RPD_{ik} = \frac{sol_{ik} - LB_i}{LB_i} \times 100 \tag{27}$$

2) TEST INSTANCES AND PARAMETER SELECTION

Since the physical examination medical service scheduling environment in this paper is similar to the open shop scheduling in production scheduling, we modify the following classic test instances for experimental design. Refer to Guéret [47], let the number of patients n equal to the number of medical equipment m, $n = m \in [3, 4, 5, 6, 7, 8, 9, 10]$, then the diagnosis time of all patients on all equipment forms a square matrix, the diagnosis time obeys uniform distribution $U[1, 999]$, 10 different test questions are set up for each group of instances, a total of 80. Refer to Taillard [48], let the number of patients n equal to the number of medical equipment m, $n = m \in [4, 5, 7, 10, 15, 20]$, the diagnosis time obeys uniform distribution $U[1, 100]$, 10 different test questions are set up for each group of instances, a total of 60. Refer to Bruckner [49], let the number of patients n equal to the number of medical equipment m, $n = m \in [3, 4, 5, 6, 7, 8]$, the diagnosis time obeys uniform distribution $U[1, 500]$, 10 different test questions are set up for each group of instances, a total of 60. These test instances did not consider the factor of setup time, so referring to the test instances proposed by Naderi [40], generating setup time randomly, the setup time of the first five question of each

TABLE 3. All experimental tests.

Test	Maxgen	Popsize	Select	Pmutation	Crossover	Restart	Replace
T1	200	30	Tournament	0.03	OX	1	Simple
T2	400	30	Roulette	0.04	PMX	1	HC
T3	600	30	Rank	0.07	CX	1	SA
T4	600	60	Tournament	0.04	OX	1	HC
T5	200	60	Roulette	0.07	PMX	1	SA
T6	400	60	Rank	0.02	CX	1	Simple
T7	600	100	Tournament	0.02	PMX	1	SA
T8	200	100	Roulette	0.04	CX	1	Simple
T9	400	100	Rank	0.07	OX	1	HC
T10	200	30	Tournament	0.07	CX	0	HC
T11	400	30	Roulette	0.02	OX	0	SA
T12	600	30	Rank	0.04	PMX	0	Simple
T13	400	60	Tournament	0.07	PMX	0	Simple
T14	600	60	Roulette	0.02	CX	0	HC
T15	200	60	Rank	0.04	OX	0	SA
T16	400	100	Tournament	0.04	CX	0	SA
T17	600	100	Roulette	0.07	OX	0	Simple
T18	200	100	Rank	0.02	PMX	0	HC

set test instances are uniformly distribution with $U[1, 499]$, and the setup times of the last five question of each set test instances uniformly distribution with $U[500, 999]$.

Choosing appropriate parameters plays a vital role for the performance of the HGA-SA. In our experimental design, the initial temperature $T_0 = 100$, the temperature attenuation rate $\rho = 0.97$, other 7 parameters listed in Table 2 are considered, each factor has two or three levels, the optimal level of each parameter is found by the Taguchi method [52], with 7 parameters and 18 tests, Table 3 describes each test. Let $n \in \{3, 4, 5, 6, 7, 8\}$, 30 test questions are randomly selected as samples from 192 available test questions.

Figure 9 uses the average relative percentage deviation (ARPD) (x-axis) to describe the main effects of each parameter (y-axis). For each test instance, the algorithm was run 10 times and the average result was tallied. The best parameters of HGA-SA are restart, Popsize = 100, tournament, CX, Pmutation = 0.07, SA replacement, Maxgen = 200. The variance analysis (ANOVA, [50]) for ARPD is shown in Table 4. We can observe that only crossover and restart operations significantly improve the quality of the solution with statistical significance.

B. EXPERIMENTAL RESULTS AND ANALYSIS

The algorithm is implemented in MATLAB 2018a and runs on a PC with a 2.6GHz Intel(R) Core (TM)i7-6700HQ CPU and 16 GB RAM memory. The implemented methods include:

TABLE 4. Analysis of variance for parameter ARPD.

Parameter	GL	QM(Aj.)	SQ(Aj.)	SQ Seq	F	p-value
Maxgen	2	67.36	121.2	121.2	0.06	0.93
Popsize	2	78.25	149.47	149.47	0.09	0.82
Select	2	582.90	1104.25	1104.25	0.37	0.53
Pmutation	2	349.57	680.27	680.27	0.39	0.71
Crossover	2	5977.21	11930.13	11930.13	4.57	**0.05**
Replace	2	487.25	983.64	983.64	0.45	0.65
Restart	1	9807.23	9807.23	9807.23	9.91	**0.04**

(1) BICH: Bounded insertion construct heuristic.
(2) MIH : Minimum insertion heuristic.
(3) HGA-SA(DD): Directly decode HGA-SA.
(4) HGA-SA(ID): Indirect decoding HGA-SA.
(5) EH: The electromagnetic heuristic, the best method so far for the similar problem [51].
(6)MILP: Mixed integer linear programming [51].

As for the meta-heuristic algorithm for comparison, compute the average result of each algorithm running 10 times for each instance. The ARPD values of all cases are given in Table 5, include low setup time and high setup time. According to the obtained results, it can be seen that the meta-heuristic methods are better than the constructive heuristic method (bold data) in all cases. Regarding the setup time, it is obvious that the solution provided by the instance with a high setup time (blue data) has a smaller ARPD value. ANOVA is used to verify the statistical significance of the difference in ARPD, and the p-values obtained in this way approach

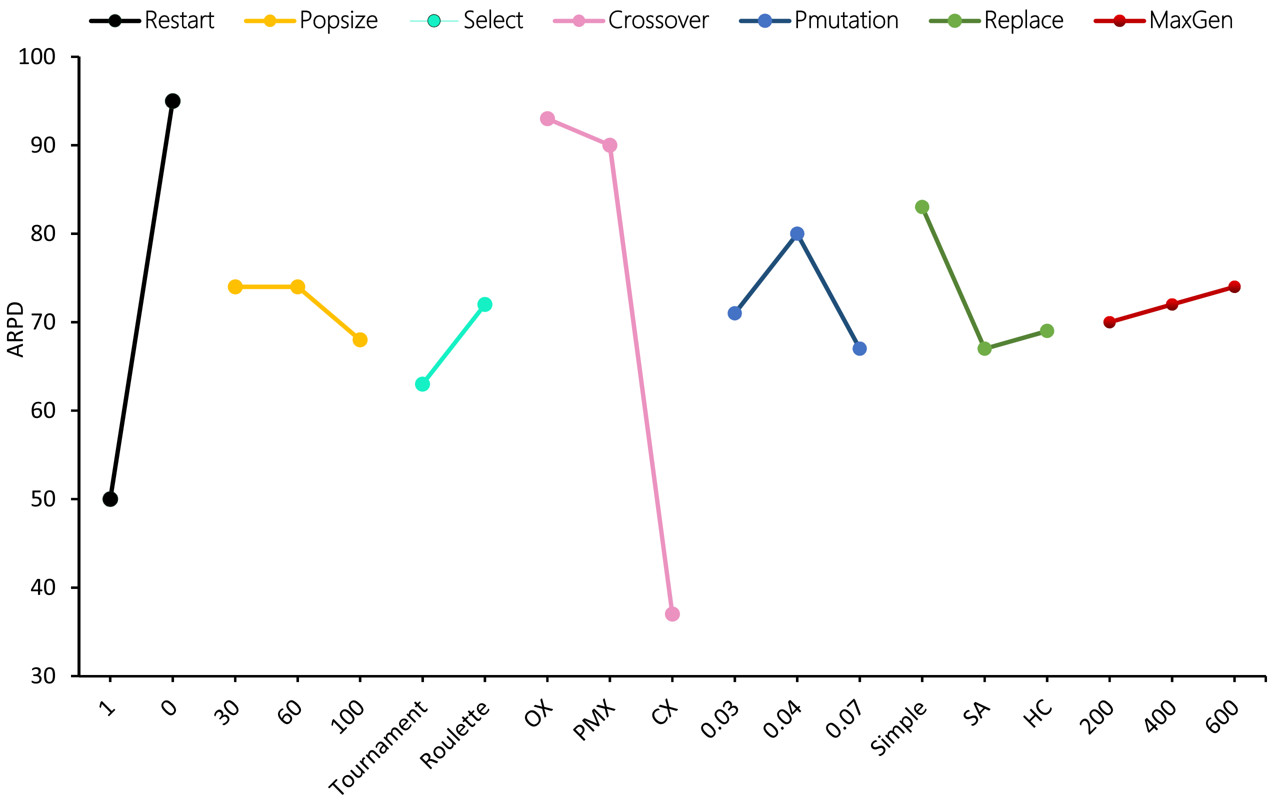


FIGURE 9. Effect of different levels of parameters.

TABLE 5. ARPD values of all methods at different setup times.

Method	MIH		BICH		EH		HGA-SA(DD)		HGA-SA(ID)		MILP	
Setup	High	low	High	Low	High	Low	High	Low	High	Low	High	Low
Guéret												
G_3*3	**28.04**	**39.26**	**48.93**	**56.96**	16.11	22.45	19.45	25.02	27.75	38.37	14.25	22.63
G_4*4	**27.23**	**54.78**	**61.82**	**70.26**	19.58	35.21	9.34	29.56	20.03	57.38	20.06	33.89
G_5*5	**29.76**	**55.44**	**74.56**	**110.23**	23.97	36.76	12.02	33.17	24.05	55.43	22.75	33.56
G_6*6	**36.78**	**69.23**	**94.35**	**161.68**	31.56	57.72	15.88	45.12	26.89	67.45	32.12	50.09
G_7*7	**34.56**	**68.34**	**116.22**	**180.45**	29.12	59.47	17.36	44.78	28.54	77.43	37.44	61.24
G_8*8	**38.83**	**83.43**	**133.12**	**200.04**	32.87	70.34	17.34	53.56	28.34	66.67	63.75	69.77
G_9*9	**37.34**	**88.42**	**135.65**	**230.47**	34.51	82.33	17.23	62.37	30.05	81.23	55.77	100.23
G_10*10	**38.56**	**95.59**	**155.97**	**311.49**	35.10	86.39	16.45	67.38	28.76	85.90	114.24	152.74
Taillard												
Tai_4*4	**26.28**	**77.57**	**55.29**	**102.44**	19.53	48.17	10.07	40.14	24.55	62.21	19.38	45.78
Tai_5*5	**39.22**	**101.67**	**79.37**	**160.88**	29.01	64.27	14.24	65.74	27.32	87.39	22.45	55.36
Tai_7*7	**39.74**	**110.45**	**106.89**	**231.92**	35.17	88.49	20.01	72.53	27.46	107.44	42.68	86.34
Tai_10*10	**42.24**	**127.24**	**142.17**	**333.47**	36.49	109.12	16.38	88.03	28.21	111.94	112.09	283.55
Tai_15*15	**36.87**	**132.69**	**219.85**	**514.74**	35.37	115.47	21.45	97.39	24.32	126.93	4.18	561.93
Tai_20*20	**34.33**	**129.66**	**267.48**	**715.35**	33.34	120.34	16.67	99.78	23.77	137.22	4.02	132.56
Brucker												
B_3*3	**52.31**	**44.77**	**95.33**	**68.49**	31.15	28.33	31.02	30.41	74.64	47.94	71.74	61.22
B_4*4	**72.42**	**61.32**	**109.84**	**104.66**	42.53	40.17	38.78	33.64	71.28	54.48	66.18	53.73
B_5*5	**92.76**	**90.58**	**166.57**	**129.27**	54.36	56.81	50.29	52.54	74.27	71.33	61.16	68.28
B_6*6	**92.04**	**96.06**	**164.33**	**190.29**	68.49	76.20	56.28	65.84	96.93	80.82	64.98	67.22
B_7*7	**108.28**	**110.23**	**227.45**	**229.59**	86.88	87.73	71.08	69.73	102.99	103.74	83.41	85.22
B_8*8	**106.57**	**113.44**	**274.48**	**230.37**	98.92	94.85	77.02	76.19	106.29	110.79	102.33	115.85
Min	**26.28**	**39.26**	**48.93**	**56.96**	16.11	22.45	9.34	25.02	20.03	38.37	4.02	22.63
Avg	**50.71**	**87.51**	**136.48**	**216.65**	39.70	69.03	27.42	57.65	44.82	81.60	50.75	107.06
Max	**108.28**	**132.69**	**274.48**	**715.35**	98.92	120.34	77.02	99.78	106.29	137.22	114.24	561.93

zero. Figure 10 shows the mean plot of all methods with a confidence interval of $\alpha = 0.05$. It can be observed that the differences between the ARPD values during all methods is statistically significant.

TABLE 6. The average computation time of each method under different instance sizes and different type setup time.

Method	EH		HGA-SA(DD)		HGA-SA(ID)		MILP	
Setup	High	Low	High	Low	High	Low	High	Low
Guéret								
G_3*3	8.97	9.02	0.21	0.21	0.34	0.34	0.10	0.08
G_4*4	9.62	9.47	0.32	0.32	0.37	0.37	9.21	1.84
G_5*5	12.37	12.78	0.97	0.81	1.52	1.52	4375.82	330.32
G_6*6	17.69	18.17	1.29	1.22	2.26	2.30	3659.72	3659.72
G_7*7	26.03	26.66	2.48	2.41	4.95	5.03	3658.43	3655.43
G_8*8	29.78	29.97	2.94	2.83	12.76	13.07	3654.24	3654.24
G_9*9	39.67	40.01	5.26	5.26	26.85	26.94	3654.02	3654.02
G_10*10	46.57	47.04	10.03	10.25	49.12	49.85	3654.87	3654.87
Taillard								
Tai_4*4	12.94	12.73	0.41	0.41	0.75	0.79	6.48	0.94
Tai_5*5	18.87	19.05	0.76	0.76	1.98	1.98	3655.07	275.82
Tai_7*7	32.15	32.26	2.98	3.02	11.89	12.02	3654.77	3654.44
Tai_10*10	57.76	58.28	17.07	17.24	93.89	95.56	3654.03	3654.74
Tai_15*15	108.69	116.79	43.48	58.68	980	998.37	3654.03	3654.14
Tai_20*20	297.95	360.75	850.28	920.36	7950	7999.74	3654.08	3654.14
Brucker								
B_3*3	8.53	8.75	0.19	0.17	0.30	0.28	0.03	0.03
B_4*4	14.01	13.18	0.30	0.29	0.68	0.65	0.10	0.16
B_5*5	18.34	19.24	0.68	0.69	1.85	1.87	10.95	11.25
B_6*6	23.59	24.17	1.47	1.41	4.74	4.71	2092.29	3083.91
B_7*7	33.02	29.95	2.92	2.89	11.62	11.53	3659.04	3659.18
B_8*8	39.79	38.28	5.76	5.71	26.37	26.16	3655.45	1271.26
Min	8.53	8.75	0.19	0.17	0.30	0.28	0.03	0.03
Avg	42.82	46.33	47.49	51.75	459.11	462.65	2518.10	2076.53
Max	297.95	360.75	870.28	920.36	7980.58	8022.74	4375.82	3659.72

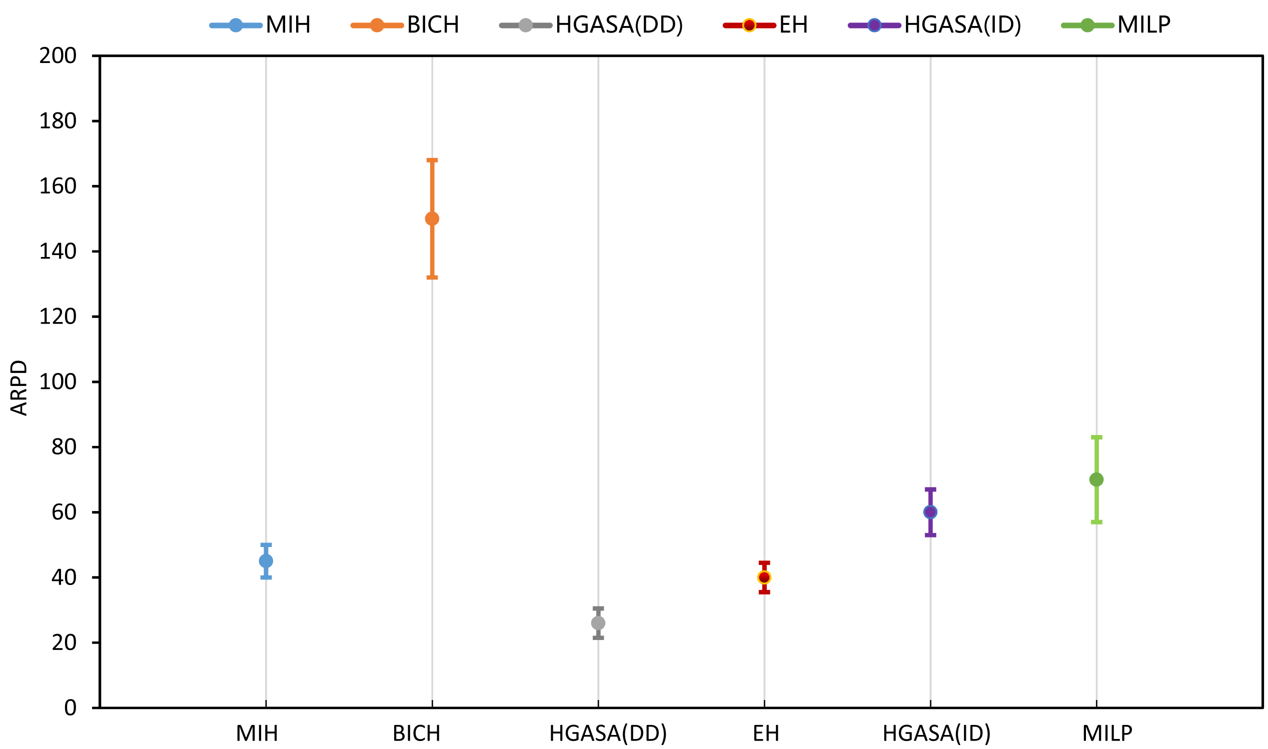


FIGURE 10. ARPD for all methods ($\alpha = 0.05$).

MIH, HGA-SA(DD), HGA-SA(ID) and EH are significantly outperform than the MILP model, even though the former has a much larger CPU effort. Regarding the constructive heuristic method, BICH gives the worst results. Although the performance of MIH is inferior to EH and HGA-SA(DD), the solution's quality obtained by MIH is similar to that obtained by the meta-heuristic algorithm, and the computation time is smaller. Therefore, if the

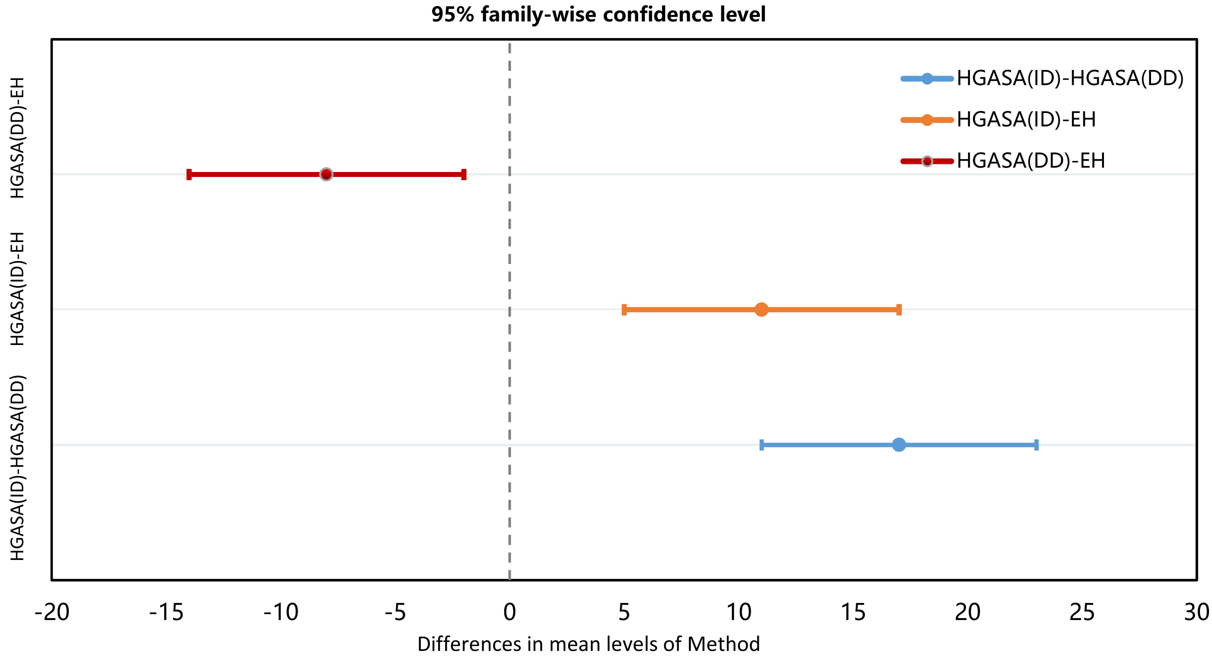


FIGURE 11. Tukey HSD intervals ($\alpha = 0.05$) for considered methods.

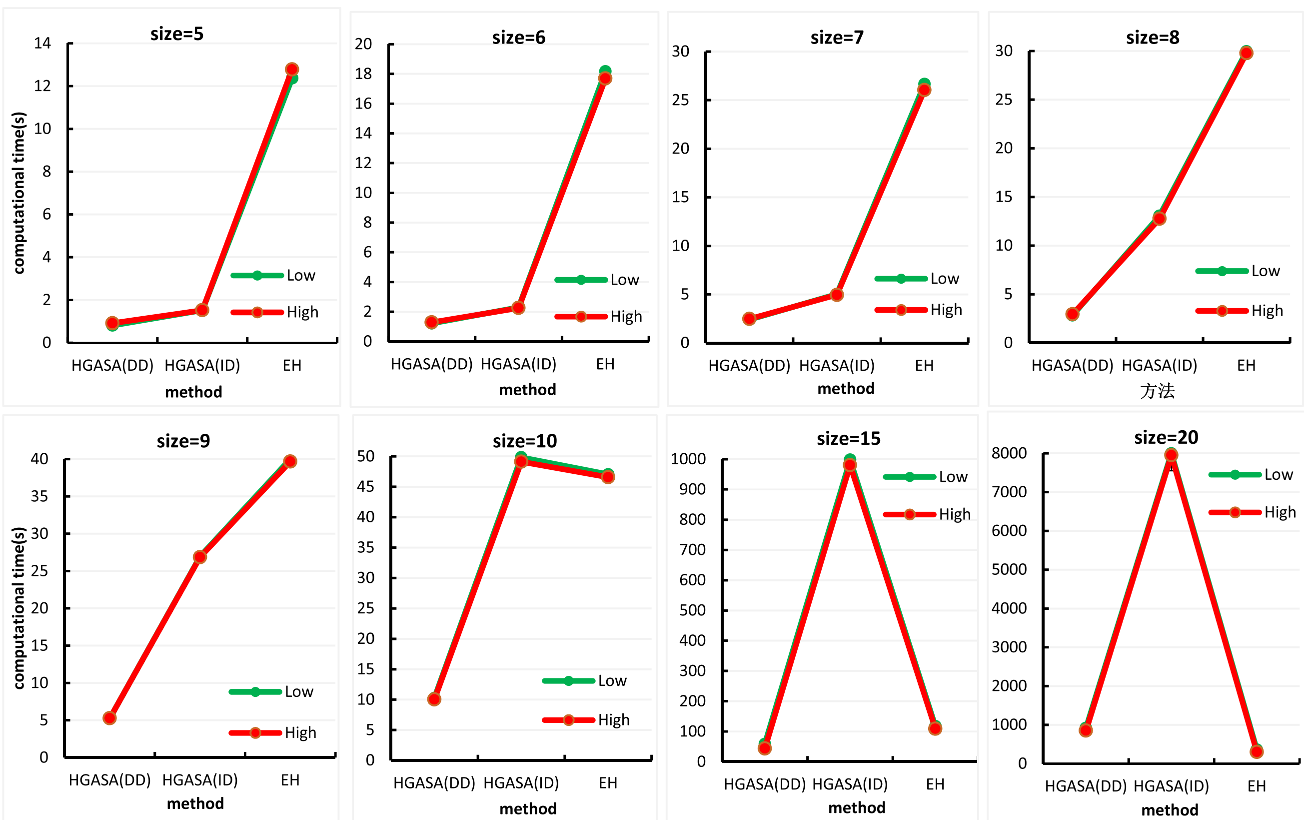


FIGURE 12. The average computation time for each method under different conditions.

solution needs to be obtained in a negligible computation time effort, the algorithm may be an effective method to solve the real-world problems in the medical examination service environment. Regarding the tested meta-heuristic

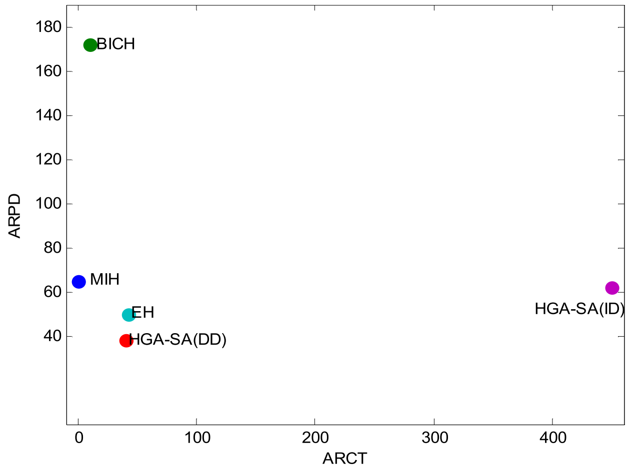


FIGURE 13. ARCT x ARPD of different methods.

algorithm tested, HGA-SA(DD) performed best, even better than the currently best algorithm EH for the problem being studied.

We performed a one-way analysis of variance (ANOVA) and gave HSD Tukey interval (ratio = 0.05) for all combined algorithms in Figure11 to further confirmed the results obtained by the meta-heuristic algorithms [50]. It can be observed from Figure11 that the gap between the average RPD values of all algorithms are statistically significant (that is, the gap between the average RPD crosses the vertical line). HGA-SA(DD) is significantly different from EH, and HGA-SA(DD) outperform EH. In terms of decoding scheme, the quality of the solution obtained by direct decoding (HGA-SA(DD)) is best.

Regarding the computation effort, Table 6 shows the average computation time (in seconds) of each methods under different conditions. We did not report the calculation time for constructive heuristic algorithms, which does not exceed 1 second, and can be negligible. From the results, it can be obtained that the MILP and HGA-SA(ID) have the most computation time, and the performance of HGA-SA(DD) is better than all other meta-heuristic methods. In order to compare the effect of instance size on computation efficiency, and to compare the effect of instance size on computation time, Figure 12 shows the average computation time for each instance size and different setup times. When the instance scale is small, the calculation performance of HGA-SA(DD) is better than that of EH, when the scale is large, the two methods has a small difference. For small-scale instances, different types of setup time have little impact on the calculation performance of the considered algorithms.

In the end, we compare the relative efficiency of ARPD and average calculation time (ARCT) of different methods to tradeoff the quality of the solution obtained by each method and its calculation requirements, shown in Figure13. We can observe that the algorithms proposed in this paper is an effective method for the problem considered. Compared with other decoding methods, HGA-SA(DD) is more efficient.

VIII. CONCLUSION AND FUTURE WORK

Due to the dynamics and inherent uncertainty in the medical service system, unexpected disturbance events will inevitably occur in the system, resulting in the initial scheduling is no longer optimal, so it is necessary to repair the initial scheduling or rescheduling. In this paper, in the physical examination medical service scheduling system involving multiple patients and multiple medical equipment, we have studied the urgent patient disturbance rescheduling problem, in the actual operation of the medical system, the setup time of medical equipment cannot be ignored, so the sequence-dependent setup time was also taken into consideration. In view of the fact that the problem we studied has the characteristics of the open shop scheduling problem, using the open shop scheduling theory, a mixed integer mathematical programming model with the goal of minimizing the total completion time was established. A new hybrid GA-SA algorithm was proposed, the hybrid algorithm used an improved genetic algorithm as global search strategy, two constructive heuristic methods were adopted to initialize the population, two different decoding schemes, three new crossover operator, and restart strategy. The local search strategy of simulated annealing was adopted, taking the minimal disturbance of the original patient's visit sequence between offspring and parent as annealing criterion. In order to evaluate the performance of the proposed algorithm, a large number of calculation experiments have been carried out. The relative deviation percentage was used as an indicator of the quality of the solution, and the average compute time was used as an indicator of the computational effort. The results showed that, in terms of the quality of the solution, the hybrid algorithm proposed in this paper was significantly better than other algorithms, and required the least calculation time.

In future, on the basis of the work in this paper, we will consider the case that the sequence of patients' physical examination items are fixed, which is similar to the flow shop scheduling environment in production scheduling. Under this environment, we will conduct further research on the physical examination medical service rescheduling problem that urgent patient disturb.

REFERENCES

[1] G. Aceto, V. Persico, and A. Pescapé, "The role of information and communication technologies in healthcare: Taxonomies, perspectives, and challenges," *J. Netw. Comput. Appl.*, vol. 107, pp. 125–154, Apr. 2018.

[2] D. D. Zhu, "Research on medical service rescheduling based on urgent patients disturbing," Ph.D. dissertation, Dept. Comput. Sci. Eng., Tianjin Technol. Univ., Tianjin, China, 2020.

[3] G. Aceto, V. Persico, and A. Pescape, "Industry 4.0 and health: Internet of Things, big data, and cloud computing for Healthcare 4.0," *J. Ind. Inf. Integr.*, vol. 18, Jun. 2020, Art. no. 100129, doi: 10.1016/j.jii.2020.100129.

[4] G. Morteza, "Industry 4.0, digitization, and opportunities for sustainability," *J. Cleaner Prod.*, vol. 252, Apr. 2020, Art. no. 119869, doi: 10.1016/j.jclepro.2019.119869.

[5] J. H. Li, "Comparison and enlightenment between German 'Industry 4.0' and 'China manufacture 2025,'" *J. China Univ. Geosci., Social Sci. Ed.*, vol. 15, no. 5, pp. 71–79, Sep. 2015.

[6] Y. P. Guo, "Research on tiered medical care system considering patients' healthcare service choice," Ph.D. dissertation, Dept. Manag. Sci. Eng., Yanshan Univ., Qingdao, China, 2018.

[7] A. Sauré, M. A. Begen, and J. Patrick, "Dynamic multi-priority, multi-class patient scheduling with stochastic service times," *Eur. J. Oper. Res.*, vol. 280, no. 1, pp. 254–265, Jan. 2020.

[8] K. Katragjini, E. Vallada, and R. Ruiz, "Flow shop rescheduling under different types of disruption," *Int. J. Prod. Res.*, vol. 51, no. 3, pp. 780–797, Feb. 2013.

[9] R. Alizadeh, J. Rezaeian, M. Abedi, and R. Chiong, "A modified genetic algorithm for non-emergency outpatient appointment scheduling with highly demanded medical services considering patient priorities," *Comput. Ind. Eng.*, vol. 139, Jan. 2020, Art. no. 106106, doi: 10.1016/j.cie.2019.106106.

[10] H. Mahmoudzadeh, A. M. Shalamzari, and H. A. Mehrizi, "Robust multi-class multi-period patient scheduling with wait time targets," *Oper. Res. Health Care*, vol. 25, Jun. 2020, Art. no. 100254, doi: 10.1016/j.orhc.2020.100254.

[11] M. Oliveira, V. Belanger, I. Marques, and A. Ruiz, "Assessing the impact of patient prioritization on operating room schedules," *Oper. Res. Health Care*, vol. 24, Mar. 2020, Art. no. 100232, doi: 10.1016/j.orhc.2019.100232.

[12] J. F. Marchesi, S. Hamacher, and J. L. Fleck, "A stochastic programming approach to the physician staffing and scheduling problem," *Comput. Ind. Eng.*, vol. 142, Apr. 2020, Art. no. 106281, doi: 10.1016/j.cie.2020.106281.

[13] G. L. Nunez, A. L. Villagra, V. Marianov, C. Obreque, F. Ramis, and L. Neriz, "Scheduling operating rooms with consideration of all resources, post anesthesia beds and emergency surgeries," *Comput. Ind. Eng.*, vol. 97, pp. 248–257, Jul. 2016, doi: 10.1016/j.cie.2016.05.016.

[14] M. Heydari and A. Soudi, "Predictive/reactive planning and scheduling of a surgical suite with emergency patient arrival," *J. Med. Syst.*, vol. 40, no. 1, p. 30, Jan. 2016, doi: 10.1007/s10916-015-0385-1.

[15] J. Castaing, A. Cohn, B. T. Denton, and A. Weizer, "A stochastic programming approach to reduce patient wait times and overtime in an outpatient infusion center," *IIE Trans. Healthcare Syst. Eng.*, vol. 6, no. 3, pp. 111–125, May 2016.

[16] B. Jiang, J. Tang, and C. Yan, "Outpatient capacity allocation considering adding capacity to match high patient demand," *J. Syst. Sci. Syst. Eng.*, vol. 26, no. 4, pp. 487–516, Aug. 2017.

[17] J. Feldman, N. Liu, H. Topaloglu, and S. Ziya, "Appointment scheduling under patient preference and no-show behavior," *Oper. Res.*, vol. 62, no. 4, pp. 794–811, Aug. 2014.

[18] Y. Wang, J. Tang, and R. Y. K. Fung, "A column-generation-based heuristic algorithm for solving operating theater planning problem under stochastic demand and surgery cancellation risk," *Int. J. Prod. Econ.*, vol. 158, pp. 28–36, Dec. 2014.

[19] J. Zhang, M. Dridi, and A. El Moudni, "Column-generation-based heuristic approaches to stochastic surgery scheduling with downstream capacity constraints," *Int. J. Prod. Econ.*, vol. 229, Nov. 2020, Art. no. 107764, doi: 10.1016/j.ijpe.2020.107764.

[20] A. Jebali and A. Diabat, "A stochastic model for operating room planning under capacity constraints," *Int. J. Prod. Res.*, vol. 53, no. 24, pp. 7252–7270, Dec. 2015.

[21] S. Neyshabouri and B. P. Berg, "Two-stage robust optimization approach to elective surgery and downstream capacity planning," *Eur. J. Oper. Res.*, vol. 260, no. 1, pp. 21–40, Jul. 2017.

[22] A. Legrain, J. Omer, and S. Rosat, "An online stochastic algorithm for a dynamic nurse scheduling problem," *Eur. J. Oper. Res.*, vol. 285, no. 1, pp. 196–210, Aug. 2020.

[23] T. I. Wickert, P. Smet, and G. V. Berghe, "The nurse rerostering problem: Strategies for reconstructing disrupted schedules," *Comput. Oper. Res.*, vol. 104, pp. 319–337, Apr. 2019.

[24] M. M. Vali-Siar, S. Gholami, and R. Ramezanian, "Multi-period and multi-resource operating room scheduling under uncertainty: A case study," *Comput. Ind. Eng.*, vol. 126, pp. 549–568, Dec. 2018.

[25] S. Lee and Y. Yih, "Reducing patient-flow delays in surgical suites through determining start-times of surgical cases," *Eur. J. Oper. Res.*, vol. 238, no. 2, pp. 620–629, Oct. 2014.

[26] A. Allahverdi, "The third comprehensive survey on scheduling problems with setup times/costs," *Eur. J. Oper. Res.*, vol. 246, no. 2, pp. 345–378, Oct. 2015.

[27] G. M. Kopanos, J. M. Laínez, and L. Puigjaner, "An efficient mixed-integer linear programming scheduling framework for addressing sequence-dependent setup issues in batch plants," *Ind. Eng. Chem. Res.*, vol. 48, no. 13, pp. 6346–6357, Jul. 2009.

[28] A. Allahverdi and F. S. Al-Anzi, "The two-stage assembly scheduling problem to minimize total completion time with setup times," *Comput. Oper. Res.*, vol. 36, no. 10, pp. 2740–2747, Oct. 2009.

[29] S. C. Trovinger and R. E. Bohn, "Setup time reduction for electronics assembly: Combining simple (SMED) and IT-based methods," *Prod. Oper. Manage.*, vol. 14, no. 2, pp. 205–217, Jan. 2009.

[30] M. T. Y. Sabouni and R. Logendran, "Carryover sequence-dependent group scheduling with the integration of internal and external setup times," *Eur. J. Oper. Res.*, vol. 224, no. 1, pp. 8–22, Jan. 2013.

[31] W. L. Pearn, S. H. Chung, M. H. Yang, and K. P. Shiao, "Solution strategies for multi-stage wafer probing scheduling problem with reentry," *J. Oper. Res. Soc.*, vol. 59, no. 5, pp. 637–651, May 2008.

[32] A. Allahverdi, H. Aydilek, and A. Aydilek, "Two-stage assembly scheduling problem for minimizing total tardiness with setup times," *Appl. Math. Model.*, vol. 40, nos. 17–18, pp. 7796–7815, Sep. 2016.

[33] Y.-Y. Chen, C.-Y. Cheng, L.-C. Wang, and T.-L. Chen, "A hybrid approach based on the variable neighborhood search and particle swarm optimization for parallel machine scheduling problems—A case study for solar cell industry," *Int. J. Prod. Econ.*, vol. 141, no. 1, pp. 66–78, Jan. 2013.

[34] N. C. Ying and M. B. Mokhtar, "Heuristic model for dynamic single machine group scheduling in laser cutting job shop to minimize the makespan," *Adv. Mater. Res.*, vols. 383–390, pp. 6236–6242, Nov. 2011.

[35] C. Park and J. Seo, "A GRASP approach to transporter scheduling for ship assembly block operations management," *Eur. J. Ind. Eng.*, vol. 7, no. 3, pp. 312–332, Feb. 2013.

[36] B. Cankaya, E. Wari, and B. E. Tokgoz, "Practical approaches to chemical tanker scheduling in ports: A case study on the port of houston," *Maritime Econ. Logistics*, vol. 21, no. 4, pp. 559–575, Apr. 2019.

[37] A. Kramer, M. Iori, and P. Lacomme, "Mathematical formulations for scheduling jobs on identical parallel machines with family setup times and total weighted completion time minimization," *Eur. J. Oper. Res.*, vol. 289, no. 3, pp. 825–840, Mar. 2021.

[38] M. Ito, S. Hara, M. Tanaka, and R. Takashima, "Examination-order scheduling for minimizing waiting time: A case study of a medical checkup," *Oper. Res. Health Care*, vol. 22, Sep. 2019, Art. no. 100190, doi: 10.1016/j.orhc.2019.100190.

[39] E. Anand and R. Panneerselvam, "Literature review of open shop scheduling problems," *Intell. Inf. Manage.*, vol. 7, no. 1, pp. 33–52, Jan. 2015.

[40] B. Naderi, S. M. T. F. Ghomi, M. Aminnayeri, and M. Zandieh, "Scheduling open shops with parallel machines to minimize total completion time," *J. Comput. Appl. Math.*, vol. 235, no. 5, pp. 1275–1287, Jan. 2011.

[41] B. Naderi, S. M. T. F. Ghomi, M. Aminnayeri, and M. Zandieh, "A contribution and new heuristics for open shop scheduling," *Comput. Oper. Res.*, vol. 37, no. 1, pp. 213–221, Jan. 2010.

[42] B. Naderi, S. M. T. F. Ghomi, M. Aminnayeri, and M. Zandieh, "Modeling and scheduling open shops with sequence-dependent setup times to minimize total completion time," *Int. J. Adv. Manuf. Technol.*, vol. 53, nos. 5–8, pp. 751–760, Mar. 2011.

[43] S. Noori-Darvish, I. Mahdavi, and N. Mahdavi-Amiri, "A bi-objective possibilistic programming model for open shop scheduling problems with sequence-dependent setup times, fuzzy processing times, and fuzzy due dates," *Appl. Soft Comput.*, vol. 12, no. 4, pp. 1399–1416, Apr. 2012.

[44] V. Fernandez-Viagas and J. M. Framinan, "Efficient non-population-based algorithms for the permutation flowshop scheduling problem with makespan minimisation subject to a maximum tardiness," *Comput. Oper. Res.*, vol. 64, pp. 86–96, Dec. 2015.

[45] L. R. Abreu, J. O. Cunha, B. A. Prata, and J. M. Framinan, "A genetic algorithm for scheduling open shops with sequence-dependent setup times," *Comput. Oper. Res.*, vol. 113, Jan. 2020, Art. no. 104793, doi: 10.1016/j.cor.2019.104793.

[46] J. M. Framinan and P. Perez-Gonzalez, "New approximate algorithms for the customer order scheduling problem with total completion time objective," *Comput. Oper. Res.*, vol. 78, pp. 181–192, Feb. 2017.

[47] C. Guéret and C. Prins, "Classical and new heuristics for the open-shop problem: A computational evaluation," *Eur. J. Oper. Res.*, vol. 107, no. 2, pp. 306–314, Jun. 1998.

[48] E. Taillard, "Benchmarks for basic scheduling problems," *Eur. J. Oper. Res.*, vol. 64, no. 2, pp. 278–285, Jan. 1993.

[49] P. Bruckner, J. Hurink, B. Jurish, and B. Wostmann, "A branch & bound algorithm for the open-shop problem," *Discrete Appl. Math.*, vol. 76, nos. 1–3, pp. 43–59, Jun. 1997.

[50] D. C. Montgomery, "Experiments with a single factor: The analysis of variance," in *Design and Analysis of Experiments*, 8th ed. Hoboken, NJ, USA: Wiley, 2017, ch. 3, pp. 66–80.

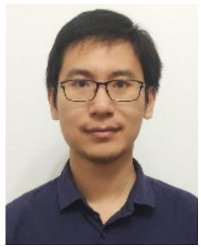
[51] B. Naderi, E. Najafi, and M. Yazdani, "An electromagnetism-like metaheuristic for open-shop problems with no buffer," *J. Ind. Eng. Int.*, vol. 8, no. 1, pp. 1–8, Dec. 2012.

[52] G. Taguchi, S. Chowdhury, and S. Taguchi, *Robust Engineering*. New York, NY, USA: McGraw-Hill, 2000, pp. 1–241.

[53] B. de Athayde Prata, "A hybrid genetic algorithm for the vehicle and crew scheduling in mass transit systems," *IEEE Latin Amer. Trans.*, vol. 13, no. 9, pp. 3020–3025, Sep. 2015.

DAN-DAN ZHU received the B.S. degree in communication engineering from Heilongjiang University, Harbin, China, and the master's degree in computer science and technology from the Tianjin University of Technology, Tianjin, China, where she is currently pursuing the Ph.D. degree with the Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology. Her research interests include modeling and simulation of complex systems, algorithm design and optimization, and scheduling theory.

JUN-QING SUN received the B.S. and master's degrees from the Harbin Institute of Technology, Harbin, China, and the Ph.D. degree from Nankai University, Tianjin, China. He is currently a Professor with the Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Tianjin University of Technology. His research interests include the modeling and simulation of complex systems, optimization of supply chains and modern logistics, service systems, and service computing.



YU ZHAO received the bachelor's degree in bioinformatics from Hebei University, Baoding, China, and the master's degree in computer science and technology from the Tianjin University of Technology, Tianjin, China. He is currently working with the People's Bank of China Tianjin Branch. His research interests include modeling and simulation of complex systems, algorithm design and optimization, and schedule theory.